Fluorescence optical tomography with a priori information (Invited Paper)

Murat Guven[a], Birsen Yazici[a], Vasilis Ntziachristos[b]

[a]Rensselaer Polytechnic Institute, Troy, NY, USA;
[b]Harvard Medical School, Charlestown, MA, USA

ABSTRACT

In this work, we discuss the incorporation of *a priori* information into the inverse problem formulation for fluorescence optical tomography. In this respect, we first formulate the inverse problem in the optimization framework which allows the incorporation of *a priori* information about the solution and its gradient. Then, we consider the variational problem, which is equivalent to the optimization problem and prove the existence and uniqueness of the solution. Finally, we discuss the design of the functions that incorporate the *a priori* information into the inverse problem formulation and present a model problem to illustrate the design procedure.

Keywords: *a priori* information, fluorescence, optical, regularization, estimation, variational formulation

1. INTRODUCTION

Fluorescence optical tomography (FOT) offers the quantification, $3D$ imaging and depth retrieval of fluorescence activity with high sensitivity, which in turn can be used for functional and molecular characterization of normal and diseased tissues.[1–3]

Under the assumption of weak fluorescence, the inverse problem in FOT is characterized by a linear integral equation, which is typical in many biological systems.[2] Despite the simplifications granted by the linear model, the ill-posed nature of this integral equation makes FOT a challenging inverse problem. In particular, due to the ill-posedness, the solution of the inverse problem is not unique. Furthermore, the ill-posedness makes the solution highly sensitive to the noise in the measurements. Hence, this linear integral equation needs to be regularized. The regularization can be within the statistical[4] or deterministic framework.[5] In general, regularization will address both the non-uniqueness problem and the potential instability of the solution due to the presence of noise. On the other hand, regularization can be more effective if *a priori* information is incorporated into the inverse problem formulation. In this respect, the *a priori* information can be used to increase the information content of the imaging problem, thereby improving the accuracy of the solution.

To tackle the ill-posed nature of the inverse problem in diffuse optical tomography (DOT), a number of approaches have been suggested to incorporate *a priori* information. See[6] and[7] and the references therein for an extensive survey on the incorporation of *a priori* information in DOT. In general, the *a priori* information in DOT is richer in content as compared to the *a priori* information in FOT. For example, in optical absorption imaging,[6,8] given an MR-image in prior, the discrimination of the tissue types is possible, which may be used to determine the average optical absorption coefficient properties of the medium under inspection. This is enabled by the high spatial correlation between the anatomical images and the corresponding optical properties of the medium.[9–11] It is not viable to obtain any average value for the fluorophore absorption value inside the optical medium, which is related to the concentration of the fluorophore. Yet, it is possible to obtain the localization information of a certain fluorophore type based on its chemical and physical properties. For example, the fluorophore may be targeted to the lungs. Then, an anatomical image can be useful to constrain the location of the fluorophore concentration.[12] In addition, the anatomical image can be used to determine the constraints on the spatial variations and smoothness of the solution. In this context, the edges in the anatomical image can be assumed to partially coincide with the edges in the fluorophore concentration image. Hence, the edges in the anatomical image can be used as *a priori* information as well.

Send correspondence to Birsen Yazici: E-mail: yazici@ecse.rpi.edu

One important restriction in the use of *a priori* information should be avoiding hard constraints, due to the limited content and uncertainty in the reliability of the *a priori* information. Constraining the location of the fluorophore concentration, thereby restricting the boundaries of the support of the solution can be considered as an example of hard constraints.[12] While computationally tractable, this approach may lead to erroneous results. Such explicit constraints can be weakened by using them implicitly in the formulation of the inverse problem. For instance, Tikhonov functionals with zeroth order spatially varying regularization terms can be used to weaken such constraints.[13]

To improve the information content of the inverse problem formulation, it is important to make use of all the available *a priori* information about the solution. The fluorophore localization (target) information, the internal edge structure of the medium, and the source-detector positions are examples of different types of information. Including as much *a priori* information will have two important outcomes: First, the resulting inverse problem formulation will be tuned to the solution of the problem at hand. Second, different types of information can complement each other and/or compensate for the correlation uncertainty of the other.

In this work, we use the *a priori* information about the localization of the fluorescent agent and internal structure of the medium to formulate the inverse FOT problem within the regularization framework. For the effective use of *a priori* information and convenient design of the penalty functions, we first formulate the inverse problem as a minimization problem where the regularization terms impose penalty simultaneously on the solution and its gradients. In this respect, we present convenient methods for the design of the regularization terms, which is driven directly by the information about the localization of the fluorescent agent and the internal edges of the medium of interest, provided by an anatomical imaging modality, such as MRI or x-ray. This makes the inverse problem formulation described in this paper different from the previously reported approaches.[12, 13] Next, to solve the inverse problem, we translate the minimization problem to the variational framework, and use finite elements to discretize the resulting variational problem. Finally, we show the uniqueness of the solutions to both the variational problem and its discretized version.

The outline of this paper is as follows: Section 2 defines the forward problem in OFT and describes the formulation of the inverse problem in the optimization and variational frameworks, respectively. In Section 3, we discuss how the available *a priori* information can be incorporated into the inverse problem formulation. In Section 4, we describe the discretization of the inverse problem using finite elements, which is followed by an example which presents the procedure for incorporating the *a priori* information into the inverse problem formulation. Following the Conclusion section, the Appendix presents proofs of the Lemmas which show the uniqueness and boundedness of the solution of the inverse problem.

2. FORWARD AND INVERSE PROBLEMS IN FOT

In this section, we first state our assumptions which simply the fluorescence imaging model. Then, we define the forward problem in FOT, which models the propagation of the excitation and emission NIR light. Next, we briefly describe the inverse FOT problem, whose objective is to estimate the unknown fluorophore absorption from the detected optical field on the boundary. Then, we formulate the inverse problem in the regularization framework which allows us to incorporate *a priori* information into the inverse problem formulation. Finally, we formulate the inverse FOT problem in the variational framework and discuss the existence and uniqueness of the solution to the resulting inverse problem.

2.1. Forward Problem

We first state the assumptions that simplify the fluorescence diffuse optical tomography (FOT) analysis:

- We focus on the estimation of the unknown fluorophore concentration. In this respect, the amplitude of the excitation light source is time-invariant, which avoids the dependence of the inverse problem formulation on the life-time of the fluorophore.
- The excitation wavelength measurements are collected before the injection of the fluorophore. Hence, the background absorption at the excitation wavelength consists of only the endogenous absorption coefficient.

- The endogenous absorption $\mu_a^{\lambda_1}$ is assumed to be equal to $\mu_a^{\lambda_2}$, assuming that the absorption does not vary in a small range of wavelength.
- For the weak fluorescence case, the contribution $\mu_{af}^{\lambda_2}$ of the fluorophore to the total absorption at the emission wavelength can be neglected, thus $\mu_a^{\lambda_2} = \mu_a^{\lambda_1} = \mu_a$.
- The diffusion coefficient $D(\mathbf{x})$ is identical during both the excitation and emission, for all $\mathbf{x}$ on the bounded domain and its boundary.

Based on these assumptions, we use the following boundary value problem to model the NIR light propagation at the excitation wavelength λ_1, on a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$[14, 15]:

$$-\nabla \cdot D(\mathbf{x})\nabla u_j^{\lambda_1}(\mathbf{x}) + \mu_a(\mathbf{x})u_j^{\lambda_1}(\mathbf{x}) = Q_j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{1}$$

$$u_j^{\lambda_1}(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial u_j^{\lambda_1}}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{2}$$

where $u_j^{\lambda_1}(\mathbf{x})$ is the isotropic photon density at the excitation wavelength λ_1 due to the j^{th} source at $\mathbf{x}$, Q_j is the point source located at $\mathbf{x}_s^j$, $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, a is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial\cdot/\partial n$ denotes the directional derivative along the unit normal vector on the boundary. Note that we assume the diffusion coefficient is isotropic. For the general anisotropic material, see.[16]

At the emission wavelength λ_2, the photon density due to the fluorophore concentration and the incident excitation field $u_j^{\lambda_1}$ is given by the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla u_j^{\lambda_2}(\mathbf{x}) + \mu_a(\mathbf{x})u_j^{\lambda_2}(\mathbf{x}) = Q_y\mu_{af}(\mathbf{x})u_j^{\lambda_1}(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{3}$$

$$u_j^{\lambda_1}(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial u_j^{\lambda_1}}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{4}$$

where $u_j^{\lambda_2}$ is the photon density at the emission wavelength λ_2 due to the fluorophore concentration excited by the photon field $u_j^{\lambda_1}$.

Finally, we consider the adjoint problem[15] associated with (1)-(2), whose solution is equivalent to the Green's function of (3)-(4)[8, 15]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^{*,\lambda_2}(\mathbf{x}) + \mu_a(\mathbf{x})g_i^{*,\lambda_2}(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \tag{5}$$

$$g_i^{*,\lambda_2}(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^{*,\lambda_2}}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \tag{6}$$

where Q_i^* is the adjoint source located at $\mathbf{x}_d^i$. In this work, we approximate the point source Q_j in (1) and the adjoint source Q_i^* in (6) by Gaussian functions with sufficiently low variance, whose centers are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively.[8]

2.2. Fluorescence Imaging Model

In the following, we state the fluorescence imaging model that gives the relationship between the unknown fluorophore absorption coefficient and the detected optical field at the emission wavelength, based on the forward model in Section 2.1.

The detected field at the emission wavelength λ_2 due to the incident field $u_j^{\lambda_1}$ is related to the unknown fluorophore concentration through the following linear integral equation:

$$\tilde{\Gamma}_{i,j} = \int_\Omega g_i^{*,\lambda_2}(\mathbf{x})u_j^{\lambda_1}(\mathbf{x})\Lambda^{i,j}(\mathbf{x})d\mathbf{x}, \tag{7}$$

where $\tilde{\Gamma}_{i,j}$ represents the emission measurement as a result of the fluorophore concentration at the i^{th} detector due to the j^{th} excitation source and $\Lambda^{i,j}(\mathbf{x})$ stands for

$$\Lambda^{i,j}(\mathbf{x}) := \left[Q_y \Theta_f Q_{eff}^{\lambda_2} \Theta_s^{\lambda_2}(\mathbf{x}_s^j) \Theta_d^{\lambda_2}(\mathbf{x}_d^i)\right] \mu_{af}^{\lambda_2}(\mathbf{x}), \tag{8}$$

where $\mu_{af}^{\lambda_2}$ is the fluorophore absorption, Q_y is the quantum yield, Θ_f is the attenuation of the filter at the detector side, $Q_{eff}^{\lambda_2}$ is the detector quantum efficiency and $\Theta_s^{\lambda_2}(\mathbf{x}_s)$, $\Theta_d^{\lambda_2}(\mathbf{x}_d)$ are the gain factors of the source and detector, respectively, modeling the light source strength, fiber coupling losses and attenuation in the system. λ_2 superscript is due to the fact that the detection is at wavelength λ_2.

Normalization of the measurement data $\tilde{\Gamma}_{i,j}$with respect to the detected field $u_{j,inc}^{\lambda_1}(\mathbf{x}_d^i)$ at the excitation wavelength leads to:

$$\Gamma_{i,j}^{norm} = \frac{\tilde{\Gamma}_{i,j}}{u_{j,inc}^{\lambda_1}(\mathbf{x}_d^i)} = \frac{1}{u_{j,inc}^{\lambda_1}(\mathbf{x}_d^i)} \int_\Omega g_i^{*,\lambda_2}(\mathbf{x}) u_j^{\lambda_1}(\mathbf{x}) \Lambda^{i,j}(\mathbf{x}) d\mathbf{x}, \tag{9}$$

where $u_{j,inc}^{\lambda_1}(\mathbf{x}_d^i)$ is given by

$$u_{j,inc}^{\lambda_1}(\mathbf{x}_d^i) = Q_{eff}^{\lambda_1} \Theta_s^{\lambda_1}(\mathbf{x}_s^j) \Theta_d^{\lambda_1}(\mathbf{x}_d^i) u_j^{\lambda_1}(\mathbf{x}_d^i). \tag{10}$$

Noting (8) and assuming the following conditions hold

$$\frac{\Theta_s^{\lambda_1}}{\Theta_s^{\lambda_2}} \approx 1, \quad \frac{\Theta_d^{\lambda_1}}{\Theta_d^{\lambda_2}} \approx 1, \quad \frac{Q_{eff}^{\lambda_1}}{Q_{eff}^{\lambda_2}} \approx 1,$$

we arrive at the following linear integral equation, that relates the fluorophore concentration to detected emission measurements:

$$\Gamma_{i,j}^{norm} = \frac{\Theta_f^{\lambda_2}}{u^{\lambda_1}(\mathbf{x}_d^i; \mathbf{x}_s^j)} \int_\Omega g_i^{*,\lambda_2}(\mathbf{x}) u_j^{\lambda_1}(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x}, \tag{11}$$

where $\alpha(\mathbf{x}) := Q_y \mu_{af}(\mathbf{x})$. We note that in the following sections, (11) will constitute the basis for the inverse problem formulation.

2.3. Inverse Problem

In this section, we first give some preliminary information about the components of the inverse problem. Next, we formulate the inverse problem in the regularization framework which provides a means to incorporate *a priori* information about the unknown image α. Then, by defining appropriate boundary conditions, we convert the inverse problem into a boundary value problem, which is followed by the variational formulation of the inverse problem. Finally, we show the existence and uniqueness of the solution to the resulting variational problem.

2.3.1. Preliminaries

We recall the integral equation (11) where we drop the superscripts λ_2 and λ_1 from g^{*,λ_2} and u^{λ_1}, respectively. Letting

$$\Gamma_{i,j} = \Gamma_{i,j}^{norm} \frac{u_j^{\lambda_1}(\mathbf{x}_d^i)}{\Theta_f^{\lambda_2}},$$

we have

$$\begin{aligned} \Gamma_{i,j} &= \int_\Omega g_i^*(\mathbf{x}) u_j(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} \\ &= \int_\Omega H_{i,j}(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} \\ &:= A_f \alpha, \end{aligned} \tag{12}$$

where $H_{i,j} = g_i^* u_j$ is the kernel of the integral operator $A_f : L^2(\Omega) \to \mathbb{R}^{N_s \times N_d}$. If we consider Gaussian approximations of the points sources at the excitation wavelength and a Gaussian approximation for the adjoint

source at the emission wavelength, then the weak solutions $g^*, u \in C^1(\Omega)$.[8] As a result, a bound for the linear integral operator A_f can be given by:

$$\|A_f\|_{L^2(\Omega)\to l^1} \leq \sum_{i,j}^{N_d,N_s} \|g_i^* u_j\|_0. \tag{13}$$

Furthermore, by the boundedness and the finite-dimensional range of the operator, A_f is compact.[8,17] As a result, the inverse problem given by (12) is ill-posed.

Let $A = A_f^* A_f$, then $A : L^2(\Omega) \to L^2(\Omega)$ is defined as follows:

$$\begin{aligned}(A\alpha)(\mathbf{x}) &:= \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) \int_\Omega H_{i,j}(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}} \\ &= \int_\Omega \kappa(\mathbf{x},\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}},\end{aligned} \tag{14}$$

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

$$\kappa(\mathbf{x},\acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) H_{i,j}(\acute{\mathbf{x}}), \tag{15}$$

and $H_{i,j}^*(\mathbf{x})$ is the kernel of the adjoint operator A_f^* given by:

$$(A_f^*\beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})u_j(\mathbf{x})\beta_{i,j}, \tag{16}$$

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

Then, equivalently, we can express (12) as follows:

$$\gamma(\mathbf{x}) = (A\alpha)(\mathbf{x}), \tag{17}$$

where $\gamma(\mathbf{x}) = A_f^* \Gamma_{i,j}$. We note that the operator $A = A_f^* A_f : L^2(\Omega) \to L^2(\Omega)$ is compact and thus bounded, which results from the compactness of the operator A_f.

2.3.2. Inverse problem as an optimization problem

In order to incorporate *a priori* information, we formulate the inverse problem as an optimization problem:

$$\hat{\alpha} = \min_{\alpha \in H^1(\Omega)} J(\alpha), \tag{18}$$

where $J(\alpha) : H^1(\Omega) \to \mathbb{R}$ is the objective functional to be minimized and $\Omega \subset \mathbb{R}^n$, $n = 2, 3$.

To incorporate *a priori* information, we consider penalty terms in addition to the data likelihood term $\int_\Omega \big[\gamma(\mathbf{x}) - (A\alpha)(\mathbf{x})\big]^2 d\mathbf{x}$ in the objective functional. As a result, the inverse problem becomes

$$\begin{aligned}\hat{\alpha} &= \min_{\alpha \in H^1(\Omega)} \int_\Omega F(\mathbf{x}, \alpha, \alpha_t) d\mathbf{x} \\ &= \min_{\alpha \in H^1(\Omega)} \int_\Omega \big[\gamma(\mathbf{x}) - (A\alpha)(\mathbf{x})\big]^2 d\mathbf{x} + \beta_1 \int_\Omega q(\mathbf{x})\big[\alpha(\mathbf{x}) - \alpha^m(\mathbf{x})\big]^2 d\mathbf{x} + \beta_2 \int_\Omega \sum_{t=1}^{n} p_t(\mathbf{x})\big[\frac{\partial \alpha}{\partial \mathbf{x}_t}(\mathbf{x}) - \acute{\alpha}^t(\mathbf{x})\big]^2 d\mathbf{x},\end{aligned} \tag{19}$$

where $\mathbf{x} = [x_1, \cdots, x_n] \in \Omega$ is the position vector, $q \in L^\infty(\Omega)$ is a positive function that weights the penalty on the error between the image and the image model α_m; $p_t \in L^\infty(\Omega)$, $t = 1, \cdots, n$ is a positive function that

weights the penalty on the error between the image gradient $\alpha_t = \partial\alpha/\partial x_t$ and the model of the image gradient $\acute{\alpha}^t$. $\beta_1, \beta_2 > 0$ are the regularization parameters.

In (19), the first penalty term can be used to constrain the solution on only admissible local regions inside the domain, by using the weighting function q properly. By the careful choice of p_t, the second penalty term can serve both to eliminate and preserve the high-frequency components of the image α. Note that p_t ensures an anisotropic penalty term for the image gradient. The scalars β_1 and β_2 control the balance between the data-fit term, the first, and the second penalty terms. We finally note that the regularization terms address the ill-posedness of (17) as well.

Note that if there is *a priori* quantitative spatial information about the fluorophore absorption value, then one can design an image model α_m and/or $\acute{\alpha}^t$. Otherwise, the models can be taken to be zero. In this case, the weighting functions q and p_t can be designed based on the available *a priori* information (see Section 3).

The convexity of the function $F(\mathbf{x}, \alpha, \alpha_t)$ with respect to the variable α_t ensures a minimizer for the optimization problem in (19).[18] Clearly

$$\begin{aligned} \big|\frac{\partial F}{\partial \alpha}(\mathbf{x}, \alpha, \alpha_t)\big| &\leq C_1(1 + |\alpha| + \big|\frac{\partial \alpha}{\partial \mathbf{x}_t}\big|^2), \\ \big|\frac{\partial F}{\partial \alpha_t}(\mathbf{x}, \alpha, \alpha_t)\big| &\leq C_2(1 + |\alpha|^2 + \big|\frac{\partial \alpha}{\partial \mathbf{x}_t}\big|), \end{aligned}$$

for some constant $C_1 > 0$ and $C_2 > 0$, as a result of the quadratic terms in the integrand $F(\mathbf{x}, \alpha, \alpha_t)$. Hence, the minimizer of (19) satisfies the following integro-differential equation. This is a necessary condition for the objective functional $J(\alpha)$ in (19) to have an extremum:

$$\frac{\partial F}{\partial \alpha} - \sum_t^n \frac{\partial}{\partial x_t}\frac{\partial F}{\partial \alpha_t} = 0. \tag{20}$$

In particular, we have the following lemma.

Lemma 1: Let

$$f(\mathbf{x}) := A^*\gamma(\mathbf{x}) + \beta_1 q(\mathbf{x})\alpha^m(\mathbf{x}) + \beta_2 \sum_{t=1}^n \big[p_t(\mathbf{x})\acute{\alpha}_t(\mathbf{x})\big]_t. \tag{21}$$

Then, the minimizer $\hat{\alpha}$ of the functional in (19) satisfies the following integro-differential equation for $\mathbf{x} \in \Omega$:

$$A^*A\alpha(\mathbf{x}) + \beta_1 q(\mathbf{x})\alpha(\mathbf{x}) - \beta_2 \sum_{t=1}^n [p_t(\mathbf{x})\frac{\partial \alpha}{\partial x_t}(\mathbf{x})]_t \quad = \quad f(\mathbf{x}). \tag{22}$$

Proof. A necessary condition for $J(\alpha)$ to have an extremum for a given function $\hat{\alpha}$ is that $\hat{\alpha}$ satisfy the Euler-Lagrange equation, which can be obtained by computing the first variation of $J(\alpha)$ in (19). Using (21) and (20) results in Lemma 1. □

Remarks:

1. Note that for $\beta_1, \beta_2 = 0$, the solution of the minimization problem satisfies (17).
2. Penalizing the image imposes a penalty on the image gradient as well; however the converse is not true.
3. The integro-differential equation requires that the first and second derivatives of α exist. We note that the solution for which the functional has its extremum may not satisfy these requirements. In the following section, we will formulate the inverse problem in the variational framework, which seeks a less smooth solution as compared to (22).
4. We note that the operator A is self-adjoint, hence $A^* = A$.

3. INCORPORATION OF THE A PRIORI INFORMATION

In order to incorporate the *a priori* information, we first turn back to the minimization problem (19). Next, we consider the corresponding weak problem with the assigned boundary conditions.

3.1. The penalty on the image α

The second term in (19) penalizes the difference between the solution α and an *a priori* image model α^m, depending on the spatially varying function q. Hence, we need to decide about q and α^m, together with the positive parameter β_1 that controls the impact of the penalty term. In general, there is no exact prior information about the magnitude of the solution, which can be used as the image model α^m. On the other hand, one can know the spatial distribution of the fluorophore concentration, that is, where the fluorophore is expected to exist and where it is expected *not* to exist. Let α_b^m be a piece-wise constant function defined as follows:

$$\alpha_b^m(\mathbf{x}) := \begin{cases} 1, & \mathbf{x} \in \Omega_e \\ 0, & \mathbf{x} \in \left(\Omega \setminus \Omega_e\right) \end{cases},$$

where Ω_e is the region where the fluorophore is expected to exist. To model the uncertainties about the boundaries of α_b^m, we use a smoothing Gaussian kernel to obtain

$$\tilde{\alpha}_b^m(\mathbf{x}) := \int_\Omega G_\sigma(\mathbf{x}\text{-}\mathbf{y})\alpha_b^m(\mathbf{y})d\mathbf{y},$$

where G_σ is a normalized Gaussian kernel with equal variance σ^2. Then, q can be defined as

$$q(\mathbf{x}) := \frac{Q}{\epsilon + \tilde{\alpha}_b^m(\mathbf{x})}, \ \mathbf{x} \in \Omega, \tag{23}$$

where, $Q > 0$ is a scalar, and $0 < \epsilon \ll 1$. Note that $q \in L^\infty(\Omega)$. We note that one way to obtain an image model α^m is to formulate the least squares problem ($\beta_1 = \beta_2 = 0$) and solve the resulting weak problem (14) on a finite-dimensional sub-space with a low dimension, such that the resulting linear equations do not pose an increase on the computational complexity.

3.2. The penalty on the image gradient $\partial\alpha/\partial x_t$

The last term in (19) penalizes the deviations of the solution gradient from a prior image gradient model, $\acute{\alpha}^t$, for $t = 1, \cdots, n$. The function p_t provides a spatially varying penalty term on the gradient. In general, it is not possible to find an exact image gradient model $\acute{\alpha}^t$, thus one can assume that $\acute{\alpha}^t = 0$. In this case, p_t can be used to suppress or allow the variations in the image gradient, by appropriately designing p_t.

Let α_s^m be a segmented image which shows the internal structure of the medium of interest. Depending on the criterion for the segmentation, $\alpha_s^m(\mathbf{x})$ may be identical to $\alpha_b^m(\mathbf{x})$, or $\alpha_s^m(\mathbf{x})$ can be given by

$$\alpha_s^m(\mathbf{x}) := r, \quad \mathbf{x} \in \Omega_r, \quad r = 0, \cdots, R-1,$$

where $\bigcup_{r=0}^{R-1} \Omega_r = \Omega$, and R is the total number of labels in the segmented image. For example, each label may indicate a different tissue type, or organ, in which the fluorophore may exist or not. In this case, $\alpha_s^m(\mathbf{x})$ differs from $\alpha_b^m(\mathbf{x})$ due to the difference in the segmentation criterion.

Next, we consider the edge map e extracted from the segmented image α_s^m, which can be obtained by applying a simple edge detection algorithm on the segmented image α_s^m:

$$e(\mathbf{x}) := \begin{cases} 1, & \|\mathbf{x} - \mathbf{x}_e\| \leq d, \ x_e \text{ on an edge} \\ 0, & \text{elsewhere} \end{cases},$$

where $\|\mathbf{x} - \mathbf{x}_e\|$ is the Euclidean distance between $\mathbf{x}$ and an edge point $\mathbf{x}_e$, and d is a predefined distance value. Then, we apply a smoothing kernel on the edge map e and obtain:

$$\tilde{e}(\mathbf{x}) := \int_\Omega G_\sigma(\mathbf{x}\text{-}\mathbf{y})e(\mathbf{y})d\mathbf{y}. \tag{24}$$

The smoothing kernel G_σ provides an uncertainty bound for the location of the edges. Furthermore, the resulting function $\tilde{e}$ becomes sufficiently smooth to take derivatives. Then, we can define p_t as follows:

$$p_t(\mathbf{x}) := P_t / \left(\epsilon + \left| \frac{\partial \tilde{e}}{\partial x_t} \right| \right), \mathbf{x} \in \Omega, \tag{25}$$

where $P_t > 0$ is scaling factor. As a result, p_t is large where there is no edge, and small where there is an edge. Note that $p_t(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega$ and $p_t \in L^\infty(\Omega)$.

3.2.1. Inverse problem formulation in the variational framework

The inverse problem in the optimization framework involves a functional minimization. Instead, we consider the variational formulation of the inverse problem, after defining appropriate boundary conditions.

Consider (22) with the boundary condition:

$$p_t(\mathbf{x}) \frac{\partial \alpha}{\partial \hat{n}_t}(\mathbf{x}) = B_t(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega. \tag{26}$$

where $B_t \in H^1(\Omega)$.

In order to obtain the variational (weak) problem, we multiply both sides of (22) by a function $\psi \in H^1(\Omega)$, and integrate on Ω:

$$\int_\Omega \psi(\mathbf{x}) \big[A^* A\alpha(\mathbf{x}) + \beta_1 q(\mathbf{x}) \alpha(\mathbf{x}) \big] d\mathbf{x} - \beta_2 \int_\Omega \psi(\mathbf{x}) \sum_{t=1}^{n} \big[p_t(\mathbf{x}) \frac{\partial \alpha}{\partial x_t}(\mathbf{x}) \big]_t d\mathbf{x} = \int_\Omega \psi(\mathbf{x}) f(\mathbf{x}). \tag{27}$$

Noting $p_t \in L^\infty(\Omega)$ and $\alpha \in H^1(\Omega)$, we can apply the integration by parts formula for the second integral on the left hand side of (27).[14] As a result

$$\begin{aligned} \int_\Omega \psi(\mathbf{x}) \big[A^* A\alpha(\mathbf{x}) + \beta_1 q(\mathbf{x}) \alpha(\mathbf{x}) \big] d\mathbf{x} \quad &+ \quad \beta_2 \sum_{t=1}^{n} \int_\Omega p_t(\mathbf{x}) \frac{\partial \alpha}{\partial x_t}(\mathbf{x}) \frac{\partial \psi}{\partial x_t}(\mathbf{x}) d\mathbf{x} \\ &- \quad \beta_2 \sum_{t=1}^{n} \int_{\partial\Omega} \psi(\mathbf{x}) p_t(\mathbf{x}) \frac{\partial \alpha}{\partial \hat{n}_t}(\mathbf{x}) ds = \int_\Omega \psi(\mathbf{x}) f(\mathbf{x}), \end{aligned} \tag{28}$$

where $\partial \cdot / \hat{n}_t$ denotes the directional derivative along the unit normal vector in x_t direction $t = 1, \cdots, n$.

Inserting (26) into (28) leads to

$$\begin{aligned} \int_\Omega \psi(\mathbf{x}) \big[A^* A\alpha(\mathbf{x}) + \beta_1 q(\mathbf{x}) \alpha(\mathbf{x}) \big] d\mathbf{x} \quad &+ \quad \beta_2 \sum_{t=1}^{n} \int_\Omega p_t(\mathbf{x}) \frac{\partial \alpha}{\partial x_t}(\mathbf{x}) \frac{\partial \psi}{\partial x_t}(\mathbf{x}) d\mathbf{x} \\ &= \quad \int_\Omega \psi(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} + \beta_2 \sum_{t=1}^{n} \int_{\partial\Omega} \psi(\mathbf{x}) B_t(\mathbf{x}) ds. \end{aligned} \tag{29}$$

Equivalently, we can express (29) as follows:

$$\Pi(\psi, \alpha) + \beta_2 \sum_{t=1}^{n} \big(\frac{\partial \psi}{\partial x_t}, p_t \frac{\partial \alpha}{\partial x_t} \big) = (\psi, f) + \beta_2 \sum_{t=1}^{n} < \psi, B_t >, \tag{30}$$

where

$$\begin{aligned} \Pi(\psi, \alpha) \quad &:= \quad (\psi, A^* A\alpha) + \beta_1 (\psi, q\alpha) = (A\psi, A\alpha) + \beta_1 (\psi, q\alpha), \\ (k, l) \quad &:= \quad \int_\Omega k(\mathbf{x}) l(\mathbf{x}) d\mathbf{x}, \\ < k, l > \quad &:= \quad \int_{\partial\Omega} k(\mathbf{x}) l(\mathbf{x}) ds. \end{aligned}$$

Let $b(\psi,\alpha): H^1(\Omega)\times H^1(\Omega)\to\mathbb{R}$ be a bilinear form given by

$$b(\psi,\alpha) := \Pi(\psi,\alpha) + \beta_2 \sum_{t=1}^{n} (\frac{\partial\psi}{\partial x_t}, p_t \frac{\partial\alpha}{\partial x_t}). \tag{31}$$

Then, we consider the following weak formulation for the inverse problem:

$$b(\psi,\alpha) = G(\psi), \tag{32}$$

where $G(\psi)\in H^{-1}(\Omega)$ is a continuous linear functional given by

$$G(\psi) := (\psi, f) - \beta_2 \sum_{t=1}^{n} < \psi, B_t > . \tag{33}$$

Remark:

1. The solution of the integro-differential equation (22) together with the boundary conditions (26) satisfies the weak problem.

In the following, we will show the existence and uniqueness of the solution to the weak formulation.

Lemma 2: Let

$$B_1 := \big(\|A^*A\|_{L^2(\Omega)\to L^2(\Omega)} + \beta_1\|q\|_{L^\infty(\Omega)} + \beta_2 \max_t \|p_t\|_{L^\infty(\Omega)}\big), \tag{34}$$

$$B_2 := \min(q_L\beta_1, p_L\beta_2), \tag{35}$$

where

$$q_L := \min_{\mathbf{x}\in\Omega} \|q(\mathbf{x})\|_{L^\infty(\Omega)},$$
$$p_L := \min_t \min_{\mathbf{x}\in\Omega} \|p_t(\mathbf{x})\|_{L^\infty(\Omega)}.$$

The bilinear form $b(\psi,\alpha)$ in (31) is bounded and coercive. Hence, $b(\psi,\alpha)$ satisfies respectively

$$|b(\psi,\alpha)| \le B_1\|\psi\|_{H^1(\Omega)}\|\alpha\|_{H^1(\Omega)},$$
$$b(\alpha,\alpha) \ge B_2\|\alpha\|^2_{H^1(\Omega)}.$$

Proof. See Appendix A. □

Lemma 2 results in important conclusions:

Corollary 1: There exists a unique solution $\alpha\in H^1(\Omega)$ for the inverse problem (32). The existence and uniqueness is a direct result of the coercivity and boundedness of the bilinear form $b(\psi,\alpha)$ and Lax-Milgram theorem.[14]

Corollary 2: The unique solution $\alpha\in H^1(\Omega)$ of the inverse problem (32) is bounded by

$$\|\alpha\|_{H^1(\Omega)} \le \frac{1}{B_2}\|G\|_{H^{-1}(\Omega)}, \tag{36}$$

where

$$\|G\|_{H^{-1}(\Omega)} \le \|f\|_{L^2(\Omega)} + C\beta_2 \sum_{t=1}^{n} \|B_t\|_{H^1(\Omega)}, \tag{37}$$

where $C>0$ is a constant.

Proof. See Appendix B. □

4. DISCRETIZATION OF THE INVERSE PROBLEM

In practice, we seek a finite-dimensional approximation to the solution of the inverse problem (32). Therefore, we discretize (32) by projecting it onto a finite dimensional subspace. In this work, we use the variational formulation for the inverse problem and thus consider the Galerkin method for projection.[17]

Let $X_n \subset X$ denote a sequence of finite-dimensional subspaces of dimension n, spanned the by piecewise linear Lagrange basis functions $\{L_1, \ldots, L_n\}$, and $\{\mathbf{x}_p\}$, $p = 1, \ldots, n$, be a set of points on Ω. Replacing ψ in (32) by $\Psi \in X_n$

$$\Psi := \sum_{i=1}^{n} P_i L_i(\mathbf{x}),$$

and α in (32) by $\alpha_n \in X_n$

$$\alpha_n := \sum_{i=1}^{n} a_i L_i(\mathbf{x})$$

results in the discretized inverse problem:

$$b(\Psi, \alpha_n) = G(\Psi). \tag{38}$$

Replacing ψ in (32) by $\Psi \in X_n$ and subtracting (38) from the resulting equation leads to

$$b(\Psi, \alpha_n - \alpha) = 0. \tag{39}$$

This last equation implies an orthogonality condition with respect to the bilinear operator $b(\psi, \alpha)$, which defines an inner product and $b(\psi, \psi)$ is a norm. As a result, the discretization in (38) defines an orthogonal projection.[17]

Lemma 3: The solution to the discretized inverse problem (38) defined by the bilinear form $b(\psi, \alpha)$ exists and is unique.

Proof. The proof uses the coercivity property of the bilinear form $b(\psi, \alpha)$ shown in Lemma 2. For the rest of the proof see Theorem 13.27 in.[17] □

5. AN EXAMPLE

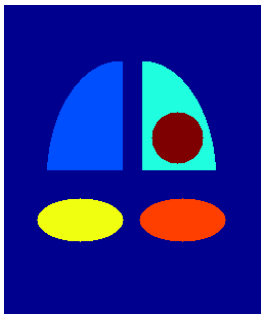
In this section, we describe the procedure to design the weighting functions q and p_t when there is only localization information (i.e. α_b^m for all $\mathbf{x} \in \Omega \subset \mathbb{R}^2$) and edge information (i.e. $e(\mathbf{x})$ for all $\mathbf{x} \in \Omega$), but no model for the image or the image gradient.

In our example study, we consider the optical medium on the bounded rectangular domain Ω with size 5 cm along x- and 6 cm along y-direction, as shown in figure 1(a), where the circular inclusion indicates the fluorophore concentration. In the following subsections, we will discuss the available *a priori* information and how we process this information to design the regularization terms, in particular, the functions q and p_t.

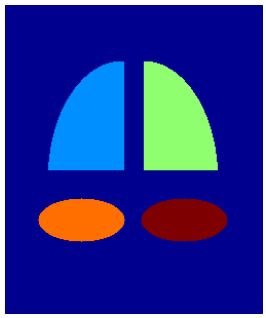
5.1. Available *a priori* information

Figure 1(b) shows the internal structure of the medium except for the inclusion which corresponds to the fluorophore concentration. This picture is analogous to an anatomical image that can be acquired from a secondary imaging modality such as MR or x-ray. In this case, the two inclusions at the top can be considered to mimic the lungs, and the inclusions below to the liver and kidneys. Assume that the fluorophore is known to exist in only the lungs. Then, figure 5.1 shows the function α_b^m, which denotes the regions where the fluorophore is expected to exist. Figure 2(b) shows the edges of the segmented image shown in figure 1(b). Note that, for the edge information, we preferred the image shown in figure 1(b), but not the function α_b^m shown in figure 5.1. Hence, we are taking into account the possibility that the actual solution can have edges outside the lungs as well.

Next, we will use the images shown in figure 2 to design the weighting functions q and p_t in the regularization terms.

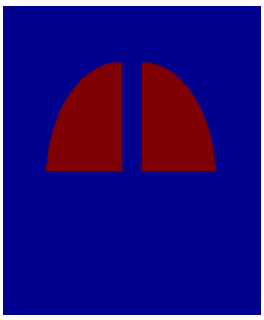


(a) The medium under inspection for the example study.

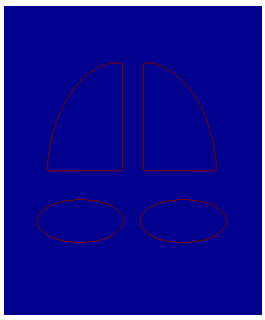


(b) The segmented image that mimics the internal organs of a mouse.

Figure 1. The medium considered in the example study (left) and the internal organ structure of the medium representing the anatomical image which can be acquired from a secondary imaging modality such as MR or x-ray.



(a) The function α_b^m.



(b) The edges of the image shown in figure 1(b).

Figure 2. The binary image that indicates the regions where the fluorophore is expected to exist and the edges of the segmented image shown in figure 1(b).

5.2. Design of the functions q and p_t

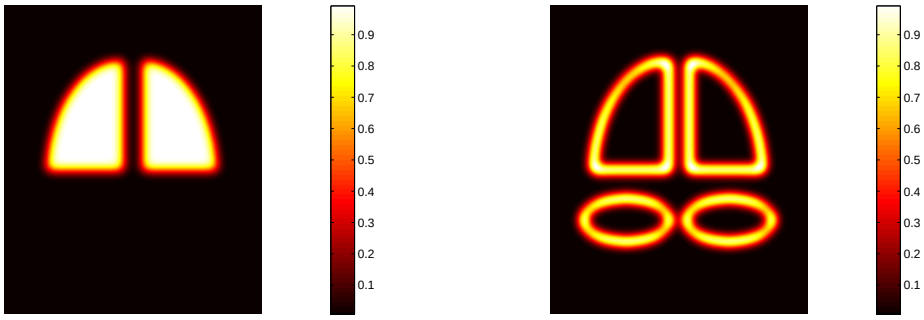
In this section, we will design the weighting functions q and p_t, which constitute the core of the regularization terms.

To start, we consider the function α_b^m shown in figure 5.1. To impose an uncertainty on the boundaries indicated by this function, we use a smoothing Gaussian kernel with equal variance $\sigma^2 = 0.1$. The resulting function $\tilde{\alpha}_b^m$ is shown in figure 5.2. We note the effect of filtering on the edges of the lungs as compared to the function α_b^m shown in figure 5.1. Similarly, we filter the edge function e by a Gaussian kernel to obtain a differentiable function and introduce an uncertainty on the exact locations of the edges.

Next, we will design the weighting functions q and p_t by using $\tilde{\alpha}_b^m$ and $\tilde{e}$, respectively.

In order to design the regularization term to penalize the solution, we define the weighting function q in (23). In this example, we set $Q = 1$ and $\epsilon = 10^{-2}$ in (23). For comparison, in figure 4, we show both the functions $\tilde{\alpha}_b^m$ and q, which is obtained by using $\tilde{\alpha}_b^m$ according to (23). In figure 4(b), it is seen that q is large where $\tilde{\alpha}_b^m$ is small and vice-versa. In the regularization framework, this corresponds penalizing the high values in the solution on regions where we do *not* expect the fluorophore to exist. Similarly, on the regions where the fluorophore is expected to exist, the penalty is less due to the low value of q on such regions.

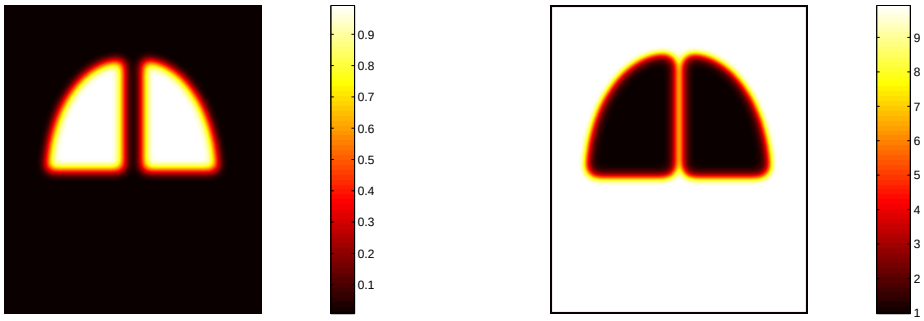
For the design of the function p_t, for $t = 1, 2$, we look into (25). First, we compute the partial derivative



(a) The function $\tilde{\alpha}_b^m$, obtained by filtering α_b^m by a Gaussian kernel.

(b) The function $\tilde{e}$, obtained by filtering the edge function e by a Gaussian kernel.

Figure 3. (a) The filtered version of α_b^m and **(b)** the edge function e after filtering.



(a) The function $\tilde{\alpha}_b^m$, obtained by filtering α_b^m by a Gaussian kernel.

(b) The function q defined by (23).

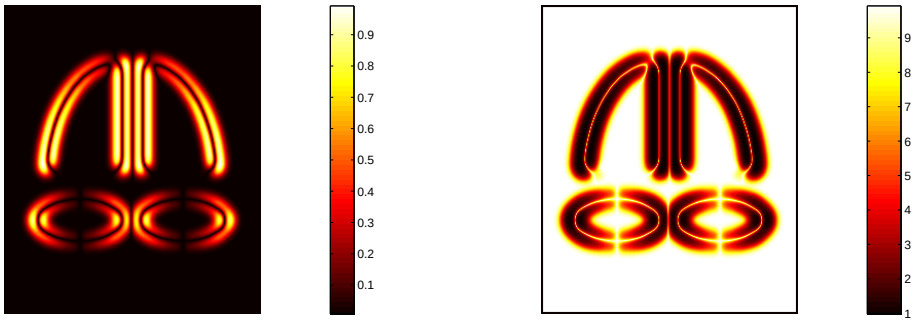
Figure 4. The functions $\tilde{\alpha}_b^m$ and q, which is obtained by using according to (23). **(a)** $\tilde{\alpha}_b^m$ and **(b)** the weighting function q.

of the differentiable function $\tilde{e}$ along x- and y-directions. We show the resulting images in figures 5.2 and 5.2, respectively. In figure 5.2, we see that the derivative of $\tilde{e}$ along x-direction is high across the edges along y-direction and small across the edges along x-direction. Similarly, in figure 5.2, we see that the derivative of $\tilde{e}$ along y-direction is high across the edges along x-direction and small across the edges along y-direction.

In order to design the functions p_1 and p_2, we note (25). In this example, we set $P_t = 1$, for $t = 1, 2$ and $\epsilon = 10^{-2}$. The resulting functions are shown in figures 5(b) and 6(b), respectively. Note that p_1 is large where the gradient of $\tilde{e}$ along x-direction is small and vice-versa. Similarly, p_2 is large where the gradient of $\tilde{e}$ along y-direction is small and vice-versa. For comparison, we display p_1 and the gradient of $\tilde{e}$ along x-direction in figure 5, and p_2 and the gradient of $\tilde{e}$ along y-direction in figure 6.

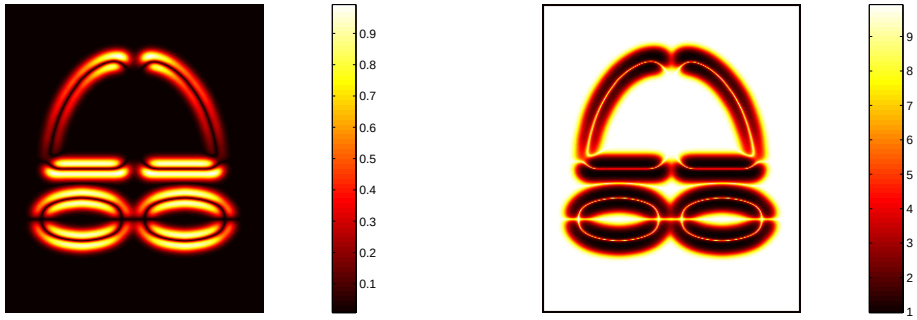
6. CONCLUSION

In this work, we discuss the incorporation of *a priori* information into the inverse problem formulation for fluorescence optical tomography. The *a priori* information consists of the information regarding the localization region of the fluorescent agent and the internal structure of the medium of interest. To make use of the available *a priori* information fully, we design regularization terms which impose spatially-varying penalties simultaneously on the solution and its gradients within the optimization framework. In this respect, we discuss the use of a



(a) The partial derivative of $\tilde{e}$ along x-direction. (b) p_t on Ω for $t = 1$ (i.e. x-direction).

Figure 5. (a) The partial derivative of $\tilde{e}$ along x-direction and **(b)** p_t computed according to (25) for $t = 1$.



(a) The partial derivative of $\tilde{e}$ along y-direction. (b) p_t on Ω for $t = 2$ (i.e. y-direction).

Figure 6. (a) The partial derivative of $\tilde{e}$ along y-direction and **(b)** p_t computed according to (25) for $t = 2$.

priori information for designing the appropriate regularization terms. We describe the design procedure in an experimental study.

Our future work will focus on the estimation of the parameters in the proposed inverse problem formulation and the validation of the proposed inverse problem formulation in experimental studies.

APPENDIX A. PROOF OF LEMMA 2

Clearly,

$$|b(\psi,\alpha)| \leq |\Pi(\psi,\alpha)| + \beta_2 \big| \sum_{t=1}^{n} (\frac{\partial \psi}{\partial x_t}, p_t \frac{\partial \alpha}{\partial x_t}) \big|,$$

and

$$\begin{aligned} |\Pi(\psi,\alpha)| &\leq \int_\Omega |(A^*A\alpha)(\mathbf{x})\beta(\mathbf{x})| d\mathbf{x} + \beta_1 \int_\Omega |q(\mathbf{x})\alpha(\mathbf{x})\beta(\mathbf{x})| d\mathbf{x} \\ &\leq \|A^*A\alpha\|_{L^2(\Omega)} \|\beta\|_{L^2(\Omega)} + \beta_1 \|q\|_{L^\infty(\Omega)} \|\alpha\|_{L^2(\Omega)} \|\beta\|_{L^2(\Omega)} \\ &\leq \|A^*A\|_{L^2(\Omega)\to L^2(\Omega)} \|\alpha\|_{L^2(\Omega)} \|\beta\|_{L^2(\Omega)} + \beta_1 \|q\|_{L^\infty(\Omega)} \|\alpha\|_{L^2(\Omega)} \|\beta\|_{L^2(\Omega)} \\ &\leq \big(\|A^*A\|_{L^2(\Omega)\to L^2(\Omega)} + \beta_1 \|q\|_{L^\infty(\Omega)} \big) \|\alpha\|_{H^1(\Omega)} \|\beta\|_{H^1(\Omega)}, \end{aligned} \tag{40}$$

and

$$\begin{aligned}\beta_2|\sum_{t=1}^{n}(\frac{\partial\psi}{\partial x_t},p_t\frac{\partial\alpha}{\partial x_t})| &\leq \beta_2\sum_{t=1}^{n}\int_\Omega|p_t(\mathbf{x})\frac{\partial\psi}{\partial x_t}(\mathbf{x})\frac{\partial\alpha}{\partial x_t}(\mathbf{x})|d\mathbf{x}\\ &\leq \beta_2\max_t\|p_t\|_{L^\infty(\Omega)}\int_\Omega|\nabla\alpha|\cdot|\nabla\beta|d\mathbf{x}\\ &\leq \beta_2\max_t\|p_t\|_{L^\infty(\Omega)}\|\alpha\|_{H^1(\Omega)}\|\beta\|_{H^1(\Omega)}.\end{aligned}$$

Hence,

$$\begin{aligned}|b(\psi,\alpha)| &\leq \big(\|A^*A\|_{L^2(\Omega)\to L^2(\Omega)}+\beta_1\|q\|_{L^\infty(\Omega)}+\beta_2\max_t\|p_t\|_{L^\infty(\Omega)}\big)\|\alpha\|_{H^1(\Omega)}\|\beta\|_{H^1(\Omega)}\\ &\leq B_1\|\alpha\|_{H^1(\Omega)}\|\beta\|_{H^1(\Omega)},\end{aligned}\tag{41}$$

where B_1 is given by (34). For the second inequality, we first write

$$b(\alpha,\alpha) = \Pi(\alpha,\alpha)+\beta_2\sum_{t=1}^{n}(\frac{\partial\alpha}{\partial x_t},p_t\frac{\partial\alpha}{\partial x_t}).$$

Since A is self-adjoint,

$$(A^*A\alpha,\alpha)=(A\alpha,A\alpha)\geq 0.$$

Thus, together with the non-negativity of the functions q and p_t, and the positive parameters β_1,β_2, we see

$$\begin{aligned}b(\alpha,\alpha) &\geq \beta_1(q\alpha,\alpha)+\beta_2\sum_{t=1}^{n}(\frac{\partial\alpha}{\partial x_t},p_t\frac{\partial\alpha}{\partial x_t})\\ &\geq q_L\beta_1(\alpha,\alpha)+p_L\beta_2\sum_{t=1}^{n}(\frac{\partial\alpha}{\partial x_t},\frac{\partial\alpha}{\partial x_t})\\ &\geq \min(q_L\beta_1,p_L\beta_2)\|\alpha\|^2_{H^1(\Omega)}\\ &\geq B_2\|\alpha\|^2_{H^1(\Omega)},\end{aligned}$$

where B_2 is given by (35).

APPENDIX B. PROOF OF COROLLARY 2

The inequality (36) is a direct result of the coercivity of the bilinear form $b(\psi,\alpha)$[19]:

$$\|G\|_{H^{-1}(\Omega)}=\sup_{0\neq\psi\in H^1(\Omega)}\frac{G(\psi)}{\|\psi\|_{H^1(\Omega)}}=\sup_{0\neq\psi\in H^1(\Omega)}\frac{b(\alpha,\psi)}{\|\psi\|_{H^1(\Omega)}}\geq\frac{b(\alpha,\alpha)}{\|\psi\|_{H^1(\Omega)}}\geq B_2\|\alpha\|_{H^1(\Omega)}.$$

The norm $\|G\|_{H^{-1}(\Omega)}$ is given by

$$\begin{aligned}\|G\|_{H^{-1}(\Omega)} &= \sup_{0\neq\psi\in H^1(\Omega)}\frac{(\psi,f)+\beta_2\sum_{t=1}^{n}<\psi,B_t>}{\|\psi\|_{H^1(\Omega)}}\\ &\leq \frac{\big(\|f\|_{L^2(\Omega)}+C\beta_2\sum_{t=1}^{n}\|B_t\|_{H^1(\Omega)}\big)\|\psi\|_{H^1(\Omega)}}{\|\psi\|_{H^1(\Omega)}},\end{aligned}$$

where the second term in the inequality results from the trace theorem[14]:

$$\|B_t\|_{L^2(\partial\Omega)}\leq C\|B_t\|^{1/2}_{L^2(\Omega)}\|B_t\|^{1/2}_{H^1(\Omega)}$$

for $C>0$ and $B_t\in H^1(\Omega)$, where the boundary $\partial\Omega$ of Ω is Lipschitz.

REFERENCES

1. A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized born ratio," *IEEE Transactions on Medical Imaging* **24**(10), pp. 1377–1386, 2005.
2. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation," *Optics Letters* **26**, pp. 893–895.
3. V. Ntziachristos, G.Turner, J. Dunham, S. Windsor, A. Soubret, J. Ripoll, and H. A. Shih, "Planar fluorescence imaging using normalized data," *Journal of Biomedical Optics* **10**(6), pp. 1–8, 2005.
4. A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," *Applied Optics* **42**, pp. 3081–3094, 2003.
5. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Adaptive finite element based tomography for fluorescence optical imaging in tissue," *Optics Express* **12**(22), pp. 5402–5417, 2004.
6. M. Guven, B. Yazici, X. Intes, and B. Chance, "Diffuse optical tomography with a priori anatomical information," *Physics in Medicine and Biology* **50**, pp. 2837–2858, 2005.
7. G. Boverman, E. L. Miller, A. Li, Q. Zhang, T. Chaves, D. H. Brooks, and D. A. Boas, "Quantitative spectroscopic diffuse optical tomography of the breast guided by imperfect a priori structural information," *Physics in Medicine and Biology* **50**, pp. 3941–3956, 2005.
8. M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes, "Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: Part I," *accepted by Inverse Problems* .
9. D. J. Cuccia, F. Bevilacqua, A. J. Durkin, S. Merritt, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, "In vivo quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration," *Applied Optics* **42**, pp. 2940–2950, 2003.
10. S. Merritt, F. Bevilacqua, A. J. Durkin, D. J. Cuccia, R. Lanning, B. J. Tromberg, G. Gulsen, G. Yu, J. Wang, and O. Nalcioglu, "Coregistration of diffuse optical spectroscopy and magnetic resonance imaging in a rat tumor model with magnetic resonance imaging coregistration," *Applied Optics* **42**, pp. 2951–2958, 2003.
11. S. Merritt, G. Gulsen, G. Chiou, Y. Chu, C. Deng, A. Cerussi, A. Durkin, B. J. Tromberg, and O. Nalcioglu, "Comparison of water and lipid content measurements using diffuse optical spectroscopy and MRI in emulsion phantoms," *Technol. Cancer Res. Treat.* **2**, pp. 563–569, 2003.
12. W. Cong, G. Wang, D. Kumar, Y. Liu, M. Jiang, L. Wang, E. Hoffman, G. McLennan, P. McCray, J. Zabner, and A. Cong, "Practical reconstruction method for bioluminescence tomography," *Optics Express* **13**(18), pp. 6756–6771, 2003.
13. A. Serdaroglu, B. Yazici, and V. Ntziachristos, "Fluorescence molecular tomography based on a priori information," in *Proceedings of Optical Society of America Biomedical Optics Meeting*, 2006.
14. S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer Verlag, 2002.
15. S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems* **15**, pp. R41–93, 1999.
16. J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, vol. 160 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 2005.
17. R. Kress, *Linear integral equations*, vol. 82 of *Applied Mathematical Sciences*, Springer-Verlag, second ed., 1999.
18. G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer Verlag, 2002.
19. F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*, Springer Verlag, 1998.