4.17 Inverse Models for Diffuse Optical Molecular Tomography

HK Kim, Columbia University College of Physicians & Surgeons, New York, NY, USA
AH Hielscher, Columbia University, New York, NY, USA

Abbreviations

BLT Bioluminescence tomography
ERT Equation of radiative transfer
FMT Fluorescence molecular tomography

4.17.1 Introduction

Optical molecular tomography is an imaging technique that makes use of light-emitting biomarkers as the source of contrast. This technique provides a three-dimensional visualization of these biomarkers inside the tissue in a noninvasive manner and can thus be used to image particular target molecules or their pathways associated with the molecular process or the development of diseases inside living organisms (Bremer et al., 2001; Contag, 2002; Galbán et al., 2010; Narsinh et al., 2009; Weissleder and Ntziachristos, 2003). Depending on the biomarker used as a contrast agent, the method can be loosely divided into two different modalities: fluorescence molecular tomography (FMT) and bioluminescence tomography (BLT). FMT uses fluorescent biomarkers called fluorophores that absorb light and reemit it at longer wavelength and has been mainly applied to preclinical studies to target specific molecules that provide functional information about biochemical processes (Ale et al., 2012; Chang et al., 1997; Chen et al., 2003; Corlu et al., 2007; Deliolanis et al., 2009; Li et al., 2012; Ntziachristos et al., 2002). BLT also makes use of light-emitting probes called luciferases that emit light when certain biochemical environments are encountered and has been widely used in preclinical setting with small animals to study various disease processes and drug effects (Choy et al, 2003; Contag, 2002, 2010; Eslami et al., 2012; Hu et al., 2012). Using these two modalities (FMT and BLT), it is thus possible to detect diseases on a molecular level before typical symptoms or macroscopic changes appear.

Tomographic images of these molecular probes are obtained through the so-called model-based image reconstruction schemes (Arridge, 1999; Hielscher and Klose, 1999) in which a forward model of light propagation is employed that leads to predictions of measured values on the boundary, assuming a certain distribution of all unknown parameters inside the medium. An objective function is defined as that quantifies the differences between predicted and actually measured values. Image reconstructions are performed iteratively to find the minimum of this objective function, by updating the parameters of the forward model. Most image reconstruction methods used for FMT and BLT rely on the validity of the diffusion approximation (DA) to the equation of radiative transfer (ERT) (Chen et al., 2001; Elaloufi et al., 2002; Kim and Ishimaru, 1998). The DA converts the difficult-to-solve integrodifferential ERT into a partial-differential diffusion equation (DE) for which various analytical solution and stable numerical methods exist.

However, it is well known that the DA becomes inaccurate if the absorption coefficient of the medium under consideration is high and becomes comparable in magnitude to the scattering coefficient. Furthermore, the approximation is inaccurate in media with small geometries where boundary effects are dominant or in fluid-filled, void-like regions (Hielscher et al., 1998). All these conditions hold true for small animal imaging for which fluorescence and bioluminescence tomographic techniques are mainly used. For example, luciferases most often employed as a light-emitting probe have a broad light spectrum with an emission peak between 460 and 630 nm (Zhao et al., 2009). In this wavelength range, the intrinsic tissue absorption is relatively high (Cheong et al., 1990). Furthermore, the optical path length in small animals such as mice and rats is relatively small and boundary effects are dominant. Hence, it is highly desirable to use ERT-based image reconstruction schemes for small-animal FMT and BLT. First, ERT-based FMT codes have been reported (Klose and Hielscher, 2003a; Klose et al., 2005). Gao and Klose have presented first ERT-based BLT schemes (Gao and Zhao, 2010; Klose, 2007). Gao et al. improved a spatial resolution of the reconstructed bioluminescence image by exploiting the sparsity of bioluminescent sources with the l_1-regularized multilevel scheme. Klose et al. employed a stochastic iterative reconstruction method to find the global minimum of a bioluminescence inverse problem, which overcame premature convergence due to infeasible initial guess. However, these ERT-based reconstruction algorithms are computationally very costly since they require solving the forward problem of ERT repeatedly until convergence, thus taking anywhere from several hours to even days to solve one single image reconstruction case.

Recently, Kim and Hielscher introduced a novel algorithm called partial differential equations (PDE)-constrained sequential quadratic programming (SQP) to solve molecular tomographic problems in a computationally efficient manner (Kim et al., 2009; Kim et al., 2010). This new algorithm increases the speed of an image reconstruction process by a factor of up to 20 as compared to traditional methods. This increase in speed is made possible by solving the forward and inverse problems simultaneously in a framework of PDE-constrained optimization. In this article, the theoretical foundation and computational efficiency for this PDE-constrained algorithm is provided with comparison to traditional unconstrained codes that make use of the limited-memory Broyden–Fletcher–Goldfarb–Shanno (lm-BFGS) method (Nocedal and Wright, 2006). The lm-BFGS method is known to be the most efficient gradient-based unconstrained optimization method and therefore provides a good standard for comparison.

4.17.2 Fluorescence Molecular Tomography

4.17.2.1 Light Propagation Model with Fluorescence Tomography

In frequency-domain (FD) FMT, 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 FD equations of radiative transfer (Kim and Charette, 2007; Kim et al., 2010; Ren et al., 2006) as

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

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

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 i is $\sqrt{-1}$; ϖ is the angular frequency at which the amplitude of the external source is modulated modulation; c is the speed of light in the medium; 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 W cm^{-2} 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 eqn [2] denotes the excitation fluence at position $\mathbf{r}$ defined by $\phi(\mathbf{r},\varpi)=\int_{4\pi}\psi^x\mathrm{d}\Omega$ in unit W cm^{-2}. For a phase function denoted by $p(\Omega',\Omega)$, we use here the Henyey–Greenstein phase function (Henyey and Greenstein, 1941) that is commonly used in tissue optics. As shown in eqn [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 $\psi^x(\mathbf{r},\Omega,\varpi)$ typically at longer wavelength (see eqn [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 equations are given:

$$\Psi_b^x\big|_{\vec{n}b\cdot\Omega<0}=\psi^{x0}\big|_{\vec{n}b\cdot\Omega<0}+R\left(\Omega',\Omega\right)\cdot\psi^x|_{\vec{n}b\cdot\Omega'<0} \quad [3]$$

$$\psi_b^m|_{\vec{n}b\cdot\Omega<0}=R\left(\Omega',\Omega\right)\cdot\psi^m|_{\vec{n}b\cdot\Omega'<0} \quad [4]$$

where $R(\Omega',\Omega)$ is the reflectivity at Fresnel interface (Modest, 2003) 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 eqns [1] and [2], one can 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 (Minkowycz et al., 2006).

Following an unstructured finite-volume discrete-ordinate method (Kim et al., 2010; Montejo et al., 2010), the discretized forms of the two ERTs given by eqns [1] and [2] are obtained by integrating eqns [1] and [2] over the control volume with a divergence theorem as

$$\sum_{j=1}^{N_{\text{surf}}}\left(\vec{n}_j\cdot\Omega^l\right)\psi_j^{x,\prime\,l}\mathrm{d}A_j+\left(\mu_a^x+\mu_s^x+\mu_a^{x\to m}+\frac{i\varpi}{c}\right)\psi_n^{x,\prime\,l}=\frac{\mu_s^x}{4\pi}\delta V_N\sum_{l'=1}^{N_\Omega}\psi_N^{x,l'}\,p^{\prime l}w^{l'} \quad [5]$$

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

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 $\psi_N^{x,l}$ 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 (Minkowycz et al., 2006). It can be easily seen that each ERT involves Nt spatial unknown intensities coupled into N_Ω directions, thus leading to the total $N_t\cdot N_\Omega$ unknowns. The system given by eqns [1] and [2] can be solved by using any iterative solvers as far as they are reliable. One can employ 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, one finally obtains two linear systems of algebraic equations as

$$A^x\psi^x=b^x \quad \text{and} \quad A^m\psi^m=b^m \quad [7]$$

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

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

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

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

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

A direct treatment of the FD equations of radiative transfer leads to the matrix formulations given by eqn [7] that contain complex-valued elements. As a result, the complex-valued algebraic linear equations can be solved with a complex version of the GMRES linear solver (Saad, 2003; Saad and Schultz, 1986).

The two equations given by eqns[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 inverse model that is used to obtain the spatial distribution of a fluorescent source inside the medium that best fits the measured emission data.

4.17.2.2 Fluorescence Tomographic Algorithms

4.17.2.2.1 Traditional approach

The optical fluorescence tomographic problem can be formulated in more general terms as

$$\min f(x,u)\ \text{ subject to }\ c(x,u)=0 \quad [9]$$

where $x\epsilon R^n$ is the vector of inverse variables, $u\epsilon Z^m$ 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 eqn [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 PDEs.

Traditional method for solving eqn [9] is to treat the forward variable u as a dependent variable of the inverse variable x, that is, $u=\tilde{c}(x)$, which makes it possible to replace the prediction vector u in f of eqn [9] by its forward solution vector. As a result, 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 eqn [10], that is, $\tilde{f}$ is now a function of x only. Thus, the forward solution $u=\tilde{c}(x)$ has to be obtained for evaluation of $\tilde{f}$, that is, 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, with respect to both 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 belong to this approach: the conjugate gradient approach, the quasi-Newton (QN) approach, the Jacobian approach (Eppstein et al., 2002; Jiang, 1998; Klose and Hielscher, 2003a; Klose et al., 2005; Lee and Sevick-Muraca, 2002; Milstein et al., 2003; O'Leary et al., 1996; Paithankar et al., 1997; Roy and Sevick-Muraca, 2001), and so on.

4.17.2.2.2 PDE-constrained approach

Another approach to solve eqn [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 (Abddoulaev et al., 2005, Bangerth and Joshi, 2008; Kim and Hielscher, 2009):

$$L(x,u;\lambda)=f+\lambda^{\mathrm{T}}c \quad [11]$$

Here, $\lambda\epsilon Z^m$ 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 where the gradient of L in eqn [11] vanishes with respect to 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 (Nocedal and Wright, 2006). 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 (Biegler et al., 2000; Biros and Ghattas, 2003, Byrd et al., 2008; Gill et al., 2005; Lalee et al., 2003).

Following this approach, the fluorescence tomographic problem given by eqn [9] can be rewritten into a problem of finding a vector of unknowns $\mu=(\mu_a^{x\to m},\tau,\eta)$, 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

$$\min\ f(\psi^m)=\frac{1}{2}\sum_{s,d}\left|Q_d\psi_s^m-z_{s,d}\right|^2 \qquad [12]$$

Subject to

$$C_s^x=A^x\psi_s^x-b_s^x=0; s=1,\ \ldots, N_s$$
$$C_s^m=A^m\psi_s^m-b_s^m=0; s=1,\ \ldots, N_s$$

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 and C^x and C^m are the discretized versions of the two coupled radiative transfer equations.

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

$$L(\mu;\psi^x,\psi^m;\lambda^x,\lambda^m)=\frac{1}{2}\left|Q\psi^m-z^2\right|+\lambda^{x\mathrm{T}}(A^x\psi^x-b^x)+\lambda^{m\mathrm{T}}(A^m\psi^m-b^m) \qquad [13]$$

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

$$L_\mu\equiv\frac{\partial L}{\partial\mu}=\left(\psi^{x\mathrm{T}}A^{x\mathrm{T}}\right)_\mu\lambda^x-\left(b^{m\mathrm{T}}\right)_\mu\lambda^m=0 \qquad [14a]$$

$$L_{\psi x}\equiv\frac{\partial L}{\partial\psi^x}=A^{x\mathrm{T}}\lambda^x-\left(b^{m\mathrm{T}}\right)_{\psi x}\lambda^m=0 \qquad [14b]$$

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

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

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

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 $\lambda^x$and λ^m, and the last two equations represent the two forward equations given by eqn [7].

The Karush–Kuhn–Tucker (KKT) system given by eqn [14] can be solved with Newton's method as

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

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 $[\Delta\mu,\Delta\psi^x,\Delta\psi^m]^{\mathrm{T}}$ and $[\Delta\lambda^x,\Delta\lambda^m]^{\mathrm{T}}$, respectively. The algebraic system given by eqn [16] can then be solved efficiently through the reduced Hessian SQP (rSQP) scheme as will be shown in the succeeding text.

4.17.2.2.3 Reduced Hessian sequential quadratic programming

The rSQP method is an established method that solves nonlinear optimization problems with relatively low cost and fast convergence (Byrd et al., 2008). Employing this rSQP method for solving the KKT system [15] in the preceding text is equivalent to finding the minimum to a quadratic approximation of the Lagrangian function L subject to the linearized constraints C^x and cm, which gives the following quadratic programming problem so that

$$\begin{aligned}&\min\ \ \Delta p^{\mathrm{T}}g_p+\frac{1}{2}\Delta p^{\mathrm{T}}W\Delta p\\ &\text{subject to } A\Delta p+c=0\end{aligned} \qquad [16]$$

where $g=\nabla f=[g_\mu g_{\psi x} g_{\psi m}]$ denotes the gradients of the objective function J with respect to each of unknowns (μ,ψ^x,ψ^m) and W^k is the full Hessian (or approximations) of the Lagrangian function. Here, 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. Following the standard rSQP method as described in the literature (Kim et al., 2010), the unknown vector Δp is given as follows:

$$\Delta\mu=-(H_r)^{-1}(g_r+d_r) \qquad [17]$$

$$\Delta\psi=\begin{bmatrix}\Delta\psi^x\\ \Delta\psi^m\end{bmatrix}=Z-Y(H_r)^{-1}(g_r+d_r) \qquad [18]$$

where $H_r=Y^{\mathrm{T}}WY$ denotes the reduced Hessian, $g_r=Y^{\mathrm{T}}g$ denotes the reduced gradient, and

$$Z=\begin{bmatrix}0\\ -(A^x)^{-1}C^x\\ -(A^m)^{-1}\left\{C^m+(b^m)_{\psi^x}(A^x)^{-1}C^x\right\}\end{bmatrix}\ \text{and}$$
$$Y=\begin{bmatrix}I\\ -(A^x)^{-1}(A^x\psi^x)_\mu\\ -(A^m)^{-1}\left\{(b^m)_\mu-(b^m)_{\psi^x}(A^x)^{-1}(A^x\psi^x)_\mu\right\}\end{bmatrix} \qquad [19]$$

following the popular choices for Z^k and Y^k (Gill et al., 2005; Kim and Hielscher, 2009; Kim et al., 2010; Nocedal and Wright, 2006).

Thus, the rSQP method requires much less memory than the full SQP one, that is, only a small $(n\times n)$ matrix needs to be

maintained and updated at each of optimization iterations. For large-scale applications, it is desirable to avoid the direct computation of the reduced Hessian Hr and its matrix inversion $(H_r^k)^{-1}$. Accordingly, the matrix–vector product of $(H_r^k)^{-1}g_r^k$ can be approximated directly by using the limited-memory updating formula (Kim and Hielscher, 2009; Klose and Hielscher, 2003a; Nocedal and Wright, 2006).

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

$$\varphi_\eta(\mu, \psi^x, \psi^m) = f(\psi^m) + \eta\|C\|_1 \qquad [20]$$

which is, at the new iterate given by $p^{k+1} = p^k + \alpha^k \Delta p$, successively monitored to ensure global progress toward the optimal solution during line search (Biegler et al., 2000). Refer to literature (Kim and Hielscher, 2009; Kim et al., 2010) for details on implementation of this rSQP technique.

4.17.2.3 Examples of Fluorescence Reconstructions

In the following, we will present some illustrative examples of fluorescence reconstructions. The examples include numerical and experimental results of a comparison between PDE-constrained and unconstrained algorithms, in terms of accuracy, computational efficiency, and effects of noise and initial guess on the reconstruction.

4.17.2.3.1 CPU times, accuracy, and effects of noise and initial guess

A numerical phantom with a diameter of 2 cm as shown in Figure 1 is used to mimic a typical fluorescence tomographic problem with small animal imaging. In the first case, a fluorescent heterogeneity with a diameter of 0.15 cm is embedded inside the cylinder at $\Gamma = \{(x,y)|(x-0.35)^2 + (y-0.35)^2 = 0.2^2\}$. 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 $\mu_a^{x\to m} = 00.5\,\text{cm}^{-1}$. 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 Figure 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 $\Gamma = \{(x,y)|x^2 + y^2 = 1\}$.

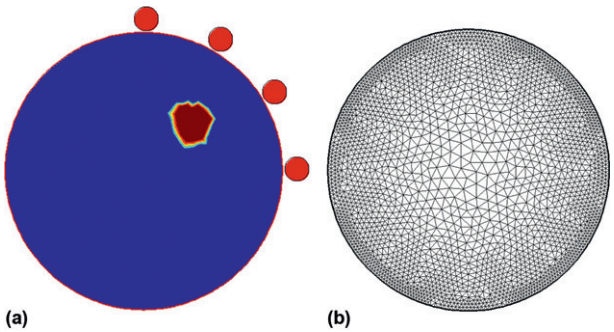


Figure 1 The schematic of test problems 1 and 2: circle with diameter of 2 cm. (a) Source–detector configuration: 4 sources (●) and 66 detectors around the surface and (b) computation domain with 4886 triangular elements.

Figure 2 shows the reconstruction results obtained with the PDE-constrained rSQP and unconstrained lm-BFGS algorithms for different levels of noise, and CPU times and accuracy are given in Table 1 for their respective case. The correlation factor $\rho(\mu^e, \mu^r)$ and the deviation factor $\delta(\mu^e, \mu^r)$ are used as the metrics of image quality as defined in the succeeding text (Kim and Hielscher, 2009; Klose and Hielscher, 2003a):

$$\rho = \frac{\sum_{i=1}^{N_t}(\mu_i^e - \bar{\mu}_i^e)(\mu_i^r - \bar{\mu}_i^r)}{(N_t - 1)\sigma(\mu_e)\sigma(\mu_r)}, \quad \delta = \frac{\sqrt{\sum_{i=1}^{N_t}(\mu_i^e - \mu_i^r)^2 N_t}}{\sigma(\mu^e)} \qquad [21]$$

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. As shown in Figure 2, the circular fluorescence perturbation as shown in Figure 1(a) is well identified by the two methods.

Table 1 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

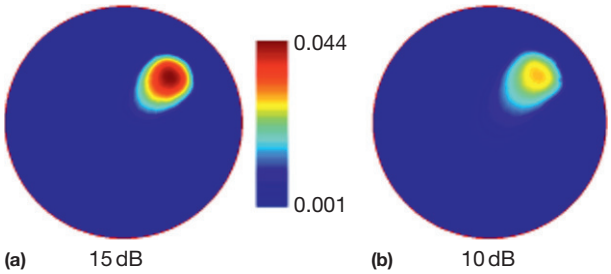


Figure 2 Reconstructed fluorescence absorption coefficients $\mu_a^{x\to m}$ obtained for the first example using the 15 and 10 dB noise data by the PDE-constrained rSQP method.

Table 1 Reconstruction quality and computation times with different noise levels

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

[a]Denotes the acceleration factor by the PDE-constrained method.

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 1). 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 of optimization iterations until it converges to the optimal solution. Indeed, the PDE-constrained method utilizes the incomplete solution of the two linearized forward equations, for example, with the loose tolerance of 10^{-2}–10^{-3}. Figure 3 illustrates this convergence behavior of the PDE-constrained method. It can be clearly seen that the forward and inverse solutions converge simultaneously toward their optimal solutions, even when a loose tolerance of 10^{-2} is used. For the reconstruction accuracy, Table 1 shows that at noise levels of 10–15 dB, the constrained and unconstrained methods make no significant difference in the accuracy.

In FMT, 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.

Table 2 and Figure 4 show the results of effects of initial guess on the fluorescence reconstruction for the two different algorithms. It can be seen from Table 2 that the PDE-constrained rSQP and unconstrained lm-BFGS method lead to similar reconstruction results. For the 20% case, $\rho(\mu^e, \mu^r) = 0.77$ and $\delta(\mu^e, \mu^r) = 0.81$ are observed by 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 10–12 times faster than the unconstrained code.

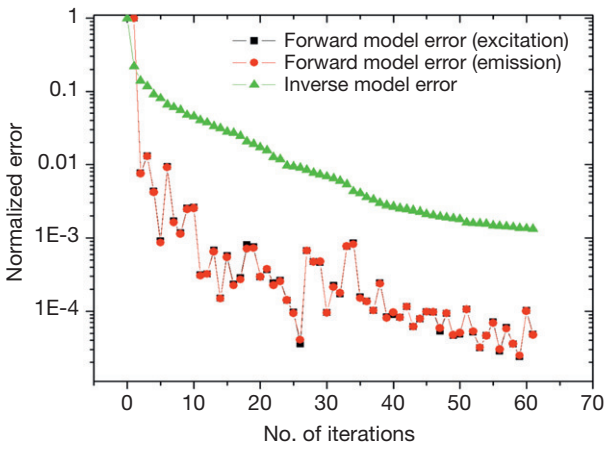


Figure 3 Convergence history of the PDE-constrained method in the forward and inverse solutions when a loose tolerance of 0.01 is used for solving the QP problems.

4.17.2.3.2 FD data versus steady-state data

FD data are considered superior to steady-state (SS) data since it provides more information. For example, FD carries a phase shift in addition to amplitude, which improves the reconstruction. A lab phantom with a square base is used here that has a size of 2.2 cm × 2.2 cm × 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 $\mu_a^{x \to m} = 0.4\ \text{cm}^{-1}, \tau = 4.0\,\text{ns}$, and $\eta = 0.93$, respectively. The fluorescent rod is positioned about 6 mm off-center and the background medium is filled only with a 2% Intralipid solution. 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. Measurements were made with the CCD camera both for SS and FD (at 150 MHz) data. Figure 5 shows the reconstruction results. As expected, the 150 MHz data give the more accurate reconstruction in the location of the fluorescent probe as compared to the SS data.

4.17.3 Bioluminescence Tomography

4.17.3.1 ERT-Based Light Propagation Model with BLT

The forward problem for light propagation in turbid media with a bioluminescent source in each spectral band ν can be accurately modeled by the ERT, given by (Kim and Hielscher, 2009; Klose, 2007)

$$(\nabla.\Omega)\psi_\nu(\mathbf{r},\Omega) + (\mu_a + \mu_s)\psi_\nu(\mathbf{r}.\Omega) = \frac{\mu_s}{4\pi}\int_{4\pi}\psi_\nu\left(\mathbf{r},\Omega'\right)p\left(\Omega',\Omega\right)d\Omega' + \frac{q_\nu(\mathbf{r})}{4\pi} \quad [22]$$

Table 2 Reconstruction quality and computation times with different initial guesses

$\mu_a^{x \to m,0}$	*Schemes*	*CPU time ([a])*	*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

[a]Denotes the acceleration factor by the PDE-constrained method.

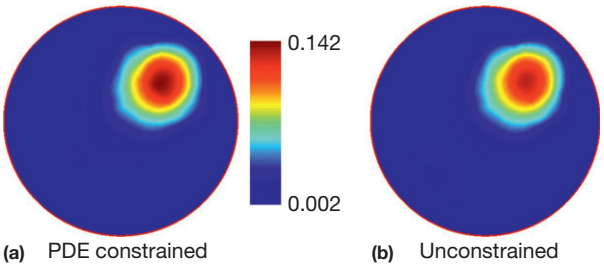


Figure 4 Reconstructed fluorescence absorption coefficients $\mu_a^{x \to m}$ obtained for the second example using the initial guess of $\mu_a^{x \to m} = 0.1$.

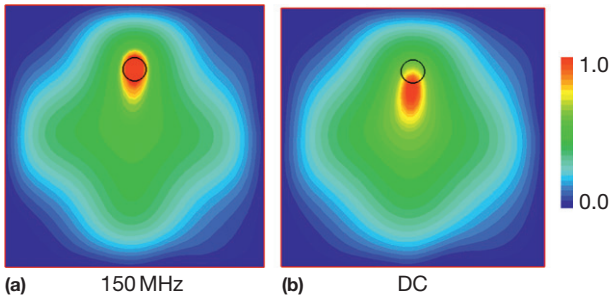


Figure 5 Reconstructed maps of fluorophore absorption coefficients $\mu_a^{x\to m}$ inside the phantom using the DC and 150 MHz data.

where ψ_v $(\mathbf{r},\Omega)$ is the spectral radiation intensity in unit (W cm^{-2} sr^{-1}), μ_a and μ_s are the absorption and scattering coefficients, respectively, in units of (cm^{-1}), $p(\Omega',\Omega)$ is the Henyey–Greenstein scattering phase function (Henyey and Greenstein, 1941) that describes scattering from incoming direction into scattering direction Ω. The source term $q_v(\mathbf{r})$ is the photon emission in the spectral band v of the total photon emission $q(\mathbf{r}) = \sum_v q_v(\mathbf{r})$ where $q_v(\mathbf{r}) = w_v q(\mathbf{r})$ where each weight w_v can be determined from a spectral analysis over the spectral energy distribution of a bioluminescent source. Thus, the total energy distribution $q(\mathbf{r})$ of a bioluminescent source is not a function of wavelength v, so we can use as many wavelength data as possible simultaneously into reconstruction as will be shown later. To be able to consider the refractive index mismatch at the air–tissue interface, one can employ the partially reflective boundary condition:

$$\psi_{v,b}\Big|_{\vec{n}_b\cdot\Omega<0} = R\left(\Omega',\Omega\right).\psi_{vb}\Big|_{\vec{n}_b\cdot\Omega'<0} \qquad [23]$$

where $R(\Omega',\Omega)$ is the reflectivity at Fresnel interface (Modest, 2003) from direction Ω' to direction Ω, ψ_{vb} is the spectral radiation intensity due to the external source function, and subscript $\vec{n}_b$ denotes the boundary surface of the medium, while m is the unit normal vector pointing outward the boundary surface.

For discretization of eqn [22], we use an unstructured node-centered finite-volume method in the spatial domain and a discrete-ordinate method in the angular domain. Following unstructured finite-volume discrete-ordinate methods (Kim et al., 2010; Montejo et al., 2010), the discretized form of eqn [22] is obtained by integrating eqn [22] over the control volume and a divergence theorem as

$$\sum_{j=1}^{N_{\text{surf}}}\left(\vec{n}_j\cdot\Omega^l\right)\psi^l_{v,j}\mathrm{d}A_j + \left(\mu_a^v+\mu_s^v\right)\psi^l{}_{v,N}$$
$$= \frac{\mu_s^v}{4\pi}\Delta V_N\sum_{l'=1}^{N_\Omega}\psi^{l'}_{v,N}p^{l'l}w^{l'} + \frac{q_v}{4\pi}\Delta V_N \qquad [24]$$

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 $\psi^l_{v,j}$ denotes the spectral radiances defined on the jth surface in direction l. Also, the surface intensities $\psi^l_{v,j}$ are related to the nodal intensities $\psi^l_{v,N}$ by the second-order spatial differencing scheme (Minkowycz et al., 2006).

4.17.3.2 Bioluminescence Tomographic Algorithms

4.17.3.2.1 Traditional approach

The most common approach to the inverse bioluminescence problem is to exploit the linear relationship between bioluminescent sources inside the medium and measurements on the tissue surface. Given the measurement z at the tissue surface, the prediction p can be represented by a linear combination of the unknown source distribution q and its Jacobian coefficient J, that is, $p=Jq$ in a matrix–vector form. As a result, the inverse bioluminescence problem reduces to solving the linear equation $Jq=z$ for the unknown vector q, and the inverse solution is obtained in a least squares sense, that is, $\min_q \frac{1}{2}\|Jq-z\|^2$, since the number m of measurements available is usually much smaller $\min_q$ than the number n of unknown sources to be estimated. Mathematically, this least squares minimization problem is, however, highly ill-posed, meaning the inverse solution is not unique and very sensitive to random errors in the measured data. This can be overcome by reformulating the originally ill-posed problem into an approximate, well-posed problem, such as $\min_q \frac{1}{2}\|Jq-z\|^2+\beta R(q)$ where $R(q)$ is a stabilizing term often represented by a classical l_1 or l_2 regularization term (Dehghani et al., 2006; Gao and Zhao, 2010; Gu et al., 2004) with a regularization parameter β. Nonlinear iterative methods are often used to solve this regularized minimization problem because in general a good choice of β is not known a priori and needs to be adjusted iteratively.

Thus, all existing BLT methods are based on the Jacobian matrix that requires a precalculation of the Jacobian coefficients before the reconstruction is started (Ahn et al., 2008; Alexandrakis et al., 2005; Chaudhari et al., 2005; Comsa et al., 2007; Cong et al., 2005; Dehghani et al., 2008; Kuo et al., 2007; Gu et al., 2004; Han et al., 2007; Klose et al., 2010; Kuo et al., 2007; Lv et al., 2007; Slavine et al., 2006; Zhang et al., 2008). The Jacobian matrix can be constructed in two different ways: the perturbation approach and the adjoint theorem. Given m measurements for n unknowns, the perturbation approach requires n forward runs, while the adjoint theorem only needs m adjoint runs $m\ll n$ (Dehghani et al., 2008). Obviously, the adjoint theorem is far more cost-efficient in this case. However, this Jacobian-based approach may lead to a highly cost-demanding process with respect to both time and memory for large-scale problems where a large number of measurements are used as input to reconstruction. This is true because in practice the CCD camera allows for several hundreds to thousands detection points and therefore the adjoint theorem will require several hundreds of forward runs only to construct J. Thus, commonly employed Jacobian methods appear to be too expensive to be applied to ERT-based image reconstruction codes since the ERT has to be solved accurately as many times as the number of detectors.

The fundamental reason for this requirement can be explained by the fact that Jacobian methods belong to a class called unconstrained. In the unconstrained approach, the inverse variable q is always treated as a function of the forward variable p, which makes it possible to eliminate the constraint $A\psi=b$ by substituting $\psi=A^{-1}b$ into $f(q,\psi)$. Then, the original constrained problem $\min f(q,\psi)$ s.t. $A\psi=b$ is now recast into the unconstrained problem $\min f(q, A^{-1}b(q))$, where f is now a

function of q only. All the existing BLT methods fall under this unconstrained category. For the inverse bioluminescence problem, unconstrained methods of this kind basically require solving the adjoint problem repeatedly as many times as detection points used. As a consequence, the reconstruction process becomes a very demanding task with respect to both time and memory. To overcome this shortcoming, gradient-based techniques often used in nonlinear optical tomographic problems can be employed in this case where the Jacobian is too expensive to obtain. Among these, QN methods (Nocedal and Wright, 2006) build up the approximate Hessian (i.e., $J^{\mathrm{T}}J$) using the gradient information through iterations, and in particular, their limited-memory version produces the product $(J^{\mathrm{T}}J+\beta R)^{-1}J^{\mathrm{T}}z$ directly, thus saving lots of time and memory. However, this QN method also requires the complete solution of two equations (the forward and adjoint equations) in order to construct the Hessian matrix iteratively, and additional function evaluations are needed as well during line search, which results in a very time-consuming process when the ERT is used as a light propagation model.

4.17.3.2.2 PDE-constrained multispectral approach

Given the predictions p_v of measurements z_v at wavelength z_v on the tissue surface, the multispectral inverse source problem is to find a vector $q(\mathbf{r})$ of spatial bioluminescent sources inside the medium that minimizes the difference between measurements and predictions as follows:

$$\min_{q,\psi v} f = \frac{1}{2}\sum_{v}(Q\psi_v - z_v)^{\mathrm{T}}(Q\psi_v - z_v) + \beta R(q)$$
$$\text{s.t. } C_v = A\psi_v - b_v = 0, \qquad [25]$$

where $A\psi_v = b_v$ denotes the discretized forward equation and Q denotes the measurement operator that projects the vector of spectral radiances onto the measureable quantity $p_{v,d} = Q_d\psi_{v,d}$ on the tissue surface. Equation [25] shows that all wavelength data can be used simultaneously into the source reconstruction q. The constrained problem [25] can be solved directly in the framework of PDE-constrained optimization where the forward and inverse variables are treated independently (Kim and Hielscher, 2009; Nocedal and Wright, 2006).

Typically, the PDE-constrained inverse problem [25] can be reformulated into the framework of the following extended objective function called 'Lagrangian' as

$$L(q;\psi_v,\lambda_v) = \frac{1}{2}|Q\psi_v - z|^2 + \beta R(q) + \lambda_v^{\mathrm{T}}(A\psi_v - b_v) \qquad [26]$$

Here, λ_v is called the vector of spectral Lagrange multipliers. The simultaneous solutions of forward and inverse problems can be achieved at points satisfying the so-called first-order KKT conditions where the gradient of L in eqn [26] vanishes with respect to q, ψ_v, and λ_v, respectively:

$$L_q \equiv \frac{\partial L}{\partial q} = \nabla f_q + \beta\nabla R + (b_v)_q\lambda_v = 0 \qquad [27a]$$

$$L_{\psi_v} \equiv \frac{\partial L}{\partial \psi_v} = \nabla f_{\psi_v} + A^{\mathrm{T}}\lambda_v = 0 \qquad [27b]$$

$$L_{\lambda_v} \equiv \frac{\partial L}{\partial \lambda_v} = A\psi_v - b_v = 0, \qquad [27c]$$

where the first eqn [27a] is the sensitivity equation and the second [27b] and third [27c] equations are the discretized adjoint and forward equations, respectively. As in the PDE-constrained approach with FMT, the rSQP method can be used to solve the original multispectral problem [25] by minimizing the multispectral Lagrangian functional eqn [26] with respect to q, ψ_v, and λ_v, respectively.

4.17.3.2.3 Reduced Hessian sequential quadratic programming

Let $x_k = (q,\psi)$ be the current iterate and $C_{k,x}$ denote $\frac{\partial C_k}{\partial x_k}$. The rSQP method solves the quadratic approximation of the Lagrangian functional equation [26] in the reduced space, subject to the linearization of the forward equations as follows (Kim and Hielscher, 2009; Nocedal and Wright, 2006):

$$\min \; g_{r,k}^{\mathrm{T}}\Delta q_k + \frac{1}{2}\Delta q^{\mathrm{T}}H_{r,qq}\Delta q_k$$
$$\text{subject to } (C_{k,x})\Delta x_k + C_k = 0 \qquad [28]$$

where $\Delta x = (\Delta q,\Delta\psi)^{\mathrm{T}}$ is the step for the next iterate, $g_{r,k}$ is the reduced gradient, and $H_{r,qq}\ (q_k)$ is the reduced Hessian of the Lagrangian function. After solving the quadratic problem [28], one finally can obtain the following expressions for the updates $\Delta x = (\Delta q,\Delta\psi)^{\mathrm{T}}$ of the forward and inverse variables as

$$\Delta q_k = -(H_{r,k})^{-1}g_{r,k} \qquad [29]$$

$$\Delta\psi_k = A^{-1}\left(C_k + C_{k,q}\Delta q_k\right) \qquad [30]$$

where the matrix–vector product $(H_{r,k})^{-1}\ g_{r,k}$ corresponds to the term $(J^{\mathrm{T}}J)^{-1}\ J^{\mathrm{T}}z$ in the unconstrained Jacobian method. QN methods can be used to build up the approximate Hessian $H_{r,k}$ (i.e., $J^{\mathrm{T}}J$) iteratively through previous gradient calculations, and in particular, the limited-memory BFGS updating scheme is employed here to directly produce the product $(H_{r,k})^{-1}\ g_{r,k}$ thus saving time and memory.

The global convergence of the PDE-constrained multispectral BLT algorithm is ensured by a line search on the following merit function that somehow balances the aims of reducing the objective function and satisfying the constraints, given as (Biegler et al., 2000; Byrd et al., 2008; Kim et al., 2010)

$$\varphi(\psi_v,q) = f(\psi_v,q) + \sum_{v}\eta_v|A\psi_v - b_v|_1 \qquad [30]$$

which is successively monitored to ensure the global progress toward the optimal forward and inverse solutions, while a line search is performed to find a step length α^k that can provide a sufficient decrease in the merit function. As a consequence, the new iterate is given as

$$q^{k+1} = q^k + \alpha^k\Delta q^k$$
$$\psi_v^{k+1} = \psi_v^k + \alpha^k\Delta\psi_v^k \qquad [31]$$

4.17.3.3 Examples of Bioluminescence Reconstructions

In the following, we will present some of the examples on fluorescence reconstructions. The results include numerical and experimental studies of CPU times, accuracy, and impacts of random noise and different initial guesses on the image reconstruction.

4.17.3.3.1 Accuracy, CPU times, and effects of noise and initial guess

The numerical phantom has a circular shape with a diameter of 2 cm, defined by $\Gamma = \{(x,y)|x^2+y^2=1\}$ and one single bioluminescent source with a diameter of 0.2 cm is embedded inside the medium defined by $\Gamma = \{(x,y)|(x-0.35)^2+(y-0.35)^2=0.2^2\}$. The bioluminescent source has a power density of 1 W cm^{-3}, which is to be found by the image reconstruction code. For numerical experiments, we considered spectrally resolved data. The chosen set of three wavelengths is 600, 650, and 700 nm, which are linearly independent of and distinguished from each other (Lu et al., 2009). The optical properties of the background medium are assumed to be $\mu_a = 1.5\,\text{cm}^{-1}, \mu_s' = 16.7\text{cm}^{-1}$ at 600 nm, $\mu_a = 0.5\,\text{cm}^{-1}, \mu_s' = 15.4\ \text{cm}^{-1}$ at 650 nm, and $\mu_a = 0.2\,\text{cm}^{-1}, \mu_s' = 14.3\ \text{cm}^{-1}$ at 700 nm. The 72 detector positions are spaced equally around the surface.

First, the effects of noise are examined on the performance of the two methods. To this end, two different SNR values (15 and 10 dB) that represent typical noise levels are used together with noise-free data. The CPU time and the deviation $\delta\epsilon\,[0,\infty]$ and correlation $\rho\epsilon\,[-1,1]$ factors are measured and given in Table 3. Figure 6 shows the maps of the reconstructed bioluminescent sources obtained for the 15 and 10 dB cases. For noise-free data, both PDE-constrained rSQP and unconstrained Jacobian methods give very similar accuracy: $\rho^{\text{SQP}}=0.71$, $\delta^{\text{SQP}}=0.92$ and $\rho^{\text{Jacobian}}=0.71$, $\delta^{\text{jacobian}}=0.92$. However, the PDE-constrained rSQP method converges faster than the unconstrained Jacobian method by a factor of about 14: the rSQP method took only 66 s to converge, while 907 s for the Jacobian method. Note that this CPU time (i.e., 66 s) by the rSQP method is equivalent to the computation effort of less than six forward runs, which in turn corresponds to the same computational effort taken to construct the Jacobian matrix for six detectors in the Jacobian method. As before, this speedup is achieved because the PDE-constrained code does not require the complete solution of a forward problem, while the inverse solution is inexact. In other words, the PDE-constrained algorithm solves the forward equation given by the RTE equation [22] with a loose tolerance in the 10^{-2}–10^{-3} range during the optimization process, which takes only a fraction of the time required for the exact solution of the forward problem with a tight tolerance ($\leq 10^{-10}$). This leads to lots of savings in the computation time as it goes toward the minimum. Figure 7 clearly shows this feature of the PDE-constrained rSQP method: both forward and inverse errors decrease in the PDE-constrained rSQP method since it solves both problems simultaneously (Figure 7(a)), which leads to the significant difference in the CPU times of the two methods (Figure 7(b)).

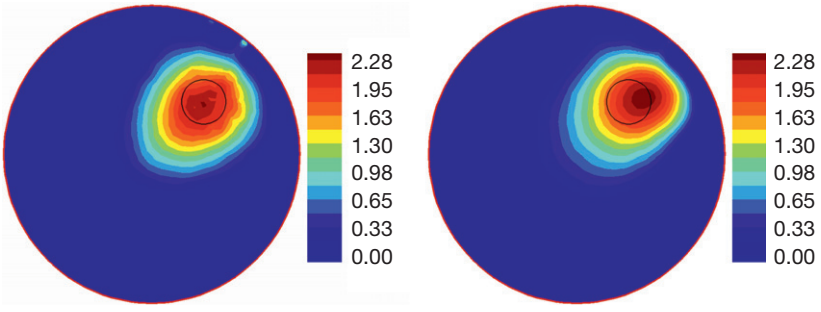
Table 3 Reconstruction quality and computation times with different noise levels

SNR	*Schemes PDE-*	*CPU time ([a])*	*Cor.* ρ	*Dev.* δ
∞	Constrained rSQP	66 s (14)	0.71	0.92
	Unconstrained Jacobian	907 s	0.71	0.92
15 dB	PDE-constrained rSQP	67 s (14)	0.74	0.90
	Unconstrained Jacobian	907 s	0.72	0.93
10 dB	PDE-constrained rSQP	79 s (12)	0.72	0.93
	Unconstrained Jacobian	920 s	0.67	0.93

[a]Denotes the acceleration factor by the PDE-constrained method.

At noise levels of 10–15 dB, the PDE-constrained rSQP method and the unconstrained Jacobian method give somewhat different results; for the 15 dB data, $\rho^{\text{SQP}}=0.74$, $\delta^{\text{SQP}}=0.90$ and $\rho^{\text{jacobian}}=0.72$, $\delta^{\text{Jacobian}}=0.93$, and for the 10 dB data, $\rho^{\text{SQP}}=0.72$, $\delta^{\text{SQP}}=0.93$ and $\rho^{\text{Jacobian}}=0.67$, $\delta^{\text{jacobian}}=0.93$. It can thus be said that the PDE-constrained rSQP method is much less influenced by random noise in the measurement than the unconstrained Jacobian method where $\rho^{\text{jacobian}}=0.72$ for 15 dB and $\delta^{\text{jacobian}}=0.67$ for 10 dB. For the CPU times, Table 3 shows that the PDE-constrained rSQP method leads to a significant saving in the computation time both for the two noise levels. The PDE-constrained rSQP method took 67 s, while the unconstrained Jacobian method requires 907 s. The similar reduction is observed for the 10 dB data: the PDE-constrained rSQP method converged after 79 s, while the unconstrained Jacobian method reached convergence after 920 min, which is already a speedup factor of about 12.

Next, the influence of initial guess on the algorithm is discussed here since gradient-based techniques including the PDE-constrained approach require making an initial guess of the unknowns to start the reconstruction and hence different initial guesses may affect the reconstruction quality. For this study, the same setup of the problem as before is considered here, and three different initial guesses ($q^0=1.0$, 3.0, and 10.0) are made that correspond to 10%, 30%, and 100%, respectively, of the actual value of a bioluminescent source. Note that a fixed noise level of 15 dB is used for the reconstructions with three different initial guesses. The reconstruction results and



(a) PDE-constrained rSQ (b) Unconstrained Jacobian

Figure 6 Reconstructed maps of bioluminescent sources q obtained with the PDE-constrained rSQP and unconstrained Jacobian methods for the 10 dB noise data.

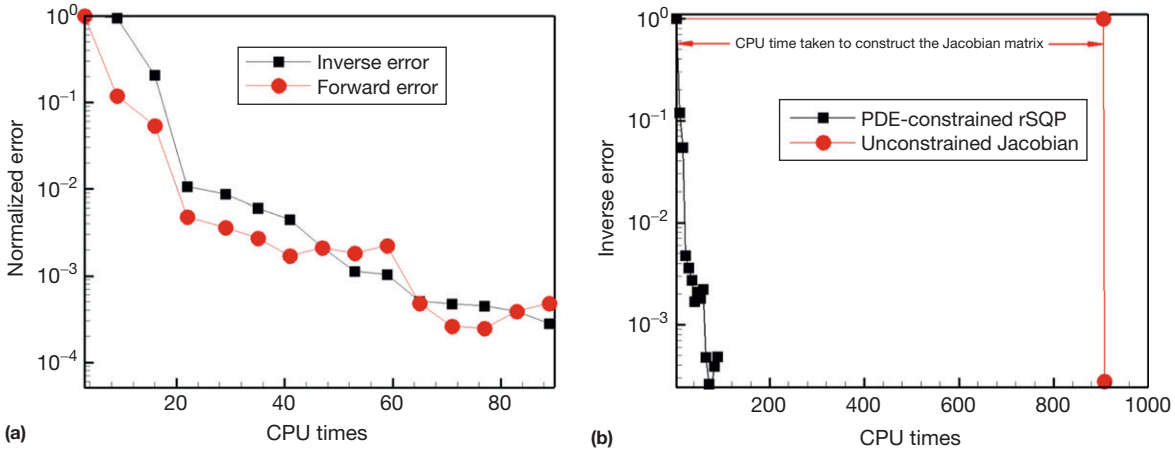


Figure 7 Objective function values of the two methods with respect to iteration numbers and CPU times.

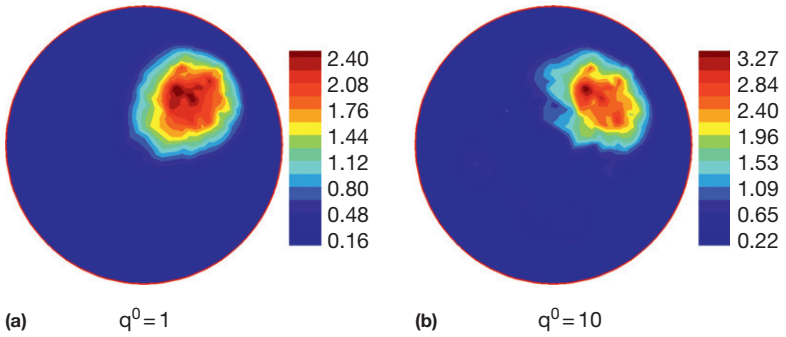


Figure 8 Reconstructed maps of bioluminescent sources obtained with the PDE-constrained rSQP method for different initial guesses of (a) $q^0 = 1.0$ and (b) $q^0 = 10.0$, respectively.

Table 4 Reconstruction quality and CPU times obtained with the PDE-constrained rSQP scheme for different initial guesses

q^0	*CPU time* [a]	*Cor.* ρ	*Dev.* δ
1	70 s (13)	0.71	0.92
3	65 s (14)	0.71	0.92
10	94 s (9.6)	0.69	0.90

[a]Denotes the acceleration factor by the PDE-constrained method.

measures of $\delta\epsilon\ [0,\infty]$ and correlation $\rho\epsilon[-1,1]$ are given in Figure 8 and Table 4, respectively. It can be seen from Table 4 that the PDE-constrained rSQP method leads to similar reconstruction results when the initial guesses are made as 10%, 30%, and 100% of the value of the target. For both of the 10% and 30% cases, the same accuracy is observed with $\rho^{SQP} = 0.71$ and $\delta^{SQP} = 0.92$, and for the 100% case, similar results are obtained with $\rho^{SQP} = 0.69$ and $\delta^{SQP} = 0.90$, although ρ^{SQP} is slightly decreased. In terms of CPU times, the PDE-constrained code took 70 s for the 10% case and 65 s for the 30% case, respectively, which are very similar to the time taken in the previous 15 dB case where the reconstruction is started with zero. Note that the 100% case took a little longer time, but this is already expected by the initial guess being too far away from the true solution.

4.17.3.3.2 Examples of experimental phantom

Experimental results of a quantum dot are presented here that allow for acquisition of spectral emission data over a broad range of wavelengths between 450 and 900 nm. With this spectral emission data alone, one seeks to find the spatial distribution of the quantum dot as a light-emitting source inside the medium. The square cylinder has a size of 25 mm × 25 mm × 100 mm ($X \times Y \times Z$) and is filled with 10% Intralipid fluid of known optical properties for all wavelengths considered here. A smaller cylinder rod with a diameter of 1.5 mm is filled with the quantum dot and embedded inside the background medium. Three sides of the phantom were imaged onto one CCD image and the 2 × 2 binning was done onto the CCD image, which leads to an image scale at which 1 pixel equals 0.167 mm. Figure 9 shows the reconstruction results of the spatial distribution of a quantum dot inside the lab phantom obtained by the two methods. The black hollow circle represents the actual location of the quantum dot embedded inside the medium. The results show that the PDE-constrained rSQP method gives very accurate localization and identification of the quantum dot, while the unconstrained Jacobian method gives similar results but seems to produce more artifacts. For CPU times, the unconstrained Jacobian method needed a total computation effort of about 2230 s taken for a precalculation of the Jacobian matrix using 123 forward runs, whereas the PDE-constrained method took only 103 s to converge, which yields a 22-fold speedup.

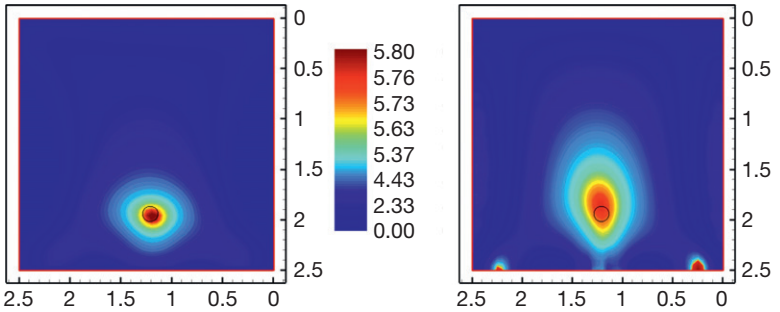


Figure 9 Reconstructed maps of a quantum dot inside the experimental phantom obtained with the spectrally resolved data. The hollow circle (O) indicates the location of the cylinder rod filled with a quantum dot.

4.17.4 Summary

Inverse models of bioluminescence and FMT are presented here. Traditional approaches to solving molecular tomographic problems are to solve a forward problem of light propagation repeatedly by updating the parameters of this forward model until the optimal values of parameters are found that minimize the objective function. As a consequence, this approach leads to a highly costly image reconstruction process, especially when the ERT is used as a forward model. On the other hand, another approach to this type of problems is to solve the forward and inverse problems simultaneously in a framework of PDE-constrained optimization, which considerably increases the image reconstruction speed with a factor of up to 25, thus saving lots of CPU times in the reconstruction. With respect to effects of noise and initial guess, both PDE-constrained and traditional unconstrained approaches show similar results.

References

Abddoulaev GS, Ren R, and Hielscher AH (2005) Optical tomography as a PDE-constrained optimization problem. *Inverse Problems* 21: 1507–1530.

Ahn S, Chaudhari AJ, Darvas F, Bouman A, and Leahy RM (2008) Fast iterative image reconstruction methods for fully 3D multispectral bioluminescence tomography. *Physics in Medicine and Biology* 53: 3921–3942.

Ale A, Ermolayev V, Herzog E, Cohrs C, Angelis MH, and Ntziachristos V (2012) FMT-XCT: In vivo animal studies with hybrid fluorescence molecular tomography-x-ray computed tomography. *Nature Methods* 9: 615–620.

Alexandrakis G, Rannou FR, and Chatziioannou AF (2005) Tomographic bioluminescence imaging by use of a combined optical-PET (OPET) system: A computer simulation feasibility study. *Physics in Medicine and Biology* 50: 4225–4241.

Arridge SR (1999) Optical tomography in medical imaging. *Inverse Problems* 15: R41–R93.

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

Biegler L, Nocedal J, Schmid C, and Ternet D (2000) Numerical experience with a reduced Hessian method for large scale constrained optimization. *Computational Optimization and Applications* 15: 45–67.

Biros G and Ghattas O (2003) Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: The Krylov-Schur solver. *SIAM Journal on Scientific Computing* 27: 687–713.

Bremer C, Ntziachristos V, and Weissleder R (2001) Optical-based molecular imaging: Contrast agents and potential medical applications. *European Radiology* 13: 231.

Byrd R, Curtis F, and Nocedal J (2008) An Inexact SQP Method for Equality Constrained Optimization. *SIAM Journal on Optimization* 19: 351–369.

Chang J, Graber HL, and Barbour RL (1997) Imaging of fluorescence in highly scattering media. *IEEE Transactions on Biomedical Engineering* 44: 810.

Chaudhari AJ, Darvas F, Bading JR, et al. (2005) Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging. *Physics in Medicine and Biology* 50: 5421–5441.

Chen B, Stamnes K, and Stamnes JJ (2001) Validity of the diffusion approximation in bio-optical imaging. *Applied Optics* 40: 6356.

Chen Y, Zheng G, Zhang ZH, et al. (2003) Metabolism-enhanced tumor localization by fluorescence imaging: In vivo animal studies. *Optics Letters* 28: 2070.

Cheong WF, Prahl SA, and Welch AJ (1990) A review of the optical properties of biological tissue. *IEEE Journal of Quantum Electronics* 26: 2166–2185.

Choy G, Choyke P, and Libutti SK (2003) Current advances in molecular imaging: Noninvasive in vivo bioluminescent and fluorescent optical imaging in cancer research. *Molecular Imaging* 2: 303.

Comsa DC, Farrell TJ, and Patterson MS (2007) Bioluminescence imaging of point sources implanted in small animals post mortem: Evaluation of a method for estimating source strength and depth. *Physics in Medicine and Biology* 52: 5415–5428.

Cong W, Wang G, Kumar D, et al. (2005) Practical reconstruction method for bioluminescence tomography. *Optics Express* 13: 6756–6771.

Contag PR (2002) Whole-animal cellular and molecular imaging to accelerate drug development. *Drug Discovery Today* 2002(7): 555–562.

Contag PR (2010) Functional imaging using bioluminescence markers. In: Weissleder R, Ross BD, Rehemtulla A, and Gambhir SS (Eds.) *Molecular Imaging*, pp. 118–138. Shelton, CT: People's Medical Publishing House Chapter 8.

Corlu A, Choe R, Durduran T, et al. (2007) Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans. *Optics Express* 15: 6696.

Dehghani H, Davis SC, Jiang S, Pogue BW, and Paulsen KD (2006) Spectrally resolved bioluminescence optical tomography. *Optics Letters* 31: 365–367.

Dehghani H, Davis SC, and Pogue BW (2008) Spectrally resolved bioluminescence tomography using the reciprocity approach. *Medical Physics* 35: 4863–4871.

Deliolanis NC, Dunham J, Wurdinger T, Figueiredo J-L, Bakhos T, and Ntziachristos V (2009) In-vivo imaging of murine tumors using complete-angle projection fluorescence molecular tomography. *Journal of Biomedical Optics* 14: 030509.

Elaloufi R, Carminati R, and Greffet JJ (2002) Time-dependent transport through scattering media: From radiative transfer to diffusion. *Journal of Optics A* 4: S103.

Eppstein MJ, Hawrysz DJ, Godavarty A, and Sevick-Muraca EM (2002) Three-dimensional Bayesian image reconstruction from sparse and noisy data sets: Near-infrared fluorescence tomography. *Proceedings of the National Academy of Sciences* 99: 9619.

Eslami S, Patterson M, Wong J, and Iordachita I (2012) SU-E-T-277: Hypothesis and design of an integrated x-ray/bioluminescent imaging (BLI) and tomography (BLT) system for the study of radiation and treatment in small animals. *Medical Physics* 39(6): 3767.

Galbán CJ, Galbán S, Van Dort ME, et al. (2010) Applications of molecular imaging. *Progress in Molecular Biology and Translational Science* 95: 237–298.

Gao H and Zhao H (2010) Multilevel bioluminescence tomography based on radiative transfer equation Part 1:l1 regularization. *Optics Express* 18: 1855.

Gill P, Murray W, and Saunders M (2005) SNOPT: An SQP algorithm for large-scale constrained optimization. *Society for Industrial and Applied Mathematics Review* 47: 99–131.

Gu X, Zhang Q, Larcom L, and Jiang H (2004) Three-dimensional bioluminescence tomography with model-based reconstruction. *Optics Express* 12: 3996–4000.

Han W, Kazmi K, Cong W, and Wang G (2007) Bioluminescence tomography with optimized optical parameters. *Inverse Problems* 23: 1215–1228.

Henyey LG and Greenstein LJ (1941) Diffuse radiation in the galaxy. *The Astrophysical Journal* 90: 70.

Hielscher AH, Alcouffe AE, and Barbour RL (1998) Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues. *Physics in Medicine and Biology* 43: 1285–1302.

Hielscher AH, Klose AD, and Hanson KM (1999) Gradient-based iterative image reconstruction scheme for time resolved optical tomography. *IEEE Transactions on Medical Imaging* 18(3): 262–271.

Hu H, Liu J, Yao L, et al. (2012) Real-time bioluminescence and tomographic imaging of gastric cancer in a novel orthotopic mouse model. *Oncology Reports* 27(6): 1937–1943.

Jiang H (1998) Frequency-domain fluorescent diffusion tomography: A finite-element-based algorithm and simulations. *Applied Optics* 37: 5337.

Joshi A, Bangerth W, and Sevick-Muraca EM (2006) Non-contact fluorescence optical tomography with scanning patterned illumination. *Optics Express* 14: 6516.

Kim HK and Charette A (2007) A sensitivity function-based conjugate gradient method for optical tomography with the frequency-domain equation of radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer* 104: 24–39.

Kim HK and Hielscher AH (2009) A PDE-constrained reduced Hessian SQP method for optical tomography based on the frequency-domain equation of radiative transfer. *Inverse Problems* 25: 015010.

Kim AD and Ishimaru A (1998) Optical diffusion of continuous-wave, pulsed, and density waves in scattering media and comparisons with radiative transfer. *Applied Optics* 37: 5313.

Kim HK, Lee JH, and Hielscher AH (2010) PDE-constrained fluorescence tomography with the frequency-domain equation of radiative transfer. *IEEE Journal of Selected Topics in Quantum Electronics* 16: 793–803.

Klose AD (2007) Transport-theory-based stochastic image reconstruction of bioluminescent sources. *Journal of the Optical Society of America A* 24(6): 1601–1608.

Klose AD and Hielscher AH (2003a) Fluorescence tomography with simulated data based on the equation of radiative transfer. *Optics Letters* 28: 1019.

Klose AD, Ntziachristos V, and Hielscher AH (2005) The inverse source problem based on the equation of radiative transfer. *Journal of Computational Physics* 202: 323–345.

Klose AD, Beattie BJ, Dehghani H, et al. (2010) In vivo bioluminescence tomography with a blocking-off finite-difference SP3 method and MRI/CT co-registration. *Medical Physics* 37: 329–338.

Kuo C, Coquoz O, Troy TL, Xu H, and Rice BW (2007) Three dimensional reconstruction of in vivo bioluminescent sources based on multispectral imaging. *Journal of Biomedical Optics* 12: 24007-1–24007-12.

Lalee M, Nocedal J, and Plantenga T (2003) On the implementation of an algorithm for large-scale equality constrained optimization. *SIAM Journal of Optimization* 8: 682–706.

Lee J and Sevick-Muraca EM (2002) Three-dimensional fluorescence enhanced optical tomography using referenced frequency-domain photon migration measurements at emission and excitation wavelengths. *Journal of the Optical Society of America A* 19: 759.

Li J, Wang X, Yi X, et al. (2012) Towards pH-sensitive imaging of small animals with photon-counting difference diffuse fluorescence tomography. *Journal of Biomedical Optics* 17(9): 096011.

Lu Y, Zhang X, Douraghy A, et al. (2009) Source reconstruction for spectrally-resolved bioluminescence tomography with sparse a priori information. *Optics Express* 17: 8062.

Lv Y, Tian J, Cong W, et al. (2007) Spectrally resolved bioluminescence tomography with adaptive finite element analysis: Methodology and simulation. *Physics in Medicine and Biology* 52: 4497–4512.

Milstein AB, Oh S, Webb KJ, et al. (2003) Fluorescence optical diffusion tomography. *Applied Optics* 42: 3081.

Minkowycz WJ, Sparrow EM, and Murthy JY (2006) *Handbook of Numerical Heat Transfer*. Hoboken, NJ: Wiley.

Modest M (2003) *Radiative Heat Transfer*. New York, NY: McGraw-Hill.

Montejo LD, Klose AD, and Hielscher AH (2010) Implementation of the equation of radiative transfer on block-structured grids for modeling light propagation in tissue. *Biomedical Optics Express* 1: 861–878.

Narsinh KH, Cao F, and Wu JC (2009) Molecular imaging of human embryonic stem cells. *Methods in Molecular Biology* 515: 13–32.

Nocedal J and Wright SJ (2006) *Numerical Optimization*. New York, NY: Springer.

Ntziachristos V and Weissleder R (2001) Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation. *Optics Letters* 26: 893.

Ntziachristos V, Tung CH, Bremer R, and Weissleder R (2002) Fluorescence molecular tomography resolves protease activity in vivo. *Nature Medicine* 8: 757.

O'Leary MA, Boas DA, Li XD, Chance B, and Yodh AG (1996) Fluorescence lifetime imaging in turbid media. *Optics Letters* 21: 158.

Paithankar DY, Chen AU, Pogue BW, Patterson MS, and Sevick-Muraca EM (1997) Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media. *Applied Optics* 36: 2260–2272.

Ren R, Bal G, and Hielscher AH (2006) Frequency domain optical tomography with the equation of radiative transfer. *SIAM Journal on Scientific Computing* 28: 1463–1489.

Roy R and Sevick-Muraca EM (2001) A numerical study of gradient-based nonlinear optimization methods for contrast enhanced optical tomography. *Optics Express* 9: 49.

Saad Y (2003) *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: SIAM.

Saad Y and Schultz MH (1986) GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 3: 856–869.

Slavine NW, Lewis MA, Richer E, and Antich PP (2006) Iterative reconstruction method for light emitting sources based on the diffusion equation. *Medical Physics* 33: 61–68.

Weissleder R and Ntziachristos V (2003) Shedding light onto live molecular targets. *Nature Medicine* 9: 123.

Zhang Q, Yin L, Tan Y, Yuan Z, and Jiang H (2008) Quantitative bioluminescence tomography guided by diffuse optical tomography. *Optics Express* 16: 1481–1486.

Zhao H, Doyle TC, Coquoz O, Kalish F, Rice BW, and Contag CH (2009) Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection in vivo. *Journal of Biomedical Optics* 10: 041210.