Published in final edited form as:

J Opt Soc Am A Opt Image Sci Vis. 2012 March 1; 29(3): 321–330.

Information loss and reconstruction in diffuse fluorescence tomography

Petra Bonfert-Taylor^{1,2,*}, Frederic Leblond², Robert W. Holt², Kenneth Tichauer², Brian W. Pogue², and Edward C. Taylor^{1,2}

¹Department of Mathematics, Wesleyan University, 265 Church Street, Middletown, Connecticut 06459, USA

²Thayer School of Engineering, Dartmouth College, 8000 Cummings Hall, Hanover, New Hampshire 03755, USA

Abstract

This paper is a theoretical exploration of spatial resolution in diffuse fluorescence tomography. It is demonstrated that, given a fixed imaging geometry, one cannot—relative to standard techniques such as Tikhonov regularization and truncated singular value decomposition—improve the spatial resolution of the optical reconstructions via increasing the node density of the mesh considered for modeling light transport. Using techniques from linear algebra, it is shown that, as one increases the number of nodes beyond the number of measurements, information is lost by the forward model. It is demonstrated that this information cannot be recovered using various common reconstruction techniques. Evidence is provided showing that this phenomenon is related to the smoothing properties of the elliptic forward model that is used in the diffusion approximation to light transport in tissue. This argues for reconstruction techniques that are sensitive to boundaries, such as L^1 -reconstruction and the use of priors, as well as the natural approach of building a measurement geometry that reflects the desired image resolution.

1. INTRODUCTION

Optical tomography with diffused light is an imaging modality that uses light in the red and near-infrared (NIR) wavelength band to transilluminate biological tissue for the purposes of medical diagnostics (see, e.g., [1–3]). Signal detection through up to several centimeters of tissue is feasible in that range because scattering dominates over absorption (e.g., from hemoglobin), with the latter being orders of magnitude smaller when compared to other wavelengths in the visible part of the electromagnetic spectrum. Optical tomography is implemented via the mathematical reconstruction of absorption and scattering tissue properties from surface measurements following excitation with light sources such as lasers [4]. Since red and NIR radiation are strongly scattered by biological tissue, one can assume photons are isotropically scattered within a few millimeters from the boundary, and as such light transport modeling can be achieved with reasonable accuracy in the diffusion approximation regime of the radiative transport equation [5,6].

Of interest here are two types of imaging modalities, namely, diffuse optical tomography (DOT) and diffuse fluorescence tomography (FT), with the latter often being referred to as fluorescence molecular tomography (FMT). The optical contrast targeted for reconstruction

in DOT is typically associated with tissue chromophores such as oxygenated and deoxygenated hemoglobin and sometimes with water and lipids [1–3]. However, FT deals with the optical contrast that is induced by fluorescent reporters following the systemic injection of a compound targeting specific molecular events associated with pathological processes [7] (see also [8] for a review). It is a well-known fact that the spatial resolution and the contrast recovery associated with optical tomography methods present engineering and mathematical challenges. Fundamental spatial resolution limits in diffused imaging are set by the mean photon free path or the inverse of the transport coefficient, which is typically smaller than 1 mm. However, although several studies have considered the actual spatial resolution and contrast recovery limits observed in re-constructed images, there is no universal criterion that has been derived to quantify them partly because they depend on the particular details of the imaging geometry and the biological application of interest.

A critical aspect impacting the imaging capabilities of optical tomography is the fact that the associated inverse problem is ill posed mathematically. More precisely, diffuse tomography models are highly ill conditioned, implying that even minute discrepancies between model predictions and experimental data might be a dominating factor in non-regularized reconstructions. Typical sources of data–model mismatch include discretization errors, stochastic noise, light transport modeling inaccuracies, and systematic errors. However, while the use of regularization allows one to maximize the amount of physical information recovered from the data, ill-posedness implies that a significant portion of the information that is available in principle can be lost in the process. As demonstrated in, e.g., [9–11], the amount of information that is lost through regularization can be minimized by improving the light detection geometry by maximizing the level of signal to noise and by using optical data types leading to less ill-posed problems. In these studies, the tool of choice for evaluating the degree of ill-posedness consists of a singular value analysis of the matrices derived from light transport solutions. Typically, the most favorable imaging problems will be those for which the singular value curves (as a function of mode order) have a slope that is as small as possible, implying that high-frequency components (e.g., associated with shot noise) will have less of a tendency to corrupt the images. A conclusion of those studies is that system design optimization should be done ensuring that image reconstruction can be achieved using as many singular modes as possible, especially when the number of available measurements is limited.

The limitations, however, associated with optical tomography methods are also impacted by the fact that the problems of interest are in most cases underdetermined [12]. In fact, owing to the sophistication of numerical methods and increasing computational power, the forward model in diffuse optical tomography is computationally deployed using a (locally) linear model with a Jacobian matrix that has a number of columns that typically exceeds by far the number of rows, and this causes the associated inverse problem to be underdetermined. The main cause for this is the fact that the number of measurements is often far smaller than the number of nodes at which one wishes to reconstruct the optical properties. One thus balances increased resolution through finer mesh sizes against the numerical costs (e.g., round-off errors, computational time) of being underdetermined [13]. In this work the focus is on a particular numerical cost to being undetermined. While the emphasis here is on aspects relating to FT, the conclusions that are reached are also believed to be generally applicable to other applications such as DOT and bioluminescence tomography because of the similarity in their mathematical formulation [14–16].

The results of this paper are as follows. It is shown, using standard structures in linear algebra, that the forward model in FT can project to its own kernel a significant portion of the vector containing the simulated optical properties in a manner that is not recoverable by standard regularization techniques such as Tikhonov regularization, truncated singular value

decomposition, and methods based on Landweber–Fridman iterations. Here this portion of the optical properties vector (i.e., the component that is projected to the kernel) is referred to as “information loss,” and this loss is quantified using numerical experiments based on diffusion light transport modeling. It is important to state that this information loss is intrinsic to optical tomography with diffused light and that it is distinct from the propagation of data–model mismatch errors due to the ill-conditioned nature of the problem. In particular, it is shown here that the magnitude of the information loss can be as high as 50% of the magnitude of the optical properties vector for typical optical tomography geometries where a limited number of measurements are available [17]. The analysis presented here demonstrates that the loss of information increases with increasing node density using a finite element discretization scheme for the forward model based on the open source diffuse optical tomography simulation software package NIRFAST [18]. The main conclusion of the paper is that one must balance the desire for higher resolution against the accumulation of this measurable destruction of information. Though this loss of information is facilitated by the fact that the inverse problem in FT is typically underdetermined, it is demonstrated mathematically that being underdetermined alone is not the controlling factor. To this effect, we present evidence showing that the smoothing properties of the elliptic forward model that is used in the diffusion approximation to light transport in tissue may be partially responsible for the information loss.

The paper is organized in the following way. In Section 2 background material is presented, including a review of the diffusion equation used to model light transport (Subsection 2.A) and a brief overview of how the FT inverse problem is formulated as a linear matrix model (2.B). Then Subsection 2.C reviews the main linear algebra concepts required to understand the results presented. In Section 3 the methods and numerical experiments are described in detail, and Section 4 contains the results. The paper is concluded with discussion in Section 5 where an analysis of the results are presented as well as their implication for the field of tomography in highly scattering media. Appendix A contains some auxiliary facts from linear algebra.

2. BACKGROUND MATERIAL

A. Light Transport and Fluorescence Model

The forward problem for diffuse optical tomography may be described as follows: given a distribution of light sources on the boundary of a medium Ω and a distribution of optical parameters in the medium (absorption and scattering properties), find the resulting measurement set on the boundary. If the magnitude of the isotropic fluence is much greater than the anisotropic fluence, then it is well established that a diffusion equation model of light transport is an accurate reduction of the radiative transfer problem [5,6,19,20]. This diffusion approximation in the frequency-domain form is given by

$$-\nabla \cdot \kappa(r) \nabla \Phi(r, \omega) + \left(\mu_a(r) + \frac{i\omega}{c_m(r)} \right) \Phi(r, \omega) = S(r, \omega), \quad (1)$$

where $\Phi(r, \omega)$ is the photon fluence rate, a position $r \in \Omega$ and modulation frequency ω , S is

an isotropic source term, $\kappa = \frac{1}{3(\mu_a + \mu'_s)}$ is the diffusion coefficient with absorption coefficient μ_a and reduced scattering coefficient μ'_s . Finally, $c_m(r)$ is the speed of light in Ω at the point r defined by $\frac{c}{n(r)}$, where $n(r)$ is the index of refraction at r and c is the speed of light in a vacuum. The air–tissue boundary is represented as a Robin boundary condition; that is, the

flux leaving the boundary of the medium is the fluence rate at the boundary weighted by a factor that accounts for some internal reflection back into the tissue:

$$\Phi(r, \omega) + 2B\kappa(r) \frac{\partial \Phi(r, \omega)}{\partial \nu} \Big|_{\partial \Omega} = 0, \quad (2)$$

where B is the term that deals with the refractive mismatch of the tissue with the surrounding media and ν is the unit surface normal of a patch on the boundary of the tissue.

The work in this paper is specifically focused on fluorescence molecular optical tomography in which fluorescent reporters are used to localize optical contrast via the accumulation or retention of the reporters; see, e.g., [8,21]. In this setting, a coupled pair of frequency-domain diffusion equations are used to predict the excitation and emission fluence rates $\Phi_x(r, \omega)$ and $\Phi_m(r, \omega)$, respectively. These equations are given as

$$\nabla \cdot \kappa_x(r) \nabla \Phi_x(r, \omega) - \left(\mu_{a,x}(r) + \frac{i\omega}{c_n(r)} \right) \times \Phi_x(r, \omega) + S_x(r, \omega) = 0, \quad (3)$$

$$\nabla \cdot \kappa_m(r) \nabla \Phi_m(r, \omega) - \left(\mu_{a,m}(r) + \frac{i\omega}{c_n(r)} \right) \times \Phi_m(r, \omega) + S_m(r, \omega) = 0, \quad (4)$$

where the diffusion coefficients are labeled x , excitation, and m , emission, and the fluorescence source term $S_m(r, \omega)$ takes the form

$$S_m(r, \omega) = \frac{Q_f \mu_{af}}{1 - i\omega\tau} \Phi_x(r, \omega). \quad (5)$$

Here Q_f is the quantum efficiency of the fluorophore, $\mu_{a,f}$ the absorption coefficient, and τ is the lifetime. Both the emission and excitation fluence equations satisfy a Robin boundary condition on Ω , as given in Eq. (2).

B. Fluorescence Tomography

In the limit where the absorption due to the fluorophores ($\mu_{a,f}$) is much smaller than that of the other chromophores (e.g., hemoglobin, water, lipids), the solution to the coupled differential equations [Eqs. (3) and (4)] can be formally written as

$$\Phi_m(r, \omega) = \int_{\Omega} d^3 r' G_m(r, r', \omega) \frac{Q_f \mu_{af}(r')}{1 - i\omega\tau} \Phi_x(r', \omega), \quad (6)$$

where G is the Green's function associated with Eq. (4). Furthermore, Eq. (6) can be formulated on a discretized mesh of the volume Ω as

$$\begin{aligned} \Phi_m(r_k, \omega) &= \Phi_{m,k}(\omega) \\ &= \sum_{i=1 \dots N_v} \Delta V_i G_m(r_k, r_i, \omega) \frac{Q_f \mu_{af}(r_i)}{1 - i\omega\tau} \Phi_x(r_i, \omega), \end{aligned}$$

where the integer k is used to label the spatial location of a measurement on the boundary Ω and the index i labels each of the N_v elements with a volume ΔV_i in the mesh (e.g., nodes in a finite element volume).

Each fluorescence measurement in a full tomography data set is modeled using this discretized equation above. In this work, those solutions are obtained using the NIRFAST software [18] for which the interrogated domain is discretized using a finite element mesh. The fluorescence tomography problem is set up by considering several measurements around the periphery of the imaging domain. Those measurements and the discretized version of the corresponding differential equation are concatenated to form the matrix problem,

$$\Phi_{m,k}(\omega) = \sum_{i=1}^{N_v} a_{ki} \mu_{af}^i, \quad (7)$$

where the measurement index k takes values up to the total number of measurements, N_m . The elements of the resulting Jacobian matrix, $A = [a_{ki}]$, are

$$a_{ki} = G_m(r_k, r_i, \omega) \frac{Q_f \mu_{af}(r_i)}{1 - i\omega\tau} \Phi_x(r_k, \omega). \quad (8)$$

In the remainder of this work, it is assumed that light transport physics in fluorescence tomography can be captured by this simple linear model, and we refer to the matrix A as either the reconstruction Jacobian or the forward model matrix, although they may differ in some cases where, for example, different meshes are used for forward modeling and image reconstruction.

C. Linear Algebra Primer: Singular Values and the Kernel Subspace

In this section the reader is reminded of certain standard linear algebra structures that we need to reference in order to establish the results in this paper. Of the multitude of possible references on the topic of linear algebra, we will use [22,23]. Only the basic definitions and results that are needed by us are established below; those who are looking for further details and arguments can find them either in the references or in Appendix A.

1. Vector Spaces and Linear Transformations—In all that follows, the symbols N_v [the number of nodes in a finite element method (FEM) mesh] and N_m (the number of measurements, i.e., source–detector pairs) are positive integers. We denote a standard real finite dimensional vector space by $\mathbb{R}^k$, where, for our purposes, k will take the values N_v or N_m . These vector spaces are endowed with the standard inner products $\langle \cdot, \cdot \rangle$ given by $\langle x, y \rangle = x^T y$. The inner product is left unlabeled, and the context should inform the reader as to whether $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^{N_v}$ or $\mathbb{R}^{N_m}$. In all instances in this paper, we will be working with a linear transformation L mapping vectors in $\mathbb{R}^{N_v}$ to vectors in $\mathbb{R}^{N_m}$; this is denoted by $L \in \mathcal{L}(\mathbb{R}^{N_v}, \mathbb{R}^{N_m})$. The action of L can be represented by the action under matrix multiplication of a matrix A associated to L (see [22] for example); by abuse of notation we will use the symbol A for both a linear transformation and its matrix representation. Using standard notation, it is seen that $A = [a_{ki}] \in M_{N_m \times N_v}$, i.e., A has N_m rows and N_v columns. In the context of this paper, this matrix is the Jacobian matrix of the forward problem. Owing to the experimental design of diffuse optical tomography devices as well as the need for increased resolution, we will only consider the case where the dimension N_v of the domain vector space $\mathbb{R}^{N_v}$ is strictly greater than the dimension N_m of the range vector space $\mathbb{R}^{N_m}$; thus, we are explicitly dealing with an underdetermined system of linear equations.

The action of A splits $\mathbb{R}^{N_v}$ into the direct sum of two canonical subspaces. The kernel of A is a vector subspace of $\mathbb{R}^{N_v}$ consisting of all vectors that are mapped to the zero vector under the action of A ; that is

$$\text{Ker}(A) = \{v \in \mathbb{R}^{N_v} : Av = 0\}.$$

The kernel of A is always nonempty as $0 \in \text{Ker}(A)$ for any choice of A . We let $\text{Ker}(A)^\perp$ be the orthogonal complement in $\mathbb{R}^{N_v}$ of $\text{Ker}(A)$, i.e.,

$$\text{Ker}(A)^\perp = \{w \in \mathbb{R}^{N_v} : \langle v, w \rangle = 0 \text{ for all } v \in \text{Ker}(A)\}.$$

Note that $\text{Ker}(A)^\perp$ is itself a vector subspace of $\mathbb{R}^{N_v}$. Then $\mathbb{R}^{N_v}$ is the direct sum of $\text{Ker}(A)$ and $\text{Ker}(A)^\perp$, i.e.,

$$\mathbb{R}^{N_v} = \text{Ker}(A) \oplus \text{Ker}(A)^\perp. \quad (9)$$

This means that every vector $x \in \mathbb{R}^{N_v}$ can be uniquely written as the sum of two vectors x_0 and $x^\perp$, where $x_0 \in \text{Ker}(A)$ and $x^\perp \in \text{Ker}(A)^\perp$. It is clear that, if v is a nonzero vector in $\text{Ker}(A)^\perp$, then $Av \neq 0$.

We recall that the *range* of $A \in \mathcal{L}(\mathbb{R}^{N_v}, \mathbb{R}^{N_m})$ is the vector subspace of $\mathbb{R}^{N_m}$ defined by

$$\text{Ran}(A) = \{w \in \mathbb{R}^{N_m} : \text{there exists } v \in \mathbb{R}^{N_v} \text{ with } Av = w\}.$$

The *rank* of the linear transformation A , denoted by $\text{Rank}(A)$, is the dimension of $\text{Ran}(A)$; we remind the reader that the dimension of a vector space is the maximal number of linearly independent vectors in that space. The kernel and the range of A are tied together by the following fundamental theorem.

Theorem 2.1 (Rank Theorem [22]): Let $A: \mathbb{R}^{N_v} \rightarrow \mathbb{R}^{N_m}$ be a linear transformation. Then

$$N_v = \dim \text{Ker}(A) + \text{Rank}(A).$$

Given the standing assumption that $N_v > N_m$, and using the Rank Theorem, it is observed that the dimension of the kernel of A is always strictly positive since the range of A is contained in $\mathbb{R}^{N_m}$, and therefore $\text{rank}(A) \leq N_m$. Finally, a useful piece of information concerning the action of A on the vector subspace $\text{Ker}(A)^\perp$ (see, e.g., [22]) is observed: the linear mapping $A: \text{Ker}(A)^\perp \rightarrow \text{Ran}(A)$ is a bijection (i.e., a one-to-one and onto correspondence).

2. Singular Value Decomposition—There exists a change of coordinates, for both $\mathbb{R}^{N_v}$ and $\mathbb{R}^{N_m}$, so that the action of A in these coordinates is a diagonal action, that is, multiplication by a (nonnegative) constant in each coordinate direction. This change of coordinates, including the nonnegative constants, is called the singular value decomposition (SVD) of A (see, e.g., [23], Chapter 2). Under the assumption that $N_m < N_v$, the singular value decomposition of $A = [a_{ki}]$, $1 \leq k \leq N_m$ and $1 \leq i \leq N_v$, is given as $\{v_i(u_k), (\sigma_k)\}$ so that (u_k) is an orthonormal basis for $\mathbb{R}^{N_m}$, (v_i) is an orthonormal basis for $\mathbb{R}^{N_v}$, and so that $A(v_k) = \sigma_k u_k$ for $1 \leq k \leq N_m$ with $\sigma_k \geq 0$ and $\sigma_{k-1} \geq \sigma_k$. Using matrix notation, we can represent $A \in \mathcal{L}(\mathbb{R}^{N_v}, \mathbb{R}^{N_m})$ as $A = UV^T$ where U and V are $N_m \times N_m$ matrices constructed

with columns from $\{u_k\}$ and $\{v_j\}$, respectively, and Σ is a $N_m \times N_v$ rectangular diagonal matrix with entries σ_j filling the diagonal.

The SVD provides a computationally convenient basis for the kernel of A (see Proposition 6.1 in Appendix A). It is with this basis that the magnitude can be computed of the projection by the FT forward model to its kernel.

3. METHODS AND NUMERICAL EXPERIMENTS

A standard numerical approach to the study of the solutions to Eqs.s (3) and (4) is the FEM [18,24,25]; in fact, the FEM is used to approximate the forward problem concerning the fluence rate Φ in any given domain. Our FEM simulation is implemented via the software package NIRFAST [18,26], which is free-ware developed at Dartmouth for the purpose of simulating light propagation in biological tissue.

The examples considered here are based on simulations for two-dimensional triangular meshes of circular tissue-simulating phantoms of diameter 5 cm. The size of the phantom has been chosen to mimic small animals such as rats [27]. For the simulations, optical properties were chosen that were similar to what is expected in a small animal, namely $\mu'_s = 1.0 \text{ mm}^{-1}$ and $\mu_a = 0.03 \text{ mm}^{-1}$. For simplicity, the optical properties were assumed to be the same at both excitation and emission wavelengths. The application of interest is the reconstruction of fluorescent sources representing tumors. As such, two inclusions of different sizes were simulated: one of radius 4 mm, and one of radius 2 mm, both with $\mu_{a,f}$ values ten times that of the simulated autofluorescence of the surrounding tissue. The images in the top row in Figs. 1 and 2 correspond to the fluorescence distribution ($\mu_{a,f}$) in the phantom.

Imaging simulations were performed by varying two different types of parameters: detection geometry and FEM mesh resolution. First, different imaging geometries were considered that are consistent with the imaging system presented in [17,28]. Briefly, in those imaging systems, illumination is achieved with a pencil-beam laser source at different points on the surface. For each illumination point, focused detection is done in transmission using five photodetection channels. This detection geometry is installed on an imaging gantry that can rotate around the specimen allowing as many optical projections as desired to be acquired. In this work, numerical simulations were done associated with the following geometries: 8 sources (40 measurements), 16 sources (80 measurements), 32 sources (160 measurements), 64 sources (320 measurements), 128 sources (640 measurements), and 360 sources (1800 measurements). The other variable parameter that is considered in the simulations is the spacial resolution of the mesh. FEM meshes of the circular phantom were created associated to eight different degrees of spatial resolution corresponding to the following number of nodes: 106, 160, 275, 499, 736, 1207, 2333, 6414.

4. RESULTS AND THEIR MATHEMATICAL EXPLANATION

As explained in Section 2, diffuse fluorescence tomography can be treated as a linear problem in the limit that the absorption associated with the fluorophores is small relative to the absorption caused by tissue chromophores such as hemoglobin.

For each combination of FEM meshes and detection geometries, a Jacobian matrix A is calculated based on solutions to the diffusion equation; see Eq. (8). Then, using a basis for its kernel from the SVD, the projection of $\mu_{a,f}$ to the kernel is found. As explained below, this projection is directly associated with information loss that cannot be reconstructed using the standard least-squares regularization techniques presented in this work. This

nonrecoverable portion of $\mu_{a,f}$ i.e., the projection of $\mu_{a,f}$ to the kernel of the Jacobian, is displayed in the bottom row in Figs. 1 and 2, whereas the portion of $\mu_{a,f}$ that does lie in the orthogonal complement of the kernel (i.e., information that is in principle available for image reconstruction using the standard techniques described below in Section 4) is displayed in the middle row.

It is remarked that the portion of $\mu_{a,f}$ that lies in the orthogonal complement of the kernel (i.e., the middle row in Figs. 1 and 2) could also be obtained by computing a data vector $y_{\text{data}} = A\mu_{a,f}$ and then performing image reconstruction using the SVD decomposition based on the simulated data vector. The projection to the kernel of $\mu_{a,f}$ (i.e., the information loss) could then be computed by taking the difference between the original $\mu_{a,f}$ and the reconstruction. Note that this last statement applies only to the situation where no errors are introduced in the data vector prior to SVD reconstruction. For example, using different meshes for forward modeling and SVD image reconstruction would lead to another type of information loss due to the propagation of discretization and model mismatch errors in the reconstructed images.

In order to quantify the amount of information loss through projection to the kernel, we show in Fig. 3 the scaled error between the target image and the recovered image, computed using the equation

$$\text{error} = \frac{\|\mu_{a,f} - \mu_{a,f}^{\text{recon}}\|}{\|\mu_{a,f}\|} = \frac{\|\mu_{a,f}^{\text{kernel}}\|}{\|\mu_{a,f}\|}$$

as a function of mesh density. Here $\mu_{a,f}^{\text{recon}}$ denotes the reconstructed image, whereas $\mu_{a,f}^{\text{kernel}}$ denotes the projection of $\mu_{a,f}$ to the kernel of A . The norm $\|\cdot\|$ is calculated as the standard Euclidean norm, that is, the square root of the sum of the squares of the values at all nodes. The mesh density (in units of nodes per millimeter) is calculated as the square root of the number of nodes divided by the area of the phantom.

As has been proposed here, the displayed errors actually correlate with the size of the nonrecoverable information (as a proportion of the size of the true solution).

Our numerical experiments for those imaging geometries where the number of measurements is small compared to the number of nodes in the mesh show that the forward model A in FT projects to its kernel a significant portion of the optical properties vector, and this projection increases as the spatial resolution of the light transport mesh increases. This projection—regarded as a vector $v_o \in \text{Ker}(A)$ in the canonical decomposition $\mathbb{R}^{N_v} = \text{Ker}(A) \oplus \text{Ker}(A)^\perp$ —has a large relative magnitude (often between 1/3 and 1/2 of the optical properties vector) and is unrecoverable via certain canonical regularization reconstruction methods. We therefore term this projection as “information loss.” Our numerical experiments using a diffuse light propagation model provide explicit examples demonstrating this loss of information in actual fluorescence reconstructions using an imaging geometry similar to what is commonly used in small-animal FMT [17]. We note that we are also considering, in our numerical experiments, detection geometries with a large enough number of source–detector pairs and a small enough number of nodes in the FEM mesh so as to no longer result in an underdetermined reconstruction problem. These correspond to those portions of the graphs in Fig. 3 that are flat and on the horizontal axis; i.e., there is no observed reconstruction error for these experiments. We have included these experiments in this exposition to demonstrate that, in order to increase the resolution in FT reconstruction, not only is a fine mesh necessary but in addition the detection geometry

needs to be such that a large enough number of measurements exist to avoid information loss through projection into the kernel.

In what follows we examine mathematically the meaning of the projection to the kernel of the forward model and explain why this projection results in information loss. By analyzing certain standard regularization methods, we demonstrate that projection into the kernel of the FT forward operator always results in reconstructed images where the information contained in the kernel has been lost and thus cannot be recovered.

We begin with the following elementary example. A restricted form of the underdetermined least-squares inversion method can be formulated as follows: for a measurement vector $y \in \text{Ran}(A)$, the so-called least-squares solution μ_{LS} is given by

$$\mu_{LS} = \min_{\mu: A\mu=y} \|\mu\|^2;$$

that is, we choose for a solution μ_{LS} out of the collection of all possible solutions of $A\mu = y$ as being the one that minimizes the square of the norm. Suppose now that a particular true optical properties distribution μ_{true} is known; that is, “true” has the meaning that, for a particular domain, the optical properties have been exactly measured. Under the forward model A the observation $y_{\text{meas}} = A\mu_{\text{true}}$ is calculated, and then, to test the efficacy of the least-squares inversion method under the assumption of this model A , a reconstruction is performed by calculating the least-squares solution μ_{LS} from the observation y_{meas} . Observe first that μ_{true} decomposes uniquely via the direct sum [Section 2.C, Eq. (9)] into

$\mu_{\text{true}} = \mu_{\text{true},0} + \mu_{\text{true}}^\perp$, where $\mu_{\text{true},0}$ lies in the kernel of A and $\mu_{\text{true}}^\perp$ is in the orthogonal complement of the kernel of A . Likewise, observe as well that *any* solution μ that satisfies $A\mu = y_{\text{meas}}$ must be of the form $\mu = \mu_0 + \mu^\perp$, where $\mu^\perp$ is in $\text{Ker}(A)^\perp$ and μ_0 is in $\text{Ker}(A)$. Note that, for any choice of solution μ in the set of all possible solutions, the projection of μ into $\text{Ker}(A)^\perp$ must be identical to $\mu_{\text{true}}^\perp$, for otherwise A would not act in a one-to-one fashion on $\text{Ker}(A)^\perp$. Using the fact that μ_0 and $\mu_{\text{true}}^\perp$ are orthogonal to each other for any choice of $\mu_0 \in \text{Ker}(A)$, then amongst all such solutions μ the minimum norm solution is $\mu_{LS} = \mu_{\text{true}}^\perp$. It is thus observed that the projection $\mu_{\text{true},0}$ of μ_{true} onto the kernel of A cannot be reconstructed in this setting, and in this sense $\mu_{\text{true},0}$ must be considered as being information that is “lost” by the forward model A under this form of the method of least squares.

In fact, for a given forward model the information associated with vectors contained in the kernel cannot always be retrieved given a particular choice of regularization. In support of this assertion, we consider several standard regularization techniques and demonstrate the phenomenon of information loss in optical tomography from first principles.

A. Tikhonov Regularization

More generally, let $A: \mathbb{R}^{N_v} \rightarrow \mathbb{R}^{N_m}$ be a linear transformation. The method of Tikhonov regularization (see [29,30]) seeks to mediate between fitting observed measurements $y_{\text{meas}} \in \mathbb{R}^{N_m}$ and fidelity to prior knowledge of some set of characteristics (e.g., size or smoothness) of the true solution $\mu \in \mathbb{R}^{N_v}$. In its simplest form, Tikhonov regularization, for $\alpha \in (0, \infty)$, is the solution μ_α of the operator equation (Theorem 2.5, [30]):

$$(A^T A + \alpha I) \mu = A^T (y_{\text{meas}}), \quad (10)$$

which is derived as the solution to the minimization problem

$$\mu_\alpha = \operatorname{argmin} \left\{ \|A\mu - y_{\text{meas}}\|^2 + \alpha \|\mu\|^2 \right\}. \quad (11)$$

We work with Eq. (10):

$$\mu_\alpha = \frac{1}{\alpha} \left(A^T y_{\text{meas}} - A^T A \mu_\alpha \right) = \frac{1}{\alpha} \left(A^T (y_{\text{meas}} - A \mu_\alpha) \right). \quad (12)$$

Recall that the range of the transpose A^T is the orthogonal complement to the kernel of A (e.g., see [30], Appendix A). Thus, the right-hand side of Eq. (12) represents an element of the orthogonal complement of the kernel of A , implying that μ_α must lie in $\operatorname{Ker}(A)^\perp$. Thus, Tikhonov regularization, as realized in Eq. (11), cannot reconstruct that part of any solution that lies in the kernel.

Two more examples are now provided of popular reconstruction techniques that produce solutions in the orthogonal complement of the kernel of A .

B. Truncated Singular Value Decomposition

A brief overview of the SVD of the matrix A has been presented in Subsection 2.C.2. The SVD can also be used as a reconstruction technique. This method can be summarized as one that uses geometric information about the matrix A to project the measurement vector y_{meas} into the range of A . One then solves the equation $A\mu = \operatorname{Proj}(y_{\text{meas}})$ in chosen a subspace of the orthogonal complement of the kernel of A and uses this value as the reconstructed solution of $A\mu = y_{\text{meas}}$.

A short description of this method is provided. Recall from our previous discussion of the SVD of $A \in M_{N_m \times N_v}$ that there exist orthogonal matrices U and V and a diagonal matrix Σ , so that

$$A = U \Sigma V^T \quad (13)$$

with

1. $U \in M_{N_m \times N_m}$ and $V \in M_{N_v \times N_v}$.
2. $\Sigma \in M_{N_m \times N_v}$ is a matrix in diagonal form and whose diagonal elements satisfy

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{N_m} \geq 0.$$

The focus remains on the underdetermined case ($N_m < N_v$). Thus, the index value p , which denotes the smallest nonzero singular value σ_p , is strictly less than N_v ; for ease of

exposition we will assume that $p = N_m$ (i.e., $\sigma_{N_m} > 0$). Via the Rank Theorem (Theorem 2.1), the kernel of A is a nontrivial subspace of $\mathbb{R}^{N_v}$. In the general implementation of the truncated singular value inversion method, an index $t \leq p$ is chosen that denotes the number of positive singular values to be used in the reconstruction; once again, for ease of exposition, we choose $t = p$. Since the columns of U are a basis for $\mathbb{R}^{N_m}$, we have that

$\sum_{j=1}^m \langle y_{\text{meas}}, u_j \rangle u_j = y_{\text{meas}}$. A consequence of Eq. (13) is that $A(v_j) = \sigma_j u_j$, $1 \leq j \leq p$ and $A^T(u_i) = \sigma_i v_i$, $1 \leq i \leq p$ (see Corollary 6.2 in Appendix A). Expressing both μ in terms of the basis $\{v_j\}$ and y_{meas} in terms of the basis $\{u_j\}$ and taking the inner product of both sides of $A\mu = y_{\text{meas}}$ with u_i it is found that the truncated singular value solution μ_p is

$$\mu_p = \sum_{i=1}^p \frac{1}{\sigma_i} \langle y_{\text{meas}}, u_i \rangle v_i,$$

which is in the orthogonal complement of the kernel of A since, as observed in Proposition 6.1 in Appendix A, $\{v_{p+1}, \dots, v_{N_v}\}$ is a basis of the kernel of A .

C. Landweber–Fridman Iteration

Landweber–Fridman iteration is a reconstruction scheme that is based on fixed point iteration. The key point is that the iteration of a contracting map on a closed and complete space contains a unique fixed point and the computed fixed point is then taken to be the reconstruction. Using the normal equations

$$A^T A \mu = A^T P y_{\text{meas}} + A^T (I - P) (y_{\text{meas}}) = A^T y_{\text{meas}} \quad (14)$$

on $y_{\text{meas}} = A\mu$, where P is the projection of y_{meas} onto the range of A in $\mathbb{R}^{N_m}$, one can show that the affine map

$$T(\mu) = \mu + \beta (A^T y_{\text{meas}} - A^T A \mu)$$

is contracting for any choice of β so that $0 < \beta < \frac{2}{\sigma_1^2}$, where σ_1 is the largest singular value in the SVD of A . If we start at $\mu^0 = 0$, it is an easy observation that $\mu_{k+1} = T(\mu_k)$ is in $\text{Ker}(A)^\perp$ for all k . As $\text{Ker}(A)^\perp$ is a vector subspace of $\mathbb{R}^{N_v}$, the fixed point will be in $\text{Ker}(A)^\perp$.

The above reconstruction methods do not represent a complete list of all methods for which the projection to the kernel is unrecoverable; another example is Kaczmarz iteration as implemented in algebraic reconstruction technique [30].

5. DISCUSSION

It is well known that optical tomography images are plagued by poor spatial resolution. Past studies have linked this limitation to the physical nature of photon transport in tissue, i.e., that it is highly scattered. The highly ill-posed nature of the diffuse tomography models means that even minute levels of noise or model mismatch can push an inversion scheme away from finding biologically reasonable solutions. The choice and use of regularization techniques in optical tomography is thus a matter of import and delicacy. In this work new light has been shed on the optical tomography problem by investigating an intrinsic maximum information content, associated with a given optical tomography problem, that can be recovered using certain standard regularization techniques. We have demonstrated that this information content is intimately related to the nature of the imaging geometry and that it is also controlled by the spatial resolution of the mesh with which reconstruction is achieved. The analysis presented here has purposefully avoided the introduction of such issues as propagation of “errors” such as stochastic noise, general discretization error, and light transport model mismatch. The goal was to focus solely on a particular meshing error associated to the discretization of the FT forward model and to analyze numerically how it is affected by varying such parameters as the number of optical measurements and the spatial resolution of the FEM mesh used to generate the forward model. The sole optical imaging model studied in this work is diffuse fluorescence tomography, though we believe that our

results are also applicable to the related diffuse optical tomography and bioluminescence tomography problems.

More precisely, we have been able to demonstrate, and quantify, that a significant portion of an optical properties vector is mapped by the FT forward model into its kernel, and examples have been given of regularization processes that cannot reconstruct the part of the vector that is mapped into the kernel. It has further been demonstrated numerically, under the assumption that the number of measurements stays fixed, that the size of the projection to the kernel grows as the mesh spatial resolution grows (number of nodes). In other words, the larger the dimension of the matrix kernel, the more information is potentially lost during the reconstruction process. It is true, as one increases the mesh density while leaving the number of measurements fixed, that the problem is becoming more underdetermined, but that in itself is not enough to mathematically guarantee that the size of the projection to the kernel will grow (see Example 6.3 in Appendix A). Thus, this property of projecting to the kernel in increasing amounts as the problem becomes more underdetermined is a fundamental property of the diffuse optical tomography forward model and not a mathematical artifact.

As mentioned earlier, we have found initial evidence that shows that the smoothing properties of the elliptic forward model that is used in the diffusion approximation to light transport in tissue may be partially responsible for the information loss. To corroborate this assertion, we show in Fig. 4 a graph that depicts the relative size of the projection to the kernel of optical properties vectors as a function of the smoothness of these vectors. The different smoothing states of the optical properties vectors were obtained by successively applying increasing amounts of Gaussian smoothing to a fixed initial vector. It is evident that the smoother the optical properties vector is, the smaller the projection into the kernel of the forward operator becomes. We plan to explore this phenomenon further in a subsequent paper.

The numerical experiments presented in Fig. 3 also demonstrate that the magnitude of the information loss (size of the projection to the kernel relative to the size of the projection to its complement) depends on the detection geometry that is considered, with preferable geometries being associated with a larger number of optical measurements. An interesting case is the 8 sources (40 measurements) curve for which the reconstruction error remains relatively flat for all node densities. In this case, because there are so few measurements, the reconstruction error remains high even when the number of nodes is at its smallest. It should be noted in closing that the work presented here is expected to have a significant impact mainly for tissue imaging geometries where the number of measurements is limited, such as point-by-point detection configurations based on photomultiplier tubes. In fact, the impact of information loss through the forward matrix kernel is not expected to be as significant when high density of measurements are available using multiple projections based on wide-field CCD detection. Several publications have demonstrated in the past that using high densities of measurements can afford significant spatial resolution gains for both FT [31] and DOT [32] instruments.

An important conclusion that comes out of the analysis presented here is that, in situations when the number of measurements is limited, one must balance the search for increased spatial resolution against the increased loss of information by the FT forward model under an increasingly fine meshing of the domain. For a fixed number of sources and detectors, as one increases mesh density in an attempt to improve resolution, it is important that one avoids certain standard regularization techniques that, for mathematical reasons, cannot reconstruct the information that is lost by the forward model. In fact, it was demonstrated here that, if image reconstruction in optical tomography is to be done using regularization

techniques such as Tikhonov, truncated SVD, or Landweber–Fridman, for example, then in order to gain a finer resolution it is not sufficient to increase the density of the reconstruction mesh, and the only available option is to increase the number of measurements. As a rule of thumb, meshes for which the number of nodes is comparable to the number of measurements will suffer less from the information loss described in this paper, and typically no gains can be expected from using significantly higher-resolution meshes. Of course, there are situations where high-resolution meshes are unavoidable when, for example, sharp boundaries need to be modeled. However, in such situations if possible the number of sources and detectors should be increased to offset the information loss associated with the higher resolution. It is of course well known that increasing the number of measurements is beneficial in optical tomography (see, e.g., [32,33]). The findings here are rather that, under limited tissue sampling conditions, increasing the spatial resolution of the reconstruction mesh can be detrimental to image quality and that this is not only caused by the underdetermined nature of the problem. Further and preliminary numerical experimentation has also revealed that boundary regions in a true solution (e.g., edges of a fluorescent lesion) account for a significant portion of the projection to the kernel. We hope to explore this issue, along with possible inverse technique remedies, in a subsequent paper.

Acknowledgments

This work has been funded by National Science Foundation grants DMS-1031954 and DMS-1031955 and National Cancer Institute grants RO1 CA120368 and K25 CA138578. The authors P. B.-T. and E. C. T. thank the Thayer School of Engineering at Dartmouth College for providing a stimulating and conducive environment for research during the academic year 2010–2011.

APPENDIX A: SINGULAR VALUE ANALYSIS AND AN EXAMPLE

This appendix has two purposes. First, in service of making the paper self contained, a pair of standard linear algebra results used in the paper is established. Second, an elementary example is provided of a matrix having a high-dimensional kernel along with a vector of large magnitude that does not project into the kernel. The point of this observation is that being underdetermined is in and of itself not enough to guarantee information loss; rather it is a property of the FT forward model itself.

The reader is reminded of the standing assumption that $N_v > N_m$, and we let σ_p ($p \leq N_m$) be the smallest positive singular value. Both results that are established are independent of the condition $N_m \leq N_v$, but both to avoid switching notation and because we are only concerned with the undetermined case, this convention is maintained. The following proposition gives a computationally convenient basis for the kernel of A , in terms of the SVD, that we used to calculate the size of the projection into the kernel.

Proposition 6.1

Let $A \in \mathcal{L}(\mathbb{R}^{N_v}, \mathbb{R}^{N_m})$. Then the collection of orthonormal vectors $\{\mathbf{v}_{p+1}, \dots, \mathbf{v}_n\}$, chosen from $(\mathbf{v}_i)_{i=1}^{N_v}$, is an orthonormal basis for the kernel of A .

Proof

Owing to the assumptions that σ_p is the smallest non-zero singular value and that the collection $(\mathbf{u}_j)$ is orthonormal, we can assert that the range of A contains a maximal collection of p linearly independent vectors, which implies that the rank of A is p . Thus, via the Rank Theorem, we have that the dimension of the kernel of A is $N_v - p$.

Let $u_k \in \{v_{p+1}, \dots, v_{N_v}\}$. Then $Av_k = U(\Sigma V^T u_k)$. Since the collection $(v_i)_{1 \leq i \leq N_v}$ is orthonormal, we know that $v_k^T v_k = 1$ for all k and furthermore that $1 \leq k \leq n$ for all $1 \leq k \leq n$. The columns of the matrix V consist of the vectors v_i , and thus $V^T v_k$ is a vector all of whose components are zero, except for the k th component, which equals 1. But multiplying this vector from the right with the diagonal matrix Σ , whose diagonal entries from place $p+1$ on (and in particular in place k) are zero, results in the zero vector. This implies that $Av_k = 0$, so that each orthonormal vector in the collection $\{v_{p+1}, \dots, v_{N_v}\}$ is in the kernel of A . As any collection of $N_v - p$ linearly independent vectors in a $N_v - p$ dimensional vector space is a basis, we observe that $\{v_{p+1}, \dots, v_{N_v}\}$ is a basis of the kernel of A .

A useful and well-known corollary to this proposition is recorded. The SVD of the matrix A also provides a change of basis that reveals that the transpose A^T also has a diagonal action.

Corollary 6.2

Let $\{(v_i), (u_j), (\sigma_j)\}$, be the SVD of A . Then for $1 \leq j \leq p$, we have

$$A^T u_j = \sigma_j v_j.$$

Proof

$A^T u_j \in \mathbb{R}^{N_v}$ of which $\{v_1, \dots, v_{N_v}\}$ is a basis, and so

$$\begin{aligned} A^T u_j &= \sum_{i=1}^{N_v} \langle A^T u_j, v_i \rangle v_i \\ &= \sum_{i=1}^p \langle A^T u_j, v_i \rangle v_i + \sum_{i=p+1}^{N_v} \langle A^T u_j, v_i \rangle v_i \\ &= \sum_{i=1}^p \langle u_j, A v_i \rangle v_i + \sum_{i=p+1}^{N_v} \langle u_j, A v_i \rangle v_i \\ &= \sum_{i=1}^p \delta(i, j) \sigma_i v_i + 0 \\ &= \sigma_j v_j. \end{aligned}$$

Finally, we finish with the promised example. Fix a positive integer N_m and let $\{N_{vi}\}$ be a sequence of integers greater than N_m for which $\lim_{i \rightarrow \infty} N_{vi} = \infty$. Furthermore, let

$A_i \in \mathcal{L}(\mathbb{R}^{N_{vi}}, \mathbb{R}^{N_m})$. Assume as well that A_i is of rank N_m for all i . The Rank Theorem would thus tell us that, for each i , the dimension of the kernel of A_i is $N_{vi} - N_m$ (and thus going to infinity as i increases). Recall that N_{vi} represents the number of nodes in our mesh at which we wish to reconstruct optical properties, whereas N_m is the number of measurements (observations) that we have made. We have shown in this paper, for in the FT tomography setting, that as N_{vi} increases, the size of the projection onto the kernel of our optical properties (and therefore the model inherent information loss) also increases.

This is an intrinsic property of the FT forward model and not a mathematical artifact, as the following simple example will show.

Example 6.3

The matrix $A_i (i \geq 2)$, as a linear transformation, is in $\mathcal{L}(\mathbb{R}^{2i}, \mathbb{R}^2)$ and explicitly has the form that the first row has ones in the odd numbered columns with zeroes in the even numbered columns while the second row has zeroes in the even numbered columns and ones in the even numbered columns. That is,

$$A_i = \begin{pmatrix} 1 & 0 & 1 & \cdots & 1 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Once again it is easy to see that the rank of A_i is two, regardless of the choice of i , and so by the Rank Theorem the dimension of the kernel of A_i is $2i - 2$ for each choice of i . In this example it is easier to focus on a particular example, as the general case follows easily from induction arguments. Thus, we will explicitly consider the $i = 3$ case A_6 , i.e.,

$$A_6 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Explicitly calculating, we find a basis for the kernel of A_6 as

$$\{(1, 0, -1, 0, 0, 0), (0, 0, 1, 0, -1, 0), \\ \times (0, 1, 0, -1, 0, 0), (0, 0, 0, 1, 0, -1)\},$$

where here the idea is to first show that the collection of vectors is linearly independent and then note, for any finite dimensional vector space of dimension k , that any collection of k linearly independent vectors is indeed a basis.

It is now an easy observation that the constant vector $x = (1, 1, 1, 1, 1, 1)$ is orthogonal to each of the basis vectors, and thus the projection of x onto the kernel of A_6 is the zero vector (that is, “no information” in $(1, 1, 1, 1, 1, 1)$ is lost to the kernel by the action of A_6).

REFERENCES

1. Shah N, Cerussi A, Eker C, Espinoza J, Butler J, Fishkin J, Hornung R, Tromberg B. Noninvasive functional optical spectroscopy of human breast tissue. *Proc. Natl. Acad. Sci. USA*. 2001; 98:4420–4425. [PubMed: 11287650]
2. Pogue BW, Poplack SP, McBride TO, Wells WA, Osterman KS, Osterberg UL, Paulsen KD. Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast. *Radiology*. 2001; 218:261–266. [PubMed: 11152812]
3. Zhang Q, Brukilacchio TJ, Li A, Stott JJ, Chaves T, Hillman E, Wu T, Chorlton M, Rafferty E, Moore RH, Kopans DB, Boas DA. Coregistered tomographic x-ray and optical breast imaging: initial results. *J. Biomed. Opt.* 2005; 10:024033. [PubMed: 15910106]
4. Arridge SR, Hebden JC. Optical imaging in medicine: II. Modelling and reconstruction. *Phys. Med. Biol.* 1997; 42:841–853. [PubMed: 9172263]
5. Arridge SR. Optical tomography in medical imaging. *Inverse Probl.* 1999; 15:R41–R93.
6. Jacques SL, Pogue BW. Tutorial on diffuse light transport. *J. Biomed. Opt.* 2008; 13:041302. [PubMed: 19021310]
7. Ntziachristos V, Tung CH, Bremer C, Weissleder R. Fluorescence molecular tomography resolves protease activity in vivo. *Nat. Med.* 2002; 8:757–761. [PubMed: 12091907]

8. Leblond F, Davis S, Vald's P, Pogue B. Pre-clinical whole-body fluorescence imaging: review of instruments methods and applications. *J. Photochem. Photobiol. B.* 2010; 98:77–94. [PubMed: 20031443]
9. Leblond F, Tichauer KM, Pogue BW. Singular value decomposition metrics show limitations of detector design in diffuse optical tomography. *Biomed. Opt. Express.* 2010; 1:1514–1531. [PubMed: 21258566]
10. Leblond F, Dehghani H, Kepshire D, Pogue BW. Early-photon fluorescence tomography: spatial resolution improvements and noise stability considerations. *J. Opt. Soc. Am. A.* 2009; 26:1444–1457.
11. Culver JP, Ntziachristos V, Holboke MJ, Yodh AG. Optimization of optode arrangements for diffuse optical tomography: a singular-value analysis. *Opt. Lett.* 2001; 26:701–703. [PubMed: 18040425]
12. Arridge SR, Lionheart WRB. Nonuniqueness in diffusion-based optical tomography. *Opt. Lett.* 1998; 23:882–884. [PubMed: 18087373]
13. Yalavarthy P, Dehghani H, Pogue B, Paulsen K. Critical computational aspects of near infrared circular tomographic imaging: analysis of measurement number, mesh resolution and reconstruction basis. *Opt. Express.* 2006; 14:6113–6127. [PubMed: 19516784]
14. Contag CH, Ross BD. It's not just about anatomy: in vivo bioluminescence imaging as an eyepiece into biology. *J. Magn. Reson. Imaging.* 2002; 16:378–387. [PubMed: 12353253]
15. Contag CH, Bachmann MH. Advances in in vivo bioluminescence imaging of gene expression. *Annu. Rev. Biomed. Eng.* 2002; 4:235–260. [PubMed: 12117758]
16. Dehghani H, Davis SC, Jiang S, Pogue BW, Paulsen KD, Patterson MS. Spectrally resolved bioluminescence optical tomography. *Opt. Lett.* 2006; 31:365–367. [PubMed: 16480210]
17. Kepshire DS, Mincu N, Hutchins M, Guber J, Dehghani H, Hynarowski J, Leblond F, Khayat M, Pogue BW. A microcomputed tomography guided fluorescence tomography system for small animal molecular imaging. *Rev. Sci. Instrum.* 2009; 80:043701. doi:10.1063/1.3109903. [PubMed: 19405660]
18. Dehghani H, Eames M, Yalavarthy P, Davis S, Srinivasan S, Carpenter C, Pogue B, Paulsen K. Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction. *Commun. Numer. Meth. Eng.* 2009; 25:711–732.
19. Groenhuis RAJ, Ferwada HA, Ten Bosch JJ. Scattering and absorption of turbid materials determined from reflection measurements. 1: Theory. *Appl. Opt.* 1983; 22:2456–2462. [PubMed: 18196156]
20. Groenhuis RAJ, Ten Bosch JJ, Ferwada HA. Scattering and absorption of turbid materials determined from reflection measurements. 2: Measuring method and calibration. *Appl. Opt.* 1983; 22:2463–2467. [PubMed: 18196157]
21. Paithankar D, Chen A, Pogue B, Patterson M, Sevick-Muraca E. Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media. *Appl. Opt.* 1997; 36:2260–2272. [PubMed: 18253202]
22. Friedberg, S.; Insel, A.; Spence, L. *Linear Algebra*. 4th ed. Prentice Hall; 2003.
23. Golub, G.; van Loan, CF. *Matrix Calculations*. Johns Hopkins University; 1989.
24. Paulsen K, Jiang H. Spatially varying optical property reconstruction using a finite element diffusion equation approximation. *Med. Phys.* 1995; 22:691–701. [PubMed: 7565358]
25. Arridge SR, Schweiger M, Hiraoka M, Delpy DT. A finite element approach for modeling photon transport in tissue. *Med. Phys.* 1993; 20:299–309. [PubMed: 8497214]
26. Zhu Q, Dehghani H, Tichauer KM, Holt RW, Vishwanath K, Leblond F, Pogue BW. A three-dimensional finite element model and image reconstruction algorithm for time-domain fluorescence imaging in highly scattering media. *Phys. Med. Biol.* 2011; 56:7419–7434. [PubMed: 22056913]
27. Leblond F, Tichauer KM, Holt RW, El-Ghoussein F, Pogue BW. Toward whole-body optical imaging of rats using single-photon counting fluorescence tomography. *Opt. Lett.* 2011; 36:3723–3725. [PubMed: 21964076]

28. Tichauer KM, Holt RW, El-Gusseini F, Zhu Q, Dehghani H, Leblond F, Pogue BW. Imaging workflow and optical data calibration for CT-guided whole-body time-domain fluorescence tomography. *Biomed. Opt. Express*. 2011; 2:3021–3036. [PubMed: 22076264]
29. Hansen, PC. *Discrete Inverse Problems: Insight and Algorithms*. SIAM; 2010.
30. Kaipio, J.; Somersalo, E. *Statistical and Computational Inverse Problems*. Springer-Verlag; 2005.
31. Chen J, Venugopal V, Lesage F, Intes X. Time-resolved diffuse optical tomography with patterned-light illumination and detection. *Opt. Lett.* 2010; 35:2121–2123. [PubMed: 20596166]
32. Konecky SD, Panasyuk GY, Lee K, Markel V, Yodh AG, Schotland JC. Imaging complex structures with diffuse light. *Opt. Express*. 2008; 16:5048–5060. [PubMed: 18542605]
33. Wang ZM, Panasyuk GY, Markel VA, Schotland JC. Experimental demonstration of an analytic method for image reconstruction in optical diffusion tomography with large data sets. *Opt. Lett.* 2005; 30:3338–3340. [PubMed: 16389824]

Fig. 1. (Color online) Image reconstruction and information loss for an imaging geometry with 32 sources and 5 detectors. The top row shows the true solutions, the middle row shows the reconstructed images using the maximal number of SVD modes, and the bottom row presents the information that is lost through regularization. The images are shown for an increasing spatial resolution from left to right.

Fig. 2.

(Color online) Image reconstruction and information loss for an imaging geometry with 128 sources and 5 detectors. The top row shows the true solutions, the middle row shows the reconstructed images using the maximal number of SVD modes, and the bottom row presents the information that is lost through regularization. The images are shown for an increasing spatial resolution from left to right.

Fig. 3. (Color online) Reconstruction error versus node density for all detection geometries considered. The error decreases as the number of measurement increases, while the error typically increases with the number of nodes.

Fig. 4.

(Color online) Reconstruction error as a function of smoothness of the optical properties vector. A value of 1 indicates that no smoothing has been applied (and thus the optical properties vector exhibits sharp edges), whereas a value of 10 indicates the largest amount of smoothing that has been applied.