PDE-Constrained Fluorescence Tomography With the Frequency-Domain Equation of Radiative Transfer

Hyun Keol Kim, Jong Hwan Lee, and Andreas H. Hielscher, *Member, IEEE*

***Abstract*—We present the first fluorescence tomography algorithm that is based on a partial differential equation (PDE) constrained approach. PDE methods have been increasingly employed in many numerical applications, as they often lead to faster and more robust solutions of many inverse problems. In particular, we use a sequential quadratic programming (SQP) method, which allows solving the two forward problems in fluorescence tomography (one for the excitation and one for the emission radiances) and one inverse problem (for recovering the spatial distribution of the fluorescent sources) simultaneously by updating both forward and inverse variables in simultaneously at each of iteration of the optimization process. We evaluate the performance of this approach with numerical and experimental data using a transport-theory frequency-domain algorithm as forward model for light propagation in tissue. The results show that the PDE-constrained approach is computationally stable and accelerates the image reconstruction process up to a factor of 15 when compared to commonly employed unconstrained methods.**

***Index Terms*—Fluorescence tomography, frequency-domain equation of radiative transfer (ERT), partial differential equation (PDE) constrained optimization, sequential quadratic programming (SQP).**

I. Introduction

OPTICAL fluorescence tomography has emerged as a new imaging modality that makes use of light-emitting biomarkers, for diagnostic imaging of pathological changes in tissues [1]–[7]. By recovering fluorescence measurement data, this technique provides a spatial distribution of fluorescent biomarkers in tissue, which can be used in preclinical and clinical studies to monitor the migration of cells, the protein–protein interaction and the progression of a disease including cancerous tumors [5]. To this end, fluorescent biomarkers have been used to target specific molecules that provide functional information about biochemical processes preceding the development of diseases [6], [7]. Thus, fluorescence tomography allows for the early detection of diseases on a molecular level before typical symptoms or macroscopic changes appear.

Many imaging codes currently used employ a so-called model-based iterative image reconstruction (MOBIIR) scheme in which a forward model of light propagation in tissue is needed [8]–[10]. The light distribution inside the medium can be described by a system of two coupled forward equations: one for the excitation radiance caused by an external light source and one for the emission radiance by an internal fluorescent source. For high-scattering, low-absorbing media, both excitation and emission radiances can be accurately modeled by the diffusion approximation (DA) to the equation of radiative transfer (ERT) [11]–[19]. It is, however, well known that the DA becomes less accurate when absorption is increased and the medium is no longer scattering dominated. Problems also arise in low-scattering regions, such as void-like, fluid-filled areas and when small tissue volumes are considered, where boundary effects dominate [20]–[24]. In small animal imaging, many of these issues are encountered and the DA is of limited applicability. The problems related to the DA can be overcome by using the ERT that provides accurate prediction of light propagation in all types of media [25], [26]. However, in general, the ERT requires much longer computation times as compared to the DA and are much more complex to implement. ERT-based image reconstruction codes can take anywhere from several hours to even days to solve one single image reconstruction case. Thus, it is highly desirable to develop computationally efficient image reconstruction schemes that allow for using the ERT for optical fluorescence tomography. Motivated by this background, we have adapted a so-called PDE-constrained approach that makes use of a reduced Hessian SQP (rSQP) method. In this approach the forward and inverse problems are solved simultaneously, which has shown in other applications [27]–[40] to lead to significant savings in the total image reconstruction time.

Recently, our group has introduced the PDE-constrained approach to the field of optical tomography. However, that work was limited to recover the absorption and scattering coefficients inside biological tissues and had not addressed the problems encountered in fluorescence tomography. First, using the frequency-domain version of the radiative transport equation as the forward model of light propagation in tissues, Abddoulaev *et al*. [41] suggested an augmented Lagrangian method (ALM) that solves the forward and inverse problems simultaneously. In numerical phantom studies, they demonstrated that PDE-constrained ALM codes can significantly reduce computation times. Subsequently, Kim and Hielscher [42] replaced the ALM

Manuscript received September 25, 2009. This work was supported in part by the National Cancer Institute under Grant NCI-4R33CA118666 and Grant NCI-U54CA126513-039001 at the National Institutes of Health. The work of A. H. Hielscher was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, the National Heart, Lung, and Blood Institute, the National Institute for Biomedical Imaging and Bioengineering, the National Cancer Institute, the Whitaker Foundation for Biomedical Engineering, and the New York State Office of Science, Technology, and Academic Research.

H. K. Kim and J. H. Lee are with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA (e-mail: hkk2107@columbia.edu; jl3132@columbia.edu).

A. H. Hielscher is with the Department of Biomedical Engineering and Radiology, Columbia University, New York, NY 10027 USA (e-mail: ahh2004@columbia.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/JSTQE.2009.2038112

by the rSQP method and showed that even faster convergence and higher numerical stability can be obtained. Using numerical and experimental data, they observed an increase in the reconstruction speed by up to a factor of 30 as compared to standard codes that do not use PDE constraints. In the study at hand, we have adapted the approach previously presented by Kim and Hielscher to the problem of fluorescence tomography.

So far only one publication exists that applies the PDE-constrained methods to fluorescence tomography. Recently, Joshi *et al*. [43], [44] have reported on a Gauss–Newton-based constrained method. However, their work used the frequency-domain DA as a forward model of light propagation and not the ERT. Moreover, the authors reported that the Gauss–Newton approach is computationally more expensive than the reduced Hessian approach with a BFGS updating formula.

In this paper, we go beyond past studies and present the first PDE-constrained fluorescence image reconstruction algorithm. The adaptation of the nonfluorescence code for problems in fluorescence tomography is nontrivial. Inverse source problems are known to be notoriously ill-posed and unlike in nonfluorescence problems, two coupled ERTs for the excitation and emission radiances have to be treated simultaneously in the framework of PDE-constrained optimization. In the following, we will first provide details on the structure and implementation of the fluorescence forward model (see Section II-A). This is followed by a description of the implementation of the PDE-constrained rSQP approach for fluorescence tomography (see Section II-B–II-D). Subsequently, we evaluate the performance of the PDE-constrained scheme using numerical and experimental results (see Section III). We focus on computational efficiency by comparing the new algorithm with an unconstrained code that makes use of the limited-memory Broyden–Fletcher–Goldfarb–Shanno (lm-BFGS) method [45]. The lm-BFGS method is known to be the most efficient gradient-based unconstrained optimization method [46], [47], and therefore, provides a good standard for comparison. Finally we draw some conclusions in Section IV.

II. Methods

A. Light Propagation Model

In frequency-domain optical fluorescence tomography, the external light source is amplitude-modulated in the 10 Hz–1 GHz frequency range and the demodulated transmitted intensities and phase shifts are measured on the tissue surface, using wavelength-dependent filters to distinguish between excitation and emission signals.

The generation and propagation of fluorescence light in biological tissue can be accurately modeled by two coupled frequency-domain equations of radiative transfer [47], [48] as

$$\left[(\nabla\cdot\Omega)+\mu_a^x+\mu_s^x+\mu_a^{x\to m}+\frac{i\varpi}{c}\right]\psi^x(\mathbf{r},\Omega,\varpi)=\int_{4\pi}p(\Omega',\Omega)\psi^x(\mathbf{r},\Omega',\varpi)d\Omega' \tag{1}$$

$$\left[(\nabla\cdot\Omega)+\mu_a^m+\mu_s^m+\frac{i\varpi}{c}\right]\psi^m(\mathbf{r},\Omega,\varpi)=\int_{4\pi}p(\Omega',\Omega)\psi^m(\mathbf{r},\Omega',\varpi)d\Omega'+\frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}. \tag{2}$$

The first equation describes the propagation of excitation light emanating from the external light source and the second represents the propagation of emission light due to the fluorescent source in tissue. Here, the superscripts *x* and *m* denote excitation and emission, respectively. Hence, μ_a^x and μ_s^x are the absorption and scattering coefficients in units of cm^{-1} at the excitation wavelength; $\mu_a^{x\to m}$ is the absorption coefficient of a fluorescent source in tissue, in unit of cm^{-1}; μ_a^x and μ_s^x are the absorption and scattering coefficients in units of cm^{-1} at the emission wavelength. $\psi^x(\mathbf{r},\Omega,\varpi)$ and $\psi^m(\mathbf{r},\Omega,\varpi)$ denote the excitation and emission radiances, respectively, in units of $\text{W}\cdot\text{cm}^{-2}\cdot\text{st}^{-1}$; η denotes the quantum yield by which the fluorescent source emits light in transit from excitation state to ground state; $\tau(\mathbf{r})$is the local lifetime of a fluorescent source. Note that $\phi(\mathbf{r},\varpi)$ appearing in (2) denotes the excitation fluence at position $\mathbf{r}$ defined by $\phi(\mathbf{r},\varpi)=\int_{4\pi}\psi^x d\Omega$ in unit W/cm^2. For a phase function denoted by $p(\Omega',\Omega)$, we use here the Henyey–Greestein phase function [49] that is commonly used in tissue optics. As shown in (1), the excitation light $\psi^x(\mathbf{r},\Omega,\varpi)$ is absorbed and scattered by the intrinsic medium, and further attenuated by the fluorophore absorption $\mu_a^{x\to m}$, and then excites a fluorochrome inside the tissue at position $\mathbf{r}$. The excited fluorophore constitutes a light source inside the medium that reemits radiation $\psi^m(\mathbf{r},\Omega,\varpi)$ typically at longer wavelength [see (2)]. The local strength of a fluorescent source is directly proportional to the local fluorophore absorption $\mu_a^{x\to m}\phi(\mathbf{r},\varpi)$ and the quantum yield η and the local fluorophore lifetime $\tau(\mathbf{r})$.

The corresponding boundary conditions for the two coupled equations are given

$$\psi_b^x|_{\vec{n}_b\cdot\Omega<0}=\psi^{x0}\big|_{\vec{n}_b\cdot\Omega<0}+R(\Omega',\Omega)\cdot\psi^x|_{\vec{n}_b\cdot\Omega'>0} \tag{3}$$

$$\psi_b^m|_{\vec{n}_b\cdot\Omega<0}=R(\Omega',\Omega)\cdot\psi^m|_{\vec{n}_b\cdot\Omega'>0} \tag{4}$$

where $R(\Omega',\Omega)$ is the reflectivity at Fresnel interface [50] from direction Ω' to direction Ω, ψ_b^{x0} is the radiation intensity due to the external source function and subscript b denotes the boundary surface of the medium, while $\vec{n}_b$ is the unit normal vector pointing outward the boundary surface. For discretization of the two equations (1) and (2), we employ a node-centered finite-volume approach in the spatial domain and a discrete ordinates method in the angular domain. The node-centered finite-volume method takes advantage of the beneficial properties of both the finite-element and finite-volume methods by combining the conservation properties of the finite-volume formulation and the geometric flexibility of the finite-element approach.

Following an unstructured finite-volume discrete-ordinates method [42], [51], the discretized forms of the two ERTs given by (1) and (2) are obtained by integrating (1) and (2) over the

control volume with a divergence theorem as

$$\sum_{j=1}^{N_{\text{surf}}} (\vec{n}_j \cdot \Omega^l)\psi_j^{x,l} dA_j + \left(\mu_a^x + \mu_s^x + \mu_a^{x\to m} + \frac{i\omega}{c}\right)\psi_N^{x,l}\Delta V_N = \mu_s^x \sum_{l'=1}^{N_\Omega} \psi_N^{x,l'} p^{l'l} w^{l'} \quad (5)$$

$$\sum_{j=1}^{N_{\text{surf}}} (\vec{n}_j \cdot \Omega^l)\psi_j^{m,l} dA_j + \left(\mu_a^m + \mu_s^m + \frac{i\omega}{c}\right)\psi_N^{m,l}\Delta V_N = \mu_s^m \Delta V_N \sum_{l'=1}^{N_\Omega} \psi_N^{m,l'} p^{l'l} w^{l'} + \frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}\Delta V_N \quad (6)$$

where N_{surf} and N_Ω are the number of surfaces surrounding the node $N(=1,\ldots,N_t)$ and the number of discrete ordinates based on the level symmetric scheme, respectively, $\vec{n}_j$ denotes the surface normal vector, and $\psi_j^{x,l}$and $\psi_j^{m,l}$ denote the excitation and emission radiances defined on the jth surface in direction l. Also the surface intensities $\psi_j^{x,l}$ and $\psi_j^{m,l}$ are related to the nodal intensities $\psi_N^{x,l}$ and $\psi_j^{m,l}$ by the second-order spatial differencing scheme [51]. It can be easily seen that each ERT involves N_t spatial unknown intensities coupled into N_Ω directions, thus leading to the total $N_t \cdot N_\Omega$ unknowns. The system of (1) and (2) can be solved by using any iterative solvers as far as they are reliable. We employ here a matrix-based iterative linear solver that enables updating all the radiation intensities $\psi_N^{x,l}$ (or $\psi_N^{m,l}$) simultaneously, which leads to fast convergence as compared to source iteration-based techniques. After discretization for all nodes, we finally obtain two linear systems of algebraic equations as

$$A^x\psi^x = b^x \quad (7a)$$

$$A^m\psi^m = b^m. \quad (7b)$$

Each line denoted by $i\,(i=1,\ldots,N_tN_\Omega)$ of the matrix A contains the coefficients of the discretized forms [see (7a) and (7b)] established at node number N and direction l.

The excitation light source comes into the term b^x after discretization on a boundary node N_b, while the fluorescent source comes into the term b^m after discretization on internal node N, as

$$b_{N_b}^{x,l} = -\sum_j \left[1 - \max(n_j\cdot\Omega^l / \left|n_j\cdot\Omega^l\right|, 0)\right](n_j\cdot\Omega^l)A_j\psi_{N_b}^{0,l} \quad (8a)$$

$$b_N^{m,l} = \frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}\Delta V_N \quad (8b)$$

where $\psi_{N_b}^{x0,l}$ is the external source function on boundary node N_b in direction l.

Since we treat the frequency-domain equations of radiative transfer directly, the matrix formulations given by (7) contain complex-valued elements. As a result, the complex-valued algebraic linear equations are solved with a complex version of the generalized minimum residual (GMRES) solver [52], [53].

The two equations given by (1) and (2) provide the predictions of the excitation radiance $\psi_j^{x,l}$ and transmitted emission radiance $\psi_j^{m,l}$ on the boundary surface. Also these two equations are used as the constraints within the PDE-constrained SQP method. Next we discuss the inverse model, which is used to obtain the spatial distribution of a fluorescent source inside the medium that best fits the measured emission data.

B. PDE-Constrained Optimization for Fluorescence Tomography

To understand the PDE-constrained approach to fluorescence tomography, we first summarize the most important aspects, previously described in [42]. The optical fluorescence tomographic problem can be formulated in more general terms as

$$\begin{aligned} \min \quad & f(x,u) \\ \text{s.t.} \quad & C(x,u) = 0 \end{aligned} \quad (9)$$

where $x \in R^n$ is the vector of inverse variables, $u \in Z^m$ is the vector of forward variables, $f(x;u)$ is an objective function that quantifies the difference between measured and predicted intensities and $c(x;u)=0$ is a discretized version of combined excitation and emission equations. The problem given by (9) is often referred to as "PDE-constrained" optimization since the optimal solution at minimum of f is constrained by equality condition $c(x;u)=0$ represented by two partial differential equations (PDEs).

Traditional methods for solving (9) is to treat the forward variable u as a dependent variable of the inverse variable x, i.e., $u=\tilde{c}(x)$, which makes it possible to replace the prediction vector u in f of (9) by its forward solution vector. As a result, the problem (9) can be reformulated as

$$\min \tilde{f}(x) = f(x,\tilde{c}(x)) \quad (10)$$

which is often referred to as "unconstrained" because the equality constraint $c(x;u)=0$ is eliminated in (10), i.e., $\tilde{f}$ is now a function of x only. Thus, the forward solution $u=\tilde{c}(x)$ has to be obtained for evaluation of $\tilde{f}$, i.e., the complete solutions of the two ERTs are required for the excitation and emission radiances at each of optimization iterations. As a consequence, the associated optimization procedure becomes a computationally very demanding process, both with respect to time and memory. Nonetheless, this approach has been widely used for the solution of optical fluorescence tomographic problems mainly because of easiness of implementation. The existing fluorescence tomographic codes [11]–[19] belong to this approach: the conjugate gradient (CG) approach [12], the quasi-Newton (QN) approach [14], [25], [26], the Jacobian approach [11], [13]–[19], etc. [15].

Another approach to solve (9) is to treat the forward variable u and the inverse variable x independently, which enables solving the PDE-constrained problem (9) directly by updating the forward and inverse variables simultaneously at each of optimization iterations. Typically, an extended objective function called "Lagrangian" is introduced as follows:

$$L(x,u;\lambda) = f + \lambda^T c. \quad (11)$$

Here, $\lambda \in Z^m$ is called the vector of Lagrange multipliers. The simultaneous solutions of forward and inverse problems can be achieved at points satisfying first-order necessary conditions [41], [42] where the gradient of L in (11) vanishes with respect λ, u, and x, respectively.

One major advantage of this PDE-constrained approach is that the complete solution of the forward problem is not required until convergence is reached. In other words, PDE-constrained methods allow for using the inexact solution of the forward problem into solving the inverse problem, which leads to a significant reduction in the total reconstruction time. The solution accuracy of the forward problem is iteratively controlled as the inverse solution goes toward the optimum. Optimization techniques of this kind have seen rapid developments mainly in applications with airfoil design, flow variable optimization, and electromagnetic inverse problems [27]–[40].

As discussed in Section II, the fluorescence tomographic problem is to find a vector of unknowns $x = (\mu_a^{x\rightarrow m}, \tau, \eta)$, assuming that all other intrinsic properties inside the medium, μ_a^x, μ_s^x and μ_a^m, μ_s^m, are already known at the excitation and emission wavelengths, such that

$$\min f(\psi^m) = \frac{1}{2}\sum_{s,d} |Q_d\psi_s^m - z_{s,d}|^2$$

subject to

$$\begin{aligned} C_s^x &= A^x\psi_s^x - b_s^x = 0, \qquad s = 1,\ldots,N_s \\ C_s^m &= A^m\psi_s^m - b_s^m = 0, \qquad s = 1,\ldots,N_s \end{aligned} \tag{12}$$

where $f(\psi_s^m)$ is the objective function that quantifies the difference between predictions and measurements of emitted light made on the tissue surface; C^x and C^m are the discretized versions of the two coupled radiative transfer equations. In this study we focus our attention to the inverse problem of reconstructing the spatial distribution of the fluorophore absorption $\mu_a^{x\rightarrow m}$ assuming a spatially homogenous τ and η distribution inside the medium, which means that no quenching occurs. Therefore, we hereafter denote $\mu_a^{x\rightarrow m}$ by μ omitting its superscript and subscript.

By introducing a Lagrangian function, the previous PDE-constrained optimization problem can be restated as

$$\begin{aligned} L(\mu;\psi^x,\psi^m;\lambda^x,\lambda^m) = \frac{1}{2}|Q\psi^m - z|^2 + \lambda^{xT}(A^x\psi^x - b^x) \\ + \lambda^{mT}(A^m\psi^m - b^m). \end{aligned} \tag{13}$$

The simultaneous solutions of the forward and inverse problems can then be obtained at points in which the following five PDEs, i.e., the first derivatives of the Lagrangian function with respect to each of variables, become zero

$$L_\mu \equiv \frac{\partial L}{\partial \mu} = (\psi^{xT}A^{xT})_\mu\lambda^x - (b^{mT})_\mu\lambda^m = 0 \tag{14a}$$

$$L_{\psi^x} \equiv \frac{\partial L}{\partial \psi^x} = A^{xT}\lambda^x - (b^{mT})_{\psi^x}\lambda^m = 0 \tag{14b}$$

$$L_{\psi^m} \equiv \frac{\partial L}{\partial \psi^m} = Q^T(Q\psi^m - z) + A^{mT}\lambda^m = 0 \tag{14c}$$

$$L_{\lambda^x} \equiv \frac{\partial L}{\partial \lambda^x} = A^x\psi^x - b^x = 0 \tag{14d}$$

$$L_{\lambda^m} \equiv \frac{\partial L}{\partial \lambda^m} = A^m\psi^m - b^m = 0 \tag{14e}$$

where the first equation represents the sensitivity equation with respect to the inverse variable μ, and the second and third equations can be viewed as the adjoint equations for λ^x and λ^m, and the last two equations represent the two forward equations given by (7).

The so-called Karush–Khun–Tucker (KKT) system given by (14) can be solved with Newton's method as

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta p \\ \Delta\lambda \end{bmatrix} = -\begin{bmatrix} L_p \\ L_\lambda \end{bmatrix} \tag{15}$$

where the block matrix W denotes the Hessian matrix of the Lagrangian function L with respect to each of unknowns $p = (\mu, \psi^x, \psi^m)$ and the block matrix A denotes the Jacobian matrix of constraints C^x and C^m with respect to each of unknowns $p = (\mu, \psi^x, \psi^m)$. Also Δp and $\Delta\lambda$ denote $[\Delta\mu, \Delta\psi^x, \Delta\psi^m]^T$ and $[\Delta\lambda^x, \Delta\lambda^m]^T$, respectively. The algebraic system given by (15) can then be solved efficiently through the reduced Hessian SQP scheme as will be shown shortly.

C. Reduced Hessian SQP

The reduced Hessian SQP method is an established method that solves nonlinear optimization problems with relatively low cost and fast convergence [27]. Employing this rSQP method for solving the previous KKT system (15) is equivalent to finding the minimum to a quadratic approximation of the Lagrangian function L subject to the linearized constraints C^x and C^m, which gives the following quadratic programming problem so that

$$\min \Delta p^T g_p + \frac{1}{2}\Delta p^T W \Delta p$$

subject to

$$A\Delta p + C = 0 \tag{16}$$

where $g = \nabla f = [g_\mu, g_{\psi^x}, g_{\psi^m}]$ denotes the gradients of the objective function f with respect to each of unknowns (μ, ψ^x, ψ^m) and W^k is the full Hessian (or approximations) of the Lagrangian function. Since the quadratic problem (16) is equivalent to the system (15), the linearized constraints in (16) correspond to the last block of the KKT system (15) as

$$L_{\mu\lambda^x}\Delta\mu + L_{\psi^x\lambda^x}\Delta\psi^x + L_{\psi^m\lambda^x}\Delta\psi^m = -L_{\lambda^x} \tag{17a}$$

$$L_{\mu\lambda^m}\Delta\mu + L_{\psi^x\lambda^m}\Delta\psi^x + L_{\psi^m\lambda^m}\Delta\psi^m = -L_{\lambda^m} \tag{17b}$$

where

$$L_{\mu\lambda^x} = (A^x\psi^x)_\mu, L_{\psi^x\lambda^x} = A^x, L_{\psi^m\lambda^x} = 0 \tag{18a}$$

$$L_{\mu\lambda^m} = (-b^m)_\mu, L_{\psi^x\lambda^m} = (-b^m)_{\psi^x}, L_{\psi^m\lambda^m} = A^m. \tag{18b}$$

Again (17) can be explicitly written into the following form:

$$A^x\Delta\psi^x = -\{C^x + (A^x\psi^x)_\mu\Delta\mu\} \tag{19a}$$

$$A^m\Delta\psi^m = -\{C^m - (b^m)_\mu\Delta\mu - (b^m)_{\psi^x}\Delta\psi^x\}. \tag{19b}$$

The full Hessian of the Lagrangian function is often difficult to obtain and its approximation by the updating schemes tends to create large dense matrix $(n+m)\times(n+m)$. These difficulties can be overcome by dropping certain noncritical second-order terms of the full Hessian matrix. In what follows, we describe the standard reduced Hessian SQP method based on separation of variables.

Since A^x and A^m are invertible, the vector Δp can be decomposed into two parts as follows:

$$\Delta p = Z + Y\Delta\mu \tag{20}$$

where

$$Z = \begin{bmatrix} 0 \\ -(A^x)^{-1}C^x \\ -(A^m)^{-1}\{C^m + (b^m)_{\psi^x}(A^x)^{-1}C^x\} \end{bmatrix} \tag{21}$$

and

$$Y = \begin{bmatrix} I \\ -(A^x)^{-1}(A^x\psi^x)_\mu \\ (A^m)^{-1}\left\{(b^m)_\mu - (b^m)_{\psi^x}(A^x)^{-1}(A^x\psi^x)_\mu\right\} \end{bmatrix}. \tag{22}$$

The choices of Z^k and Y^k is one challenging problem arising in practical implementation of the reduced SQP scheme. In this study, we have followed the popular choices for Z^k and Y^k as used by many authors [27], [31], [36], [37], [40].

By substituting (20) into (16) and differentiating the resulting expression with respect to $\Delta\mu$, and using the identity $Y^TA^T = 0$, we can obtain the reduced Hessian of the form:

$$H_r\Delta\mu = -(g_r + d) \tag{23}$$

where $H_r = Y^TWY$ denotes the reduced Hessian, $g_r = Y^Tg$ denotes the reduced gradient, and $d = Y^TWZ$ is called the cross-term. Thus, the reduced SQP method requires much less memory than the full SQP one, i.e., only a small $(n\times n)$ matrix needs to be maintained and updated at each of optimization iterations. From (23), the inverse solution $\Delta\mu$ and the forward solutions $\Delta\psi^x$ and $\Delta\psi^x$ can be obtained, respectively, as

$$\Delta\mu = -(H_r)^{-1}(g_r + d_r) \tag{24}$$

$$\Delta\psi = \begin{bmatrix} \Delta\psi^x \\ \Delta\psi^m \end{bmatrix} = Z - Y(H_r)^{-1}(g_r + d_r). \tag{25}$$

At the new iterate, the Lagrangian multiplier vectors are updated from

$$\lambda^m = \left[A^{mT}\right]^{-1}Q^T(Q\psi^m - z) \tag{26a}$$

$$\lambda^x = \left[A^{xT}\right]^{-1}(b^{mT})_{\psi^x}\lambda^m. \tag{26b}$$

Using (22), we can rewrite the reduced gradient as

$$\begin{aligned} g_r &= Y^Tg \\ &= (b^{mT})_\mu\lambda^m - (\psi^{xT}A^{xT})_\mu\lambda^x. \end{aligned} \tag{27}$$

For large-scale applications, it is desirable to avoid the direct computation of the reduced Hessian H_r^k and its matrix inversion $(H_r^k)^{-1}$. Accordingly, we approximate the matrix-vector product of $(H_r^k)^{-1}g_r^k$ directly by using the limited-memory BFGS updating formula [42], [45], which is another important feature of our study that enables the proposed rSQP scheme to be applied to large-scale optimization problems. Also note that we ignore the cross-term vector $d_r^k = Y^{kT}W^kZ^k$ as in other works [27], [36], [37].

The global convergence of the rSQP scheme is ensured by line search on the following real-valued l_1 merit function defined as:

$$\varphi_\eta(\mu,\psi^x,\psi^m) = f(\psi^m) + \eta\left\|C\right\|_1 \tag{28}$$

which is chosen here for its simplicity and low computational cost. The directional derivative of $\varphi_\eta(\mu,\psi^x,\psi^m)$ along the descent direction Δp is given by

$$D\varphi_\eta = g^{kT}\Delta p - \eta\left\|C\right\|_1. \tag{29}$$

Thus, the descent property $D\varphi_\eta(\mu,\psi^x,\psi^m) < 0$ can be maintained by choosing [50]

$$\eta > \frac{g^T\Delta p}{\left\|C\right\|_1}. \tag{30}$$

At the new iterate given by $p^{k+1} = p^k + \alpha^k\Delta p$, the merit function (28) is successively monitored to ensure the global progress toward the solution while line search is performed to find a step length that can provide the sufficient decrease in the merit function as [40]

$$\varphi_{\eta_k}(p^k + \alpha^k\Delta p^k) \le \varphi_{\eta_k}(p^k) + \beta\alpha_k D_{\eta_k}(p^k;\Delta p^k) \tag{31}$$

where β is a small value that determines the line search accuracy, typically set to 0.01 (or smaller).

D. Computational Algorithm of the rSQP Scheme

The algorithm that makes use of the reduced Hessian SQP method, as described in Section II-B and II-C has the following structure:

1) Set $k=0$ and initialize $\mu^0, \psi^{x,0}, \psi^{m,0}$, and $H_r^0 = I$.
2) Solve $A^{mT}\lambda^{m,0} = -Q^T(Q\psi^{m,0} - z)$ for λ^{m0}, and solve $A^{xT}\lambda^{x,0} = (-b^{mT})_{\psi^{x0}}\lambda^{m,0}$ for λ^{x0}.
3) Compute $g_r^0 = (b^{mT})_\mu\lambda^{m,0} - (\psi^{xT}A^{xT})_\mu\lambda^{x,0}$.
4) Check the stopping criteria: if satisfied, stop.
5) Solve $H_r^k\Delta\mu^k = -g_r^k$ for $\Delta\mu$ via a damped BFGS updating formula.
6) Solve the two QP problems for $\Delta\psi^{xk}$ and $\Delta\psi^{mk}$.

$$A^x\Delta\psi^{x,k} = -\left\{C^x + (A^x\psi^x)_\mu\Delta\mu^k\right\}$$

$$A^m\Delta\psi^{m,k} = -\left\{C^m - (b^m)_\mu\Delta\mu^k - (b^m)_{\psi^x}\Delta\psi^x\right\}.$$

7) Set $\alpha_k = 1$ and check if it ensures the sufficient decrease in the merit function

$$\varphi_{\eta_k}(p^k + \alpha^k\Delta p^k) \le \varphi_{\eta_k}(p^k) + \beta\alpha_k D_{\eta_k}(p^k;\Delta p^k)$$

where $\varphi_\eta(\mu,\psi^x,\psi^m) = f(\psi^m) + \eta\left\|C\right\|_1$

8) If the sufficient decreasing condition is satisfied by the searched step length, then set

$$\mu^{k+1} = \mu^k + \alpha^k \Delta\mu^k$$
$$\psi^{x,k+1} = \psi^{x,k} + \alpha^k \Delta\psi^{x,k}$$
$$\psi^{m,k+1} = \psi^{m,k} + \alpha^k \Delta\psi^{m,k}$$

otherwise the following restricted quadratic interpolation is performed:

$$\alpha^k = \max\left\{\frac{-0.5D\varphi_{\eta_k}(\alpha^k)^2}{\varphi_{\eta_k}(p^k + \alpha^k \Delta p^k) - \varphi_{\eta_k}(p^k) - \alpha^k D\varphi_{\eta_k}}, \alpha^{\min}\right\}$$

with a minimum step size $\alpha^{\min}$.

9) Evaluate C^{k+1}, g^{k+1} and A^{k+1}, and compute Y^{k+1} and Z^{k+1}.

10) Solve for $\lambda^{x,k+1}$ and $\lambda^{m,k+1}$ with a GMRES solver

$$\lambda^{m,k+1} = \left[A^{mT}\right]^{-1} Q^T (Q\psi^m - z)$$
$$\lambda^{x,k+1} = \left[A^{xT}\right]^{-1} (b^{mT})_{\psi^x} \lambda^{m,k+1}$$

and update the merit function parameter η_k by

$$\eta_{k+1} = 1.001 + \|\lambda_{k+1}\|_\infty .$$

11) Evaluate $g_r^{k+1} = (b^{mT})_\mu \lambda^{m,k+1} - (\psi^{xT} A^{xT})_\mu \lambda^{x,k+1}$.

12) Get $y^k = Y^{T,k+} g^{k+} - Y^{T,k} g^k$ and $s^k = \mu^{k+1} - \mu^k$.

13) Set $k = k + 1$ and return to Step 4 to check the convergence.

Note that our algorithm stated previously is similar to the reduced Hessian SQP algorithm applied to solving nonfluorescence optical tomographic problems [42].

As mentioned earlier, the rSQP algorithm does not require the exact solutions of the two radiative transfer equations involving the excitation and emission radiances during the reconstruction process. Instead, the rSQP scheme solves the two linearized forward equations as described in step 6), which allows us to utilize the incomplete solution obtained with the loose tolerance (10^{-2}–10^{-3}). Accordingly, we stop the GMRES iteration for the linearized forward solution of the rSQP scheme when its relative residual becomes smaller than 10^{-2}–10^{-3}, whereas for the forward solution of the lm-BFGS method, the tolerance of 10^{-10}–10^{-14} is used [41], [42].

III. Results and Discussion

A. Numerical Experiments

In this section, we show numerical results for reconstructions of the fluorescence absorption distribution $\mu_a^{x\to m}$ inside the medium by using the rSQP method and the lm-BFGS method. To illustrate the code performance, we compare the results of both schemes on two types of problems, with an emphasis on computational efficiency.

1) Setup of the Test Problems: We consider two different problems, which both involve a cylinder with an outer diameter of 2 cm. The chosen dimensions mimic fluorescence tomographic problems typical for small animal imaging. In the first case a fluorescent heterogeneity with a diameter of 0.15 cm is embedded inside the cylinder at $\Gamma = \{(x, y)|(x - 0.35)^2 + (y - 0.35)^2 = 0.2^2\}$ (see Fig. 1). The background medium has optical parameters of $\mu_a^x = 0.4\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the excitation wavelength and $\mu_a^x = 0.4\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the emission wavelength. The fluorescence heterogeneity has the absorbing coefficient of $\mu_a^{x\to m} = 0.05\ \text{cm}^{-1}$. In the second example, a stronger absorbing medium is considered: the optical parameters of the background medium are $\mu_a^x = 1.0\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the excitation wavelength and $\mu_a^x = 1.0\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the emission wavelength, and the absorption coefficient of the fluorescence heterogeneity is $\mu_a^{x\to m} = 0.5\ \text{cm}^{-1}$. In both examples, the quantum yield and fluorescence lifetime are assumed spatially constant at the values of $\tau = 4$ ns and $\eta = 0.95$, which is taken to match the corresponding properties of fluorophore (Fluorescein) used later in the experimental study. As shown in Fig. 1, four external sources are located on the tissue boundary close to the target fluorescent source and 66 detectors are equally distributed around the circular circumference of the medium defined by $\Gamma = \{(x, y)|x^2 + y^2 = 1\}$. The corresponding values of optical parameters for the two cases are given in Table I.

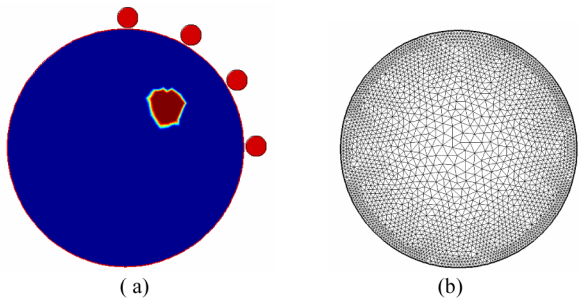


Fig. 1. Schematic of the test problems 1 and 2: circle with diameter of 2 cm. (a) Source–detector configuration: four sources (●) and 66 detectors around the surface; (b) computation domain with 4886 triangular elements.

TABLE I
Parameters Used in the Two Examples

	Problem 1	Problem 2
μ_a^x (cm^{-1})	0.4	1.0
$\mu_s'^x$ (cm^{-1})	15	15
μ_a^m (cm^{-1})	0.4	1.0
$\mu_s'^m$ (cm^{-1})	15	15
$\mu_a^{x\to m}$ (cm^{-1})	0.05	0.5
ω (MHz)	150	150
N_{source}	4	4
$N_{detector}$	66	66

For numerical experiments, we obtain synthetic data at detector locations for the emission wavelength by solving the two frequency-domain ERTs given by (1) and (2) with an original distribution of fluorescent sources and intrinsic absorption and scattering coefficients inside the medium. All synthetic data are generated on the mesh that is two times finer than that used for the reconstructions. The solution of the forward problem at detector locations provides the exact measurements z^{ex}. Since measurements always contain some noise, we add an error term to z^{ex} in the form $z^{obs} = z^{ex} + \varpi\sigma$, where σ is the estimated

standard deviation of measurement errors and ϖ is the random variable with normal distribution, zero mean and unitary standard deviation. With the use of such noisy data as the input to the reconstruction code, we examine the stability of the algorithms with respect to noises.

We stop the optimization process of the schemes when the following stopping criteria are satisfied:

$$\left|\mathcal{J}^{k+1}(u;x)-\mathcal{J}^{k}(u;x)\right|/\mathcal{J}^{k}(u;x)\le\varepsilon_1 \tag{32a}$$

$$\mathcal{J}^{k}(u;x)\le\varepsilon_2 \tag{32b}$$

$$r^{x}=\|A^{x}\psi_s^{x}-b_s^{x}\|<\varepsilon_3 \tag{32c}$$

$$r^{m}=\|A^{m}\psi_s^{m}-b_s^{m}\|<\varepsilon_4 \tag{32d}$$

where ε_1 and ε_2 are the small tolerances and ε_1 is set to 10^{-5} throughout this study, and ε_2 is chosen to have the same order of magnitude of measurement errors, which leads to sufficiently stable results in the *principle of discrepancy* [54]. Also the stopping criteria [see (33c) and (33d)] indicate the accuracy of the forward solutions: therefore ε_3 and ε_4 are set to the initial residual multiplied by a small value 10^{-2}–10^{-3} for the PDE-constrained method and 10^{-10} for the unconstrained method, i.e., $\varepsilon_{3(4)}=r_0^{x(m)}\cdot(10^{-2}\text{–}10^{-10})$. When the noise-free data is considered, the tolerance ε_3 is assigned a sufficiently small number (typically 10^{-6}). The four stopping criteria given by (32) are applied to both of the PDE-constrained (rSQP) and unconstrained (lm-BFGS) methods.

To measure the quality of reconstruction, we introduce the correlation factor $\rho(\mu^e,\mu^r)$ and the deviation factor $\delta(\mu^e,\mu^r)$ as used in [42], to give

$$\rho=\frac{\sum_{i=1}^{N_t}(\mu_i^e-\bar{\mu}_i^e)(\mu_i^r-\bar{\mu}_i^r)}{(N_t-1)\sigma(\mu^e)\sigma(\mu^r)},\quad \delta=\frac{\sqrt{\sum_{i=1}^{N_t}(\mu_i^e-\mu_i^r)^2N_t}}{\sigma(\mu^e)} \tag{33}$$

where $\bar{\mu}$ and $\sigma(\mu)$ are the mean value and the standard deviation for the spatial function of the fluorescence absorption coefficient. Similarly, μ^e and μ^r are the exact and reconstructed distributions of fluorescent sources, respectively. In terms of quality of the reconstruction results, the correlation coefficient indicates the degree of correlation between exact and estimated quantities while the deviation factor describes the discrepancy in absolute values of exact and estimated quantities. Accordingly, the closer $\rho(\mu^e,\mu^r)$ gets to unity, and the closer $\delta(\mu^e,\mu^r)$ gets to zero, the better is quality of reconstruction.

In the following sections, the PDE-constrained rSQP method and the unconstrained lm-BFGS method [32], [33] are applied to functional estimations of unknown fluorescence absorption coefficients for the two test problems, as given in Table I. All the simulations are carried out on a Pentium IV 3.0 GHz CPU processor.

2) Case 1 (CPU Times and Influence of Noise): With the first example, we compare the CPU time and the influence of noise in rSQP-based and lm-BFGS-based algorithms. To examine the effects of noise on the algorithm we use different SNR values varying from infinity (no noise) to 15 dB and 10 dB, with the later two representing typical noise levels encountered in [41]. Fig. 2 shows the maps of the reconstructed fluorescence absorption coefficients obtained for the 15 dB case. As shown in figures, the circular fluorescence perturbation, as shown in Fig. 1(a), is well identified by the two methods.

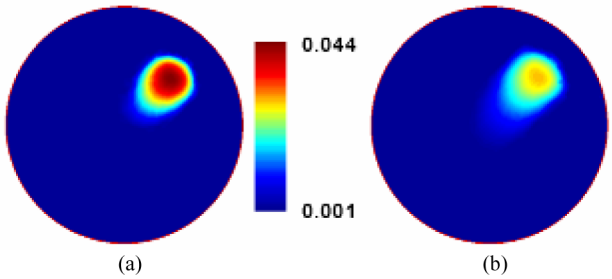


Fig. 2. Reconstructed fluorescence absorption coefficients $\mu_a^{x\to m}$ obtained for the first example using the 15 and 10 dB noise data by the PDE-constrained rSQP method. (a) 15 dB. (b) 10 dB.

TABLE II
RECONSTRUCTION QUALITY AND COMPUTATION TIMES WITH DIFFERENT NOISE LEVELS

SNR	Schemes	CPU time (*)	Cor. ρ	Dev. δ
∞	PDE-constr.	0.31 h (13)	0.78	**0.60**
	Unconstr.	3.98 h	0.77	0.60
15dB	PDE-constr.	0.33 h (13)	0.78	**0.70**
	Unconstr.	4.28 h	0.75	0.69
10dB	PDE-const.	0.35 h (17)	0.66	**0.81**
	Unconstr.	5.89 h	0.64	0.78

*Denotes the acceleration factor by the PDE-constrained method.

First, the CPU times are measured for the two methods. Table II shows that the PDE-constrained rSQP method leads to a significant reduction in the computation time in all cases considered here. For the case of the noise-free data, the PDE-constrained method takes only 0.31 h to converge, while the unconstrained lm-BFGS method takes about 3.98 h to meet the same convergence criterion. Therefore, the PDE-constrained method reduces the reconstruction time by a factor of about 13. A similar reduction is observed in the two other cases of different noise levels (see Table II). The 15 dB data take 0.33 h using the PDE-constrained method, while the unconstrained methods require 4.28 h. With the 10 dB data, the PDE-constrained code requires 0.35 h, while the unconstrained codes take 5.89 h to converge, which is approximately 17 times slower.

The main reason for this significant reduction in the CPU time can be explained by the fact that the PDE-constrained rSQP method does not require the exact solution of the forward problem at each iteration of optimization until it converges to the optimal solution, as mentioned earlier. Indeed, the PDE-constrained method utilizes the incomplete solution of the two linearized forward equations, obtained with the loose tolerance of 10^{-2} to 10^{-3} empirically chosen from our extensive study. We found that even a tolerance of 10^{-2} can generate a sequence of intermediate solutions satisfying the first order necessary conditions, while making reasonable progress toward the true solution through the optimization process. Fig. 3 illustrates this convergence behavior of the PDE-constrained method. It can be seen that the forward and inverse solutions converge simultaneously toward their optimal solutions, even when a loose tolerance is used. We usually terminate the GMRES iteration

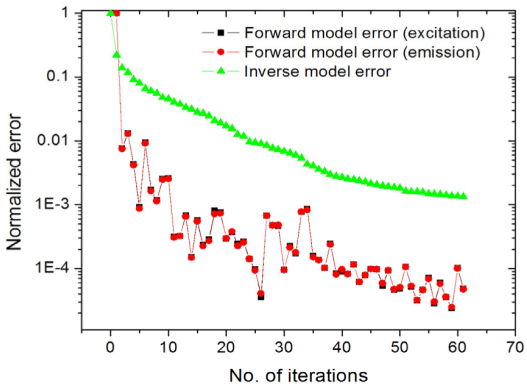


Fig. 3. Convergence history of the PDE-constrained method in the forward and inverse solutions when a loose tolerance of 0.01 is used for solving the QP problems.

process when the relative error in the solution becomes smaller than 10^{-10}, which is necessary for accurate evaluation of the objective function in the unconstrained method. Thus, the GMRES iteration process essentially has to perform a sufficient number of matrix-vector multiplications to reach the desired accuracy, which is the case using the unconstrained lm-BFGS method. On the other hand, the PDE-constrained rSQP scheme uses a less strict tolerance and stops the GMRES iteration at a much earlier stage of the iterations process. As a consequence, this scheme requires a smaller number of matrix-vector multiplications, which leads to significant time savings [42].

In addition to the CPU time, we measured the accuracy of reconstruction as a function of the SNR. The correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$, as defined in (33), are computed and the corresponding values are given in Table II. At noise levels of 10 to 15dB, the constrained and unconstrained methods make no significant difference in the accuracy; both schemes show a decrease in the correlation factor and an increase in the deviation factor as compared to the corresponding values of the noise-free data (∞ dB; see Table II). Therefore, it can be stated that the constrained and unconstrained methods are similar to each other in terms of response to noise in the data.

3) Case 2 (Effect of Initial Guess): In fluorescence tomography, the optimization scheme starts with a homogeneous initial guess of unknown fluorescence absorption coefficients, which is typically zero. However, it is well known that different initial guesses affect the reconstruction accuracy. For this reason, we discuss the robustness of the PDE-constrained rSQP code to initial guess.

For this study, we consider problem 2 (see Table I). The initial guesses of $\mu_a^{x\to m,0} = 0.1$ and 0.2 are made for the entire medium that correspond to 20% and 40%, respectively, of the absorption coefficient of the fluorescence target. We performed the reconstruction simulations with 20 dB noisy data and measured the image quality with the correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$. The reconstruction results and measures of $\rho(\mu^e, \mu^r)$ and $\delta(\mu^e, \mu^r)$ are given in Fig. 4 and Table III, respectively. It can be seen from Table III that the PDE-constrained rSQP and unconstrained lm-BFGS method lead to similar reconstruction results when the initial guesses of the fluorescence

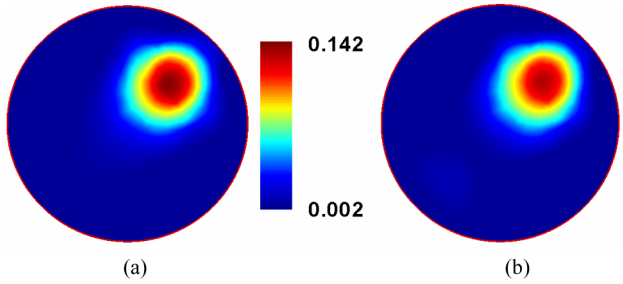


Fig. 4. Reconstructed fluorescence absorption coefficients $\mu_a^{x\to m}$ obtained for the second example using the initial guess of $\mu_a^{x\to m,0} = 0.1$. (a) PDE constrained. (b) Unconstrained.

TABLE III
RECONSTRUCTION QUALITY AND COMPUTATION TIMES WITH DIFFERENT INITIAL GUESSES

$\mu_a^{x\to m,0}$	Schemes	CPU time (*)	Cor. ρ	Dev. δ
0	PDE-constr.	0.45 h (12)	0.78	**0.84**
	Unconstr.	5.4 h	0.78	0.88
0.1 (20%)	PDE-constr.	0.47 (11)	0.77	**0.81**
	Unconstr.	5.2 h	0.76	0.83
0.2 (40%)	PDE-constr.	0.45 h (11)	0.75	**0.79**
	Unconstr.	4.9 h	0.78	0.81

*Denotes the acceleration factor by the PDE-constrained method.

absorption in the background medium are 20% and 40% of that of the target. For the 20% case, we observe $\rho(\mu^e, \mu^r) = 0.77$ and $\delta(\mu^e, \mu^r) = 0.81$ using the PDE-constrained method and $\rho(\mu^e, \mu^r) = 0.76$ and $\delta(\mu^e, \mu^r) = 0.83$ using the constrained lm-BFGS method. In the 40% case, $\rho(\mu^e, \mu^r) = 0.75$ and $\delta(\mu^e, \mu^r) = 0.79$, and $\rho(\mu^e, \mu^r) = 0.78$ and $\delta(\mu^e, \mu^r) = 0.81$ employing the constrained and unconstrained approach, respectively. However, the PDE-constrained code yields these results about 10 to 12 times faster than the unconstrained code. In the 20% case, the times to convergence are 0.47hrs and 5.2 h, respectively, and in the 40% case 0.45 h versus 4.9 h, respectively. We made similar observations for cases where the background optical properties differ as much as 50% or 100% from the target value.

B. Applications to the Experimental Data

In addition to numerical results, we studied the reconstruction of fluorescence absorption coefficients inside the medium using experimental data.

The lab phantom with a square base has a size of 2.2 cm $\times$ 2.2 cm $\times$ 10.3 cm ($X \times Y \times Z$). A fluorescent cylinder rod of $d = 2$ mm is filled with Fluorescein and embedded inside the medium. The optical properties of Fluorescein are $\mu_a^{x\to m} = 0.4\ \text{cm}^{-1}$, $\tau = 4.0$ ns, and $\eta = 0.93$ (see Table IV). With this phantom, we consider two setups. In both cases, a single fluorescent source (a cylindrical rod) is used, but it is placed in a different location inside the medium. Furthermore, the background optical properties are different. All other experimental conditions are kept the same. In the first case, a cylindrical rod that contains the fluorophore is placed in the center of the phantom. The background optical properties (a mixture of Intralipid and India ink) are $\mu_a^x = 1.04\ \text{cm}^{-1}$, $\mu_s'^x = 5.83\ \text{cm}^{-1}$ at excitation

TABLE IV
OPTICAL PARAMETERS USED FOR THE PHANTOM EXPERIMENT

	λ^1 = 475 nm	λ^2 = 515 nm
μ_a^x	1.04 cm^{-1}	0.97 cm^{-1}
$\mu_s'^x$	5.83 cm^{-1}	5.36 cm^{-1}
$\mu_a^{x\to m}$	0.38 cm^{-1}	0.06 cm^{-1}
τ	4.00 ns	
η	0.93	

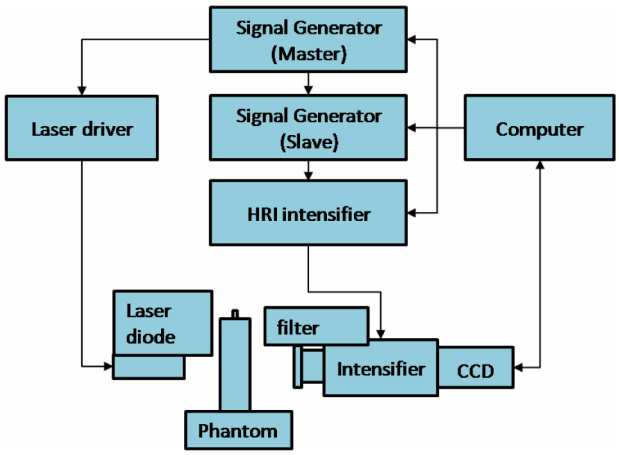


Fig. 5. Experimental setup for the phantom experiment.

wavelength of 475 nm and $\mu_a^x = 0.97$ cm^{-1}, $\mu_s'^x = 5.36$ cm^{-1} at emission wavelength of 515 nm. In the second example, the fluorescent rod is positioned about 6 mm off-center and the background medium is filled only with a 2% Intralipid solution.

Fig. 5 shows a schematic of the experimental setup. A frequency-domain system is used for fast 2-D imaging of modulated light transmitted through small-tissue volume. The main components of the system are the illumination module, the light detection module, and the modulation sources for the light source and detector. The master signal generator provides a sinusoidal ac input to the laser diode driver that supplies the laser diode (wavelength $\lambda = 475$ nm) with a bias and ac current. The modulated light that is transmitted through the phantom is imaged by a lens to an intensified CCD (ICCD) camera.

The system operates in homodyne mode, i.e., the gain of the ICCD is modulated by a slave signal generator at the same frequency as the laser [55]. As a result, a steady-state image at the intensifier output is imaged to the CCD. The signal in every pixel depends on the phase between source and detector modulation. Master and slave signal generators are linked together and the phase delay is adjustable. To detect the complete oscillation of the modulation, multiple images are taken at phase delays covering the range of 2π and are transferred to a computer.

Each side of the phantom was illuminated by a focused light source ($\lambda^1 = 475$ nm) and measurements of transmitted light intensities were made on the opposite side of the illumination side at $\lambda^2 = 515$ nm, for which Fluorescein emission is largest. The position of the laser spot ($\sim$1 mm in diameter) is at the center of each side of the phantom. We used one transmission side for the measurement: each side is illuminated and its opposite side only is measured with the CCD camera. The absolute phase delay can be accurately calibrated; however, the amplitude signal can only be determined relative to its absolute value. Therefore, for the image reconstruction, we chose to use the rescaled data that can eliminate this ambiguity: the measured data z_f due to fluorophores at the emission wavelength were rescaled to the reference data. In other words, we use the measured data z_f multiplied by the rescaling factor $r = p_r/z_r$, where z_r denotes the measurement of the background medium without a fluorescent probe at excitation wavelength and p_r denotes the prediction data obtained from the forward model for the same excitation wavelength [26]. For the second example, we measured the emission data only (i.e., we did not take the measurement of the background medium) and calibration procedure was different. In this case, we use the normalized data for reconstruction. We obtained the measurement data for dc and 150 MHz frequency.

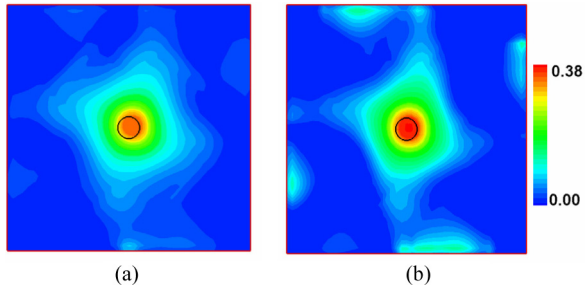


Fig. 6. Reconstructed maps of fluorophore absorption coefficients $\mu_a^{x\to m}$ (in centimeters inverse) inside the phantom using the dc and 150 MHz data. (a) 150 MHz. (b) DC.

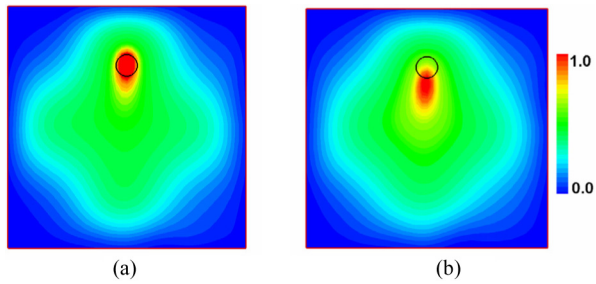


Fig. 7. Reconstructed maps of fluorophore absorption coefficients $\mu_a^{x\to m}$ (in centimeters inverse) inside the phantom using the dc and 150 MHz data. (a) 150 MHz. (b) DC.

We performed the reconstruction for a 2-D plane where the laser illumination and detection is performed. The spatial domain is discretized into a grid of 47×47 points and the angular domain into 24 discrete ordinates (S_4) to solve the radiative transfer equation. We started the reconstruction from an initial guess $\mu_a^{x\to m,0} = 0$ for both of the two cases. Figs. 6 and 7 show the reconstruction results for the first and second cases. Note that the reconstruction results of the constrained SQP method only are presented here because the unconstrained lm-BFGS method yielded almost same results with respect to the size and location of a fluorescent probe, as in the numerical study. Also the different metrics, i.e., the location and size of a reconstructed fluorescent probe, are used to assess the reconstruction results (see Table V) since, for the second case, there are some unknown effects caused by the unknown strength of a source.

For the first case, the difference between the dc and 150 MHz data is not significant: both methods yielded almost same

TABLE V
MEASURES OF ACCURACY IN THE RECONSTRUCTED IMAGES WITH THE EXPERIMENTAL DATA (UNIT: cm^{-1})

	Target	DC	150 MHz
$\mu_a^{x\to m}$ (case 1)	0.38	0.40	0.34
(x, y) (case 1)	(1.12, 1.12)	(1.14, 1.15)	(1.15, 1.14)
(x, y) (case 2)	(1.12, 1.7)	(1.11, 1.54)	(1.13, 1.72)

results with respect to the location and size of a fluorescent probe: $\mu_{a,\mathrm{dc}}^{x\to m} = 0.40\ \mathrm{cm}^{-1}$ and $\mu_{a,150\,\mathrm{MHz}}^{x\to m} = 0.34\ \mathrm{cm}^{-1}$ (see Table V). However, in the second case, with the 150 MHz data we obtained the more accurate reconstruction in the location of the fluorescent probe as compared to the dc data, which shows that the frequency data give better results than the dc data, especially in the identification of the target location. Note that since the source strength is not known, the reconstructed image is shown in a normalized scale [0, 1]. Again, CPU times are compared between the two methods. Here, we observed a similar speedup factor as previously determined in the numerical studies. The PDE-constrained rSQP method took about 6 min, while the unconstrained lm-BFGS method required 58 min to reach the same accuracy; hence a $\sim$10-fold acceleration was achieved.

IV. CONCLUSION

To accelerate the image reconstruction process for optical fluorescence tomography, we present here a PDE-constrained reduced Hessian SQP (rSQP) method. The proposed algorithm solves the two forward problems and one inverse problem, encountered in fluorescence tomography, at once by updating the excitation and emission radiances and the fluorescence absorption coefficient simultaneously in one iteration process. The frequency-domain ERT was employed as a light transport model for both excitation and emission radiances. To evaluate the performance of the proposed PDE-constrained scheme, we performed numerical experiments varying the optical parameters of the test problem and compared the results of the constrained approach with an unconstrained lm-BFGS method (the fastest unconstrained scheme in OT) in terms of computation time, accuracy, and robustness. Furthermore, we applied our code to the experimental data obtained for the lab phantom with a fluorescent probe inside.

In general, we found that the PDE-constrained method leads to a significant reduction in the computation time. We typically observe acceleration factors between 10 and 15 as compared to the unconstrained lm-BFGS method. We also studied the impact of noise on the image quality in the reconstruction process. While the image quality depends on the signal-to-noise level of the data, there appear to be little difference between the constrained and unconstrained methods. Furthermore, we found that the two approaches respond similarly to the choice of the initial guess. Therefore, the image quality deteriorates in a similar fashion as the initial guess is chosen further and further away from the actual background optical properties. However, the PDE-constrained codes yields results approximately ten times faster than the unconstrained code.

ACKNOWLEDGMENT

We also acknowledge Dr. A. D. Klose, Columbia University, for his many useful discussions on this study.

REFERENCES

[1] N. C. Deliolanis, J. Dunham, T. Wurdinger, J.-L. Figueiredo, T. Bakhos, and V. Ntziachristos, "In-vivo imaging of murine tumors using complete-angle projection fluorescence molecular tomography," *J. Biomed. Opt.*, vol. 14, pp. 030509-1–030509-3, 2009.

[2] A. Garofalakis, G. Zacharakis, H. Meyer, E. N. Economou, C. Mamalaki, J. Papamatheakis, D. Kioussis, V. Ntziachristos, and J. Ripoll, "Three-dimensional in vivo imaging of green fluorescent protein-expressing T cells in mice with noncontact fluorescence molecular tomography," *Mol. Imag.*, vol. 6, pp. 96–107, 2007.

[3] A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans," *Opt. Exp.*, vol. 15, pp. 6696–6716, 2007.

[4] R. Weissleder and V. Ntziachristos, "Shedding light onto live molecular targets," *Nat. Med.*, vol. 9, pp. 123–128, 2003.

[5] G. Choy, P. Choyke, and S. K. Libutti, "Current advances in molecular imaging: Noninvasive in vivo bioluminescent and fluorescent optical imaging in cancer research," *Mol. Imag.*, vol. 2, pp. 303–312, 2003.

[6] C. Bremer, V. Ntziachristos, and R. Weissleder, "Optical-based molecular imaging: Contrast agents and potential medical applications," *Eur. Radiol.*, vol. 13, pp. 231–243, 2001.

[7] D. J. Bornhop, C. H. Contag, K. Licha, and C. J. Murphy, "Advances in contrast agents, reporters, and detection," *J. Biomed. Opt.*, vol. 6, pp. 106–110, 2001.

[8] V. Ntziachristos, C. H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity in vivo," *Nat. Med.*, vol. 8, pp. 757–761, 2002.

[9] V. Ntziachristos, R. Bremer, and R. Weissleder, "Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging," *Eur. Radiol.*, vol. 13, pp. 195–208, 2003.

[10] Y. Chen, G. Zheng, Z. H. Zhang, D. Blessington, M. Zhang, H. Li, Q. Liu, L. Zhou, X. Intes, S. Achilefu, and B. Chance, "Metabolism-enhanced tumor localization by fluorescence imaging: in vivo animal studies," *Opt. Lett.*, vol. 28, pp. 2070–2072, 2003.

[11] D. Y. Paithankar, A. U. Chen, B. W. Pogue, M. S. Patterson, and E. M. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," *Appl. Opt.*, vol. 36, pp. 2260–2272, 1997.

[12] J. Chang, H. L. Graber, and R. L. Barbour, "Imaging of fluorescence in highly scattering media," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 810–822, Sep. 1997.

[13] H. Jiang, "Frequency-domain fluorescent diffusion tomography: A finite-element-based algorithm and simulations," *Appl. Opt.*, vol. 37, pp. 5337–5343, 1998.

[14] R. Roy and E. M. Sevick-Muraca, "A numerical study of gradient-based nonlinear optimization methods for contrast enhanced optical tomography," *Opt. Exp.*, vol. 9, pp. 49–65, 2001.

[15] M. A. O_Leary, D. A. Boas, X. D. Li, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.*, vol. 21, pp. 158–160, 1996.

[16] V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," *Opt. Lett.*, vol. 26, pp. 893–895, 2001.

[17] J. Lee and E. M. Sevick-Muraca, "Three-dimensional fluorescence enhanced optical tomography using referenced frequency-domain photon migration measurements at emission and excitation wavelengths," *J. Opt. Soc. Amer. A*, vol. 19, pp. 759–771, 2002.

[18] M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. SevickMuraca, "Three-dimensional, Bayesian image reconstruction from sparse and noisy data sets: Near-infrared fluorescence tomography," *Proc. Nat. Acad. Sci. USA*, vol. 99, pp. 9619–9624, 2002.

[19] 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.

[20] A. D. Kim and A. Ishimaru, "Optical diffusion of continuous-wave, pulsed, and density waves in scattering media and comparisons with radiative transfer," *Appl. Opt.*, vol. 37, pp. 5313–5319, 1998.

[21] R. Aronson and N. Corngold, "Photon diffusion coefficient in an absorbing medium," *J. Opt. Soc. Amer. A*, vol. 16, pp. 1066–1071, 1999.

[22] B. Chen, K. Stamnes, and J. J. Stamnes, "Validity of the diffusion approximation in bio-optical imaging," *Appl. Opt.*, vol. 40, pp. 6356–6366, 2001.
[23] R. Elaloufi, R. Carminati, and J.-J. Greffet, "Time-dependent transport through scattering media: from radiative transfer to diffusion," *J. Opt. A*, vol. 4, pp. S103–S108, 2002.
[24] A. H. Hielscher, A. E. Alcouffe, and R. L. Barbour, "Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues," *Phys. Med. Biol.*, vol. 43, pp. 1285–1302, 1998.
[25] A. D. Klose and A. H. Hielscher, "Fluorescence tomography with simulated data based on the equation of radiative transfer," *Opt. Lett.*, vol. 28, pp. 1019–1021, 2003.
[26] A. D. Klose, V. Ntziachristos, and A. H. Hielscher, "The inverse source problem based on the equation of radiative transfer," *J. Comput. Phys.*, vol. 202, pp. 323–345, 2005.
[27] L. Biegler, O. Ghattas, M. Heinkenschloss, and B. Bloemen, *Large-Scale PDE-Constrained Optimization*. New York: Springer-Verlag, 2003.
[28] R. Byrd, F. Curtis, and J. Nocedal, "An inexact SQP method for equality constrained optimization," *SIAM J. Optim.*, vol. 19, pp. 351–369, 2008.
[29] P. Gill, W. Murray, and M. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization SIAM review," vol. 47, pp. 99–131, 2005.
[30] Biros and O. Ghattas, "Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver," *SIAM J. Sci. Comput.*, vol. 27, pp. 687–713, 2003.
[31] L. Biegler, C. Schmid, and D. Ternet, *A Multiplier-Free, Reduced Hessian Method For Process Optimization, Large-Scale Optimization with Applications, Part II: Optimal Design and Control*. New York: Springer-Verlag, 1997, p. 101.
[32] E. Haber and U. Ascher, "Preconditioned all-at-once methods for large, sparse parameter estimation problems," *Inverse Probl.*, vol. 17, pp. 1847–1864, 2001.
[33] J. Bonnas, J. Gilbert, C. Lemaréchal, and C. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*. New York: Springer-Verlag, 2003.
[34] M. Burger and W. Muhlhuber, "Iterative regularization of parameter identification problems by sequential quadratic programming methods," *Inverse Probl.*, vol. 18, pp. 943–969, 2002.
[35] P. Boggs and J. Tolle, "Sequential quadratic programming for large-scale nonlinear optimization," *J. Comput. Appl. Math.*, vol. 124, pp. 123–137, 2000.
[36] J. L. Hu, Z. Wu, H. McCann, L. E. Davis, and C. G. Xie, "Sequential quadratic programming method for solution of electromagnetic inverse problems," *IEEE Trans. Antennas Propag.*, vol. 53, no. 8, pp. 2680–2687, Aug. 2005.
[37] D. Feng and T. Pulliam "All at-at-once reduced Hessian SQP scheme for aerodynamics design optimization," Tech. Rep., NASA Ames Research Center, Mountain View, CA, 1995.
[38] T. Coleman, J. Liu, and W. Yuan, "A Quasi-Newton quadratic penalty method for minimization subject to nonlinear equality constraints," *Comput. Optim. Appl.*, vol. 15, pp. 103–123, 2000.
[39] M. Lalee, J. Nocedal, and T. Plantenga, "On the implementation of an algorithm for large-scale equality constrained optimization," *SIAM J. Optim.*, vol. 8, pp. 682–706, 2003.
[40] L. Biegler, J. Nocedal, C. Schmid, and D. Ternet, "Numerical experience with a reduced Hessian method for large scale constrained optimization," *Comput. Optim. Appl.*, vol. 15, pp. 45–67, 2000.
[41] G. S. Abddoulaev, K. Ren, and A. H. Hielscher, "Optical tomography as a PDE-constrained optimization problem," *Inverse Probl.*, vol. 21, pp. 1507–1530, 2005.
[42] H. K. Kim and A. H. Hielscher, "A PDE-constrained reduced Hessian SQP method for optical tomography based on the frequency-domain equation of radiative transfer," *Inverse Probl.*, vol. 25, pp. 015010-1–015010-20, 2009.
[43] W. Bangerth and A. Joshi, "Adaptive finite element methods for the solution of inverse problems in optical tomography," *Inverse Probl.*, vol. 24, pp. 034011-1–034011-22, 2008.
[44] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Non-contact fluorescence optical tomography with scanning patterned illumination," *Opt. Exp.*, vol. 14, pp. 6516–6534, 2006.
[45] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer-Verlag, 2006.
[46] A. D. Klose and A. H. Hielscher, "Quasi-Newton methods in optical tomographic image reconstruction," *Inverse Probl.*, vol. 19, pp. 387–409, 2003.
[47] R, K. Ren, G. Bal, and A. H. Hielscher, "Frequency domain optical tomography with the equation of radiative transfer," *SIAM J. Sci. Comput.*, vol. 28, pp. 1463–1489, 2006.
[48] H. K. Kim and A. Charette, "A sensitivity function-based conjugate gradient method for optical tomography with the frequency-domain equation of radiative transfer," *J. Quant. Spec. Rad. Trans.*, vol. 104, pp. 24–39, 2007.
[49] L. G. Henyey and L. J. Greenstein, "Diffuse radiation in the galaxy," *Astrophys*, vol. 90, pp. 70–83, 1941.
[50] M. Modest, *Radiative Heat Transfer*. New York: MacGraw-Hill, 2003.
[51] S. Minkowycz, E. Sparrow, and J. Murthy, *Handbook of Numerical Heat Transfer*. Hoboken, NJ: Wiley, 2006.
[52] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: SIAM, 2003.
[53] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, vol. 3, pp. 856–869, 1986.
[54] O. M. Alifanov, *Inverse Heat Transfer Problems*. New York: Springer-Verlag, 1994.
[55] U. J. Netz, J. Beuthan, and A. H. Hielscher, "Multipixel system for gigahertz frequency-domain optical imaging of finger joints," *Rev. Sci. Instrum.*, vol. 79, no. 8, pp. 034301-1–034301-14, 2008.

Hyun Keol Kim received the B.S. degree from Seoul National University, Seoul, Korea, in 1999, and the M. S. and Ph.D. degrees in mechanical engineering from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2001 and 2004, respectively.

In 2005, he was with the Department of Applied Sciences, Quebec University at Chicoutimi, QC, Canada, as a Postdoctoral Research Fellow. In 2006, he joined the Department of Biomedical Engineering, Columbia University, New York, NY, where he spent three years as a Postdoctoral Research Scientist and is currently an Associate Research Scientist with the Biophotonics and Optical Radiology Laboratory. His current research interests include the development of state-of-the-art imaging software for optical tomography including bioluminescence and fluorescence tomography.

Dr. Kim was awarded the Oversea Postdoctoral Fellowship from Korea Research Foundation in 2005.

Jong Hwan Lee received the B.S. degree in electrical engineering from Kyungpook National University, Daegu, Korea, in 1999, and the M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 2006. He is currently working toward the Ph.D. degree with the Department of Biomedical Engineering, Columbia University, New York, NY.

His research interests include biomedical imaging and instrumentation. His current research focuses on developing molecular imaging and optical-tomography-imaging systems.

Andreas H. Hielscher (S'93–M'95) received the Ph.D. degree in electrical and computer engineering from Rice University, Houston, TX, in 1995.

After spending two years as a Postdoctoral Fellow with the Los Alamos National Laboratory, Los Alamos, NM, he joined the faculty at the State University of New York Downstate Medical Center, Stony Brook, NY. In September 2001, he moved to Columbia University, New York, NY, where he is currently the Director of the Biophotonics and Optical Radiology Laboratory. He holds joint appointments as an Associate Professor with the Departments of Biomedical Engineering and Radiology, Columbia University. His work focuses on the development of state-of-the-art imaging software and hardware for optical tomography. He applies this emerging technology to the imaging of cancer and joint diseases and uses it in support of drug development. He has authored or coauthored more than 120 articles in peer-reviewed scientific journals and conference proceedings.

Dr. Hielscher is currently an Associate Editor of the IEEE TRANSACTIONS OF MEDICAL IMAGING.

PDE-Constrained Fluorescence Tomography With the Frequency-Domain Equation of Radiative Transfer

Hyun Keol Kim, Jong Hwan Lee, and Andreas H. Hielscher, *Member, IEEE*

***Abstract*—We present the first fluorescence tomography algorithm that is based on a partial differential equation (PDE) constrained approach. PDE methods have been increasingly employed in many numerical applications, as they often lead to faster and more robust solutions of many inverse problems. In particular, we use a sequential quadratic programming (SQP) method, which allows solving the two forward problems in fluorescence tomography (one for the excitation and one for the emission radiances) and one inverse problem (for recovering the spatial distribution of the fluorescent sources) simultaneously by updating both forward and inverse variables in simultaneously at each of iteration of the optimization process. We evaluate the performance of this approach with numerical and experimental data using a transport-theory frequency-domain algorithm as forward model for light propagation in tissue. The results show that the PDE-constrained approach is computationally stable and accelerates the image reconstruction process up to a factor of 15 when compared to commonly employed unconstrained methods.**

***Index Terms*—Fluorescence tomography, frequency-domain equation of radiative transfer (ERT), partial differential equation (PDE) constrained optimization, sequential quadratic programming (SQP).**

I. Introduction

OPTICAL fluorescence tomography has emerged as a new imaging modality that makes use of light-emitting biomarkers, for diagnostic imaging of pathological changes in tissues [1]–[7]. By recovering fluorescence measurement data, this technique provides a spatial distribution of fluorescent biomarkers in tissue, which can be used in preclinical and clinical studies to monitor the migration of cells, the protein–protein interaction and the progression of a disease including cancerous tumors [5]. To this end, fluorescent biomarkers have been used to target specific molecules that provide functional information about biochemical processes preceding the development of diseases [6], [7]. Thus, fluorescence tomography allows for the early detection of diseases on a molecular level before typical symptoms or macroscopic changes appear.

Many imaging codes currently used employ a so-called model-based iterative image reconstruction (MOBIIR) scheme in which a forward model of light propagation in tissue is needed [8]–[10]. The light distribution inside the medium can be described by a system of two coupled forward equations: one for the excitation radiance caused by an external light source and one for the emission radiance by an internal fluorescent source. For high-scattering, low-absorbing media, both excitation and emission radiances can be accurately modeled by the diffusion approximation (DA) to the equation of radiative transfer (ERT) [11]–[19]. It is, however, well known that the DA becomes less accurate when absorption is increased and the medium is no longer scattering dominated. Problems also arise in low-scattering regions, such as void-like, fluid-filled areas and when small tissue volumes are considered, where boundary effects dominate [20]–[24]. In small animal imaging, many of these issues are encountered and the DA is of limited applicability. The problems related to the DA can be overcome by using the ERT that provides accurate prediction of light propagation in all types of media [25], [26]. However, in general, the ERT requires much longer computation times as compared to the DA and are much more complex to implement. ERT-based image reconstruction codes can take anywhere from several hours to even days to solve one single image reconstruction case. Thus, it is highly desirable to develop computationally efficient image reconstruction schemes that allow for using the ERT for optical fluorescence tomography. Motivated by this background, we have adapted a so-called PDE-constrained approach that makes use of a reduced Hessian SQP (rSQP) method. In this approach the forward and inverse problems are solved simultaneously, which has shown in other applications [27]–[40] to lead to significant savings in the total image reconstruction time.

Recently, our group has introduced the PDE-constrained approach to the field of optical tomography. However, that work was limited to recover the absorption and scattering coefficients inside biological tissues and had not addressed the problems encountered in fluorescence tomography. First, using the frequency-domain version of the radiative transport equation as the forward model of light propagation in tissues, Abddoulaev *et al*. [41] suggested an augmented Lagrangian method (ALM) that solves the forward and inverse problems simultaneously. In numerical phantom studies, they demonstrated that PDE-constrained ALM codes can significantly reduce computation times. Subsequently, Kim and Hielscher [42] replaced the ALM

Manuscript received September 25, 2009. This work was supported in part by the National Cancer Institute under Grant NCI-4R33CA118666 and Grant NCI-U54CA126513-039001 at the National Institutes of Health. The work of A. H. Hielscher was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, the National Heart, Lung, and Blood Institute, the National Institute for Biomedical Imaging and Bioengineering, the National Cancer Institute, the Whitaker Foundation for Biomedical Engineering, and the New York State Office of Science, Technology, and Academic Research.

H. K. Kim and J. H. Lee are with the Department of Biomedical Engineering, Columbia University, New York, NY 10027 USA (e-mail: hkk2107@columbia.edu; jl3132@columbia.edu).

A. H. Hielscher is with the Department of Biomedical Engineering and Radiology, Columbia University, New York, NY 10027 USA (e-mail: ahh2004@columbia.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/JSTQE.2009.2038112

by the rSQP method and showed that even faster convergence and higher numerical stability can be obtained. Using numerical and experimental data, they observed an increase in the reconstruction speed by up to a factor of 30 as compared to standard codes that do not use PDE constraints. In the study at hand, we have adapted the approach previously presented by Kim and Hielscher to the problem of fluorescence tomography.

So far only one publication exists that applies the PDE-constrained methods to fluorescence tomography. Recently, Joshi *et al*. [43], [44] have reported on a Gauss–Newton-based constrained method. However, their work used the frequency-domain DA as a forward model of light propagation and not the ERT. Moreover, the authors reported that the Gauss–Newton approach is computationally more expensive than the reduced Hessian approach with a BFGS updating formula.

In this paper, we go beyond past studies and present the first PDE-constrained fluorescence image reconstruction algorithm. The adaptation of the nonfluorescence code for problems in fluorescence tomography is nontrivial. Inverse source problems are known to be notoriously ill-posed and unlike in nonfluorescence problems, two coupled ERTs for the excitation and emission radiances have to be treated simultaneously in the framework of PDE-constrained optimization. In the following, we will first provide details on the structure and implementation of the fluorescence forward model (see Section II-A). This is followed by a description of the implementation of the PDE-constrained rSQP approach for fluorescence tomography (see Section II-B–II-D). Subsequently, we evaluate the performance of the PDE-constrained scheme using numerical and experimental results (see Section III). We focus on computational efficiency by comparing the new algorithm with an unconstrained code that makes use of the limited-memory Broyden–Fletcher–Goldfarb–Shanno (lm-BFGS) method [45]. The lm-BFGS method is known to be the most efficient gradient-based unconstrained optimization method [46], [47], and therefore, provides a good standard for comparison. Finally we draw some conclusions in Section IV.

II. Methods

A. Light Propagation Model

In frequency-domain optical fluorescence tomography, the external light source is amplitude-modulated in the 10 Hz–1 GHz frequency range and the demodulated transmitted intensities and phase shifts are measured on the tissue surface, using wavelength-dependent filters to distinguish between excitation and emission signals.

The generation and propagation of fluorescence light in biological tissue can be accurately modeled by two coupled frequency-domain equations of radiative transfer [47], [48] as

$$\left[(\nabla\cdot\Omega)+\mu_a^x+\mu_s^x+\mu_a^{x\to m}+\frac{i\varpi}{c}\right]\psi^x(\mathbf{r},\Omega,\varpi)=\int_{4\pi}p(\Omega',\Omega)\psi^x(\mathbf{r},\Omega',\varpi)d\Omega' \tag{1}$$

$$\left[(\nabla\cdot\Omega)+\mu_a^m+\mu_s^m+\frac{i\varpi}{c}\right]\psi^m(\mathbf{r},\Omega,\varpi)=\int_{4\pi}p(\Omega',\Omega)\psi^m(\mathbf{r},\Omega',\varpi)d\Omega'+\frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}. \tag{2}$$

The first equation describes the propagation of excitation light emanating from the external light source and the second represents the propagation of emission light due to the fluorescent source in tissue. Here, the superscripts *x* and *m* denote excitation and emission, respectively. Hence, μ_a^x and μ_s^x are the absorption and scattering coefficients in units of cm^{-1} at the excitation wavelength; $\mu_a^{x\to m}$ is the absorption coefficient of a fluorescent source in tissue, in unit of cm^{-1}; μ_a^x and μ_s^x are the absorption and scattering coefficients in units of cm^{-1} at the emission wavelength. $\psi^x(\mathbf{r},\Omega,\varpi)$ and $\psi^m(\mathbf{r},\Omega,\varpi)$ denote the excitation and emission radiances, respectively, in units of $\text{W}\cdot\text{cm}^{-2}\cdot\text{st}^{-1}$; η denotes the quantum yield by which the fluorescent source emits light in transit from excitation state to ground state; $\tau(\mathbf{r})$is the local lifetime of a fluorescent source. Note that $\phi(\mathbf{r},\varpi)$ appearing in (2) denotes the excitation fluence at position $\mathbf{r}$ defined by $\phi(\mathbf{r},\varpi)=\int_{4\pi}\psi^x d\Omega$ in unit W/cm^2. For a phase function denoted by $p(\Omega',\Omega)$, we use here the Henyey–Greestein phase function [49] that is commonly used in tissue optics. As shown in (1), the excitation light $\psi^x(\mathbf{r},\Omega,\varpi)$ is absorbed and scattered by the intrinsic medium, and further attenuated by the fluorophore absorption $\mu_a^{x\to m}$, and then excites a fluorochrome inside the tissue at position $\mathbf{r}$. The excited fluorophore constitutes a light source inside the medium that reemits radiation $\psi^m(\mathbf{r},\Omega,\varpi)$ typically at longer wavelength [see (2)]. The local strength of a fluorescent source is directly proportional to the local fluorophore absorption $\mu_a^{x\to m}\phi(\mathbf{r},\varpi)$ and the quantum yield η and the local fluorophore lifetime $\tau(\mathbf{r})$.

The corresponding boundary conditions for the two coupled equations are given

$$\psi_b^x|_{\vec{n}_b\cdot\Omega<0}=\psi^{x0}\big|_{\vec{n}_b\cdot\Omega<0}+R(\Omega',\Omega)\cdot\psi^x|_{\vec{n}_b\cdot\Omega'>0} \tag{3}$$

$$\psi_b^m|_{\vec{n}_b\cdot\Omega<0}=R(\Omega',\Omega)\cdot\psi^m|_{\vec{n}_b\cdot\Omega'>0} \tag{4}$$

where $R(\Omega',\Omega)$ is the reflectivity at Fresnel interface [50] from direction Ω' to direction Ω, ψ_b^{x0} is the radiation intensity due to the external source function and subscript b denotes the boundary surface of the medium, while $\vec{n}_b$ is the unit normal vector pointing outward the boundary surface. For discretization of the two equations (1) and (2), we employ a node-centered finite-volume approach in the spatial domain and a discrete ordinates method in the angular domain. The node-centered finite-volume method takes advantage of the beneficial properties of both the finite-element and finite-volume methods by combining the conservation properties of the finite-volume formulation and the geometric flexibility of the finite-element approach.

Following an unstructured finite-volume discrete-ordinates method [42], [51], the discretized forms of the two ERTs given by (1) and (2) are obtained by integrating (1) and (2) over the

control volume with a divergence theorem as

$$\sum_{j=1}^{N_{\text{surf}}} (\vec{n}_j \cdot \Omega^l)\psi_j^{x,l} dA_j + \left(\mu_a^x + \mu_s^x + \mu_a^{x\to m} + \frac{i\omega}{c}\right)\psi_N^{x,l}\Delta V_N = \mu_s^x \sum_{l'=1}^{N_\Omega} \psi_N^{x,l'} p^{l'l} w^{l'} \quad (5)$$

$$\sum_{j=1}^{N_{\text{surf}}} (\vec{n}_j \cdot \Omega^l)\psi_j^{m,l} dA_j + \left(\mu_a^m + \mu_s^m + \frac{i\omega}{c}\right)\psi_N^{m,l}\Delta V_N = \mu_s^m \Delta V_N \sum_{l'=1}^{N_\Omega} \psi_N^{m,l'} p^{l'l} w^{l'} + \frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}\Delta V_N \quad (6)$$

where N_{surf} and N_Ω are the number of surfaces surrounding the node $N(=1,\ldots,N_t)$ and the number of discrete ordinates based on the level symmetric scheme, respectively, $\vec{n}_j$ denotes the surface normal vector, and $\psi_j^{x,l}$and $\psi_j^{m,l}$ denote the excitation and emission radiances defined on the jth surface in direction l. Also the surface intensities $\psi_j^{x,l}$ and $\psi_j^{m,l}$ are related to the nodal intensities $\psi_N^{x,l}$ and $\psi_j^{m,l}$ by the second-order spatial differencing scheme [51]. It can be easily seen that each ERT involves N_t spatial unknown intensities coupled into N_Ω directions, thus leading to the total $N_t \cdot N_\Omega$ unknowns. The system of (1) and (2) can be solved by using any iterative solvers as far as they are reliable. We employ here a matrix-based iterative linear solver that enables updating all the radiation intensities $\psi_N^{x,l}$ (or $\psi_N^{m,l}$) simultaneously, which leads to fast convergence as compared to source iteration-based techniques. After discretization for all nodes, we finally obtain two linear systems of algebraic equations as

$$A^x\psi^x = b^x \quad (7a)$$

$$A^m\psi^m = b^m. \quad (7b)$$

Each line denoted by $i\,(i=1,\ldots,N_tN_\Omega)$ of the matrix A contains the coefficients of the discretized forms [see (7a) and (7b)] established at node number N and direction l.

The excitation light source comes into the term b^x after discretization on a boundary node N_b, while the fluorescent source comes into the term b^m after discretization on internal node N, as

$$b_{N_b}^{x,l} = -\sum_j \left[1 - \max(n_j\cdot\Omega^l / \left|n_j\cdot\Omega^l\right|, 0)\right](n_j\cdot\Omega^l)A_j\psi_{N_b}^{0,l} \quad (8a)$$

$$b_N^{m,l} = \frac{1}{4\pi}\frac{\eta\mu_a^{x\to m}\phi(\mathbf{r},\varpi)}{(1-\varpi\tau(\mathbf{r}))}\Delta V_N \quad (8b)$$

where $\psi_{N_b}^{x0,l}$ is the external source function on boundary node N_b in direction l.

Since we treat the frequency-domain equations of radiative transfer directly, the matrix formulations given by (7) contain complex-valued elements. As a result, the complex-valued algebraic linear equations are solved with a complex version of the generalized minimum residual (GMRES) solver [52], [53].

The two equations given by (1) and (2) provide the predictions of the excitation radiance $\psi_j^{x,l}$ and transmitted emission radiance $\psi_j^{m,l}$ on the boundary surface. Also these two equations are used as the constraints within the PDE-constrained SQP method. Next we discuss the inverse model, which is used to obtain the spatial distribution of a fluorescent source inside the medium that best fits the measured emission data.

B. PDE-Constrained Optimization for Fluorescence Tomography

To understand the PDE-constrained approach to fluorescence tomography, we first summarize the most important aspects, previously described in [42]. The optical fluorescence tomographic problem can be formulated in more general terms as

$$\begin{aligned} \min \quad & f(x,u) \\ \text{s.t.} \quad & C(x,u) = 0 \end{aligned} \quad (9)$$

where $x \in R^n$ is the vector of inverse variables, $u \in Z^m$ is the vector of forward variables, $f(x;u)$ is an objective function that quantifies the difference between measured and predicted intensities and $c(x;u)=0$ is a discretized version of combined excitation and emission equations. The problem given by (9) is often referred to as "PDE-constrained" optimization since the optimal solution at minimum of f is constrained by equality condition $c(x;u)=0$ represented by two partial differential equations (PDEs).

Traditional methods for solving (9) is to treat the forward variable u as a dependent variable of the inverse variable x, i.e., $u=\tilde{c}(x)$, which makes it possible to replace the prediction vector u in f of (9) by its forward solution vector. As a result, the problem (9) can be reformulated as

$$\min \tilde{f}(x) = f(x,\tilde{c}(x)) \quad (10)$$

which is often referred to as "unconstrained" because the equality constraint $c(x;u)=0$ is eliminated in (10), i.e., $\tilde{f}$ is now a function of x only. Thus, the forward solution $u=\tilde{c}(x)$ has to be obtained for evaluation of $\tilde{f}$, i.e., the complete solutions of the two ERTs are required for the excitation and emission radiances at each of optimization iterations. As a consequence, the associated optimization procedure becomes a computationally very demanding process, both with respect to time and memory. Nonetheless, this approach has been widely used for the solution of optical fluorescence tomographic problems mainly because of easiness of implementation. The existing fluorescence tomographic codes [11]–[19] belong to this approach: the conjugate gradient (CG) approach [12], the quasi-Newton (QN) approach [14], [25], [26], the Jacobian approach [11], [13]–[19], etc. [15].

Another approach to solve (9) is to treat the forward variable u and the inverse variable x independently, which enables solving the PDE-constrained problem (9) directly by updating the forward and inverse variables simultaneously at each of optimization iterations. Typically, an extended objective function called "Lagrangian" is introduced as follows:

$$L(x,u;\lambda) = f + \lambda^T c. \quad (11)$$

Here, $\lambda \in Z^m$ is called the vector of Lagrange multipliers. The simultaneous solutions of forward and inverse problems can be achieved at points satisfying first-order necessary conditions [41], [42] where the gradient of L in (11) vanishes with respect λ, u, and x, respectively.

One major advantage of this PDE-constrained approach is that the complete solution of the forward problem is not required until convergence is reached. In other words, PDE-constrained methods allow for using the inexact solution of the forward problem into solving the inverse problem, which leads to a significant reduction in the total reconstruction time. The solution accuracy of the forward problem is iteratively controlled as the inverse solution goes toward the optimum. Optimization techniques of this kind have seen rapid developments mainly in applications with airfoil design, flow variable optimization, and electromagnetic inverse problems [27]–[40].

As discussed in Section II, the fluorescence tomographic problem is to find a vector of unknowns $x = (\mu_a^{x\rightarrow m}, \tau, \eta)$, assuming that all other intrinsic properties inside the medium, μ_a^x, μ_s^x and μ_a^m, μ_s^m, are already known at the excitation and emission wavelengths, such that

$$\min f(\psi^m) = \frac{1}{2}\sum_{s,d} |Q_d\psi_s^m - z_{s,d}|^2$$

subject to

$$\begin{aligned} C_s^x &= A^x\psi_s^x - b_s^x = 0, \qquad s = 1,\ldots,N_s \\ C_s^m &= A^m\psi_s^m - b_s^m = 0, \qquad s = 1,\ldots,N_s \end{aligned} \tag{12}$$

where $f(\psi_s^m)$ is the objective function that quantifies the difference between predictions and measurements of emitted light made on the tissue surface; C^x and C^m are the discretized versions of the two coupled radiative transfer equations. In this study we focus our attention to the inverse problem of reconstructing the spatial distribution of the fluorophore absorption $\mu_a^{x\rightarrow m}$ assuming a spatially homogenous τ and η distribution inside the medium, which means that no quenching occurs. Therefore, we hereafter denote $\mu_a^{x\rightarrow m}$ by μ omitting its superscript and subscript.

By introducing a Lagrangian function, the previous PDE-constrained optimization problem can be restated as

$$\begin{aligned} L(\mu;\psi^x,\psi^m;\lambda^x,\lambda^m) = \frac{1}{2}|Q\psi^m - z|^2 + \lambda^{xT}(A^x\psi^x - b^x) \\ + \lambda^{mT}(A^m\psi^m - b^m). \end{aligned} \tag{13}$$

The simultaneous solutions of the forward and inverse problems can then be obtained at points in which the following five PDEs, i.e., the first derivatives of the Lagrangian function with respect to each of variables, become zero

$$L_\mu \equiv \frac{\partial L}{\partial \mu} = (\psi^{xT}A^{xT})_\mu\lambda^x - (b^{mT})_\mu\lambda^m = 0 \tag{14a}$$

$$L_{\psi^x} \equiv \frac{\partial L}{\partial \psi^x} = A^{xT}\lambda^x - (b^{mT})_{\psi^x}\lambda^m = 0 \tag{14b}$$

$$L_{\psi^m} \equiv \frac{\partial L}{\partial \psi^m} = Q^T(Q\psi^m - z) + A^{mT}\lambda^m = 0 \tag{14c}$$

$$L_{\lambda^x} \equiv \frac{\partial L}{\partial \lambda^x} = A^x\psi^x - b^x = 0 \tag{14d}$$

$$L_{\lambda^m} \equiv \frac{\partial L}{\partial \lambda^m} = A^m\psi^m - b^m = 0 \tag{14e}$$

where the first equation represents the sensitivity equation with respect to the inverse variable μ, and the second and third equations can be viewed as the adjoint equations for λ^x and λ^m, and the last two equations represent the two forward equations given by (7).

The so-called Karush–Khun–Tucker (KKT) system given by (14) can be solved with Newton's method as

$$\begin{bmatrix} W & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta p \\ \Delta\lambda \end{bmatrix} = -\begin{bmatrix} L_p \\ L_\lambda \end{bmatrix} \tag{15}$$

where the block matrix W denotes the Hessian matrix of the Lagrangian function L with respect to each of unknowns $p = (\mu, \psi^x, \psi^m)$ and the block matrix A denotes the Jacobian matrix of constraints C^x and C^m with respect to each of unknowns $p = (\mu, \psi^x, \psi^m)$. Also Δp and $\Delta\lambda$ denote $[\Delta\mu, \Delta\psi^x, \Delta\psi^m]^T$ and $[\Delta\lambda^x, \Delta\lambda^m]^T$, respectively. The algebraic system given by (15) can then be solved efficiently through the reduced Hessian SQP scheme as will be shown shortly.

C. Reduced Hessian SQP

The reduced Hessian SQP method is an established method that solves nonlinear optimization problems with relatively low cost and fast convergence [27]. Employing this rSQP method for solving the previous KKT system (15) is equivalent to finding the minimum to a quadratic approximation of the Lagrangian function L subject to the linearized constraints C^x and C^m, which gives the following quadratic programming problem so that

$$\min \Delta p^T g_p + \frac{1}{2}\Delta p^T W \Delta p$$

subject to

$$A\Delta p + C = 0 \tag{16}$$

where $g = \nabla f = [g_\mu, g_{\psi^x}, g_{\psi^m}]$ denotes the gradients of the objective function f with respect to each of unknowns (μ, ψ^x, ψ^m) and W^k is the full Hessian (or approximations) of the Lagrangian function. Since the quadratic problem (16) is equivalent to the system (15), the linearized constraints in (16) correspond to the last block of the KKT system (15) as

$$L_{\mu\lambda^x}\Delta\mu + L_{\psi^x\lambda^x}\Delta\psi^x + L_{\psi^m\lambda^x}\Delta\psi^m = -L_{\lambda^x} \tag{17a}$$

$$L_{\mu\lambda^m}\Delta\mu + L_{\psi^x\lambda^m}\Delta\psi^x + L_{\psi^m\lambda^m}\Delta\psi^m = -L_{\lambda^m} \tag{17b}$$

where

$$L_{\mu\lambda^x} = (A^x\psi^x)_\mu, L_{\psi^x\lambda^x} = A^x, L_{\psi^m\lambda^x} = 0 \tag{18a}$$

$$L_{\mu\lambda^m} = (-b^m)_\mu, L_{\psi^x\lambda^m} = (-b^m)_{\psi^x}, L_{\psi^m\lambda^m} = A^m. \tag{18b}$$

Again (17) can be explicitly written into the following form:

$$A^x\Delta\psi^x = -\{C^x + (A^x\psi^x)_\mu\Delta\mu\} \tag{19a}$$

$$A^m\Delta\psi^m = -\{C^m - (b^m)_\mu\Delta\mu - (b^m)_{\psi^x}\Delta\psi^x\}. \tag{19b}$$

The full Hessian of the Lagrangian function is often difficult to obtain and its approximation by the updating schemes tends to create large dense matrix $(n+m)\times(n+m)$. These difficulties can be overcome by dropping certain noncritical second-order terms of the full Hessian matrix. In what follows, we describe the standard reduced Hessian SQP method based on separation of variables.

Since A^x and A^m are invertible, the vector Δp can be decomposed into two parts as follows:

$$\Delta p = Z + Y\Delta\mu \tag{20}$$

where

$$Z = \begin{bmatrix} 0 \\ -(A^x)^{-1}C^x \\ -(A^m)^{-1}\{C^m + (b^m)_{\psi^x}(A^x)^{-1}C^x\} \end{bmatrix} \tag{21}$$

and

$$Y = \begin{bmatrix} I \\ -(A^x)^{-1}(A^x\psi^x)_\mu \\ (A^m)^{-1}\left\{(b^m)_\mu - (b^m)_{\psi^x}(A^x)^{-1}(A^x\psi^x)_\mu\right\} \end{bmatrix}. \tag{22}$$

The choices of Z^k and Y^k is one challenging problem arising in practical implementation of the reduced SQP scheme. In this study, we have followed the popular choices for Z^k and Y^k as used by many authors [27], [31], [36], [37], [40].

By substituting (20) into (16) and differentiating the resulting expression with respect to $\Delta\mu$, and using the identity $Y^TA^T = 0$, we can obtain the reduced Hessian of the form:

$$H_r\Delta\mu = -(g_r + d) \tag{23}$$

where $H_r = Y^TWY$ denotes the reduced Hessian, $g_r = Y^Tg$ denotes the reduced gradient, and $d = Y^TWZ$ is called the cross-term. Thus, the reduced SQP method requires much less memory than the full SQP one, i.e., only a small $(n\times n)$ matrix needs to be maintained and updated at each of optimization iterations. From (23), the inverse solution $\Delta\mu$ and the forward solutions $\Delta\psi^x$ and $\Delta\psi^x$ can be obtained, respectively, as

$$\Delta\mu = -(H_r)^{-1}(g_r + d_r) \tag{24}$$

$$\Delta\psi = \begin{bmatrix} \Delta\psi^x \\ \Delta\psi^m \end{bmatrix} = Z - Y(H_r)^{-1}(g_r + d_r). \tag{25}$$

At the new iterate, the Lagrangian multiplier vectors are updated from

$$\lambda^m = \left[A^{mT}\right]^{-1}Q^T(Q\psi^m - z) \tag{26a}$$

$$\lambda^x = \left[A^{xT}\right]^{-1}(b^{mT})_{\psi^x}\lambda^m. \tag{26b}$$

Using (22), we can rewrite the reduced gradient as

$$\begin{aligned} g_r &= Y^Tg \\ &= (b^{mT})_\mu\lambda^m - (\psi^{xT}A^{xT})_\mu\lambda^x. \end{aligned} \tag{27}$$

For large-scale applications, it is desirable to avoid the direct computation of the reduced Hessian H_r^k and its matrix inversion $(H_r^k)^{-1}$. Accordingly, we approximate the matrix-vector product of $(H_r^k)^{-1}g_r^k$ directly by using the limited-memory BFGS updating formula [42], [45], which is another important feature of our study that enables the proposed rSQP scheme to be applied to large-scale optimization problems. Also note that we ignore the cross-term vector $d_r^k = Y^{kT}W^kZ^k$ as in other works [27], [36], [37].

The global convergence of the rSQP scheme is ensured by line search on the following real-valued l_1 merit function defined as:

$$\varphi_\eta(\mu,\psi^x,\psi^m) = f(\psi^m) + \eta\left\|C\right\|_1 \tag{28}$$

which is chosen here for its simplicity and low computational cost. The directional derivative of $\varphi_\eta(\mu,\psi^x,\psi^m)$ along the descent direction Δp is given by

$$D\varphi_\eta = g^{kT}\Delta p - \eta\left\|C\right\|_1. \tag{29}$$

Thus, the descent property $D\varphi_\eta(\mu,\psi^x,\psi^m) < 0$ can be maintained by choosing [50]

$$\eta > \frac{g^T\Delta p}{\left\|C\right\|_1}. \tag{30}$$

At the new iterate given by $p^{k+1} = p^k + \alpha^k\Delta p$, the merit function (28) is successively monitored to ensure the global progress toward the solution while line search is performed to find a step length that can provide the sufficient decrease in the merit function as [40]

$$\varphi_{\eta_k}(p^k + \alpha^k\Delta p^k) \le \varphi_{\eta_k}(p^k) + \beta\alpha_k D_{\eta_k}(p^k;\Delta p^k) \tag{31}$$

where β is a small value that determines the line search accuracy, typically set to 0.01 (or smaller).

D. Computational Algorithm of the rSQP Scheme

The algorithm that makes use of the reduced Hessian SQP method, as described in Section II-B and II-C has the following structure:

1) Set $k=0$ and initialize $\mu^0, \psi^{x,0}, \psi^{m,0}$, and $H_r^0 = I$.
2) Solve $A^{mT}\lambda^{m,0} = -Q^T(Q\psi^{m,0} - z)$ for $\lambda^{m\,0}$, and solve $A^{xT}\lambda^{x,0} = (-b^{mT})_{\psi^{x\,0}}\lambda^{m,0}$ for $\lambda^{x\,0}$.
3) Compute $g_r^0 = (b^{mT})_\mu\lambda^{m,0} - (\psi^{xT}A^{xT})_\mu\lambda^{x,0}$.
4) Check the stopping criteria: if satisfied, stop.
5) Solve $H_r^k\Delta\mu^k = -g_r^k$ for $\Delta\mu$ via a damped BFGS updating formula.
6) Solve the two QP problems for $\Delta\psi^{x\,k}$ and $\Delta\psi^{m\,k}$.

$$A^x\Delta\psi^{x,k} = -\left\{C^x + (A^x\psi^x)_\mu\Delta\mu^k\right\}$$

$$A^m\Delta\psi^{m,k} = -\left\{C^m - (b^m)_\mu\Delta\mu^k - (b^m)_{\psi^x}\Delta\psi^x\right\}.$$

7) Set $\alpha_k = 1$ and check if it ensures the sufficient decrease in the merit function

$$\varphi_{\eta_k}(p^k + \alpha^k\Delta p^k) \le \varphi_{\eta_k}(p^k) + \beta\alpha_k D_{\eta_k}(p^k;\Delta p^k)$$

where $\varphi_\eta(\mu,\psi^x,\psi^m) = f(\psi^m) + \eta\left\|C\right\|_1$

8) If the sufficient decreasing condition is satisfied by the searched step length, then set

$$\mu^{k+1} = \mu^k + \alpha^k \Delta\mu^k$$
$$\psi^{x,k+1} = \psi^{x,k} + \alpha^k \Delta\psi^{x,k}$$
$$\psi^{m,k+1} = \psi^{m,k} + \alpha^k \Delta\psi^{m,k}$$

otherwise the following restricted quadratic interpolation is performed:

$$\alpha^k = \max\left\{\frac{-0.5D\varphi_{\eta_k}(\alpha^k)^2}{\varphi_{\eta_k}(p^k + \alpha^k \Delta p^k) - \varphi_{\eta_k}(p^k) - \alpha^k D\varphi_{\eta_k}}, \alpha^{\min}\right\}$$

with a minimum step size $\alpha^{\min}$.

9) Evaluate C^{k+1}, g^{k+1} and A^{k+1}, and compute Y^{k+1} and Z^{k+1}.

10) Solve for $\lambda^{x,k+1}$ and $\lambda^{m,k+1}$ with a GMRES solver

$$\lambda^{m,k+1} = \left[A^{mT}\right]^{-1} Q^T (Q\psi^m - z)$$
$$\lambda^{x,k+1} = \left[A^{xT}\right]^{-1} (b^{mT})_{\psi^x} \lambda^{m,k+1}$$

and update the merit function parameter η_k by

$$\eta_{k+1} = 1.001 + \|\lambda_{k+1}\|_\infty .$$

11) Evaluate $g_r^{k+1} = (b^{mT})_\mu \lambda^{m,k+1} - (\psi^{xT} A^{xT})_\mu \lambda^{x,k+1}$.

12) Get $y^k = Y^{T,k+} g^{k+} - Y^{T,k} g^k$ and $s^k = \mu^{k+1} - \mu^k$.

13) Set $k = k + 1$ and return to Step 4 to check the convergence.

Note that our algorithm stated previously is similar to the reduced Hessian SQP algorithm applied to solving nonfluorescence optical tomographic problems [42].

As mentioned earlier, the rSQP algorithm does not require the exact solutions of the two radiative transfer equations involving the excitation and emission radiances during the reconstruction process. Instead, the rSQP scheme solves the two linearized forward equations as described in step 6), which allows us to utilize the incomplete solution obtained with the loose tolerance (10^{-2}–10^{-3}). Accordingly, we stop the GMRES iteration for the linearized forward solution of the rSQP scheme when its relative residual becomes smaller than 10^{-2}–10^{-3}, whereas for the forward solution of the lm-BFGS method, the tolerance of 10^{-10}–10^{-14} is used [41], [42].

III. Results and Discussion

A. Numerical Experiments

In this section, we show numerical results for reconstructions of the fluorescence absorption distribution $\mu_a^{x\to m}$ inside the medium by using the rSQP method and the lm-BFGS method. To illustrate the code performance, we compare the results of both schemes on two types of problems, with an emphasis on computational efficiency.

1) Setup of the Test Problems: We consider two different problems, which both involve a cylinder with an outer diameter of 2 cm. The chosen dimensions mimic fluorescence tomographic problems typical for small animal imaging. In the first case a fluorescent heterogeneity with a diameter of 0.15 cm is embedded inside the cylinder at $\Gamma = \{(x, y)|(x - 0.35)^2 + (y - 0.35)^2 = 0.2^2\}$ (see Fig. 1). The background medium has optical parameters of $\mu_a^x = 0.4\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the excitation wavelength and $\mu_a^x = 0.4\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the emission wavelength. The fluorescence heterogeneity has the absorbing coefficient of $\mu_a^{x\to m} = 0.05\ \text{cm}^{-1}$. In the second example, a stronger absorbing medium is considered: the optical parameters of the background medium are $\mu_a^x = 1.0\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the excitation wavelength and $\mu_a^x = 1.0\ \text{cm}^{-1}$ and $\mu_s^x = 15\ \text{cm}^{-1}$ at the emission wavelength, and the absorption coefficient of the fluorescence heterogeneity is $\mu_a^{x\to m} = 0.5\ \text{cm}^{-1}$. In both examples, the quantum yield and fluorescence lifetime are assumed spatially constant at the values of $\tau = 4$ ns and $\eta = 0.95$, which is taken to match the corresponding properties of fluorophore (Fluorescein) used later in the experimental study. As shown in Fig. 1, four external sources are located on the tissue boundary close to the target fluorescent source and 66 detectors are equally distributed around the circular circumference of the medium defined by $\Gamma = \{(x, y)|x^2 + y^2 = 1\}$. The corresponding values of optical parameters for the two cases are given in Table I.

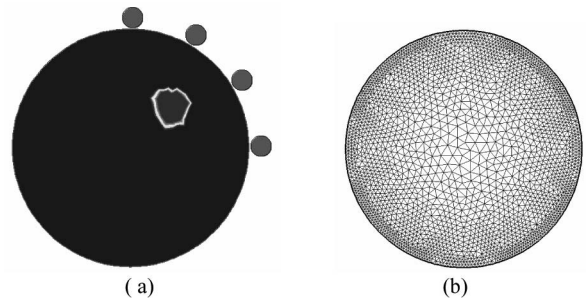


Fig. 1. Schematic of the test problems 1 and 2: circle with diameter of 2 cm. (a) Source–detector configuration: four sources (●) and 66 detectors around the surface; (b) computation domain with 4886 triangular elements.

TABLE I
Parameters Used in the Two Examples

	Problem 1	Problem 2
μ_a^x (cm^{-1})	0.4	1.0
$\mu_s'^x$ (cm^{-1})	15	15
μ_a^m (cm^{-1})	0.4	1.0
$\mu_s'^m$ (cm^{-1})	15	15
$\mu_a^{x\to m}$ (cm^{-1})	0.05	0.5
ω (MHz)	150	150
N_{source}	4	4
$N_{detector}$	66	66

For numerical experiments, we obtain synthetic data at detector locations for the emission wavelength by solving the two frequency-domain ERTs given by (1) and (2) with an original distribution of fluorescent sources and intrinsic absorption and scattering coefficients inside the medium. All synthetic data are generated on the mesh that is two times finer than that used for the reconstructions. The solution of the forward problem at detector locations provides the exact measurements z^{ex}. Since measurements always contain some noise, we add an error term to z^{ex} in the form $z^{obs} = z^{ex} + \varpi\sigma$, where σ is the estimated

standard deviation of measurement errors and ϖ is the random variable with normal distribution, zero mean and unitary standard deviation. With the use of such noisy data as the input to the reconstruction code, we examine the stability of the algorithms with respect to noises.

We stop the optimization process of the schemes when the following stopping criteria are satisfied:

$$\left|\mathcal{J}^{k+1}(u;x) - \mathcal{J}^{k}(u;x)\right| / \mathcal{J}^{k}(u;x) \leq \varepsilon_1 \tag{32a}$$

$$\mathcal{J}^{k}(u;x) \leq \varepsilon_2 \tag{32b}$$

$$r^x = \|A^x \psi_s^x - b_s^x\| < \varepsilon_3 \tag{32c}$$

$$r^m = \|A^m \psi_s^m - b_s^m\| < \varepsilon_4 \tag{32d}$$

where ε_1 and ε_2 are the small tolerances and ε_1 is set to 10^{-5} throughout this study, and ε_2 is chosen to have the same order of magnitude of measurement errors, which leads to sufficiently stable results in the *principle of discrepancy* [54]. Also the stopping criteria [see (33c) and (33d)] indicate the accuracy of the forward solutions: therefore ε_3 and ε_4 are set to the initial residual multiplied by a small value 10^{-2}–10^{-3} for the PDE-constrained method and 10^{-10} for the unconstrained method, i.e., $\varepsilon_{3(4)} = r_0^{x(m)} \cdot (10^{-2}$–$10^{-10})$. When the noise-free data is considered, the tolerance ε_3 is assigned a sufficiently small number (typically 10^{-6}). The four stopping criteria given by (32) are applied to both of the PDE-constrained (rSQP) and unconstrained (lm-BFGS) methods.

To measure the quality of reconstruction, we introduce the correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$ as used in [42], to give

$$\rho = \frac{\sum_{i=1}^{N_t} (\mu_i^e - \bar{\mu}_i^e)(\mu_i^r - \bar{\mu}_i^r)}{(N_t - 1)\sigma(\mu^e)\sigma(\mu^r)}, \quad \delta = \frac{\sqrt{\sum_{i=1}^{N_t} (\mu_i^e - \mu_i^r)^2 N_t}}{\sigma(\mu^e)} \tag{33}$$

where $\bar{\mu}$ and $\sigma(\mu)$ are the mean value and the standard deviation for the spatial function of the fluorescence absorption coefficient. Similarly, μ^e and μ^r are the exact and reconstructed distributions of fluorescent sources, respectively. In terms of quality of the reconstruction results, the correlation coefficient indicates the degree of correlation between exact and estimated quantities while the deviation factor describes the discrepancy in absolute values of exact and estimated quantities. Accordingly, the closer $\rho(\mu^e, \mu^r)$ gets to unity, and the closer $\delta(\mu^e, \mu^r)$ gets to zero, the better is quality of reconstruction.

In the following sections, the PDE-constrained rSQP method and the unconstrained lm-BFGS method [32], [33] are applied to functional estimations of unknown fluorescence absorption coefficients for the two test problems, as given in Table I. All the simulations are carried out on a Pentium IV 3.0 GHz CPU processor.

2) Case 1 (CPU Times and Influence of Noise): With the first example, we compare the CPU time and the influence of noise in rSQP-based and lm-BFGS-based algorithms. To examine the effects of noise on the algorithm we use different SNR values varying from infinity (no noise) to 15 dB and 10 dB, with the later two representing typical noise levels encountered in [41]. Fig. 2 shows the maps of the reconstructed fluorescence absorption coefficients obtained for the 15 dB case. As shown in figures, the circular fluorescence perturbation, as shown in Fig. 1(a), is well identified by the two methods.

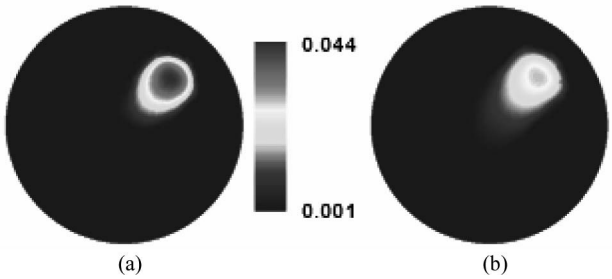


Fig. 2. Reconstructed fluorescence absorption coefficients $\mu_a^{x\to m}$ obtained for the first example using the 15 and 10 dB noise data by the PDE-constrained rSQP method. (a) 15 dB. (b) 10 dB.

TABLE II
RECONSTRUCTION QUALITY AND COMPUTATION TIMES WITH DIFFERENT NOISE LEVELS

SNR	Schemes	CPU time (*)	Cor. ρ	Dev. δ
∞	PDE-constr.	0.31 h (13)	0.78	**0.60**
	Unconstr.	3.98 h	0.77	0.60
15dB	PDE-constr.	0.33 h (13)	0.78	**0.70**
	Unconstr.	4.28 h	0.75	0.69
10dB	PDE-const.	0.35 h (17)	0.66	**0.81**
	Unconstr.	5.89 h	0.64	0.78

*Denotes the acceleration factor by the PDE-constrained method.

First, the CPU times are measured for the two methods. Table II shows that the PDE-constrained rSQP method leads to a significant reduction in the computation time in all cases considered here. For the case of the noise-free data, the PDE-constrained method takes only 0.31 h to converge, while the unconstrained lm-BFGS method takes about 3.98 h to meet the same convergence criterion. Therefore, the PDE-constrained method reduces the reconstruction time by a factor of about 13. A similar reduction is observed in the two other cases of different noise levels (see Table II). The 15 dB data take 0.33 h using the PDE-constrained method, while the unconstrained methods require 4.28 h. With the 10 dB data, the PDE-constrained code requires 0.35 h, while the unconstrained codes take 5.89 h to converge, which is approximately 17 times slower.

The main reason for this significant reduction in the CPU time can be explained by the fact that the PDE-constrained rSQP method does not require the exact solution of the forward problem at each iteration of optimization until it converges to the optimal solution, as mentioned earlier. Indeed, the PDE-constrained method utilizes the incomplete solution of the two linearized forward equations, obtained with the loose tolerance of 10^{-2} to 10^{-3} empirically chosen from our extensive study. We found that even a tolerance of 10^{-2} can generate a sequence of intermediate solutions satisfying the first order necessary conditions, while making reasonable progress toward the true solution through the optimization process. Fig. 3 illustrates this convergence behavior of the PDE-constrained method. It can be seen that the forward and inverse solutions converge simultaneously toward their optimal solutions, even when a loose tolerance is used. We usually terminate the GMRES iteration

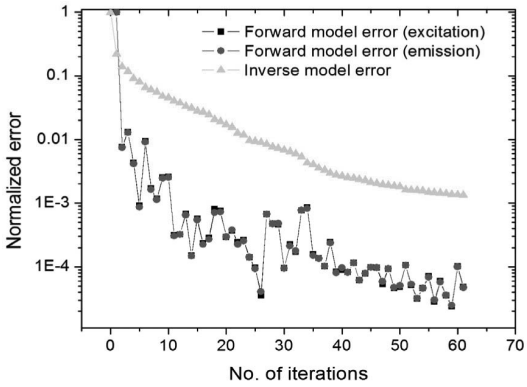


Fig. 3. Convergence history of the PDE-constrained method in the forward and inverse solutions when a loose tolerance of 0.01 is used for solving the QP problems.

process when the relative error in the solution becomes smaller than 10^{-10}, which is necessary for accurate evaluation of the objective function in the unconstrained method. Thus, the GMRES iteration process essentially has to perform a sufficient number of matrix-vector multiplications to reach the desired accuracy, which is the case using the unconstrained lm-BFGS method. On the other hand, the PDE-constrained rSQP scheme uses a less strict tolerance and stops the GMRES iteration at a much earlier stage of the iterations process. As a consequence, this scheme requires a smaller number of matrix-vector multiplications, which leads to significant time savings [42].

In addition to the CPU time, we measured the accuracy of reconstruction as a function of the SNR. The correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$, as defined in (33), are computed and the corresponding values are given in Table II. At noise levels of 10 to 15dB, the constrained and unconstrained methods make no significant difference in the accuracy; both schemes show a decrease in the correlation factor and an increase in the deviation factor as compared to the corresponding values of the noise-free data (∞ dB; see Table II). Therefore, it can be stated that the constrained and unconstrained methods are similar to each other in terms of response to noise in the data.

3) Case 2 (Effect of Initial Guess): In fluorescence tomography, the optimization scheme starts with a homogeneous initial guess of unknown fluorescence absorption coefficients, which is typically zero. However, it is well known that different initial guesses affect the reconstruction accuracy. For this reason, we discuss the robustness of the PDE-constrained rSQP code to initial guess.

For this study, we consider problem 2 (see Table I). The initial guesses of $\mu_a^{x\to m,0} = 0.1$ and 0.2 are made for the entire medium that correspond to 20% and 40%, respectively, of the absorption coefficient of the fluorescence target. We performed the reconstruction simulations with 20 dB noisy data and measured the image quality with the correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$. The reconstruction results and measures of $\rho(\mu^e, \mu^r)$ and $\delta(\mu^e, \mu^r)$ are given in Fig. 4 and Table III, respectively. It can be seen from Table III that the PDE-constrained rSQP and unconstrained lm-BFGS method lead to similar reconstruction results when the initial guesses of the fluorescence

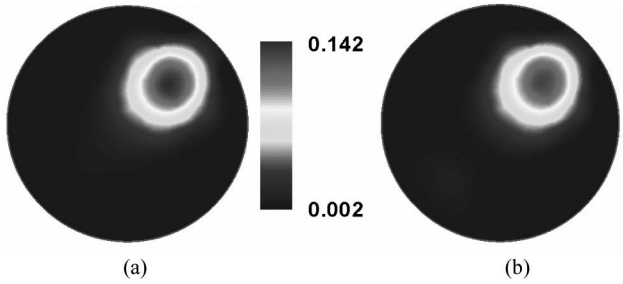


Fig. 4. Reconstructed fluorescence absorption coefficients $\mu_a^{x\to m}$ obtained for the second example using the initial guess of $\mu_a^{x\to m,0} = 0.1$. (a) PDE constrained. (b) Unconstrained.

TABLE III
RECONSTRUCTION QUALITY AND COMPUTATION TIMES WITH DIFFERENT INITIAL GUESSES

$\mu_a^{x\to m,0}$	Schemes	CPU time (*)	Cor. ρ	Dev. δ
0	PDE-constr.	0.45 h (12)	0.78	**0.84**
	Unconstr.	5.4 h	0.78	0.88
0.1 (20%)	PDE-constr.	0.47 (11)	0.77	**0.81**
	Unconstr.	5.2 h	0.76	0.83
0.2 (40%)	PDE-constr.	0.45 h (11)	0.75	**0.79**
	Unconstr.	4.9 h	0.78	0.81

*Denotes the acceleration factor by the PDE-constrained method.

absorption in the background medium are 20% and 40% of that of the target. For the 20% case, we observe $\rho(\mu^e, \mu^r) = 0.77$ and $\delta(\mu^e, \mu^r) = 0.81$ using the PDE-constrained method and $\rho(\mu^e, \mu^r) = 0.76$ and $\delta(\mu^e, \mu^r) = 0.83$ using the constrained lm-BFGS method. In the 40% case, $\rho(\mu^e, \mu^r) = 0.75$ and $\delta(\mu^e, \mu^r) = 0.79$, and $\rho(\mu^e, \mu^r) = 0.78$ and $\delta(\mu^e, \mu^r) = 0.81$ employing the constrained and unconstrained approach, respectively. However, the PDE-constrained code yields these results about 10 to 12 times faster than the unconstrained code. In the 20% case, the times to convergence are 0.47hrs and 5.2 h, respectively, and in the 40% case 0.45 h versus 4.9 h, respectively. We made similar observations for cases where the background optical properties differ as much as 50% or 100% from the target value.

B. Applications to the Experimental Data

In addition to numerical results, we studied the reconstruction of fluorescence absorption coefficients inside the medium using experimental data.

The lab phantom with a square base has a size of 2.2 cm $\times$ 2.2 cm $\times$ 10.3 cm ($X \times Y \times Z$). A fluorescent cylinder rod of $d = 2$ mm is filled with Fluorescein and embedded inside the medium. The optical properties of Fluorescein are $\mu_a^{x\to m} = 0.4\ \text{cm}^{-1}$, $\tau = 4.0$ ns, and $\eta = 0.93$ (see Table IV). With this phantom, we consider two setups. In both cases, a single fluorescent source (a cylindrical rod) is used, but it is placed in a different location inside the medium. Furthermore, the background optical properties are different. All other experimental conditions are kept the same. In the first case, a cylindrical rod that contains the fluorophore is placed in the center of the phantom. The background optical properties (a mixture of Intralipid and India ink) are $\mu_a^x = 1.04\ \text{cm}^{-1}$, $\mu_s'^x = 5.83\ \text{cm}^{-1}$ at excitation

TABLE IV
OPTICAL PARAMETERS USED FOR THE PHANTOM EXPERIMENT

	λ^1 = 475 nm	λ^2 = 515 nm
μ_a^x	1.04 cm^{-1}	0.97 cm^{-1}
$\mu_s'^x$	5.83 cm^{-1}	5.36 cm^{-1}
$\mu_a^{x\to m}$	0.38 cm^{-1}	0.06 cm^{-1}
τ	4.00 ns	
η	0.93	

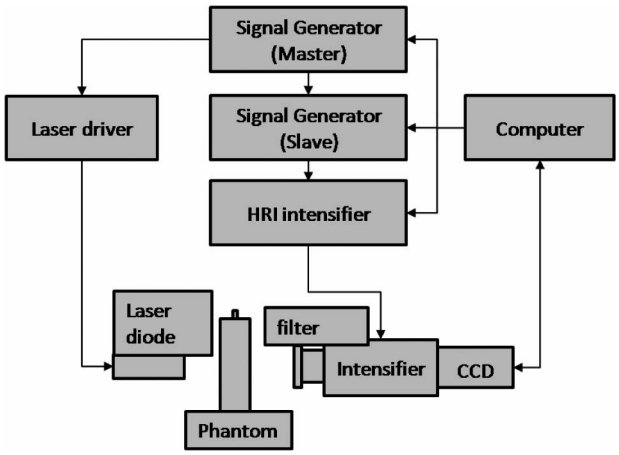


Fig. 5. Experimental setup for the phantom experiment.

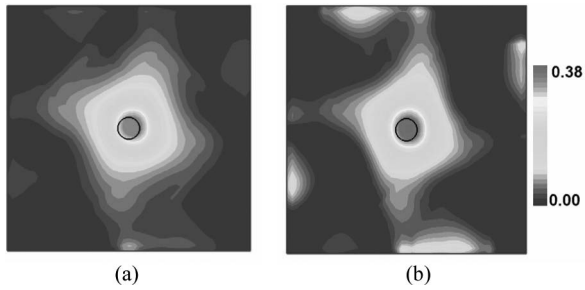


Fig. 6. Reconstructed maps of fluorophore absorption coefficients $\mu_a^{x\to m}$ (in centimeters inverse) inside the phantom using the dc and 150 MHz data. (a) 150 MHz. (b) DC.

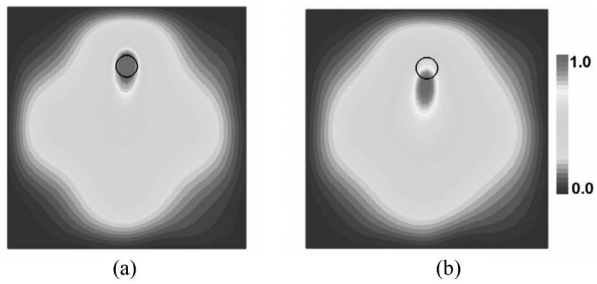


Fig. 7. Reconstructed maps of fluorophore absorption coefficients $\mu_a^{x\to m}$ (in centimeters inverse) inside the phantom using the dc and 150 MHz data. (a) 150 MHz. (b) DC.

wavelength of 475 nm and $\mu_a^x = 0.97$ cm^{-1}, $\mu_s'^x = 5.36$ cm^{-1} at emission wavelength of 515 nm. In the second example, the fluorescent rod is positioned about 6 mm off-center and the background medium is filled only with a 2% Intralipid solution.

Fig. 5 shows a schematic of the experimental setup. A frequency-domain system is used for fast 2-D imaging of modulated light transmitted through small-tissue volume. The main components of the system are the illumination module, the light detection module, and the modulation sources for the light source and detector. The master signal generator provides a sinusoidal ac input to the laser diode driver that supplies the laser diode (wavelength $\lambda = 475$ nm) with a bias and ac current. The modulated light that is transmitted through the phantom is imaged by a lens to an intensified CCD (ICCD) camera.

The system operates in homodyne mode, i.e., the gain of the ICCD is modulated by a slave signal generator at the same frequency as the laser [55]. As a result, a steady-state image at the intensifier output is imaged to the CCD. The signal in every pixel depends on the phase between source and detector modulation. Master and slave signal generators are linked together and the phase delay is adjustable. To detect the complete oscillation of the modulation, multiple images are taken at phase delays covering the range of 2π and are transferred to a computer.

Each side of the phantom was illuminated by a focused light source ($\lambda^1 = 475$ nm) and measurements of transmitted light intensities were made on the opposite side of the illumination side at $\lambda^2 = 515$ nm, for which Fluorescein emission is largest. The position of the laser spot ($\sim$1 mm in diameter) is at the center of each side of the phantom. We used one transmission side for the measurement: each side is illuminated and its opposite side only is measured with the CCD camera. The absolute phase delay can be accurately calibrated; however, the amplitude signal can only be determined relative to its absolute value. Therefore, for the image reconstruction, we chose to use the rescaled data that can eliminate this ambiguity: the measured data z_f due to fluorophores at the emission wavelength were rescaled to the reference data. In other words, we use the measured data z_f multiplied by the rescaling factor $r = p_r/z_r$, where z_r denotes the measurement of the background medium without a fluorescent probe at excitation wavelength and p_r denotes the prediction data obtained from the forward model for the same excitation wavelength [26]. For the second example, we measured the emission data only (i.e., we did not take the measurement of the background medium) and calibration procedure was different. In this case, we use the normalized data for reconstruction. We obtained the measurement data for dc and 150 MHz frequency.

We performed the reconstruction for a 2-D plane where the laser illumination and detection is performed. The spatial domain is discretized into a grid of 47×47 points and the angular domain into 24 discrete ordinates (S_4) to solve the radiative transfer equation. We started the reconstruction from an initial guess $\mu_a^{x\to m,0} = 0$ for both of the two cases. Figs. 6 and 7 show the reconstruction results for the first and second cases. Note that the reconstruction results of the constrained SQP method only are presented here because the unconstrained lm-BFGS method yielded almost same results with respect to the size and location of a fluorescent probe, as in the numerical study. Also the different metrics, i.e., the location and size of a reconstructed fluorescent probe, are used to assess the reconstruction results (see Table V) since, for the second case, there are some unknown effects caused by the unknown strength of a source.

For the first case, the difference between the dc and 150 MHz data is not significant: both methods yielded almost same

TABLE V
MEASURES OF ACCURACY IN THE RECONSTRUCTED IMAGES WITH THE EXPERIMENTAL DATA (UNIT: cm^{-1})

	Target	DC	150 MHz
$\mu_a^{x\to m}$ (case 1)	0.38	0.40	0.34
(x, y) (case 1)	(1.12, 1.12)	(1.14, 1.15)	(1.15, 1.14)
(x, y) (case 2)	(1.12, 1.7)	(1.11, 1.54)	(1.13, 1.72)

results with respect to the location and size of a fluorescent probe: $\mu_{a,\mathrm{dc}}^{x\to m} = 0.40\ \mathrm{cm}^{-1}$ and $\mu_{a,150\,\mathrm{MHz}}^{x\to m} = 0.34\ \mathrm{cm}^{-1}$ (see Table V). However, in the second case, with the 150 MHz data we obtained the more accurate reconstruction in the location of the fluorescent probe as compared to the dc data, which shows that the frequency data give better results than the dc data, especially in the identification of the target location. Note that since the source strength is not known, the reconstructed image is shown in a normalized scale [0, 1]. Again, CPU times are compared between the two methods. Here, we observed a similar speedup factor as previously determined in the numerical studies. The PDE-constrained rSQP method took about 6 min, while the unconstrained lm-BFGS method required 58 min to reach the same accuracy; hence a $\sim$10-fold acceleration was achieved.

IV. CONCLUSION

To accelerate the image reconstruction process for optical fluorescence tomography, we present here a PDE-constrained reduced Hessian SQP (rSQP) method. The proposed algorithm solves the two forward problems and one inverse problem, encountered in fluorescence tomography, at once by updating the excitation and emission radiances and the fluorescence absorption coefficient simultaneously in one iteration process. The frequency-domain ERT was employed as a light transport model for both excitation and emission radiances. To evaluate the performance of the proposed PDE-constrained scheme, we performed numerical experiments varying the optical parameters of the test problem and compared the results of the constrained approach with an unconstrained lm-BFGS method (the fastest unconstrained scheme in OT) in terms of computation time, accuracy, and robustness. Furthermore, we applied our code to the experimental data obtained for the lab phantom with a fluorescent probe inside.

In general, we found that the PDE-constrained method leads to a significant reduction in the computation time. We typically observe acceleration factors between 10 and 15 as compared to the unconstrained lm-BFGS method. We also studied the impact of noise on the image quality in the reconstruction process. While the image quality depends on the signal-to-noise level of the data, there appear to be little difference between the constrained and unconstrained methods. Furthermore, we found that the two approaches respond similarly to the choice of the initial guess. Therefore, the image quality deteriorates in a similar fashion as the initial guess is chosen further and further away from the actual background optical properties. However, the PDE-constrained codes yields results approximately ten times faster than the unconstrained code.

ACKNOWLEDGMENT

We also acknowledge Dr. A. D. Klose, Columbia University, for his many useful discussions on this study.

REFERENCES

[1] N. C. Deliolanis, J. Dunham, T. Wurdinger, J.-L. Figueiredo, T. Bakhos, and V. Ntziachristos, "In-vivo imaging of murine tumors using complete-angle projection fluorescence molecular tomography," *J. Biomed. Opt.*, vol. 14, pp. 030509-1–030509-3, 2009.

[2] A. Garofalakis, G. Zacharakis, H. Meyer, E. N. Economou, C. Mamalaki, J. Papamatheakis, D. Kioussis, V. Ntziachristos, and J. Ripoll, "Three-dimensional in vivo imaging of green fluorescent protein-expressing T cells in mice with noncontact fluorescence molecular tomography," *Mol. Imag.*, vol. 6, pp. 96–107, 2007.

[3] A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans," *Opt. Exp.*, vol. 15, pp. 6696–6716, 2007.

[4] R. Weissleder and V. Ntziachristos, "Shedding light onto live molecular targets," *Nat. Med.*, vol. 9, pp. 123–128, 2003.

[5] G. Choy, P. Choyke, and S. K. Libutti, "Current advances in molecular imaging: Noninvasive in vivo bioluminescent and fluorescent optical imaging in cancer research," *Mol. Imag.*, vol. 2, pp. 303–312, 2003.

[6] C. Bremer, V. Ntziachristos, and R. Weissleder, "Optical-based molecular imaging: Contrast agents and potential medical applications," *Eur. Radiol.*, vol. 13, pp. 231–243, 2001.

[7] D. J. Bornhop, C. H. Contag, K. Licha, and C. J. Murphy, "Advances in contrast agents, reporters, and detection," *J. Biomed. Opt.*, vol. 6, pp. 106–110, 2001.

[8] V. Ntziachristos, C. H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity in vivo," *Nat. Med.*, vol. 8, pp. 757–761, 2002.

[9] V. Ntziachristos, R. Bremer, and R. Weissleder, "Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging," *Eur. Radiol.*, vol. 13, pp. 195–208, 2003.

[10] Y. Chen, G. Zheng, Z. H. Zhang, D. Blessington, M. Zhang, H. Li, Q. Liu, L. Zhou, X. Intes, S. Achilefu, and B. Chance, "Metabolism-enhanced tumor localization by fluorescence imaging: in vivo animal studies," *Opt. Lett.*, vol. 28, pp. 2070–2072, 2003.

[11] D. Y. Paithankar, A. U. Chen, B. W. Pogue, M. S. Patterson, and E. M. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," *Appl. Opt.*, vol. 36, pp. 2260–2272, 1997.

[12] J. Chang, H. L. Graber, and R. L. Barbour, "Imaging of fluorescence in highly scattering media," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 810–822, Sep. 1997.

[13] H. Jiang, "Frequency-domain fluorescent diffusion tomography: A finite-element-based algorithm and simulations," *Appl. Opt.*, vol. 37, pp. 5337–5343, 1998.

[14] R. Roy and E. M. Sevick-Muraca, "A numerical study of gradient-based nonlinear optimization methods for contrast enhanced optical tomography," *Opt. Exp.*, vol. 9, pp. 49–65, 2001.

[15] M. A. O_Leary, D. A. Boas, X. D. Li, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.*, vol. 21, pp. 158–160, 1996.

[16] V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," *Opt. Lett.*, vol. 26, pp. 893–895, 2001.

[17] J. Lee and E. M. Sevick-Muraca, "Three-dimensional fluorescence enhanced optical tomography using referenced frequency-domain photon migration measurements at emission and excitation wavelengths," *J. Opt. Soc. Amer. A*, vol. 19, pp. 759–771, 2002.

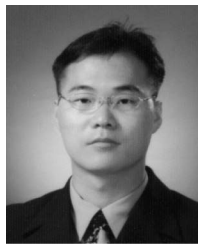
[18] M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. SevickMuraca, "Three-dimensional, Bayesian image reconstruction from sparse and noisy data sets: Near-infrared fluorescence tomography," *Proc. Nat. Acad. Sci. USA*, vol. 99, pp. 9619–9624, 2002.

[19] 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.

[20] A. D. Kim and A. Ishimaru, "Optical diffusion of continuous-wave, pulsed, and density waves in scattering media and comparisons with radiative transfer," *Appl. Opt.*, vol. 37, pp. 5313–5319, 1998.

[21] R. Aronson and N. Corngold, "Photon diffusion coefficient in an absorbing medium," *J. Opt. Soc. Amer. A*, vol. 16, pp. 1066–1071, 1999.

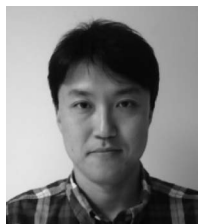
[22] B. Chen, K. Stamnes, and J. J. Stamnes, "Validity of the diffusion approximation in bio-optical imaging," *Appl. Opt.*, vol. 40, pp. 6356–6366, 2001.
[23] R. Elaloufi, R. Carminati, and J.-J. Greffet, "Time-dependent transport through scattering media: from radiative transfer to diffusion," *J. Opt. A*, vol. 4, pp. S103–S108, 2002.
[24] A. H. Hielscher, A. E. Alcouffe, and R. L. Barbour, "Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues," *Phys. Med. Biol.*, vol. 43, pp. 1285–1302, 1998.
[25] A. D. Klose and A. H. Hielscher, "Fluorescence tomography with simulated data based on the equation of radiative transfer," *Opt. Lett.*, vol. 28, pp. 1019–1021, 2003.
[26] A. D. Klose, V. Ntziachristos, and A. H. Hielscher, "The inverse source problem based on the equation of radiative transfer," *J. Comput. Phys.*, vol. 202, pp. 323–345, 2005.
[27] L. Biegler, O. Ghattas, M. Heinkenschloss, and B. Bloemen, *Large-Scale PDE-Constrained Optimization*. New York: Springer-Verlag, 2003.
[28] R. Byrd, F. Curtis, and J. Nocedal, "An inexact SQP method for equality constrained optimization," *SIAM J. Optim.*, vol. 19, pp. 351–369, 2008.
[29] P. Gill, W. Murray, and M. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization SIAM review," vol. 47, pp. 99–131, 2005.
[30] Biros and O. Ghattas, "Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver," *SIAM J. Sci. Comput.*, vol. 27, pp. 687–713, 2003.
[31] L. Biegler, C. Schmid, and D. Ternet, *A Multiplier-Free, Reduced Hessian Method For Process Optimization, Large-Scale Optimization with Applications, Part II: Optimal Design and Control*. New York: Springer-Verlag, 1997, p. 101.
[32] E. Haber and U. Ascher, "Preconditioned all-at-once methods for large, sparse parameter estimation problems," *Inverse Probl.*, vol. 17, pp. 1847–1864, 2001.
[33] J. Bonnas, J. Gilbert, C. Lemaréchal, and C. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*. New York: Springer-Verlag, 2003.
[34] M. Burger and W. Muhlhuber, "Iterative regularization of parameter identification problems by sequential quadratic programming methods," *Inverse Probl.*, vol. 18, pp. 943–969, 2002.
[35] P. Boggs and J. Tolle, "Sequential quadratic programming for large-scale nonlinear optimization," *J. Comput. Appl. Math.*, vol. 124, pp. 123–137, 2000.
[36] J. L. Hu, Z. Wu, H. McCann, L. E. Davis, and C. G. Xie, "Sequential quadratic programming method for solution of electromagnetic inverse problems," *IEEE Trans. Antennas Propag.*, vol. 53, no. 8, pp. 2680–2687, Aug. 2005.
[37] D. Feng and T. Pulliam "All at-at-once reduced Hessian SQP scheme for aerodynamics design optimization," Tech. Rep., NASA Ames Research Center, Mountain View, CA, 1995.
[38] T. Coleman, J. Liu, and W. Yuan, "A Quasi-Newton quadratic penalty method for minimization subject to nonlinear equality constraints," *Comput. Optim. Appl.*, vol. 15, pp. 103–123, 2000.
[39] M. Lalee, J. Nocedal, and T. Plantenga, "On the implementation of an algorithm for large-scale equality constrained optimization," *SIAM J. Optim.*, vol. 8, pp. 682–706, 2003.
[40] L. Biegler, J. Nocedal, C. Schmid, and D. Ternet, "Numerical experience with a reduced Hessian method for large scale constrained optimization," *Comput. Optim. Appl.*, vol. 15, pp. 45–67, 2000.
[41] G. S. Abddoulaev, K. Ren, and A. H. Hielscher, "Optical tomography as a PDE-constrained optimization problem," *Inverse Probl.*, vol. 21, pp. 1507–1530, 2005.
[42] H. K. Kim and A. H. Hielscher, "A PDE-constrained reduced Hessian SQP method for optical tomography based on the frequency-domain equation of radiative transfer," *Inverse Probl.*, vol. 25, pp. 015010-1–015010-20, 2009.
[43] W. Bangerth and A. Joshi, "Adaptive finite element methods for the solution of inverse problems in optical tomography," *Inverse Probl.*, vol. 24, pp. 034011-1–034011-22, 2008.
[44] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Non-contact fluorescence optical tomography with scanning patterned illumination," *Opt. Exp.*, vol. 14, pp. 6516–6534, 2006.
[45] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer-Verlag, 2006.
[46] A. D. Klose and A. H. Hielscher, "Quasi-Newton methods in optical tomographic image reconstruction," *Inverse Probl.*, vol. 19, pp. 387–409, 2003.
[47] R, K. Ren, G. Bal, and A. H. Hielscher, "Frequency domain optical tomography with the equation of radiative transfer," *SIAM J. Sci. Comput.*, vol. 28, pp. 1463–1489, 2006.
[48] H. K. Kim and A. Charette, "A sensitivity function-based conjugate gradient method for optical tomography with the frequency-domain equation of radiative transfer," *J. Quant. Spec. Rad. Trans.*, vol. 104, pp. 24–39, 2007.
[49] L. G. Henyey and L. J. Greenstein, "Diffuse radiation in the galaxy," *Astrophys*, vol. 90, pp. 70–83, 1941.
[50] M. Modest, *Radiative Heat Transfer*. New York: MacGraw-Hill, 2003.
[51] S. Minkowycz, E. Sparrow, and J. Murthy, *Handbook of Numerical Heat Transfer*. Hoboken, NJ: Wiley, 2006.
[52] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: SIAM, 2003.
[53] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, vol. 3, pp. 856–869, 1986.
[54] O. M. Alifanov, *Inverse Heat Transfer Problems*. New York: Springer-Verlag, 1994.
[55] U. J. Netz, J. Beuthan, and A. H. Hielscher, "Multipixel system for gigahertz frequency-domain optical imaging of finger joints," *Rev. Sci. Instrum.*, vol. 79, no. 8, pp. 034301-1–034301-14, 2008.



Hyun Keol Kim received the B.S. degree from Seoul National University, Seoul, Korea, in 1999, and the M. S. and Ph.D. degrees in mechanical engineering from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2001 and 2004, respectively.

In 2005, he was with the Department of Applied Sciences, Quebec University at Chicoutimi, QC, Canada, as a Postdoctoral Research Fellow. In 2006, he joined the Department of Biomedical Engineering, Columbia University, New York, NY, where he spent three years as a Postdoctoral Research Scientist and is currently an Associate Research Scientist with the Biophotonics and Optical Radiology Laboratory. His current research interests include the development of state-of-the-art imaging software for optical tomography including bioluminescence and fluorescence tomography.

Dr. Kim was awarded the Oversea Postdoctoral Fellowship from Korea Research Foundation in 2005.



Jong Hwan Lee received the B.S. degree in electrical engineering from Kyungpook National University, Daegu, Korea, in 1999, and the M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 2006. He is currently working toward the Ph.D. degree with the Department of Biomedical Engineering, Columbia University, New York, NY.

His research interests include biomedical imaging and instrumentation. His current research focuses on developing molecular imaging and optical-tomography-imaging systems.

Andreas H. Hielscher (S'93–M'95) received the Ph.D. degree in electrical and computer engineering from Rice University, Houston, TX, in 1995.

After spending two years as a Postdoctoral Fellow with the Los Alamos National Laboratory, Los Alamos, NM, he joined the faculty at the State University of New York Downstate Medical Center, Stony Brook, NY. In September 2001, he moved to Columbia University, New York, NY, where he is currently the Director of the Biophotonics and Optical Radiology Laboratory. He holds joint appointments as an Associate Professor with the Departments of Biomedical Engineering and Radiology, Columbia University. His work focuses on the development of state-of-the-art imaging software and hardware for optical tomography. He applies this emerging technology to the imaging of cancer and joint diseases and uses it in support of drug development. He has authored or coauthored more than 120 articles in peer-reviewed scientific journals and conference proceedings.

Dr. Hielscher is currently an Associate Editor of the IEEE TRANSACTIONS OF MEDICAL IMAGING.