Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 187 (2003) 597–619
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Coupled complex adjoint sensitivities for frequency-domain fluorescence tomography: theory and vectorized implementation

F. Fedele [a], J.P. Laible [a], M.J. Eppstein [b,*]

[a] *Department of Civil and Environmental Engineering, University of Vermont, Burlington, VT 05405, USA*
[b] *Department of Computer Science, University of Vermont, Burlington, VT 05405, USA*

Received 10 December 2002; received in revised form 27 February 2003; accepted 28 February 2003

Abstract

We present a computationally efficient and accurate adjoint method for calculating coupled sensitivities of complex frequency-domain excitation and emission fluence to any underlying optical parameters in highly scattering media. The method is shown to be general and accurate. Novel vectorized implementations for finite element global matrix assembly and adjoint sensitivity calculations are shown to speed up calculations by orders of magnitude over traditional loop implementations, thereby making least-squares approaches to fluorescence tomography computationally practical.

Keywords: Adjoint sensitivities; Vectorization; Fluorescence tomography; Coupled elliptic equations

1. Introduction

Over the past decade there has been considerable progress made in near infrared (NIR) optical probing of biological tissues as a means of imaging endogenous differences in tissue properties such as absorption and scattering [1], particularly for breast cancer imaging [2,3]. The hope is that optical methods will provide functional information regarding local biochemical environments that will complement structural information already achievable by other imaging modalities, such as X-ray and magnetic resonance imaging. Recently, there has been a growing appreciation for the potential benefits of using exogenously introduced fluorescing dyes as an efficient means to improve contrast in order to discern small inclusions of disease in thick tissues [4,5], and work in fluorescence tomography has proceeded in a number of laboratories [6–28]. A variety of receptor-mediated fluorescent dyes under development offer the potential of highly selective

* Corresponding author. Tel.: +1-802-656-1918; fax: +1-802-656-0696.
E-mail address: Maggie.Eppstein@uvm.edu (M.J. Eppstein).

0021-9991/03/$ - see front matter
doi:10.1016/S0021-9991(03)00150-5

targeting of diseased tissues that will then fluoresce when excited by impinging NIR light [29–31]. In NIR fluorescence tomography, the goal is to successfully reconstruct the source of these fluorescent emissions and thereby characterize the size and location of diseased tissues. We restrict our discussions to frequency-domain techniques, although our results are readily translated into the time domain via Fourier transforms.

In frequency-domain photon migration through tissues, sinusoidally intensity-modulated NIR "excitation" light is launched into the tissue at the surface. During transit through the tissues, photons are absorbed and scattered owing to the local tissue optical properties and their spatial variations. The intensity wave that is detected at a number of positions on the tissue surface is therefore phase-delayed (θ) and amplitude attenuated (α) relative to the incident light. In the near-infrared range (700–900 nm) there exists a so-called "therapeutic window" where optical absorption, due mainly to water and hemoglobin, is relatively low, and the photon density wave can travel several centimeters before it is completely attenuated, although it is still highly scattered. When light is absorbed by fluorophore that is present in the tissue, the fluorophore is elevated to an excited state and remains there for some period of time (the fluorescence lifetime, τ). Some proportion of the excited molecules (the fluorescence quantum efficiency, ϕ) will ultimately release their excess energy by emitting a photon as they drop back to the ground state. This creates an "emission" photon density wave that is also scattered and absorbed before it reaches the detectors on the tissue surface, where it can be separated from the excitation photon density wave via interference filters.

In addition to an accurate forward model of coupled excitation and emission light propagation through highly scattering media, frequency-domain fluorescence tomography in tissues requires an inverse method for using noisy measurements of excitation and/or emission phase-delays and/or amplitude attenuation to reconstruct interior optical property maps of the tissues. Because of the high degree of scattering, most approaches to fluorescence tomography in large tissue volumes are based on regularized nonlinear least-squares optimization, such as the Levenberg–Marquardt method [11] or the Bayesian approximate extended Kalman filter [8]. Central to these methods is the repeated computation of Jacobian sensitivity matrices quantifying the effects of local changes in optical properties on the detected fluence. The focus of this paper is on development of a rapid and accurate methodology for computation of these sensitivities.

Previous approaches to computing emission sensitivities include first-order finite differences [6], second-order finite differences [7], and an approximate adjoint approach [8,10,11]. In finite difference approximations to the Jacobian, an optical parameter p is perturbed by some small amount δp and the resulting perturbed fluence $(\Phi + \delta\Phi)$ is explicitly evaluated by a forward simulator. For example, a first-order finite difference approximation is computed by

$$\frac{\partial \Phi}{\partial p} \overset{1^{\text{st}}\ \text{order}}{\approx} \frac{\Phi(p+\delta p) - \Phi(p)}{\delta p}. \tag{1}$$

If the size of the perturbation is chosen carefully, finite difference approximations can be very accurate, but in either case if the parameter p has been spatially discretized (e.g., into nodes or elements), then Eq. (1) must be recalculated for each discrete location in the domain. For large, 3-D domains this is computationally impractical. Nonetheless, finite difference approaches are very flexible and easily implemented for any optical property and any (directly or indirectly) measurable quantity, and they have been used for estimating the sensitivities of various components (e.g., phase and/or amplitude) of fluence at excitation and/or emission wavelengths, relative to absorption, scattering, fluorescence lifetime, and fluorescence quantum efficiency [6].

Adjoint methods [32,33] have been previously applied separately to the excitation equation [34,35] and to the emission equation [10,11]. This latter results in an approximate adjoint Jacobian that can be used for fluorescence tomography. While relatively fast, it makes the potentially limiting assumptions that the diffusion coefficient at emission wavelengths is spatially smooth and that the sensitivity is linear with respect to fluorescence absorption. In this paper, we apply the adjoint approach to the more complicated coupled system of excitation and emission equations describing fluorescent light generation and emission, in order

to derive the exact adjoint sensitivities of the coupled complex frequency-domain excitation and emission fluence relative to any of the underlying optical properties. We develop computational forms of the adjoint sensitivity equations using the Galerkin finite element method, and propose a novel computational implementation of these equations that is highly vectorized and manages memory requirements through domain decomposition. Computational results validate the accuracy and computational efficiency of our approach. Finally, we mention other physical problems modeled with coupled elliptic partial differential equations that could benefit from a similar approach.

2. Governing equations

The generation and propagation of fluorescent light through highly scattering media (such as biological tissues) is often modeled by a pair of second order, coupled, elliptic, partial differential equations [36–38]. The first equation represents propagation of excitation light (subscript x) and the second models the generation and propagation of fluorescently emitted light (subscript m). In the frequency domain, these diffusion approximations to the coupled radiative transport equation over a three-dimensional (3-D) bounded domain Ω are

$$\begin{cases} -\nabla\cdot(D_x\nabla\Phi_x)+k_x\Phi_x=S_x \\ -\nabla\cdot(D_m\nabla\Phi_m)+k_m\Phi_m=\beta\Phi_x \end{cases} \quad \text{on } \Omega \tag{2}$$

subject to the Robin boundary conditions on the domain boundary $\partial\Omega$ of

$$\begin{cases} \overrightarrow{n}\cdot(D_x\nabla\Phi_x)+b_x\Phi_x=0 \\ \overrightarrow{n}\cdot(D_m\nabla\Phi_m)+b_m\Phi_m=0 \end{cases} \quad \text{on } \partial\Omega, \tag{3}$$

where ∇ is the 3×1 grad operator and $\overrightarrow{n}$ is the 3×1 vector normal to the boundary. The excitation light source S_x (W/cm^3) is intensity modulated with sinusoidal frequency ω (rad/s), and propagates through the media resulting in the AC component of complex photon fluence at the excitation wavelength of Φ_x(W/cm^2), where $\Phi_x=\alpha_x e^{i\theta_x}$. Some of this excitation light may be absorbed by fluorophore in the media and reemitted, resulting in complex photon fluence at the emission wavelength $\Phi_m=\alpha_m e^{i\theta_m}$. The diffusion $(D_{x,m})$, decay $(k_{x,m})$, and emission source (β) coefficients, as shown below

$$\begin{cases} D_x=\frac{1}{3\left(\mu_{axi}+\mu_{axf}+\mu'_{sx}\right)}, \\ D_m=\frac{1}{3\left(\mu_{ami}+\mu_{amf}+\mu'_{sm}\right)}, \end{cases} \quad \begin{cases} k_x=\frac{i\omega}{c}+\mu_{axi}+\mu_{axf}, \\ k_m=\frac{i\omega}{c}+\mu_{ami}+\mu_{amf}, \end{cases} \quad \beta=\frac{\phi\mu_{axf}}{1-i\omega\tau}, \tag{4}$$

are functions of absorption due to non-fluorescing chromophore (μ_{axi},μ_{ami}), absorption due to fluorophore (μ_{axf},μ_{amf}), and isotropic (reduced) scattering (μ'_{sx},μ'_{sm}) at the two wavelengths (all in units of cm^{-1}), fluorescence quantum efficiency (ϕ), and fluorescence lifetime (τ, in s). Here, $i=\sqrt{-1}$, and c is the speed of light in the media (cm/s). The Robin boundary coefficients (b_x,b_m) are governed by the reflection coefficients (R_x,R_m), which range from 0 (no reflectance) to 1 (total reflectance)

$$b_x=\frac{1-R_x}{2(1+R_x)}, \quad b_m=\frac{1-R_m}{2(1+R_m)}. \tag{5}$$

In NIR fluorescence imaging, where excitation and emission wavelengths are relatively close to one another, the absorption and scattering at emission wavelengths can be approximated as linear functions of absorption and scattering at excitation wavelengths. Although not explicitly shown in these equations, all optical properties and field variables are understood to be potentially variable in Cartesian space.

2.1. Matrix notation for the coupled equations

For brevity, we represent the coupled governing equations (2) with the following single matrix equation:

$$-\underline{\underline{\nabla}}^{\mathrm{T}}(\underline{\underline{\mathbf{D}}}\,\underline{\underline{\nabla}}\,\underline{\mathbf{\Phi}}) + \underline{\underline{\mathbf{k}}}\,\underline{\mathbf{\Phi}} = \underline{\mathbf{S}} \quad \text{on } \Omega. \tag{6}$$

Similarly, the boundary conditions (3) are represented by the matrix equation

$$\underline{\underline{\mathbf{n}}}^{\mathbf{T}}(\underline{\underline{\mathbf{D}}}\,\underline{\underline{\nabla}}\,\underline{\mathbf{\Phi}}) + \underline{\underline{\mathbf{b}}}\,\underline{\mathbf{\Phi}} = \underline{\mathbf{0}} \quad \text{on } \partial\Omega. \tag{7}$$

In (6) and (7), we use the following matrix definitions (sizes of each matrix are shown for clarity):

$$\begin{cases} \underset{(6\times 2)}{\underline{\underline{\nabla}}} = \begin{bmatrix} \nabla & \underline{0} \\ \underline{0} & \nabla \end{bmatrix}, \quad \underset{(6\times 2)}{\underline{\underline{\mathbf{n}}}} = \begin{bmatrix} \overrightarrow{n} & \underline{0} \\ \underline{0} & \overrightarrow{n} \end{bmatrix}, \quad \underset{(6\times 6)}{\underline{\underline{\mathbf{D}}}} = \begin{bmatrix} D_x\underline{\underline{\mathbf{I}}} & \underline{\underline{0}} \\ \underline{\underline{0}} & D_m\underline{\underline{\mathbf{I}}} \end{bmatrix}, \\ \underset{(2\times 2)}{\underline{\underline{\mathbf{k}}}} = \begin{bmatrix} k_x & 0 \\ -\beta & k_m \end{bmatrix}, \quad \underset{(2\times 2)}{\underline{\underline{\mathbf{b}}}} = \begin{bmatrix} b_x & 0 \\ 0 & b_m \end{bmatrix}, \\ \underset{(2\times 1)}{\underline{\mathbf{\Phi}}} = \begin{bmatrix} \Phi_x \\ \Phi_m \end{bmatrix}, \quad \underset{(2\times 1)}{\underline{\mathbf{S}}} = \begin{bmatrix} S_x \\ 0 \end{bmatrix}. \end{cases} \tag{8}$$

3. Adjoint sensitivity formulation

3.1. Perturbation equations

Consider that p is any optical property either directly or indirectly embedded in Eq. (6) or (7). E.g.,

$$p \in \{\mu_{axf}, \mu_{axi}, \mu'_{sx}, \mu_{amf}, \mu_{ami}, \mu'_{sm}, \tau, \phi, R_x, R_m, \ldots\}. \tag{9}$$

An infinitesimally small perturbation of the parameter δp will cause a corresponding variation in fluence $\underline{\delta\mathbf{\Phi}}$

$$p \to p + \delta p \Rightarrow \underline{\mathbf{\Phi}} \to \underline{\mathbf{\Phi}} + \underline{\delta\mathbf{\Phi}} \tag{10}$$

such that the system

$$-\underline{\underline{\nabla}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}(p+\delta p)\underline{\underline{\nabla}}(\underline{\mathbf{\Phi}} + \underline{\delta\mathbf{\Phi}})\right) + \underline{\underline{\mathbf{k}}}(p+\delta p)(\underline{\mathbf{\Phi}} + \underline{\delta\mathbf{\Phi}}) = \underline{\mathbf{S}} \quad \text{on } \Omega, \tag{11}$$

$$\underline{\underline{\mathbf{n}}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}(p+\delta p)\underline{\underline{\nabla}}(\underline{\mathbf{\Phi}} + \underline{\delta\mathbf{\Phi}})\right) + \underline{\underline{\mathbf{b}}}(p+\delta p)(\underline{\mathbf{\Phi}} + \underline{\delta\mathbf{\Phi}}) = \underline{\mathbf{0}} \quad \text{on } \partial\Omega \tag{12}$$

is satisfied. Using a Taylor series expansion of all terms in (11) and (12) yields the following first-order perturbation equations governing the variation $\underline{\delta\mathbf{\Phi}}$, where we have ignored higher order terms $\mathrm{O}(\delta p^2)$:

$$-\underline{\underline{\nabla}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}\,\underline{\underline{\nabla}}\,\underline{\delta\mathbf{\Phi}}\right) + \underline{\underline{\mathbf{k}}}\,\underline{\delta\mathbf{\Phi}} = \underline{\underline{\nabla}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\,\underline{\underline{\nabla}}\,\underline{\mathbf{\Phi}}\right) - \frac{\partial\underline{\underline{\mathbf{k}}}}{\partial p}\delta p\,\underline{\mathbf{\Phi}} \quad \text{on } \Omega, \tag{13}$$

$$\underline{\underline{\mathbf{n}}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}\,\underline{\underline{\nabla}}\,\underline{\delta\mathbf{\Phi}}\right) + \underline{\underline{\mathbf{b}}}\,\underline{\delta\mathbf{\Phi}} = -\underline{\underline{\mathbf{n}}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\,\underline{\underline{\nabla}}\,\underline{\mathbf{\Phi}}\right) - \frac{\partial\underline{\underline{\mathbf{b}}}}{\partial p}\delta p\,\underline{\mathbf{\Phi}} \quad \text{on } \partial\Omega. \tag{14}$$

One could solve the system (13) subject to (14) in order to get a first-order approximation of the sensitivity of $\underline{\mathbf{\Phi}}$ with respect to p. As with the finite difference approach, if p has been spatially discretized this would require solving system (13) once for each locally discrete region of p. We will now apply the adjoint method [32,33] in order to reduce the number and size of the systems of equations that must be solved.

3.2. Analytical adjoint sensitivities

To derive the sensitivities by an adjoint method, we first define a matrix of functions $\underline{\underline{\mathbf{\Psi}}}$ as

$$\underset{(2\times 2)}{\underline{\underline{\mathbf{\Psi}}}} = \begin{bmatrix} \Psi_{xx} & \Psi_{xm} \\ \Psi_{mx} & \Psi_{mm} \end{bmatrix}, \tag{15}$$

and choose $\underline{\underline{\mathbf{\Psi}}}$ such that it satisfies the following system of equations that we will refer to as the adjoint system of (6)

$$\begin{cases} -\underline{\nabla}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}^{\mathrm{T}}\underline{\underline{\nabla\mathbf{\Psi}}}\right) + \underline{\underline{\mathbf{k}}}^{\mathrm{T}}\underline{\underline{\mathbf{\Psi}}} = \underline{\underline{\Delta_d}} & \text{on } \Omega, \\ \underline{\mathbf{n}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}^{\mathrm{T}}\underline{\underline{\nabla\mathbf{\Psi}}}\right) + \underline{\underline{\mathbf{b}}}^{\mathrm{T}}\underline{\underline{\mathbf{\Psi}}} = \underline{0} & \text{on } \partial\Omega, \end{cases} \quad \text{where } \underset{(2\times 2)}{\underline{\underline{\Delta_d}}} = \begin{bmatrix} \Delta_d & 0 \\ 0 & \Delta_d \end{bmatrix}. \tag{16}$$

We multiply the system (13) by $\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}$ and integrate over the entire domain Ω to get

$$\int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(-\underline{\nabla}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}\underline{\underline{\nabla\delta\mathbf{\Phi}}}\right) + \underline{\underline{\mathbf{k}}}\underline{\delta\mathbf{\Phi}}\right) = \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(\underline{\nabla}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right) - \frac{\partial\underline{\underline{\mathbf{k}}}}{\partial p}\delta p\underline{\mathbf{\Phi}}\right). \tag{17}$$

Integrating by parts twice, applying the boundary conditions (14), and rearranging terms, we get

$$\begin{aligned} \int_\Omega \overbrace{\left(-\underline{\nabla}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}^{\mathrm{T}}\underline{\underline{\nabla\mathbf{\Psi}}}\right) + \underline{\underline{\mathbf{k}}}^{\mathrm{T}}\underline{\underline{\mathbf{\Psi}}}\right)^{\mathrm{T}}}^{\underline{\underline{\Delta_d}}}\underline{\delta\mathbf{\Phi}} &= \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathbf{T}}\left(\underline{\nabla}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right)\right) - \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\frac{\partial\underline{\underline{\mathbf{k}}}}{\partial p}\delta p\underline{\mathbf{\Phi}} \\ &\quad + \int_{\partial\Omega}\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(-\underline{\mathbf{n}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right)\right) - \int_{\partial\Omega}\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\frac{\partial\underline{\underline{\mathbf{b}}}}{\partial p}\delta p\underline{\mathbf{\Phi}} \\ &\quad - \int_{\partial\Omega}\overbrace{\underline{\underline{\mathbf{n}}}^{\mathrm{T}}\left(\underline{\underline{\mathbf{D}}}^{\mathrm{T}}\underline{\underline{\nabla\mathbf{\Psi}}}\right) + \underline{\underline{\mathbf{b}}}^{\mathrm{T}}\underline{\underline{\mathbf{\Psi}}}}^{\underline{\underline{0}}}\underline{\delta\mathbf{\Phi}}. \end{aligned} \tag{18}$$

As indicated above the terms in Eq. (18), we can simplify by substituting Eqs. (16) to get the following simplified equation:

$$\underline{\delta\mathbf{\Phi}} = \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(\underline{\nabla}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right)\right) - \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\frac{\partial\underline{\underline{\mathbf{k}}}}{\partial p}\delta p\underline{\mathbf{\Phi}} + \int_{\partial\Omega}\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(-\underline{\mathbf{n}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right)\right) - \int_{\partial\Omega}\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\frac{\partial\underline{\underline{\mathbf{b}}}}{\partial p}\delta p\underline{\mathbf{\Phi}}. \tag{19}$$

Integrating the first term of Eq. (19) by parts once more, rearranging, and canceling terms, we are left with the following three terms:

$$\underline{\delta\mathbf{\Phi}} = -\int_\Omega \left(\underline{\underline{\nabla\mathbf{\Psi}}}\right)^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{D}}}}{\partial p}\delta p\underline{\nabla\mathbf{\Phi}}\right) - \int_\Omega \underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{k}}}}{\partial p}\delta p\underline{\mathbf{\Phi}}\right) - \int_{\partial\Omega}\underline{\underline{\mathbf{\Psi}}}^{\mathrm{T}}\left(\frac{\partial\underline{\underline{\mathbf{b}}}}{\partial p}\delta p\underline{\mathbf{\Phi}}\right). \tag{20}$$

Since the solution to system (16) depends on p, but not on the variation δp, this system only needs to be solved once for $\underline{\underline{\Psi}}$ for a given set of parameter values. Eq. (20) can then be used to compute desired sensitivities to any underlying optical properties. The matrix $\underline{\underline{\Psi}}$ thus represents the solutions to the adjoint system (16), in response to a Dirac source located at each of the detectors d. We refer to $\underline{\underline{\Psi}}$ as the adjoint or Green matrix of the coupled system.

Now, let us take a closer look at the matrix $\underline{\underline{\Psi}}$. The matrix system (16) is simply a compact representation of the following four boundary value problems:

$$\begin{cases} -\nabla\cdot(D_x\nabla\Psi_{xx}) + k_x\Psi_{xx} - \beta\Psi_{mx} = \Delta_d & \text{on } \Omega, \\ \vec{n}\cdot(D_x\nabla\Psi_{xx}) + b_x\Psi_{xx} = 0 & \text{on } \partial\Omega, \end{cases} \tag{21}$$

$$\begin{cases} -\nabla\cdot(D_x\nabla\Psi_{xm}) + k_x\Psi_{xm} = \beta\Psi_{mm} & \text{on } \Omega, \\ \vec{n}\cdot(D_x\nabla\Psi_{xm}) + b_x\Psi_{xm} = 0 & \text{on } \partial\Omega, \end{cases} \tag{22}$$

$$\begin{cases} -\nabla\cdot(D_m\nabla\Psi_{mx}) + k_m\Psi_{mx} = 0 & \text{on } \Omega, \\ \vec{n}\cdot(D_m\nabla\Psi_{mx}) + b_m\Psi_{mx} = 0 & \text{on } \partial\Omega, \end{cases} \tag{23}$$

$$\begin{cases} -\nabla\cdot(D_m\nabla\Psi_{mm}) + k_m\Psi_{mm} = \Delta_d & \text{on } \Omega, \\ \vec{n}\cdot(D_m\nabla\Psi_{mm}) + b_m\Psi_{mm} = 0 & \text{on } \partial\Omega. \end{cases} \tag{24}$$

Note that Eq. (23) admits only the trivial solution; that is, $\Psi_{mx} = 0$. It can admit a non-vanishing solution if and only if the coefficient k_m is an eigenvalue λ of the equation, i.e., $-\nabla\cdot(D_m\nabla\Psi_{mx}) + \lambda\Psi_{mx} = 0$. The latter equation admits real positive eigenvalues since it represents wave phenomena where resonance can occur. Since we are dealing with a diffusion process, k_m is a complex number and cannot be an eigenvalue, therefore only the trivial solution is admissible. The fact that $\Psi_{mx} = 0$ reflects the asymmetry in the coupled governing equations (2); that is, Φ_x affects Φ_m, but not vice versa.

The adjoint sensitivities represented in Eq. (20) can be decomposed into the sensitivity equations for excitation and emission fluence. For example, the sensitivity of excitation fluence is

$$\delta\Phi_x = -\int_\Omega \nabla\Psi_{xx}\frac{\partial D_x}{\partial p}\delta p\nabla\Phi_x - \int_\Omega \Psi_{xx}\frac{\partial k_x}{\partial p}\delta p\Phi_x - \int_{\partial\Omega}\Psi_{xx}\frac{\partial b_x}{\partial p}\delta p\Phi_x, \tag{25}$$

which is identical to the adjoint formulation for excitation fluence reported elsewhere (where $\partial b_x/\partial p = 0$) [1]. The sensitivity for emission fluence in the coupled system is a new result, as follows:

$$\begin{aligned}\delta\Phi_m = &-\int_\Omega \nabla\Psi_{mm}\frac{\partial D_m}{\partial p}\delta p\nabla\Phi_m - \int_\Omega \Psi_{mm}\frac{\partial k_m}{\partial p}\delta p\Phi_m - \int_{\partial\Omega}\Psi_{mm}\frac{\partial b_m}{\partial p}\delta p\Phi_m - \int_\Omega \nabla\Psi_{xm}\frac{\partial D_x}{\partial p}\delta p\nabla\Phi_x \\ &- \int_\Omega \Psi_{xm}\frac{\partial k_x}{\partial p}\delta p\Phi_x - \int_{\partial\Omega}\Psi_{xm}\frac{\partial b_x}{\partial p}\delta p\Phi_x + \int_\Omega \Psi_{mm}\frac{\partial\beta}{\partial p}\delta p\Phi_x.\end{aligned} \tag{26}$$

The previously proposed approximate adjoint approach [10,11] is equivalent to only the last term in Eq. (26) when D_m is smooth. We shall refer to the sensitivies (25) and (26) as analytical adjoint sensitivities because they are the exact analytical sensitivities assuming the exact Green matrix is known. However, analytical solutions for the Green matrix are only available for simplified geometries and constant coefficients, so if a discrete numerical formulation is adopted in order to approximate the Green matrix, the resulting sensitivities also become only approximate.

4. Finite element formulation

4.1. FEM for governing equations of excitation and fluorescence

The method of *weighted residuals* is used as the basis for the *finite element method* (FEM) [39]. We introduce on the domain Ω a three-dimensional grid Υ with n nodes and boundary $\partial\Omega$. Over the entire domain Ω, we approximate a continuous (possibly complex) field variable u with the finite functional space $V(\Omega)$ as

$$V(\Omega) = \{\widehat{\boldsymbol{u}} \in \mathscr{C}^0(\Omega),\ \widehat{\boldsymbol{u}}|_e. \text{ is a polynomial},$$

where e is a generic element in Υ. We define a basis for $V_h(\Omega)$ as $[\mathbf{N}] = [N_1, N_2, \ldots, N_n]$. The generic basis elements are defined such that $N_i(x_j, y_j, z_j) = \delta_{ij}$. By means of these basis functions, we approximate a generic field variable u as $u \approx \widehat{\boldsymbol{u}} = [\mathbf{N}][\mathbf{u}]$, where $[\mathbf{u}]$ is a column vector of the nodal values of the function u. In particular, the complex field variables Φ_x and Φ_m are approximated as

$$\begin{aligned} \Phi_x &\approx \hat{\boldsymbol{\Phi}}_x = [\mathbf{N}][\boldsymbol{\Phi}_x], \\ \Phi_m &\approx \hat{\boldsymbol{\Phi}}_m = [\mathbf{N}][\boldsymbol{\Phi}_m]. \end{aligned} \tag{27}$$

We use the Galerkin Finite Element Method (GFEM) [39], in which the weight functions are $[\mathbf{N}]^{\mathrm{T}}$, and hence the weighted residual forms of the governing equations (2) are

$$\begin{aligned} \int_\Omega [\mathbf{N}]^{\mathrm{T}} \left(-\nabla \cdot \left(D_x \nabla \hat{\boldsymbol{\Phi}}_x \right) + k_x \hat{\boldsymbol{\Phi}}_x \right) &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} S_x, \\ \int_\Omega [\mathbf{N}]^{\mathrm{T}} \left(-\nabla \cdot \left(D_m \nabla \hat{\boldsymbol{\Phi}}_m \right) + k_m \hat{\boldsymbol{\Phi}}_m \right) &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} \beta \hat{\boldsymbol{\Phi}}_x. \end{aligned} \tag{28}$$

Integration by parts of the second-order terms yields:

$$\begin{aligned} \int_\Omega (\nabla[\mathbf{N}])^{\mathrm{T}} D_x \nabla \hat{\boldsymbol{\Phi}}_x + \int_\Omega [\mathbf{N}]^{\mathrm{T}} k_x \hat{\boldsymbol{\Phi}}_x - \int_{\partial\Omega} [\mathbf{N}]^{\mathrm{T}} \left(\vec{n} \cdot \left(D_x \nabla \hat{\boldsymbol{\Phi}}_x \right) \right) &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} S_x, \\ \int_\Omega (\nabla[\mathbf{N}])^{\mathrm{T}} D_m \nabla \hat{\boldsymbol{\Phi}}_m + \int_\Omega [\mathbf{N}]^{\mathrm{T}} k_m \hat{\boldsymbol{\Phi}}_m - \int_{\partial\Omega} [\mathbf{N}]^{\mathrm{T}} \left(\vec{n} \cdot \left(D_m \nabla \hat{\boldsymbol{\Phi}}_m \right) \right) &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} \beta \hat{\boldsymbol{\Phi}}_x. \end{aligned} \tag{29}$$

Introducing the Robin Boundary conditions of Eq. (3) yields

$$\begin{aligned} \int_\Omega (\nabla[\mathbf{N}])^{\mathrm{T}} D_x \nabla \hat{\boldsymbol{\Phi}}_x + \int_\Omega [\mathbf{N}]^{\mathrm{T}} k_x \hat{\boldsymbol{\Phi}}_x + \int_{\partial\Omega} [\mathbf{N}]^{\mathrm{T}} b_x \hat{\boldsymbol{\Phi}}_x &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} S_x, \\ \int_\Omega (\nabla[\mathbf{N}])^{\mathrm{T}} D_m \nabla \hat{\boldsymbol{\Phi}}_m + \int_\Omega [\mathbf{N}]^{\mathrm{T}} k_m \hat{\boldsymbol{\Phi}}_m + \int_{\partial\Omega} [\mathbf{N}]^{\mathrm{T}} b_m \hat{\boldsymbol{\Phi}}_m &= \int_\Omega [\mathbf{N}]^{\mathrm{T}} \beta \hat{\boldsymbol{\Phi}}_x. \end{aligned} \tag{30}$$

Introducing the approximations (27) and $S_x = [\mathbf{N}][\mathbf{S}_x]$ (where $[\mathbf{S}_x]$ represents nodal values of S_x) after collecting common terms we finally have in block form:

$$[\mathbf{A}][\boldsymbol{\Phi}] = [\mathbf{S}], \tag{31}$$

where we have defined the block vectors

$$[\boldsymbol{\Phi}] = \begin{bmatrix} [\boldsymbol{\Phi}_x], \\ [\boldsymbol{\Phi}_m] \end{bmatrix}, \quad [\mathbf{S}] = \begin{bmatrix} [\mathbf{M}][\mathbf{S}_x] \\ [\mathbf{0}] \end{bmatrix} \tag{32}$$

and block matrix

$$[\mathbf{A}] = \begin{bmatrix} [\mathbf{A}_x] & [\mathbf{0}] \\ -[\mathbf{M}_\beta] & [\mathbf{A}_m], \end{bmatrix} \tag{33}$$

where

$$[\mathbf{A}_x] = \mathscr{A}(D_x, k_x, b_x), \tag{34}$$

$$[\mathbf{A}_m] = \mathscr{A}(D_m, k_m, b_m), \tag{35}$$

$$[\mathbf{M}] = \mathscr{M}(1), \tag{36}$$

$$[\mathbf{M}_\beta] = \mathscr{M}(\beta) \tag{37}$$

with the following matrix function definitions:

$$\mathscr{A}(D,k,b) = \left[\int_\Omega (\nabla[\mathbf{N}])^{\mathrm{T}} D \nabla[\mathbf{N}]\right] + \left[\int_\Omega [\mathbf{N}]^{\mathrm{T}} k [\mathbf{N}]\right] + \left[\int_{\partial\Omega} [\mathbf{N}]^{\mathrm{T}} b [\mathbf{N}]\right], \tag{38}$$

$$\mathscr{M}(\beta) = \left[\int_\Omega [\mathbf{N}]^{\mathrm{T}} \beta [\mathbf{N}]\right]. \tag{39}$$

These equations are general in the sense that any type of finite element can be used. In order to model tissue domains that are 3-D and irregular in shape, we employ linear tetrahedral elements each of which have four nodes. At the element level, the explicit matrices have the following dimensions:

$$\underset{(1\times 4)}{[\mathbf{N}]} = [\,N_1 \quad N_2 \quad N_3 \quad N_4\,], \quad \underset{(3\times 4)}{\nabla[\mathbf{N}]} = \begin{bmatrix} \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial x} & \frac{\partial N_4}{\partial x} \\ \frac{\partial N_1}{\partial y} & \frac{\partial N_2}{\partial y} & \frac{\partial N_3}{\partial y} & \frac{\partial N_4}{\partial y} \\ \frac{\partial N_1}{\partial z} & \frac{\partial N_2}{\partial z} & \frac{\partial N_3}{\partial z} & \frac{\partial N_4}{\partial z} \end{bmatrix}. \tag{40}$$

In general, any of the nodally discretized coefficients ($D_{x,m}, k_{x,m}, b_{x,m}$, and β) can have directional dependency and are thus properly represented by 3×3 matrices (second-order tensors). For our application we assume isotropic (but not homogeneous) conditions. As such, we implement nodal values of $D_{x,m}$, $k_{x,m}$, $b_{x,m}$, and β as scalars. The inhomogeneity of optical properties means that even the scalar representations of $D_{x,m}$, $k_{x,m}$, $b_{x,m}$, and β can vary over an element. If these coefficients are represented by nodal values, then it is customary to use the same basis function expansions N_i to represent the value of the coefficients at any point in the element, e.g. $D_x \approx [\mathbf{N}][\mathbf{D}_x]$, where $[\mathbf{D}_x]$ is a 4×1 vector of nodal values for the tetrahedral element. However, if the elements are sufficiently small they can be treated as homogeneous in order to simplify the computational procedures, as described later.

4.2. FEM for continuous adjoint equations of the coupled governing equations

We again employ the GFEM method, this time to the adjoint Eqs. (21)–(24). We use the approximation for the Green matrix $\mathbf{\Psi}$ as

$$\mathbf{\Psi} \approx \begin{bmatrix} \hat{\boldsymbol{\Psi}}_{xx} & \hat{\boldsymbol{\Psi}}_{xm} \\ \hat{\boldsymbol{\Psi}}_{mx} & \hat{\boldsymbol{\Psi}}_{mm} \end{bmatrix} = \begin{bmatrix} [\mathbf{N}] & [\mathbf{0}] \\ [\mathbf{0}] & [\mathbf{N}] \end{bmatrix} \begin{bmatrix} [\mathbf{\Psi}_{xx}] & [\mathbf{\Psi}_{xm}] \\ [\mathbf{\Psi}_{mx}] & [\mathbf{\Psi}_{mm}] \end{bmatrix}, \tag{41}$$

where $[\mathbf{\Psi}_{xx}], [\mathbf{\Psi}_{xm}], [\mathbf{\Psi}_{mx}], [\mathbf{\Psi}_{mm}]$ are the nodal values of the respective field approximations $\hat{\mathbf{\Psi}}_{xx}$, $\hat{\mathbf{\Psi}}_{xm}$, $\hat{\mathbf{\Psi}}_{mx}$, $\hat{\mathbf{\Psi}}_{mm}$. We thus have the following set of discretized equations:

$$[\mathbf{A}_x][\mathbf{\Psi}_{xx}] = [\mathbf{\Delta}_d], \tag{42}$$

$$[\mathbf{A}_x][\mathbf{\Psi}_{xm}] = \left[\mathbf{M}_\beta\right][\mathbf{\Psi}_{mm}], \tag{43}$$

$$[\mathbf{A}_m][\mathbf{\Psi}_{mm}] = [\mathbf{\Delta}_d], \tag{44}$$

or in block form

$$\left[\tilde{\mathbf{A}}\right][\mathbf{\Psi}] = [\mathbf{\Delta}], \tag{45}$$

where we have defined the block matrices

$$\left[\tilde{\mathbf{A}}\right] = \begin{bmatrix} [\mathbf{A}_x] & -\left[\mathbf{M}_\beta\right] \\ [\mathbf{0}] & [\mathbf{A}_m] \end{bmatrix} \tag{46}$$

and

$$[\mathbf{\Psi}] = \begin{bmatrix} [\mathbf{\Psi}_{xx}] & [\mathbf{\Psi}_{xm}] \\ [\mathbf{0}] & [\mathbf{\Psi}_{mm}] \end{bmatrix}, \quad [\mathbf{\Delta}] = \begin{bmatrix} [\mathbf{\Delta}_d] & [\mathbf{0}], \\ [\mathbf{0}] & [\mathbf{\Delta}_d] \end{bmatrix}. \tag{47}$$

In agreement with the previous observation that $\Psi_{mx} = 0$ for the continuous equation (23), $[\mathbf{\Psi}_{mx}] = [\mathbf{0}]$ for the discrete equation, as explicitly indicated in (47). Because each individual partial differential equation of the coupled system (2) is self-adjoint, and because the finite element matrices generated by (38) and (39) are symmetric, the kernal (46) of the adjoint system (45) is the transpose of the kernal (33) of the forward system (31). That is,

$$\left[\tilde{\mathbf{A}}\right] = [\mathbf{A}]^{\mathrm{T}}. \tag{48}$$

4.3. FEM formulation of the analytical sensitivity of Φ_x and Φ_m

Applying the GFEM to Eq. (20) yields

$$[\boldsymbol{\delta\Phi}] = -[\mathbf{\Psi}]^{\mathrm{T}}[\boldsymbol{\delta A}][\mathbf{\Phi}] \tag{49}$$

in which

$$[\boldsymbol{\delta\Phi}] = \begin{bmatrix} [\boldsymbol{\delta\Phi}_x]_d \\ [\boldsymbol{\delta\Phi}_m]_d \end{bmatrix} \quad \text{and} \quad [\boldsymbol{\delta A}] = \begin{bmatrix} \mathscr{A}(\delta D_x, \delta k_x, \delta b_x) & [\mathbf{0}] \\ -\mathscr{M}(\delta\beta) & \mathscr{A}(\delta D_m, \delta k_m, \delta b_m) \end{bmatrix} \tag{50}$$

at detectors d, where

$$\delta D_{x,m} = \frac{\partial D_{x,m}}{\partial p}\delta p, \quad \delta k_{x,m} = \frac{\partial k_{x,m}}{\partial p}\delta p, \quad \delta b_{x,m} = \frac{\partial b_{x,m}}{\partial p}\delta p, \quad \delta\beta = \frac{\partial\beta}{\partial p}\delta p. \tag{51}$$

In Eqs. (51), the analytic expressions for the partial derivatives are employed. For example, the analytic expression for $\partial D_x/\partial\mu_{axf}$ is

$$\frac{\partial D_x}{\partial \mu_{axf}} = \frac{\partial}{\partial \mu_{axf}}\left(\frac{1}{3\left(\mu_{axi}+\mu_{axf}+\mu'_{sx}\right)}\right) = -\frac{1}{3\left(\mu_{axi}+\mu_{axf}+\mu'_{sx}\right)^2}. \tag{52}$$

4.4. Adjoint sensitivities by the discrete formulation

The expression (49) has been obtained by applying the GFEM to the continuous equation (20) of the analytical variation $\underline{\boldsymbol{\delta\Phi}}$, where a GFEM approximation of the analytical Green matrix $\underline{\underline{\boldsymbol{\Psi}}}$ has been used; we will refer to this as the "continuous" formulation of the adjoint sensitivities. Alternatively, we can apply the adjoint method directly to the discretized forward equations (31), a method we will refer to as the "discrete" formulation of the adjoint sensitivities.

The finite element matrices in (33) depend upon the generic parameter p; if we vary p as $p+\delta p$, $[\boldsymbol{\Phi}]$ will vary as $[\boldsymbol{\Phi}]+[\widetilde{\delta\boldsymbol{\Phi}}]$, and the latter still must satisfy Eq. (31). Using a Taylor-series expansion on (31) with respect to the parameter p and neglecting higher order terms one gets

$$([\mathbf{A}]+[\boldsymbol{\delta A}])\left([\boldsymbol{\Phi}]+\left[\widetilde{\boldsymbol{\delta\Phi}}\right]\right) = [\mathbf{S}]+\mathrm{O}(\delta p), \tag{53}$$

where $[\boldsymbol{\delta A}]$is expressed as in (50). This gives, up to first order in δp,

$$[\mathbf{A}]\left[\widetilde{\boldsymbol{\delta\Phi}}\right] = -[\boldsymbol{\delta A}][\boldsymbol{\Phi}]. \tag{54}$$

Multiplying by the transpose of an arbitrary block matrix $[\tilde{\boldsymbol{\Psi}}]$ one gets

$$\left[\tilde{\boldsymbol{\Psi}}\right]^{\mathrm{T}}[\mathbf{A}]\left[\widetilde{\boldsymbol{\delta\Phi}}\right] = -\left[\tilde{\boldsymbol{\Psi}}\right]^{\mathrm{T}}[\boldsymbol{\delta A}][\boldsymbol{\Phi}] \tag{55}$$

or equivalently

$$\left[\widetilde{\boldsymbol{\delta\Phi}}\right]^{\mathrm{T}}[\mathbf{A}]^{\mathrm{T}}\left[\tilde{\boldsymbol{\Psi}}\right] = -\left[\tilde{\boldsymbol{\Psi}}\right]^{\mathrm{T}}[\boldsymbol{\delta A}][\boldsymbol{\Phi}]. \tag{56}$$

Choosing

$$[\mathbf{A}]^{\mathrm{T}}\left[\tilde{\boldsymbol{\Psi}}\right] = [\boldsymbol{\Delta}], \tag{57}$$

where $[\boldsymbol{\Delta}]$ is the discrete delta Dirac matrix as in (47), gives

$$\left[\widetilde{\boldsymbol{\delta\Phi}}\right] = -\left[\tilde{\boldsymbol{\Psi}}\right]^{\mathrm{T}}[\boldsymbol{\delta A}][\boldsymbol{\Phi}]. \tag{58}$$

Because property (48) holds, this implies from (57) that

$$\tilde{\boldsymbol{\Psi}} = \Psi$$

and consequently from (58)

$$\widetilde{\boldsymbol{\delta\Phi}} = \boldsymbol{\delta\Phi}. \tag{59}$$

Thus, the continuous and discrete FEM formulations of the adjoint sensitivities are, in this case, equivalent. If the continuous operators we are dealing with were not self-adjoint, or if the discretized matrices were not symmetric, the continuous and discrete formulations of the sensitivities would differ. If the spatial discretization is done using finite differences rather than finite elements, certain discretizations of

the boundary elements can lead to asymmetric matrices. For example, let us consider a 1-D problem on the domain $\Omega = [0, L]$; let $\Delta x = L/n_x$, where n_x is the number of sub-intervals. We now define a uniform mesh $\Omega_x = \{x_i, 0 \leqslant i \leqslant n_x\}$, where $x_i = \mathrm{i}\Delta x$. The first-order finite difference approximation of the Robin boundary conditions $(\mathrm{d}f/\mathrm{d}x) + rf|_{0,L} = 0$gives the following boundary matrix:

$$\begin{bmatrix} -\frac{1}{\Delta x} + r & \frac{1}{\Delta x} & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ & \cdots & \cdots & \cdots & & \\ & \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & -\frac{1}{\Delta x} & \frac{1}{\Delta x} + r \end{bmatrix} \tag{60}$$

which is clearly asymmetric implying the breakdown of the property (48), with the result that the discrete and continuous formulations will not be equivalent for this finite difference formulation.

5. Vectorized implementation

Both the forward finite element equation (31) and adjoint finite element equation (45) are implemented using a highly vectorized approach in Matlab (Version 6.1 [40]), optimizing for computational speed at the expense of memory. Memory demands are kept manageable through domain decomposition. We outline the main points of our vectorization strategy here.

5.1. Combined forward and adjoint solutions

We can combine the forward and adjoint problems into the following cascade of solutions:

$$\begin{cases} \underset{(n\times n)}{[\mathbf{A}_m]}\,\underset{(n\times d)}{[\boldsymbol{\Psi}_{mm}]} = \underset{(n\times d)}{[\boldsymbol{\Delta}_d]}, \\ \underset{(n\times n)}{[\mathbf{A}_x]}\,[\,\underset{}{\boldsymbol{\Phi}_x},\ \underset{(n\times(1+2d))}{\boldsymbol{\Psi}_{xx}},\ \boldsymbol{\Psi}_{xm}\,] = [\,[\mathbf{M}][\mathbf{S}_x],\ \underset{(n\times(1+2d))}{\boldsymbol{\Delta}_d},\ \underset{}{\left[\mathbf{M}_\beta\right]}[\boldsymbol{\Psi}_{mm}]\,], \\ \underset{(n\times n)}{[\mathbf{A}_m]}\,\underset{(n\times 1)}{[\boldsymbol{\Phi}_m]} = [\,\underset{(n\times 1)}{\left[\mathbf{M}_\beta\right]}[\boldsymbol{\Phi}_x]]. \end{cases} \tag{61}$$

The dimensions are indicated below each matrix, for a system with d detectors, a single excitation light source, and where the dimensionality of the parameterization is n nodes. Note that $[\boldsymbol{\Delta}_d]$ represents the discrete Dirac delta matrix for the detectors; i.e., each column is all zeros except for a 1 in the row corresponding to a detector location. We solve the complex systems of equations in (61) using Matlab's conjugate gradient squared method (cgs.m).

5.2. Vectorization of global FEM matrix assembly

In any finite element method, the global matrices must be assembled from the local element matrices (in the case of tetrahedral elements, these are 4×4 matrices for internal volume elements and 3×3 matrices for triangular surface elements), as indicated by the domain and surface integrals required for computation of Eqs. (34), (35), and (37). In conventional implementations, this assembly is performed by looping over each local element matrix and adding its contents into the appropriate locations in the global matrix. For realistically sized 3-D problems, the loop method of global matrix assembly is a computational bottleneck, especially in parameter estimation problems such as tomography, where the global matrices must be reassembled every time the underlying parameters are updated.

We propose a more computationally efficient approach that uses vectorization to facilitate pipelining in the microprocessor architecture. In this approach, we perform a one-time construction of 3-D matrices that contain the constant kernals K_{type} of each of the distinct types of element matrices as follows:

$$\begin{cases} \underset{(4\times4\times nvel)}{K_{\text{stiffness}}} = \left[\underset{(4\times4)}{\int_{\Omega_e} (\nabla[\mathbf{N}])^{\mathrm{T}} \nabla[\mathbf{N}]} \right]_{(1..nvel)} \ldots^{]}, \\ \underset{(4\times4\times nvel)}{K_{\text{mass}}} = \left[\underset{(4\times4)}{\int_{\Omega_e} [\mathbf{N}]^{\mathrm{T}} [\mathbf{N}]} \right]_{(1..nvel)} \ldots^{]}, \\ \underset{(3\times3\times nsel)}{K_{\text{boundary}}} = \left[\underset{(3\times3)}{\int_{\partial\Omega_e} [\mathbf{N}]^{\mathrm{T}} [\mathbf{N}]} \right]_{(1..nsel)} \ldots^{]}, \end{cases} \tag{62}$$

where the volume and surface integrals appearing in (62) have been evaluated exactly and *nvel* is the number of volume elements, *nsel* is the number of surface elements, Ω_e is the volume of element e, and $\partial\Omega_e$ is the surface area of element e. Note that in Eqs. (62) each 2-D plane of these 3-D matrices corresponds to a (volume or surface) element in the finite element mesh. This idea of pre-calculating the kernal matrices is similar to an approach that has been previously applied to piecewise constant elemental coefficients [41]. We also prespecify index vectors that correctly associate each position in these 3-D kernal matrices with their corresponding nodal locations in the global matrices. For example, I_v and J_v, each of length $(16 * nvel)$, are replicated and expanded from the adjacency matrix denoting which four nodes are connected to each element, as shown in Fig. 1. Expanded index vectors I_s and J_s of length $(9 * nsel)$ are constructed similarly from the matrix specifying which three nodes are connected to each surface element. Whenever it is time to assemble a global matrix, we do the following: (i) convert nodally based coefficients C, where $C \in \{D_{x,m}, k_{x,m}, b_{x,m}, \beta\}$, to elementally based coefficients C^e, (ii) multiply the kernals defined in Eqs. (62) by the appropriate elemental parameter coefficients C^e, as specified in Eqs. (34), (35), (37), and sum them up to yield the 3-D local element matrices, and (iii) use the preconstructed index vectors to add the values from all local element matrices directly into the sparse global matrix.

Construction of a generic local 3-D element matrix L from elemental coefficients C^e and generic kernal matrix K assembly of a generic global matrix G can be written in pseudo-code (using the Matlab notations for element-by-element multiplication (.*), matrix slicing (:), and creation of a sparse matrix) as follows:

$$\begin{cases} \underset{(\text{for } i=1..4,\ \text{for } j=1..4)}{L(i,j,:)} = \underset{(1\times nvel)}{C^e} .* \underset{(1\times nvel)}{K(i,j,:)} \\ \underset{(n\times n)}{G} = \text{sparse}(I_v, J_v, \underset{(1\times 4*4*nvel)}{L(:)}, n, n) \end{cases} \tag{63}$$

and is graphically depicted in Fig. 1.

Assembly of specific global matrices generalizes the method shown in pseudo-code (63) for all three kernal matrices (62). For example, to construct the global matrix A_x (34), the pseudo-code is

$$\begin{cases} \underset{(\text{for } i=1..4,\ \text{for } j=1..4)}{L_v(i,j,:)} = \underset{(1\times nvel)}{D_x} .* \underset{(1\times nvel)}{K_{\text{stiffness}}(i,j,:)} + \underset{(1\times nvel)}{k_x} .* \underset{(1\times nvel)}{K_{\text{mass}}(i,j,:)}, \\ \underset{(\text{for } i=1..3,\ \text{for } j=1..3)}{L_s(i,j,:)} = \underset{(1\times svel)}{b_x} .* \underset{(1\times svel)}{K_{\text{boundary}}(i,j,:)}, \\ \underset{(n\times n)}{A_x} = \text{sparse}(I_v, J_v, \underset{(1\times 4*4*nvel)}{L_v(:)}, n, n) + \text{sparse}(I_s, J_s, \underset{(1\times 3*3*nvel)}{L_s(:)}, n, n). \end{cases}$$

It should be noted that the assumption of homogeneity within elements introduces a small numerical error into the mass and boundary matrices that is linearly proportional to the difference in nodal values attached to each element. If elements are not small enough to be safely approximated as homogeneous, there is no need to convert nodally based coefficients to elementally based coefficients as we have described

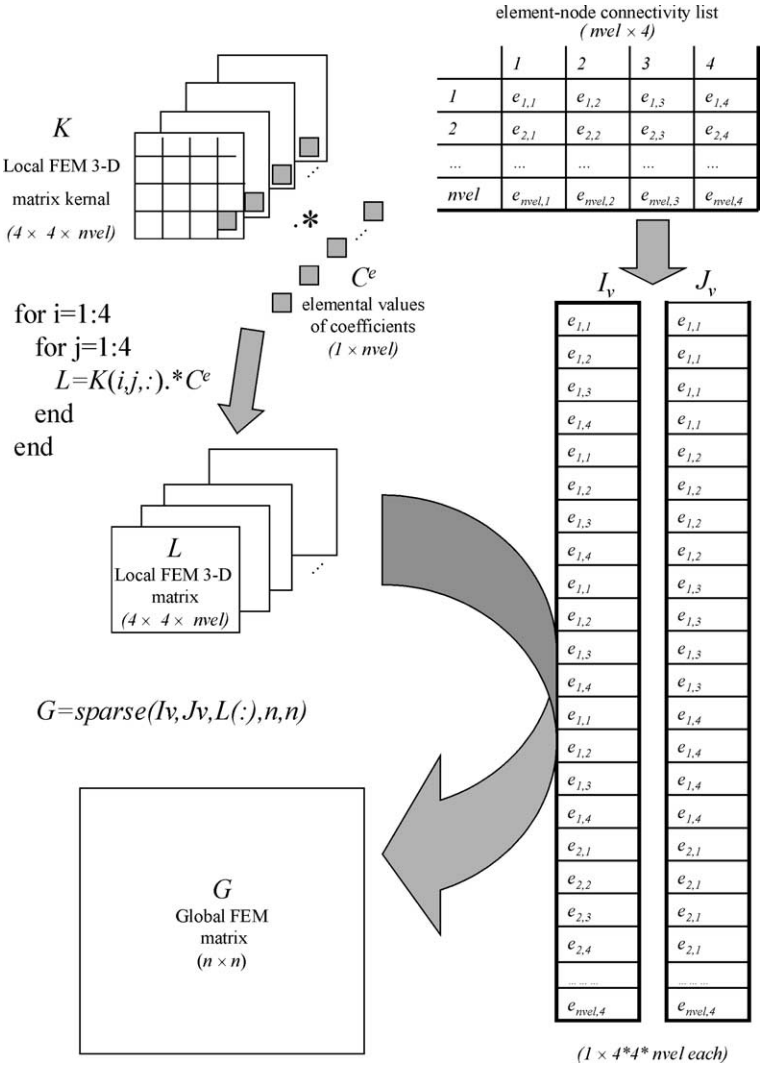

Fig. 1. A depiction of the canonical vectorized operations implemented for assembly of local 4×4 FEM matrices into an $n \times n$ global FEM matrix G. The preconstructed generic kernal matrix K represents either $K_{\text{stiffness}}$ or K_{mass}. The operations for K_{boundary} are similar but with dimensions 3 rather than 4.

here. Rather, four versions of the K_{mass} and K_{boundary} kernal matrices can be preconstructed, one for each of the four basis functions, as follows:

$$\left.\begin{aligned} \underset{(4\times4\times nvel)}{K^i_{\text{mass}}} &= \Big[\int_{\Omega_e} [\mathbf{N}]^{\mathrm{T}} N_i [\mathbf{N}]\Big]_{\substack{(4\times4)\\(1..nvel)}} \text{...}^{]} \\ \underset{(3\times3\times nsel)}{K^i_{\text{boundary}}} &= \Big[\int_{\partial\Omega_e} [\mathbf{N}]^{\mathrm{T}} N_i [\mathbf{N}]\Big]_{\substack{(3\times3)\\(1..nsel)}} \text{...}^{]} \end{aligned}\right\} \quad \text{for } i = 1..4. \tag{64}$$

Each of the matrices in (64) are then multiplied by the corresponding nodal values associated with each element, and summed up to yield the 3-D local element matrices L. This latter method introduces no

numerical errors in the integrations at the expense of slightly greater computational requirements; the trade-off between these two methods must be determined by the applications programmer.

5.3. Vectorization of sensitivity calculations

As indicated by Eqs. (49) and (50), after solution for the adjoint variables $\mathbf{\Psi}_{xx}, \mathbf{\Psi}_{xm}, \mathbf{\Psi}_{mm}$ calculation of any desired sensitivities reduces to a series of matrix multiplications of the general form

$$\underset{(d\times 1)}{\frac{\partial \Phi}{\partial p(i)}} = \underset{(d\times n)}{[\mathbf{\Psi}]^{\mathrm{T}}}\,\underset{(n\times n)}{[\mathbf{G}(\delta C)]}\,\underset{(n\times 1)}{[\mathbf{\Phi}]} \quad \text{for } i = 1..n, \tag{65}$$

where $\mathbf{G}(\delta C)$ is a generic global FEM matrix constructed from the analytical derivatives of generic coefficient C, where $C \in \{D_{x,m}, k_{x,m}, b_{x,m}, \beta\}$, as shown in Eq. (51). We shall use nodally based parameters. Assuming that the parameter p has been discretized for n nodes (so $p(i)$ is the discrete value of the parameter for node i), the sizes of the matrix factors in the terms of the form (65) are as indicated above, for a system with d detectors and 1 source. The sensitivities $\partial\Phi/\partial p(i)$ need to be evaluated for each of the $i = 1..n$ nodes in order to compute each of the n columns of the final $(d \times n)$ Jacobian matrix. One could simply loop through for each node, assembling each matrix of the form $\mathbf{G}(\delta C)$ and evaluating each term of form (65). However, as with the loop method for matrix assembly, this can be very slow for large meshes. We therefore consider an alternative vectorized approach for computing these matrix products over the entire domain.

While each assembled factor $\mathbf{G}(\delta C)$ is a sparse matrix populated by just the element matrices for those elements attached to node i, the number of such elements varies for each node i, and consequently the structure of the sparse matrix varies for each node evaluation. For example, the locations of non-zero entries in global matrices $\mathbf{G}$ for two arbitrary nodes in a breast-shaped finite element mesh are shown in Fig. 2. This variability in the number of elements attached to each node inhibits direct vectorization of the loop in (65) that performs the matrix multiplications for all nodes. On the other hand, each volume element is attached to exactly four nodes. Consequently, we developed a vectorized method for first computing the elemental sensitivities and then converting them to nodal sensitivities. We compute sensitivities for all volume elements (assuming, for simplicity in this pseudo-code, that $\partial b/\partial p = 0$) as follows:

$$\underset{(1\times nvel)}{L(i,j,:)} = \underset{(1\times nvel)}{K(i,j,:)} .* \underset{(1\times nvel)}{\left.\frac{\partial C}{\partial p}\right|_e} \quad \text{for } i = 1..4, \ \text{ for } j = 1..4, \tag{66}$$

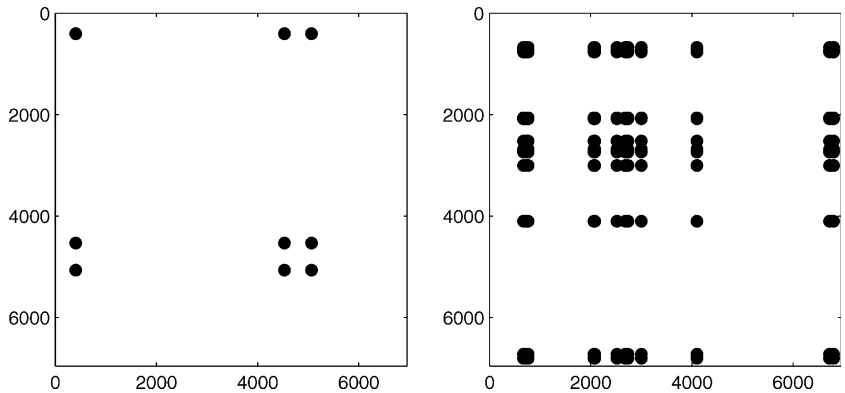

Fig. 2. Locations of non-zero entries in global matrices of two arbitrarily selected nodes in a breast-shaped finite element mesh, highlighting the irregular structure of these sparse matrices.

$$\underset{(1\times nvel)}{T(i,:)} = \sum_{j=1:4} \underset{(1\times nvel)}{L(i,j,:)}.* \underset{(1\times nvel)}{\Phi^e(j,:)} \quad \text{for } i = 1..4, \tag{67}$$

$$\underset{(d\times nvel)}{\frac{\partial \Phi}{\partial p_e}} = \sum_{i=1..4} \underset{(d\times nvel)}{[\Psi^e(i,:,:)]}.* \underset{(d\times nvel)}{\begin{bmatrix} T(i,:) \\ T(i,:) \\ \cdots \\ T(i,:) \end{bmatrix}}, \tag{68}$$

where K is the appropriate $4 \times 4 \times nvel$ kernal matrix from (62), $(\partial C/\partial p)|_e$ is the appropriate analytical derivative of a coefficient C (as identified in Eqs. (49)–(51)) evaluated at each element, L and T are temporary matrices, Φ^e is a $4 \times nvel$ matrix such that the four rows contain the four nodal values of Φ for each of the $nvel$ elements in the FE mesh, and Ψ^e is a $4 \times d \times nvel$ matrix such that the four rows contain the

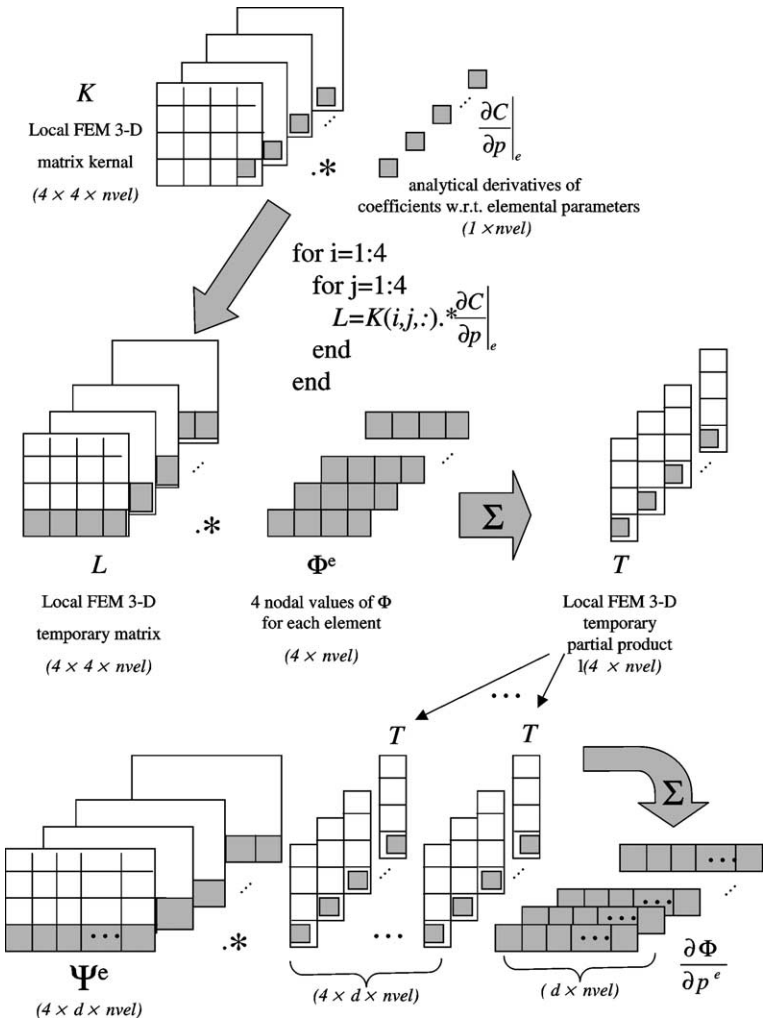

Fig. 3. A pictorial depiction of the vectorized operations in equations (66) (first row), (67) (second row), and (68) (third row). Note that in some cases the graphical orientation of the matrices has been altered to make the depiction easier to follow.

four nodal vectors of Ψ (each of length d), for each of the *nvel* elements in the FE mesh. As before, in Eqs. (66)–(68) we have adopted the Matlab notations for element-by-element multiplication (.*) and matrix slicing along an entire dimension (:). Note that, in contrast to Eq. (65) where we must loop n times ($n =$ number of nodes), in this implementation we now loop only 16 times for the pseudo-code specified in (66) and (67) and only four times for the pseudo-code shown in (68), regardless of the size or geometry of the mesh. All operations are completely vectorized over the largest dimension (*nvel*). These vectorized operations for computing elemental sensitivities are depicted in Fig. 3. In order to convert back to nodal sensitivities, we recognize that the total variation at a node is the sum of one quarter of the variations of each element attached to that node. We thus trivially implement the conversion from elemental to nodal sensitivities by multiplying the elemental sensitivity matrix $(d \times nvel)$ by a precomputed constant sparse

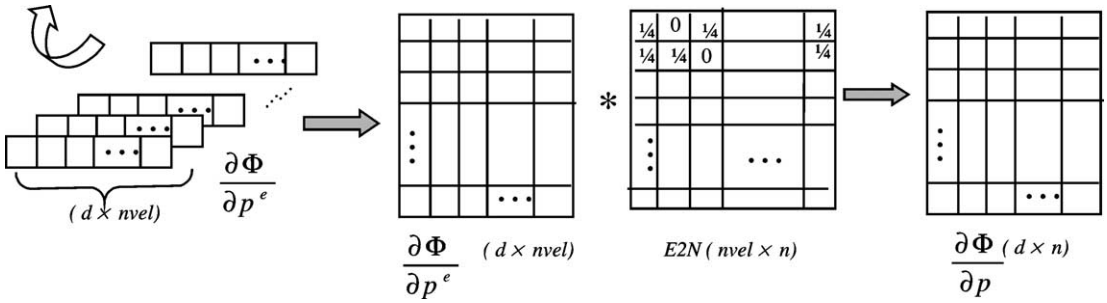

Fig. 4. Elemental sensitivities are converted to nodal sensitivities by simply multiplying by an adjacency matrix E2N that contains the value 0.25 in the positions corresponding to element-node adjacencies.

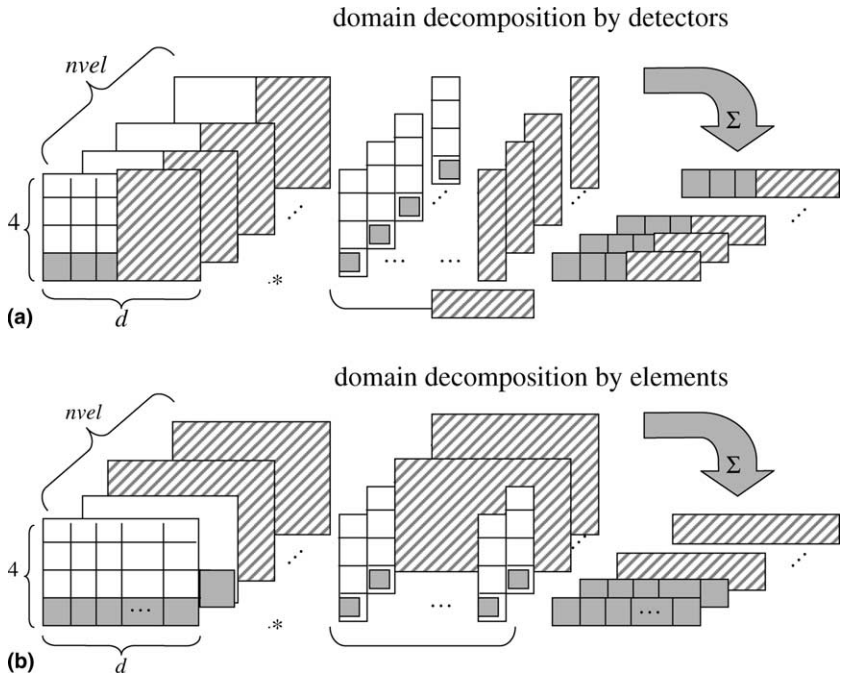

Fig. 5. The depiction of the vectorized operations for Eq. (68) shows that the unhatched and hatched regions may be computed independently, making it easy to decompose the computations into arbitrarily small vectorized subsets of (a) detectors and/or (b) elements.

adjacency matrix ($nvel \times n$) that has the value 0.25 in locations corresponding to element-node adjacencies and is 0 elsewhere (Fig. 4). Experiments confirm that nodal sensitivities computed by the nodal loop in (65) are identical to elemental sensitivities computed by the approach of (67) that are then transformed back to nodal sensitivities. Although the vectorized approach is more space intensive than the loop approach, the memory requirements are easily managed by decomposing the computations for pseudo-code (68) into smaller independent chunks of elements and/or of detectors,as shown in Fig. 5.

The vectorized implementations described here were implemented in Matlab, which facilitates vectorization by providing users with high-level programming abstractions of whole-matrix operations. These vectorized operations are efficiently implemented by calling hardware-optimized BLAS routines [42] that take advantage of pipelining in the microprocessor architecture. Vectorization using the same algorithms described here is also possible when programming in C or Fortran, although there the burden is on the applications programmer to access the appropriate BLAS routines.

6. Numerical experiments

In order to compare the accuracy and efficiency of the various methods for computing the sensitivities we compared results on a $4 \times 8 \times 8$ cm^3 synthetic domain, with coordinate ranges $\langle -2..2, -4..4, -4..4 \rangle$ cm. The size, shape, and optical properties of this domain were chosen to model an existing experimental tissue-mimicking phantom used for experiments in fluorescence tomography using the fluorophore indocyanine green [8], with excitation $\lambda = 785$ nm and emission $\lambda = 830$ nm at $\omega = 100$ MHz modulation frequency. Background optical properties were modeled as homogeneous with the following values: $\mu_{axf} = 0.006$ cm^{-1}, $\mu_{amf} = 0.0846 \cdot \mu_{axf}$, $\mu_{axi} = 0.031$ cm^{-1}, $\mu_{ami} = 0.7987 \cdot \mu_{axi}$, $\mu'_{sx} = 10.95$ cm^{-1}, $\mu'_{sm} = 0.732 \cdot \mu'_{sx}$, $\tau = 0.56$ ns, $\phi = 0.016$, $R_{x,m} = 0.431$ on the top and $R_{x,m} = 0.0222$ on the other five sides.

A laser excitation light source was modeled on the surface at $\langle -2, -1, 1 \rangle$ cm. Simulated measurements of emission fluence were predicted at 50 detector locations on the exterior. Experiments were run using our finite element model with three different levels of discretization in the mesh, as shown in Table 1.

The size of the large mesh is commensurate with realistic mesh sizes for 3-D problems of clinical interest.

6.1. Timing studies

Three sets of timing studies on the small, medium, and large meshes are presented. All timings were performed on a 2.2 GHz Pentium IV with 2 GB RAM.

6.1.1. Vectorization of global finite element matrix assembly

Timings for matrix assembly of the A_x matrix using both the vectorized and loop implementations are shown in Table 2 and Fig. 6(a).

Not only was the vectorized implementation much faster than the loop method, but the speedup due to vectorization continued to increase with increasing mesh size (Table 2, Fig. 6(b). On the large mesh, the speedup of 6.5 means being able to assemble each global matrix in under 5 s, rather than over 30 s. Since

Table 1
Three finite element meshes used to discretize the domain

FE mesh	Spacing (cm)	# Nodes	# Elements
Small	0.5	405	1536
Medium	0.25	2601	12,288
Large	0.125	18,513	98,304

Table 2
Computation times for global matrix assembly using vectorized and loop implementations

FE mesh	Vectorized (s)	Loop (s)	Speedup
Small	0.079	0.171	2.2
Medium	0.578	1.691	2.9
Large	4.828	31.219	6.5

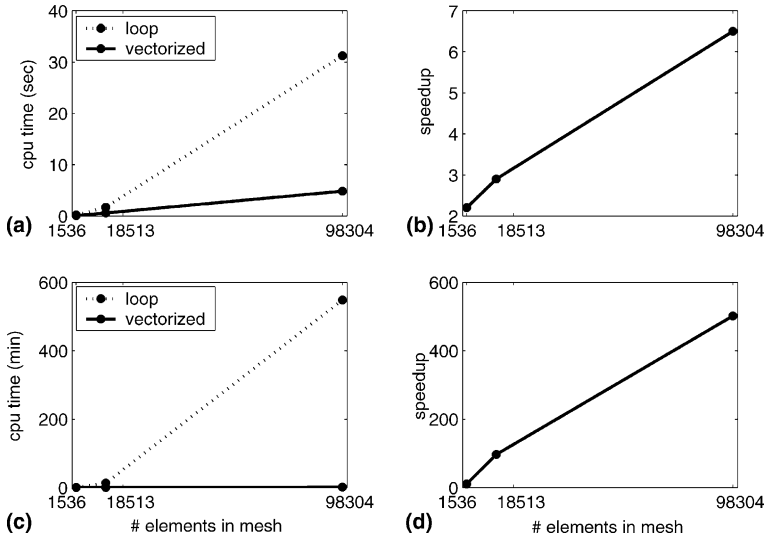

Fig. 6. (a) CPU times for assembly of one global matrix; (b) speedup of global matrix assembly due to vectorization; (c) CPU times for calculation of sensitivity; (d) speedup of sensitivity calculation due to vectorization.

several global matrices must be reassembled during each iteration of a tomography code, vectorization of matrix assembly can indeed yield significant savings in computation time.

6.1.2. Vectorization of adjoint sensitivity calculations

Computations for the adjoint sensitivities for emission fluence with respect to fluorescence absorption $\partial\Phi_m/\partial\mu_{axf}$, from the already computed adjoint field variables $\mathbf{\Psi}$, were timed for both the vectorized (66)–(68) and loop (65) implementations (Table 3, Fig. 6(c). The loop implementation did utilize vectorized matrix assembly for global matrices in $\boldsymbol{\delta A}$ (50), so that these times reflect only the speedup due to (indirectly) vectorizing the loop shown in statement (65).

Appropriate selection of the number of subdomains of elements used in computing the vectorized code is determined by the amount of available memory and the size of the problem (elements times detectors). On

Table 3
Computation times for the sensitivity calculations using vectorized and loop implementations

FE mesh	Vectorized (min)	# Subdomains	Loop (min)	Speedup
Small	0.019	2	0.191	10.6
Medium	0.137	13	13.227	96.5
Large	1.092	104	548.180	502.0

our system, optimal times were achieved with the vectorized code when the domain was decomposed into subdomains of elements such that the number of subdomain elements times the number of detectors was approximately 47,300. The number of subdomains used in timing the vectorized code for each mesh size is indicated in Table 3. Run times for the vectorized code took up to 50% longer when too many or too few subdomains were employed, presumably cache and/or virtual memory was not optimally utilized. Memory constraints precluded the largest problem from running with fewer than three subdomains, but the use of domain decomposition prevents memory becoming a limiting factor, regardless of the mesh size. The non-vectorized code was more memory efficient and thus did not require any decomposition into subdomains.

Vectorization of the matrix products required for computing sensitivities from solutions to the adjoint field variables $\mathbf{\Psi}$ yielded even more dramatic improvements in computation time than did vectorization of matrix assembly. Here, speedups due to vectorization increased from 10.6 on the small mesh to 502 on the large mesh (Table 3, Fig. 6(d). On the large mesh, computation of emission sensitivities with respect to fluorescence absorption $\partial\Phi_m/\partial\mu_{axf}$ took only about 1 min using the vectorized approach, but required over 9 h using the loop approach, even though the loop method was implemented to utilize the vectorized matrix assembly! As with matrix assembly, these sensitivity calculations must be performed during each iteration of a tomography algorithm, so the computational savings of the vectorized approach become enormous.

This *superscalar* execution on a single processor system seems quite remarkable, but is readibly achievable by writing code that takes maximal advantage of low-level microprocessor architecture via hardware-optimized BLAS routines [42]. The vectorization and domain decomposition approaches presented here are also naturally amenable to parallelization, and parallel implementations could offer even further computational benefits.

6.1.3. Adjoint vs. finite difference methods

Numerical results of the emission sensitivities computed by the adjoint method and the finite difference method were identical, to within the accuracy specified for the iterative conjugate gradient squared solver ($1e-9$), for a variety of parameters tested ($\mu_{axf}, \mu_{axi}, \mu'_{sx}, \tau, \phi,$). While our main intention in using the finite difference approach to computing sensitivities was to thus validate the accuracy of our adjoint sensitivity equations, it is also informative to compare computation times in order to appreciate just how much is gained by the adjoint approach. Timings for computing the sensitivities of emission fluence with respect to fluorescence absorption using both the adjoint method and a first-order finite difference method (1) are shown in Table 4. Both methods utilized vectorization of matrix assembly, and the adjoint method also utilized vectorization of the matrix products required for the adjoint sensitivitiy calculations, so that these times reflect the comparison between our fastest implementations of each approach.

The finite difference approach required nearly 6 days of computation time to compute a Jacobian for the large mesh. Recalling that these sensitivities must be recomputed for each source and each iteration of a tomography application, these times are indeed prohibitive. In contrast, our vectorized adjoint method required just over 3 min for computation of this large (50×98304) Jacobian sensitivity matrix, rendering the least-squares tomography problem computationally feasible.

Table 4
Total computation times for calculating sensitivities using our implementations of the adjoint method and the first-order finite difference method

FE mesh	Adjoint (min)	Finite difference (min)	Speedup
Small	0.033	5.697	170.8
Medium	0.290	94.807	327.4
Large	3.247	8557.635	2635.4

6.2. Application of the adjoint sensitivities

We have incorporated our vectorized implementation of the full adjoint sensitivity calculations for complex emission fluence Φ_m with respect to fluorescence absorption μ_{axf} into a 3-D fluorescence tomography code and have used it to reconstruct 3-D maps of fluorescence absorption from emission measurements taken on 3-D tissue-mimicking phantoms, as reported in [9]. In these experiments, a 1087 cc tissue-mimicking phantom was modeled with a 6956 node (34,413 element) FEM mesh, with an average node spacing of 0.5 cm in the 10 cm diameter hemispherical "breast" portion of the model. In one representative experiment, we used a fast version of the Bayesian approximate extended Kalman filter [43] to invert 429 frequency-domain emission measurements of emission fluence collected on a phantom containing a 1 cc embedded fluorescent target with a 100:1 target:background ratio in Indocyanine Green. The correct target location was identifiable after the first iteration (7 min), and this inversion converged (exhibited less that a 1% decrease in model prediction error) after 18 iterations (2 h). For this problem, there were a total of 224 forward solutions required during each iteration (for the predictions of fluence at both excitation (Φ_x) and emission (Φ_m) wavelengths for seven excitation sources, and for the adjoint variables Ψ_{xm} and Ψ_{mm} at 105 detector locations); these solutions were obtained sequentially by calling Matlab's conjugate gradients squared function (cgs.m). Each subsequent iteration of the inverse algorithm required 6.5 min, partioned between the various tasks as follows: 1.0% for building all global FEM matrices, 15.1% for solving for all primary and adjoint variables, 81.3% for performing the adjoint sensitivity calculations, 0.6% for the filter update, and

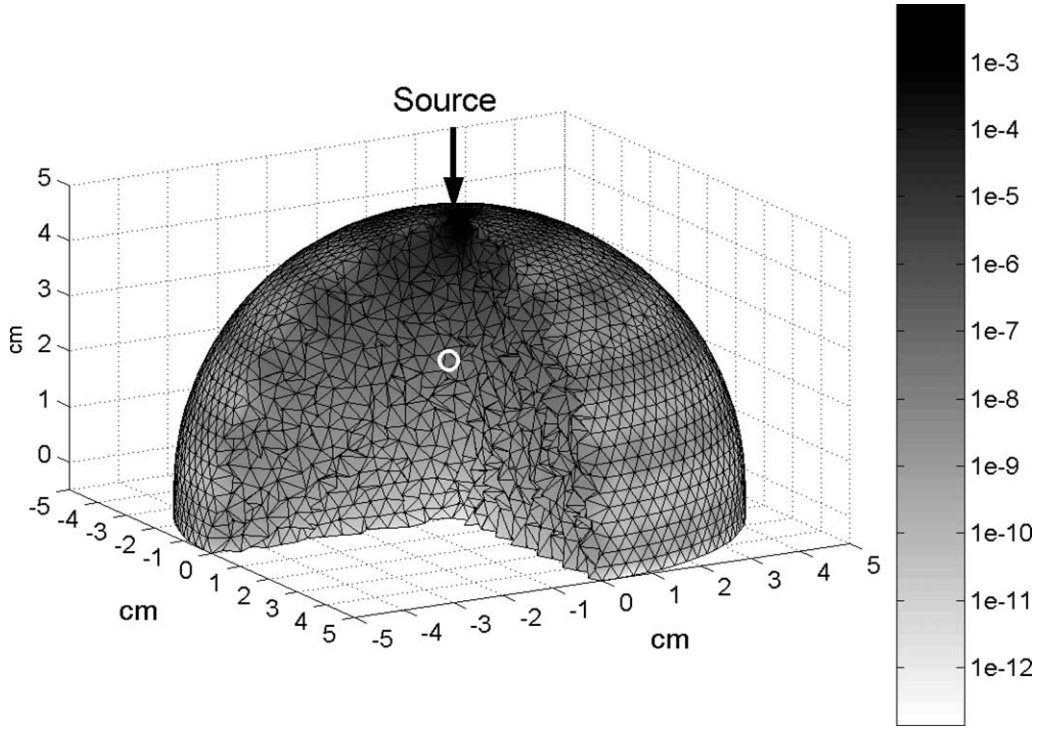

Fig. 7. A 12,657 node (65,509 element) hemispherical 3-D finite element mesh with a uniformly low fluorescence absorption μ_{axf} of 0.006 cm^{-1}. We have plotted the maximum absolute value of complex sensitivity of emission fluence Φ_m with respect to fluorescence absorption μ_{axf} at each node, for the closest of 129 detectors evenly spaced around the surface of the hemisphere, in response to an excitation light source at the top of the hemisphere (i.e., we have plotted the column-wise maxima of the absolute value of $\delta\Phi_m/\delta\mu_{axf}$). The node at $\langle 0, 0, 2.5\rangle$ is marked by a white circle and is further described in Fig. 8. Note that the gray scale for sensitivities is logarithmic.

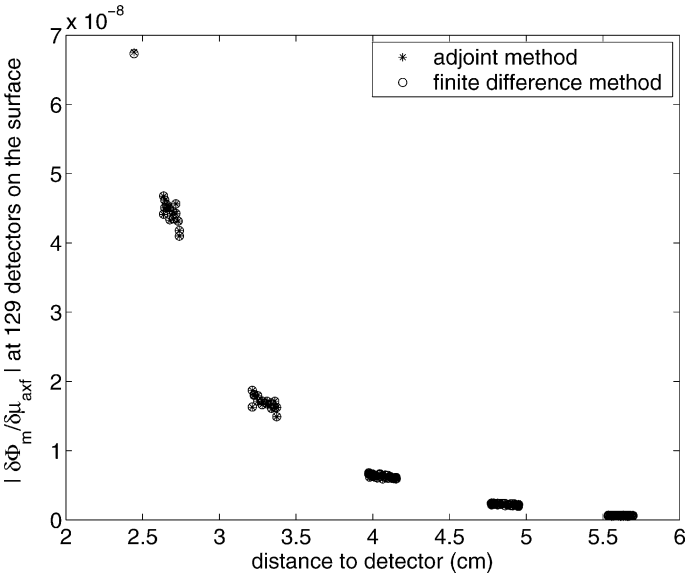

Fig. 8. Absolute values of complex sensitivities, for the data point marked with the white circle in the hemispherical breast model shown in Fig. 7, at all 129 surface detector locations are seen to drop off exponentially with distance between the node at the detectors (since distance to source is constant). Adjoint sensitivities matched finite difference sensitivity approximations to within the accuracy of the solver (1e − 9).

2.0% for other miscellaneous tasks. Clearly, the Jacobian calculations remain the computational bottleneck. This model was decomposed into 120 subdomains to avoid exceeding available memory. Details of the data collection and tomographic reconstruction are in [9]. Elsewhere [44], we show that the previously proposed approximate adjoint method [10] overestimates the sensitivity of emission fluence Φ_m with respect to fluorescence absorption μ_{axf}, and that the full adjoint sensitivities derived here result in more accurate reconstructions of fluorescent targets. Sensitivities for a hemi-spherical model are shown in Figs. 7 and 8. These already low sensitivites are shown to fall off exponentially with distance from source and detectors, contributing to the difficulty of accurately solving the inverse problem of fluorescence tomography in large tissue volumes. The close agreement between finite difference and adjoint sensitivity calculations (Fig. 8) validates the accuracy of the derived adjoint sensitivity equations presented herein.

7. Summary

In order for fluorescence tomography using nonlinear least-squares approaches to become a practical modality for imaging large tissue volumes, there must be an accurate and computationally efficient means of computing coupled sensitivities of excitation and emission fluence with respect to various optical properties that are to be estimated. In this paper, we make several contributions. (1) We develop full adjoint solutions for the coupled complex sensitivities of excitation and emission fluence with respect to any arbitrary optical parameters of interest, (2) we develop a finite element discretization of the adjoint sensitivity equations, (3) we show that the continuous and discrete FEM formulations of the sensitivities are equivalent for the fluorescence equations, (4) we validate our adjoint sensitivity calculations by comparison with sensitivities calculated by finite differences, (5) we describe a vectorized implementation of global finite element matrix assembly that is applicable to any finite element model, and we show that the vectorized implementation is 6.5 times faster (on a 2.2 GHz Pentium IV) than a traditional loop implementation on a

98,304 element mesh, and (6) we propose a novel implementation of adjoint sensitivity calculations that is highly vectorized and uses domain decomposition to keep memory requirements manageable, and show that our vectorized implementation is 502 times faster than a traditional loop implementation on a 98,304 element mesh with 50 detector locations.

The superscalar speedup achieved by our vectorized code underscores how software development has lagged behind hardware development; traditional software implementations often drastically underutilize current microprocessor capabilities. While vectorizing code is often non-intuitive and may entail the use of extra computer memory and program development and debugging time, the resulting runtime speedups can be dramatic and may be well worth the up-front costs for programs with repeated time-intensive calculations. Programming environments like Matlab that provide high-level programmer interfaces to low-level vectorized operations can help to dramatically reduce the program development costs associated with vectorization significantly, but the vectorized algorithms described herein may also be implemented in C or Fortran for significant computational savings, by calling the appropriate optimized BLAS routines. Regardless of the language, the vectorized algorithm design, like parallel algorithm design, must still come from the programmer.

While the methodologies proposed in this paper are developed in the context of fluorescence tomography, they are easily generalized for computing sensitivities of other coupled elliptic equations. For example, in the context of steady state transport, the zeroth-order terms in our general equations can represent the production and decay of chemicals. Similarly, the equations for groundwater flow of immiscible fluids can be modeled by these equations, as can solid mechanics problems of connected materials with different properties. The coupling of heat transport and electric potential also have their basic form described by these equations. We envision numerous applications of our methodology, with only minor changes to the present code, that can be useful in a broad range of physical phenomena.

References

[1] S.R. Arridge, Optical tomography in medical imaging, Inverse Probl. 15 (1999) R41.
[2] B.W. Pogue, S.P. Poplack, T.O. Mcbride, W.A. Wells, K.S. Osterman, U.L. Osterberf, K.D. Paulsen, Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast, Radiology 218 (2001) 261.
[3] N. Shah, A. Cerussi, C. Eker, J. Espinoza, J. Butler, J. Fishkin, R. Hornung, B. Tromberg, Noninvasive functional optical spectroscopy and human breast tissue, Proc. Natl. Acad. Sci. USA 98 (2001) 4420.
[4] D.J. Hawrysz, E.M. Sevick-Muraca, Developments toward diagnostic breast cancer imaging using near-infrared optical measurements and fluorescent contrast agents, Neoplasia 2 (2000) 388.
[5] E.M. Sevick-Muraca, A. Godavarty, J.P. Houston, A.B. Thompson, R. Roy, Near-infrared imaging with fluorescent contrast agents, in: M.-A. Mycek, B.W. Pogue (Eds.), Handbook of Biomedical Fluorescence, Biomedical Optics Series, Marcel-Dekker, Inc., New York, 2003, pp. 445–528.
[6] M.J. Eppstein, D.E. Dougherty, T.L. Troy, E.M. Sevick-Muraca, Biomedical optical tomography using dynamic parameterization and Bayesian conditioning on photon migration measurements, Appl. Opt. 38 (1999) 2138.
[7] M.J. Eppstein, D.E. Dougherty, D.J. Hawrysz, E.M. Sevick-Muraca, Three-dimensional Bayesian optical image reconstruction with domain decomposition, IEEE Trans. Med. Imag. 20 (2001) 147.
[8] M.J. Eppstein, D.J. Hawrysz, A. Godavarty, E.M. Sevick-Muraca, Three-dimensional, Bayesian image reconstruction from sparse and noisy data sets: near-infrared fluorescence tomography, Proc. Natl. Acad. Sci. USA 99 (2002) 9619.
[9] A. Godavarty, M.J. Eppstein, C. Zhang, A.B. Thompson, M. Gurfinkel, S. Theru, E.M. Sevick-Muraca, Fluorescence-enhanced optical imaging in large tissue volumes using a gain modulated ICCD camera, Phys. Med. Biol. (submitted).
[10] J. Lee, E.M. Sevick-Muraca, Fluorescence-enhanced absorption imaging using frequency-domain photon migration: tolerance to measurement error, J. Biomed. Opt. 6 (2001) 58.
[11] J. Lee, 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. Am. A 19 (2002) 759.
[12] M.A. O'Leary, D.A. Boas, B. Chance, A.G. Yodh, Reradiation and imaging of diffuse photon density waves using fluorescent inhomogeneities, J. Lumin. 60/61 (1994) 281.
[13] J. Wu, Y. Wang, L. Perleman, I. Itzkan, R.R. Dasari, M.S. Feld, Time-resolved multichannel imaging of fluorescent objects embedded in turbid media, Opt. Lett. 20 (1995) 489.

[14] J. Chang, R.L. Barbour, H. Graber, R. Aronson, Fluorescence optical tomography, Proc. SPIE 2570 (1995) 59.
[15] M.A. O'Leary, D.A. Boas, X.D. Li, B. Chance, A.G. Yodh, Fluorescence lifetime imaging in turbid media, Opt. Lett. 21 (1996) 158.
[16] D.Y. Paithankar, A.U. Chen, B.W. Pogue, M.S. Patterson, E.M. Sevick-Muraca, Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media, Appl. Opt. 36 (1997) 2260.
[17] J. Wu, L. Perelman, R.R. Dasari, M.S. Feld, Fluorescence tomographic imaging in turbid media using early-arriving photons and Laplace transforms, Proc. Natl. Acad. Sci. USA 94 (1997) 8783.
[18] J. Chang, H.L. Graber, R.L. Barbour, Improved reconstruction algorithm for luminescence when background luminophore is present, Appl. Opt. 37 (1998) 3547.
[19] H. Jiang, Frequency-domain fluorescent diffusion tomography: a finite-element-based algorithm and simulations, Appl. Opt. 37 (1998) 5337.
[20] E.L. Hull, M.G. Nichols, T.H. Foster, Localization of luminescent inhomogeneities in turbid media with spatially resolved measurements of cw diffuse luminescence emittance, Appl. Opt. 37 (1998) 2755.
[21] V. Chenomordik, D. Hattery, I. Gannot, A.H. Gandjbakhche, Inverse method 3-D reconstruction of localized in vivo fluorescence – application to Sjogren syndrome, IEEE J. Selec. Top. Quantum Electron. 54 (1999) 930.
[22] R. Roy, E.M. Sevick-Muraca, Truncated Newton's optimization scheme for absorption and fluorescence optical tomography: Part II. Reconstruction from synthetic measurements, Opt. Exp. 4 (1999) 372.
[23] R. Roy, E.M. Sevick-Muraca, Active constrained truncated Newton method for simple-bound optical tomography, J. Opt. Soc. Am. A 17 (2000) 1627.
[24] Y. Yang, N. Iftimia, Y. Xu, H. Jiang, Frequency-domain fluorescent diffusion tomography of turbid media and in vivo tissues, Proc. SPIE 4250 (2001) 537.
[25] D.J. Hawrysz, M.J. Eppstein, J. Lee, E.M. Sevick-Muraca, Error consideration in contrast-enhanced three-dimensional optical tomography, Opt. Lett. 26 (2001) 704.
[26] R. Roy, E.M. Sevick-Muraca, Three-dimensional unconstrained and constrained image-reconstruction techniques applied to fluorescence, frequency-domain photon migration, Appl. Opt. 40 (2001) 2206.
[27] R. Roy, E.M. Sevick-Muraca, A numerical study of gradient-based nonlinear optimization methods for contrast-enhanced optical tomography, Opt. Exp. 9 (2001) 49.
[28] V. Ntziachristos, R. Weissleder, Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation, Opt. Lett. 26 (2001) 893.
[29] S. Achilefu, H.N. Jimenez, R.B. Dorshow, J.E. Bugaj, E.G. Webb, R.R. Wilhelm, R. Rajagopalan, J. Johler, J.L. Erion, Synthesis, in vitro receptor binding, and in vivo evaluation of fluorescein and carbocyanin peptide-based optical contrast agents, J. Med. Chem. 45 (2002) 2003.
[30] A. Becker, C. Hessenius, K. Licha, B. Ebert, U. Sukowski, W. Semmler, B. Wiedenmann, C. Grotzinger, Receptor-targeted optical imaging of tumors with near-infrared fluorescent ligands, Nat. Biotech. 19 (2001) 327–331.
[31] Y. Lin, R. Weissleder, C.-H. Tung, Novel near-infrared cyanine fluorochromes: synthesis, properties, and bioconjugation, Bioconj. Chem. 13 (2002) 605.
[32] G.I. Marchuk, Adjoint Equations and Analysis of Complex Systems, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.
[33] G.I. Marchuk, V.I. Agoshkov, V.P. Shutyaev, Adjoint Equations and Perturbation Algorithms in Nonlinear Problems, CRC Press, Boca Raton, FL, 1996.
[34] S.R. Arridge, Photon-measurement density functions. Part 1: Analytical forms, Appl. Opt. 34 (1995) 7395.
[35] S.R. Arridge, M. Schweiger, Photon-measurement density functions. Part 2: Finite-element-method calculations, Appl. Opt. 34 (1995) 8026.
[36] M.S. Patterson, B.W. Pogue, Mathematical model for time-resolved and frequency-domain fluorescence spectroscopy in biological tissues, Appl. Opt. 33 (1994) 1963.
[37] E.M. Sevick-Muraca, C.L. Burch, Origin of phosphorescence signals re-emitted from tissues, Opt. Lett. 19 (1994) 1928.
[38] C.L. Hutchinson, T.L. Troy, E.M. Sevick-Muraca, Fluorescence-lifetime determination in tissues or other scattering media from measurement of excitation and emission kinetics, Appl. Opt. 35 (1996) 2325.
[39] H. Grandin Jr., Fundamentals of the Finite Element Method, Macmillan, New York, 1986.
[40] The Mathworks, 24 Prime Park Way, Natick, MA 01760-1500.
[41] N. Polydorides, W.R.B. Lionheart, A Matlab toolkit for three-dimensional electrical impedence tomography: a contribution to the electrical impedence and diffuse optical reconstruction software project, Meas. Sci. Technol. 13 (2002) 1871.
[42] Basic Linear Algebra Subroutines. Available from http://www.netlib.org/blas/.
[43] C. Zhang, M.J. Eppstein, A. Godavarty, E.M. Sevick-Muraca, A hybrid approach to Bayesian image reconstruction, Proc SPIE 4955 (2003), in press.
[44] M.J. Eppstein, F. Fedele, J.P. Laible, C. Zhang, A. Godavarty, E.M. Sevick-Muraca, A comparison of exact and approximate adjoint sensitivities in fluorescence tomography, IEEE Trans. Med. Imag. (submitted).