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

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

***Abstract*—For imaging problems in which numerical solutions need to be computed for both the inverse and the underlying forward problems, discretization can be a major factor that determines the accuracy of imaging. In this work, we analyze the effect of discretization on the accuracy of fluorescence diffuse optical tomography. We model the forward problem by a pair of diffusion equations at the excitation and emission wavelengths and consider a finite element discretization method for the numerical solution of the forward problem. For the inverse problem, we use an optimization framework which allows incorporation of *a priori* information in the form of zeroth- and first-order Tikhonov regularization terms. Next, we convert the inverse problem into a variational problem and use Galerkin projection to discretize the inverse problem. Following the discretization, we analyze the error in reconstructed images due to the discretization of the forward and inverse problems and present two theorems which point out the factors that may lead to high error such as the mutual dependence of the forward and inverse problems, the number of sources and detectors, their configuration and their positions with respect to fluorophore concentration, and the formulation of the inverse problem. Finally, we demonstrate the results and implications of our error analysis by numerical experiments. In the second part of the paper, we apply our results to design novel adaptive discretization algorithms.**

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

I. Introduction

THERE HAS been a growing interest in fluorescence imaging due to its potential for the characterization of biological processes at cellular and molecular levels [1]–[4]. In particular, fluorescence diffuse optical tomography (FDOT) offers the quantification, 3-D imaging, and depth retrieval of fluorescence activity with high sensitivity, which can be used for functional and molecular characterization of normal and diseased tissues [5]–[7].

Like its analogue diffuse optical tomography, (see [8] and [9], and the references therein), FDOT poses a computationally intense imaging problem. This stems from the necessity of solving a forward problem comprised of a pair of coupled partial differential equations (PDEs) and an ill-posed integral equation resulting from the linearization of the inverse problem whose formulation relies on the solutions of the forward problem. In general, no closed form solutions can be computed except for specific domain geometries and optical medium properties [10], [11]. Hence, one has to consider numerical solutions of the forward and inverse problems. However, numerical solutions are merely approximations to the actual solutions and they possess error as a result of the discretization involved in the process of solving them. Due to the interdependence of the forward and inverse problems, discretization of each problem typically leads to error in the reconstructed images. While the effect of inverse problem discretization on the imaging accuracy can be deduced rather intuitively [12], errors due to the discretization error in the forward problem solutions can also result in severe imaging artifacts in optical tomography [12], [13].

There is a vast degree of work on the estimation and analysis of discretization error in the solutions of partial differential equations (PDEs) [14]–[19]. A different approach is followed in [20], [21] in which error in quantities of interest is related to the discretization of the second-order elliptic PDEs. In the area of parameter estimation problems governed by PDEs, relatively little has been published. See for example [22], [23] for an *a posteriori* error estimate for the Lagrangian in the inverse scattering problem for the time-dependent acoustic wave equation and [24] for a similar approach, and [25] for *a posteriori* error estimates for distributed elliptic optimal control problems. In the area of DOT, it was numerically shown that the approximation errors resulting from the discretization of the forward problem can lead to significant errors in the reconstructed optical images [13]. In [26], we presented an approach for the analysis of the error in reconstructed optical absorption images due to discretization, which led to the development of new adaptive mesh generation techniques [12]. In these studies [12], [26], the ill-posed nature of the inverse problem was addressed by zeroth-order Tikhonov regularization [27] and a collocation method was used for the discretization of the inverse problem.

Manuscript received June 07, 2009; revised July 31, 2009; accepted August 25, 2009. Current version published February 03, 2010. This work was supported in part by U.S. Army Medical Research under Grant 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, Santa Clara, CA 95054 USA (e-mail: guven@rpi.edu).

L. Reilly-Raska is with Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02421 USA (e-mail: reilll@rpi.edu).

L. Zhou is with the Departments of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: zhoul2@rpi.edu).

*B. Yazıcı is with the Departments of Electrical, Computer and Systems Engineering and Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 02180 USA (e-mail: 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.2009.2031492

The main premise in the first part of this work is to analyze the error in fluorescence optical imaging due to discretization. In this respect, we identify the key factors specific to the imaging problem that show how discretization of the forward and inverse problems impacts the accuracy of the reconstructed optical image. In particular, we focus on the estimation of the fluorophore absorption coefficient in a bounded optical domain where the light propagation at the excitation and emission wavelengths is modeled by a pair of coupled frequency-domain diffusion equations with appropriate boundary conditions. We consider an iterative linearization algorithm based on first-order Fréchet derivatives to address the nonlinearity of the inverse problem, and use the gradient descent algorithm to solve the resulting optimization problem. Following each linearization step, we formulate the inverse problem in the optimization framework which enables us to incorporate *a priori* information in the form of zeroth- and first-order Tikhonov regularization terms. For the discretization of the forward problem, we use finite elements with first-order Lagrange basis functions. Before we discretize the inverse problem, we first convert the optimization problem into a variational formulation which can be regarded as a boundary value problem with the assignment of suitable boundary conditions. Then we use projection by Galerkin method with first-order Lagrange basis functions to discretize the resulting inverse problem. Following the discretization of the forward and inverse problems, we derive two new error estimates which show respectively the effect of forward and inverse problem discretizations on the accuracy of reconstructed optical absorption coefficient of the fluorophore. Next we discuss the major implications established by these estimates.

The error analysis presented in this work motivates the development of novel adaptive discretization schemes. In the second part of this work, we propose two novel adaptive mesh generation algorithms for the discretization of the forward and inverse problems [28], and discuss and validate the potential computational savings and improvements in the accuracy of optical fluorescence imaging.

The outline of this paper is as follows. We describe the forward model in the next section. In Section III, we state the inverse problem formulation under iterative linearization and discuss regularization to address ill-posedness. We next pose the inverse problem as an optimization problem incorporating *a priori* information and derive the variational formulation from this optimization problem. Section IV details the discretization of forward and inverse problems, respectively. In Section V, we present the theorems which describe the error in fluorescence imaging arising from these discretizations. In Section VI, we demonstrate the dependency of the error to a number of specific factors in numerical experiments. Finally, in Section VII, we conclude our discussion.

II. Forward Problem

A. Notational Conventions

In this paper, we denote operators by capital cursive Latin letters ($\mathcal{A}$) and matrices by bold capital Latin letters ($\mathbf{A}$). 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 $\boldsymbol{r}$, $\boldsymbol{\Gamma}$. Table I provides a summary of key variables and function spaces and norms used throughout the paper.

TABLE I
Definition of Function Spaces and Norms

Notation	Explanation
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \vert \text{ess sup}_\Omega \vert f(\mathbf{x})\vert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f\vert\ (\int_\Omega \vert f(\mathbf{x})\vert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1,\infty)$
$H^p(\Omega)$	$H^p(\Omega) = \{f\vert\ (\sum_{\vert z\vert \le p} \Vert D_w^z f\Vert_0^2)^{1/2} < \infty\ \}$, $p \in [1,\infty)$
$\Vert f\Vert_0$	The $L^2(\Omega)$ norm of f
$\Vert f\Vert_p$	The $H^p(\Omega)$ norm of f
$\Vert f\Vert_{p*}$	The dual norm of f in the dual space $H^{p*}(\Omega)$
$\Vert f\Vert_\infty$	The $L^\infty(\Omega)$ norm of f
$\Vert f\Vert_{0,m}$	The L^2 norm of f over the mth finite element Ω_m
$\Vert f\Vert_{p,m}$	The H^p norm of f over the mth finite element Ω_m
$\Vert f\Vert_{\infty,m}$	The L^∞ norm of f over the mth finite element Ω_m

B. Forward Problem Derivation

We start with the coupled diffusion equations which describe the light transport in a fluorescent medium of a bounded domain $\Omega \subset R^3$ with Lipschitz boundary $\partial\Omega$

$$-\nabla \cdot D_x(\boldsymbol{r})\nabla\phi_x(\boldsymbol{r},\omega) + \left(\mu_{ax}(\boldsymbol{r}) + \frac{j\omega}{c}\right)\phi_x(\boldsymbol{r},\omega) = S_i(\boldsymbol{r},\omega) \tag{1}$$

$$-\nabla \cdot D_m(\boldsymbol{r})\nabla\phi_m(\boldsymbol{r},\omega) + \left(\mu_{am}(\boldsymbol{r}) + \frac{j\omega}{c}\right)\phi_m(r,\omega) = \phi_x(\boldsymbol{r})\eta\mu_{\text{axf}}(\boldsymbol{r})\frac{1-j\omega\tau(\boldsymbol{r})}{1+(\omega\tau(\boldsymbol{r}))^2} \tag{2}$$

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 optical fields, $D_{x,m}$ represents the isotropic diffusion coefficients, and S_i is the ith excitation source. We assume that the diffusion coefficients are known and they are identical at both the excitation and emission wavelengths in the closed domain; this implies $D(\boldsymbol{r}) := D_x(\boldsymbol{r}) = D_m(\boldsymbol{r})$, $\boldsymbol{r} \in \Omega \bigcup \partial\Omega$. The quantum efficiency is denoted by η; μ_{axf} is the absorption coefficient of the fluorophore and τ is the lifetime of the fluorophore. For the sake of exposition, we make the following simplifying assumption that the frequency $\omega = 0$. Subsequent developments can be extended to include multiple frequencies where τ is known. The quantities μ_{ax} and μ_{am} represent the absorption coefficient of the medium at the excitation and emission wavelengths, respectively. Typically these are represented as

$$\mu_{ax}(\boldsymbol{r}) = \mu_{axe}(\boldsymbol{r}) + \mu_{\text{axf}}(\boldsymbol{r}) \tag{3}$$

$$\mu_{am}(\boldsymbol{r}) = \mu_{ame}(\boldsymbol{r}) + \mu_{amf}(\boldsymbol{r}) \tag{4}$$

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 nonnegative and bounded on Ω.

Let N_S be the number of point sources at position $\boldsymbol{r}_i$ for $i = 1, \ldots, N_S$. Based on the assumptions stated above, we use

the following boundary value problem to model NIR light propagation at the excitation wavelength due to the ith source at $\boldsymbol{r}_i$

$$-\nabla\cdot D(\boldsymbol{r})\nabla\phi_x(\boldsymbol{r},\boldsymbol{r}_i)+\mu_{ax}(\boldsymbol{r})\phi_x(\boldsymbol{r},\boldsymbol{r}_i)=S_i(\boldsymbol{r}) \quad (5)$$

$$-\nabla\cdot D(\boldsymbol{r})\nabla\phi_m(\boldsymbol{r},\boldsymbol{r}_i)+\mu_{am}(\boldsymbol{r})\phi_m(\boldsymbol{r},\boldsymbol{r}_i) =\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\mu_{\mathrm{axf}}(\boldsymbol{r}) \quad (6)$$

where $\boldsymbol{r}\in\Omega$. The Robin-type boundary conditions are

$$2D(\boldsymbol{r})\frac{\partial\phi_x(\boldsymbol{r},\boldsymbol{r}_i)}{\partial n}+\rho\phi_x(\boldsymbol{r},\boldsymbol{r}_i)=0 \quad (7)$$

$$2D(\boldsymbol{r})\frac{\partial\phi_m(\boldsymbol{r},\boldsymbol{r}_i)}{\partial n}+\rho\phi_m(\boldsymbol{r},\boldsymbol{r}_i)=0 \quad (8)$$

where $\boldsymbol{r}\in\partial\Omega$, 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 this work, S_i represents the ith point source which is modeled by a Gaussian function centered at source position $\boldsymbol{r}_i$ [26].

In order to simplify the analysis of later sections, we make use of the adjoint problem associated with (6) and (8). Let N_D be the number of detectors. Then, for a detector located at $\boldsymbol{r}_j\in\partial\Omega$, $j=1,\ldots,N_D$

$$-\nabla\cdot D(\boldsymbol{r})\nabla g_m^*(\boldsymbol{r},\boldsymbol{r}_j)+\mu_{am}(\boldsymbol{r})g_m^*(\boldsymbol{r},\boldsymbol{r}_j)=0,\ \boldsymbol{r}\in\Omega \quad (9)$$

$$2D(\boldsymbol{r})\frac{\partial g_m^*(\boldsymbol{r},\boldsymbol{r}_j)}{\partial n}+\rho g_m^*(\boldsymbol{r},\boldsymbol{r}_j)=S_j^*(\boldsymbol{r}),\ \boldsymbol{r}\in\partial\Omega \quad (10)$$

where S^* is the adjoint source. For a point adjoint source located at the detector position $\boldsymbol{r}_j$, the following holds [8]:

$$g_m^*(\boldsymbol{r},\boldsymbol{r}_j)=g_m(\boldsymbol{r}_j,\boldsymbol{r}),\quad \boldsymbol{r}\in\Omega \quad (11)$$

where g_m is the Green's function of (6) and (8). Note that in this paper, we model the point adjoint source by a Gaussian function with sufficiently low variance, centered at $\boldsymbol{r}_j$ [26].

The emission field at $\boldsymbol{r}_j$ due to the source at $\boldsymbol{r}_i$ is given by the following nonlinear integral equation:

$$\phi_m(\boldsymbol{r}_j,\boldsymbol{r}_i)=\int_\Omega g_m^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}. \quad (12)$$

The relationship between ϕ_m and μ_{axf} defined in (12) is nonlinear because g_m^* is dependent on μ_{amf} which in turn is related to μ_{axf} and the dependence of ϕ_x on μ_{axf} is clear. In the next section, we formally state the inverse problem and address the nonlinearity by using an iterative linearization scheme based on first-order Fréchet derivatives.

III. Inverse Problem

Given N_S sources and N_D detectors, we define $\Gamma_{i,j}$ to be the measurement for a detector at position r_j, $j=1,\ldots,N_D$, due to a source at r_i, $i=1,\ldots,N_S$. The individual measurements can be grouped into the vector form

$$\boldsymbol{\Gamma}:=[\Gamma_{1,1},\ldots,\Gamma_{1,N_D},\Gamma_{2,1},\ldots,\Gamma_{N_S,N_D}]^T \quad (13)$$

where the (i,j)th measurement satisfies the following model:

$$\Gamma_{i,j}=\int_\Omega g_m^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}. \quad (14)$$

Our objective is to recover the quantity μ_{axf} using the measurement vector $\boldsymbol{\Gamma}$ based on the nonlinear integral equation (14) for each (i,j)th pair. In this model, we assume that the measurements are noise-free. This assumption allows us to eliminate the effect of noise in our error analysis, and to focus primarily on the effect of the discretization error.

In the next subsection, to address the problem of nonlinearity in (14), we select an iterative linearization scheme based on first-order Fréchet derivatives. Next, to address the ill-posedness, we discuss regularization in an optimization framework and incorporation of *a priori* information about the unknown image μ_{axf}. Then by taking the derivative of the resulting optimization problem and defining appropriate boundary conditions for it, we convert this optimization problem into a boundary value problem. In the final subsection, we show the variational formulation of the boundary value problem and comment on the existence and uniqueness of the solution.

A. Iterative Linearization

Consider an infinitesimal perturbation on μ_{axf}, [3], [29]

$$\mu_{\mathrm{axf}}\leftarrow\mu_{\mathrm{axf}}+\delta\mu_{\mathrm{axf}}. \quad (15)$$

Then at each linearization step, the corresponding perturbation $\delta\phi_m$ in the emission field at detector position $\boldsymbol{r}_j$ due to the ith source at $\boldsymbol{r}_i$ is given by the following linear integral equation:

$$\begin{aligned}\delta\phi_m(\boldsymbol{r}_j,\boldsymbol{r}_i)=&\int_\Omega g_m^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\delta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}\\&-\int_\Omega g_m^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_m(\boldsymbol{r},\boldsymbol{r}_i)\frac{\partial\mu_{amf}}{\partial\mu_{\mathrm{axf}}}\delta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}\\&-\int_\Omega g_{mx}^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\delta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}\end{aligned} \quad (16)$$

where g_m^* is the solution to the boundary value problem (9), (10) and g_{mx}^* is the solution to the boundary value problem (5) and (7) where $S_i(\boldsymbol{r})$ in (5) is replaced by $\eta\mu_{\mathrm{axf}}(\boldsymbol{r})g_m^*(\boldsymbol{r},\boldsymbol{r}_j)$. In (16), the first integral results from the right-hand side of (6), while the second and third integrals originate from the dependence of respectively μ_{am} and μ_{ax} on the unknown fluorophore absorption coefficient. We note that the kernels of the second and third integrals are much smaller than the kernel of the first integral. Therefore, the first integral in (16) dominates and the last two terms can be neglected. As a result, inverse problem at each linearization step reads

$$\delta\phi_m(\boldsymbol{r}_j,\boldsymbol{r}_i)=\int_\Omega g_m^*(\boldsymbol{r},\boldsymbol{r}_j)\phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\delta\mu_{\mathrm{axf}}(\boldsymbol{r})d\boldsymbol{r}. \quad (17)$$

To simplify the notation, we introduce $\delta\mu(\boldsymbol{r}) := \eta\delta\mu_{\text{axf}}(\boldsymbol{r})$ which represents the unknown perturbed fluorophore absorption coefficient scaled by the quantum efficiency and assume that μ_{axe} is known and $\mu_{ame} \approx \mu_{axe}$. Furthermore, noting that the emission and excitation subscripts are fixed for the duration of this analysis, we represent $g_j^*(\boldsymbol{r}) := g_m^*(\boldsymbol{r}, \boldsymbol{r}_j)$ and $\phi_i(\boldsymbol{r}) := \phi_x(\boldsymbol{r}, \boldsymbol{r}_i)$, suppressing the x, m dependence of these functions.

We define $\delta\Gamma_{i,j}$ to be the differential measurement at the jth detector due to ith source normalized with respect to a known background fluorophore absorption. Let $H_{ij}(\boldsymbol{r}) = g_j^*(\boldsymbol{r})\phi_i(\boldsymbol{r})$. Using (17) we model $\delta\Gamma_{i,j}$ as follows:

$$\begin{aligned}\delta\Gamma_{i,j} &= \int_\Omega H_{ij}(\boldsymbol{r})\delta\mu(\boldsymbol{r})d\boldsymbol{r} \\ &:= (\mathcal{A}\delta\mu)_{ij}. \end{aligned} \tag{18}$$

We represent the differential measurements corresponding to individual source-detector pairs as elements of a vector

$$\delta\boldsymbol{\Gamma} := [\delta\Gamma_{1,1}, \ldots, \delta\Gamma_{1,N_D}, \delta\Gamma_{2,1}, \ldots, \delta\Gamma_{N_S,N_D}]^T.$$

Then

$$\delta\boldsymbol{\Gamma} = \mathcal{A}\delta\mu \tag{19}$$

where $\mathcal{A} : L^2(\Omega) \to \mathbb{R}^{N_S \times N_D}$ is a vector of operators whose (i,j)th entry acting on $\delta\mu$ corresponds to (18). Although all norms on a finite-dimensional space are equivalent, we select the norm on the range of $\mathcal{A}$ to be the l^1 norm as this proves useful in later analysis. Then an upper bound for the linear operator can be given by

$$\|\mathcal{A}\|_{L^2(\Omega)\to l^1} \leq \sum_{i,j}^{N_S,N_D} \left\|g_j^*\phi_i\right\|_0. \tag{20}$$

The boundedness and the finite-dimensional range of operator $\mathcal{A}$ means it is compact [30]. We define the adjoint operator $\mathcal{A}^* : \mathbb{R}^{N_S \times N_D} \to L^2(\Omega)$ acting on the vector $w \in \mathbb{R}^{N_S \times N_D}$ as

$$[\mathcal{A}^* w](\boldsymbol{r}) = \left[H_{11}^*(\boldsymbol{r}), \ldots, H_{N_S1}^*(\boldsymbol{r}), \ldots, H_{N_SN_D}^*(\boldsymbol{r})\right] w \tag{21}$$

where $H_{ij}^*(\boldsymbol{r}) = g_j^*(\boldsymbol{r})\phi_i(\boldsymbol{r})$ for $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$.

Let $\mathcal{B} = \mathcal{A}^*\mathcal{A} : L^2(\Omega) \to L^2(\Omega)$, we have

$$(\mathcal{B}\delta\mu)(\boldsymbol{r}) := \int_\Omega \kappa(\boldsymbol{r}, \acute{\boldsymbol{r}})\delta\mu(\acute{\boldsymbol{r}})d\acute{\boldsymbol{r}} \tag{22}$$

where

$$\kappa(\boldsymbol{r}, \acute{\boldsymbol{r}}) := \sum_{i,j}^{N_S,N_D} H_{ij}^*(\boldsymbol{r})H_{ij}(\acute{\boldsymbol{r}}). \tag{23}$$

Then an alternate form of (19) can be expressed as follows:

$$\gamma(\boldsymbol{r}) = (\mathcal{B}\delta\mu)(\boldsymbol{r}) \tag{24}$$

where $\gamma = \mathcal{A}^*\delta\boldsymbol{\Gamma}$. Note that $\mathcal{B}$ is compact since $\mathcal{A}$ is compact. Therefore, (24) is ill-posed.

In the next subsection, we address the ill-posedness in an optimization framework by incorporating regularization terms.

B. Inverse Problem as an Optimization Problem and Regularization

In this subsection, we address the ill-posedness of (24) through regularization in the optimization framework which provides a suitable means for the incorporation of *a priori* information about the solution. In this respect, we consider the following minimization problem where we seek a solution $\delta\hat{\mu} \in H^1(\Omega)$

$$\delta\hat{\mu} = \min_{\delta\mu \in H^1(\Omega)} J(\delta\mu, \nabla\delta\mu) \tag{25}$$

where the $H^1(\Omega)$ smoothness on the solution is imposed through the use of appropriate regularization terms. The functional J in (25) can be decomposed into two parts, J_L and J_R as follows:

$$J(\delta\mu, \nabla\delta\mu) = J_L(\delta\mu) + J_R(\delta\mu, \nabla\delta\mu) \tag{26}$$

where J_L measures the difference between the predicted and actual measurements

$$J_L(\delta\mu) = \|\delta\boldsymbol{\Gamma} - \mathcal{A}\delta\mu\|_{l^2}^2 \tag{27}$$

and the regularization term J_R contains the *a priori* information. In this work, we assume that *a priori* information on the image and image gradient is available. Let $\beta_1(\boldsymbol{r})$ denote the *a priori* information on the image 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 the image gradient. We incorporate the *a priori* information via zeroth- and first-order Tikhonov regularization terms as follows [31]:

$$\begin{aligned} J_R(\delta\mu, \nabla\delta\mu) = \lambda_1 &\int_\Omega \left[\delta\mu(\boldsymbol{r}) - \beta_1(\boldsymbol{r})\right]^2 d\boldsymbol{r} \\ &+\lambda_2 \int_\Omega \left|\nabla\delta\mu(\boldsymbol{r}) - \boldsymbol{\beta}_2(\boldsymbol{r})\right|^2 d\boldsymbol{r} \end{aligned} \tag{28}$$

where $\nabla\delta\mu$ is the image gradient and $\lambda_1, \lambda_2 > 0$ are regularization parameters. Using (27) and (28), the minimization problem (25) can be rewritten as follows:

$$\begin{aligned} \delta\hat{\mu} = \min_{\delta\mu \in H^1(\Omega)} \Bigg(&\sum_{i,j}^{N_S,N_D} \left[\delta\Gamma_{i,j} - (\mathcal{A}\delta\mu)_{i,j}\right]^2 \\ &+ \lambda_1 \int_\Omega \left[\delta\mu(\boldsymbol{r}) - \beta_1(\boldsymbol{r})\right]^2 d\boldsymbol{r} \\ &+\lambda_2 \int_\Omega \left|\nabla\delta\mu(\boldsymbol{r}) - \boldsymbol{\beta}_2(\boldsymbol{r})\right|^2 d\boldsymbol{r} \Bigg). \end{aligned} \tag{29}$$

There are a number of methods in choosing appropriate regularization parameters, see [32]–[36]. In this work, we assume that λ_1 and λ_2 are properly chosen and focus on deriving discretization error estimates. In the next subsection, after defining appropriate boundary conditions, we consider the equivalent variational formulation of the minimization problem in (29).

C. Inverse Problem as Boundary Value Problem and Variational Formulation

In this work, we follow a finite element method for the discretization of the inverse problem. Due to incorporation of the regularization term on the gradient of the solution, a natural step is to formulate the minimization problem as a variational one. In this subsection, we describe the derivation of the variational problem formulation of the inverse problem by first considering the first-order optimality condition for the minimization problem (29). Next, with the aid of properly chosen boundary conditions, we transform the optimization problem into a boundary value problem (BVP), which is followed by the variational formulation of the BVP. Finally, we show that a unique solution exists to the variational formulation of the regularized inverse problem.

The solution of (29) satisfies $\partial J/\partial\delta\mu(\delta\mu, \nabla_q\delta\mu) = 0$, where ∇_q is the gradient with respect to the r_qth direction for $q = 1, 2, 3$. In particular, if $J = \int u(\boldsymbol{r}, \delta\mu, \partial\delta\mu/\partial r_q)d\boldsymbol{r}$, the Gâteaux derivative [37] is defined by

$$\frac{\partial J}{\partial \delta\mu} = \frac{\partial u}{\partial \delta\mu} - \sum_q \frac{\partial}{\partial r_q}\left(\frac{\partial u}{\partial \delta\mu_q}\right). \tag{30}$$

Taking the Gâteaux derivative of (29) with respect to $\delta\mu$ and setting it equal to zero yields

$$\mathcal{B}\delta\mu(\boldsymbol{r}) + \lambda_1\delta\mu(\boldsymbol{r}) - \lambda_2\nabla^2\delta\mu(\boldsymbol{r}) = f(\boldsymbol{r}) \tag{31}$$

where

$$f(\boldsymbol{r}) := \gamma(\boldsymbol{r}) + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2\sum_{q=1}^{3}\beta_{2q}(\boldsymbol{r}). \tag{32}$$

Note that $f(\boldsymbol{r})$ is composed of known terms from *a priori* information and measurements.

We consider (31) with the following Neumann boundary condition:

$$\frac{\partial\delta\mu}{\partial\hat{n}}(\boldsymbol{r}) = 0, \quad \boldsymbol{r} \in \partial\Omega \tag{33}$$

where $\partial\delta\mu/\partial\hat{n}$ is the directional derivative of $\delta\mu$ along the unit normal vector at the boundary $\partial\Omega$. The boundary condition in (33) implies that no changes in the perturbed fluorophore concentration occur across the boundary.

At this point, one can consider a finite difference scheme for the solution of the inverse problem which is posed as a boundary value problem (31)(33). However, as our goal in this paper is to apply a finite element scheme for the discretization of the BVP, we obtain the corresponding variational (weak) problem. Hence, we multiply both sides of (31) by a test function $\psi \in H^1(\Omega)$, and integrate over Ω. Applying Green's first theorem to the last term on the left and using the boundary condition in (33), we obtain

$$\int_\Omega \psi(\boldsymbol{r})\left[(\mathcal{B}\delta\mu)(\boldsymbol{r}) + \lambda_1\delta\mu(\boldsymbol{r})\right]d\boldsymbol{r} + \lambda_2\int_\Omega \nabla\psi(\boldsymbol{r})\cdot\nabla\delta\mu(\boldsymbol{r})d\boldsymbol{r} = \int_\Omega \psi(\boldsymbol{r})f(\boldsymbol{r})d\boldsymbol{r}. \tag{34}$$

A more convenient way to express (34) is through a bilinear form. Thus, we define

$$\mathcal{F}(\psi, \delta\mu) := (\psi, \mathcal{B}\delta\mu) + \lambda_1(\psi, \delta\mu) + \lambda_2(\nabla\psi, \nabla\delta\mu), \tag{35}$$

$$\mathcal{G}(\psi) := (\psi, f) \tag{36}$$

where the inner product is defined by

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

Hence, (34) can be expressed as

$$\mathcal{F}(\psi, \delta\mu) = \mathcal{G}(\psi). \tag{37}$$

It can be shown that the bilinear form (35) is bounded and coercive for regularization parameters $\lambda_1, \lambda_2 > 0$ [30]. Thus, by the Lax–Milgram lemma, a unique solution exists for the regularized inverse problem (37) for each pair of $\lambda_1, \lambda_2 > 0$ used in the formulation of the Galerkin problem (34) [30], [38]. For an explicit statement of the Lax–Milgram lemma, see Appendix A.

In the following section, we describe the discretization methods selected in this paper for each of the separate forward and inverse discretizations as well as the combined forward and inverse discretization.

IV. Discretization by Finite Element Method of the Forward and Inverse Problems

In the following subsections, we first discuss the variational formulation and finite element discretization of the forward problem. In practice, for arbitrary domains and background optical properties, no analytical solutions exist for the forward problem when it is defined in a variational form. Thus, we discretize the forward problem and obtain finite-dimensional approximations of g_j^* and ϕ_i, for $j = 1, \ldots, N_D$, $i = 1, \ldots, N_S$.

Next, we use the finite element solutions of the forward problem in the inverse problem formulation, which implies an approximation to the inverse problem. The resulting inverse problem in general does not possess a closed form solution. Therefore, finding the solution calls for numerical techniques. We discuss the discretization of the resulting approximate inverse problem using projection by Galerkin method.

A. Discretization of the Forward Problem

In this subsection, we discuss the forward problem discretization. We express the coupled PDEs in their variational form in order to apply a finite element method.

To do so, we multiply (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 (7) we have

$$\int_\Omega (\nabla\xi_1\cdot D\nabla\phi_i + \mu_{ax}\xi_1\phi_i)d\boldsymbol{r} + \frac{1}{2\rho}\int_{\partial\Omega}\xi_1\phi_i dl = \int_\Omega \xi_1 S_i d\boldsymbol{r}. \tag{38}$$

It can be shown that a unique solution for (38) exists and is bounded [38], [39]. Similarly, for a test function $\xi_2 \in H^1(\Omega)$,

the variational form for the adjoint forward problem (9), (10) becomes

$$\int_{\Omega} \left(\nabla\xi_2 \cdot D\nabla g_j^* + \mu_{am}\xi_2 g_j^*\right) d\boldsymbol{r} + \frac{1}{2\rho}\int_{\partial\Omega} \xi_2 g_j^* dl = \frac{1}{2\rho}\int_{\partial\Omega} \xi_2 S_j^* dl \tag{39}$$

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

Next, the functions ξ_1, ϕ_i in (38) and ξ_2, g_j^* in (39) are replaced by their finite-dimensional counterparts

$$\Xi_1^{N_i}(\boldsymbol{r}) := \sum_{k=1}^{N_i} p_k L_k(\boldsymbol{r}), \quad \Phi_i^{N_i} := \sum_{k=1}^{N_i} c_k L_k(\boldsymbol{r}). \tag{40}$$

$$\Xi_2^{N_j}(\boldsymbol{r}) := \sum_{k=1}^{N_j} p_k L_k(\boldsymbol{r}), \quad G_j^{*,N_j} := \sum_{k=1}^{N_j} d_k L_k(\boldsymbol{r}). \tag{41}$$

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

Substitution of (40), (41) into the variational forward problem (38), (39) yields the matrix equations

$$\mathbf{M}\mathbf{c}_i = \mathbf{q}_i \tag{42}$$

$$\mathbf{M}^*\mathbf{d}_j^* = \mathbf{q}_j^* \tag{43}$$

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

Let $\{\Omega_{ni}\}$ denote the set of linear elements used to discretize (38) for $n = 1,\ldots,N_\Delta^i$, where N_Δ^i is the number of elements for the ith source such that $\bigcup_n^{N_\Delta^i} \Omega_{ni} = \Omega$ for all $i = 1,\ldots,N_S$. Similarly, let $\{\Omega_{mj}\}$ denote the set of linear elements used to discretize (39) for $m = 1,\ldots,N_\Delta^{*j}$ where N_Δ^{*j} is the number of elements for the jth detector such that $\bigcup_m^{N_\Delta^{*j}} \Omega_{mj} = \Omega$ for all $j = 1,\ldots,N_D$.

A bound for error e_j^* and e_i between g_j^*, ϕ_i and G_j^*, Φ_i respectively on each finite element can be given by [38]

$$\left\|e_j^*\right\|_{0,mj} \le C \left\|g_j^*\right\|_{1,mj} h_{mj} \tag{44}$$

$$\|e_i\|_{0,ni} \le C\|\phi_i\|_{1,ni} h_{ni} \tag{45}$$

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

In the next subsection, these approximate solutions to the forward problem are substituted into the inverse problem operator. The error is estimated based on the resulting operators with approximations.

B. Simultaneous Discretization of the Forward and Inverse Problems

We substitute the finite element solutions $\Phi_i, i = 1,\ldots,N_S$ and $G_j^*, j = 1,\ldots,N_D$ of the forward problem into $H_{i,j}, H_{i,j}^*$ in the operators $\mathcal{A}, \mathcal{A}^*$ defined by (18) and (21). The resulting approximate operators are denoted by tildes $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{A}}^*$, indicating that the finite element solutions of the forward problem are used. By so doing, we arrive at the approximate variational problem formulation

$$\widetilde{\mathcal{F}}(\psi, \widetilde{\delta\mu}) = \tilde{\mathcal{G}}(\psi). \tag{46}$$

In (46), $\widetilde{\mathcal{F}}(\psi, \widetilde{\delta\mu})$ and $\tilde{\mathcal{G}}(\psi)$ are given, respectively, by

$$\widetilde{\mathcal{F}}(\psi, \widetilde{\delta\mu}) := (\psi, \tilde{\mathcal{B}}\widetilde{\delta\mu}) + \lambda_1(\psi, \widetilde{\delta\mu}) + \lambda_2(\nabla\psi, \nabla\widetilde{\delta\mu}) \tag{47}$$

$$\tilde{\mathcal{G}}(\psi) := (\psi, \tilde{f}) \tag{48}$$

where

$$(\tilde{\mathcal{B}}\widetilde{\delta\mu})(\boldsymbol{r}) := \int_{\Omega} \tilde{\kappa}(\boldsymbol{r}, \boldsymbol{r}')\widetilde{\delta\mu}(\boldsymbol{r}')d\boldsymbol{r}' = \int_{\Omega} \sum_{i,j}^{N_S,N_D} G_j^*(\boldsymbol{r})\Phi_i(\boldsymbol{r})G_j^*(\boldsymbol{r}')\Phi_i(\boldsymbol{r}')\widetilde{\delta\mu}(\boldsymbol{r}')d\boldsymbol{r}'$$

and

$$\tilde{f}(\boldsymbol{r}) := \tilde{\gamma}(\boldsymbol{r}) + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2\sum_{q=1}^{3}\beta_{2q}(\boldsymbol{r}) = \sum_{i,j}^{N_S,N_D} G_j^*(\boldsymbol{r})\Phi_i(\boldsymbol{r})\Gamma_{i,j} + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2\sum_{q=1}^{3}\beta_{2q}(\boldsymbol{r}).$$

Next, we discretize the functions ψ and $\widetilde{\delta\mu}$. 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 $\{L_1,\ldots,L_N\}$ which are associated with the set of points $\{\boldsymbol{r}_p\}$, $p = 1,\ldots,N$, on Ω. We replace ψ and $\widetilde{\delta\mu}$ in (37) by their respective, finite-dimensional counterparts $\Psi^N \in V_N$ and $\widetilde{\Delta\mu}^N \in V_N$

$$\Psi^N := \sum_{k=1}^{N} p_k L_k(\boldsymbol{r}), \tag{49}$$

$$\widetilde{\Delta\mu}^N := \sum_{k=1}^{N} m_k L_k(\boldsymbol{r}) \tag{50}$$

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 (49), (50) into (46) we arrive at

$$\widetilde{\mathcal{F}}(\Psi, \widetilde{\Delta\mu}) = \tilde{\mathcal{G}}(\Psi). \tag{51}$$

This can be transformed to a matrix equation

$$\mathbf{F}_N \boldsymbol{m} = \mathbf{G}_N \tag{52}$$

where $\boldsymbol{m} = [m_1, \cdots, m_N]^T$ represents the unknown coefficients in the finite expansion of (50) and $\mathbf{F}_N$ and $\mathbf{G}_N$ are respectively the finite element matrix and the load vector resulting from the projection of (46) by Galerkin method.

Let $\{\Omega_t\}$ denote the set of linear elements used to discretize (46) for $t = 1, \ldots, N_\Delta$, 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 discretize the forward problem. Similar to the forward problem, a traditional error estimate for error e_μ between $\widetilde{\delta\mu}$ and $\widetilde{\Delta\mu}$ on each finite element can be given by

$$\|e_\mu\|_{0,t} \leq C \|\widetilde{\delta\mu}\|_{1,t} h_t \tag{53}$$

where C is a positive constant, $\|\cdot\|_{0,t}$ and $\|\cdot\|_{1,t}$ denote the L^2 and H^1 norms, respectively on $\{\Omega_t\}$; and h_t is the diameter of the smallest ball containing the finite element $\{\Omega_t\}$ in the solution $\widetilde{\Delta\mu}$.

V. Analysis of the Error in Fluorescence Imaging Due to Discretization

In this work, we consider the solution of the problem stated in (37) to be exact solution since neither the forward problem nor the inverse problem is discretized. Our objective is to examine the error in fluorescence absorption 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 meshes that could reduce the total error in the reconstructed images due to discretization.

We have divided this section into two subsections: In the first subsection, we derive an estimate for the error in fluorescence absorption imaging due to the forward problem discretization as described in the previous section. Thus, the first error we find is the difference $e := \delta\mu - \widetilde{\delta\mu}$, where $\delta\mu$ satisfies the exact inverse problem (37) and $\widetilde{\delta\mu}$ satisfies the approximate inverse problem (46). Note that the inverse problem is not discretized for this case.

In the second subsection, we analyze the error in the reconstructed fluorophore absorption coefficient resulting from the finite element discretization of the inverse problem. Therefore, we examine the error E between the solution $\widetilde{\delta\mu}$ of (46) and the solution $\widetilde{\Delta\mu}$ of (51), i.e., $E = \widetilde{\delta\mu} - \widetilde{\Delta\mu}$. In this case, the error is due entirely to the discretization of the approximate inverse problem (46).

Finally, we define the total error as the difference between the exact solution $\delta\mu$ of (37) and $\widetilde{\Delta\mu}$ of (51) in terms of the two contributors

$$\delta\mu - \widetilde{\Delta\mu} = e + E. \tag{54}$$

In the following, we analyze each of the error contributors and derive estimates in the form of upper bounds of the $H^1(\Omega)$ norm of these errors.

A. Error in Fluorescence Imaging Due to Forward Problem Discretization

The following theorem presents a bound for the $H^1(\Omega)$ of the error $e := \delta\mu - \widetilde{\delta\mu}$, where $\delta\mu$ satisfies the exact inverse problem (37) and $\widetilde{\delta\mu}$ satisfies the approximate inverse problem (46).

Theorem 1: Let $\{\Omega_{ni}\}$ denote the set of linear elements used to discretize (38) for $n = 1, \ldots, N_\Delta^i$; such that $\bigcup_n^{N_\Delta^i} \Omega_{ni} = \Omega$ and h_{ni} is the diameter of the smallest ball that contains the nth element in the solution Φ_i, for all $i = 1, \ldots, N_S$. Similarly, let $\{\Omega_{mj}\}$ denote the set of linear elements used to discretize (39) for $m = 1, \ldots, N_\Delta^{*j}$; such that $\bigcup_m^{N_\Delta^{*j}} \Omega_{mj} = \Omega$ and h_{mj} is the diameter of the smallest ball that contains the mth element in the solution G_j^*, for all $j = 1, \ldots, N_D$. Then a bound for the error between the solution $\delta\mu$ of (37) and the solution $\widetilde{\delta\mu}$ of (46) due to the approximations $\widetilde{\mathcal{F}}$ and $\tilde{\mathcal{G}}$ is given by

$$\begin{aligned}
&\|\delta\mu - \widetilde{\delta\mu}\|_1 \\
&\quad \leq \frac{C}{\min(\lambda_1, \lambda_2)} \max_{i,j} \left\| g_j^* \phi_i \right\|_0 \\
&\qquad \cdot \Bigg(\sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} \left(2 \left\| g_j^* \delta\mu \right\|_{0,ni} + \left\| g_j^* \right\|_{\infty,ni} \|\delta\mu\|_0 \right) \\
&\qquad\quad \cdot \|\phi_i\|_{1,ni} h_{ni} \\
&\qquad\quad + \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} (2\|\phi_i \delta\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta\mu\|_0) \\
&\qquad\quad \times \left\| g_j^* \right\|_{1,mj} h_{mj} \Bigg)
\end{aligned} \tag{55}$$

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

Proof: The proof of this theorem is given in Appendix D.□

The error estimate (55) 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 criterion for the forward problem which includes the accuracy of the inverse problem solution, in addition to the accuracy of the forward problem solution. First, the forward problem discretization should include the discretization of each solution ϕ_i and g_j^*. Second, to keep the total error bound low, the h_{ni} of the nth element in solution ϕ_i has to be small when $\sum_j^{N_D} (2\|g_j^* \delta\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\delta\mu\|_0) \|\phi_i\|_{1,ni}$ is large on this element; and the h_{mj} of the mth element in solution g_j^* has to be small when $\sum_i^{N_S} (2\|\phi_i \delta\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta\mu\|_0) \|g_j^*\|_{1,mj}$ is large on this element. Note that $\|\phi_i\|_{1,ni}$ or $\|g_j^*\|_{1,mj}$ is large on the element close to the ith source or the jth detector, respectively, and the value of the terms $\sum_j^{N_D} (2\|g_j^* \delta\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\delta\mu\|_0)$ and $\sum_i^{N_S} (2\|\phi_i \delta\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta\mu\|_0)$ depend on the absorption coefficient of the fluorophore 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 traditional error estimates (44) and (45) only depend on the smoothness and support of ϕ_i and g_j^* and the finite-dimensional space of approximating functions. Traditional error estimates 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 (44) and (45) 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 lead to a higher error bound. The regularization parameters scale the sum of the terms. Thus, choosing smaller values for λ_1, λ_2 results in a higher error estimate. Note, too, that since the error bound is a sum over all sources and detectors, increasing the number of either can have an impact on the error estimate.

B. Error in Fluorescence Imaging Due to Inverse Problem Discretization

In the following theorem, we present a bound for the $H^1(\Omega)$ norm of the error E between the solution $\widetilde{\delta\mu}$ of (46) and the solution $\widetilde{\Delta\mu}$ of (51).

Theorem 2: Consider the Galerkin projection of the variational problem (46) on a finite-dimensional subspace $V_N \subset H^1(\Omega)$ using a set of linear 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 tth element. Assume that the solution $\widetilde{\delta\mu}$ of (46) also satisfies $\widetilde{\delta\mu} \in H^2(\Omega)$. Then a bound for the error E in the solution $\widetilde{\Delta\mu}$ of (51) with respect to the solution $\widetilde{\delta\mu}$ of (46) can be given by

$$\begin{aligned}
&\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \\
&\quad \le \frac{C}{\min(\lambda_1, \lambda_2)} \\
&\qquad \times \left(\max_{i,j} \left\|G_j^* \Phi_i\right\|_0 \cdot \sum_t^{N_\Delta} \sum_{i,j}^{N_S, N_D} \left\|G_j^* \Phi_i\right\|_{0,t} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 \right. \\
&\qquad\quad \left. + \lambda_1 \sum_t^{N_\Delta} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 + \lambda_2 \sum_t^{N_\Delta} \|\widetilde{\delta\mu}\|_{2,t} h_t \right) \qquad (56)
\end{aligned}$$

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

Proof: The proof of this theorem is given in Appendix E.□

The error estimate (56) shows that the error bound for the discretization of inverse problem 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 (56) shows that the term $\|\widetilde{\delta\mu}\|_{2,t}$ 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 error bound 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 $\|\widetilde{\delta\mu}\|_{2,t}$ is large, can help to lower the error bound, but it also depends on the regularization parameters λ_1 and λ_2. Comparing (56) with the traditional error estimate (53), (56) suggests a discretization criterion based on the inverse problem solution $\widetilde{\delta\mu}$, the forward problem solutions Φ_i and G_j^*, as well as the spatial relationship among these solutions. This also implies that simply keeping the mesh parameter small over regions where $\|\widetilde{\delta\mu}\|_{2,t}$ is large, as indicated in (53), may not ensure a lower error bound, thereby a reduction of the error in the reconstructed image because of the dependence of the bound on the location of the heterogeneity with respect to the sources and detectors.

Similar to Theorem 1, the regularization parameters λ_1, λ_2 and the number of sources and detectors also affect the error bound.

Combining results of Theorems 1 and 2 and rearranging the terms, both error estimates can be viewed in a single bound. Let $\|G_j^* \Phi_i\|_0 \le \|g_j^* \phi_i\|_0$ for all $i = 1, \cdots, N_S$ and $j = 1, \cdots, N_D$. Then

$$\begin{aligned}
&\|\delta\mu - \widetilde{\Delta\mu}\|_1 \\
&\quad \le \frac{C}{\min(\lambda_1, \lambda_2)} \max_{i,j} \left\|g_j^* \phi_i\right\|_0 \\
&\qquad \cdot \sum_{i,j}^{N_S, N_D} \left(\sum_n^{N_\Delta^i} \left(2 \left\|g_j^* \delta\mu\right\|_{0,ni} + \left\|g_j^*\right\|_{\infty,ni} \|\delta\mu\|_0 \right) \right. \\
&\qquad\qquad \times \|\phi_i\|_{1,ni} h_{ni} \\
&\qquad\qquad + \sum_m^{N_\Delta^{*j}} \left(2\|\phi_i \delta\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\delta\mu\|_0\right) \\
&\qquad\qquad \left. \times \left\|g_j^*\right\|_{1,mj} h_{mj} + \sum_t^{N_\Delta} \left\|G_j^* \Phi_i\right\|_{0,t} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 \right) \\
&\qquad + \frac{C}{\min(\lambda_1, \lambda_2)} \\
&\qquad \times \left(\lambda_1 \sum_t^{N_\Delta} \|\widetilde{\delta\mu}\|_{2,t} h_t^2 + \lambda_2 \sum_t^{N_\Delta} \|\widetilde{\delta\mu}\|_{2,t} h_t \right). \qquad (57)
\end{aligned}$$

We note that the error estimates in Theorem 1 and 2 can be extended to show the effect of noise as an additional error source [26].

C. Error in Iterative Linearization

The error bounds described in the previous subsections only address one linearized step of the nonlinear inverse problem. To consider all iterative linearization steps, we use the following method to describe the propagation of errors at each linearization step:

Let $\delta\mu^{(k)}$ and $\widetilde{\Delta\mu}^{(k)}$ represent the solution to (37) and (51) at the kth linearization step, respectively. The absorption coefficient as given in (15) at the end of the $(r-1)$th iteration step is $\tilde{\mu}_{\text{axf}}^{(r-1)} = \mu_{\text{axf}}^{(0)} + \sum_{k=1}^{r-1} \widetilde{\Delta\mu}^{(k)}/\eta$, where $\mu_{\text{axf}}^{(0)}$ is an initial guess for the value of the background absorption coefficient, and $\widetilde{\Delta\mu}^{(k)}$ contains discretization error with respect to the exact solution $\delta\mu^{(k)}$. Three errors are introduced at the rth iteration step. The first is due to the finite element discretization of the inverse problem. The next is due to finite element discretization in the forward problem by the error $(\mathcal{A} - \tilde{\mathcal{A}})^{(r-1)}$ in operators $\tilde{\mathcal{A}}^{(r-1)}$ and the corresponding measurement vector error $(\gamma - \tilde{\gamma})^{(r-1)}$ in $\tilde{\gamma}^{(r-1)}$. Finally, μ_{axf} and μ_{amf}, which are related to each other, appear as the coefficient in the forward problem (5)–(7) and the

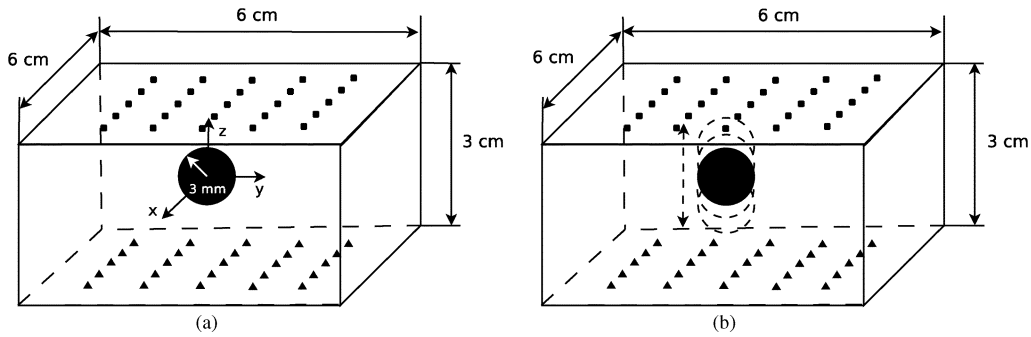

Fig. 1. Setups used for the simulation studies 1 and 2. The squares and triangles denote the detectors and sources, respectively. (a) Optical domain and source-detector configuration for Simulation study 1. (b) Optical domain and source-detector configuration for Simulation study 2. The radius of the circles is 3 mm.

adjoint problem (9), (10), respectively. Then, the error in the solution $\mu_{\text{axf}}^{(k)}$ at each step will propagate and lead to an additional error in Φ_i and G_j^* at the rth step in addition to the discretization error analyzed before.

If the exact solution (without any discretization error) after $(r-1)$th iteration step is given by $\mu_{\text{axf}}^{(r-1)} = \mu_{\text{axf}}^{(0)} + \sum_{k=1}^{r-1} \delta\mu^{(k)}/\eta$, then the error in $\tilde{\mu}_{\text{axf}}^{(r)}$ at the rth iteration is bounded by

$$\left\|\mu_{\text{axf}}^{(r)} - \tilde{\mu}_{\text{axf}}^{(r)}\right\|_1 = \left\|\sum_{k=1}^{r}\left[\delta\mu^{(k)} - \widetilde{\Delta\mu}^{(k)}\right]/n\right\|_1 \leq \sum_{k=1}^{r}\left\|\delta\mu^{(k)} - \widetilde{\Delta\mu}^{(k)}\right\|_1/n. \quad (58)$$

VI. Numerical Experiments

To demonstrate the implications of Theorem 1 and 2, we performed a series of numerical experiments. In the first simulation, we considered a series of image reconstructions to show how the error, in reconstructed images due to discretization, changes as the absorption coefficient of the fluorophore increases. In the next set of experiments, we showed the effect of the relative position of the fluorophore concentration with respect to the sources and detectors on the accuracy of reconstructed images.

We first describe the numerical phantom, data generation and image reconstruction procedures, and next, describe the results of numerical simulations.

A. Description of the Phantom, Data Generation and Reconstruction

For both simulation studies, we considered a $6\,\text{cm} \times 6\,\text{cm} \times 3\,\text{cm}$ cubic domain Ω shown in Fig. 1(a). We placed 25 sources and 25 detectors evenly on two 5×5 grids at the bottom and top surface of the domain, respectively. The circular heterogeneity represents the fluorophore concentration with constant absorption coefficient μ_{axf} and radius 3 mm, embedded in an optically homogeneous background with $\mu_{ae} = 0.05\ \text{cm}^{-1}$ at both excitation and emission wavelengths. We set the diffusion coefficient $D(\boldsymbol{r}) = 0.0410\,\text{cm}$ for $\boldsymbol{r} \in \Omega \cup \partial\Omega$ and the refractive index mismatch parameter $\rho = 3$ for the boundary. Using the parameters above, we simulated the fluorescence data by solving the coupled diffusion equations (5) and (6) with their corresponding boundary conditions (7) and (8) on a fine uniform grid with $81 \times 81 \times 41$ nodes.

To demonstrate the effect of discretization, we considered three image reconstruction scenarios corresponding to three different meshing schemes.

1) To obtain an accurate solution, we solved the forward problem on a very fine uniform mesh with $61 \times 61 \times 31$ nodes and obtained the corresponding finite element solutions, which we assume has negligible error due to discretization. Next, we formulated the inverse problem with these finite element solutions and discretized the resulting variational inverse problem on the same uniform mesh. As a result of using very fine meshes, we assume that the solution of the inverse problem on this mesh, denoted by $\delta\mu$, has negligible error that can be attributed to discretization. Therefore, we used $\delta\mu$ as a baseline for comparison.
2) To demonstrate the effect of discretization of the forward problem, we solved the forward problem on relatively coarse mesh shown in Fig. 2(a). Then, we used the fine uniform mesh with $61 \times 61 \times 31$ nodes to compute the inverse problem solution. We denote this solution by $\widetilde{\delta\mu}$ and assume that it possesses error with respect to the baseline image $\delta\mu$ due to only the forward problem discretization.
3) To demonstrate the effect of discretization of the inverse problem, we discretized the forward and inverse problem on two coarse meshes shown in Fig. 2(a) and (b), respectively, and computed the solution of the resulting inverse problem. We denote the solution by $\widetilde{\Delta\mu}$ and assume that it possesses error with respect to the solution $\widetilde{\delta\mu}$ due to only the inverse problem discretization.

We note that for each case the forward problem was solved with the same parameters as the data generation with the exception of the mesh size.

For the inverse problem, we chose the regularization parameters 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 $\lambda_2 = 1 \times 10^{-8}$. We assumed that no *a priori* information were available about the image or its gradient, hence we set $\beta_1(\boldsymbol{r}) = 0$ and $\boldsymbol{\beta}_2(\boldsymbol{r}) = \boldsymbol{0}$.

Note that our simulation study is performed using a C++ finite element library—deal.II [40]. We used hexahedral finite elements associated with trilinear Lagrange basis functions to dis-

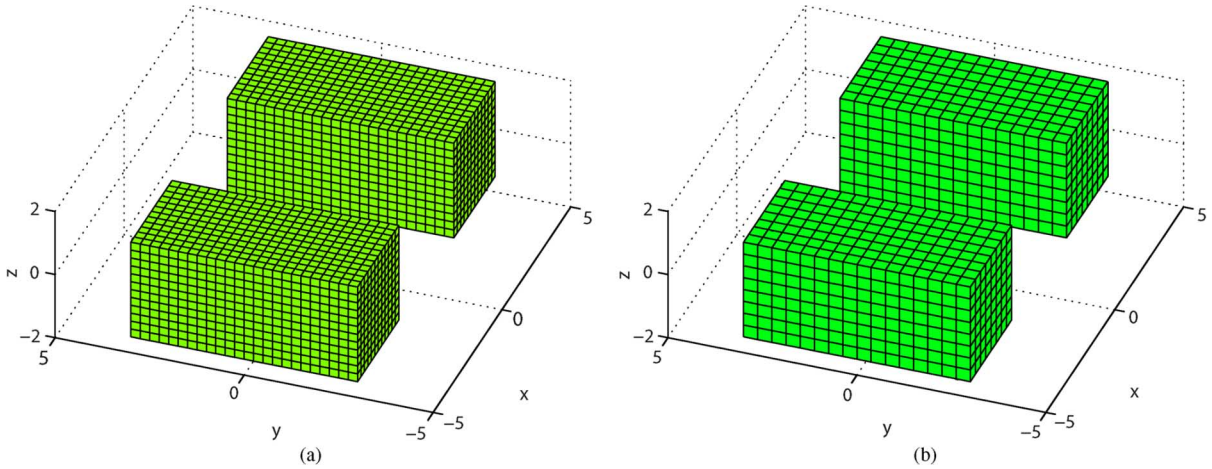

Fig. 2. Coarse uniform meshes used to discretize the forward and inverse problems in the Simulation studies 1 and 2. The mesh is cut through to show the mesh structure inside. (a) The coarse uniform mesh with $25 \times 25 \times 13$ nodes used to discretize the forward problem. (b) Coarse uniform mesh with $17 \times 17 \times 9$ nodes used to discretize the inverse problem.

TABLE II
$H^1(\Omega)$ NORM OF THE ERROR RESULTING FROM FORWARD AND INVERSE PROBLEM DISCRETIZATIONS IN THE RECONSTRUCTED OPTICAL IMAGES IN SIMULATION STUDY 1. ABSORPTION COEFFICIENT $\mu_{\rm axf}$ OF THE FLUOROPHORE IS GIVEN IN $\rm cm^{-1}$

	Case 1	Case 2	Case 3	Case 4	Case 5
μ_{axf}:	0.005	0.010	0.015	0.020	0.025
$\|\delta\mu - \widetilde{\delta\mu}\|_1$ $(\times 10^{-4})$:	0.715	1.362	1.969	2.551	3.029
$\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ $(\times 10^{-4})$:	0.3965	0.7197	0.9834	1.3605	1.5977
$\|\delta\mu\|_0$ $(\times 10^{-4})$:	1.427	2.850	3.858	5.057	6.531
$\|\delta\mu\|_1$ $(\times 10^{-4})$:	1.965	3.924	5.915	7.789	9.086

TABLE III
$H^1(\Omega)$ NORM OF THE ERROR RESULTING FROM FORWARD AND INVERSE PROBLEM DISCRETIZATIONS IN THE RECONSTRUCTED OPTICAL IMAGES IN SIMULATION STUDY 2

	Case 1	Case 2	Case 3	Case 4	Case 5
center located at:	(0,0,0.8)	(0,0,0.4)	(0,0,0)	(0,0,-0.4)	(0,0,-0.8)
$\|\delta\mu - \widetilde{\delta\mu}\|_1$ $(\times 10^{-4})$:	2.378	2.102	1.969	2.003	2.182
$\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ $(\times 10^{-4})$:	1.172	1.058	0.983	0.996	1.075
$\|\delta\mu\|_0$ $(\times 10^{-4})$:	4.021	3.940	3.858	3.897	3.914
$\|\delta\mu\|_1$ $(\times 10^{-4})$:	6.327	6.123	5.915	6.073	6.114

cretize both the forward and inverse problems. Then we used the Gaussian quadrature method to evaluate the integrals in the variational problems given in (38), (39) and (51) [41]. 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.

B. Simulation Results

We performed two sets of numerical experiments to show the effect of fluorophore concentration and fluorophore location on the error in reconstructed images due to discretization. In the first simulation study, we consider the geometry shown in Fig. 1(a). To show the effect of fluorophore concentration on the error in reconstructed images resulting form discretization, we chose 5 data sets corresponding to 5 different values for $\mu_{\rm axf}$: $0.005, 0.010, 0.015, 0.020$ and $0.025\ \rm cm^{-1}$. In the second simulation study, we considered the geometry shown in Fig. 1(b). To show the effect of the fluorophore location on the error due to discretization in reconstructed images, we simulated 5 data sets corresponding to 5 different positions of the circle centered at: (0,0,−0.8), (0,0,−0.4), (0,0,0), (0,0,0.4) and (0,0,0.8), respectively. Note that the center of the square domain is positioned at (0,0,0). The absorption coefficient of the fluorophore concentration was set to $\mu_{\rm axf} = 0.015\ \rm cm^{-1}$. Tables II and III show the error in the solutions $\widetilde{\delta\mu}$ and $\widetilde{\Delta\mu}$ due to forward and inverse problem discretizations, respectively, in two sets of simulations.

The results in Tables II and III show the following.

i) The error due to both the forward and inverse problem discretization is comparable to the norm of the accurate solution $\|\delta\mu\|_1$. This indicates that the error due to discretization can be a significant factor that determines the accuracy of fluorescence imaging.

ii) The error in reconstructed images due to the forward problem discretization is larger as compared to the error due to the inverse problem discretization. This shows the effect of the accuracy of the forward problem solutions in reconstructed images. This result is consistent with Theorem 1 which states the interrelatedness of the forward and inverse problem discretization in terms of the error in reconstructed images.

iii) The error due to discretization increases nonlinearly with the increasing absorption coefficient $\mu_{\rm axf}$ of the fluorophore, and also increases when the fluorophore concentration moves closer to the detectors. The results show the dependency of the discretization error on image and/or geometry specific factors such as the value of fluorophore absorption coefficient and its position. The effect of image and geometry specific factors is not indicated by the traditional error estimate, but clearly spelled out by the error estimates in Theorems 1 and 2.

In the second part of this work [28], we present two novel adaptive mesh generation algorithms developed based on the theorems presented in this paper.

VII. Conclusion

In this work, we analyzed the error in fluorescence diffuse optical tomography resulting from the discretization. We presented the results of the error analysis in two theorems which provide two estimates for the error in the reconstructed absorption coefficient of the fluorophore resulting from the discretization of the forward and inverse problems, respectively. These theorems show that the error in the reconstructed optical image due to the discretization of each problem is affected by the absorption coefficient value of the fluorophore, the number of source-detector pairs, their locations with respect to the fluorophore concentration and the regularization parameters that are used to address the ill-posedness of the inverse problem.

To demonstrate the dependency of the error due to discretization on various medium parameters, we conducted a series of simulation experiments. The first set of experiments shows the dependence of the error due to discretization on the absorption coefficient of the fluorophore. The second set of experiments show that the location of the fluorophore concentration with respect to the sources and detectors can be an important factor that determines the extent of the error resulting from discretization in the reconstructed optical image.

Appendix A
Lax–Milgram Lemma

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

$$\mathcal{A}(v,u) = \mathcal{F}(v), \quad \forall v \in V. \tag{59}$$

The proof may be found in [38]. As a direct consequence of this lemma, an upper bound on u may be established as

$$\|u\|_V \leq \frac{1}{Q}\|\mathcal{F}\|_{V^*} \tag{60}$$

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

Appendix B
Definition of Dual Norm

The dual norm of $\mathcal{F}$ is defined by [38]

$$\|(\mathcal{F})(\delta\mu,\psi)\|_{1*} := \sup_{\psi\neq 0}\frac{|(\mathcal{F})(\delta\mu,\psi)|}{\|\psi\|_1} \tag{61}$$

where $\|\cdot\|_{1*}$ denotes the norm of $H^{1*}(\Omega)$ which is the dual space of $H^1(\Omega)$.

Appendix C
Coercivity of $\mathcal{F}$

Recall $\mathcal{B} = \mathcal{A}^*\mathcal{A}$. Thus $(\mathcal{B}\theta,\theta) \geq 0$ and from the positive parameters λ_1, λ_2, we see

$$\begin{aligned}|\mathcal{F}(\theta,\theta)| &\geq \lambda_1(\theta,\theta) + \lambda_2(\nabla\theta,\nabla\theta)\\ &\geq \min(\lambda_1,\lambda_2)\|\theta\|_1^2\\ &\geq Q\|\theta\|_1^2\end{aligned} \tag{62}$$

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

Appendix D
Proof of Theorem 1: Error Estimate Due to Forward Problem Discretization

Subtracting $\mathcal{F}(\psi,\delta\mu)$ from both sides of (46) yields

$$\begin{aligned}\widetilde{\mathcal{F}}(\widetilde{\delta\mu},\psi) - \widetilde{\mathcal{F}}(\delta\mu,\psi) &= \tilde{\mathcal{G}}(\psi) - \widetilde{\mathcal{F}}(\delta\mu,\psi),\\ \widetilde{\mathcal{F}}(\widetilde{\delta\mu}-\delta\mu,\psi) &= \tilde{\mathcal{G}}(\psi) - \widetilde{\mathcal{F}}(\delta\mu,\psi).\end{aligned} \tag{63}$$

Adding and subtracting $\mathcal{G}(\psi)$ on the right hand side of (63) leads to

$$\begin{aligned}\widetilde{\mathcal{F}}(\widetilde{\delta\mu}-\delta\mu,\psi) &= \tilde{\mathcal{G}}(\psi) - \widetilde{\mathcal{F}}(\delta\mu,\psi) + \mathcal{G}(\psi) - \mathcal{G}(\psi)\\ &= \tilde{\mathcal{G}}(\psi) - \widetilde{\mathcal{F}}(\delta\mu,\psi) + \mathcal{F}(\delta\mu,\psi) - \mathcal{G}(\psi)\\ &= (\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi) + (\tilde{\mathcal{G}}-\mathcal{G})(\psi).\end{aligned} \tag{64}$$

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

$$\begin{aligned}\|\delta\mu-\widetilde{\delta\mu}\|_1 &\leq \frac{1}{\min(\lambda_1,\lambda_2)}\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi)+(\tilde{\mathcal{G}}-\mathcal{G})(\psi)\right\|_{1*}\\ &\leq \frac{1}{\min(\lambda_1,\lambda_2)}\left(\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi)\right\|_{1*}\right.\\ &\qquad \left. + \left\|(\tilde{\mathcal{G}}-\mathcal{G})(\psi)\right\|_{1*}\right)\end{aligned} \tag{65}$$

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

Clearly,

$$\begin{aligned}(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi) &= (\mathcal{B}\delta\mu,\psi) + \lambda_1(\delta\mu,\psi) - \lambda_2(\nabla\delta\mu,\nabla\psi)\\ &\quad - (\tilde{\mathcal{B}}\delta\mu,\psi) - \lambda_1(\delta\mu,\psi) + \lambda_2(\nabla\delta\mu,\nabla\psi)\\ &= \left((\mathcal{B}-\tilde{\mathcal{B}})\delta\mu,\psi\right).\end{aligned} \tag{66}$$

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

$$\begin{aligned}\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi)\right\|_{1*} &:= \sup_{\psi\neq 0}\frac{\left|(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi)\right|}{\|\psi\|_1}\\ &= \sup_{\psi\neq 0}\frac{\left|\left((\mathcal{B}-\tilde{\mathcal{B}})\delta\mu,\psi\right)\right|}{\|\psi\|_1}\\ &\leq \sup_{\psi\neq 0}\frac{\left\|(\mathcal{B}-\tilde{\mathcal{B}})\delta\mu\right\|_0\|\psi\|_0}{\|\psi\|_0}.\end{aligned} \tag{67}$$

Hence,

$$\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\delta\mu,\psi)\right\|_{1*} \leq \left\|(\mathcal{B}-\tilde{\mathcal{B}})\delta\mu\right\|_0. \tag{68}$$

Following [26], we express

$$\begin{aligned}\left\|(\mathcal{B}-\tilde{\mathcal{B}})\delta\mu\right\|_0 &\approx 2\left\|\mathcal{A}^*(\mathcal{A}-\tilde{\mathcal{A}})\delta\mu\right\|_0\\ &\approx 2\left\|\sum_{i,j}^{N_S,N_D} g_j^*\phi_i\int_\Omega\left(g_j^*e_i+\phi_i e_j^*\right)\delta\mu d\boldsymbol{r}\right\|_0\\ &\leq 2\max_{i,j}\left\|g_j^*\phi_i\right\|_0\\ &\quad\times\sum_{i,j}^{N_S,N_D}\int_\Omega\left|\left(g_j^*e_i+\phi_i e_j^*\right)\delta\mu\right|d\boldsymbol{r}\end{aligned} \tag{69}$$

where $e_i = \phi_i - \Phi_i$ and $e_j^* = g_j^* - G_j^*$ are the discretization errors.

Then if we expand the integral on Ω as a summation of the integral on the finite element Ω_{mj}, $m = 1, \cdots, N_\Delta^{*j}$ and Ω_{ni}, $n = 1, \cdots, N_\Delta^i$, an upper bound for $\|(\mathcal{B} - \tilde{\mathcal{B}})\delta\mu\|_0$ becomes

$$\begin{aligned}
&\left\|(\mathcal{B} - \tilde{\mathcal{B}})\delta\mu\right\|_0 \\
&\quad \le 2 \max_{i,j} \left\|g_j^* \phi_i\right\|_0 \\
&\qquad \times \sum_{i,j}^{N_S, N_D} \left(\sum_{n=1}^{N_\Delta^i} \|e_i\|_{0,ni} \left\|g_j^* \delta\mu\right\|_{0,ni} \right. \\
&\qquad\qquad \left. + \sum_{m=1}^{N_\Delta^{*j}} \left\|e_j^*\right\|_{0,mj} \|\phi_i \delta\mu\|_{0,mj} \right) \\
&\quad \le 2 \max_{i,j} \left\|g_j^* \phi_i\right\|_0 \\
&\qquad \times \left(\sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} \|e_i\|_{0,ni} \left\|g_j^* \delta\mu\right\|_{0,ni} \right. \\
&\qquad\qquad \left. + \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} \left\|e_j^*\right\|_{0,mj} \|\phi_i \delta\mu\|_{0,mj} \right). \quad (70)
\end{aligned}$$

Next, following a similar argument as in (67) and (68), we obtain

$$\left\|(\tilde{\mathcal{G}} - \mathcal{G})(\psi)\right\|_{1*} \le \|\tilde{f} - f\|_0. \quad (71)$$

Using the definition of f in (32), an upper bound for $\|\tilde{f} - f\|_0$ is given by [26]

$$\begin{aligned}
\|\tilde{f} - f\|_0 &= \left\|(\tilde{\mathcal{A}}^* - \mathcal{A}^*)\delta\Gamma\right\|_0 \\
&\approx \left\| \sum_{i,j}^{N_S, N_D} \left(g_j^* e_i + \phi_i e_j^*\right) \delta\Gamma_{ij} \right\|_0 \\
&\le \max_{i,j} |\delta\Gamma_{ij}| \sum_{i,j}^{N_S, N_D} \left\|g_j^* e_i + \phi_i e_j^*\right\|_0. \quad (72)
\end{aligned}$$

A bound for $\max_{i,j} |\delta\Gamma_{ij}|$ can be obtained by using (18)

$$\max_{i,j} |\delta\Gamma_{ij}| \le \max_{i,j} \left\|g_j^* \phi_i\right\|_0 \|\delta\mu\|_0. \quad (73)$$

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}, $n = 1, \cdots, N_\Delta^i$

$$\begin{aligned}
\|\tilde{f} - f\|_0 \le \max_{i,j} \left\|g_j^* \phi_i\right\|_0 \|\delta\mu\|_0 &\left(\sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} \|e_i\|_{0,ni} \left\|g_j^*\right\|_{\infty,ni} \right. \\
&\left. + \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} \left\|e_j^*\right\|_{0,mj} \|\phi_i\|_{\infty,mj} \right). \quad (74)
\end{aligned}$$

In the end, using the discretization error estimates (44) and (45) leads to the theorem.

Appendix E
Proof of Theorem 2: Error Estimate Due to Inverse Problem Discretization

Taking (46) as the starting point, by coercivity we can write

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

Let $\varphi \in V_N$. Then the above inequality is equivalent to

$$\begin{aligned}
\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1^2 &\le \frac{1}{\min(\lambda_1, \lambda_2)} \Big(\widetilde{\mathcal{F}}(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \widetilde{\delta\mu} - \varphi) \\
&\qquad + \widetilde{\mathcal{F}}(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \varphi - \widetilde{\Delta\mu}) \Big) \\
&\le \frac{1}{\min(\lambda_1, \lambda_2)} \widetilde{\mathcal{F}}(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \widetilde{\delta\mu} - \varphi)
\end{aligned}$$

since $\varphi - \widetilde{\Delta\mu} \in V_N$ and the error $\widetilde{\delta\mu} - \widetilde{\Delta\mu}$ is orthogonal to the finite-dimensional subspace with respect to the norm induced by the bilinear form. Noting

$$\begin{aligned}
&\widetilde{\mathcal{F}}(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \widetilde{\delta\mu} - \varphi) = \Big(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \tilde{\mathcal{B}}(\widetilde{\delta\mu} - \varphi)\Big) \\
&+\lambda_1(\widetilde{\delta\mu} - \widetilde{\Delta\mu}, \widetilde{\delta\mu} - \varphi) + \lambda_2 \Big(\nabla(\widetilde{\delta\mu} - \widetilde{\Delta\mu}), \nabla(\widetilde{\delta\mu} - \varphi)\Big) \quad (75)
\end{aligned}$$

it is clear

$$\begin{aligned}
&\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1^2 \\
&\quad \le \frac{1}{\min(\lambda_1, \lambda_2)} \Big\{ \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_0 \left\|\tilde{\mathcal{B}}(\widetilde{\delta\mu} - \varphi)\right\|_0 \\
&\qquad + \lambda_1 \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_0 \|\widetilde{\delta\mu} - \varphi\|_0 \\
&\qquad + \lambda_2 \|\nabla\widetilde{\delta\mu} - \nabla\widetilde{\Delta\mu}\|_0 \|\nabla\widetilde{\delta\mu} - \nabla\varphi\|_0 \Big\} \\
&\quad \le \frac{1}{\min(\lambda_1, \lambda_2)} \Big\{ \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \|\tilde{\mathcal{B}}(\widetilde{\delta\mu} - \varphi)\|_0 \\
&\qquad + \lambda_1 \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \|\widetilde{\delta\mu} - \varphi\|_0 \\
&\qquad + \lambda_2 \|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \|\nabla\widetilde{\delta\mu} - \nabla\varphi\|_0 \Big\}. \quad (76)
\end{aligned}$$

Cancel $\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1$ terms

$$\begin{aligned}
\|\widetilde{\delta\mu} - \widetilde{\Delta\mu}\|_1 \le \frac{1}{\min(\lambda_1, \lambda_2)} &\Big\{ \left\|\tilde{\mathcal{B}}(\widetilde{\delta\mu} - \varphi)\right\|_0 + \lambda_1 \|\widetilde{\delta\mu} - \varphi\|_0 \\
&+ \lambda_2 \|\nabla\widetilde{\delta\mu} - \nabla\varphi\|_0 \Big\}. \quad (77)
\end{aligned}$$

Let $\varphi \in V_N$ be the interpolant of $\widetilde{\delta\mu}$ and $e_\mu = \widetilde{\delta\mu} - \varphi$ be the interpolation error. Then the first term in the bound (77) can be expanded as follows:

$$\begin{aligned}
&\left\|\tilde{\mathcal{B}}(\widetilde{\delta\mu} - \varphi)\right\|_0 \\
&\quad = \|\tilde{\mathcal{B}} e_\mu\|_0 = \left\| \int_\Omega \tilde{\kappa}(\cdot, \boldsymbol{r}') e_\mu(\boldsymbol{r}') d\boldsymbol{r}' \right\|_0 \\
&\quad = \left\| \int_\Omega \sum_{i,j}^{N_S, N_D} G_j^*(\cdot) \Phi_i(\cdot) G_j^*(\boldsymbol{r}') \Phi_i(\boldsymbol{r}') e_\mu(\boldsymbol{r}') d\boldsymbol{r}' \right\|_0 \\
&\quad \le \sum_{i,j}^{N_S, N_D} \left\|G_j^* \Phi_i\right\|_0 \int_\Omega \left| G_j^*(\boldsymbol{r}') \Phi_i(\boldsymbol{r}') e_\mu(\boldsymbol{r}') \right| d\boldsymbol{r}'.
\end{aligned}$$

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

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

The remaining two terms in (77) can be expressed in a straightforward way

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

Assume that our solution also satisfies $\widetilde{\delta\mu} \in H^2(\Omega)$. Then a bound for the interpolation error and its gradient on each element can be given by [38]

$$\|e_\mu\|_{0,t} \leq C\|\mu\|_{2,t}h_t^2 \tag{78}$$
$$\|\nabla e_\mu\|_{0,t} \leq C\|\mu\|_{2,t}h_t \tag{79}$$

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 (78) and (79) into (77) proves the theorem.

REFERENCES

[1] 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.*, vol. 42, pp. 3081–3094, 2003.

[2] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Adaptive finite element based tomography for fluorescence optical imaging in tissue," *Opt. Expr.*, vol. 12, no. 22, pp. 5402–5417, 2004.

[3] M. J. Eppstein, F. Fedele, J. P. Laible, C. Zhang, A. Godavarty, and E. M. Sevick-Muraca, "A comparison of exact and approximate adjoint sensitivities in fluorescence tomography," *IEEE Trans. Med. Imag.*, vol. 22, no. 10, pp. 1215–1223, Oct. 2003.

[4] R. Roy, A. Godavarty, and E. M. Sevick-Muraca, "Fluorescence-enhanced optical tomography using referenced measurements of heterogeneous media," *IEEE Trans. Med. Imag.*, vol. 22, no. 7, pp. 824–836, Jul. 2003.

[5] A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized born ratio," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1377–1386, Oct. 2005.

[6] V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation," *Opt. Lett.*, vol. 26, no. 12, pp. 893–895, 2001.

[7] V. Ntziachristos, G. Turner, J. Dunham, S. Windsor, A. Soubret, J. Ripoll, and H. A. Shih, "Planar fluorescence imaging using normalized data," *J. Biomed. Opt.*, vol. 10, no. 6, pp. 1–8, 2005.

[8] S. R. Arridge, "Optical tomography in medical imaging," *Inv. Probl.*, vol. 15, pp. R41–R93, 1999.

[9] A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Phys. Med. Biol.*, vol. 50, pp. R1–R43, 2005.

[10] V. A. Markel and J. C. Schotland, "Inverse problem in optical diffusion tomography. II. Role of boundary conditions," *J. Opt. Soc. Am. A*, vol. 19, no. 3, pp. 1–9, 2002.

[11] V. A. Markel, V. Mital, and J. C. Schotland, "Inverse problem in optical diffusion tomography. III. Inversion formulas and singular-value decomposition," *J. Opt. Soc. Am. A*, vol. 20, no. 5, pp. 890–902, 2003.

[12] 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," *Inv. Probl.*, vol. 23, pp. 1135–1160, 2007.

[13] 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," *Inv. Probl.*, vol. 22, pp. 175–195, 2006.

[14] 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.

[15] I. Babuška and W. C. Rheinboldt, "Error estimates for adaptive finite element computations," *SIAM J. Numer. Anal.*, vol. 15, pp. 736–754, 1978.

[16] I. Babuška, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira, *Accuracy Estimates and Adaptive Refinements in Finite Element Computations*. New York: Wiley, 1986.

[17] R. E. Bank and A. Weiser, "Some a posterior error estimators for elliptic partial differential equations," *Math. Comput.*, vol. 44, pp. 283–301, 1985.

[18] T. Strouboulis and K. A. Hague, "Recent experiences with error estimation and adaptivity, Part I: Review of error estimators for scalar elliptic problems," *Comput. Meth. Appl. Mech. Eng.*, vol. 97, pp. 399–436, 1992.

[19] R. Verfurth, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*. New York: Teubner-Wiley, 1996.

[20] J. T. Oden and S. Prudhomme, "Goal-oriented error estimation and adaptivity for the finite element method," *Comput. Math. Applicat.*, vol. 41, pp. 735–756, 2000.

[21] V. Heuveline and R. Rannacher, "Duality-based adaptivity in the hp-finite element method," *J. Numer. Math.*, vol. 11, pp. 95–113, 2003.

[22] L. Beilina and C. Johnson, "A posteriori error estimation in computational inverse scattering," *Math. Models Meth. Appl. Sci.*, vol. 15, pp. 23–37, 2005.

[23] L. Beilina and C. Johnson, "Adaptive finite element/difference method for inverse elastic scattering waves," *Appl. Comput. Math.*, vol. 2, pp. 158–174, 2003.

[24] W. Bangerth, "Adaptive finite element methods for the identification of distributed parameters in partial differential equations," Ph.D. dissertation, Univ. Heidelberg, Heidelberg, Germany, 2002.

[25] R. Li, W. Liu, H. Ma, and T. Tang, "Adaptive finite element approximation for distributed elliptic optimal control problems," *SIAM J. Contr. Optim.*, vol. 41, pp. 1321–1349, 2002.

[26] 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," *Inv. Probl.*, vol. 23, pp. 1115–1133, 2007.

[27] J. Kaipio and E. Somersalo, *Computational and Statistical Inverse Problems*, ser. Applied Mathematical Sciences. New York: Springer-Verlag, 2005, vol. 160.

[28] M. Guven, L. Zhou, L. Reilly-Raska, and B. Yazici, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography: Part II," *IEEE Trans. Med. Imag.*, to be published.

[29] 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.*, vol. 187, pp. 597–619, 2003.

[30] R. Kress, *Linear Integral Equations*, ser. Applied Mathematical Sciences, 2nd ed. New York: Springer-Verlag, 1999, vol. 82.

[31] M. Guven, B. Yazici, and V. Ntziachristos, "Optical fluorescence tomography with a priori information," *Proc. SPIE*, vol. 6431, 2007.

[32] 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 Process.*, vol. 1, pp. 322–336, Jul. 1992.

[33] M. Hanke and T. Raus, "A general heuristic for choosing the regularization parameter in ill-posed problems," *SIAM J. Sci. Comput.*, vol. 17, pp. 956–972, 1996.

[34] R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and regularization methods for hyperparameter estimation in image restoration," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 231–246, Feb. 1999.

[35] M. Heath, G. Golub, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, p. 215C223, 1979.

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

[37] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*. New York: Springer Verlag, 2002.

[38] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*. New York: Springer-Verlag, 2002.

[39] D. Daners, "Robin boundary value problems on arbitrary domains," *Trans. Am. Math. Soc.*, vol. 352, no. 9, pp. 4207–4236, 2000.

[40] 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.

[41] I. M. Smith and D. V. Griffiths, *Programming the Finite Element Method*, 4th ed. New York: Wiley, 2004.