

Discretization Error Analysis and Adaptive Meshing Algorithms for Fluorescence Diffuse Optical Tomography: Part II

Murat Guven, Lu Zhou, Laurel Reilly-Raska, and Birsen Yazıcı*, *Senior Member, IEEE*

Abstract—In the first part of this work, we analyze the effect of discretization on the accuracy of fluorescence diffuse optical tomography (FDOT). Our error analysis provides two new error estimates which present a direct relationship between the error in the reconstructed fluorophore concentration and the discretization of the forward and inverse problems. In this paper, based on these error estimates, we develop two new adaptive mesh generation algorithms for the numerical solutions of the forward and inverse problems in FDOT, with the objective of error reduction in the reconstructed optical images due to discretization while keeping the size of the discretized forward and inverse problems within the allowable limits. We present three-dimensional numerical simulations to demonstrate the improvements in accuracy, resolution and detectability of small heterogeneities in reconstructed images provided by the use of the new adaptive mesh generation algorithms. Finally, we compare our algorithms both analytically and numerically with the existing conventional adaptive mesh generation algorithms.

Index Terms—Adaptive meshing algorithms, error analysis, fluorescence diffuse optical tomography.

I. INTRODUCTION

MOST practical applications in fluorescence diffuse optical tomography (FDOT) require numerical solutions for the forward and inverse problems which in turn require discretization of the domain. These numerics frequently present a tradeoff between accuracy in the final reconstructed image and the computational complexity. One way to achieve greater gains in accuracy is to reduce the size of the mesh used to discretize the forward and inverse problems. However, reducing the

mesh size may imply a substantial increase in the size of the resulting discrete problems, thereby reducing the computational efficiency of image formation. A solid understanding of the key factors in the error in FDOT imaging due to discretization of the forward and inverse problems can motivate the development of adaptive mesh generation algorithms which address the computational complexity versus accuracy tradeoff.

In this paper, based on the error analysis presented in Part I [1], we present two new adaptive mesh generation algorithms for the numerical solutions of the forward and inverse problems. The resulting locally refined adaptive meshes aim to reduce the error in the reconstructed optical images due to discretization while keeping the size of the discrete forward and inverse problems within the limits of allowable nodes number in the mesh.

There has been extensive research on adaptive mesh generation for the numerical solution of partial differential equations [2]–[7] and inverse parameter estimation problems to reduce the undesired effect of discretization error [8]–[11]. For example, in [10], [11], an *a posteriori* error estimate is derived and an adaptive mesh refinement method is presented for the solution of an inverse acoustic scattering problem where the inverse problem is formulated as an optimal control problem. In the area of DOT, several investigators have reported on adaptive meshing schemes for the forward and inverse problems to address the optical image degradation due to discretization [12]–[16]. In [17], the authors numerically showed that approximation errors resulting from the discretization of the forward problem can lead to significant artifacts in the reconstructed optical images. In that work, these artifacts are minimized by modeling and treating the approximation error within the Bayesian framework. In the area of FDOT, it has been shown that using adaptively refined meshes results in better accuracy and higher resolution in reconstructed images than that of uniform meshes when computational resources are constrained [18], [19]. In [18], the image reconstruction problem is formulated in a PDE-constrained optimization framework, and the mesh for solving this optimization problem is updated and refined based on the criterion suggested in dual weighted residual framework [20]. In [21], an algorithm to identify and resolve intersection of tetrahedral elements was developed to achieve fast and robust parameter mapping between the adaptively refined/derefining meshes of PDE-based forward and inverse problems. In [19] this algorithm was utilized for FDOT image reconstruction within a dual adaptive mesh scheme that are independently refined based on *a posteriori* error estimates. Note that in all of these studies, with the exception of [10], [11], [16], [18], [20], [22], the mesh refinement criterion is consid-

Manuscript received June 07, 2009; accepted July 31, 2009. First published August 25, 2009; current version published February 03, 2010. This work was supported in part by U.S. Army Medical Research-W81XWH-04-1-0559 and in part by the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation under Award EEC-9986821. Asterisk indicates corresponding author.

M. Guven is with Intel Corporation USA, Santa Clara, CA 95050 USA (e-mail: guven@rpi.edu).

L. Zhou is with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: zhou2@rpi.edu; yazici@ecse.rpi.edu).

L. Reilly-Raska is with Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420 USA (e-mail: reilll@rpi.edu).

*B. Yazıcı is with the Department of Electrical, Computer and Systems Engineering and the Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY USA (e-mail: zhou2@rpi.edu; yazici@ecse.rpi.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2009.2029855

ered separately for each problem disregarding its impact on the solution of the other problem.

In this two-part study, as well as in [16] and [22], we introduced a framework to analyze the error in the reconstructed images due to discretization of the forward and inverse problems and associated adaptive meshing algorithms that take into account domain-specific factors and the interdependence of the solutions of the forward and inverse problems. While both the present and our earlier study use the same framework, they differ in a number of ways: Apart from the obvious difference between the imaging modalities, the present study formulates the inverse problem in an optimization framework as opposed to the matrix inversion formulation used in [22]. This new formulation can incorporate wide-range of *a priori* information in the form of penalty functions. Additionally, after converting the optimization problem to a second-order partial differential equation, both the forward and inverse problems are treated in a FEM framework, where both problems are discretized by the Galerkin method. The new formulation of the inverse problem together with the new discretization result in new and easy-to-evaluate error estimates for practical applications. The adaptive mesh generation algorithms described in our present work are based on the new error estimates, and therefore significantly different from those in [16].

In this work, we introduce adaptive mesh generation algorithms for the finite element solutions of the forward and inverse problems in FDOT, based on the two new error estimates presented in Part I [1]. We note that the mesh refinement criterion for each problem takes into account the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. Thus, the presented adaptive mesh generation algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the relative location of the sources and detectors with respect to the fluorophore concentration. We consider two approaches for the mesh generation, first of which uses a predefined error bound tolerance as a constraint for adaptive refinement. Next, we present a more practical approach which uses the number of nodes as a constraint for the number of discretization points of the final adaptive mesh.

We present three-dimensional numerical simulations to illustrate the practical advantages of our adaptive mesh generation algorithms in FDOT image reconstruction. Our numerical simulations show that the new algorithms significantly improve the accuracy, resolution, and the detectability of small scale heterogeneities as compared to that of conventional adaptive mesh generation algorithms when the number of unknowns in the forward and inverse problems is constrained. We also show that the computational complexity of adaptive mesh generation is less than that of solving the resulting discrete problems.

The outline of the paper is as follows. In Section II, we provide an overview of our results from Part I and recall the notation adapted. In Section III, we present two new adaptive mesh generation algorithms. In Section IV, we provide a three-dimensional numerical study to demonstrate the performance of our adaptive mesh generation algorithms. Section V concludes our discussion.

TABLE I
DEFINITION OF FUNCTION SPACES AND NORMS

Notation	Explanation
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \text{ess sup}_\Omega f(\mathbf{x}) < \infty\}$
$L^p(\Omega)$	$L^p(\Omega) = \{f (\int_\Omega f(\mathbf{x}) ^p d\mathbf{x})^{1/p} < \infty\}, p \in [1, \infty)$
$H^p(\Omega)$	$H^p(\Omega) = \{f (\sum_{ z \leq p} \ D_z^* f\ _0^2)^{1/2} < \infty\}, p \in [1, \infty)$
$\ f\ _0$	The $L^2(\Omega)$ norm of f
$\ f\ _p$	The $L^p(\Omega)$ norm of f
$\ f\ _\infty$	The $L^\infty(\Omega)$ norm of f
$\ f\ _{0,m}$	The L^2 norm of f over the m th finite element Ω_m
$\ f\ _{p,m}$	The L^p norm of f over the m th finite element Ω_m
$\ f\ _{\infty,m}$	The L^∞ norm of f over the m th finite element Ω_m

II. OVERVIEW

A. Notational Conventions

In this paper, we denote operators by capital cursive Latin letters (\mathcal{A}), \mathcal{A}^* denotes the adjoint. Functions are denoted by lowercase Latin or Greek letters. For a function f , F denotes its finite element approximation. We use bold to denote vectorized quantities such as \mathbf{r} . Table I provides a summary of function spaces and norms used throughout the paper.

B. Forward and Inverse Problems in FDOT

In this work, we use the coupled diffusion equations to model the light transport in a fluorescent medium of a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$

$$-\nabla \cdot D(\mathbf{r}) \nabla \phi_x(\mathbf{r}, \mathbf{r}_i) + \mu_{ax}(\mathbf{r}) \phi_x(\mathbf{r}, \mathbf{r}_i) = S_i(\mathbf{r}) \quad (1)$$

$$-\nabla \cdot D(\mathbf{r}) \nabla \phi_m(\mathbf{r}, \mathbf{r}_i) + \mu_{am}(\mathbf{r}) \phi_m(\mathbf{r}, \mathbf{r}_i) = \phi_x(\mathbf{r}, \mathbf{r}_i) \eta \mu_{axf}(\mathbf{r}) \quad (2)$$

where $\mathbf{r} = [r_1, r_2, r_3] \in \Omega$, S_i is the i th excitation source for $i = 1, \dots, N_S$, N_S is the number of sources. The sources are modeled by a Gaussian function centered at source position \mathbf{r}_i with angular frequency $\omega = 0$ [1]. $\phi_{x,m}$ represent the optical fields at the excitation and emission wavelengths, respectively, D is the isotropic diffusion coefficient, μ_{ax} and μ_{am} represent the absorption coefficient of the medium at the excitation and emission wavelengths, respectively, η is the quantum efficiency, and μ_{axf} is the absorption coefficient of the fluorophore concentration. The Robin-type boundary conditions associated with (1) and (2) are

$$2 - D(\mathbf{r}) \frac{\partial \phi_x(\mathbf{r}, \mathbf{r}_i)}{\partial n} + \rho \phi_x(\mathbf{r}, \mathbf{r}_i) = 0 \quad (3)$$

$$2 - D(\mathbf{r}) \frac{\partial \phi_m(\mathbf{r}, \mathbf{r}_i)}{\partial n} + \rho \phi_m(\mathbf{r}, \mathbf{r}_i) = 0 \quad (4)$$

where $\mathbf{r} \in \partial\Omega$, 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 of later sections, we make use of the adjoint problem associated with (2) and (4)

$$-\nabla \cdot D(\mathbf{r}) \nabla g_m^*(\mathbf{r}, \mathbf{r}_j) + \mu_{am}(\mathbf{r}) g_m^*(\mathbf{r}, \mathbf{r}_j) = 0, \quad \mathbf{r} \in \Omega, \quad (5)$$

$$2 - D(\mathbf{r}) \frac{\partial g_m^*(\mathbf{r}, \mathbf{r}_j)}{\partial n} + \rho g_m^*(\mathbf{r}, \mathbf{r}_j) = S_j^*(\mathbf{r}), \quad \mathbf{r} \in \partial\Omega \quad (6)$$

where g_j^* is the solution of the adjoint problem for the j th adjoint source S_j^* located at the detector position $\mathbf{r}_j \in \partial\Omega$, $j = 1, \dots, N_D$, N_D is the number of detectors. For a point adjoint source located at the detector position \mathbf{r}_j , $g_m^*(\mathbf{r}, \mathbf{r}_j) = g_m(\mathbf{r}_j, \mathbf{r})$ holds where $g_m(\cdot, \mathbf{r})$ is the Green's function of (2) and (4). Note that we model the adjoint source by a Gaussian function centered at \mathbf{r}_j [1], [22]. For the rest of the paper, we represent $g_m^*(\mathbf{r}, \mathbf{r}_j)$ by $g_j^*(\mathbf{r})$ and $\phi_x(\mathbf{r}, \mathbf{r}_i)$ by $\phi_i(\mathbf{r})$, suppressing the x, m dependence of these functions.

Next, we give the variational formulations of the boundary value problems (1) and (3) which are used to discretize the forward problem by finite element method

$$\int_{\Omega} (\nabla \xi_1 \cdot D \nabla \phi_i + \mu_{\text{axf}} \xi_1 \phi_i) d\mathbf{r} + \frac{1}{2\rho} \int_{\partial\Omega} \xi_1 \phi_i dl = \int_{\Omega} \xi_1 S_i d\mathbf{r} \quad (7)$$

where ξ_1 is a test function. Similarly for a test function ξ_2 , the variational form for the adjoint problem (5), (6) is

$$\int_{\Omega} (\nabla \xi_2 \cdot D \nabla g_j^* - \mu_{\text{am}} \xi_2 g_j^*) d\mathbf{r} + \frac{1}{2\rho} \int_{\partial\Omega} \xi_2 g_j^* dl = \frac{1}{2\rho} \int_{\partial\Omega} \xi_2 S_j^* dl. \quad (8)$$

In this work, we focus on the estimation of the unknown optical absorption coefficient, μ_{axf} . We use an iterative linearization scheme based on first order Fréchet derivatives. As a result, at each linearization step, the following linear integral equation relates the perturbation $\delta\mu$ in the absorption coefficient of the fluorophore concentration (scaled by the quantum efficiency η) to the differential emission measurements $\delta\Gamma_{i,j}$ [1]

$$\begin{aligned} \delta\Gamma_{i,j} &= \int_{\Omega} g_m^*(\mathbf{r}, \mathbf{r}_j) \phi_x(\mathbf{r}, \mathbf{r}_i) \eta \delta\mu_{\text{axf}}(\mathbf{r}) d\mathbf{r} \\ &:= \int_{\Omega} H_{ij}(\mathbf{r}) \delta\mu(\mathbf{r}) d\mathbf{r} \\ &:= (\mathcal{A}\delta\mu)_{ij}. \end{aligned} \quad (9)$$

Let \mathcal{A}^* denote the adjoint of \mathcal{A} [1], and $\mathcal{B} := \mathcal{A}^* \mathcal{A}$. Representing the differential measurements corresponding to individual source-detector pairs as elements of the vector $\delta\mathbf{\Gamma} = [\delta\Gamma_{1,1}, \dots, \delta\Gamma_{N_S, N_D}]^T$, and defining $\gamma = \mathcal{A}^* \delta\mathbf{\Gamma}$, (9) can be equivalently represented as follows:

$$\gamma(\mathbf{r}) = (\mathcal{B}\delta\mu)(\mathbf{r}). \quad (10)$$

We address the ill-posedness of (10) using zeroth- and first-order Tikhonov regularization terms in an optimization framework. With a similar reasoning as in the forward problem, we formulate the regularized inverse problem in the variational framework. Let $\psi \in H^1(\Omega)$ be a test function and

$$\mathcal{F}(\psi, \delta\mu) := (\psi, \mathcal{B}\delta\mu) + \lambda_1(\psi, \delta\mu) + \lambda_2(\nabla\psi, \nabla\delta\mu) \quad (11)$$

$$\mathcal{G}(\psi) := (\psi, \gamma) \quad (12)$$

where

$$(\mathbf{k}, \mathbf{l}) := \int_{\Omega} \mathbf{k}(\mathbf{r}) \cdot \mathbf{l}(\mathbf{r}) d\mathbf{r}$$

and $\lambda_1, \lambda_2 > 0$ are regularization parameters. Then, the regularized inverse problem in the variational framework becomes [1]

$$\mathcal{F}(\psi, \delta\mu) = \mathcal{G}(\psi). \quad (13)$$

Next, we briefly describe the discretization of the forward and inverse problems and present the effect of discretization on the accuracy of FDOT imaging in two theorems derived in [1]. First, we consider the effect of forward problem discretization on accuracy of the reconstructed optical image and then we discuss the effect of the inverse problem discretization.

C. Error in FDOT Due to Forward Problem Discretization

Let L_k , $k = 1, \dots, N_i$ denote the first-order Lagrange basis functions. We define $Y_i \subset H^1(\Omega)$, $i = 1, \dots, N_S$ as the finite-dimensional subspace spanned by $\{L_k\}$ which are associated with the set of points $\{\mathbf{r}_p\}$, $p = 1, \dots, N_i$, on Ω . Note that the dimension of Y_i is N_i . Similarly, we define $Y_j^* \subset H^1(\Omega)$, $j = 1, \dots, N_D$, as the finite-dimensional subspace spanned by $\{L_k\}$, $k = 1, \dots, N_j$, which are associated with the set of N_j points $\{\mathbf{r}_p\}$. We discretize the variational problems (7) and (8) by replacing the functions ξ_1 , ϕ_i in (7) and ξ_2 , g_j^* in (8) with their finite-dimensional counterparts

$$\Xi_2(\mathbf{r}) := \sum_{k=1}^{N_i} p_k L_k(\mathbf{r}), \quad \Phi_i := \sum_{k=1}^{N_i} c_k L_k(\mathbf{r}). \quad (14)$$

$$\Xi_1(\mathbf{r}) := \sum_{k=1}^{N_j} p_k L_k(\mathbf{r}), \quad G_j^* := \sum_{k=1}^{N_j} d_k L_k(\mathbf{r}). \quad (15)$$

We substitute the finite element solutions (14), (15) into (10) and denote the resulting operator and function by respectively $\tilde{\mathcal{B}}$ and $\tilde{\gamma}$. Then, the approximate variational formulation can be written as follows:

$$\tilde{\mathcal{F}}(\psi, \tilde{\delta\mu}) = \tilde{\mathcal{G}}(\psi) \quad (16)$$

where $\tilde{\mathcal{F}}$ indicates that the finite element solutions (14), (15) are used, and $\tilde{\delta\mu}$ denotes the approximation to the actual solution $\delta\mu$. Theorem 1 presents a bound for the error between $\delta\mu$ and $\tilde{\delta\mu}$, which can be attributed to the discretization of the forward problem.

Theorem 1: Let $\{\Omega_{ni}\}$ denote the set of linear elements used to discretize (7) for $n = 1, \dots, N_{\Delta}^i$; 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 $i = 1, \dots, N_S$. Similarly, let $\{\Omega_{mj}\}$ denote the set of linear elements used to discretize (8) for $m = 1, \dots, N_{\Delta}^{*j}$; 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 $j = 1, \dots, N_D$. Then a bound for the

error between the solution $\delta\mu$ of (13) and the solution $\widetilde{\delta\mu}$ of (16) due to the approximations $\widetilde{\mathcal{F}}$ and $\widetilde{\mathcal{G}}$ is given by

$$\begin{aligned} & \|\delta\mu - \widetilde{\delta\mu}\|_1 \\ & \leq \frac{C}{\min(\lambda_1, \lambda_2)} \max_{i,j} \|g_j^* \phi_i\|_0 \\ & \quad \cdot \left(\sum_{i=1}^{N_S} \sum_{n,j}^{N_{\Delta}^i, N_D} \left(2 \|g_j^* \delta\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\delta\mu\|_0 \right) \right. \\ & \quad \cdot \|\phi_i\|_{1,ni} h_{ni} \\ & \quad + \sum_{j=1}^{N_D} \sum_{m,i}^{N_{\Delta}^j, N_S} \left(2 \|\phi_i \delta\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta\mu\|_0 \right) \\ & \quad \left. \times \|g_j^*\|_{1,mj} h_{mj} \right) \end{aligned} \quad (17)$$

where $\lambda_1, \lambda_2 > 0$ are regularization parameters.

Proof: See [1]

D. Error in FDOT Due to Inverse Problem Discretization

Next, we discretize the inverse problem approximation (16) using Galerkin discretization. Let $V_N \subset H^1(\Omega)$ denote a sequence of finite-dimensional subspaces of dimension N , spanned by the first-order Lagrange basis functions $\{L_1, \dots, L_N\}$ which are associated with the set of points $\{\mathbf{r}_p\}$, $p = 1, \dots, N$, on Ω . We replace ψ and $\widetilde{\delta\mu}$ in (16) by their respective finite dimensional counterparts $\Psi \in V_N$ and $\widetilde{\Delta\mu} \in V_N$

$$\Psi := \sum_{k=1}^N p_k L_k(\mathbf{r}) \quad (18)$$

$$\widetilde{\Delta\mu} := \sum_{k=1}^N m_k L_k(\mathbf{r}) \quad (19)$$

where p_k and m_k are unknown coefficients. Substituting (18), (19) into (16), we arrive at the fully discrete inverse problem

$$\widetilde{\mathcal{F}}(\Psi, \widetilde{\Delta\mu}) = \widetilde{\mathcal{G}}(\Psi). \quad (20)$$

We define the error due to the inverse problem discretization to be the difference between the solution $\widetilde{\delta\mu}$ of (16) and the solution $\widetilde{\Delta\mu}$ of (20). The following theorem presents a bound for the $H^1(\Omega)$ norm of this error.

Theorem 2: Consider the Galerkin projection of the variational problem (16) on a finite dimensional subspace $V_N \subset H^1(\Omega)$ using a set of linear finite elements $\{\Omega_t\}$, for $t = 1, \dots, N_{\Delta}$ whose vertices are at $\{r_p\}$, $p = 1, \dots, N$ 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 $\widetilde{\delta\mu}$ of (16) also satisfies $\widetilde{\delta\mu} \in H^2(\Omega)$. Then, a bound for the error

E in the solution $\widetilde{\Delta\mu}$ of (20) with respect to the solution $\widetilde{\delta\mu}$ of (16) can be given by

$$\begin{aligned} & \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \\ & \leq \frac{C}{\min(\lambda_1, \lambda_2)} \\ & \quad \times \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \cdot \sum_t^{N_{\Delta}} \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 \right. \\ & \quad \left. + \lambda_1 \sum_t^{N_{\Delta}} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 + \lambda_2 \sum_t^{N_{\Delta}} \|\widetilde{\delta\mu}\|_{2,t} h_t \right) \end{aligned} \quad (21)$$

where $\lambda_1, \lambda_2 > 0$ are regularization parameters.

Proof: See [1].

Part I of this study contains the full details of the error analysis due to discretization of the forward and inverse problems. Note that the analysis of the error between the true solution $\delta\mu$ of (10) and the solution $\widetilde{\Delta\mu}$ of the fully discretized inverse problem (20) is divided into two steps: Theorem 1 describes the error difference due to discretization of the forward problem while Theorem 2 focuses on the inverse problem discretization. This convenient representation allows for designing an adaptive mesh for each of the forward and inverse problems. In the next section, we discuss application of these error bounds as an error criterion in designing novel and effective mesh generation algorithms.

III. ADAPTIVE MESH GENERATION ALGORITHMS

We seek to develop algorithms for the discretization of the forward and inverse problems using two theorems proposed before. By selecting $\|\delta\mu - \widetilde{\delta\mu}\|_1$ and $\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ as the figures of merit and reducing the respective bounds (17) and (21), we aim to minimize the error between the fully discretized solution $\widetilde{\Delta\mu}$ and the true solution $\delta\mu$. Clearly, we can reduce the both error bounds by reducing the mesh size parameter [h_{ni} and h_{mj} in (17) and h_t in (21)], 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 result in an increase in the size of the discrete forward and/or inverse problems. In this respect, we consider two h -refinement approaches, first of which targets to achieve a certain predetermined error bound, while the latter is constrained by the maximum number of discretization points (i.e., the size of discrete problems) permitted by the computing resources at hand.

The first approach we discuss sets a predetermined error tolerance value for the error bound computed on each finite element and reduce the mesh size parameter by refining the element until this predetermined tolerance value is reached. This approach relies on the knowledge of the value for the predetermined error bound which gives the desired level of accuracy of discretization for the reconstructed image. From a practical standpoint, defining the ideal predetermined value for the error tolerance bound can be a complex undertaking, since the presence of unknown factor and functions such as C , g_j^* , ϕ_i , $\delta\mu$, and $\widetilde{\delta\mu}$ in both theorems.

The second approach, which is more easily applied in practice, limits the total number of discretization points in the finite dimensional approximations (14), (15), and (19). Instead of using a predetermined error tolerance bound, the average of the error bounds computed on the finite elements is calculated at each iteration in the refinement process. Then, the elements with computed error bounds larger than this average are refined until the limit of the number of discretization points is reached.

In this respect, the first approach can be viewed to rely on *a priori* information while the second approach makes use of *a posteriori* information on the error bounds. In the next two subsections, we design the adaptive mesh generation algorithms that make use of the error estimates derived previously for the forward and inverse problems, respectively. First, we explore the mesh generation for the forward problem and describe the algorithm which uses a predetermined error bound before discussing the second approach. Next, we present the adaptive mesh generation algorithm for the inverse problem in a similar way.

A. Adaptive Mesh Generation for Forward Problem

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 (17) contains known or estimable factors multiplying a summation represented by two terms. The first term arises from the discretization of the boundary value problem (1) and (3) for each source, and the second term arises from the boundary value problem (5), (6) 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 in the reconstructed optical image is less than $\tilde{\varepsilon}_f$.

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

$$\varepsilon_f := \frac{\min(\lambda_1, \lambda_2)}{C \max_{i,j} \|g_j^* \phi_i\|_0} \left(\sum_i^{N_S} N_{\Delta}^i + \sum_j^{N_D} N_{\Delta}^{*j} \right)^{-1} \tilde{\varepsilon}_f \quad (22)$$

where C is the positive constant in (17) and N_{Δ}^i (N_{Δ}^{*j}) are the number of elements in the finite element approximations Φ_i (C_j^*). Let B_{ni} and B_{mj} for $i = 1, \dots, N_S$ and $j = 1, \dots, N_D$ be defined as

$$B_{ni} := \frac{\varepsilon_f}{\sum_j^{N_D} \left(2 \|g_j^* \delta \mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\delta \mu\|_0 \right) \|\phi_i\|_{1,ni}} \quad (23)$$

$$B_{mj}^* := \frac{\varepsilon_f}{\sum_i^{N_S} \left(2 \|\phi_i \delta \mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta \mu\|_0 \right) \|g_j^*\|_{1,mj}} \quad (24)$$

Then, 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 $\|\delta \mu - \widetilde{\delta \mu}\|_1 \leq \tilde{\varepsilon}_f$.

The algorithm is initiated starting with a coarse uniform mesh. Each mesh size parameter h_{ni} (h_{mj}) is checked against

the corresponding bound B_{ni} (B_{mj}^*). Then, each element with mesh size parameter h_{ni} (h_{mj}) larger than the bound B_{ni} (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} (h_{mj}) for each element is less than the bound B_{ni} (B_{mj}^*). In other words, the tolerance error bound $\tilde{\varepsilon}_f$ is uniformly distributed over all finite elements. Note that B_{ni} (B_{mj}^*) are computed for a given set of meshes $\{\Delta^i\}$, $i = 1, \dots, N_S$ ($\{\Delta^{*j}\}$, $j = 1, \dots, N_D$), and the associated number of elements is N_{Δ}^i (N_{Δ}^{*j}) in each of these meshes.

The practical implementation of the algorithm requires several adjustments, because B_{ni} and B_{mj}^* in (23) and (24) can not be computed exactly due to the unknown actual values of g_j^* , ϕ_i , $\delta \mu$ and unknown constant C . First, ε_f can be scaled by C to eliminate the unknown constant C , and B_{ni} , B_{mj}^* can be estimated using analytical solutions for g_j^* , ϕ_i (see Appendix II) and approximations of $\delta \mu$ based on *a priori* information [16]. Alternatively, these functions can also be estimated using the updated finite dimensional solutions of g_j^* , ϕ_i and $\delta \mu$ from the most recent solution updates.

Finally, in addition to these practical modifications, we 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 discrete 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 terms in summation in (17) and define two error indicators

$$\varepsilon_f^i(n) = \sum_j^{N_D} \left(2 \|g_j^* \delta \mu\|_{0,ni} + \|\delta \mu\|_0 \|g_j^*\|_{\infty,ni} \right) \times \|\phi_i\|_{1,ni} h_{ni} \quad (25)$$

$$\varepsilon_f^j(m) = \sum_i^{N_S} \left(2 \|\phi_i \delta \mu\|_{0,mj} + \|\delta \mu\|_0 \|\phi_i\|_{\infty,mj} \right) \times \|g_j^*\|_{1,mj} h_{mj} \quad (26)$$

Then, for each source (detector), we compute $\varepsilon_f^i(n)$ ($\varepsilon_f^j(m)$) on every finite element and compute the average value $\bar{\varepsilon}_f^i$ ($\bar{\varepsilon}_f^j$) of them. Every element with $\varepsilon_f^i(n) > \bar{\varepsilon}_f^i$ ($\varepsilon_f^j(m) > \bar{\varepsilon}_f^j$) is refined so that the new $\varepsilon_f^i(n)$ ($\varepsilon_f^j(m)$) value for that element will be smaller. The algorithm must be stopped when the number of discretization points reaches the allowable limit. This method is further described in the Algorithm 1 by pseudocode.

B. Adaptive Mesh Generation for Inverse Problem

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 bound to be distributed over the whole domain. Then, we discuss the practical approach through limiting the number of discretization points.

Let $\tilde{\varepsilon}_i$ be the *predetermined* allowable error bound on the reconstructed image due to the Galerkin projection of (16). We

Algorithm 1 The pseudocode for the mesh generation algorithm for at every linearization step of the forward problem.

◇ Generate an initial uniform mesh:
 $(\Delta^{*j}, N_{\Delta}^{*j}), \Delta^{*j} = \bigcup_{m=1}^{N_{\Delta}^{*j}} \{\Delta_m\}$
 $((\Delta^i, N_{\Delta}^i), \Delta^i = \bigcup_{n=1}^{N_{\Delta}^i} \{\Delta_n\})$

◇ Set the maximum number of nodes N_{max}^f
while Number of nodes in Δ^{*j} (Δ^i) less than N_{max}^f
 for each element $\Delta_m \in \Delta^{*j}$ with mesh parameter h_{mj} (for $\Delta_n \in \Delta^i$ with h_{ni})
 if first linearization
 • Use analytical solutions for ϕ_i and g_j^* and *a priori* information about $\delta\mu$ to compute $\varepsilon_f^j(m)$ in (25) ($\varepsilon_f^i(n)$ in (26))
 else
 • Use current solution updates G_j, G_i^* and $\widetilde{\Delta\mu}$ to compute $\varepsilon_f^j(m)$ in (25) ($\varepsilon_f^i(n)$ in (26))
 end
 • Compute $\bar{\varepsilon}_f^j$ ($\bar{\varepsilon}_f^i$)
 • Refine the elements with $\varepsilon_f^j(m) > \bar{\varepsilon}_f^j$ ($\varepsilon_f^i(n) > \bar{\varepsilon}_f^i$)
 • Update the mesh Δ^{j*} (Δ^i)
end
 ◇ Solve for G_j^* (Φ_i)

desire to distribute this error bound evenly over all elements in the domain. Let

$$\varepsilon_i := \frac{\min(\lambda_1, \lambda_2)}{CN_{\Delta}} \tilde{\varepsilon}_i \quad (27)$$

and

$$a(t) := \max_{i,j} \|G_j^* \Phi_i\|_0 \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} \|\widetilde{\delta\mu}\|_{2,t} + \lambda_1 \|\widetilde{\delta\mu}\|_{2,t} \quad (28)$$

$$b(t) := \lambda_2 \|\widetilde{\delta\mu}\|_{2,t} \quad (29)$$

such that Theorem 2 can be rewritten as follows:

$$\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \leq \frac{C}{\min(\lambda_1, \lambda_2)} \sum_t^{N_{\Delta}} a(t)h_t^2 + b(t)h_t. \quad (30)$$

Next, we define the parameter B_t as follows:

$$B_t := \frac{-b(t) + \sqrt{b(t)^2 + 4a(t)\varepsilon_i}}{2a(t)}. \quad (31)$$

Then for $0 < h_t \leq B_t$, by (30) and (27), we have

$$\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \leq \frac{CN_{\Delta}}{\min(\lambda_1, \lambda_2)} \varepsilon_i \leq \tilde{\varepsilon}_i. \quad (32)$$

Determining the bound in practice from (31) requires knowledge of $\widetilde{\delta\mu}$. In this respect, we can either estimate $\delta\mu$ using *a priori*

information, or use the most recent updates of the inverse problem solutions $\widetilde{\Delta\mu}$ as the approximation to it.

The second more practical approach limits the number of discretization points in the adaptive mesh (thereby determining the size of the discrete inverse problem) rather than relying on a predetermined error tolerance bound. As in the previous subsection, this algorithm uses an error bound computed after each sweep of refinement on every finite element

$$\varepsilon_i(t) := a(t)h_t^2 + b(t)h_t \quad (33)$$

where $a(t)$ and $b(t)$ are defined in (28) and (29), respectively. Given an initial uniform mesh, we compute the value of ε_i for each element with parameter h_t and compute the average value $\bar{\varepsilon}_i$ of these ε_i . Next, we refine those elements with $\varepsilon_i(t) > \bar{\varepsilon}_i$, 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. The pseudocode for this algorithm is outlined in Algorithm 2.

C. A Comparison to Conventional Techniques

There are various *a priori* error estimates used for adaptive mesh generation, which are developed in particular to estimate discretization error in the numerical solutions of partial differential equations (see [7] for a survey on error estimation procedures). These estimates are in general adopted for adaptive mesh generation in inverse problems, such as parameter estimation problems, where a partial differential equation governs the underlying forward problem.

One commonly used estimate is the interpolation error estimate [22]–[24], which depends on the smoothness of the solution, the mesh size, and the basis function (first-order Lagrangian basis in this case) used for approximation

$$\|\widetilde{\delta\mu} - \widetilde{\delta\mu}_I\|_{1,t} \leq C \|\widetilde{\delta\mu}\|_{2,t} h_t \quad (34)$$

where $\widetilde{\delta\mu}_I$ is the interpolant of $\widetilde{\delta\mu}$, and C is a constant similar as those in (17) and (21). Adoption of such error estimates 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 (21) in Theorem 2, takes into account these factors addition to the smoothness of the inverse problem solution, the mesh size and approximation basis function. In the following discussion, we compare the reliability of the *a priori* interpolation error estimates 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 (16) on a finite dimensional subspace $V_N \subset H^1(\Omega)$ spanned by a set of linear finite elements $\{\Omega_t\}$, for $t = 1, \dots, N_{\Delta}$, such that $\bigcup_t^{N_{\Delta}} \Omega_t = \Omega$. The vertices of $\{\Omega_t\}$ are at $\{r_p\}$, $p = 1, \dots, N$, and h_t is the diameter of the smallest ball that contains the t th

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

◇ Generate an initial uniform mesh
 (Δ, N_Δ) , $\Delta = \bigcup_{t=1}^{N_\Delta} \{\Delta_t\}$
 ◇ Set the maximum number of nodes N_{max}^i
while Number of nodes N less than N_{max}^i
 for each element $\Delta_t \in \Delta$ with mesh size parameter h_t
 if first linearization
 • Use current solution updates G_j^* and Φ_i and *a priori* information about $\widetilde{\delta\mu}$ to compute $\varepsilon_i(t)$ in (33)
 else
 • Use current solution updates G_j^* and Φ_i and $\widetilde{\Delta\mu}$ to compute $\varepsilon_i(t)$ in (33)
 end
 • Compute $\bar{\varepsilon}_i$
 • Refine the elements with $\varepsilon_i(t) > \bar{\varepsilon}_i$
 • Update the mesh Δ
 end
 ◇ Solve for $\widetilde{\Delta\mu}$.

element Ω_t . Further assume that, the following constraint is satisfied for the *a priori* interpolation error estimate on the linear finite elements $\{\Omega_t\}$ of the mesh for $t = 1, \dots, N_\Delta$:

$$C \|\widetilde{\delta\mu}\|_{2,t} h_t \leq \epsilon \quad (35)$$

so that

$$\|\delta\mu - \widetilde{\delta\mu}\|_1 \leq C \sum_{t=1}^{N_\Delta} \|\widetilde{\delta\mu}\|_{2,t} h_t \leq N_\Delta \epsilon. \quad (36)$$

In other words, if the mesh to discretize the inverse problem (16) is generated with every element subject to (35) which is based on conventional error estimate, then the upper bound on the $H^1(\Omega)$ norm of the error $\delta\mu - \widetilde{\delta\mu}$ is expected *not to exceed* $N_\Delta \epsilon$.

In the following, we show how a constraint on the conventional error estimate as in (35) leads to a bound on $\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ 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 high error in the reconstructed optical image $\widetilde{\Delta\mu}$.

Substituting the constraint (35) into the bound (21) in Theorem 2, we obtain a new upper bound Θ for $\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ and a lower bound on Θ can be given by (see Appendix A)

$$\begin{aligned} \Theta &\geq N_\Delta \epsilon \left(1 + \min_t h_t + \frac{\max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \right. \\ &\quad \left. \times \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right) \\ &\geq N_\Delta \epsilon. \end{aligned} \quad (37)$$

Depending on the choice of λ_1, λ_2 , the last term in brackets can dominate the lower bound on Θ , or can be neglected. In either condition, the lower bound in (37) implies that mesh generation based on conventional error estimates indicates an increase in the error bound relative to the expected bound of $N_\Delta \epsilon$, which may result in error in the reconstructed optical image higher than intended for.

D. Toy Model Adaptive Mesh

We demonstrate the refinement process 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 Fig. 1(a). The sources and detectors are indicated by the solid triangles and squares, respectively. In this toy problem, we use triangular finite element with first-order Lagrange basis functions. Using the practical adaptive mesh generation algorithm, we want to generate an adaptive mesh to solve the boundary value problem defined by the diffusion equation and the associated boundary conditions for a point source located at $(-2.5, -3)$. We constrain the number of vertices of the final mesh not to exceed 32, which defines the computational resources available at hand.

We assume that we do not have the finite element solutions of the forward problem, i.e., $\Phi_i, i = 1, \dots, N_S$ and $G_j^*, j = 1, \dots, N_D$. Thus, we use the analytical solution of the diffusion equation on an unbounded domain to approximate Φ_i and G_j^* , (see Appendix B) and use an *a priori* image model given by

$$\delta\mu_{\text{prior}}(x, y) = \mu_0 e^{-x^2 - (\frac{y}{2})^2} + 0.001 \quad (38)$$

where $\mu_0 = 0.005 \text{ cm}^{-1}$ in an approximation to $\delta\mu$ [shown in Fig. 1(d)]. Note that $\delta\mu_{\text{prior}}$ is a two-dimensional Gaussian shaped fluorophore concentration centered at $(0, 0)$. We compute the error indicators (25) on each of the 18 elements of the initial uniform mesh. In Fig. 1(a), the computed error indicator values are shown inside the corresponding finite elements. The elements with error indicator values higher than the average are selected for refinement. Using Rivara's algorithm [25], the adaptive refinement results in a mesh with 21 nodes and 28 elements, shown in Fig. 1(b). We notice that the refinement occurs around the source and around the detectors which are relatively close to the source. The error indicator values are recomputed on the refined mesh, which are shown inside the elements in Fig. 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 iteration, we see that the mesh is refined around the source, the closest detectors, and the fluorophore concentration [see Fig. 1(c)]. At this iteration, 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.

E. Computational Complexity of the Adaptive Mesh Generation Algorithms

In this subsection, we briefly discuss the computational complexity of the adaptive mesh generation algorithms described in the previous subsections. We first start with the adaptive mesh generation algorithm for the forward problem discretization,

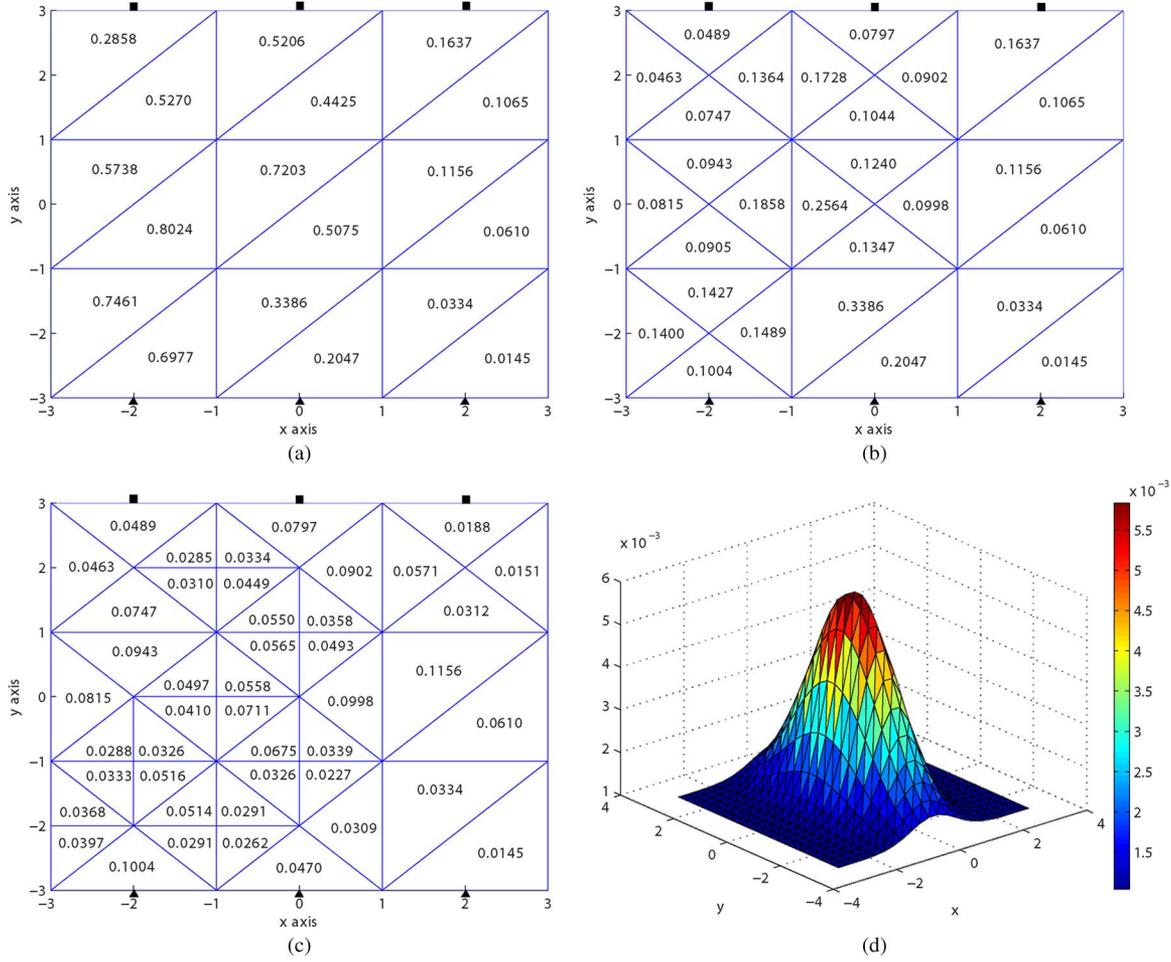


Fig. 1. The initial uniform mesh and the following adaptive meshes after the first and second refinements, respectively, generated for a point source located at $(-2.5, -3)$ and fluorophore concentration with *a priori* image model shown in the last figure. The error indicator values computed for each element based on (25) are shown inside the elements. (a) The initial uniform mesh with 16 vertices 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 $\delta\mu$.

which is followed by the adaptive mesh generation algorithm for the inverse problem discretization.

Using triangular finite elements with first order Lagrange basis functions and an analytical (exact) integration on each finite element, we assume the number of multiplications required to compute the L^2 or H^1 norm of a finite dimensional function on each triangular element is of $\mathcal{O}(1)$ complexity. Then for two finite dimensional functions multiplied with each other, it becomes $\mathcal{O}(10)$ [16]. Then, the computational complexity of computing (25) becomes $\mathcal{O}(N_D)$ for each element. For all elements, the computational complexity amounts to $\mathcal{O}(N_\Delta^i N_D)$. Similarly, one can obtain the same computational complexity for computing the error indicator (26).

It is possible to reduce the computational complexity by making the following approximations on the error indicators (25) and (26):

$$\varepsilon_f^i(n) \approx \left(2 \left\| \sum_j^{N_D} g_j^* \delta\mu \right\|_{0,ni} + \|\delta\mu\|_0 \sum_j^{N_D} \|g_j^*\|_{\infty,ni} \right) \cdot \|\phi_i\|_{1,ni} h_{ni}, \quad (39)$$

$$\varepsilon_f^j(m) \approx \left(2 \left\| \sum_i^{N_S} \phi_i \delta\mu \right\|_{0,mj} + \|\delta\mu\|_0 \sum_i^{N_S} \|\phi_i\|_{\infty,mj} \right) \cdot \|g_j^*\|_{1,mj} h_{mj}. \quad (40)$$

Then, the computational complexity of adaptive mesh generation for each source and detector reduces to $\mathcal{O}(N_\Delta^i)$ and $\mathcal{O}(N_\Delta^{*j})$, respectively.

Next, we discuss the computational complexity of adaptive mesh generation for the discretization of the inverse problem. Based on the same arguments, the computational complexity of computing (33) becomes $\mathcal{O}(N_S N_D)$. For all elements, the computational complexity amounts to $\mathcal{O}(N_\Delta N_S N_D)$. Similarly, it is possible to reduce the complexity by making the following approximation on the error indicator (33):

$$\varepsilon_i(t) \approx \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \left\| \sum_{i,j}^{N_S, N_D} G_j^* \Phi_i \right\|_{0,t} \|\delta\mu\|_{2,t} + \lambda_1 \|\delta\mu\|_{2,t} \right) h_t^2 + \lambda_2 \|\delta\mu\|_{2,t} h_t. \quad (41)$$

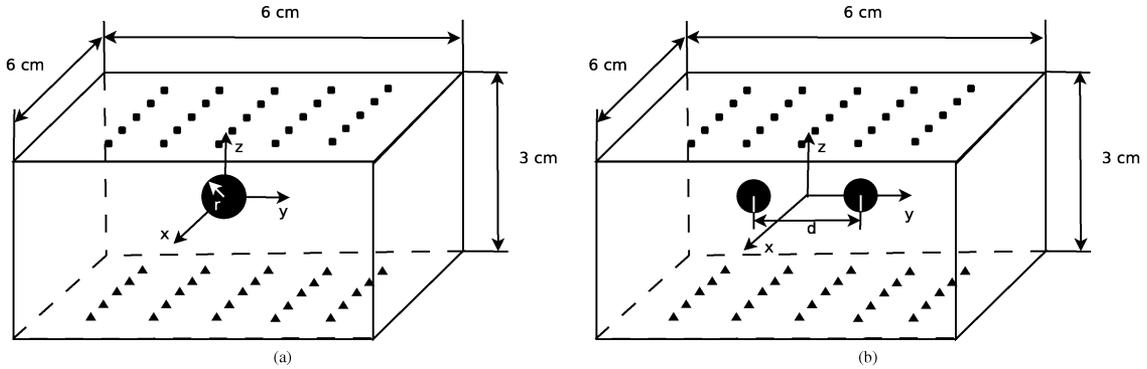


Fig. 2. The setups used for the simulation studies 1 and 2. The squares and triangles denote the detectors and sources, respectively. (a) The optical domain and source-detector configuration for simulation study 1. (b) The optical domain and source-detector configuration for simulation study 2. The radius of the circles is 2 mm.

Then, the computational complexity of adaptive mesh generation for inverse problem reduces to $\mathcal{O}(N_{\Delta})$. Note that, for conventional error estimate, the computational complexity for one sweep of the refinement is always $\mathcal{O}(N_{\Delta}^i)$, $\mathcal{O}(N_{\Delta}^{*j})$, or $\mathcal{O}(N_{\Delta})$. Although the computational resources needed to compute the conventional error estimate is less than our estimate, the resulting meshes may not be as effective as the adaptive meshes developed based on our estimate in reducing the error in reconstructed images.

IV. NUMERICAL RESULTS

To demonstrate the improvements in the reconstructed optical image quality due to our adaptive mesh generation algorithms, we performed a series of numerical experiments. Specifically, we evaluated the quality of the reconstructed images in terms of discretization error, detectability of small heterogeneities and resolution and compared our results with that of uniform and conventional adaptive meshes when the number of unknowns is constrained. We performed our simulation study using deal.II finite element C++ library [26], and used hexahedral finite elements with trilinear Lagrange basis functions to discretize both the forward and inverse problem. Note that we used the Gaussian quadrature method to evaluate the integrals in the variational problems given in (7), (8) and (20) [27]. 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.

For consistency, we considered the three-dimensional bounded domain shown in Fig. 2(a) in our numerical experiments, which we also used for numerical experiments in the Part I of this work [1]. The circular heterogeneity with radius r in the figure denotes the concentration of the fluorophore with constant absorption coefficient μ_{axf} , embedded in an optically homogeneous background with $\mu_a = 0.05 \text{ cm}^{-1}$ at both the excitation and emission wavelengths. We set the diffusion coefficient $D(\mathbf{r}) = 0.0410 \text{ cm}^{-1}$ for $\mathbf{r} \in \Omega \cup \partial\Omega$, the refractive index mismatch parameter $\rho = 3$ for the boundary, and placed 25 sources and 25 detectors evenly at the bottom and top surface of the domain. Using the parameters above, we simulated the fluorescence data by solving the coupled diffusion equations

(1) and (2) with their corresponding boundary conditions (3) and (4) on a fine uniform grid with $81 \times 81 \times 41$ nodes.

In the first set of simulations, we considered the geometry shown in Fig. 2(a). We set $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$ and simulated 5 different data sets corresponding to 5 different values of the radius r : 1, 2, 3, 4, and 5 mm, respectively. We evaluated the accuracy of the reconstructed images and the detectability of the heterogeneities with respect to their size using the adaptive meshes generated based on Theorem 1 and 2 in comparison with the uniform meshes and the conventional adaptive meshes generated based on the error estimates (36).

In the second set of simulations, we considered the geometry shown in Fig. 2(b), and set $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$, $r = 2 \text{ mm}$ for both fluorophore concentrations 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 five different distances, d , between the two heterogeneities, i.e., $d = 0.5, 0.75, 1.0, 1.25, 1.5 \text{ cm}$. We evaluated the accuracy as well as the resolution of the reconstructed images by varying the distance between the two heterogeneities when the adaptive meshes generated based on Theorem 1 and 2, and uniform and conventional adaptive meshes were used.

To obtain the solution of the variational problem, we used a fine uniform mesh with $61 \times 61 \times 31$ nodes to discretize the forward problem (1)–(4) and the inverse problem (13). We assumed that the error due to discretization in the resulting image, denoted by $\delta\mu$, is negligible with respect to the images reconstructed using the coarse meshes, and used this image as a baseline for comparison.

We used three different types of coarse meshes: Uniform mesh, the adaptive mesh generated based on the conventional error estimate, and the adaptive mesh generated based on Theorem 1 and 2, to discretize the forward and inverse problems. For the forward problem, the number of nodes for the coarse mesh ranges from 7000 to 9000; and for the inverse problem, it ranges from 2000 to 3000. For performance evaluation, we considered five different image reconstruction scenarios corresponding to five different meshes:

- 1) We used the coarse uniform mesh shown in Fig. 3(a) to discretize the forward problem and the coarse uniform mesh shown in Fig. 3(b) to discretize the inverse problem (20).

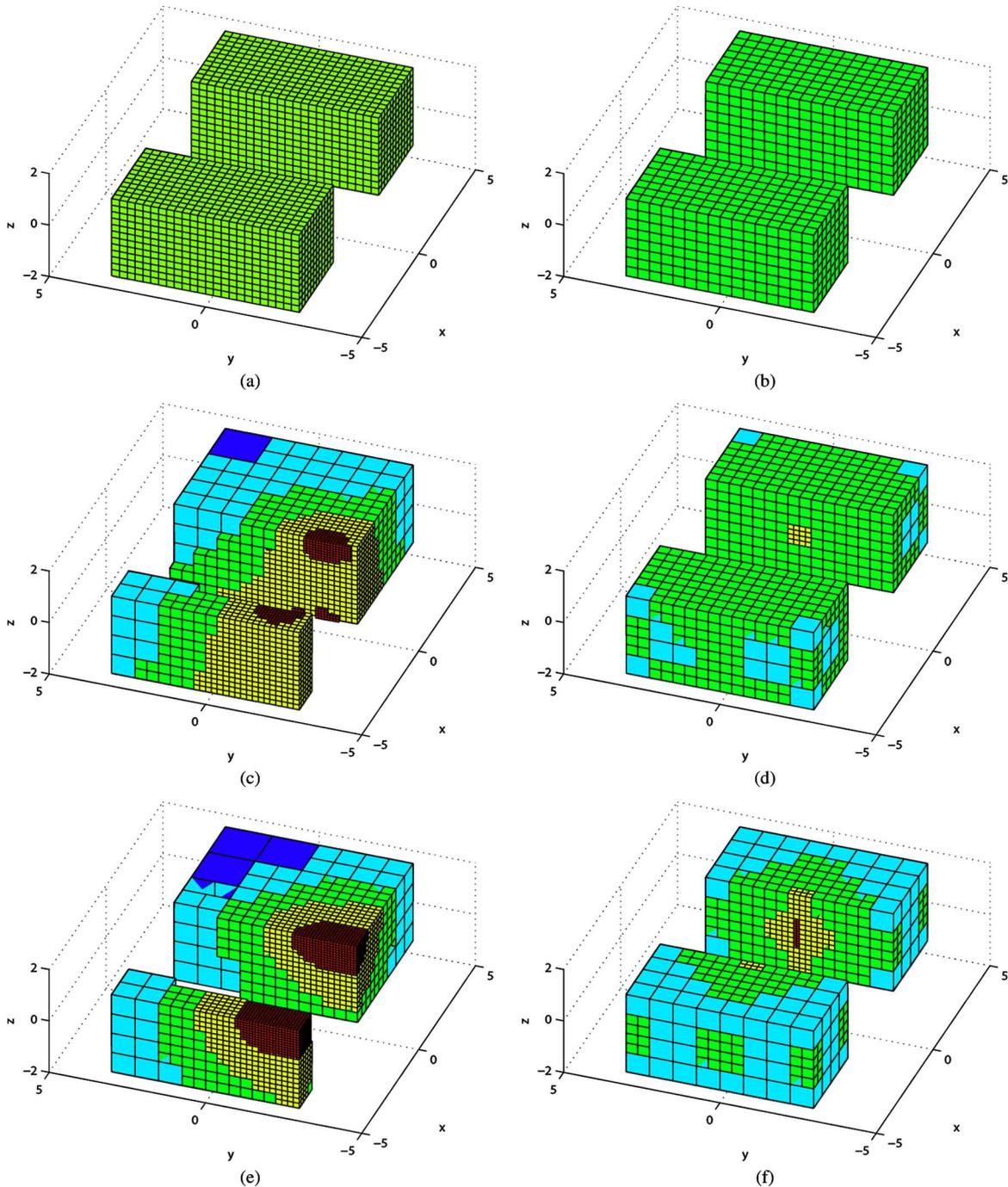


Fig. 3. Examples of meshes used in the simulation study. The mesh is cut through to show the mesh structure inside. (a) The coarse uniform mesh with 8125 nodes used to discretize the forward problem. (b) The coarse uniform mesh with 2601 nodes used to discretize the inverse problem. (c) The adaptive mesh with 8141 nodes generated for the forward problem for the detector located at $(-2.0, -2.0, 1.5)$ in simulation study 1, case 5, with $\mu_{af} = 0.015 \text{ cm}^{-1}$, $r = 5 \text{ mm}$. (d) The adaptive mesh with 2583 nodes generated for the inverse problem in simulation study 1, case 5, with $\mu_{af} = 0.015 \text{ cm}^{-1}$, $r = 5 \text{ mm}$. (e) The conventional adaptive mesh with 8756 nodes generated for the forward problem for the detector located at $(-2.0, -2.0, 1.5)$ in simulation study 1, case 5, with $\mu_{af} = 0.015 \text{ cm}^{-1}$, $r = 5 \text{ mm}$. (f) The conventional adaptive mesh with 2691 nodes generated for the inverse problem in simulation study 1, case 5, with $\mu_{af} = 0.015 \text{ cm}^{-1}$, $r = 5 \text{ mm}$.

We compared the resulting image, denoted by $\widetilde{\Delta\mu_{UU}}$, with the image reconstructed by using the adaptive meshes generated by our algorithms.

- 2) We used the adaptive meshes generated based on Theorem 1 [see Fig. 3(c)] to discretize the forward problem and the coarse uniform mesh shown in Fig. 3(b) to discretize the inverse problem. We used the resulting image, denoted by

$\widetilde{\Delta\mu_{AU}}$, to show the effectiveness of our adaptive mesh generation algorithm based on Theorem 1.

- 3) We used the adaptive meshes generated based on Theorem 1 [see Fig. 3(c)] to discretize the forward problem and the adaptive mesh generated based on Theorem 2 [see Fig. 3(d)] to discretize the inverse problem. We used the resulting image, denoted by $\widetilde{\Delta\mu_{AA}}$, to show the effective-

TABLE II
 $H^1(\Omega)$ NORM OF THE ERROR IN THE IMAGES RECONSTRUCTED BY USING DIFFERENT MESHES IN SIMULATION STUDY 1

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

TABLE III
 $H^1(\Omega)$ NORM OF THE ERROR IN THE IMAGES RECONSTRUCTED BY USING DIFFERENT MESHES IN SIMULATION STUDY 2

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

ness of our adaptive mesh generation algorithms based on Theorem 1 and 2 combined.

- 4) We used the adaptive meshes generated based on the conventional error estimates (36) [see Fig. 3(e)] to discretized the forward problem and the coarse uniform mesh shown in Fig. 3(b) to discretize the inverse problem. We compared the resulting image, denoted by $\widetilde{\Delta\mu}_{CAU}$, with the image reconstructed by using the adaptive meshes generated by our algorithms.
- 5) We used the adaptive meshes generated based on the conventional error estimates (36) to discretized the forward and inverse problems [see Fig. 3(f)]. We compared the resulting image, denoted by $\widetilde{\Delta\mu}_{CAA}$, with image reconstructed by using the adaptive meshes generated by our algorithms.

We note that the meshes we used to solve the forward and inverse problem are independent. The coarse uniform mesh used for solving the forward problem shown in Fig. 3(a) has $25 \times 25 \times 13$ nodes and the uniform mesh used for solving the inverse problem shown in Fig. 3(b) has $17 \times 17 \times 9$ nodes. For adaptive meshes, we first generated the meshes for the forward and inverse problems using Algorithm 1 and 2 described in Section III. Then, we generated the conventional adaptive meshes for the forward and inverse problems based on the error estimate (36).

Fig. 3(c) and (e) shows examples of our adaptive mesh and the conventional adaptive mesh generated for the forward problem when the detector is located at $(-2.0, -2.0, 1.5)$. We observe that our algorithm refines the mesh around the designated detector at $(-2.0, -2.0, 1.5)$, fluorophore concentration, as well as some sources close to this detector on the opposite side of the phantom. However, the conventional method only refines the mesh near the designated detector. Fig. 3(d) and (f) show examples of our adaptive mesh and the conventional adaptive mesh generated for the inverse problem when $\mu_{axf} = 0.015 \text{ cm}^{-1}$ and $r = 1 \text{ mm}$. We observe that the conventional method refines the mesh only based on the fluorophore concentration. Our algorithm, on the other hand, performs the refinement based on

TABLE IV
 SBR OF THE IMAGES RECONSTRUCTED BY USING DIFFERENT MESHES IN SIMULATION STUDY 1

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

TABLE V
 FWHM OF THE IMAGES RECONSTRUCTED BY USING DIFFERENT MESHES IN SIMULATION STUDY 1

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

the value of fluorophore concentration as well as the position of the fluorophore with respect to both sources and detectors.

Finally, we note that for each scenario, we solved the forward problem on the corresponding uniform mesh or adaptive meshes with the same parameter values, ρ , D , μ_a , used in the data generation. For the inverse problem, 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, appropriate values for the regularization parameters were empirically selected as $\lambda_1 = 1 \times 10^{-8}$ and $\lambda_2 = 1 \times 10^{-8}$. To eliminate bias due to regularization from our performance evaluation, we compare the reconstructed images using different meshing schemes with the solution of the variational problem solved on very fine meshes while evaluating the discretization errors.

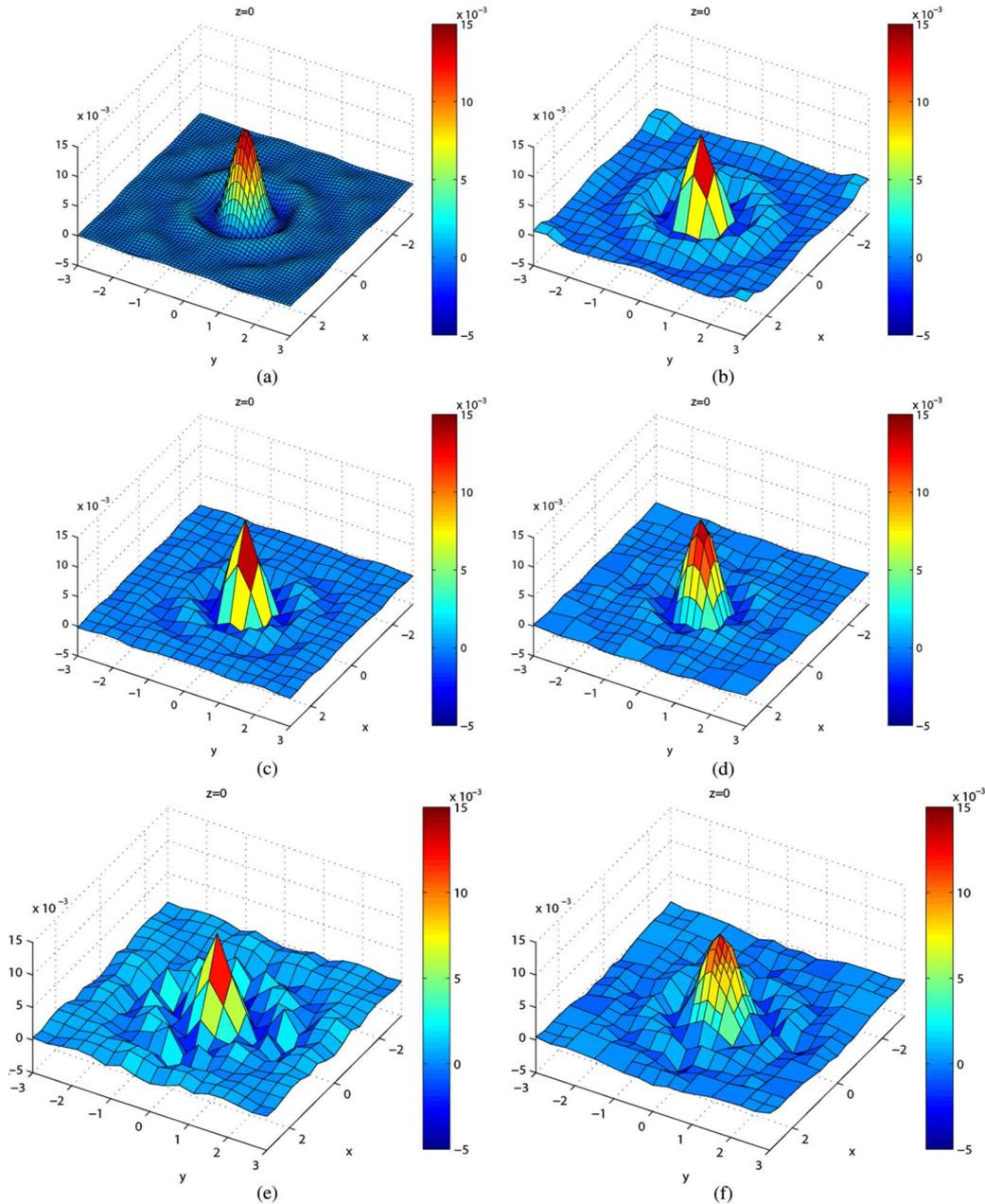


Fig. 4. The reconstruction results of simulation study 1, case 1, on plane $z = 0$, with $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$, $r = 1 \text{ mm}$. (a) The baseline image. (b) The image reconstructed using the uniform meshes for both forward and inverse problems. (c) The image reconstructed using 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.

In the following subsections, we discuss the simulation results and compare our adaptive meshing algorithm with the uniform and conventional meshing schemes in terms of discretization accuracy, detectability, and resolution.

A. Discretization Accuracy

For both sets of simulations, we summarized the $H^1(\Omega)$ norm of the error in the reconstructed images with respect to the baseline image in Tables II and III. Additionally, we tabulated the percentage of error as compared to the error in the image reconstructed by using the coarse and uniform meshes in the forward and inverse problems. In Tables II and III the left column under

each case is the absolute error value, and the right column under each case is the percentage of error. The results show that the error in the images reconstructed by using the uniform meshes for both the forward and inverse problems is significantly reduced when the adaptive meshes generated by our algorithm are used. We note that the adaptive meshes generated by our algorithm reduces the error up to around 60% when the radius of heterogeneity is close to 1 mm, and the error increases as the size of heterogeneity increases, but the total error reduction is roughly around 50%. On the other hand, the conventional adaptive mesh generation algorithm even increases the error by roughly 5% when the size of the heterogeneity is small, and provides only

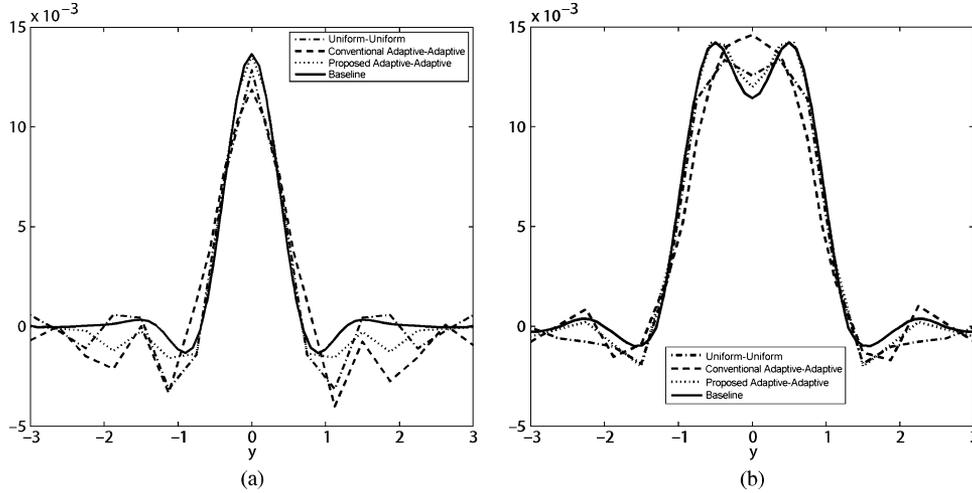


Fig. 5. The reconstructed images $\delta\mu$, $\widetilde{\Delta\mu}_{UU}$, $\widetilde{\Delta\mu}_{AA}$ and $\widetilde{\Delta\mu}_{CAA}$ along the y -axis in simulation study 1, case 1 and simulation study 2, case 3. (a) The reconstructed images along the y -axis, in simulation 1, case 1, with $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$, $r = 1 \text{ mm}$. (b) The reconstructed images along the y -axis, in simulation 2, case 3, with $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$, $r = 2 \text{ mm}$, $d = 1 \text{ cm}$.

about 15% error reduction when the radius of the heterogeneity increases to 5 mm.

B. 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 define

$$\text{SBR}(\widetilde{\Delta\mu}) = 20 \log_{10} \left(\frac{\left(\|\widetilde{\Delta\mu}\|_{L^2(\Omega_F)}^2 / V_{\Omega_F} \right)}{\left(\|\widetilde{\Delta\mu}\|_{L^2(\Omega_B)}^2 / V_{\Omega_B} \right)} \right) \quad (42)$$

where Ω_F and Ω_B denote the target and background regions and 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 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 FWHM along the x - and y -axis.

For the first set of the simulations, we summarized the SBR and FWHM of the reconstructed images in Tables IV and V. The results show that the images reconstructed using the adaptive meshing algorithms based on Theorem 1 and 2 have higher SBR and lower FWHM than the ones reconstructed using the uniform and conventional meshing schemes.

Fig. 4 shows the cross sections of the reconstructed images on $z = 0$ plane for the fluorophore concentration with 1 mm radius. The cross section of the baseline image, $\delta\mu$, which was used to compute the error values in Table II, is shown in Fig. 4(a). Fig. 4(b), (c), and (e) shows the cross sections of the images $\widetilde{\Delta\mu}_{UU}$, $\widetilde{\Delta\mu}_{AU}$, and $\widetilde{\Delta\mu}_{CAU}$, reconstructed when the coarse

TABLE VI
PVR OF THE IMAGES RECONSTRUCTED BY USING DIFFERENT MESHES IN SIMULATION STUDY 2

r :	Case 1	Case 2	Case 3	Case 4	Case 5
	0.5 cm	0.75 cm	1 cm	1.25 cm	1.5 cm
$\delta\mu$:	-	-	1.227	1.957	3.638
$\widetilde{\Delta\mu}_{UU}$:	-	-	1.071	1.617	3.021
$\widetilde{\Delta\mu}_{AU}$:	-	-	1.136	1.818	3.529
$\widetilde{\Delta\mu}_{AA}$:	-	-	1.136	1.819	3.533
$\widetilde{\Delta\mu}_{CAU}$:	-	-	-	1.175	1.723
$\widetilde{\Delta\mu}_{CAA}$:	-	-	-	1.184	1.730

uniform mesh was used in solving the inverse problem. We see that the absorption coefficient of fluorophore concentration were reconstructed in a pyramid shape in all images. Since in this case the actual size of the fluorophore concentration 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 concentration. Fig. 4(d) and (f) shows the images of $\widetilde{\Delta\mu}_{AA}$ and $\widetilde{\Delta\mu}_{CAA}$, reconstructed using the adaptive meshes for both the forward and inverse problems. Since our mesh generation algorithm and the conventional method both adaptively refined the inverse problem mesh around the fluorophore concentration, its shape is better resolved than that in $\widetilde{\Delta\mu}_{UU}$, $\widetilde{\Delta\mu}_{AU}$, and $\widetilde{\Delta\mu}_{CAU}$. Furthermore, comparing all the images reconstructed using the coarse meshes, we observe that the cross sections of $\widetilde{\Delta\mu}_{AU}$ and $\widetilde{\Delta\mu}_{AA}$ have the largest SBR as compared to the other images, which is consistent with the results in Tables IV and V.

Fig. 5(a) shows the reconstructed images of $\delta\mu$, $\widetilde{\Delta\mu}_{UU}$, $\widetilde{\Delta\mu}_{AA}$ and $\widetilde{\Delta\mu}_{CAA}$ along the y -axis. The solid line in Fig. 5(a) shows the baseline image $\delta\mu$ which is assumed to have no or negligible error. We observe that the image, $\widetilde{\Delta\mu}_{AA}$, is the closest one to $\delta\mu$ among all three reconstructed images. In particular, we observe that this image has higher response at the center of the fluorophore concentration and lower sidelobe magnitude at the background region, as compared to those reconstructed using the conventional adaptive meshes and uniform meshes. This indicates that our adaptive mesh generation algorithms can effectively improve the detectability of small targets as compared to the other meshing schemes.

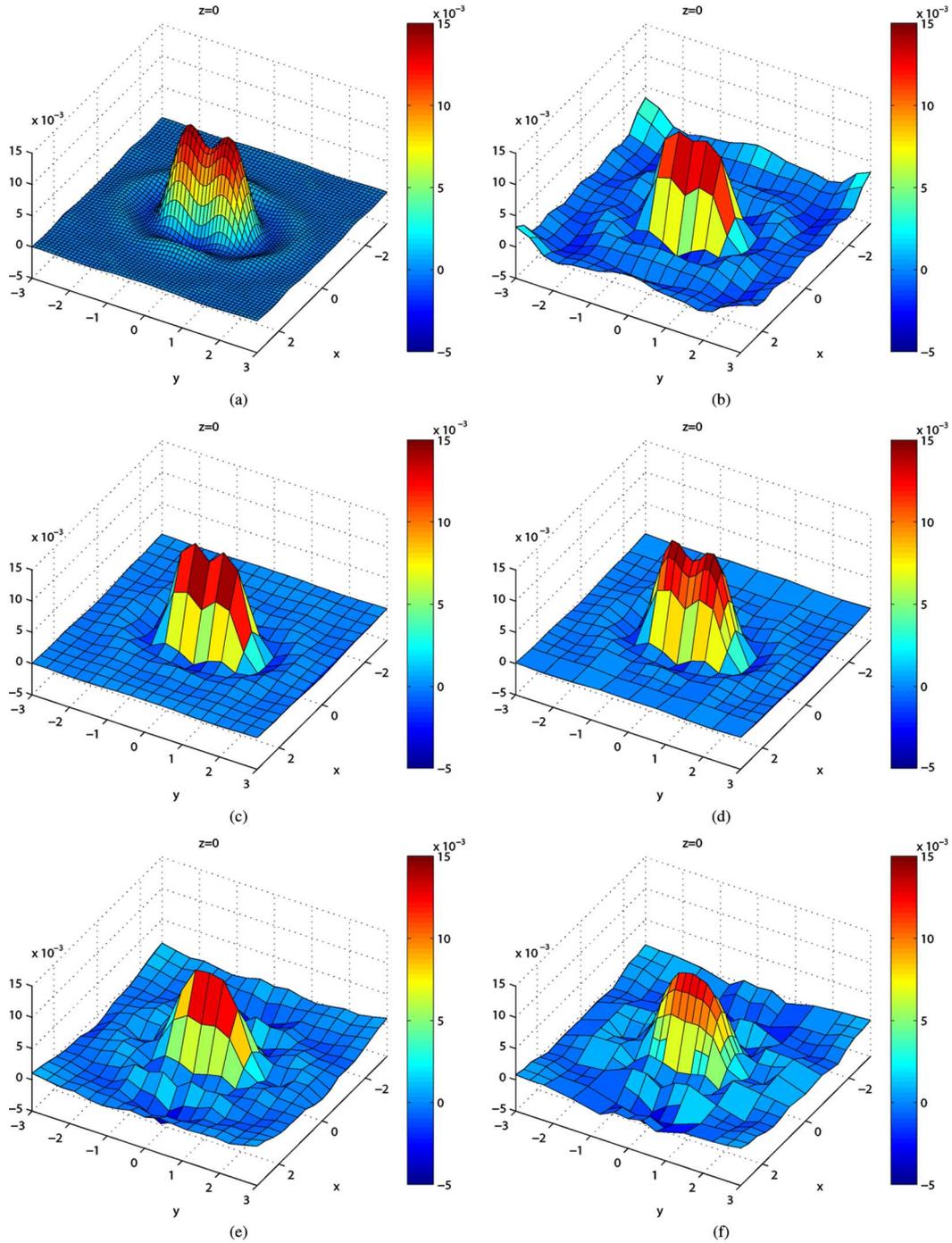


Fig. 6. The reconstruction results of simulation study 2, case 3, on plane $z = 0$, with $\mu_{\text{axf}} = 0.015 \text{ cm}^{-1}$, $r = 2 \text{ mm}$, $d = 1 \text{ cm}$. (a) The baseline image. (b) The image reconstructed using the uniform meshes for both forward and inverse problems. (c) The image 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.

C. 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 absorption coefficient 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 VI. We observe that the two fluorophore con-

centrations 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 the two heterogeneities 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 concentrations are distinguishable, the quantitative results show that those reconstructed using our adaptive meshing algorithms have higher PVR than the ones reconstructed using the uniform and conventional meshing schemes.

In Fig. 6, we present the cross sections of the reconstructed images on $z = 0$ plane when the distance between the two fluorophores is 1 cm. Fig. 6(a) shows the cross section of the baseline image, $\delta\mu$, where the two fluorophore concentrations are clearly distinguishable. In Fig. 6(b)–(d), we can also clearly see the two peaks of the absorption coefficient of the fluorophore concentration, while in the cross sections of $\widetilde{\Delta\mu}_{CAU}$ and $\widetilde{\Delta\mu}_{CAA}$ shown in Fig. 6(e) and (f), these two peaks appear merged together.

For a close examination, in Fig. 5(b), we also present the reconstructed images along the y -axis on $z = 0$ plane. The cross section of $\Delta\mu_{CAA}$ reveals that the conventional adaptive meshing method fails to produce an image that can resolve the two distinct fluorophore concentrations [see dashed line in Fig. 5(b)]. On the other hand, $\widetilde{\Delta\mu}_{AA}$ reconstructed by using the adaptive meshes based on our approach, produces two distinguishable peaks of the absorption coefficient of the fluorophore concentration with higher PVR than that of $\Delta\mu_{UU}$ reconstructed using uniform meshes. This indicates that our adaptive mesh generation algorithm can effectively improve the resolution of the reconstructed images as compared to other meshing schemes.

V. CONCLUSION

Following the analysis in [1], in this work, we developed two new adaptive mesh generation algorithms for the numerical solutions of the forward and inverse problems in FDOT. These adaptive mesh generation algorithms aim to reduce the error in the reconstructed optical images while constraining the number of vertices in the adaptive meshes (i.e., the number of unknowns in the discretized forward and inverse problems) to a predefined number. The key feature that distinguishes these algorithms from the existing conventional adaptive mesh generation algorithms is the fact that they take into account the interdependence between the solutions of the forward and inverse problems, the source-detector configuration, the location of the fluorophore concentration with respect to the sources and detectors as well as the formulation of the inverse problem. In addition, we analytically showed that the conventional adaptive mesh generation algorithms relying on the interpolation error estimates may lead to higher error in the reconstructed images than expected.

We also showed that the computational complexity of the our adaptive mesh generation algorithms can be reduced from $\mathcal{O}(N_{\Delta}^i N_D)$ or $\mathcal{O}(N_{\Delta}^i N_S)$ to $\mathcal{O}(N_{\Delta}^i)$ or $\mathcal{O}(N_{\Delta}^j)$, for the forward problem. For the inverse problem, it can be reduced from $\mathcal{O}(N_{\Delta} N_S N_D)$ to $\mathcal{O}(N_{\Delta})$. Hence, the cost of the adaptive mesh generation is negligible as compared to the cost of the matrix inversion involved in the numerical solutions of the forward and inverse problems on the generated adaptive meshes.

Our numerical experiments show that using the adaptive meshes generated by the conventional adaptive meshing method which do not include domain specific factors may lead to higher errors in reconstructed images, especially when the fluorophore

concentration is small. However, given a limited number of unknowns in the discrete forward and inverse problems, our mesh generation algorithms can significantly improve the accuracy of the reconstructed images in terms of the improved detectability of small fluorescent targets as well as the improved image resolution, as compared with uniform and conventional adaptive meshing method.

From the perspective of mean-square-error estimation, the regularization parameters provide a trade-off between the bias and the variance of reconstructed images [28], [29]. It is well-known that the bias due to regularization can sometimes overwhelm the error due to other sources in the reconstructed image. Our discretization error estimates given in Theorems 1 and 2 are functions of the regularization parameters and are given with respect to the solution of the variational formulation of the inverse problem. In the development of adaptive meshing schemes, we assume that the appropriate regularization parameters for the variational formulation of the inverse problem are determined *a priori* and therefore keep these parameters fixed for different meshing schemes. However, since our error estimates involve regularization parameters, it is possible to study the interplay between the choice of regularization parameters and the mesh size and adaptively select both in an attempt to reduce the overall error in reconstructed images. We leave for the future the exploration of the interplay between the adaptive selection of regularization parameters and the mesh size.

In this paper, as well as in our previous publications [16], [22], we introduced an approach to develop discretization error estimates and adaptive meshing algorithms that take into account domain-specific factors as well as the interdependence of the solutions of the forward and inverse problems. This approach is not limited to single frequency FDOT imaging system, and can be extended to multiple frequency and time-domain optical imaging systems. Additionally, a similar approach can also be utilized in developing discretization error estimates and adaptive meshing algorithms for other inverse coefficient estimation problems involving PDEs, such as electrical impedance tomography and microwave tomography. However, specific discretization error estimates and adaptive meshing algorithms may be different in different inverse coefficient estimation problems depending on the particular PDE and the inverse problem formulation. Furthermore, our framework also provides an approach which can be used to analyze other types of errors involved in diffuse optical tomographic imaging. For example, the error due to Born approximation, the error due to background heterogeneities (incorrect assumptions about the background optical properties), and the error in reconstructed images due to variations in the source and detector magnitudes and positions.

APPENDIX A

LOWER BOUND ON Θ

Substituting the constraint (35) into the bound (21), we get

$$\begin{aligned} \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \\ &\times \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \epsilon \sum_t \sum_{i,j}^{N_{\Delta}, N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right. \\ &\quad \left. + \lambda_1 \sum_t \epsilon h_t + \lambda_2 \sum_t \epsilon \right) \\ &:= \Theta. \end{aligned} \quad (43)$$

Next, we compute a lower bound for Θ in (43).

$$\begin{aligned}
 \Theta &= \frac{1}{\min(\lambda_1, \lambda_2)} \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \epsilon \sum_t^{N_\Delta} \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right. \\
 &\quad \left. + \lambda_1 \sum_t^{N_\Delta} \epsilon h_t + \lambda_2 \sum_t^{N_\Delta} \epsilon \right) \\
 &\geq \frac{N_\Delta \epsilon \max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \\
 &\quad + \frac{\lambda_1}{\min(\lambda_1, \lambda_2)} N_\Delta \epsilon \min_t h_t + \frac{\lambda_2}{\min(\lambda_1, \lambda_2)} N_\Delta \epsilon \\
 &\geq N_\Delta \epsilon \left(1 + \min_t h_t + \frac{\max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \right. \\
 &\quad \left. \times \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right). \quad (44)
 \end{aligned}$$

APPENDIX B

SOLUTION OF THE MODEL PROBLEM

In order to initialize the adaptive mesh for the solution of the forward problem (provided $D(\mathbf{x}) = \bar{D}$ and $\mu_a(\mathbf{x}) = \bar{\mu}_a$ are spatially constant), we use an analytical solution to compute the approximations of ϕ_i and g_j^* . Below, we give the solution in 2-D for the forward problem (1) [16]

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

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 [30], and $k_\Omega = \sqrt{-K_\Omega^2}$. Under the same conditions, an analytical solution for the adjoint problem (5) can be obtained in a similar way. The solution of the problem in 3-D can be derived in a similar manner [31], [32].

REFERENCES

- [1] M. Guven, L. Reilly-Raska, L. Zhou, and B. Yazici, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography: Part I," *IEEE Trans. Medical Imaging*, to be published.
- [2] M. Ainsworth and J. T. Oden, "A unified approach to a posteriori error estimation using elemental residual methods," *Numerische Mathematik*, vol. 65, pp. 23–50, 1993.
- [3] I. Babuška and W. C. Rheinboldt, "Error estimates for adaptive finite element computations," *SIAM Journal on Numerical Analysis*, vol. 15, pp. 736–754, 1978.
- [4] 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.
- [5] R. E. Bank and A. Weiser, "Some a posteriori error estimators for elliptic partial differential equations," *Mathematics of Computation*, vol. 44, pp. 283–301, 1985.
- [6] 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*, vol. 97, pp. 399–436, 1992.

- [7] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*. : Teubner-Wiley, 1996.
- [8] J. G. Brankov, Y. Yang, and M. N. Wernick, "Tomographic image reconstruction based on a content-adaptive mesh model," *IEEE Trans. Medical Imaging*, vol. 23, no. 2, pp. 202–212, 2004.
- [9] M. Molinari, S. J. Cox, B. H. Blott, and G. J. Daniell, "Adaptive mesh refinement techniques for electrical impedance tomography," *Physiological Measurement*, vol. 22, pp. 91–96, 2001.
- [10] L. Beilina and C. Johnson, "A posteriori error estimation in computational inverse scattering," *Mathematical Models and Methods in Applied Sciences*, vol. 15, pp. 23–37, 2005.
- [11] L. Beilina and C. Johnson, "Adaptive finite element/difference method for inverse elastic scattering waves," *Applied and Computational Mathematics*, vol. 2, pp. 158–174, 2003.
- [12] M. Torregrossa, C. V. Zint, and P. Poulet, "Image reconstruction in optical tomography: Mesh influence," in *IV International Workshop, 2002, Computational Problems of Electrical Engineering*, pp. 183–186.
- [13] 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, 2003, vol. 2, Proc. of IEEE*, pp. 823–826.
- [14] X. Gu, Y. Xu, and H. Jiang, "Mesh-based enhancement schemes in diffuse optical tomography," *Medical Physics*, vol. 30, no. 5, pp. 861–869, 2003.
- [15] M. Huang and Q. Zhu, "Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information," *Applied Optics*, vol. 43, no. 8, pp. 1654–1662, 2004.
- [16] 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*, vol. 23, pp. 1135–1160, 2007.
- [17] 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*, vol. 22, pp. 175–195, 2006.
- [18] W. Bangerth and A. Joshi, "Adaptive finite element methods for the solution of inverse problems in optical tomography," *Inverse Problems*, vol. 24, p. 034011, 2008.
- [19] 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," *Optics Express*, vol. 15, no. 11, p. 6955C6975, 2007.
- [20] W. Bangerth, "Adaptive Finite Element Methods for the Identification of Distributed Parameters in Partial Differential Equations," Ph.D. dissertation, University of Heidelberg, , 2002.
- [21] 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*, vol. 227, pp. 5778–5798, 2008.
- [22] 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*, vol. 23, pp. 1115–1133, 2007.
- [23] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*. : Springer Verlag, 2002.
- [24] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation*. Orlando, FL: Academic Press, 1984.
- [25] M. C. Rivara, "Mesh refinement processes based on the generalized bisection of simplices," *SIAM Journal on Numerical Analysis*, vol. 21, no. 3, pp. 604–613, 1984.
- [26] W. Bangerth, R. Hartmann, and G. Kanschat, "deal.II—A general-purpose object-oriented finite element library," *ACM Trans. Math. Softw.*, vol. 33, no. 4, 2007.
- [27] I. M. Smith and D. V. Griffiths, *Programming the Finite Element Method*, 4th ed. : John Wiley & Sons, 2004.
- [28] T. A. Johansen, "On tikhonov regularization, bias and variance in nonlinear system identification," *Automatica*, vol. 33, pp. 441–446, 1997.
- [29] 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 Trans. Image Processing*, vol. 1, pp. 322–336, 1992.
- [30] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. : Dover, 1968.
- [31] J. D. Jackson, *Classical Electrodynamics*. : John Wiley and Sons, 1962.
- [32] S. Walker, D. A. Boas, and E. Gratton, "Photon density waves scattered from cylindrical inhomogeneities: Theory and experiments," *Applied Optics*, vol. 37, no. 10, pp. 1935–1944, 1998.