

# A Statistical Approach to Inverting the Born Ratio

Damon Hyde\*, Eric Miller, Dana H. Brooks, and Vasilis Ntziachristos

**Abstract**—We examine the problem of fluorescence molecular tomography using the normalized Born approximation, termed herein the *Born ratio*, from a statistical perspective. Experimentally verified noise models for received signals at the excitation and emission wavelengths are combined to generate a stochastic model for the Born ratio. This model is then utilized within a maximum likelihood framework to obtain an inverse solution based on a fixed point iteration. Results are presented for three experimental scenarios: phantom data with a homogeneous background, phantoms implanted within a small animal, and *in vivo* data using an exogenous probe.

**Index Terms**—Fluorescence, inverse problems, optical imaging.

## I. INTRODUCTION

IN recent years there has been a growing interest in the use of fluorescence in tomographic biological imaging [1]–[12]. While planar epi-illumination imaging has long been used in combination with fluorescence, its utility, especially in imaging of intact animals, is limited by the lack of three-dimensional information in the resulting images; the planar images acquired are projections of the 3-D fluorochrome distribution onto the 2-D imaging plane, blurred by the diffusive nature of tissue. Moreover, differences in tissue optical properties result in a modulation of signal intensity that further complicates quantification. This results in an image with no depth resolution, from which it is difficult to obtain quantitative information. Fluorescence tomography, by contrast, combines measured data with a model of the diffusion of light through tissue to obtain a three-dimensional reconstruction of the fluorochrome distribution through the solution of an inverse problem.

The near infrared range, between 650 and 900 nm, is commonly used for optical imaging of tissue due in part to a window in the absorption spectrum of hemoglobin and water that allows light to penetrate through depths of several centimeters in biological tissue. In fluorescence tomography, a laser with a wavelength within this range is chosen such that it matches the excitation wavelength of a known fluorescing agent present in the

tissue. In vivo, this agent can be present either as a result of direct introduction of an exogenous probe [13]–[15] or through genetic manipulation, producing transgenic species with naturally occurring levels of fluorescent proteins [16].

The targeting ability and specificity with which these fluorescence contrast agents can be designed gives fluorescence molecular tomography a distinct advantage over traditional diffuse optical tomography (DOT). The resulting contrast can be significantly higher than that naturally present in the tissue due to endogenous chromophores such as oxy- and deoxyhemoglobin, fat, and water, or fluorescent metabolites and other structural elements such as NADH or elastin and collagen. This higher contrast enables fluorescence tomography to more easily differentiate between the desired target and surrounding tissue.

Conversion of the received data into a three-dimensional tomographic image is achieved through solution of an inverse problem, which is nonlinear and ill-posed. In the case of fluorescence tomography, the system to be inverted is a set of two coupled differential equations that describe the behavior of light within tissue at both the fluorochrome's excitation and emission wavelengths. One method for obtaining a solution is to explicitly construct forward models for each wavelength and directly solve the joint inverse problem [11], [17]. As an intermediary, these methods often necessitate solving for the intrinsic optical properties of the tissue being imaged—parameters that may not be of interest.

To address problems in which one is only interested in the distribution of fluorochrome, a data normalization scheme was previously introduced, referred to as the normalized Born approximation, or, in short, the *Born ratio* [2], [18], [19]. By dividing measurements at the fluorescence wavelength by their corresponding excitation wavelength measurements, an effective approximation to the joint problem can be obtained, which at the same time grants a number of benefits [19]. In particular, the Born ratio has been shown to achieve a considerable degree of invariance to nonlinearities created by imaging through an inhomogeneous medium. These factors include effects such as the boundary conditioning of the diffusion approximation, the knowledge of the exact background optical properties of the region in question, and the variation in background optical properties [19]. This invariance allows us to obtain accurate fluorescence images without directly solving the set of two coupled differential equations, which is a considerably more complicated computational problem. Additionally, the Born ratio is independent of source intensities and detector coupling coefficients, which otherwise must be explicitly solved for before obtaining an image. While there exist methods to directly solve for these parameters, they involve the inclusion of additional parameters in the inversion, which have the potential to make the problem more ill posed and thus more difficult to solve [20].

Manuscript received April 12, 2006; revised January 23, 2007. Asterisk indicates corresponding author.

\*D. Hyde is with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA, and the Center for Molecular Imaging Research, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129 USA (e-mail: dhyde@ece.neu.edu).

E. Miller and D. H. Brooks are with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA (e-mail: elmiller@ece.neu.edu; brooks@ece.neu.edu).

V. Ntziachristos is with the Center for Molecular Imaging Research, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129 USA (e-mail: vasilis@helix.mgh.harvard.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2007.895467

However, the Born ratio approach has its own drawbacks. In particular, the information content, or signal-to-noise ratio, of the resulting normalized measurements is difficult to quantify. A large value of the ratio for a particular source–detector pair can be the result of a large fluorescence signal, but as well can be the result of a small excitation signal, while a small ratio value can mean either weak fluorescence or strong excitation. Current methods use empirical threshold levels to reject data pairs with sufficiently low excitation and/or fluorescence signals. While useful results can be obtained, these thresholds introduce user variability into the reconstructions, and as such do not necessarily make optimal use of the available data. The ability to automatically determine and retain only the significant Born ratio values for subsequent inversion, based on an objective criterion, would be of significant utility in order to optimize imaging performance, and thus would offer an improvement over previously applied empirical thresholds based on calibration studies.

In this paper, we introduce a statistical method that allows us to determine an automatic data-weighting scheme based on an approximate statistical model of the Born ratio data and of its relationship to the unknown fluorescence distribution. As we report below, experimental results in a variety of phantom and animal preparations show that this approach provides consistent, repeatable, and operator-independent reconstructions using the Born ratio.

Our model begins with the use of the central limit theorem to approximate the Poisson distribution of the received signals using Gaussian random variables. However, the *ratio* of two Gaussian random variables (RVs), which models the Born ratio, does not lead to an easily tractable distribution. The most commonly known case is when both random variables are zero mean with unit variance, which leads to a Cauchy density. The presence of nonzero means further complicates the distribution, although results exist in the literature that give the density of such a ratio as a function of the means. However, this density does not lend itself to a simple relationship between the data and the unknowns. Here, by using a combination of approximations, point estimates and a fixed point iteration, we obtain an approximate maximum-likelihood (ML) estimate as the solution to a sequence of iteratively reweighted linear least squares problems, each of which is solved using a standard conjugate gradient algorithm. The weights in this formulation play a role equivalent to the case-by-case thresholds used to date but are determined automatically via the statistical model.

We describe in detail below our statistical model for the measurements, based on experimentally measured noise characteristics of a previously described fluorescence molecular tomography system [2] and how it relates to the data and the forward model employed. We then present inversion results obtained from a variety of different data sets. These include results for simple phantoms in a homogeneous background, phantoms placed in a mouse body (and thus with an inhomogeneous background), and *in vivo* mouse experiments. Our results indicate that the fixed point iteration converges quickly and that the resulting reconstructions are comparable to the best hand-tuned Born ratio reconstructions across a variety of experimental scenarios, but without the need for empirically set thresholds.

Section II gives an overview of the forward modeling associated with this system, whose result is a linear, matrix-based forward model. Section III discusses statistical models for each of the individual received signals and then uses them to develop a statistical model for the Born ratio itself. The resulting model is used in Section IV to derive an inverse solution. Results for the various data sets are presented in Section V. In Section VI, we offer some conclusions as well as suggestions for continuation of this research.

## II. ANALYTIC MODELING

The propagation of light through a turbid medium such as tissue can be reasonably approximated by a diffusion equation for a source-detector separation of more than a few mean scattering lengths [21]. Using Green's function theory, the equation for the observed fluorescence  $U_{\text{fl}}(r_d, r_s)$  at a detector location  $r_d$ , given a source at location  $r_s$ , can be modeled as [18]

$$U_{\text{fl}}(r_d, r_s) = Q_E^{\lambda_{\text{fl}}} \Theta_f^{\text{fl}} \times \int d^3r \Theta_{\text{det}}(r_d) G_o^{\lambda_{\text{fl}}}(r_d, r) \times \frac{\nu}{D^{\lambda_{\text{fl}}}} x(r) G_o^{\lambda_{\text{exc}}}(r, r_s) \Theta_{\text{src}}(r_s). \quad (1)$$

This equation is the integral form of the Born approximation for the received fluorescence signal, which assumes that the total received signal is the sum of the first-order scattering responses across the entire volume. Here,  $Q_E^{\lambda_{\text{fl}}}$  is the quantum efficiency of the detector at the fluorescence wavelength,  $\Theta_f^{\text{fl}}$  is the attenuation of the fluorescence filter,  $\Theta_{\text{det}}(r)$ , and  $\Theta_{\text{src}}(r)$  are the detector coupling and source intensity coefficients respectively, while  $\nu$  is the speed of light within the medium and  $D^{\lambda_{\text{fl}}}$  is the diffusion coefficient at the fluorescence wavelength. The parameter  $x(r)$  represents the spatially varying fluorochrome concentration that we want to recover. Finally,  $G_o^{\lambda_{\text{exc}}}(r, r_s)$  and  $G_o^{\lambda_{\text{fl}}}(r_d, r)$  are the two point Green's functions describing the transport of light from source to voxel and voxel to detector at the corresponding wavelengths. The exact values for these two functions can be determined analytically either by assuming the background can be treated as homogeneous or through various nonlinear methods incorporating information about background inhomogeneities [5]–[7]. In the work presented here, we employ the homogeneous background assumption.

### A. The Born Ratio

As described in Section I, to avoid the necessity of determining a number of parameters that are not of direct interest, we employ the data normalization method known as the Born ratio [2], [18], [19]. This method makes use of a second set of data, collected at the excitation wavelength. By dividing the fluorescence wavelength measurement by the corresponding excitation measurement, these variables can be eliminated. If we write the equation for the excitation field  $U_{\text{exc}}(r_d, r_s)$  as

$$U_{\text{exc}}(r_d, r_s) = Q_E^{\lambda_{\text{exc}}} \Theta_f^{\text{exc}} \Theta_{\text{det}}(r_d) \times G_o^{\lambda_{\text{exc}}}(r_d, r_s) \Theta_{\text{src}}(r_s) \quad (2)$$

with quantum efficiency  $Q_E^{\lambda_{\text{exc}}}$  and excitation filter attenuation  $\Theta_f^{\text{exc}}$ , then the so-called Born ratio relationship can be written as

$$\begin{aligned} U(r_s, r_d) &\triangleq \frac{U_{\text{fl}}(r_d, r_s)}{U_{\text{exc}}(r_d, r_s)} \\ &= \alpha \frac{\int d^3r G_o^{\lambda_{\text{fl}}}(r_d, r) x(r) G_o^{\lambda_{\text{exc}}}(r, r_s)}{G_o^{\lambda_{\text{exc}}}(r_d, r_s)} \end{aligned} \quad (3)$$

with

$$\alpha = \frac{Q_E^{\lambda_{\text{fl}}}}{Q_E^{\lambda_{\text{exc}}}} \frac{\Theta_f^{\text{fl}}}{\Theta_f^{\text{exc}}} \frac{\nu}{D^{\lambda_{\text{fl}}}}. \quad (4)$$

The parameter  $\alpha$  relies only upon basic system characteristics and can be easily determined experimentally from measurements of a fluorochrome of known concentration [19]. Thus we are left with an equation containing no excess free parameters.

### B. Discretization

The forward model given by (3) is then discretized and put into matrix form for inversion. The result is a forward model of the form

$$\hat{\mathbf{u}} = \mathbf{W}\mathbf{x} \quad (5)$$

where  $\hat{\mathbf{u}}$  is the vector of measurement ratios predicted by the Born ratio model, given a fluorescence distribution defined by  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$  and a matrix operator  $\mathbf{W}$ , discretized from (3), relating the two. The individual components of the matrix are defined by

$$\mathbf{W}_{ij} = \Delta_v \frac{\alpha G_o^{\lambda_{\text{fl}}}(r_d^i, r^j) G_o^{\lambda_{\text{exc}}}(r^j, r_s^i)}{G_o^{\lambda_{\text{exc}}}(r_d^i, r_s^i)}. \quad (6)$$

Here,  $\Delta_v$  is the volume of each individual voxel and results from the discretization of the volume integral in (3). The index  $i$  denotes the particular source–detector pair corresponding to the  $i$ th element of  $\hat{\mathbf{u}}$  and the index  $j$  corresponds to a particular voxel in the discretized geometry.

## III. STATISTICAL MODELING

### A. Signal Modeling

To begin our statistical analysis, we assume probability models for each of the two received signals. Traditionally, optical measurements are modeled using Poisson statistics. For sufficiently high-rate photon counts, such as those typically present in the systems of interest here, Poisson distributed random variables can be approximated as Gaussian RVs with variances equal to their means. In practice, the nature of the experimental setup and the measurement system alters this relationship slightly, making the variance proportional to the mean rather than equal to it. The proportionality constant is constant across all data points and can be determined experimentally, as we now describe.

We performed a simple experiment to determine the value of this constant. A single source was used to illuminate a chamber

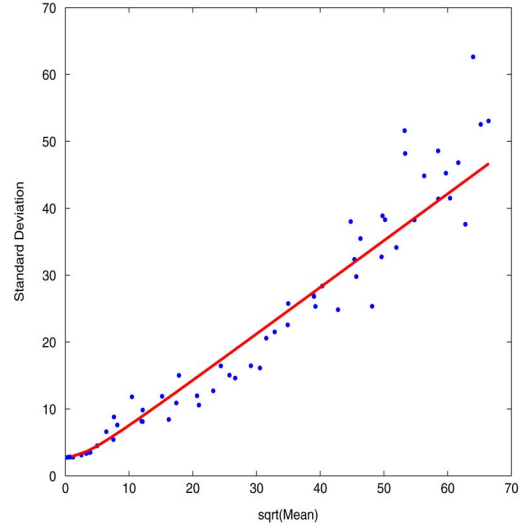


Fig. 1. Scatter plot of experimentally collected noise data, along with a regression line from the model in (7). The  $R^2$  value for this fit is 0.86. In order to better display the results at all intensity levels, the plot shows standard deviation plotted against the square root of the mean.

full of 1% intralipid solution and 50 ppm India ink, and 100 frames of data were taken with a charge-coupled device (CCD) camera at each of four different laser intensities. Radial bins were then defined as concentric rings centered on the location of the source in the camera plane. Data from equivalent bins were pooled across all 100 frames, and the resulting data were used to compute experimental mean and variance values. Plotting the mean and variance pairs against each other in a scatter plot, a linear regression line was calculated, with the slope of that line giving the value of the proportionality constant. Means and signal levels here, and through the rest of this paper, are reported in counts, representing the digital output of the CCD's 16-bit analog-to-digital converter.

We note then an additional perturbation to the Gaussian approximation of the Poisson comes from the addition of nonnegligible Gaussian readout noise from the CCD. As it turns out, we can also estimate the variance of this noise (it is well approximated at zero mean) from the experiment just described, as the  $y$ -axis intercept of the regression line.

The data from this experiment are displayed in Fig. 1, and the linear fit ( $R^2 = 0.86$ ) resulted in a proportionality constant  $\beta = 0.4915$  and read noise standard deviation  $\sigma_r = 2.8$ . It should be noted that while the proportionality factor is less than one, this does not indicate a reduction in noise levels below Poisson noise. It is a result of the fact that each increment of the CCD output corresponds to more than a single received photon. Examining the experimental values, at low signal intensities the readout noise contributes enough power to the measurement compared to the signal that we explicitly include it as an additive noise factor. Thus our model for the variance of the received signal at either wavelength can be written as a function of the mean  $m$

$$\sigma^2(m) = \beta m + \sigma_r^2. \quad (7)$$

Note that the variance at each spatial location depends on the mean of the “signal” component.

In order to write an expression for the joint density of the fluorescence and excitation measurements, so that we can then derive an expression for the density of their ratio, we need to model their covariance both across space at a fixed wavelength and also across wavelengths. In the work described in this paper, we make the simplifying assumption that the measurements are uncorrelated both in space and across wavelengths. This results in a joint density for the received signals of

$$\begin{bmatrix} \mathbf{u}_{\text{fl}} \\ \mathbf{u}_{\text{exc}} \end{bmatrix} \sim N \left( \begin{bmatrix} \mathbf{m}_{\text{fl}} \\ \mathbf{m}_{\text{exc}} \end{bmatrix}, \text{diag} \left( \beta \begin{bmatrix} \mathbf{m}_{\text{fl}} \\ \mathbf{m}_{\text{exc}} \end{bmatrix} + \sigma_r^2 \right) \right). \quad (8)$$

While these independence assumptions are not physically accurate, they greatly simplify the process of finding an inverse solution. As we will see in our results section, in practice they do not prevent us from obtaining useful solutions.

### B. Ratio Distribution

Given a stochastic model for our received signals, we now write the Born ratio as

$$u_i = \frac{u_{\text{fl},i}}{u_{\text{exc},i}} \quad (9)$$

where the densities for  $u_{\text{fl},i}$  and  $u_{\text{exc},i}$  are as previously defined. The work in [22] describes an explicit derivation of the density associated with the ratio of two arbitrary Gaussian random variables. The resulting model incorporates correlation between the numerator and denominator in addition to arbitrary means and variances. Since we assume independence, we can remove terms involving the correlation coefficient and write that density as

$$\begin{aligned} p(u_i) = & \frac{b(u_i)d(u_i)}{\sqrt{2\pi}\sigma(m_{\text{fl},i})\sigma(m_{\text{exc},i})a(u_i)^3} \\ & \times \left[ \Phi \left\{ \frac{b(u_i)}{a(u_i)} \right\} - \Phi \left\{ -\frac{b(u_i)}{a(u_i)} \right\} \right] \\ & + \frac{1}{\pi\sigma(m_{\text{fl},i})\sigma(m_{\text{exc},i})a(u_i)^2} e^{-0.5c} \end{aligned} \quad (10)$$

where  $a(u_i)$ ,  $b(u_i)$ ,  $c$ , and  $d(u_i)$  are defined as

$$\begin{aligned} a(u_i) &= \left( \frac{u_i^2}{\sigma^2(m_{\text{fl},i})} + \frac{1}{\sigma^2(m_{\text{exc},i})} \right)^{1/2} \\ b(u_i) &= \frac{m_{\text{fl},i}u_i}{\sigma^2(m_{\text{fl},i})} + \frac{m_{\text{exc},i}}{\sigma^2(m_{\text{exc},i})} \\ c &= \frac{m_{\text{fl},i}^2}{\sigma^2(m_{\text{fl},i})} + \frac{m_{\text{exc},i}^2}{\sigma^2(m_{\text{exc},i})} \\ d(u_i) &= \exp \left\{ -0.5 \frac{m_{\text{exc},i}^2 \left( \frac{m_{\text{fl},i}}{m_{\text{exc},i}} - u_i \right)^2}{(\sigma^2(m_{\text{exc},i})u_i + \sigma^2(m_{\text{fl},i})))} \right\}. \end{aligned} \quad (11)$$

$\Phi(u)$  is the usual integral of the unit normal density  $\phi(y)$

$$\begin{aligned} \Phi(u) &= \int_{-\infty}^u \phi(y) dy \\ \phi(u) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}. \end{aligned} \quad (12)$$

We next simplify the expression in (10). Our first step is to eliminate the terms involving  $\Phi$  by making the approximation

$$\left[ \Phi \left\{ \frac{b(u_i)}{a(u_i)} \right\} - \Phi \left\{ -\frac{b(u_i)}{a(u_i)} \right\} \right] \simeq 1. \quad (13)$$

We show in Appendix I that when either  $m_{\text{fl},i} > 18$  counts or  $m_{\text{exc},i} > 18$  counts is satisfied,<sup>1</sup> the value of  $(b(u_i))/(a(u_i))$  is sufficiently large to make the approximation valid with a probability of  $(1 - 2 \times 10^{-6})$ . Thus we simply remove from the reconstruction all data pairs that do not satisfy at least one of these two inequalities.<sup>2</sup>

In Appendix II, we show that by slightly expanding our previous minimum signal requirements to be  $m_{\text{fl},i} > 18$  counts and  $m_{\text{exc},i} > 20$  counts, we can ensure that the second term in (10) is several orders of magnitude smaller than the first. Thus we also can drop the second term. This leaves

$$p(u_i) \simeq \left[ \frac{b(u_i)d(u_i)}{\sqrt{2\pi}\sigma(m_{\text{fl},i})\sigma(m_{\text{exc},i})a(u_i)^3} \right] \quad (14)$$

as our model for the statistical distribution of the normalized Born ratio data.

### C. Incorporating a Forward Model

We are now interested in relating the parameters of (14) to the collected data and the spatial fluorescence distribution. Given that the standard deviations of each density depend only upon the mean, there are two parameters that must be solved for:  $m_{\text{exc},i}$  and  $m_{\text{fl},i}$ .

Because  $m_{\text{exc},i}$  is the mean of the excitation signal, it is independent of the spatial distribution of fluorochrome. In order to evaluate it explicitly, however, a separate forward model is required, as well as knowledge about the background optical inhomogeneities of the medium. The construction of such a model, and determination of the associated parameters, is the problem of traditional diffuse optical tomography. Due to the complexity of such problems, and the fact that we are only interested in the fluorescence distribution, we avoid the requirement for a separate forward model by making the approximation

$$m_{\text{exc},i} \simeq u_{\text{exc},i}. \quad (15)$$

Thus, we assume that the mean of the excitation signal can be approximated by the measured values. Next, we make the as-

<sup>1</sup>We note that full range is  $2^{16} = 65536$  counts, so 18 counts is  $<0.3\%$  of the available dynamic range.

<sup>2</sup>Note that the introduction of this threshold does not contradict our desire to eliminate existing threshold-based techniques because the threshold here is fixed at an analytically determined level and is not operator dependant.

sumption that our forward model (15) can be related to the statistics as

$$\mathbf{W}_i \mathbf{x} = \frac{m_{\text{fl},i}}{m_{\text{exc},i}} \quad (16)$$

that is, we assume that the forward model generates the ratio of the means of the collected data. Thus, we approximate  $m_{\text{fl},i}$  as

$$m_{\text{fl},i}(\mathbf{x}) \simeq u_{\text{exc},i} \mathbf{W}_i \mathbf{x}. \quad (17)$$

Equations (15) and (17) now allow us to rewrite (14) in a form conditioned upon the image  $\mathbf{x}$

$$p(u_i | \mathbf{x}) = \left[ \frac{b(u_i, \mathbf{x}) d(u_i, \mathbf{x})}{\sqrt{2\pi} \sigma(m_{\text{fl},i}(\mathbf{x})) \sigma(m_{\text{exc},i}) a(u_i, \mathbf{x})^3} \right] \quad (18)$$

which allows us to proceed with a statistical inverse solution.

#### IV. STATISTICAL DERIVATION

We seek an inverse solution to the Born ratio using a ML approach. The general form for an ML solution is

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} p(\mathbf{u} | \mathbf{x}). \quad (19)$$

Using (18) as the density for each individual point, and given our independence assumptions, the right-hand side of (19) becomes

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} \prod_i \left[ \frac{b(u_i, \mathbf{x}) d(u_i, \mathbf{x})}{\sqrt{2\pi} \sigma(m_{\text{fl},i}(\mathbf{x})) \sigma(m_{\text{exc},i}) a(u_i, \mathbf{x})^3} \right]. \quad (20)$$

Taking the log of the right-hand side and dropping terms that do not depend on  $\mathbf{x}$  leaves

$$\begin{aligned} \hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} \sum_i & [\log(b(u_i, \mathbf{x})) \\ & - \log(\sigma(m_{\text{fl},i}(\mathbf{x}))) \\ & 3 \log(a(u_i, \mathbf{x})) + \log(d(u_i, \mathbf{x}))]. \end{aligned} \quad (21)$$

To simplify notation, we define a function

$$g_i(u_i, \mathbf{x}) = \log(b(u_i, \mathbf{x})) - \log(\sigma(m_{\text{fl},i}(\mathbf{x}))) - 3 \log(a(u_i, \mathbf{x})) \quad (22)$$

and a vector  $\mathbf{q}(\mathbf{u}, \mathbf{x})$  with elements

$$q_i(u_i, \mathbf{x}) = \frac{m_{\text{exc},i}}{\sqrt{\sigma(m_{\text{exc},i})^2 u_i + \sigma(m_{\text{fl},i}(\mathbf{x}))^2}} \quad (23)$$

and obtain

$$\begin{aligned} \hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} \sum_i & \left[ g_i(u_i, \mathbf{x}) - 0.5 \right. \\ & \left. \times \left( q_i(u_i, \mathbf{x}) \left( \frac{m_{\text{fl},i}(\mathbf{x})}{m_{\text{exc},i}} - u_i \right) \right)^2 \right]. \end{aligned} \quad (24)$$

Rewriting the second term in (24) into matrix form, we get

$$\begin{aligned} \hat{\mathbf{x}}_{\text{ML}} = \arg \max_{\mathbf{x}} & \left[ \sum_i g_i(u_i, \mathbf{x}) \right. \\ & \left. - 0.5 \|\mathbf{Q}(\mathbf{u}, \mathbf{x})(\mathbf{W}\mathbf{x} - \mathbf{u})\|_2^2 \right] \end{aligned} \quad (25)$$

where we have defined the matrix  $\mathbf{Q}(\mathbf{u}, \mathbf{x}) = \text{diag}(\mathbf{q}(\mathbf{u}, \mathbf{x}))$ .

Now we want to take the gradient of (25) with respect to the image  $\mathbf{x}$  and set it equal to zero. If we define the cost function over which we are maximizing to be  $f(\mathbf{x})$ , we can use the results of Appendices III-A and III-B to write the gradient of that cost function as

$$\begin{aligned} \frac{df(\mathbf{x})}{d\mathbf{x}} = & \left[ \sum_i \frac{dg_i(u_i, \mathbf{x})}{d\mathbf{x}} \right] \\ & - \mathbf{W}^T \mathbf{Q}(\mathbf{x})^T \mathbf{Q}(\mathbf{x}) (\mathbf{W}\mathbf{x} - \mathbf{u}) \\ & - \left( \frac{d\mathbf{q}(\mathbf{x})}{d\mathbf{x}} \right)^T A(\mathbf{x})^T A(\mathbf{x}) \mathbf{q}(\mathbf{u}, \mathbf{x}) \end{aligned} \quad (26)$$

where we have defined  $A(\mathbf{x})$  in Appendix III-A to be

$$A(\mathbf{u}, \mathbf{x}) = \text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u}). \quad (27)$$

Setting (26) to zero and rearranging, we obtain

$$\begin{aligned} \mathbf{x} = & [\mathbf{W}^T \mathbf{Q}(\mathbf{x})^T \mathbf{Q}(\mathbf{x}) \mathbf{W}]^{-1} \\ & \times \left[ \sum_i \frac{dg_i(u_i, \mathbf{x})}{d\mathbf{x}} + \right. \\ & \times \mathbf{W}^T \mathbf{Q}(\mathbf{x})^T \mathbf{Q}(\mathbf{x}) \mathbf{u} \\ & \left. - \left( \frac{d\mathbf{q}(\mathbf{x})}{d\mathbf{x}} \right)^T A(\mathbf{x})^T A(\mathbf{x}) \mathbf{q}(\mathbf{x}) \right]. \end{aligned} \quad (28)$$

Given the dependence on the solution  $\mathbf{x}$  of the terms on the right, this result suggests a fixed point iteration, which, calling the right-hand side  $\zeta(\mathbf{x})$ , can be written as

$$\mathbf{x}^{(n+1)} = \zeta(\mathbf{x}^{(n)}). \quad (29)$$

Solutions for each iteration of (29) were obtained using a conjugate gradient (CG) solver, chosen because the matrix being inverted in (28) is symmetric positive definite. The CG algorithm utilizes the forward matrix  $\mathbf{W}$  only in matrix vector multiplications of the form  $\mathbf{W}\mathbf{x}$  or  $\mathbf{W}^T \mathbf{x}$ , which allows for the matrix to be utilized without ever being explicitly stored in memory.

To obtain a value for  $\mathbf{x}^{(0)}$ , an arbitrarily high threshold, based solely on the excitation signal, was set such that 75% of the data pairs were removed from the reconstruction, and a solution was then obtained using a CG algorithm applied to the normal equations, without the use of any statistical models. This provided a suitable initial estimate for the fixed point iteration.

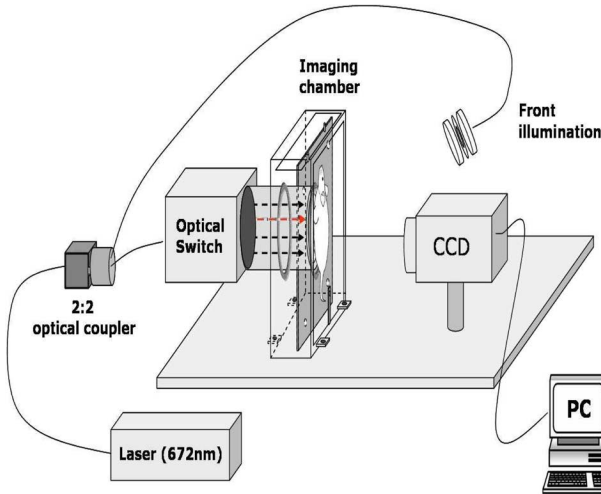


Fig. 2. Diagram of imaging system.

## V. EXPERIMENTAL METHODS AND RESULTS

The experimental setup we utilized is based on a slab geometry of 1.3 cm thickness. A diagram of the system is shown in Fig. 2. The illumination consisted of a laser diode at 672 nm, passed through an optical switch (DiCon Fiberoptics, Richmond, CA) to select from an array of 46 source fibers arranged on a  $1.8 \times 1.2 \text{ cm}^2$  area. Output power at the imaging chamber was approximately 10 mW. For detection, a CCD camera (Roper Scientific, Trenton, NJ), electrically cooled to  $-35^\circ\text{C}$ , was used. Selection of excitation and emission wavelengths was achieved by appropriate filters: bandpass (three-cavity interference at  $670 \pm 5 \text{ nm}$  for excitation measurements and  $710 \pm 10 \text{ nm}$  for emission measurements (Andover, Salem, NH) and long-pass (emission wavelength: cutoff 695 nm; Omega Optical, Brattleboro, VT). A grid of  $13 \times 13$  virtual detector locations was defined, corresponding to uniform spacing over a  $2.2 \times 2.2 \text{ cm}^2$  region centered around the source array, on the opposite side of the illuminated volume from the sources. Individual detector values were obtained by integrating the values of CCD pixels corresponding to 1-mm-diameter equivalent circular detectors on the camera. The solution space, defined to be the  $2.4 \times 2.4 \times 1.3 \text{ cm}^3$  region centered around the sources, was discretized into  $20 \times 20 \times 21 = 8400$  volume elements (voxels), each of dimension  $0.12 \times 0.12 \times 0.0619 \text{ cm}^3$ . On all reconstruction images, this region is demarcated by a red box.

In all cases described below, the imaging chamber was filled with a solution of 1% intralipid and 0.5% India ink, which has optical properties  $\mu_a = 0.58 \text{ cm}^{-1}$  and  $\mu_s = 10 \text{ cm}^{-1}$ . These values were chosen to correspond closely with the bulk background optical parameters of mice.

The first set of data employed came from phantom measurements, providing a simple scenario to validate our method. The phantom consisted of two 8 mm long and 2.5-mm-wide tubes, each containing 400 nM Cy5.5, suspended in a tank of intralipid solution as described. The two tubes were placed 5 mm apart along the lateral axis, with one tube located against the detector side of the volume and the other in its center, corresponding to distances of 0.6 and 1.2 cm from the source plane. Fig. 3(a) and

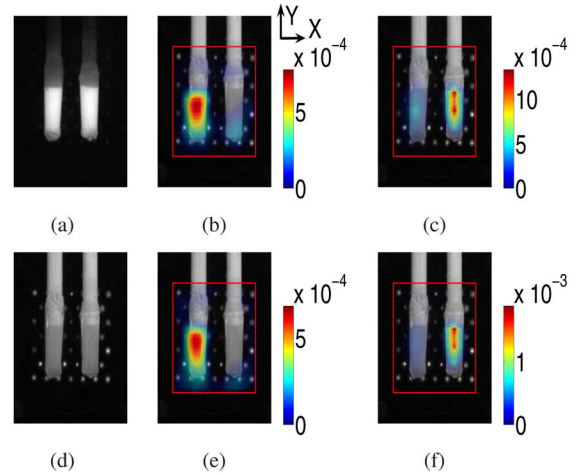


Fig. 3. Dual tube phantom. (a) Reflectance fluorescence image. (b) Manual threshold: coronal slice through tube at  $z = 0.42 \text{ cm}$ . (c) Manual threshold: coronal slice through tube at  $z = 1.25 \text{ cm}$ . (d) Flat image. (e) Fixed point: coronal slice through  $z = 0.42 \text{ cm}$ . (f) Fixed point: coronal slice through  $z = 1.25 \text{ cm}$ . In all images,  $z = 0 \text{ cm}$  corresponds to the source side of the slab, with a total slab thickness of 1.3 cm. The red box in the image denotes the boundary of the solution space, and the orientation of the coordinate system specified in (b) is shared by all images.

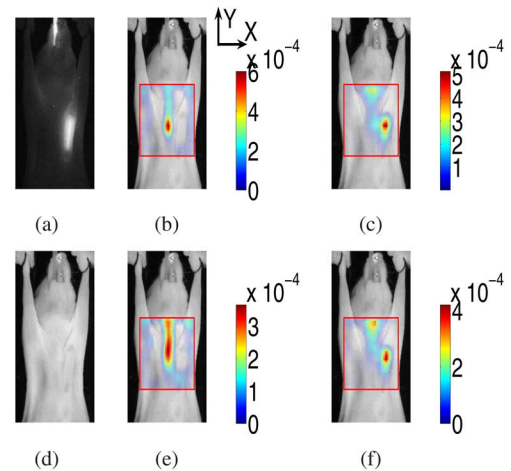


Fig. 4. Euthanized mouse. (a) Fluorescence reflectance. (b) Manual threshold: coronal slice through esophageal tube. (c) Manual threshold: coronal slice through subcutaneous tube. (d) Flat image. (e) Fixed point: coronal slice through esophageal tube. (f) Fixed point: coronal slice through subcutaneous tube. In all images,  $z = 0 \text{ cm}$  corresponds to the source side of the slab, with a total slab thickness of 1.3 cm. The red box denotes the boundary of the solution space. (a) and (b)  $z = 0.59 \text{ cm}$ , (c)  $z = 1.25 \text{ cm}$ , (d) and (e)  $z = 0.59 \text{ cm}$ , (f)  $z = 1.25 \text{ cm}$ .

(d) shows flat images at the fluorescence and excitation wavelength, respectively. They were taken prior to the chamber's being filled with intralipid, and while the objects appear to be at the same distance from the camera, they are in fact separated by 0.4 cm along the source-camera axis.

The second set of data examined consisted of two plastic catheter tubes with an inner diameter of 0.8 mm, containing 1  $\mu\text{M}$  Cy5.5, implanted in the body of an euthanized mouse. As can be seen in the fluorescence reflectance and flat images [Fig. 4(a) and (d) respectively], one tube was inserted down the esophagus while the other was implanted subcutaneously on the

anterior side of the torso. The goal of this experiment was to provide an easily identifiable and quantifiable target (the tubes) under conditions where the background optical parameters are similar to those of an *in vivo* experiment.

The third study presented considered an *in vivo* experiment. A Her2-neu transgenic mouse was selected, exhibiting a spontaneous breast tumor. The tumor was located in the left mammary pad as indicated by an arrow on Fig. 5(a) and was approximately  $5 \times 6$  mm in size, as measured externally with calipers. The animal was injected with 2 nmol of the ProSense680 cathepsin-sensitive activatable probe (VisEn Medical, Woburn, MA) 24 h prior to imaging. Accompanying the reconstructions is a flat image at the excitation wavelength in Fig. 5(a), on which the tumor is visible as a dark region in the left mammary pad (right side of image).

For each data set, we present and compare two methods. In the first method we solved using the normalized Born ratio after thresholding to reject all ratio data with an excitation wavelength intensity less than a fixed value. The threshold values used for each data set were determined by obtaining inverse solutions for thresholds set between the zeroth and ninety-fifth percentile of the excitation wavelength measurements, in 5% increments. The “optimal” solution was then chosen from among these 20 results, based on prior knowledge of the target being imaged. The chosen thresholds varied from 15% to 35% of the Born ratio values being retained for use in the inversion algorithm.

The second method utilized the statistical methods derived in Section IV. The previously established threshold requiring  $\mathbf{u}_{\text{exc},i} > 20$  counts or  $\mathbf{u}_{\text{fl},i} > 18$  counts was applied to the data prior to inversion, resulting in between 0.1% and 32% of the data’s being removed. For both reconstruction methods, depth-dependent regularization was achieved by making the stopping point of the iterations increase exponentially as a function of depth from either four or ten iterations, which optimally reconstructed surface activity, to 25 for the middle reconstructed slice. Further iterations resulted in only minor changes in the reconstruction. We note that compared to utilizing a depth-independent number of iterations, the depth-dependent regularization employed does not alter the relative performance of the statistical versus the automatic method but yields a more accurate reconstruction of the physical dimensions of the fluorescence activity, as was also noted in [23].

#### A. Comparison Metrics

In addition to subjective visual analysis of results, we also used two metrics to compare our statistical method to existing threshold based methods. These are used to quantify the location and size of the objects.

In order to compute these two metrics, we first needed to determine the extent of the reconstructed object. Given the differences in intensity owing to the different depths at which fluorescent objects are located, we started this process by manually selecting, for each object, a slice that contained a clear cross-section of that object. We then took the voxel in that slice with the highest amplitude to be the initial location of the detected object. Selecting a threshold of 50% of this initial value, we

allowed the object to grow by iteratively incorporating neighboring voxels in three dimensions whose amplitude were above the threshold. In this manner, we determined which voxels were considered to lie within the resolved object.

The two metrics used to quantify the reconstructions were the centroid of the reconstructed object and its average dimensions. The centroid was computed as the mean of the objects constituent voxel locations, weighted by the associated image intensities. Average dimensions were determined by first computing the average location of each side of the object, then taking the distance between the two sides associated with each of the  $X$ ,  $Y$ , and  $Z$  axes.

#### B. Phantoms in Intralipid

In addition to the reconstructions, Fig. 3 displays two images of the phantoms without intralipid present. Fig. 3(d), the flat image, is a front illuminated image of the chamber, while Fig. 3(a), the fluorescence reflectance, is an image captured at the fluorescence wavelength while front illuminating the chamber with light at the excitation wavelength. Note that the reconstruction slices displayed in Fig. 3 are chosen at the depth of maximum intensity in the statistical reconstruction and were obtained using four iterations for the surface object and 25 for the central one.

Using the manually optimized method, the threshold was set to retain 15% of the data for use in the reconstruction. Fig. 3(b) and (c) depicts two coronal slices of the reconstructed three-dimensional image. In each slice, the solution clearly shows a fluorescing object present at the location of the phantoms. The slice located deeper within the volume, shown in Fig. 3(b), resolves the object to be larger in size compared to the slice in (c), a result of decreasing resolution at greater depths.

The results with our statistical inversion are illustrated in Fig. 3(e) and (f), where very similar reconstructions are seen. The object in Fig. 3(e) is slightly larger than that in (b) but remains consistent with the size of the phantoms seen in the flat and fluorescence images. Fig. 3(f) shows an image very similar to that in (c). These images were achieved with the use of 99.9% of the collected data. Only about 100 of the 14 000 data points collected did not satisfy the minimum signal requirements.

The sizes reconstructed with the two methods are reasonably consistent with one another. Notably, in both cases, the resolution is poorest along the  $Z$ -axis between the source and detector planes. Both methods overestimate the object size along this dimension. Additionally, while the thickness ( $Z$ -axis measurement) of the central tube is overestimated, that of the detector side tube is underestimated. This variation in reconstructed object thickness is seen for both methods, with all data sets, and is due to the restrictions on source and detector locations when using a slab geometry.

#### C. Euthanized Mouse

As with the tubes in intralipid, flat and reflectance fluorescence images are displayed in Fig. 4(d) and (a), alongside reconstructions corresponding to the slice at which the maximum intensity was obtained with the statistical reconstruction method. As with the tube phantom, these reconstructions use four iterations for the subcutaneous tube and 25 for the esophageal one.



TABLE I

TABLE OF ANALYTIC METRICS COMPUTED FOR EACH SOLUTION IMAGE. CENTROID AND MAXIMUM DIMENSION ARE GIVEN IN ( $X$ ,  $Y$ ,  $Z$ ) COORDINATES EXCEPT FOR THE SUBCUTANEOUS TUBE IN THE EUTHANIZED MOUSE, WHERE WIDTH, LENGTH, AND DEPTH OF THE TUBE ARE GIVEN, AS IT IS NOT PARALLEL WITH ANY OF THE THREE PRIMARY AXES. MEASUREMENTS ON THE RECONSTRUCTIONS ARE COMPUTED AS DESCRIBED IN THE TEXT. ALL SIZE MEASUREMENTS ARE GIVEN IN CENTIMETERS

Data Set	Object	Inversion Method	Centroid Location	Average Dimensions
Phantom	Central	Threshold	(-0.53,-0.18,0.55) cm	$0.47 \times 0.66 \times 0.71$
		Statistical	(-0.54,-0.20,0.53) cm	$0.45 \times 0.73 \times 0.69$
		Actual Value	(-0.51,-0.23,0.6) cm	$0.48 \times 0.8 \times 0.48$
	Side	Threshold	(0.56,-0.18,1.15) cm	$0.27 \times 0.62 \times 0.26$
		Statistical	(0.55,-0.03,1.18) cm	$0.27 \times 0.56 \times 0.21$
		Actual Value	(0.57, -0.23, 1.2) cm	$0.48 \times 0.8 \times 0.48$
Euthanized	Esophageal	Threshold	(0.05,-0.17,0.64) cm	$0.28 \times 0.53 \times 0.59$
		Statistical	(0.06,-0.06,0.61) cm	$0.36 \times 1.35 \times 0.76$
		Actual Values	Unknown	$0.2 \times 1.5 \times 0.2$
	Subcutaneous	Threshold	(0.60,-0.21,1.14) cm	$0.34 \times 0.61 \times 0.24$
		Statistical	(0.66,-0.14,1.12) cm	$0.30 \times 0.56 \times 0.30$
		Actual Values	Unknown	$0.2 \times 1.0 \times 0.2$
In-Vivo	Tumor	Threshold	(0.34,0.20,1.22) cm	$0.20 \times 0.63 \times 0.13$
		Statistical	(0.56,0.32,1.14) cm	$0.32 \times 0.59 \times 0.28$

The results using the threshold method are shown in Fig. 4(b) and (c). These were obtained by retaining 35% of the data. As with the phantom, two coronal slices are shown, one through each of the implanted tubes. Fig. 4(c) shows a vaguely tube-like object located just under the surface. Its location corresponds well with the fluorescence image seen in Fig. 4(a) but is not nearly as large as would be expected. The object in the center of the volume can be seen in Fig. 4(b). However, rather than appearing to be a uniform tube, the object appears to be more centrally concentrated.

Fig. 4(b) and (c) contains the results for this data set using our fixed point iteration. The subcutaneous tube is resolved in a manner similar to that using the thresholding technique. It is, however, slightly larger, more tube-like, and more consistent with Fig. 4(a). Looking at the esophageal tube, it is resolved to be a much more uniform tube than it was in Fig. 4(b). For this data set, the minimum signal requirements resulted in 32% of the data's being left out of the reconstruction, as with the threshold method a significantly higher number than for the phantom experiment. This seems to be a result of a higher level of absorption within the mouse body. While for the other two data sets, fewer than 200 excitation measurements failed the minimum criterion, for the euthanized mouse, nearly 3700 points failed the test. The additional absorption also seems to have affected the resulting fluorescence levels, as a significantly higher number of fluorescence measurements failed as well.

Observing the metrics in Table I, we see that the esophageal tube, as with the central tube of the previous data set, is resolved to have a much larger  $X$ - $Z$  cross-section than the subcutaneous tube. However, the length ( $Y$ -dimension) of the esophageal tube is much more accurately resolved using the statistical reconstruction. While the statistical method does result in average  $X$  and  $Z$  dimensions that are slightly larger than the threshold method, the difference is small and comparable to the size of a single voxel.

Looking at the subcutaneous tube, the location shift was less than a fraction of a voxel, while the sizes along the three dimensions are very similar between the two methods. In both cases,

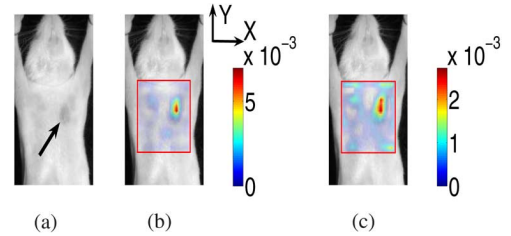


Fig. 5. Results for *in vivo* experiment. (a) Flat image. (b) Manually selected threshold. (c) Fixed point iteration. (a) and (b)  $z = 1.25$  cm, (c)  $z = 1.25$  cm.

while the width and depth are slightly overestimated, the length is underestimated, and the reconstruction does not appear to extend the full length of the tube seen in Fig. 4(a).

#### D. In-Vivo Mouse

Fig. 5(a) shows a flat image of the mouse being imaged. In it, the tumor is clearly visible as a dark spot in the left mammary pad (right side of image). This is indicative of the increased angiogenesis and blood concentration expected in the tumor. This reconstruction used ten iterations to reconstruct the tumor.

For this data set, the manually selected threshold was set to retain 25% of the data, based on results from previous studies [2]. The resulting inversion is seen in Fig. 5(b). A fluorescing mass is clearly present on the anterior surface of the mouse, just to the right of center. Results with the fixed point iteration are shown in Fig. 5(c). There, the tumor is resolved to be slightly larger than with the threshold, and its spatial extent is more consistent with the dark appearance of the tumor in the flat image. As with the phantom experiment, less than 0.1% of the data was dropped from the reconstruction.

Looking at the analytic metrics, we can see that the statistical method resolves the tumor to be slightly larger in width and depth than the threshold method. This is more consistent with the approximate size of the tumor of  $5 \times 6 \text{ mm}^2$  and represents a more ellipsoidal and less flattened tumor than the threshold method.



## VI. DISCUSSION

This paper has presented a new method by which fluorescence molecular tomography solutions can be obtained based on the normalized Born approximation. This method incorporates statistical properties of the measurements at the excitation and emission wavelengths in order to eliminate the use of operator defined thresholds. The method performed consistently and automatically across a variety of experimental scenarios, something that was not possible with previous thresholding techniques. Examining the results, we see that they compare favorably with those obtained using empirically defined thresholds, even when the thresholds were tuned using knowledge of the correct result.

Possible future extensions to this work include the use of fully nonlinear inversion methods, more complex regularization methods, the inclusion of correlation between the various data points, and spatially variant regularization to improve the dependence of size estimates on depth.

One of the advantages of using repeated solution of linear systems, such as we have done here, is that such systems are in general efficient and easily implemented. If instead we were to apply the conjugate gradient method directly to our nonlinear cost function in (25), we could eliminate the need for the fixed point iteration, at the cost of greater algorithmic and computational complexity.

Another possible extension of this work is the inclusion of correlation into the equations. Two independence assumptions were made to arrive at the weights detailed in this paper. The first of these was that each fluorescence measurement was independent of its associated excitation measurement. The second was that each fluorescence measurement was independent of every other fluorescence measurement, and likewise for the excitation measurements. Neither of these conditions is truly accurate. The intensity of excitation light reaching a detector is highly predictive of the intensity of fluorescence that the detector may see, although the exact value seen is still primarily dependent upon the physical distribution of the fluorochrome.

Overall, the method presented offers an inversion scheme that can enable more robust tomographic performance, by reducing image accuracy dependencies on calibration studies and user defined thresholds. Guided by *a priori* information based on the statistical nature of the measurements, this approach potentially offers a generalized and transferable inversion method across different platforms appropriate for standardization of optical tomography techniques. Further studies will involve application of the method to a larger number of *in vivo* data to confirm the validity of application across a large number of different potential targets, organs, and tissue sizes.

## APPENDIX I

APPROXIMATION OF TERMS INVOLVING  $\Phi$ 

In Section IV, our first step in simplifying the statistical density for the data was to make the approximation

$$\left[ \Phi \left\{ \frac{b(u_i)}{a(u_i)} \right\} - \Phi \left\{ -\frac{b(u_i)}{a(u_i)} \right\} \right] \simeq 1 \quad (30)$$

in order to eliminate the integrals from the density. We show here the conditions under which this approximation is valid. Clearly, as  $(b(u_i))/(a(u_i)) \rightarrow \infty$ , (30) is valid, and equivalently the integral

$$\int_0^{b(u_i)/a(u_i)} \phi(u) du \quad (31)$$

is equal to 0.5. Using the definition of  $\text{erf}()$

$$\int_0^q \phi(t) dt = \frac{1}{2} \text{erf} \left( \frac{q}{\sqrt{2}} \right) \quad (32)$$

we can write the inequality

$$0.5 - \frac{1}{2} \text{erf} \left( \frac{b(u_i)}{\sqrt{2}a(u_i)} \right) < \frac{\epsilon}{2} \quad (33)$$

where we define  $\epsilon$  as the tolerance allowed in the final approximation. From this we can obtain the values for the ratio  $(a(u_i))/(b(u_i))$  that satisfy the inequality. We rearrange the inequality to obtain

$$\frac{b(u_i)}{a(u_i)} > \sqrt{2} \text{erf}^{-1}(1 - \epsilon) = \delta \quad (34)$$

where  $\delta$  is set equal to the bound. Setting  $\epsilon = 2 * 10^{-5}$  in (33) yields  $\delta = 4.42$  as the required condition for (30) to be an appropriate approximation. Even increasing the precision to  $\epsilon = 2 * 10^{-16}$  only requires  $\delta = 8.21$ .

However,  $a(u_i)$  and  $b(u_i)$  are compound quantities that depend on the physical quantities  $m_{\text{fl},i}$  and  $m_{\text{exc},i}$  in a complicated fashion [see (11)]. Thus, we want to know equivalent bounds on the values of  $m_{\text{exc},i}$  and  $m_{\text{fl},i}$ , given a specified value for  $\delta$ . We start with

$$\frac{b(u_i)}{a(u_i)} = \frac{\frac{m_{\text{fl},i} u_i}{\sigma^2(m_{\text{fl},i})} + \frac{m_{\text{exc},i}}{\sigma^2(m_{\text{exc},i})}}{\left( \frac{u_i^2}{\sigma^2(m_{\text{fl},i})} + \frac{1}{\sigma^2(m_{\text{exc},i})} \right)^{1/2}}. \quad (35)$$

By placing this into the equation  $(b(u_i))/(a(u_i)) > \delta$ , squaring both sides, and rearranging, we can obtain

$$\begin{aligned} & \sigma^2(m_{\text{exc},i}) u_i^2 \times \left( \frac{m_{\text{fl},i}^2}{\sigma^2(m_{\text{fl},i})} - \delta^2 \right) \\ & + \sigma^2(m_{\text{fl},i}) \left( \frac{m_{\text{exc},i}^2}{\sigma^2(m_{\text{exc},i})} - \delta^2 \right) \\ & > -2m_{\text{fl},i} u_i m_{\text{exc},i}. \end{aligned} \quad (36)$$

When both  $(m_{\text{fl},i})/(\sigma(m_{\text{fl},i}))$  and  $(m_{\text{exc},i})/(\sigma(m_{\text{exc},i}))$  are greater than  $\delta$ , the inequality is satisfied because the left-hand side will be positive while the right-hand side will be negative (given that  $u_i$ ,  $m_{\text{exc},i}$  and  $m_{\text{fl},i}$  are all positive quantities). For  $\delta = 4.42$ , this occurs when both  $m_{\text{exc},i}$  and  $m_{\text{fl},i}$  are greater than 18 counts. In general, however, we want to be able to use as much of our data as possible, so we consider the case where one of the two values is less than  $\delta$ . This will result in one of the two left-hand terms being negative rather than positive. Using

the approximation  $u_i \simeq (m_{\text{fl},i})/(m_{\text{exc},i})$ , we can rewrite the inequality as

$$\begin{aligned} & \frac{\sigma^2(m_{\text{exc},i})}{m_{\text{exc},i}^2} \times \left( \frac{m_{\text{fl},i}^2}{\sigma^2(m_{\text{fl},i})} - \delta^2 \right) \\ & + \frac{\sigma^2(m_{\text{fl},i})}{m_{\text{fl},i}^2} \\ & \times \left( \frac{m_{\text{exc},i}^2}{\sigma^2(m_{\text{exc},i})} - \delta^2 \right) > -2. \end{aligned} \quad (37)$$

Given that the means and variances are all positive quantities, it is easily seen that the values of the quantities inside the parentheses must be greater than  $-\delta^2$ . If, as assumed, only one of the two quantities is greater than zero (i.e., the associated signal is higher than 18), then it can quickly be proven that the inequality must be satisfied. Because of this, we now choose to utilize all data points where at least one of the two signals is greater than 18.

## APPENDIX II

### ELIMINATION OF LOW MAGNITUDE TERM

Equation (10) in Section III gives the density of the data as the sum of two terms. However, it turns out that the second term is significantly smaller than the first, given reasonable conditions, thereby greatly simplifying the log-likelihood function. These conditions turn out to be achievable by minor additional requirements on the values of  $m_{\text{exc},i}$  and  $m_{\text{fl},i}$  compared to those obtained in Appendix I.

In the course of these derivations, we will make two primary assumptions.

- 1)  $0 \leq u_i \leq 1.5$ : The lower end of this bound is a result of both received signals inherently being positive quantities, while the upper bound arises from analysis of many data sets. In all data sets examined, the inequality  $m_{\text{fl},i} < 1.5m_{\text{exc},i}$  is satisfied by all data points. Thus, when taking the Born ratio,  $u_i = (m_{\text{fl},i})/(m_{\text{exc},i})$ , the inequality  $u_i < 1.5$  will be satisfied.
- 2)  $\mathbf{W}_i \mathbf{x} = (m_{\text{exc},i})/(m_{\text{fl},i}) \simeq u_i$ : The equality portion of this statement is the same assumption that was made in relating our statistical model to our forward model. We simply restate it here and further assume that the value of  $\mathbf{x}$  that corresponds to the actual distribution of fluorochrome will result in a value of  $\mathbf{W}\mathbf{x}$  that closely matches the received data  $\mathbf{u}$ .

What we explicitly need to show is that the term

$$\frac{1}{\pi\sigma(m_{\text{fl},i})\sigma(m_{\text{exc},i})} e^{-c/2} \quad (38)$$

is small enough to be dropped from the density. We assume that it can be removed when it is several orders of magnitude smaller than the other term in the density. We take the ratio of the two terms and study the conditions under which it will be much larger than one

$$\frac{\pi b(u_i) d(u_i)}{\sqrt{2\pi} a(u_i)} e^{c/2} \gg 1. \quad (39)$$

We specify the explicit bound at which (39) is satisfied to be  $\alpha$ , take the log of both sides, and rearrange to get

$$\log(d(u_i)) + c/2 \geq \beta + \log(a(u_i)) - \log(b(u_i)) \quad (40)$$

with

$$\beta = \log(\alpha) + 0.5 \log(\pi/2). \quad (41)$$

Now, given that

$$c/2 = 0.5 \left( \frac{m_{\text{fl},i}^2}{\sigma^2(m_{\text{fl},i})} + \frac{m_{\text{exc},i}^2}{\sigma^2(m_{\text{exc},i})} \right) \quad (42)$$

and:

$$\log(d(u_i)) = \left\{ -0.5 \frac{m_{\text{exc},i}^2 \left( \frac{m_{\text{fl},i}}{m_{\text{exc},i}} - u_i \right)^2}{(\sigma^2(m_{\text{exc},i})u_i + \sigma^2(m_{\text{fl},i}))} \right\} \quad (43)$$

we can now move the first term of  $c/2$  to the right-hand side and combine the second term with  $\log(d(u_i))$  in order to obtain

$$0.5 \left[ \frac{m_{\text{exc},i}^2}{\sigma(m_{\text{exc},i})^2} \left( 1 - \frac{(m_{\text{fl},i} - u_i)^2}{u_i + \frac{\sigma^2(m_{\text{fl},i})}{\sigma^2(m_{\text{exc},i})}} \right) \right] \geq \quad (44)$$

$$\beta + \log(a(u_i)) - \log(b(u_i)) - 0.5 \frac{m_{\text{fl},i}^2}{\sigma(m_{\text{fl},i})^2}. \quad (45)$$

Rearranging things a bit yields

$$\begin{aligned} & \left( 1 - \frac{\left( \frac{m_{\text{fl},i}}{m_{\text{exc},i}} - u_i \right)^2}{u_i + \frac{\sigma^2(m_{\text{fl},i})}{\sigma^2(m_{\text{exc},i})}} \right) \geq \frac{\sigma(m_{\text{exc},i})^2}{m_{\text{exc},i}^2} \\ & \times \left[ 2\beta + 2[\log(a(u_i)) - \log(b(u_i))] - \frac{m_{\text{fl},i}^2}{\sigma(m_{\text{fl},i})^2} \right]. \end{aligned} \quad (46)$$

Looking at the terms on the right-hand side, we want to find an upper bound for  $a(u_i)$  and a lower bound for  $b(u_i)$  in order to maximize the overall value of the right hand side. Looking at  $\log(b(u_i))$ , we have

$$\log(b(u_i)) = \log \left( \frac{m_{\text{fl},i} u_i}{\sigma^2(m_{\text{fl},i})} + \frac{m_{\text{exc},i}}{\sigma^2(m_{\text{exc},i})} \right) \quad (47)$$

$$\geq \log(m_{\text{exc},i}) - \log(\sigma^2(m_{\text{exc},i})) \quad (48)$$

where we have used the lower bound on  $u_i$  to establish the inequality.

For the term  $a(u_i)$ , we want to determine an upper bound. We use our assumed upper bound on  $u_i$  to obtain the inequality

$$\log(a(u_i)) \leq 0.5 \times \log \left( \frac{1.5^2}{\sigma^2(m_{\text{fl},i})} + \frac{1}{\sigma^2(m_{\text{exc},i})} \right). \quad (49)$$

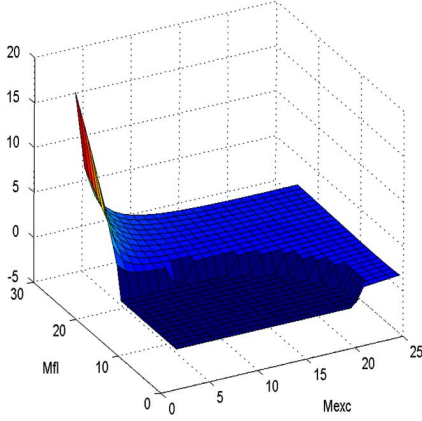


Fig. 6. Plot of the right-hand side of (50). Those points with magnitude less than one have been set equal to  $-1$  to give a better idea of the shape of the region in question.

By substituting these bounds in place of  $\log(b(u_i))$  and  $\log(a(u_i))$  in (46) and rearranging, we get

$$\begin{aligned} & \left( \frac{m_{fl,i}}{m_{exc,i}} - u_i \right)^2 \\ & \leq \frac{\sigma^2(m_{fl,i})}{\sigma^2(m_{exc,i})} \left[ 1 - \frac{\sigma^2(m_{exc,i})}{m_{exc,i}^2} \right] \\ & \quad \times \left[ 2\beta + \log \left( \frac{1.5^2}{\sigma^2(m_{fl,i})} + \frac{1}{\sigma^2(m_{exc,i})} \right) \right. \\ & \quad \left. - 2(\log(m_{exc,i}) - \log(\sigma^2(m_{exc,i}))) - \frac{m_{fl,i}^2}{\sigma^2(m_{fl,i})} \right]. \end{aligned} \quad (50)$$

This puts the left-hand side in terms of the mismatch between the estimate and the collected data and makes the right-hand side independent of  $u_i$ . Now, we use our assumption that  $(m_{exc,i})/(m_{fl,i}) \simeq u_i$  in order to assume that the left-hand side of the inequality is less than or equal to one. This allows for a significant mismatch between the received data and our model and provides a convenient threshold for determining a bound on which data pairs are usable. We establish this bound in a graphical fashion by simply computing the right-hand side for a range of  $m_{fl,i}$  and  $m_{exc,i}$  values and plotting to determine where the value is greater than one. The resulting surface plot is shown in Fig. 6. From this plot, we established that all points with  $m_{fl,i} > 15$  or  $m_{exc,i} > 20$  will satisfy the criterion. Combining this requirement with the previously established requirement that one of the two signals be greater than 18 counts, we determine that all data point pairs with either  $m_{fl,i} > 18$  or  $m_{exc,i} > 20$  will be retained for use in the reconstruction.

### APPENDIX III

#### DERIVING THE GRADIENT OF THE COST FUNCTION

##### A. Primary Derivations

In the course of obtaining a ML solution in Section IV, we derived the cost function

$$f(\mathbf{u}, \mathbf{x}) = \sum_i g_i(x) - 0.5 \|\mathbf{Q}(\mathbf{u}, \mathbf{x})(\mathbf{W}\mathbf{x} - \mathbf{u})\|_2^2 \quad (51)$$

which we want to maximize in order to obtain our solution. In order to simplify the notation involved in finding the gradient of (51), we write

$$s(\mathbf{u}, \mathbf{x}) = 0.5 \|\mathbf{Q}(\mathbf{u}, \mathbf{x})(\mathbf{W}\mathbf{x} - \mathbf{u})\|_2^2. \quad (52)$$

We then write an equation for the gradient as

$$\frac{df(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} = \left[ \sum_i \frac{dg_i(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} \right] - \frac{ds(\mathbf{u}, \mathbf{x})}{d\mathbf{x}}. \quad (53)$$

Given the complexity of the  $g_i(\mathbf{u}, \mathbf{x})$  terms, we derive the gradient of those components separately in Appendix III-B and make use of those results at the end of this section. Turning our attention to the second term of the gradient, we can write that gradient as

$$\begin{aligned} \frac{ds(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} &= \left[ \frac{d(\mathbf{W}\mathbf{x} - \mathbf{u})}{d\mathbf{x}} \right]^T \frac{ds(\mathbf{u}, \mathbf{x})}{d(\mathbf{W}\mathbf{x} - \mathbf{u})} \\ &\quad + \left[ \frac{d\mathbf{Q}(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} \right]^T \frac{ds(\mathbf{x})}{d\mathbf{Q}(\mathbf{u}, \mathbf{x})}. \end{aligned} \quad (54)$$

The first term of the sum can then be computed as

$$\begin{aligned} \left[ \frac{d(\mathbf{W}\mathbf{x} - \mathbf{u})}{d\mathbf{x}} \right]^T \frac{ds(\mathbf{x})}{d(\mathbf{W}\mathbf{x} - \mathbf{u})} \\ = 2\mathbf{W}^T C(\mathbf{x})^T C(\mathbf{x})(\mathbf{W}\mathbf{x} - \mathbf{u}). \end{aligned} \quad (55)$$

To compute the second term, we use the fact that  $\mathbf{q}(\mathbf{u}, \mathbf{x}) = \text{diag}(q(\mathbf{u}, \mathbf{x}))$  in order to rewrite  $s(\mathbf{x})$  as

$$s(\mathbf{u}, \mathbf{x}) = \mathbf{q}(\mathbf{u}, \mathbf{x})^T \text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u})^T \text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u}) \mathbf{q}(\mathbf{x}). \quad (56)$$

This rearrangement allows us to write the gradient as

$$\frac{ds(\mathbf{u}, \mathbf{x})}{d\mathbf{q}(\mathbf{x})} = 2\text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u})^T \text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u}) \mathbf{q}(\mathbf{x}). \quad (57)$$

For notational simplicity, we now define

$$A(\mathbf{x}) = \text{diag}(\mathbf{W}\mathbf{x} - \mathbf{u}) \quad (58)$$

and rewrite the gradient as

$$\frac{ds(\mathbf{u}, \mathbf{x})}{d\mathbf{Q}(\mathbf{x})} = 2A(\mathbf{x})^T A(\mathbf{x}) \mathbf{q}(\mathbf{x}). \quad (59)$$

Now we derive the value of  $(d\mathbf{q}(x))/(d\mathbf{x})$ . Given that

$$\begin{aligned} \mathbf{q}_i(\mathbf{x}) &= m_{exc,i} (\beta m_{exc,i} (u_i + W_i \mathbf{x}) \\ &\quad + (1 + u_i) \sigma_r^2)^{-1/2} \end{aligned} \quad (60)$$

we can write the derivative as

$$\begin{aligned} \frac{d\mathbf{q}_i(\mathbf{x})}{d\mathbf{x}} &= \frac{-m_{exc,i}}{2} (\beta m_{exc,i} (u_i + W_i \mathbf{x}) \\ &\quad + (1 + u_i) \sigma_r^2)^{-3/2} \beta m_{exc,i} W_i. \end{aligned} \quad (61)$$

$$\frac{1}{b(u_i, \mathbf{x})} \frac{db(u_i, \mathbf{x})}{d\mathbf{x}} = \frac{(m_{\text{exc},i} W_{ij}) \sigma(m_{\text{exc},i})^2 [u_i \sigma(m_{\text{fl},i}(\mathbf{x})) - \beta m_{\text{fl},i}(\mathbf{x}) u_i]}{\sigma(m_{\text{fl},i}(\mathbf{x}))^2 [\sigma(m_{\text{exc},i})^2 m_{\text{fl},i}(\mathbf{x}) u_i + \sigma(m_{\text{fl},i}(\mathbf{x}))^2 m_{\text{exc},i}]} \quad (67)$$

This gives us a vector, representing the gradient of a particular  $\mathbf{q}_i(\mathbf{x})$  with respect to  $\mathbf{x}$ . Obtaining one for each  $\mathbf{q}_i(\mathbf{x})$ , we can then arrange all of these vectors into a matrix as

$$\frac{d\mathbf{q}(\mathbf{x})}{d\mathbf{x}} = \begin{bmatrix} \frac{d\mathbf{q}_1(\mathbf{x})^T}{d\mathbf{x}} & \frac{d\mathbf{q}_2(\mathbf{x})^T}{d\mathbf{x}} & \dots & \frac{d\mathbf{q}_N(\mathbf{x})^T}{d\mathbf{x}} \end{bmatrix}. \quad (62)$$

This, combined with the gradient of  $g_i(\mathbf{x})$  derived in Appendix III-B, allows us to write the full gradient of the cost function as

$$\begin{aligned} \frac{df(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} = & \left[ \sum_i \frac{dg_i(\mathbf{u}, \mathbf{x})}{d\mathbf{x}} \right] \\ & - \mathbf{W}^T \mathbf{Q}(\mathbf{x})^T \mathbf{Q}(\mathbf{x}) (\mathbf{W}\mathbf{x} - \mathbf{u}) \\ & - \left( \frac{d\mathbf{q}(\mathbf{x})}{d\mathbf{x}} \right)^T A(\mathbf{x})^T A(\mathbf{x}) \mathbf{q}(\mathbf{x}) \end{aligned} \quad (63)$$

which we use to derive our fixed point iteration in the text.

#### B. Derivation of One of the Gradient Components

In this section, we explicitly determine the gradient of the  $g_i(u_i, \mathbf{x})$  terms. We repeat (22) to begin

$$g_i(u_i, \mathbf{x}) = \log(b(u_i, \mathbf{x})) - \log(\sigma(m_{\text{fl},i}(\mathbf{x}))) - 3 \log(a(u_i, \mathbf{x})). \quad (64)$$

We explicitly write the gradient of this equation

$$\begin{aligned} \frac{dg_i(x)}{d\mathbf{x}} = & \frac{1}{b(u_i, \mathbf{x})} \frac{db(u_i, \mathbf{x})}{d\mathbf{x}} \\ & - \frac{1}{\sigma(m_{\text{fl},i}(\mathbf{x}))} \frac{d\sigma(m_{\text{fl},i}(\mathbf{x}))}{d\mathbf{x}} \\ & - 3 \frac{1}{a(u_i, \mathbf{x})} \frac{da(u_i, \mathbf{x})}{d\mathbf{x}}. \end{aligned} \quad (65)$$

This leaves us with a gradient that is the sum of three terms, each of which consists of the gradient and the reciprocals of known functions. We treat each term separately. First, we write the gradient of  $b(u_i, \mathbf{x})$  with respect to  $\mathbf{x}$  as

$$\begin{aligned} \frac{db(u_i, \mathbf{x})}{d\mathbf{x}} = & \frac{db(u_i, \mathbf{x})}{dm_{\text{fl},i}(\mathbf{x})} \frac{dm_{\text{fl},i}(\mathbf{x})}{d\mathbf{x}} \\ = & \left[ \frac{u_i \sigma(m_{\text{fl},i}(\mathbf{x}))^2 - \beta m_{\text{fl},i}(\mathbf{x}) u_i}{\sigma(m_{\text{fl},i}(\mathbf{x}))^4} \right] \\ & \times [m_{\text{exc},i} W_{ij}]. \end{aligned} \quad (66)$$

Combining this with (11) yields (67) as shown at the top of the page. Next, given the definition for  $\sigma(m)$  in (7), we can write the second term in the gradient of  $g_i(u_i, \mathbf{x})$  as

$$\frac{1}{\sigma(m_{\text{fl},i}(\mathbf{x}))} \frac{d\sigma(m_{\text{fl},i})}{d\mathbf{x}} = \frac{1}{2} \frac{\beta m_{\text{exc},i} W_{ij}}{\beta m_{\text{exc},i} W_{ij} \mathbf{x} + \sigma_r^2}. \quad (68)$$

Finally, we look at the third term, which depends on  $a(u_i, \mathbf{x})$  and contains the gradient

$$\begin{aligned} \frac{da(u_i, \mathbf{x})}{d\mathbf{x}} = & \frac{da(u_i, \mathbf{x})}{dm_{\text{fl},i}(\mathbf{x})} \frac{dm_{\text{fl},i}(\mathbf{x})}{d\mathbf{x}} \\ = & \left[ \frac{-1}{2} \left( \frac{u_i^2}{\sigma(m_{\text{fl},i}(\mathbf{x}))^2} + \frac{1}{\sigma(m_{\text{exc},i})^2} \right)^{-1/2} \right. \\ & \left. \times \frac{\beta u_i^2}{\sigma(m_{\text{fl},i}(\mathbf{x}))^4} \right] [m_{\text{exc},i} W_{ij}]. \end{aligned} \quad (69)$$

Thus, again using (11), we get

$$\begin{aligned} \frac{1}{a(u_i, \mathbf{x})} \frac{da(u_i, \mathbf{x})}{d\mathbf{x}} = & \left[ \frac{1}{2} \frac{1}{a(u_i, \mathbf{x})^2} \frac{\beta u_i^2}{\sigma(m_{\text{fl},i}(\mathbf{x}))^4} \right] [m_{\text{exc},i} W_{ij}]. \end{aligned} \quad (70)$$

The values of (67), (70), and (68) allow us to write the full gradient of  $g_i(u_i, \mathbf{x})$ . With that in hand, we are able to explicitly evaluate the full gradient of our cost function.

#### REFERENCES

- [1] V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder, "Looking and listening to light: The evolution of whole-body photonic imaging," *Nature Biotechnol.*, vol. 23, no. 3, pp. 313–320, Mar. 2005.
- [2] E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A sub-millimeter resolution fluorescence molecular imaging system for small animal imaging," *Med. Phys.*, vol. 30, no. 5, pp. 901–911, May 2003.
- [3] V. Ntziachristos, E. A. Schellenberger, J. Ripoll, D. Yessayan, E. Graves, J. A. Bogdanov, L. Josephson, and R. Weissleder, "Visualization of antitumor treatment by means of fluorescence molecular tomography with an annexin v-cy5.5 conjugate," *Proc. Nat. Acad. Sci.*, vol. 101, no. 33, pp. 12 294–12 299, Aug. 2004.
- [4] A. Godavarty, M. J. Eppstein, C. Zhang, and E. M. Sevick-Muraca, "Detection of single and multiple targets in tissue phantoms with fluorescence-enhanced optical images: Feasibility study," *Radiology*, vol. 235, no. 1, pp. 148–154, Apr. 2005.
- [5] J. Lee and E. M. Sevick-Muraca, "Three-dimensional fluorescence enhanced optical tomography using referenced frequency-domain photon migration measurements at emission and excitation wavelengths," *J. Opt. Soc. Amer. A*, vol. 19, no. 4, pp. 759–771, Apr. 2002.
- [6] A. D. Klose and A. H. Hielscher, "Fluorescence tomography with simulated data based on the equation of radiative transfer," *Opt. Lett.*, vol. 28, no. 12, pp. 1091–1021, Jun. 2003.
- [7] M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. Sevick-Muraca, "Three-dimensional, bayesian image reconstruction from sparse and noisy data sets: Near-infrared fluorescence tomography," *Proc. Nat. Acad. Sci.*, vol. 99, no. 15, pp. 9619–9624, Jul. 2002.
- [8] H. Jiang, "Frequency-domain fluorescent diffusion tomograph: A finite-element-based algorithm and simulations," *Appl. Opt.*, vol. 37, no. 22, pp. 5337–5343, Aug. 1998.
- [9] M. A. O'Leary, D. A. Boas, X. D. Li, B. Chance, and A. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.*, vol. 21, no. 2, pp. 158–160, Jan. 1996.
- [10] M. Eppstein, F. Fedele, J. Laible, C. Zhang, A. Godavarty, and E. Sevick-Muraca, "A comparison of exact and approximate adjoint sensitivities in fluorescence tomography," *IEEE Trans. Med. Imag.*, vol. 22, no. 10, pp. 1215–1223, Oct. 2003.

- [11] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. Millane, "Fluorescence optical diffusion tomography," *Appl. Opt.*, vol. 42, no. 16, pp. 3081–3093, Jun. 2003.
- [12] J. Chang, H. L. Graber, and R. L. Barbour, "Imaging of fluorescence in highly scattering media," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 810–822, Sep. 1997.
- [13] A. Becher, 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," *Nature Biotechnol.*, vol. 19, pp. 327–331, Apr. 2001.
- [14] S. Achilefu, S. Bloch, M. Markiewicz, T. Zhong, Y. Ye, R. Dorshow, B. Change, and K. Liang, "Synergistic effects of light-emitting probes and peptides for targeting and monitoring integrin expression," *Prof. Nat. Acad. Sci. USA*, vol. 102, pp. 7976–81, 2005.
- [15] C. Tung, U. Mahmood, S. Bredow, and R. Weissleder, "In-vivo imaging of proteolytic enzyme activity using a novel molecular reporter," *Cancer Res.*, vol. 60, pp. 4953–8, 2000.
- [16] F. S. Wouters, P. J. Verveer, and P. I. Bastiaens, "Imaging biochemistry inside cells," *Trends Cell Biol.*, vol. 11, no. 5, pp. 203–211, May 2001.
- [17] R. Rot and E. M. Sevick-Muraca, "Three-dimensional unconstrained and constrained image-reconstruction techniques applied to fluorescence, frequency-domain photon migration," *Appl. Opt.*, vol. 40, no. 13, pp. 2206–2214, May 2001.
- [18] V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of the normalized Born approximation," *Opt. Lett.*, vol. 26, no. 12, pp. 893–895, June 2001.
- [19] A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized Born ratio," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1377–1386, Oct. 2005.
- [20] D. Boas, T. Gaudette, and S. Arridge, "Simultaneous imaging and optode calibration with diffuse optical tomography," *Opt. Expr.*, no. 8, pp. 263–270, 2001.
- [21] A. Kienle and M. S. Patterson, "Improved solutions of the steady-state and the time-resolved diffusion equations for reflectance from a semi-infinite turbid medium," *J. Opt. Soc. Amer. A*, vol. 14, pp. 246–254, 1997.
- [22] D. V. Hinkley, "On the ratio of two correlated normal random variables," *Biometrika*, vol. 56, no. 3, pp. 635–639, 1969.
- [23] B. W. Pogue, T. O. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, "Spatially variant regularization improves diffuse optical tomography," *Appl. Opt.*, vol. 38, no. 13, 1999.