Fluorescence Optical Diffusion Tomography

Adam B. Milstein[a], Seungseok Oh[a], Kevin J. Webb[a],

Charles A. Bouman[a], Quan Zhang[b], David A. Boas[b], and

R. P. Millane[c]

[a]School of Electrical and Computer Engineering,

Purdue University, West Lafayette, IN, 47907-1285

[b]NMR Center, Massachusetts General Hospital,

Harvard Medical School, Charlestown, MA 02129

[c]Dept. of Electrical and Computer Engineering,

University of Canterbury, Christchurch, New Zealand

Correspondence to: *Kevin J. Webb*

School of Electrical and Computer Engineering

Purdue University

West Lafayette, IN 47907-1285

webb@ecn.purdue.edu

Tel.: 765-494-3373, Fax: 765-494-2706

Date: February 5, 2003

OCIS codes: 170.6280, 290.7050, 100.3010, 100.3190, 100.6950, 170.3010, 290.3200

Abstract - A nonlinear, Bayesian optimization scheme is presented for reconstructing fluorescent yield and lifetime, the absorption coefficient, and the diffusion coefficient in turbid media, such as biological tissue. The method utilizes measurements at both the excitation and emission wavelengths for reconstructing all unknown parameters. The effectiveness of the reconstruction algorithm is demonstrated by simulation and by application to experimental data from a tissue phantom containing the fluorescent agent indocyanine green.

1. INTRODUCTION

Optical diffusion tomography (ODT) is emerging as a powerful tissue imaging modality.[1,2] In ODT, images are comprised of the spatially dependent absorption and scattering properties of the tissue. Boundary measurements from several sources and detectors are used to recover the unknown parameters from a scattering model described by a partial differential equation. Contrast between the properties of diseased and healthy tissue might then be used in clinical diagnosis. In principle, sinusoidally modulated, continuous-wave (CW), or pulsed excitation light is launched into the biological tissue, where it undergoes multiple scattering and absorption before exiting. The measured intensity and phase (or delay) information may be used to reconstruct three-dimensional maps of the absorption and scattering properties by optimizing a fit to diffusion model computations. As a result of the nonlinear dependence of the diffusion equation photon flux on the unknown parameters and the inherently three-dimensional nature of photon scattering, this inverse problem is computationally intensive and must be solved iteratively.

A relatively modest intrinsic contrast between the optical parameters of diseased and healthy breast tissue has been reported in some studies.[3,4] The use of exogenous fluorescent agents has the potential to improve the contrast and thus to facilitate early diagnosis. In recent years, the use of fluorescent indicators as exogenous contrast agents for *in vivo* imaging of tumors with near-infrared (NIR) or visible light has shown great promise, attracting considerable interest.[5–14] In experimental studies with animal subjects,[5–7,9,10,13,14] fluorescence has been successfully used to visualize cancerous tissue *in vivo* near the skin surface. In addition, Ntziachristos *et al.*[12] have used optical diffusion tomography after indocyanine green (ICG) injection to image the absorption of a malignant breast tumor in a human subject. The injected fluorophore may preferentially accumulate in diseased tissue due to increased blood flow from tumor neovascularization.[9] Alternatively, the agent may have different decay properties in diseased tissue, which could be useful in localizing

tumors independently of fluorophore concentration.[7] In addition, contrast between tumors and surrounding tissue may be substantially improved by the use of diagnostic agents that selectively target receptors specific to cancer cells.[8,10,13,14]

In frequency-domain fluorescence optical diffusion tomography, sinusoidally modulated light at the fluorophore's excitation wavelength is launched into the tissue. The excited fluorophore, when it decays to the ground state, emits light at a longer (emission) wavelength, and this emission is measured by an array of detection devices. These emission data are then used to perform a volumetric reconstruction of the yield (a measure of the fluorescence efficiency) and the lifetime (the fluorescent decay parameter). However, the multiple scattering in tissue complicates the reconstruction.[15,16] The emission intensity of the fluorophore is proportional to the optical intensity at the excitation wavelength at that position, which depends, in turn, on the optical parameters of the scattering domain at the excitation wavelength. A rigorous reconstruction of fluorescence property maps should also therefore include reconstructions of absorption and scattering parameters at the excitation and emission wavelengths. In addition, reconstruction of the unknown absorption and scattering coefficients by use of ODT can function as an adjunct image to the fluorescence image in screening for tumors.

Fluorescence imaging simulations with three-dimensional (3-D)[17] and two-dimensional (2-D)[18–20] geometries have reconstructed fluorescence yield and lifetime parameters. These simulations have generally assumed that the absorption and scattering parameters are known in advance, except for Roy and Sevick-Muraca,[17] who also reconstructed the excitation wavelength absorption. In an early experimental result, Chang *et al.*[21] used a transport theory model to reconstruct fluorescent yield in a heterogeneous tissue phantom containing Rhodamine 6G. Their study used CW data recorded in a 2-D plane geometry. Recently, Ntziachristos and Weissleder[22] used a normalized Born approximation to reconstruct 3-D, fluorescent heterogeneities containing the NIR cyanine dye

Cy5.5 embedded in a tissue phantom. Under the assumption of known background optical properties and absorbers limited to a perturbative regime, their technique can circumvent the need for recording background measurements before contrast agent administration.

The development of nonlinear inversion methods for optical diffusion tomography is necessary due to the fundamentally limited accuracy of methods which linearize the forward model.[23] Previously, we have presented a nonlinear Bayesian approach[24–26] and shown that it produces high quality images compared to previous methods such as the distorted Born iterative method.[27] The method formulates the inversion as the optimization of an objective function which incorporates a model of the detection system and *a priori* knowledge about the image properties. We have found that a neighborhood regularization scheme used in a Bayesian framework reduces artifacts characteristic of previous approaches which impose a penalty on the norm of the image updates.[24] The inversion can be made more computationally efficient by multigrid techniques.[25]

Here, we extend our previous approach to include fluorescence yield and lifetime in the inverse problem. We present a new inversion algorithm and a measurement scheme for reconstructing all the unknown fluorescence, absorption, and diffusion parameters. Numerical simulations validate the scheme and demonstrate its computational efficacy. We use the method to image a spherical heterogeneity in a tissue phantom by use of transmission data collected by a CW imaging device. The heterogeneity contains ICG, a fluorescent diagnostic agent approved by the FDA for use in the NIR range, where biomedical imaging with light is most practical.

2. FLUORESCENCE DIFFUSION TOMOGRAPHY PROBLEM

The transport of modulated light (at modulation angular frequency ω, i.e., $e^{j\omega t}$ variation) in a fluorescent, highly scattering medium with an external source at the excitation wavelength is modeled

by using the coupled diffusion equations:[15,16,28]

$$\nabla \cdot [D_x(\mathbf{r})\nabla\phi_x(\mathbf{r},\omega)] - [\mu_{a_x}(\mathbf{r}) + j\omega/c]\,\phi_x(\mathbf{r},\omega) = -\delta(\mathbf{r}-\mathbf{r}_{s_k}) \quad (1)$$

$$\nabla \cdot [D_m(\mathbf{r})\nabla\phi_m(\mathbf{r},\omega)] - [\mu_{a_m}(\mathbf{r}) + j\omega/c]\,\phi_m(\mathbf{r},\omega) = -\phi_x(\mathbf{r},\omega)\eta\mu_{a_f}(\mathbf{r})\frac{1-j\omega\tau(\mathbf{r})}{1+[\omega\tau(\mathbf{r})]^2}, \quad (2)$$

where the subscripts x and m, respectively, denote excitation and emission wavelengths λ_x and λ_m, $\phi(\mathbf{r},\omega)$ is the complex modulation envelope of the photon flux, $\delta(\mathbf{r})$ is the Dirac function, and $\mathbf{r}_{s_k}$ is the location of the excitation point source. We also assume single exponential decay in this model. The optical parameters are the diffusion coefficients $D(\mathbf{r})$ and the absorption coefficients $\mu_a(\mathbf{r})$. The fluorescence parameters are the lifetime $\tau(\mathbf{r})$ and the fluorescent yield $\eta\mu_{a_f}(\mathbf{r})$. The fluorescent yield incorporates the fluorophore's quantum efficiency η (which depends on the type of fluorophore and the chemical environment) and its absorption coefficient, μ_{a_f} (which depends on the fluorophore concentration). Note the right hand side of (2), where the light absorbed by fluorophores and subsequently emitted at the emission wavelength, is incorporated into an effective source term. In the case of an external point source at the emission wavelength, the flux is governed by

$$\nabla \cdot [D_m(\mathbf{r})\nabla\phi_m(\mathbf{r},\omega)] - [\mu_{a_m}(\mathbf{r}) + j\omega/c]\,\phi_m(\mathbf{r},\omega) = -\delta(\mathbf{r}-\mathbf{r}_{s_k}). \quad (3)$$

In the most general case, the unknown parameters in (1) and (2) are μ_{a_x}, μ_{a_m}, D_x, D_m, τ, and $\eta\mu_{a_f}$. Reconstructions of the D_x and μ_{a_x} images may be obtained using data from sources and detectors at the excitation wavelength λ_x. Similarly, D_m and μ_{a_m} may be obtained using data from sources and detectors at the emission wavelength λ_m. Finally, having found these parameters, using sources at λ_x and detectors filtered at λ_m will yield the fluorescence parameters. Figure 1 depicts this measurement approach schematically.

After discretizing the domain into N voxels of equal size, the unknown parameters can be regarded as three image vectors, each corresponding to a measurement set. Let $\mathbf{r}_i$ denote the position

of the i^{th} voxel centroid, i.e., the location of a node in a Cartesian finite difference representation of (1)-(3). We define the image vectors as

$$\begin{aligned}
\mathbf{x}_x &= \begin{bmatrix} \mathbf{x}_{xa} \\ \\ \mathbf{x}_{xb} \end{bmatrix} = [\mu_{a_x}(\mathbf{r}_1) \cdots \mu_{a_x}(\mathbf{r}_N), D_x(\mathbf{r}_1) \cdots D_x(\mathbf{r}_N)]^T \\
\mathbf{x}_m &= \begin{bmatrix} \mathbf{x}_{ma} \\ \\ \mathbf{x}_{mb} \end{bmatrix} = [\mu_{a_m}(\mathbf{r}_1) \cdots \mu_{a_m}(\mathbf{r}_N), D_m(\mathbf{r}_1) \cdots D_m(\mathbf{r}_N)]^T \\
\mathbf{x}_f &= \begin{bmatrix} \mathbf{x}_{fa} \\ \\ \mathbf{x}_{fb} \end{bmatrix} = [\gamma(\mathbf{r}_1) \cdots \gamma(\mathbf{r}_N), \tau(\mathbf{r}_1) \cdots \tau(\mathbf{r}_N)]^T,
\end{aligned} \tag{4}$$

where the subscript f denotes the fluorescence image and the superscript T denotes the transpose operation. Note that the three image vectors are each of size $2N$, consisting of two unknown parameter vectors of size N. In addition, we reparameterize the fluorescence unknowns $\{\eta\mu_{a_f}, \tau\}$ to $\{\gamma, \tau\}$ using

$$\gamma(\mathbf{r}, \omega) = \eta\mu_{a_f}(\mathbf{r}) \frac{1}{1 + [\omega\tau(\mathbf{r})]^2}, \tag{5}$$

which, when substituted into (2), gives

$$\nabla \cdot [D_m(\mathbf{r}) \nabla \phi_m(\mathbf{r}, \omega)] - [\mu_{a_m}(\mathbf{r}) + j\omega/c] \, \phi_m(\mathbf{r}, \omega) = -\phi_x(\mathbf{r}, \omega)\gamma(\mathbf{r}, \omega) \left[1 - j\omega\tau(\mathbf{r})\right]. \tag{6}$$

As explained in Appendix A, this new parameterization is useful because, in a sequential optimization scheme, it takes advantage of the inherent linearity of the fluorescence inverse problem while allowing regularization to be applied to τ directly. The sets of flux measurements corresponding to the above image vectors may be defined, respectively, as $\mathbf{y}_x$, $\mathbf{y}_m$, and $\mathbf{y}_f$.

3. INVERSION

The estimation of each of the unknown images $\{\mathbf{x}_x, \mathbf{x}_m, \mathbf{x}_f\}$ from the corresponding observations $\{\mathbf{y}_x, \mathbf{y}_m, \mathbf{y}_f\}$ is an ill-posed, typically underdetermined, inverse problem. As in previous work,[24–26,29]

we address this by formulating the inverse problem in a Bayesian framework. This framework allows the incorporation of *a priori* information, and it encapsulates all available information about the problem model into an objective function to be optimized. Let $\mathbf{x}$ denote one of the images of (4), and let $\mathbf{y}$ denote its corresponding observations. We use Bayes' rule to compute the maximum *a posteriori* (MAP) estimate, given by

$$\hat{\mathbf{x}}_{MAP} \quad = \quad \arg\max_{\mathbf{x}\geq 0}\{\ p(\mathbf{y}|\mathbf{x}) + p(\mathbf{x})\ \}, \tag{7}$$

where $p(\mathbf{y}|\mathbf{x})$ is the data likelihood and $p(\mathbf{x})$ is the prior density for the image. The data likelihood can be formed from a Gaussian model by considering, for example, the physical properties of a photocurrent shot noise-limited measurement system.[24] This gives

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(\pi\alpha)^P|\mathbf{\Lambda}|^{-1}} \exp\left[-\frac{||\mathbf{y}-\mathbf{f}(\mathbf{x})||^2_{\mathbf{\Lambda}}}{\alpha}\right], \tag{8}$$

where P is the number of measurements, $\mathbf{f}$ is the appropriate forward operator, α is a scalar parameter that scales the noise variance, and, for an arbitrary vector $\mathbf{w}$, $||\mathbf{w}||^2_{\mathbf{\Lambda}} = \mathbf{w}^H\mathbf{\Lambda}\mathbf{w}$ (where H denotes Hermitian transpose), and $\alpha\mathbf{\Lambda}^{-1}$ is the covariance matrix. In a small signal shot noise model, the measurements are independent and normally distributed with a mean equal to the exact (noiseless) measurement and a variance proportional to the exact measurement at a modulation frequency of zero (DC). Following Ye *et al.*,[24] we approximate the DC flux for the i^{th} datum as $|y_i|$. The resulting covariance matrix is given by

$$\alpha\mathbf{\Lambda}^{-1} \quad = \quad diag[|y_1|, |y_2|, ...|y_P|]. \tag{9}$$

For the prior density $p(\mathbf{x})$, we use the generalized Gaussian Markov random field (GGMRF) model, which enforces smoothness in the solution while preserving sharp edge transitions.[24,30] For each node (representing a voxel) inside the image, we form a three-dimensional neighborhood from the 26 adjacent nodes. Let $\mathbf{x}^T = [\mathbf{x}_a^T \quad \mathbf{x}_b^T]$, as in (4). Assuming independence of $\mathbf{x}_a$ and $\mathbf{x}_b$, the

density function is given by

$$p(\mathbf{x}) = p(\mathbf{x}_a) \cdot p(\mathbf{x}_b) \tag{10}$$

$$= \left[\frac{1}{\sigma_a^N z(p_a)} \exp\left(-\frac{1}{p_a \sigma_a^{p_a}} \sum_{\{i,j\} \in \mathcal{N}_a} b_{i-j} \left| x_i - x_j \right|^{p_a} \right) \right] \cdot \left[\frac{1}{\sigma_b^N z(p_b)} \exp\left(-\frac{1}{p_b \sigma_b^{p_b}} \sum_{\{i,j\} \in \mathcal{N}_b} b_{i-j} \left| x_i - x_j \right|^{p_b} \right) \right], \tag{11}$$

where the subscripts a and b have the same meaning as in (4), x_i denotes the i^{th} node of $\mathbf{x}$, the set $\mathcal{N}$ consists of all pairs of neighboring nodes, and b_{i-j} is the weighting coefficient corresponding to the i^{th} and j^{th} nodes. The coefficients b_{i-j} are assigned to be inversely proportional to the node separation in a cube-shaped node layout, with the requirement that that $\sum_j b_{i-j} = 1$. The constants p and σ control the shape and scale of the distribution, and the factor $z(p)$ is a normalization term.

As in previous work,[25] we incorporate α into the inverse problem as an unknown for each image. We have found that this tends to improve the robustness and speed of convergence. As a result, we perform a joint MAP estimation of both $\mathbf{x}$ and α for each image:

$$\hat{\mathbf{x}}_x = \arg \max_{\mathbf{x}_x \geq \mathbf{0}, \alpha_x} \{ \, p(\mathbf{x}_x | \mathbf{y}_x, \alpha_x) \, \} \tag{12}$$

$$\hat{\mathbf{x}}_m = \arg \max_{\mathbf{x}_m \geq \mathbf{0}, \alpha_m} \{ \, p(\mathbf{x}_m | \mathbf{y}_m, \alpha_m) \, \} \tag{13}$$

$$\hat{\mathbf{x}}_f = \arg \max_{\mathbf{x}_f \geq \mathbf{0}, \alpha_f} \{ \, p(\mathbf{x}_f | \mathbf{y}_f, \alpha_f, \hat{\mathbf{x}}_x, \hat{\mathbf{x}}_m) \, \}. \tag{14}$$

The estimations of $\mathbf{x}_x$ and $\mathbf{x}_m$ are performed independently of each other, using (1) and (3) as the respective forward models. Subsequently, these estimates are incorporated into the coupled diffusion equations (1) and (2) to estimate $\mathbf{x}_f$.

Let $\mathbf{x}$ and α correspond to one of the images in (12)-(14). Ye *et al.*[25] showed that the above reconstructions are equivalent to maximizing the log posterior probability $l(\mathbf{x})$, which can be derived using (7), (8), and (11):

$$l(\mathbf{x}) = -P \ln ||\mathbf{y} - \mathbf{f}(\mathbf{x})||_{\mathbf{\Lambda}}^2 - \frac{1}{p_a \sigma_a^{p_a}} \sum_{\{i,j\} \in \mathcal{N}_a} b_{i-j} \left| x_i - x_j \right|^{p_a} - \frac{1}{p_b \sigma_b^{p_b}} \sum_{\{i,j\} \in \mathcal{N}_b} b_{i-j} \left| x_i - x_j \right|^{p_b} \tag{15}$$

Optimizing $l(\mathbf{x})$ may be implemented by alternating closed form updates of $\hat{\alpha}$ with updates of $\hat{\mathbf{x}}$:[25]

$$\hat{\alpha} = \frac{1}{P} ||\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}})||^2_{\mathbf{\Lambda}} \tag{16}$$

$$\hat{\mathbf{x}} \simeq \arg\max_{\mathbf{x}\geq 0}\{ \ln p(\mathbf{y}|\mathbf{x}, \hat{\alpha}) + \ln p(\mathbf{x}|\hat{\alpha})\}, \tag{17}$$

where $\simeq$ implies an update iteration, rather than a full optimization. The $\hat{\mathbf{x}}$ updates represent more computationally expensive steps toward optimizing (7) than the $\hat{\alpha}$ updates. For each image, we form an objective function from (8) and (11):

$$c(\mathbf{x}_x, \hat{\alpha}_x) = \frac{1}{\hat{\alpha}_x} ||\mathbf{y}_x - \mathbf{f}_x(\mathbf{x}_x)||^2_{\mathbf{\Lambda}_x} + \frac{1}{p_{xa}\sigma_{xa}^{p_{xa}}} \sum_{\{i,j\}\in\mathcal{N}_{xa}} b_{i-j} \left| x_{xa_i} - x_{xa_j} \right|^{p_{xa}} + \frac{1}{p_{xb}\sigma_{xb}^{p_{xb}}} \sum_{\{i,j\}\in\mathcal{N}_{xb}} b_{i-j} \left| x_{xb_i} - x_{xb_j} \right|^{p_{xb}} \tag{18}$$

$$c(\mathbf{x}_m, \hat{\alpha}_m) = \frac{1}{\hat{\alpha}_m} ||\mathbf{y}_m - \mathbf{f}_m(\mathbf{x}_m)||^2_{\mathbf{\Lambda}_m} + \frac{1}{p_{ma}\sigma_{ma}^{p_{ma}}} \sum_{\{i,j\}\in\mathcal{N}_{ma}} b_{i-j} \left| x_{ma_i} - x_{ma_j} \right|^{p_{ma}} + \frac{1}{p_{mb}\sigma_{mb}^{p_{mb}}} \sum_{\{i,j\}\in\mathcal{N}_{mb}} b_{i-j} \left| x_{mb_i} - x_{mb_j} \right|^{p_{mb}} \tag{19}$$

$$c(\mathbf{x}_f, \hat{\mathbf{x}}_x, \hat{\mathbf{x}}_m, \hat{\alpha}_f) = \frac{1}{\hat{\alpha}_f} ||\mathbf{y}_f - \mathbf{f}_f(\mathbf{x}_f, \hat{\mathbf{x}}_x, \hat{\mathbf{x}}_m)||^2_{\mathbf{\Lambda}_f} + \frac{1}{p_{fa}\sigma_{fa}^{p_{fa}}} \sum_{\{i,j\}\in\mathcal{N}_{fa}} b_{i-j} \left| x_{fa_i} - x_{fa_j} \right|^{p_{fa}} + \frac{1}{p_{fb}\sigma_{fb}^{p_{fb}}} \sum_{\{i,j\}\in\mathcal{N}_{fb}} b_{i-j} \left| x_{fb_i} - x_{fb_j} \right|^{p_{fb}}. \tag{20}$$

The variables have the same meaning as in (8) and (11), and their subscripts have the same meaning as in (4). Note that forward operator $\mathbf{f}_f$ is a function of $\mathbf{x}_f$ and the estimates $\hat{\mathbf{x}}_x$ and $\hat{\mathbf{x}}_m$. In principle, one could jointly optimize (18)-(20) over $\mathbf{x}_x$, $\mathbf{x}_m$, and $\mathbf{x}_f$, but for computational simplicity, we first optimize (18) and (19) and subsequently incorporate the estimates into (20). With the objective functions (18)-(20) established, an optimization algorithm to minimize these costs is needed, which is described in the next section.

4. ITERATIVE COORDINATE DESCENT OPTIMIZATION

The optimizations of (18)-(20) are performed using the iterative coordinate descent (ICD) algorithm,[24,26,31] a sequential single-site update scheme similar to the Gauss-Seidel method used in

other problems. One ICD scan consists of forming a local quadratic approximation to the cost function, followed by an update of each image element individually to minimize the approximate objective function. On each subsequent scan, the Fréchet derivative of the nonlinear forward operator is recomputed, and a new quadratic approximation is made.

Once again, let $\mathbf{x}$ denote one of the three images to be optimized. During the scan, the individual voxels of $\mathbf{x}$ are sequentially updated in random order. At the beginning of the scan, $\mathbf{f}(\mathbf{x})$ is first expressed using a Taylor expansion as

$$||\mathbf{y} - \mathbf{f}(\mathbf{x})||^2_{\mathbf{\Lambda}} \simeq ||\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}) - \mathbf{F}'(\hat{\mathbf{x}})\mathbf{\Delta x}||^2_{\mathbf{\Lambda}} \ , \tag{21}$$

where $\mathbf{\Delta x} = \mathbf{x} - \hat{\mathbf{x}}$, and $\mathbf{F}'(\hat{\mathbf{x}})$ represents the Fréchet derivative of $\mathbf{f}(\mathbf{x})$ with respect to $\mathbf{x}$ at $\mathbf{x} = \hat{\mathbf{x}}$. Using (21), we formulate the approximate cost function

$$c(\mathbf{x}, \hat{\alpha}) \simeq \frac{1}{\hat{\alpha}} ||\mathbf{z} - \mathbf{F}'(\hat{\mathbf{x}})\mathbf{x}||^2_{\mathbf{\Lambda}} + \frac{1}{p_a \sigma_a{}^{p_a}} \sum_{\{i,j\} \in \mathcal{N}_a} b_{i-j} \left| x_i - x_j \right|^{p_a} + \frac{1}{p_b \sigma_b{}^{p_b}} \sum_{\{i,j\} \in \mathcal{N}_b} b_{i-j} \left| x_i - x_j \right|^{p_b} \ , \tag{22}$$

where

$$\mathbf{z} \ = \ \mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}) + \mathbf{F}'(\hat{\mathbf{x}})\hat{\mathbf{x}} \ . \tag{23}$$

With the other image elements fixed, the ICD update for $\hat{x}_i$ is given by

$$\hat{x}_i \ = \ \arg\min_{x_i \geq 0} \left\{ \frac{1}{\hat{\alpha}} \left|\left| \mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}) - \left[\mathbf{F}'(\hat{\mathbf{x}})\right]_{*(i)} (x_i - \hat{x_i}) \right|\right|^2_{\mathbf{\Lambda}} + \frac{1}{p\sigma^p} \sum_{j \in \mathcal{N}_i} b_{i-j} \left| x_i - \hat{x}_j \right|^p \right\} \ , \tag{24}$$

where $\left[\mathbf{f}'(\hat{\mathbf{x}})\right]_{*(i)}$ is the i^{th} column of the Fréchet matrix and $\mathcal{N}_i$ is the set of nodes neighboring node i, and p and σ are chosen appropriately from $\{p_a, p_b\}$ and $\{\sigma_a, \sigma_b\}$. This one-dimensional minimization is solved by use of a simple half-interval search.[24] The Fréchet matrices used for each image are given in Appendix A. Appendix B summarizes the ICD optimization algorithm in pseudocode form.

Previously, we have found that multiresolution techniques can reduce the computational burden and improve robustness of convergence for the optical diffusion tomography problem.[25] Hence, for

large computational domains, it may be beneficial to perform several ICD scans at a reduced resolution followed by interpolation as an initialization step for the full-resolution problem.

5. SIMULATIONS

Figure 2 shows cross section images of a $17.3 \times 17.3 \times 6$ cm tissue phantom having background values $\mu_{a_{x,m}} = 0.01$ cm^{-1}, $D_{x,m} = 0.047$ cm, $\tau = 0$ ns and $\eta\mu_{a_f} = 0$ cm^{-1}. A slightly off-center spherical heterogeneity with diameter of roughly 3 cm is present, with $\mu_{a_x} = 0.05$ cm^{-1}, $\mu_{a_m} = 0.01$ cm^{-1}, D_x and $D_m = 0.30$ cm, $\tau = 0.55$ ns, and $\eta\mu_{a_f} = 0.02$ cm^{-1}. Figure 2(g) shows the location and size of the fluorophore as the $\eta\mu_{a_f} = 0.01$ cm^{-1} isosurface. As shown in Figure 3, the bottom face of the domain contains 16 sources (modulated at 70 MHz) arranged in a 4×4 grid pattern. On the top face, 16 detectors are placed in an identical grid. Using multigrid finite differences[32] to solve the diffusion equations, we generated synthetic measurements. Additive noise was introduced using the approximate shot noise model of (8) and (9), giving an average signal-to-noise ratio (SNR) of 34 dB and a maximum SNR of 41 dB. In the forward solution, an extrapolated zero-flux boundary condition[33] was used to model the free space absorbing boundaries.

For each of the $\mathbf{x}_x$, $\mathbf{x}_m$, and $\mathbf{x}_f$ inversions, 20 ICD iterations at a resolution of $17 \times 17 \times 9$ nodes, followed by 20 ICD iterations at a resolution of $33 \times 33 \times 17$ nodes, were performed. For the nonlinear $\mathbf{x}_x$ and $\mathbf{x}_m$ problems, multigrid finite differences were used to solve the forward model prior to each ICD image update. During the inversions, the log posterior probability was evaluated as the convergence criterion. For each image, convergence (with subsequent iterations changing the images very little) was obtained in approximately 10 minutes of computation on a AMD Athlon 1333 MHz workstation. While automatic estimation of the GGMRF hyperparameters p and σ is in principle possible using a maximum likelihood estimation technique,[34] we follow Ye *et al.*[24] and use parameter values which empirically give good results. For each reconstruction, the solutions were

initialized with the correct constant background value, as the ICD method's convergence is slow for low spatial frequency image components.[31] Multigrid inversion methods in conjunction with ICD updates alleviate this difficulty,[25,35] but, again, we do not address them in this investigation.

Reconstructions of $\mu_{a_{x,m}}$, $D_{x,m}$, τ, and $\eta\mu_{a_f}$ are shown in Figure 4. We used $p = 2.0$ in all of the reconstructions, and σ values of 0.015 cm^{-1}, 0.02 cm, 0.5 ns, and 0.002 cm^{-1} in the reconstructions of μ_a, D, τ, and γ, respectively. The reconstructions are qualitatively and quantitatively accurate.

In some cases, changes in the lifetime parameter τ may be used to distinguish between diseased and healthy tissue environments. Hence, it is useful to determine if such changes are within the accuracy of the reconstruction algorithm. The simulation was repeated for four fluorophores, each with a different value of τ: 0.1375 ns, 0.275 ns, 0.55 ns, and 1.10 ns. The procedure outlined above was performed for each image. To determine single value of $\hat{\tau}_{avg}$ for each reconstructed image, we used a weighted average:

$$\hat{\tau}_{avg} \quad = \quad \frac{\sum_{i=0}^{N-1} \hat{\gamma}(\mathbf{r}_i)\hat{\tau}(\mathbf{r}_i)}{\sum_{i=0}^{N-1} \hat{\gamma}(\mathbf{r}_i)}. \tag{25}$$

The weighted average is reasonable, as the reconstruction $\hat{\tau}(\mathbf{r})$ may conceivably be large in regions where $\eta\hat{\mu}_{a_f} \approx 0$. It is also similar to the weighting that occurs in the source term of (6), which represents the effect of τ on the data. Figure 5 shows a plot of $\hat{\tau}_{avg}$ as a function of the true value. The result suggests that the method can track even small changes in diagnostic lifetime imaging applications.

We also investigated the propagation of error from the $\hat{\mathbf{x}}_x$ and $\hat{\mathbf{x}}_m$ calculations into the $\hat{\mathbf{x}}_f$ result. Rather than computing reconstructions $\hat{\mathbf{x}}_x$ and $\hat{\mathbf{x}}_m$, we initialized the fluorescence inversion with several constant background values for μ_{a_x}, D_x, μ_{a_m}, and D_m and computed an error metric. For the $\hat{\eta\mu}_{a_f}$ results, the error metric was the normalized root mean squared error (NRMSE),

defined as

$$\text{NRMSE} \quad = \quad \left[\frac{\sum_{i=0}^{N-1} |\hat{\eta\mu}_{a_f}(\mathbf{r}_i) - \eta\mu_{a_f}(\mathbf{r}_i)|^2}{\sum_{i=0}^{N-1} |\eta\mu_{a_f}(\mathbf{r}_i)|^2} \right]^{1/2} . \tag{26}$$

For the $\hat{\tau}$ results, the error metric is the fractional error, $|\hat{\tau}_{avg} - \tau_{true}|/\tau_{true}$. The plots in Figure 6(a)-(d) show the NRMSE for $\hat{\eta\mu}_{a_f}$, as a function of background μ_{a_x}, D_x, μ_{a_m}, and D_m, respectively. The plots in Figure 7(a)-(d) show the fractional error for $\hat{ta}u_{avg}$, also as a function of background μ_{a_x}, D_x, μ_{a_m}, and D_m, respectively. Whenever one parameter was varied, the others were all set to the correct image's background value. In all plots, an "X" symbol shows the error metric value resulting from computing $\hat{\mathbf{x}}_x$ and $\hat{\mathbf{x}}_m$ in advance. For $\hat{\eta\mu}_{a_f}$, the NRMSE is above 0.5 in all cases, even when the full reconstruction was done. This is likely due to blurring of the sharp edges as a result of using the GGMRF prior model with $p = 2$.[30] Ignoring the heterogeneities, but using the correct background values, substantially increases the NRMSE. However, the NRMSE changes relatively little over a range of incorrect background values for all of the parameters, μ_{a_x}, D_x, μ_{a_m}, and D_m. This suggests that the reconstructed yield is fairly robust to initialization errors. However, $\hat{\tau}$ is highly sensitive to initialization with incorrect backgrounds.

6. EXPERIMENTAL RESULTS

To further evaluate the proposed reconstruction method, we performed fluorescence measurements. The data were recorded using a CW imaging device and a versatile phantom box (shown in Figure 8), both described in detail elsewhere.[36] In the absence of phase information, we did not reconstruct τ, and we assumed D to be constant and known based on physical considerations. The instrument has laser diode sources available at both 690 nm and 830 nm, and it has avalanche photodiode (APD) detectors. While 690 nm excitation is not ideal for ICG, a published excitation spectrum[37] indicates that the detected 830 nm emission intensity, using 690 nm excitation, is about 30% of the maximum value (obtained using 780 nm excitation), making it acceptable for this demonstration.

The box had internal dimensions of $16 \times 16 \times 3.8$ cm, where the last dimension is the vertical thickness. As shown in Figure 9, 9 source fibers were connected to the bottom plate (at $z = -1.9$ cm), and 14 detector fibers were connected to the top plate (at $z = 1.9$ cm). A hollow, surface frosted glass sphere of outer diameter 2 cm and thickness of about 2 mm was mounted with its center near $z = 0.7$ cm, as depicted in Figure 10. This sphere was mounted on a small plastic stand on the bottom of the box. It was also connected to a closed circulation channel via thin, translucent rubber tubes leading outside the box, allowing fluorophore solution to be titrated into the sphere from an external reservoir over the course of the experiment. The titration allowed comparable data to be taken both without and with the fluorophore present for analysis and calibration purposes.

The box was filled with a suspension of 0.4% Intralipid. Assuming 690 nm excitation and 830 nm emission, this results in background values of approximately $D_x = 0.071$ cm, $D_m = 0.082$ cm, $\mu_{a_x} = 0.0052$ cm^{-1}, and $\mu_{a_m} = 0.03$ cm^{-1}.[38,39] The sphere was initially filled with the same suspension, creating an essentially homogeneous domain (apart from the glass sphere, rubber tubes, and plastic stand). Before administration of the ICG, two sets of measurements, $\mathbf{y}_x^{(base)}$ and $\mathbf{y}_m^{(base)}$, were recorded using sources at 690 nm and at 830 nm, respectively. We refer to the measurements recorded before ICG administration as baseline measurements. Subsequently, ICG was introduced into the sphere at a concentration of 1.0 μmol/L. For the purposes of reconstructing μ_{a_x}, measurements, which we define as $\mathbf{y}_x^{(uncal)}$, were recorded using 690 nm sources with no optical bandpass filters installed over the detectors. We neglect the fluorescence signal in these measurements, as published quantum efficiency values[16,40] imply that its effect on the data is two or three orders of magnitude below the effects due to absorption. Following a similar procedure, 830 nm data, which we call $\mathbf{y}_m^{(uncal)}$, were recorded (with no filters) for reconstructing μ_{a_m}. Upon completion of these measurements, 830 nm bandpass filters with 12 nm FWHM (Newport 10LF10-830) were installed in front of the detectors to perform the fluorescence measurements. Due to a limited number of

filters, only 9 of the 14 detectors were used for recording the filtered fluorescence measurements (as shown in Figure 9).

The installation of the filters required disconnection of the detector fibers from the detection devices. In principle, the disconnection and subsequent reconnection of the fibers could invalidate the previous baseline calibrations by potentially changing the detectors' coupling efficiencies. Such effects might adversely affect the $\mathbf{y}_x$ and $\mathbf{y}_m$ measurements. Hence, a new baseline calibration procedure was performed. This need for multiple calibrations is a limitation in the design of the experiment which could be alleviated by using a different detection scheme. For example, the instrument used by Ntziachristos and Weissleder,[22] in which a CCD camera imaged a detection fiber array, required only the installation of a single fluorescence filter without perturbing the detection fibers. Incorporating the unknown calibration parameters into the inverse problem[29,41] may also alleviate this difficulty.

To perform the new calibration, the ICG-Intralipid mixture was pumped out of the sphere and replaced with new Intralipid without ICG. Baseline measurements $\mathbf{y}_f^{(base)}$ with 690 nm sources and 830 nm detection were made. Subsequently, a new ICG-Intralipid mixture identical in concentration to the previous one was titrated into the sphere. With the ICG now present, fluorescence measurements $\mathbf{y}_f^{(uncal)}$ were recorded using 690 nm source excitation and 830 nm detection.

Before applying the reconstruction algorithm for $\mathbf{x}_x$ and $\mathbf{x}_m$, calibrations using the baseline data were performed. Synthetic data $\mathbf{y}_x^{(comp)}$ (for 690 nm sources) and $\mathbf{y}_m^{(comp)}$ (for 830 nm sources) were computed for a homogeneous phantom with $D_x = 0.071$ cm, $D_m = 0.082$ cm $\mu_{a_x} = 0.0052$ cm^{-1}, and $\mu_{a_m} = 0.03$ cm^{-1} on a $33 \times 33 \times 17$ grid. Calibrations were performed by normalizing to these computed data:

$$y_{x_i} = y_{x_i}^{(uncal)} \frac{y_{x_i}^{(comp)}}{y_{x_i}^{(base)}} \tag{27}$$

$$y_{m_i} \quad = \quad y_{m_i}^{(uncal)} \frac{y_{m_i}^{(comp)}}{y_{m_i}^{(base)}}, \tag{28}$$

where the subscript i represents the i^{th} component of the data vector. This baseline calibration procedure estimates the unknown scaling and coupling efficiencies in the measurements. For the much dimmer fluorescence measurements, the baseline data $\mathbf{y}_f^{(base)}$ contained significant background signal. Calibrations were performed to account for the unknown coupling efficiencies and to remove these background components from the fluorescence data:

$$y_{f_i} \quad = \quad \left(y_{f_i}^{(uncal)} - y_{f_i}^{(base)}\right) \frac{y_{x_i}^{(comp)}}{y_{x_i}^{(base)}}, \tag{29}$$

where we have used the 690 nm calibration factors. The resulting fluorescence data contain an unknown scale factor, due to the unknown filter attenuation of the 690 nm excitation light relative to the 830 nm fluorescence light.

The reconstructions of μ_{a_x} and μ_{a_m} are shown in Figures 11 and 12, respectively. For each inversion, a volume representing the whole box was discretized into $33 \times 33 \times 17$ voxels. The μ_{a_x} inversion used $\sigma = 0.015$ cm^{-1} and $p = 2$, and the μ_{a_m} inversion used $\sigma = 0.03$ cm^{-1} and $p = 2$. For both images, the ICD algorithm was run for 20 iterations on a 927 MHz Pentium III workstation, taking approximately 10 minutes. The resulting μ_{a_x} and μ_{a_m} reconstructions show a heterogeneity with accurate shape, though with artifacts present in the region close to the top plate. In both images, the sphere's vertical positions are similar, but below the true location by approximately 4 or 5 mm. The similarity of the two reconstructions, despite the fact that they are based on independent data sets, suggests that this error is due to a systematic effect in the reconstruction method. They may be a result of calibration errors, as the assumption of a diffuse, homogeneous medium in the baseline calibrations neglected the presence of the low-scattering glass sphere, the plastic stand used to hold the sphere, and the thin rubber tubes used for pumping in ICG solution. Small errors in the assumed D_x and D_m values might also contribute to artifacts in the reconstructions. In

addition, placing the sphere close to the detectors may have resulted in modeling errors under the diffusion approximation. The reconstructed ICG μ_{a_x} is slightly smaller than the predicted value of 0.039 cm^{-1} which one would expect from the results of Sevick-Muraca *et al.*,[16] after correcting for the use of 690 nm, rather than 780 nm, excitation with the aforementioned 30% factor.[37] The μ_{a_m} image has higher contrast than the μ_{a_x} image, in contrast to a published absorption spectrum for ICG 6.5 μmol/L, which shows higher absorption at 690 nm than at 830 nm. It is possible that ICG's instability in aqueous solution causes some variability in its optical spectrum, as Landsman *et al.*[42] observed a shift in the absorption peak toward longer wavelengths with decreasing concentration. In addition, the effect of an Intralipid suspension on ICG's absorption spectrum has not been investigated in detail, to our knowledge.

Figure 13 shows the reconstructed fluorescent yield $\eta\mu_{a_f}$. As a result of the unknown scale factor in the fluorescence data, the image is in arbitrary units (AU). Making use of the reconstructed μ_{a_x} and μ_{a_m}, the ICD algorithm, using $p = 2.0$ and $\sigma = 5.0$ AU , was run for 20 iterations (about 3 minutes). The iterations were computationally inexpensive due to the linearity of the fluorescence inverse problem. In contrast to the absorption reconstructions, the reconstructed fluorophore's center is slightly higher than that of the true fluorophore. Though no quantitative information is available, the size and shape are approximately correct.

7. CONCLUSION

We have presented a computationally efficient Bayesian inversion strategy for reconstructing fluorescence, absorption, and scattering properties, and demonstrated the method in a simulation study and in a tissue phantom experiment to image fluorescence from ICG in a spherical heterogeneity. The results show potential for use of optical diffusion tomography with fluorescence as a tool for localizing fluorescent contrast agents in clinical diagnostic applications.

It is worth noting that our experimental calibration procedure simplifies the problem substantially by choosing a background absorption value in advance and observing changes with respect to baseline measurements. We have found that this procedure reconstructs similar-looking absorbers, over a wide range of background values. Ongoing work with nonlinear multigrid reconstruction algorithms[25] and more accurate transport models[43] may improve reconstruction accuracy with less favorable calibrations.

The error propagation study suggests that qualitative, if not necessarily quantitative, results for the yield imaging problem may be obtainable without doing the full $\mathbf{x}_x$ and $\mathbf{x}_m$ inversions. Although the reconstructed lifetime was highly sensitive to incorrectly initialized background properties, the reconstructed yield was not. Hence, as Ntziachristos and Weissleder[22] have observed, simplified first-order models are more easily applied to the fluorescent yield imaging problem than to the full absorption imaging problem. This enables a simpler experimental approach which requires no baseline data. For qualitative localization of tumors, this could prove to be a decisive advantage of fluorescence imaging over absorption imaging.

8. APPENDICES

A. Fréchet Derivatives

Here we describe the computation of the Fréchet derivatives of the forward operators used in this study. Let $g(\mathbf{r}_{src}, \mathbf{r}_{obs}; \mathbf{x})$ be the diffusion equation Green's function for the problem domain computed using the image vector $\mathbf{x}$ and a numerical forward solver, with $\mathbf{r}_{src}$ as the source location and $\mathbf{r}_{obs}$ as the observation point. In addition, suppose that for a particular image $\mathbf{x}$ there are K sources and M detectors, and a total of $P = KM$ measurements. Let $\mathbf{r}_{s_k}$ represent the position of the k^{th} source and let $\mathbf{r}_{d_{m'}}$ represent the position of the $\{m'\}^{th}$ detector. * It follows that the

*Here, we use the letter m to denote detector number, as in our previous publications,[24–26,29] but with a prime mark to avoid confusion with the fluorescence emission subscript.

computed data vector $\mathbf{f}(\mathbf{x})$ is given by

$$\mathbf{f}(\mathbf{x}) = \left[\begin{array}{cccccccc} g(\mathbf{r}_{s_1}, \mathbf{r}_{d_1}; \mathbf{x}) & g(\mathbf{r}_{s_1}, \mathbf{r}_{d_2}; \mathbf{x}) & \cdots & g(\mathbf{r}_{s_1}, \mathbf{r}_{d_M}; \mathbf{x}) & g(\mathbf{r}_{s_2}, \mathbf{r}_{d_1}; \mathbf{x}) & \cdots & g(\mathbf{r}_{s_K}, \mathbf{r}_{d_M}; \mathbf{x}) \end{array} \right]^T . \quad (30)$$

For image vectors $\mathbf{x}$ of size $2N$, the Fréchet derivative is the $P \times 2N$ complex matrix given by

$$\mathbf{f}'(\mathbf{x}) = \left[\begin{array}{cccc} \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_1} & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_{2N}} \\ \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_2}; \mathbf{x})}{\partial x_1} & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_2}; \mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_2}; \mathbf{x})}{\partial x_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_1} & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g(\mathbf{r}_{s_1}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_{2N}} \\ \frac{\partial g(\mathbf{r}_{s_2}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_1} & \frac{\partial g(\mathbf{r}_{s_2}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g(\mathbf{r}_{s_2}, \mathbf{r}_{d_1}; \mathbf{x})}{\partial x_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g(\mathbf{r}_{s_K}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_1} & \frac{\partial g(\mathbf{r}_{s_K}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_2} & \cdots & \frac{\partial g(\mathbf{r}_{s_K}, \mathbf{r}_{d_M}; \mathbf{x})}{\partial x_{2N}} \end{array} \right] . \quad (31)$$

For the absorption and scattering coefficients, the discrete representations of the Fréchet derivative matrix elements have been derived and reported previously[27,44] as

$$\frac{\partial g(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x})}{\partial \mu_a(\mathbf{r}_i)} \simeq -g(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}) g(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}) V \quad (32)$$

$$\frac{\partial g(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x})}{\partial D(\mathbf{r}_i)} \simeq -\nabla g(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}) \cdot \nabla g(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}) V \ , \quad (33)$$

where $\simeq$ is used due to domain discretization errors, V is the voxel volume, $\mathbf{r}_i$ is the position of the i^{th} voxel, and reciprocity[45] (which allows replacement of $g(\mathbf{r}_{src}, \mathbf{r}_{obs}; \mathbf{x})$ with $g(\mathbf{r}_{obs}, \mathbf{r}_{src}; \mathbf{x})$) has been used to reduce computation. Here, ∇ is the spatial gradient operator, which, in our computations, is evaluated numerically as a symmetric first difference. The separability of (32) and (33) with respect to source index and detector index enables additional savings in computation and in storage.[29] Rather than creating the entire $KM \times 2N$ matrix, it suffices to initially compute and

store two Green's function matrices of sizes $K \times N$ and $M \times N$, respectively:

$$\mathbf{G}^{(s)} = \begin{bmatrix} g(\mathbf{r}_{s_1}, \mathbf{r}_1; \mathbf{x}) & \cdots & g(\mathbf{r}_{s_1}, \mathbf{r}_N; \mathbf{x}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{r}_{s_K}, \mathbf{r}_1; \mathbf{x}) & \cdots & g(\mathbf{r}_{s_K}, \mathbf{r}_N; \mathbf{x}) \end{bmatrix} \tag{34}$$

$$\mathbf{G}^{(d)} = \begin{bmatrix} g(\mathbf{r}_{d_1}, \mathbf{r}_1; \mathbf{x}) & \cdots & g(\mathbf{r}_{d_1}, \mathbf{r}_N; \mathbf{x}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{r}_{d_M}, \mathbf{r}_1; \mathbf{x}) & \cdots & g(\mathbf{r}_{d_1}, \mathbf{r}_N; \mathbf{x}) \end{bmatrix}. \tag{35}$$

During the ICD scan, when the i^{th} voxel of $\mathbf{x}$ is to be modified, the i^{th} column of $\mathbf{f}'(\mathbf{x})$ can be formed from (34) and (35).

For the fluorescence problem, more specific notation is needed. Let $g_x(\mathbf{r}_{src}, \mathbf{r}_{obs}; \mathbf{x}_x)$ denote the λ_x Green's function obtained by solving (1), and let $g_m(\mathbf{r}_{src}, \mathbf{r}_{obs}; \mathbf{x}_m)$ denote the λ_m Green's function obtained by solving (3). We denote the Green's function matrices accordingly:

$$\mathbf{G}_x^{(s)} = \begin{bmatrix} g_x(\mathbf{r}_{s_1}, \mathbf{r}_1; \mathbf{x}_x) & \cdots & g_x(\mathbf{r}_{s_1}, \mathbf{r}_N; \mathbf{x}_x) \\ \vdots & \ddots & \vdots \\ g_x(\mathbf{r}_{s_K}, \mathbf{r}_1; \mathbf{x}_x) & \cdots & g_x(\mathbf{r}_{s_K}, \mathbf{r}_N; \mathbf{x}_x) \end{bmatrix} \tag{36}$$

$$\mathbf{G}_x^{(d)} = \begin{bmatrix} g_x(\mathbf{r}_{d_1}, \mathbf{r}_1; \mathbf{x}_x) & \cdots & g_x(\mathbf{r}_{d_1}, \mathbf{r}_N; \mathbf{x}_x) \\ \vdots & \ddots & \vdots \\ g_x(\mathbf{r}_{d_M}, \mathbf{r}_1; \mathbf{x}_x) & \cdots & g_x(\mathbf{r}_{d_M}, \mathbf{r}_N; \mathbf{x}_x) \end{bmatrix} \tag{37}$$

$$\mathbf{G}_m^{(s)} = \begin{bmatrix} g_m(\mathbf{r}_{s_1}, \mathbf{r}_1; \mathbf{x}_m) & \cdots & g_m(\mathbf{r}_{s_1}, \mathbf{r}_N; \mathbf{x}_m) \\ \vdots & \ddots & \vdots \\ g_m(\mathbf{r}_{s_K}, \mathbf{r}_1; \mathbf{x}_m) & \cdots & g_m(\mathbf{r}_{s_K}, \mathbf{r}_N; \mathbf{x}_m) \end{bmatrix} \tag{38}$$

$$\mathbf{G}_m^{(d)} = \begin{bmatrix} g_m(\mathbf{r}_{d_1}, \mathbf{r}_1; \mathbf{x}_m) & \cdots & g_m(\mathbf{r}_{d_1}, \mathbf{r}_N; \mathbf{x}_m) \\ \vdots & \ddots & \vdots \\ g_m(\mathbf{r}_{d_M}, \mathbf{r}_1; \mathbf{x}_m) & \cdots & g_m(\mathbf{r}_{d_M}, \mathbf{r}_N; \mathbf{x}_m) \end{bmatrix}. \quad (39)$$

Consider one reparameterization of the right hand side of (2):

$$\eta\mu_{a_f}(\mathbf{r})\frac{1 - j\omega\tau(\mathbf{r})}{1 + [\omega\tau(\mathbf{r})]^2} = \beta_R(\mathbf{r}) - j\beta_I(\mathbf{r}). \quad (40)$$

It follows immediately that the inverse problem for β_R and β_I is linear. Let $g_f(\mathbf{r}_{src}, \mathbf{r}_{obs}; \mathbf{x}_x, \mathbf{x}_m)$ denote the fluorescence observed at $\mathbf{r}_{obs}$ emitted in response to excitation at $\mathbf{r}_{src}$. The Fréchet derivatives for β_I and β_R are given by

$$\frac{\partial g_f(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x}_x, \mathbf{x}_m)}{\partial \beta_R(\mathbf{r}_i)} \simeq g_m(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}_m) g_x(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}_x) V \quad (41)$$

$$\frac{\partial g_f(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x}_x, \mathbf{x}_m)}{\partial \beta_I(\mathbf{r}_i)} \simeq -j g_m(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}_m) g_x(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}_x) V. \quad (42)$$

It is possible to solve the fluorescence inverse problem using this parameterization, and then convert the result into the physical parameters $\eta\mu_{a_f}$ and τ. However, the computation of τ requires a division of β_I by β_R, an operation which could result in large noise artifacts in regions where β_R is small. To circumvent this problem, we use the γ and τ parameterization of (6), permitting us to apply regularization directly to τ. In our sequential update scheme, τ is assumed constant while updates of γ are performed, and vice versa. As a result, we use the following Fréchet derivative expressions:

$$\frac{\partial g_f(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x}_x, \mathbf{x}_m)}{\partial \gamma(\mathbf{r}_i)} \simeq g_m(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}_m) g_x(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}_x)(1 - j\omega\hat{\tau}(\mathbf{r}_i)) V \quad (43)$$

$$\frac{\partial g_f(\mathbf{r}_{s_k}, \mathbf{r}_{d_{m'}}; \mathbf{x}_x, \mathbf{x}_m)}{\partial \tau(\mathbf{r}_i)} \simeq -j\omega\hat{\gamma}(\mathbf{r}_i) g_m(\mathbf{r}_{d_{m'}}, \mathbf{r}_i; \mathbf{x}_m) g_x(\mathbf{r}_{s_k}, \mathbf{r}_i; \mathbf{x}_x) V. \quad (44)$$

After the reconstructions of $\mathbf{x}_x$ and $\mathbf{x}_m$ are obtained, $\mathbf{G}_x(\mathbf{r}_s, \mathbf{r}; \mathbf{x}_x)$ and $\mathbf{G}_m(\mathbf{r}_d, \mathbf{r}; \mathbf{x}_m)$ have already been stored, and the Green's functions of (43) and (44) need not be recomputed. As the estimates $\hat{\gamma}$ and $\hat{\tau}$ are updated, they are incorporated into the derivative expressions.

B. Pseudocode for Inversion Algorithm

main {

1. Initialize $\hat{\mathbf{x}}_x$, $\hat{\mathbf{x}}_m$, and $\hat{\mathbf{x}}_f$ with background estimates.

2. Repeat until converged: {

   (a) $\hat{\alpha}_x \leftarrow \dfrac{1}{P_x} \; || \; \mathbf{y}_x - \mathbf{f}_x(\hat{\mathbf{x}}_x) \; ||^2_{\mathbf{\Lambda}_x}$

   (b) For $k = 1 : K$ {

   Compute $g_x(\mathbf{r}_{s_k}, \mathbf{r}; \hat{\mathbf{x}}_x)$ by solving Eq. (1) with source at $\mathbf{r}_{s_k}$

   }

   (c) For $m' = 1 : M$ {

   Compute $g_x(\mathbf{r}_{d_{m'}}, \mathbf{r}; \hat{\mathbf{x}}_x)$ by solving Eq. (1) with source at $\mathbf{r}_{d_{m'}}$

   }

   (d) Form $\mathbf{G}_x^{(s)}$ and $\mathbf{G}_x^{(d)}$ using Eq. (36) and Eq. (37)

   (e) $\hat{\mathbf{x}}_x \leftarrow$ ICD_update$(\hat{\mathbf{x}}_x, \hat{\alpha}_x, \mathbf{G}_x^{(s)}, \mathbf{G}_x^{(d)})$

}

3. Repeat until converged: {

   (a) $\hat{\alpha}_m \leftarrow \dfrac{1}{P_m} \; || \; \mathbf{y}_m - \mathbf{f}_m(\hat{\mathbf{x}}_m) \; ||^2_{\mathbf{\Lambda}_m}$

   (b) For $k = 1 : K$ {

   Compute $g_m(\mathbf{r}_{s_k}, \mathbf{r}; \hat{\mathbf{x}}_m)$ by solving Eq. (3) with source at $\mathbf{r}_{s_k}$

   }

   (c) For $m' = 1 : M$ {

   Compute $g_m(\mathbf{r}_{d_{m'}}, \mathbf{r}; \hat{\mathbf{x}}_m)$ by solving Eq. (3) with source at $\mathbf{r}_{d_{m'}}$

   }

   (d) Form $\mathbf{G}_m^{(s)}$ and $\mathbf{G}_m^{(d)}$ using Eq. (38) and Eq. (39)

(e) $\hat{\mathbf{x}}_m \leftarrow \text{ICD\_update}(\hat{\mathbf{x}}_m, \hat{\alpha}_m, \mathbf{G}_m^{(s)}, \mathbf{G}_m^{(d)})$

4. Repeat until converged: {

(a) $\hat{\alpha}_f \leftarrow \frac{1}{P_f} \; || \; \mathbf{y}_f - \mathbf{f}_f(\hat{\mathbf{x}}_f, \hat{\mathbf{x}}_x, \hat{\mathbf{x}}_m) \; ||^2_{\mathbf{\Lambda}_f}$

(b) $\hat{\mathbf{x}}_f \leftarrow \text{ICD\_update}(\hat{\mathbf{x}}_f, \hat{\alpha}_f, \mathbf{G}_x^{(s)}, \mathbf{G}_m^{(d)})$

}

}

$\hat{\mathbf{x}} \leftarrow \text{ICD\_update}(\hat{\mathbf{x}}, \hat{\alpha}, \mathbf{G}^{(s)}, \mathbf{G}^{(d)}; \mathbf{x}))$ {

1. For $i = 1, \ldots, N$ (in random order), {

(a) Compute $[\mathbf{f}'(\hat{\mathbf{x}})]_{*(i)}$, as described in Appendix A

(b) Update x_i, as described by Ye *et al.*:[24]

$$\hat{x}_i \leftarrow \arg\min_{x_i \geq 0} \left\{ \frac{1}{\hat{\alpha}} \left|\left| \mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}) - \left[\mathbf{F}'(\hat{\mathbf{x}})\right]_{*(i)} (x_i - \hat{x_i}) \right|\right|^2_{\mathbf{\Lambda}} + \frac{1}{p_a \sigma_a{}^{p_a}} \sum_{j \in \mathcal{N}_i} b_{i-j} \left| x_i - \hat{x}_i \right|^{p_a} \right\}$$

}

2. For $i = N+1, \ldots, 2N$ (in random order), {

(a) Compute $[\mathbf{f}'(\hat{\mathbf{x}})]_{*(i)}$, as described in Appendix A

(b) Update x_i, as described by Ye *et al.*:[24]

$$\hat{x}_i \leftarrow \arg\min_{x_i \geq 0} \left\{ \frac{1}{\hat{\alpha}} \left|\left| \mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}) - \left[\mathbf{F}'(\hat{\mathbf{x}})\right]_{*(i)} (x_i - \hat{x_i}) \right|\right|^2_{\mathbf{\Lambda}} + \frac{1}{p_b \sigma_b{}^{p_b}} \sum_{j \in \mathcal{N}_i} b_{i-j} \left| x_i - \hat{x}_i \right|^{p_b} \right\}$$

}

3. Return $\hat{\mathbf{x}}$.

}

9. ACKNOWLEDGMENTS

This work was funded by the National Science Foundation under contract CCR-0073357. Quan Zhang and David A. Boas acknowledge support from Advanced Research Technologies Inc. In addition, we thank Tina Chaves for her assistance with the laboratory equipment.

References

1. S. R. Arridge, "Optical tomography in medical imaging," Inverse Problems **15,** R41–R93 (1999).

2. D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," IEEE Signal Processing Magazine **18,** 57–75 (2001).

3. V. G. Peters, D. R. Wyman, M. S. Patterson, and G. L. Frank, "Optical properties of normal and diseased human breast tissues in the visible and near infrared," Phys. Med. Biol. **35,** 1317–1334 (1990).

4. T. L. Troy, D. L. Page, and E. M. Sevick-Muraca, "Optical properties of normal and diseased breast tissue: prognosis for optical mammography," J. Biomed. Opt. **1,** 342–355 (1996).

5. A. Pèlegrin, S. Folli, F. Buchegger, J. Mach, G. Wagnières, and H. van den Bergh, "Antibody-fluorescein conjugates for photoimmunodiagnosis of human colon carcinoma in nude mice," Cancer **67,** 2529–2537 (1991).

6. B. Ballou, G. W. Fisher, T. R. Hakala, and D. L. Farkas, "Tumor detection and visualization using cyanine fluorochrome-labeled antibodies," Biotechnol. Prog. **13,** 649–658 (1997).

7. R. Cubeddu, G. Canti, A. Pifferi, P. Taroni, and G. Valentini, "Fluorescence lifetime imaging of experimental tumors in hematoporphyrin derivative-sensitized mice," Photochem. Photobiol. **66,** 229–236 (1997).

8. J. A. Reddy and P. S. Low, "Folate-mediated targeting of therapeutic and imaging agents to cancers," Critical Reviews in Therapeutic Drug Carrier Systems **15,** 587–627 (1998).

9. J. S. Reynolds, T. L. Troy, R. H. Mayer, A. B. Thompson, D. J. Waters, K. K. Cornell, P. W. Snyder, and E. M. Sevick-Muraca, "Imaging of spontaneous canine mammary tumors using fluorescent contrast agents," Photochem. Photobiol. **70,** 87–94 (1999).

10. U. Mahmood, C. Tung, J. A. Bogdanov, and R. Weissleder, "Near-infrared optical imaging of protease activity for tumor detection," Radiology **213,** 866–870 (1999).

11. K. Licha, B. Riefke, V. Ntziachristos, A. Becker, B. Chance, and W. Semmler, "Hydrophilic cyanine dyes as contrast agents for near-infrared tumor imaging: synthesis, photophysical properties and spectroscopic *in vivo* characterization," Photochem. Photobiol. **72,** 392–398 (2000).

12. V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance, "Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement," Proc. Natl. Acad. Sci **97,** 2767–2772 (2000).

13. A. Becker, C. Hessenius, K. Licha, B. Ebert, U. Sukowski, W. Semmler, B. Wiedenmann, and C. Grotzinger, "Receptor-targeted optical imaging of tumors with near-infrared fluorescent ligands," Nat. Biotech. **19,** 327–331 (2001).

14. J. E. Bugaj, S. Achilefu, R. B. Dorshow, and R. Rajagopalan, "Novel fluorescent contrast agents for optical imaging of *in vivo* tumors based on a receptor-targeted dye-peptide conjugate platform," J. Biomed. Opt. **6,** 122–133 (2001).

15. M. S. Patterson and B. W. Pogue, "Mathematical model for time-resolved and frequency-domain fluorescence spectroscopy in biological tissues," Appl. Opt. **33,** 1963–1974 (1994).

16. E. M. Sevick-Muraca, G. Lopez, J. S. Reynolds, T. L. Troy, and C. L. Hutchinson, "Fluorescence and absorption contrast mechanisms for biomedical optical imaging using frequency-domain techniques,"

Photochem. Photobiol. **66,** 55–64 (1997).

17. R. Roy and E. M. Sevick-Muraca, "Three-dimensional unconstrained and constrained image-reconstruction techniques applied to fluorescence, frequency-domain photon migration," Appl. Opt. **40,** 2206–2215 (2001).

18. M. A. O'Leary, D. A. Boas, X. D. Li, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," Opt. Lett. **21,** 158–160 (1996).

19. D. Paithankar, A. Chen, B. Pogue, M. Patterson, and E. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," Appl. Opt. **36,** 2260–2272 (1997).

20. H. Jiang, "Frequency-domain fluorescent diffusion tomography: a finite-element-based algorithm and simulations," Appl. Opt. **37,** 5337–5343 (1998).

21. J. Chang, H. L. Graber, and R. L. Barbour, "Luminescence optical tomography of dense scattering media," J. Opt. Soc. Am. A **14,** 288–299 (1997).

22. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," Opt. Lett. **26,** 893–895 (2001).

23. D. A. Boas, "A fundamental limitation of linearized algorithms for diffuse optical tomography," Opt. Express **1,** 404–413 (1997).

24. J. C. Ye, K. J. Webb, C. A. Bouman, and R. P. Millane, "Optical diffusion tomography using iterative coordinate descent optimization in a Bayesian framework," J. Opt. Soc. Am. A **16,** 2400–2412 (1999).

25. J. C. Ye, C. A. Bouman, K. J. Webb, and R. P. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," IEEE Trans. Image Processing **10,** 909–922 (2001).

26. A. B. Milstein, S. Oh, J. S. Reynolds, K. J. Webb, C. A. Bouman, and R. P. Millane, "Three-dimensional Bayesian optical diffusion tomography with experimental data," Opt. Lett. **27,** 95–97 (2002).

27. J. C. Ye, K. J. Webb, R. P. Millane, and T. J. Downar, "Modified distorted Born iterative method with an approximate Fréchet derivative for optical diffusion tomography," J. Opt. Soc. Am. A **16,** 1814–1826 (1999).

28. J. S. Reynolds, C. A. Thompson, K. J. Webb, F. P. LaPlant, and D. Ben-Amotz, "Frequency domain modeling of reradiation in highly scattering media," Appl. Opt. **36,** 2252–2259 (1997).

29. S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector calibration in three-dimensional Bayesian optical diffusion tomography," J. Opt. Soc. Am. A (in press).

30. C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," IEEE Trans. Image Processing **2,** 296–310 (1993).

31. K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," IEEE Trans. Signal Processing **41,** 534–548 (1993).

32. J. C. Adams, "MUDPACK: Multigrid portable FORTRAN software for the efficient solution of linear elliptic partial differential equations," Appl. Math. Comput. **34,** 113–146 (1989).

33. J. J. Duderstadt and L. J. Hamilton, *Nuclear Reactor Analysis* (Wiley, New York, 1976).

34. S. S. Saquib, C. A. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," IEEE Trans. Image Processing **7,** 1029–1044 (1998).

35. K. Sauer and C. Bouman, "Bayesian Estimation of Transmission Tomograms Using Segmentation Based Optimization," IEEE Trans. on Nuclear Science **39,** 1144–1152 (1992).

36. Q. Zhang, T. J. Brukilacchio, T. Gaudett, L. Wang, A. Li, and D. A. Boas, "Experimental comparison of using continuous-wave and frequency-domain diffuse optical imaging systems to detect heterogeneities," In *Proc. SPIE*, **4250,** 219–238 (2001).

37. R. H. Mayer, J. S. Reynolds, and E. M. Sevick-Muraca, "Measurement of the fluorescence lifetime in scattering media by frequency-domain photon migration," Appl. Opt. **38,** 4930–4938 (1999).

38. H. J. van Staveren, C. J. M. Moes, J. van Marie, S. A. Prahl, and M. J. C. van Gemert, "Light scattering in Intralipid-10% in the wavelength range of 400-1100 nm," Appl. Opt. **30,** 4507–4514 (1991).

39. G. M. Hale and M. R. Querry, "Optical constants of water in the 200-nm to 200-μm wavelength region," Appl. Opt. **12,** 555–563 (1973).

40. R. C. Benson and H. A. Kues, "Fluorescence properties of indocyanine green as related to angiography," Phys. Med. Biol. **23,** 159–163 (1978).

41. D. Boas, T. Gaudette, and S. Arridge, "Simultaneous imaging and optode calibration with diffuse optical tomography," Opt. Express **8,** 263–270 (2001).

42. M. L. J. Landsman, G. Kwant, G. A. Mook, and W. G. Zijlstra, "Light-absorbing properties, stability, and spectral stabilization of indocyanine green," J. Appl. Physiol. **40,** 575–583 (1976).

43. A. D. Klose and A. H. Hielscher, "A transport-theory based reconstruction algorithm for optical tomography," In *Optical Tomography and Spectroscopy of Tissue III*, **Proc. 3597,** 26–35 (1999).

44. S. R. Arridge, "Photon-measurement density functions. Part 1: Analytical forms," Appl. Opt. **34,** 7395–7409 (1995).

45. W. C. Chew, *Waves and Fields in Inhomogeneous Media* (Van Nostrand Reinhold, New York, 1990).

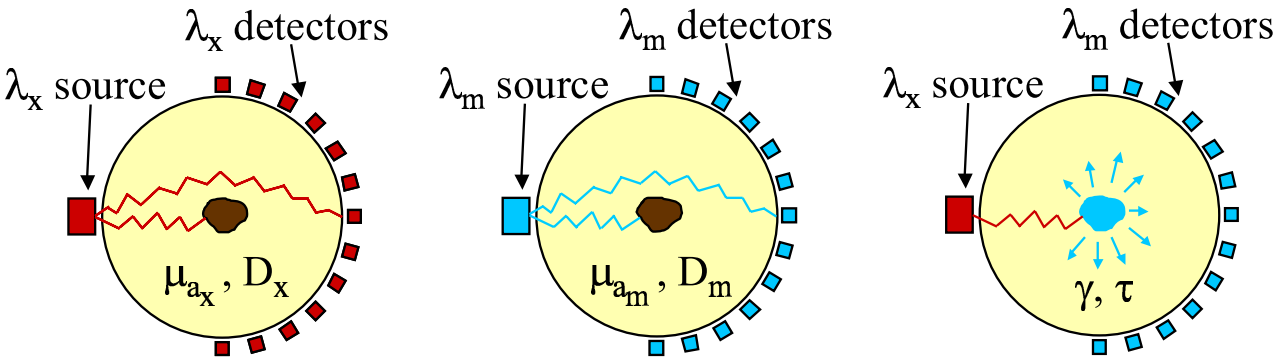


Fig. 1. Proposed measurement scheme.

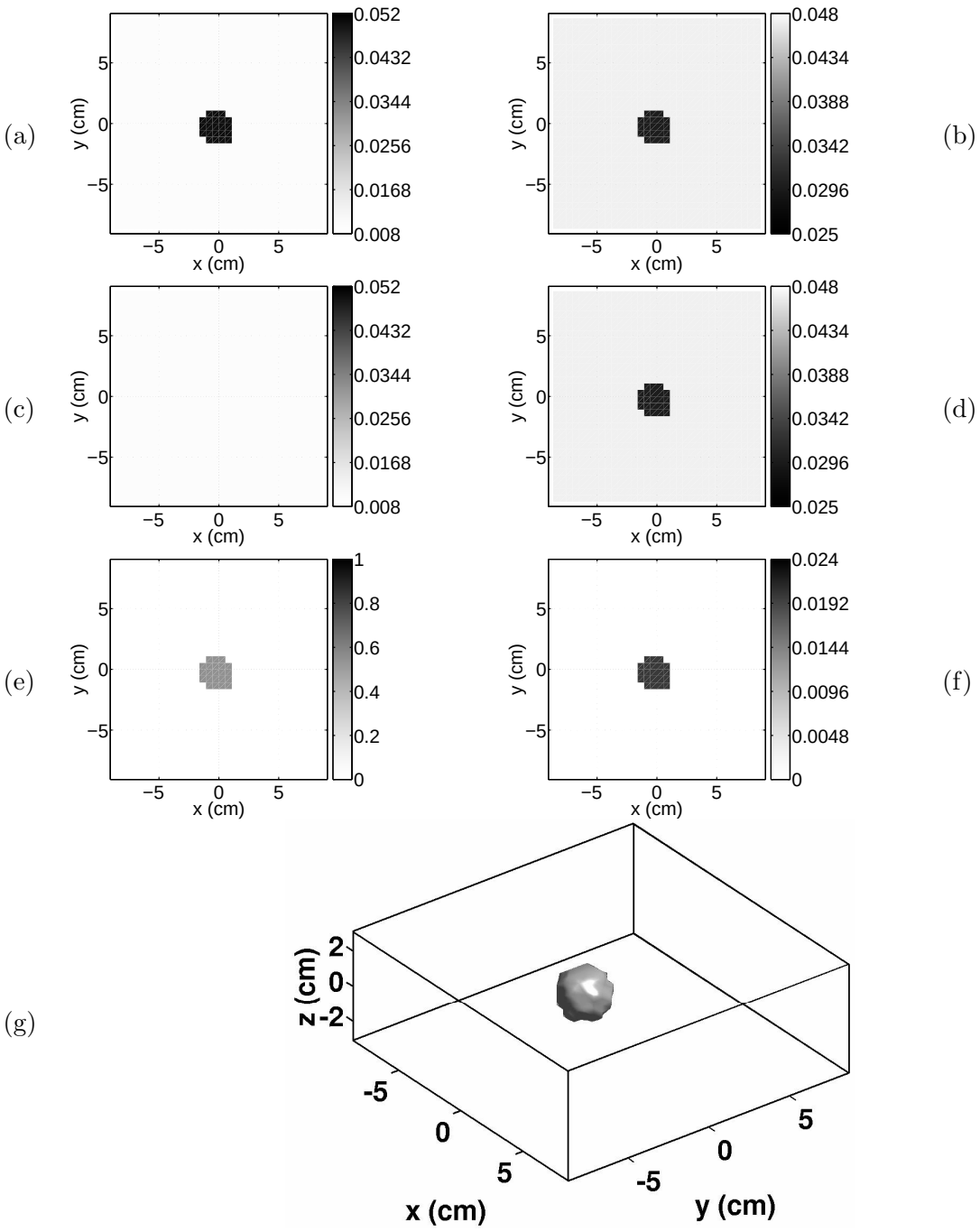


Fig. 2. True phantom, with cross sections of the widest part of the heterogeneity: (a) μ_{a_x} in cm^{-1} (b) D_x in cm (c) μ_{a_m} in cm^{-1} (d) D_m in cm (e) τ in ns (f) $\eta\mu_{a_f}$ in cm^{-1} (g) $\eta\mu_{a_f} = 0.01$ cm^{-1} isosurface

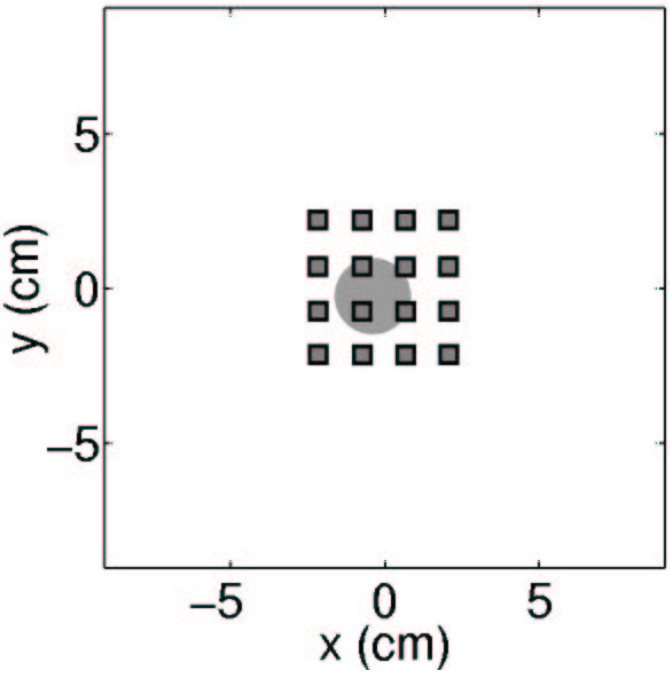


Fig. 3. Grid used for both sources and detectors in the simulation. with the relative location of the sphere depicted.

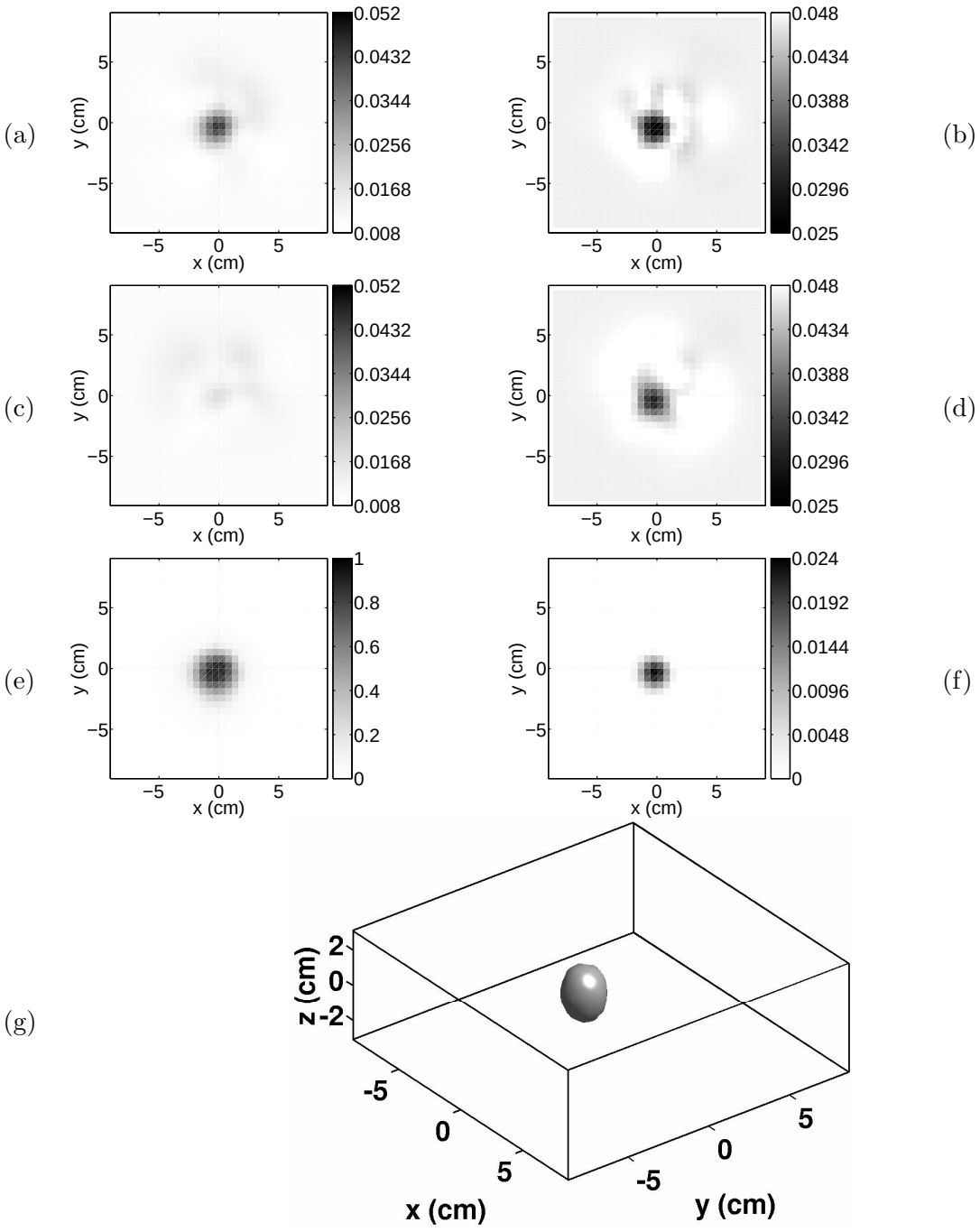


Fig. 4. Reconstructed phantom: (a) μ_{a_x} in cm^{-1} (b) D_x cm (c) μ_{a_m} in cm^{-1} (d) D_m in cm (e) τ in ns (f) $\eta\mu_{a_f}$ in cm^{-1} (g) $\eta\mu_{a_f} = 0.01$ cm^{-1} isosurface

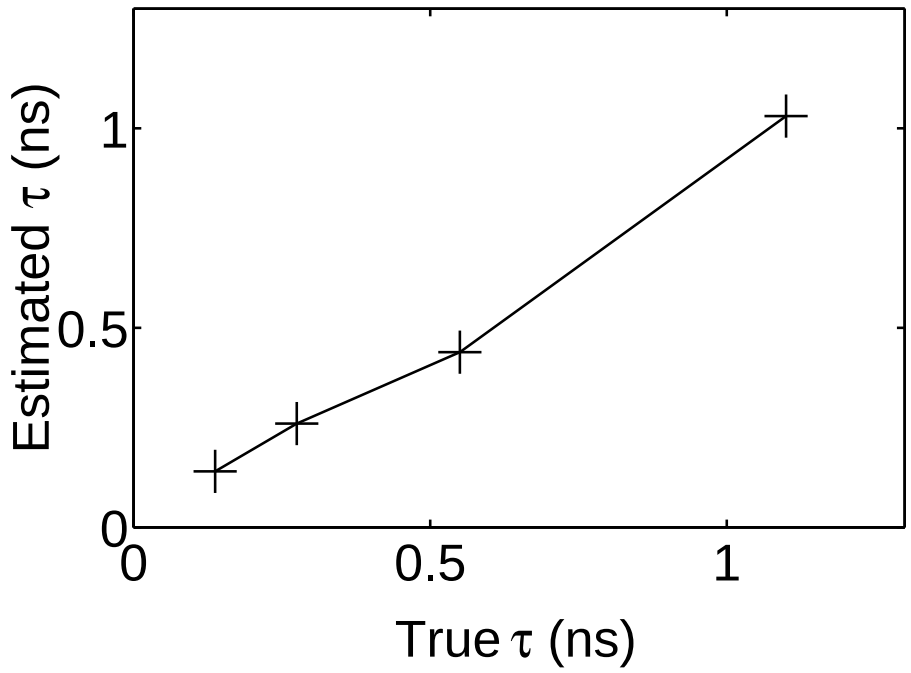


Fig. 5. Plot of estimate $\hat{\tau}_{avg}$ versus the true value of τ. The trend is almost linear, as desired.

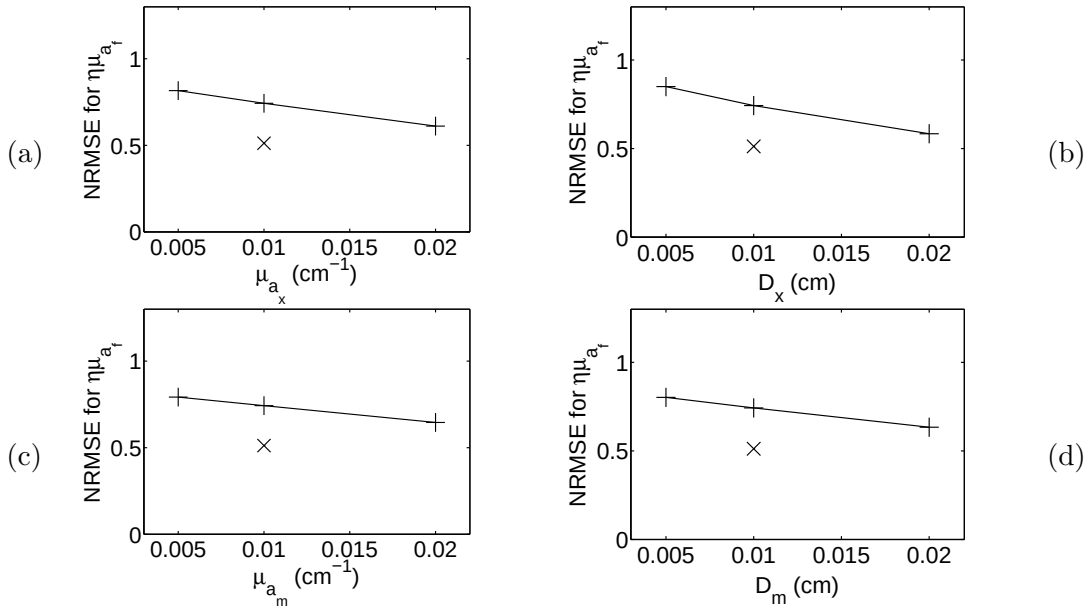


Fig. 6. NRMSE for $\hat{\eta\mu}_{a_f}$, due to changes in assumed constant background values for: (a) μ_{a_x} in cm^{-1} (b) D_x cm (c) μ_{a_m} in cm^{-1} (d) D_m in cm

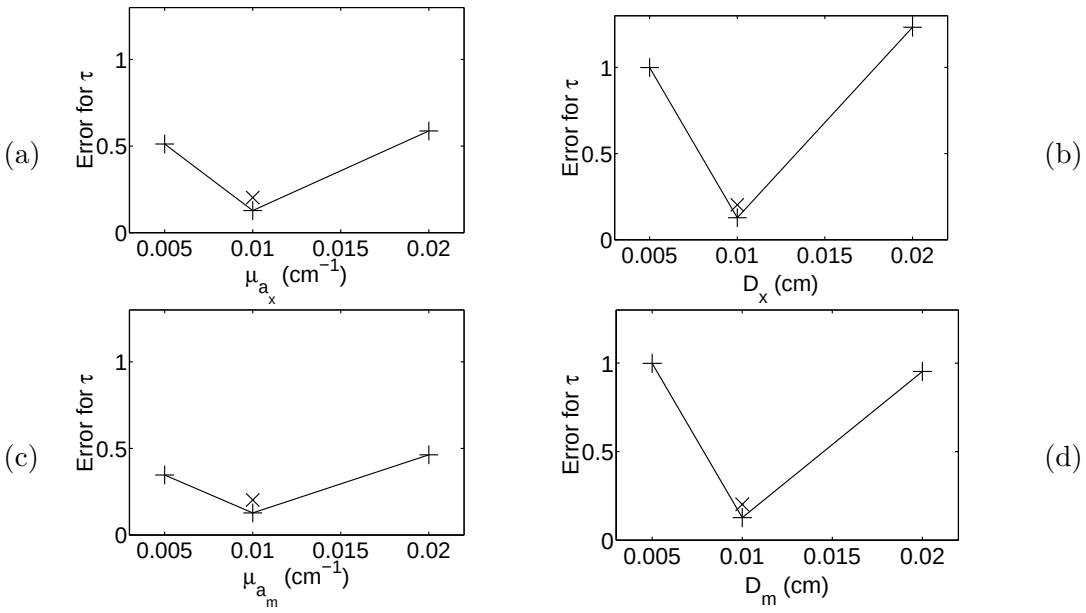


Fig. 7. Fractional error for $\hat{\tau}_{avg}$, due to changes in assumed constant background values for: (a) μ_{a_x} in cm^{-1} (b) D_x cm (c) μ_{a_m} in cm^{-1} (d) D_m in cm

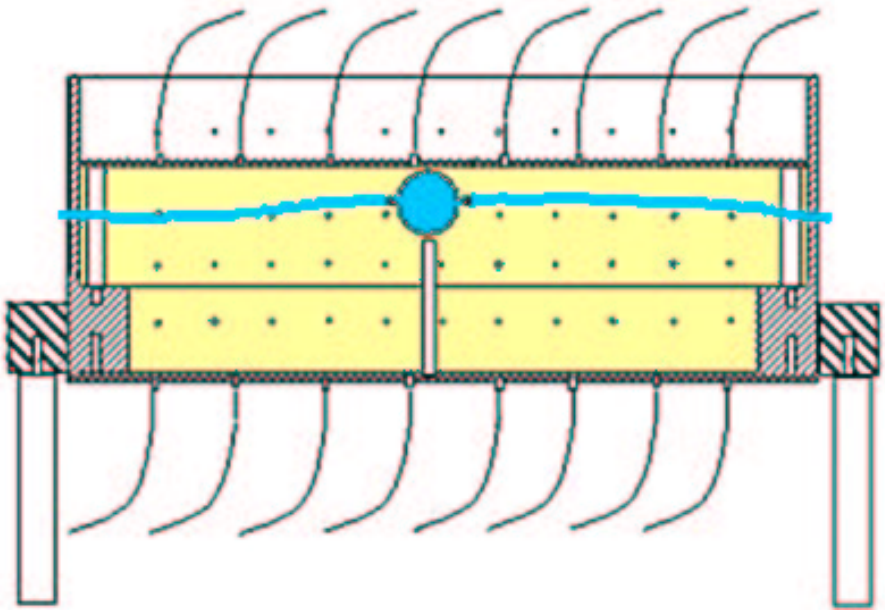
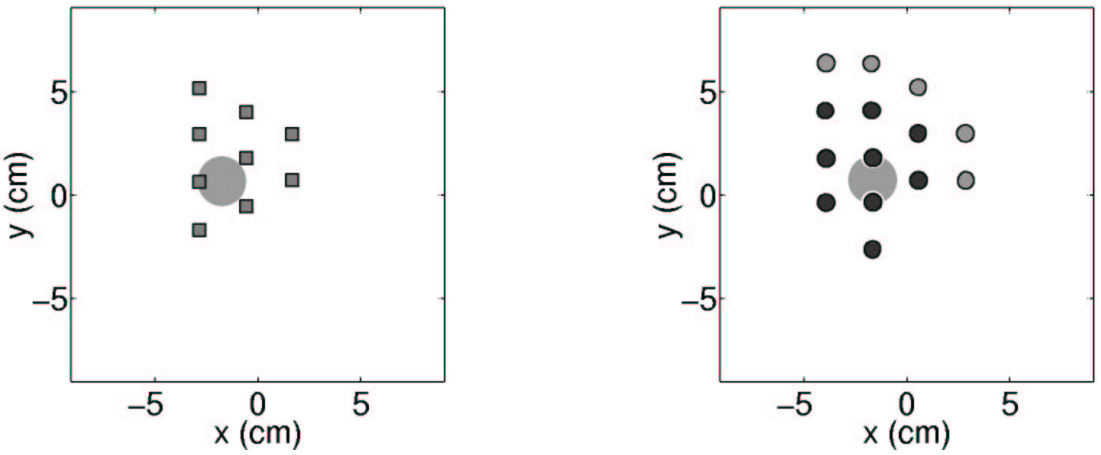


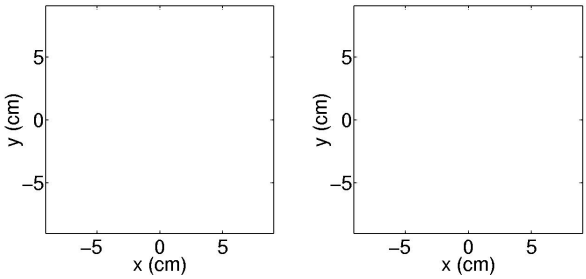
Fig. 8. Phantom box schematic, showing the fibers and the spherical heterogeneity.



(a) Bottom plate (sources)

(b) Top plate (detectors)

Fig. 9. Source and detector layout for experiment. The blackened detector symbols represent detector positions used in the fluorescence measurements. The relative location of the sphere is also depicted.



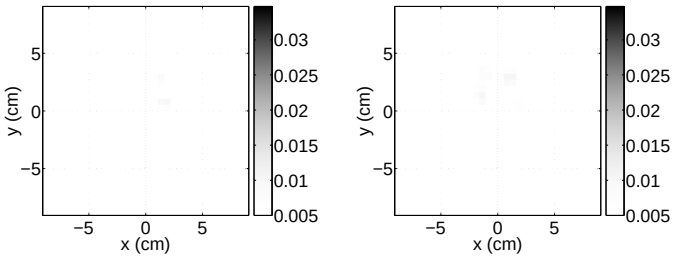
(a) z=-1.82 cm (b) z=-1.30 cm

(c) z=-0.78 cm (d) z=-0.26 cm

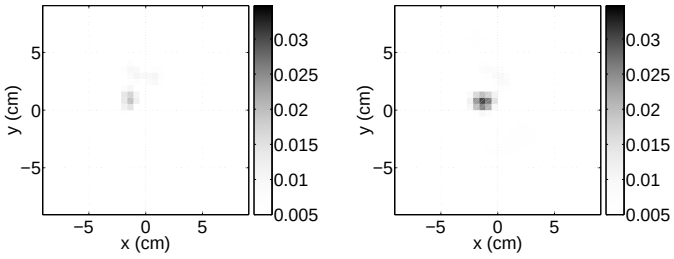
(e) z=0.26 cm (f) z=0.78 cm

(g) z=1.30 cm (h) z=1.82 cm

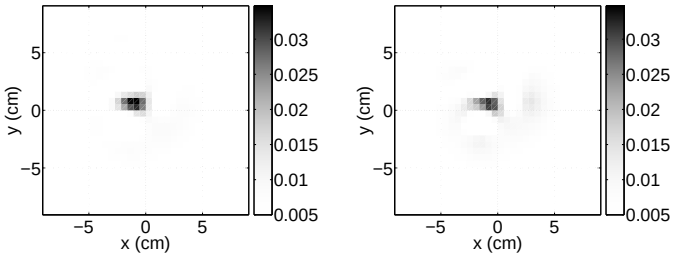
Fig. 10. True fluorophore location



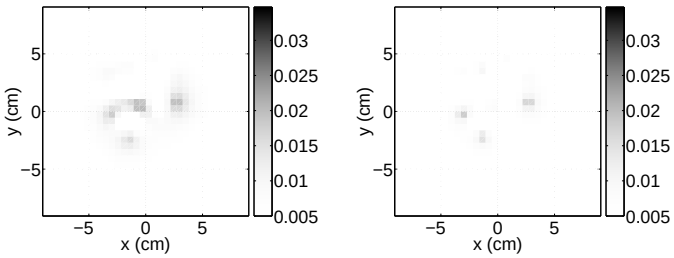
(a) z=-1.82 cm (b) z=-1.30 cm



(c) z=-0.78 cm (d) z=-0.26 cm

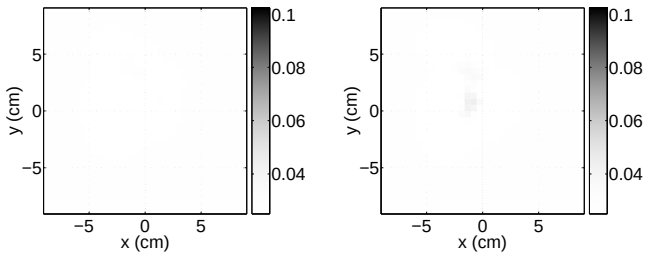


(e) z=0.26 cm (f) z=0.78 cm



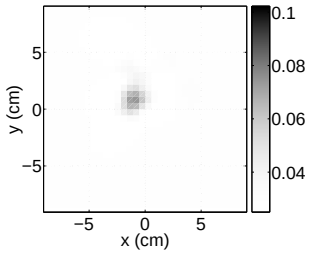
(g) z=1.30 cm (h) z=1.82 cm

Fig. 11. Reconstructions of μ_{a_x} in cm^{-1}

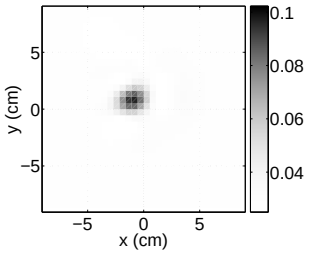


(a) z=-1.82 cm

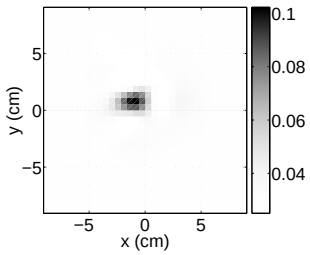
(b) z=-1.30 cm



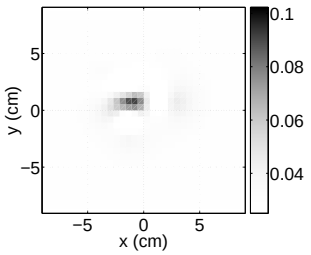
(c) z=-0.78 cm



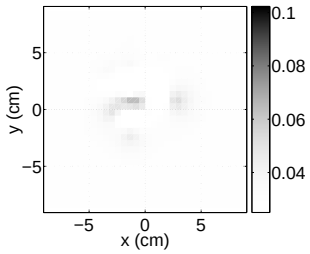
(d) z=-0.26 cm



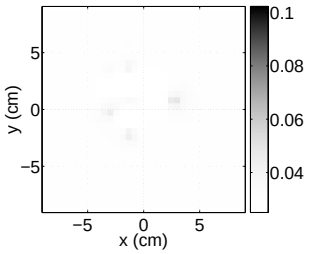
(e) z=0.26 cm



(f) z=0.78 cm

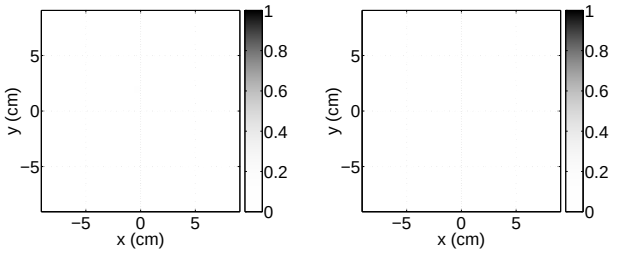


(g) z=1.30 cm

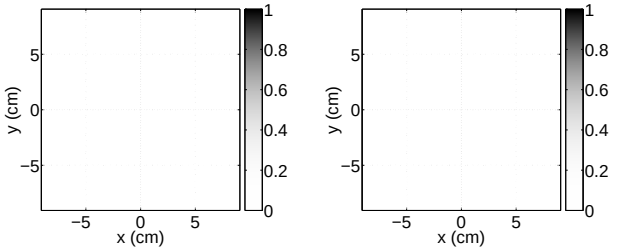


(h) z=1.82 cm

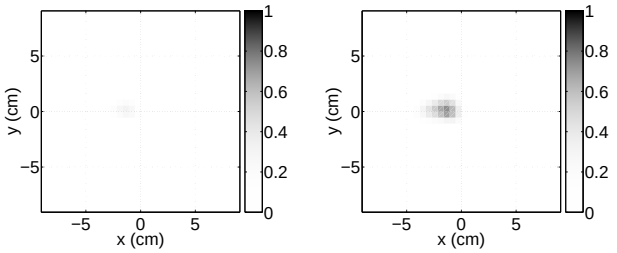
Fig. 12. Reconstructions of μ_{a_m} in cm^{-1}



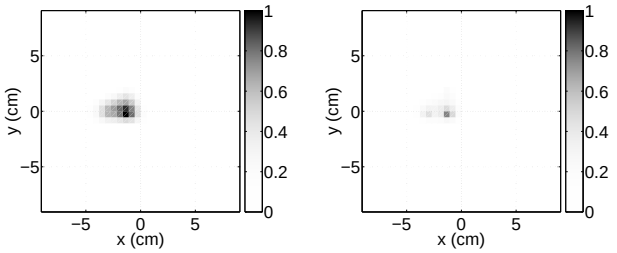
(a) z=-1.82 cm (b) z=-1.30 cm



(c) z=-0.78 cm (d) z=-0.26 cm



(e) z=0.26 cm (f) z=0.78 cm



(g) z=1.30 cm (h) z=1.82 cm

Fig. 13. Reconstructions of $\eta\mu_{a_f}$ in arbitrary units