This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

Light Illumination and Detection Patterns for Fluorescence Diffuse Optical Tomography based on Compressive Sensing

An Jin *Student Member, IEEE*, Birsen Yazıcı* *Senior Member, IEEE*, and Vasilis Ntziachristos *Member, IEEE*

***Abstract*—Fluorescence diffuse optical tomography (FDOT) is an emerging molecular imaging modality that uses near infrared light to excite the fluorophore injected into tissue; and to reconstruct the fluorophore concentration from boundary measurements. The FDOT image reconstruction is a highly ill-posed inverse problem due to a large number of unknowns and limited number of measurements. However, the fluorophore distribution is often very sparse in the imaging domain since fluorophores are typically designed to accumulate in relatively small regions.**

In this paper, we use compressive sensing (CS) framework to design light illumination and detection patterns to improve the reconstruction of sparse fluorophore concentration. Unlike the conventional FDOT imaging where spatially distributed light sources illuminate the imaging domain one at a time and the corresponding boundary measurements are used for image reconstruction, we assume that the light sources illuminate the imaging domain simultaneously several times and the corresponding boundary measurements are linearly filtered prior to image reconstruction. We design a set of optical intensities (illumination patterns) and a linear filter (detection pattern) applied to the boundary measurements to improve the reconstruction of sparse fluorophore concentration maps. We show that the FDOT sensing matrix can be expressed as a columnwise Kronecker product of two matrices determined by the excitation and emission light fields. We derive relationships between the incoherence of the FDOT forward matrix and these two matrices and use these results to reduce the incoherence of the FDOT forward matrix. We present extensive numerical simulation and the results of a real phantom experiment to demonstrate the improvements in image reconstruction due to CS-based light illumination and detection patterns in conjunction with relaxation and greedy-type reconstruction algorithms.

I. Introduction

Fluorescence diffuse optical tomography (FDOT) is an imaging modality that uses near infrared (NIR) light to measure the 3D fluorophore activity inside biological tissue [1]–[6]. FDOT is minimally invasive, relatively inexpensive and portable imaging modality. Furthermore, it provides information that is not available in conventional imaging modalities [1], [2].

Yazici and Jin are with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180 USA E-mail: B. Y: yazici@ecse.rpi.edu, Phone: (518)-276 2905, Fax: (518)-276 6261

Ntziachristos is with Institute for Biological and Medical Imaging (IBMI), Helmholtz Zentrum München, D-85764 Neuherberg, Germany

This work was partly supported by the National Science Foundation Engineering Research Center Censsis under the Grant No. ERC-08030672.

* Corresponding author.

The inverse problem in FDOT involves reconstruction of the three-dimensional (3D) fluorophore distribution inside the imaging domain using the boundary measurements obtained at the emission and excitation wavelengths [3], [4], [6]. The FDOT inverse problem is highly ill-posed since the number of measurements available is typically much smaller than the number of unknowns [7], [8].

Most optical fluorophore are "targeted probes", in that, they are designed to accumulate in relatively small, specific regions in tissue such as tumors. As a result, the fluorophore distribution in the imaging domain is often very sparse. Compressive sensing (CS) theory provides exact recovery of sparse signals from a small number of measurements under certain conditions. These conditions are given in terms of the incoherence of the sensing matrix that maps the unknown signal of interest to measurements [9]–[12]. In many applications, the forward matrix can be designed to improve the recovery of sparse signals [10]–[16].

In this paper, we design light illumination and detection patterns for FDOT based on the CS theory to improve the quality of reconstructed images.

A. Related Literature

In FDOT, the imaging domain is often illuminated sequentially by a number of point sources placed at the boundary of the imaging domain. However, recent work has demonstrated that the "structured illumination" can be useful in reducing the ill-posedness of the optical tomography inverse problem, and to improve the spatial resolution of the reconstructed images [17]–[23]. In [17], a method to design distribution of light intensities over a predetermined set of locations and modulation frequencies to maximize the detection of heterogeneities is presented. In [18], different illumination light patterns including line sources, Gaussian spots, and their combination are explored to reduce the ill-posedness of the FDOT problem. In [21] and [22], spatially modulated frequency diversity is explored to improve image reconstruction. In [23], the superiority of wide-field patterned light sources over point sources is demonstrated using phantom experiments. In [24], the measurements were optimized using the same method as in this paper, however with a different inverse problem formulation. Randomly located point light sources were used to maximize the rank of the forward matrix. Note that the concept of structured illumination is similar to the coded aperture imaging technique where predetermined temporal or spatial masks/patterns are used to encode the illumination sources to optimize imaging processes, including the ones in medical imaging [25], spectral imaging [26], and others [27].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

Recently, CS theory has been applied to optical tomography [24], [28]–[32]. In [28]–[31], l_1-norm constraint regularization methods are used in FDOT and endogenous diffuse optical tomography (DOT) inverse problems. In [32], improvements in DOT image reconstruction is demonstrated in numerical simulations using a greedy-type algorithm and a small number of measurements. In [24], a non-iterative inversion method based on generalized MUSIC algorithm was introduced and improvements in DOT image reconstruction are shown in 3D numerical simulations. Although CS techniques have shown promising results in optical tomography, it has not been explored to design light illumination and detection patterns.

B. Our Approach

In this paper, we design light illumination and detection patterns to improve the incoherence of the FDOT forward matrix, and hence, to improve the reconstruction of fluorophore concentration images. We assume that the endogenous diffusion and absorption coefficients are known.

Unlike the conventional FDOT imaging where point sources with constant intensity are time-multiplexed, i.e., turned on one at a time, we assume that multiple spatially distributed light sources illuminate the medium simultaneously and the corresponding boundary measurements are linearly filtered prior to image reconstruction. In general the FDOT image reconstruction is a nonlinear inverse problem. The non-linearity can be addressed by employing an iterative perturbation approach wherein the reconstruction of the fluorophore map perturbation becomes a linear problem at each iteration. In this paper, under the weak fluorophore assumption [4], we formulate the FDOT image reconstruction as a linear inverse problem.

In CS theory, accurate recovery of a sparse signal from an underdetermined linear system requires the underlying forward matrix to be incoherent. We show that the FDOT forward matrix can be expressed as a columnwise Kronecker product of two matrices. The first one is determined by the excitation light field which depends on the spatial configuration of light sources, their intensities, and the Green's function of the endogenous background. The second matrix is determined by the emission light field which depends on the spatial configuration of the detectors and the Green's function of the emission light field. The incoherence of the FDOT forward sensing matrix is related to the incoherence of these two underlying matrices. We design two preconditioners, which we refer to as optical and measurement masks, to reduce the coherence of these two matrices. The optical mask is an intensity matrix, with each of its rows being an intensity/illumination pattern applied to the point sources. The measurement masks is a linear filter applied to the measured data collected by the detectors in each illumination.

We present extensive numerical simulations and a real silicon phantom experiment to show that the application of optical and measurement masks reduces the coherence of the FDOT forward matrix. We use a number of sparse signal recovery techniques, including greedy type and convex relaxation algorithms to demonstrate the improvements in fluorophore reconstruction when optical and measurement masks are applied.

Our approach can be used to determine not only the optimal source intensities, but also the location of sources and detectors. Given a fluorophore designed to accumulate in certain regions and anatomical a priori information, optimal source detector locations and source intensities can be determined via numerical simulations assuming a large array of sources and detectors prior to the imaging process. This may eliminate unnecessary illumination, optimize imaging process and result in better image quality than the conventional FDOT imaging.

While our primary interest is in FDOT imaging, similar approach can be applied to other wave based imaging problems, such as DOT, electromagnetic and acoustic imaging.

C. The Organization of the Paper

The rest of the paper is organized as follows: Section II presents the FDOT forward problem. Section III presents the FDOT inverse problem formulation as a sparse signal recovery problem and briefly reviews the CS theory from a perspective relevant to the subsequent discussion. In Section IV, we derive bounds on the incoherence of the FDOT forward matrix. In V, we present the design of the optical and measurement masks. In Sections VI and VII, we present the performance of our approach in 3D numerical simulations and in a real silicon phantom experiment. Finally, Section VIII concludes the paper.

II. FDOT Imaging Problem

We use the following notational conventions throughout the paper. The bold symbol, bold italic lower-case letters, $\mathbf{\Gamma}, \boldsymbol{y}$, etc. are used to denote vector variables, italic symbol or Roman letters are used to denote scalar variables. The calligraphic letters, $\mathcal{A}, \mathcal{B}$, etc. are used to denote operators. The bold Roman upper-case letters are used to denote matrices, $\boldsymbol{A}, \boldsymbol{B}$, etc. We use $E_{x,m}$ as a shorthand notation for the quantity E at either excitation or emission wavelengths.

A. Light Propagation Model for FDOT

We model the excitation and emission light propagation in a bounded region $\Omega \subset \mathbb{R}^3$ by the following coupled diffusion equations [3]:

$$\begin{aligned} - &\quad \nabla \cdot D_x(\boldsymbol{r}) \nabla \phi_x^i(\boldsymbol{r},\omega) + \left(\mu_{ax}(\boldsymbol{r}) + \frac{j\omega}{c}\right) \phi_x^i(\boldsymbol{r},\omega) \\ = &\quad s^i(\boldsymbol{r},\omega), \quad \boldsymbol{r} \in \Omega, \end{aligned} \tag{1}$$

$$\begin{aligned} - &\quad \nabla \cdot D_m(\boldsymbol{r}) \nabla \phi_m^i(\boldsymbol{r},\omega) + \left(\mu_{am}(\boldsymbol{r}) + \frac{j\omega}{c}\right) \phi_m^i(r,\omega) \\ = &\quad \phi_x^i(\boldsymbol{r},\omega) \eta \mu_{axf}(\boldsymbol{r}) \frac{1 - j\omega\tau(\boldsymbol{r})}{1 + (\omega\tau(\boldsymbol{r}))^2}, \quad \boldsymbol{r} \in \Omega \end{aligned} \tag{2}$$

where the subscripts x and m denote the excitation and emission wavelengths, respectively. $\boldsymbol{r} \in \Omega$ denotes location; ω is the angular modulation frequency. s^i denotes the i^{th} source. $\phi_{x,m}^i$ denotes the photon density field due to the i^{th} source. c is the speed of light in Ω. $D_{x,m}$ is the diffusion coefficient and $\mu_{ax,m}$ is the absorption coefficient. μ_{axf} is the absorption coefficient of the fluorophore at the excitation wavelength; η is the quantum yield; $\eta\mu_{axf}$ is the fluorophore yield; τ is the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

fluorescence lifetime. We assume that $D_{x,m} \approx \frac{1}{3\mu'_{sx,m}}$, where $\mu'_{sx,m}$ is the reduced scattering coefficient. Furthermore, $D_{x,m}$ is known, and can be spatially varying in Ω.

We use Robin type boundary conditions

$$2D_x(\boldsymbol{r})\hat{n}\cdot\nabla\phi^i_x(\boldsymbol{r},\omega)+\rho\phi^i_x(\boldsymbol{r},\omega) = 0, \quad \boldsymbol{r}\in\partial\Omega, \quad (3)$$

$$2D_m(\boldsymbol{r})\hat{n}\cdot\nabla\phi^i_m(\boldsymbol{r},\omega)+\rho\phi^i_m(\boldsymbol{r},\omega) = 0, \quad \boldsymbol{r}\in\partial\Omega, \quad (4)$$

where $\partial\Omega$ denotes the boundary of Ω, $\hat{n}$ denotes the outward normal of the boundary and ρ is the boundary mismatch parameter due to light reflection at the boundary.

For the sake of exposition, we use a single modulation frequency measurement model. We choose $\omega = 0$ and drop ω variable from $\phi_{x,m}$ and other terms. Note that when $\omega \neq 0$ and τ is known, the source term in (2) includes $\frac{1-j\omega\tau(\boldsymbol{r})}{1+(\omega\tau(\boldsymbol{r}))^2}$. However, this modification does not change the key aspects of our development. The subsequent problem formulation and results can be extended in a straightforward way to include multiple frequencies wherever τ is known.

The absorption coefficients $\mu_{ax,m}$ of the medium is composed of the endogenous absorption $\mu_{aex,m}$ and the exogenous absorption $\mu_{afx,m}$ of the fluorophore:

$$\mu_{ax,m}(\boldsymbol{r}) = \mu_{aex,m}(\boldsymbol{r}) + \mu_{afx,m}(\boldsymbol{r}). \quad (5)$$

Since the excitation and emission wavelengths are typically close to each other, in many applications, the absorption coefficients at both wavelengths are assumed to be approximately equal [4]. Hence, we set

$$\mu_a \approx \mu_{ax} \approx \mu_{am}, \quad \mu_{af} \approx \mu_{axf} \approx \mu_{amf}. \quad (6)$$

B. FDOT Forward and Inverse Problems

We assume that there are N_s light sources and N_d detectors located at the boundary of the imaging domain. From (1) and (3), the excitation light field due to the i^{th} source is given by

$$\phi^i_x(\boldsymbol{r}) = \int_\Omega g_x(\boldsymbol{r},\boldsymbol{r}')s^i(\boldsymbol{r}')d\boldsymbol{r}', \ i=1,...,N_s \quad (7)$$

where g_x is the Green's function of (1) and (3). The excitation light interacts with the fluorophore inside the imaging domain and generates the emission light with a slightly longer wavelength

$$\phi^i_m(\boldsymbol{r}) = \int_\Omega g_m(\boldsymbol{r},\boldsymbol{r}')\phi^i_x(\boldsymbol{r}')\eta\mu_{af}(\boldsymbol{r}')d\boldsymbol{r}', \ i=1,...,N_s \quad (8)$$

where g_m is the Green's function of (2) and (4). Note that $\phi^i_x\eta\mu_{af}$ acts as the light source for the photon density field at the emission wavelength.

In general, the optical properties of tissue are wavelength dependent. However, they are usually assumed to be approximately equal at both excitation and emission wavelengths for many of the fluorophore probes used in practice [1]. However, this assumption is not a limitation for our method, since $\mu_\lambda = C\epsilon^\lambda$, where C is the concentration, λ is the wavelength and ϵ is the wavelength independent extinction coefficient. As a result, $\mu_{\lambda_1} = \mu_{\lambda_2}\epsilon^{\lambda_1-\lambda_2}$. It is straightforward to include this known linear relationship between the absorption coefficients at the excitation and emission wavelengths into our model.

We write the measurement $\Gamma_{i,j}$ at the j^{th} detector due to the i^{th} source as

$$\Gamma_{i,j} = \int_\Omega g^j_m(\boldsymbol{r})\phi^i_x(\boldsymbol{r})\eta\mu_{af}(\boldsymbol{r})d\boldsymbol{r}, \ i=1,...,N_s, \ j=1,...,N_d \quad (9)$$

where g^j_m is a shorthand notation for the Green's function of the emission light field evaluated at the j^{th} detector.

In general, (9) is nonlinear, since both g^j_m and ϕ^i_x depend on μ_a, which in turn, depends on μ_{af}. However, for the case of weak fluorophore [4], (9) can be linearized by assuming that the contribution of the fluorophore absorption to the total absorption is negligible, i.e.,

$$\mu_a = \mu_{ae} + \mu_{af} \approx \mu_{ae}. \quad (10)$$

Using (10), (9) can be linearized as

$$\Gamma_{i,j} = \int_\Omega g^j_{em}(\boldsymbol{r})\phi^i_{ex}(\boldsymbol{r})\eta\mu_{af}(\boldsymbol{r})d\boldsymbol{r} \quad (11)$$

where g^j_{em} is the Green's function of (2) and (4), and ϕ^i_{ex} is the excitation field when $\mu_a \approx \mu_{ae}$. We organize all the measurements $\Gamma_{i,j}$, $i=1,...,N_s$, $j=1,...,N_d$, into a vector $\boldsymbol{\Gamma}$ of length $M = N_s \times N_d$

$$\boldsymbol{\Gamma} = [\Gamma_{1,1},...,\Gamma_{1,N_d},\Gamma_{2,1},...,\Gamma_{N_s,N_d}]^T \in \mathbb{R}^M. \quad (12)$$

We define the operator $\mathcal{A}$: $L(\Omega) \to \mathbb{R}^M$ as

$$(\mathcal{A}\mu)_{i,j} := \int_\Omega g^j_{em}(\boldsymbol{r})\phi^i_{ex}(\boldsymbol{r})\mu(\boldsymbol{r})d\boldsymbol{r} \quad (13)$$

where $\mu = \eta\mu_{af}$ denotes the fluorophore yield. Note that $\mathcal{A}$ linearly maps the fluorophore yield μ to the measurement vector $\boldsymbol{\Gamma}$:

$$\boldsymbol{\Gamma} = \mathcal{A}\mu. \quad (14)$$

We discretize the domain Ω into N voxels, $x_i = \mu(\boldsymbol{r}_i)$, $i=1,...,N$, where $\boldsymbol{r}_i$ is the center of the i^{th} voxel; and form the image vector

$$\boldsymbol{x} = [x_1, x_2, ..., x_N]^T \in \mathbb{R}^N.$$

Discretizing (14) by the finite element method, we obtain:

$$\boldsymbol{\Gamma} = \boldsymbol{A}\boldsymbol{x}, \quad (15)$$

where

$$\boldsymbol{A} = \begin{bmatrix} g^1_{em,1}\phi^1_{ex,1} & \cdots & g^1_{em,N}\phi^1_{ex,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^1_{ex,1} & \cdots & g^{N_d}_{em,N}\phi^1_{ex,N} \\ g^1_{em,1}\phi^2_{ex,1} & \cdots & g^1_{em,N}\phi^2_{ex,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^{N_s}_{ex,1} & \cdots & g^{N_d}_{em,N}\phi^{N_s}_{ex,N} \end{bmatrix} \in \mathbb{R}^{M\times N} \quad (16)$$

is the vector-valued forward operator. $g^j_{em,k}$ is the discretized Green's function of (2) and (4) evaluated at the j^{th} detector due to the k^{th} voxel; and $\phi^i_{ex,k}$ is the discretized excitation light field at the k^{th} voxel due to the i^{th} source. Note that we refer to $\boldsymbol{A}$ as the *sensing or forward operator* interchangeably for the rest of the paper. For a detailed discussion on the discretization of the FDOT forward and inverse problems, see [3].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

III. FDOT Inverse Problem with Sparsity Constraint and Sparse Signal Reconstruction

A. FDOT Inverse Problem with Sparsity Constraint

In this section, we present the FDOT inverse problem formulation with sparsity constraint in the presence of additive noise and briefly describe the solutions and performance guarantees offered by sparse signal recovery techniques.

The idealized FDOT inverse problem involves the reconstruction of the fluorophore yield μ from the boundary measurement $\mathbf{\Gamma}$ based on the model (14) or (15). Taking into account the measurement noise, (15) can be modified to

$$\boldsymbol{y} = \mathbf{\Gamma} + \boldsymbol{\epsilon} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\epsilon} \tag{17}$$

where $\boldsymbol{y}$ denotes the noisy measurements, $\boldsymbol{\epsilon} = [\epsilon_1, ..., \epsilon_n]^T \in \mathbb{R}^M$ is the measurement noise with $\|\boldsymbol{\epsilon}\|_2 \leq \varepsilon$, for some $\varepsilon > 0$, and $\|\cdot\|_2$ denotes the l_2-norm.

Since $\boldsymbol{x}$ is sparse, the FDOT image reconstruction can be formulated as the following constrained optimization problem:

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \;\text{ such that }\; \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2 \leq \varepsilon \tag{18}$$

where $\|\boldsymbol{x}\|_0$ denotes the ℓ_0-norm of $\boldsymbol{x}$ defined as the number of non-zero entries of $\boldsymbol{x}$. ℓ_0-norm is a commonly used measure of sparsity [33].

The optimization problem in (18) is NP hard. Therefore, approximate methods are used to address (18). These methods can be roughly categorized into two classes [33]. The first class is based on greedy algorithms which is a group of heuristic iterative algorithms that seek an optimal solution in each iteration [34]. The second class is based on the relaxation technique which replaces the l_0-norm constraint in (18) with l_1-norm constraint:

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \;\text{ such that }\; \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2 \leq \varepsilon \tag{19}$$

where $\|\boldsymbol{x}\|_1$ is the ℓ_1-norm of $\boldsymbol{x}$. The optimization problem in (19) is also known as the basis pursuit (BP) [35]. BP is a convex optimization problem that can be solved by a number of linear programming techniques [35].

B. Mutual Coherence and Performance Guarantees in Sparse Signal Reconstruction

In general, the solution of an underdetermined linear system is non-unique. However, CS theory provides sufficient conditions under which exact recovery of the sparse signals from limited number of measurements is possible [9], [10], [33]. In this subsection, we review some of the results on performance guarantees in CS theory relevant to our subsequent discussion.

Many of the sufficiency conditions proposed in the literature specify the degree to which the forward matrix is orthogonal. One of these conditions is given in terms of the concept of *mutual coherence* [33], which is defined as

$$\mathcal{M}(\boldsymbol{A}) = \max_{p,q,p\neq q} \frac{|\langle \boldsymbol{a}_p, \boldsymbol{a}_q \rangle|}{\|\boldsymbol{a}_p\|_2 \|\boldsymbol{a}_q\|_2} \tag{20}$$

where $\boldsymbol{a}_p$ and $\boldsymbol{a}_q$ are two different columns in $\boldsymbol{A}$. Note that $\mathcal{M}(\boldsymbol{A})$ is the largest normalized inner product of two different columns; and has a small value only if the columns of $\boldsymbol{A}$ are almost orthogonal to each other.

It was shown in [9] that, if $\boldsymbol{x}_0$ satisfies (17) with

$$\|\boldsymbol{x}_0\|_0 \leq \frac{1}{2}(1 + \frac{1}{\mathcal{M}(\boldsymbol{A})}), \tag{21}$$

then, $\boldsymbol{x}_0$ is necessarily the unique vector that satisfies the linear system in (17) whose number of nonzero entries satisfies (21). Furthermore, the difference between $\boldsymbol{x}_0$ and the solutions of the sparsity constraint optimization problems in (18) and (19) are only within a constant multiple of ε [9], [11].

The mutual coherence of $\boldsymbol{A}$ is determined by the two least "incoherent" columns, which shows the worst-case analysis of the orthogonality of the columns of $\boldsymbol{A}$. A number of alternative measures to quantify the "average" coherence of a sensing matrix has been proposed [10], [15], [16], [36]. One of these is the *cumulative coherence* or the *Babel function*, which is defined as [10]

$$\mathcal{M}_1(k, \boldsymbol{A}) = \max_{p} \max_{|Q|=k, p\notin Q} \sum_{q\in Q} \frac{|<\boldsymbol{a}_p, \boldsymbol{a}_q>|}{\|\boldsymbol{a}_p\|_2 \|\boldsymbol{a}_q\|_2}, \tag{22}$$

where Q is a subset of the columns in $\boldsymbol{A}$. If

$$\mathcal{M}_1(k, \boldsymbol{A}) + \mathcal{M}_1(k+1, \boldsymbol{A}) < 1, \tag{23}$$

then, the signal $\boldsymbol{x}_0$ with $\|\boldsymbol{x}_0\|_0 \leq k$ can be exactly recovered using greedy algorithms or basis pursuit [10]. Note that $\mathcal{M}_1(k, \boldsymbol{A})$ is an monotonically non-decreasing function in k, and $\mathcal{M}_1(1, \boldsymbol{A}) = \mathcal{M}(\boldsymbol{A})$.

More generally, a matrix with N columns have $\frac{N(N-1)}{2}$ normalized inner products between different pairs of columns. It was observed that as the average value of the normalized inner products between different columns decreases, the performance of sparse signal recovery methods typically improves [10]–[12], [15], [16], [33]. In particular, it was shown that a well-designed preconditioning matrix can reduce the normalized inner products between the columns of the sensing matrix, and as a consequence, reduce the mean square error of the reconstructed images [11]–[14].

In the following section, we present the coherence of the FDOT forward matrix with respect to a number of measures summarized in this section.

IV. Bounds on the Coherence of the FDOT Forward Matrix

In this section, we present the inherent structure of the FDOT forward matrix and show that it is the column-wise Kronecker product of two matrices: A matrix that is composed of the discretized excitation light fields; and another matrix composed of the discretized Green's function of the diffusion equation governing the emission light field propagation. We show that the normalized inner product between two different columns of the FDOT forward matrix can be expressed as a product of the normalized inner products between the columns of these two underlying matrices. We next derive upper bounds on the mutual coherence and cumulative coherence of the FDOT forward matrix. These observations are essential in the design and optimization of the FDOT forward matrix.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

Recall that the FDOT forward matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \begin{bmatrix} g^1_{em,1}\phi^1_{ex,1} & g^1_{em,2}\phi^1_{ex,2} & \cdots & g^1_{em,N}\phi^1_{ex,N} \\ \vdots & & & \vdots \\ g^{N_d}_{em,1}\phi^1_{ex,1} & g^{N_d}_{em,2}\phi^1_{ex,2} & \cdots & g^{N_d}_{em,N}\phi^1_{ex,N} \\ g^1_{em,1}\phi^2_{ex,1} & g^1_{em,2}\phi^2_{ex,2} & \cdots & g^1_{em,N}\phi^2_{ex,N} \\ \vdots & & & \vdots \\ g^{N_d}_{em,1}\phi^{N_s}_{ex,1} & g^{N_d}_{em,2}\phi^{N_s}_{ex,2} & \cdots & g^{N_d}_{em,N}\phi^{N_s}_{ex,N} \end{bmatrix}.$$

We see that the entries of the $\boldsymbol{A}$ is the product of two quantities: (1) the discretized excitation light field; and (2) the discretized Green's function of the emission light field. To further analyze the structure of the FDOT forward matrix, we define two new matrices. The first one is an $N_s \times N$ matrix with its i^{th} row being the excitation light field due to the i^{th} source:

$$\boldsymbol{\Phi} := \begin{bmatrix} \phi^1_{ex,1} & \cdots & \phi^1_{ex,N} \\ \vdots & \cdots & \vdots \\ \phi^{N_s}_{ex,1} & \cdots & \phi^{N_s}_{ex,N} \end{bmatrix} \in \mathbb{R}^{N_s \times N}. \quad (24)$$

The second one is an $N_d \times N$ matrix with its j^{th} row being the Green's function of the emission light field corresponding to the j^{th} detector:

$$\boldsymbol{G} := \begin{bmatrix} g^1_{em,1} & \cdots & g^1_{em,N} \\ \vdots & \cdots & \vdots \\ g^{N_d}_{em,1} & \cdots & g^{N_d}_{em,N} \end{bmatrix} \in \mathbb{R}^{N_d \times N}. \quad (25)$$

Let $\boldsymbol{a}_k$, $k = 1, ..., N$ be the k^{th} column of the FDOT forward matrix $\boldsymbol{A}$; and let

$$\begin{aligned} \boldsymbol{\phi}_k &= [\phi^1_{ex,k}, ..., \phi^{N_s}_{ex,k}]^T \\ \boldsymbol{g}_k &= [g^1_{em,k}, ..., g^{N_d}_{em,k}]^T \end{aligned} \quad (26)$$

be the k^{th} column of $\boldsymbol{\Phi}$ and $\boldsymbol{G}$, respectively. We observe $\boldsymbol{a}_k$ is the Kronecker product of $\boldsymbol{\phi}_k$ and $\boldsymbol{g}_k$, and write

$$\boldsymbol{a}_k = \boldsymbol{\phi}_k \otimes \boldsymbol{g}_k, \quad k = 1, ..., N. \quad (27)$$

Thus, $\boldsymbol{A}$ is the columnwise Knocker product of $\boldsymbol{\Phi}$ and $\boldsymbol{G}$:

$$\boldsymbol{A} = \boldsymbol{\Phi} * \boldsymbol{G} = [\boldsymbol{\phi}_1 \otimes \boldsymbol{g}_1, \boldsymbol{\phi}_2 \otimes \boldsymbol{g}_2, ..., \boldsymbol{\phi}_N \otimes \boldsymbol{g}_N] \quad (28)$$

where $*$ denotes the columnwise Kronecker product. As a result of (27), we have

$$\begin{aligned} \|\boldsymbol{a}_k\|_2^2 &= \boldsymbol{a}_k \cdot \overline{\boldsymbol{a}_k} = (\boldsymbol{\phi}_k \otimes \boldsymbol{g}_k) \cdot (\overline{\boldsymbol{\phi}_k \otimes \boldsymbol{g}_k}) \\ &= \|\boldsymbol{\phi}_k\|_2^2 \|\boldsymbol{g}_k\|_2^2. \end{aligned} \quad (29)$$

Furthermore,

$$\begin{aligned} \langle \boldsymbol{a}_p,\ \boldsymbol{a}_q \rangle &= \boldsymbol{a}_p \cdot \overline{\boldsymbol{a}_q} = (\boldsymbol{\phi}_p \otimes \boldsymbol{g}_p) \cdot (\overline{\boldsymbol{\phi}_q \otimes \boldsymbol{g}_q}) \\ &= (\boldsymbol{\phi}_p \cdot \overline{\boldsymbol{\phi}_q})(\boldsymbol{g}_p \cdot \overline{\boldsymbol{g}_q}) \\ &= \langle \boldsymbol{\phi}_p,\ \boldsymbol{\phi}_q \rangle \langle \boldsymbol{g}_p,\ \boldsymbol{g}_q \rangle. \end{aligned} \quad (30)$$

Let $r_{A_{p,q}}$, $r_{\phi_{p,q}}$ and $r_{g_{p,q}}$ be the normalized inner product of the p^{th} and q^{th} columns of $\boldsymbol{A}$, $\boldsymbol{\Phi}$ and $\boldsymbol{G}$, respectively. Combining (29) and (30), we obtain

$$r_{A_{p,q}} = r_{\phi_{p,q}} r_{g_{p,q}}. \quad (31)$$

Below we summarize two results on the coherence of the FDOT forward matrix.

Observation 1[1]*:*

The mutual coherence of $\boldsymbol{A}$ is upper bounded by the product of the mutual coherence of $\boldsymbol{G}$ and $\boldsymbol{\Phi}$.

$$\begin{aligned} \mathcal{M}(\boldsymbol{A}) &= \max_{p,q} |r_{A_{p,q}}| = \max_{p,q} |r_{\phi_{p,q}} r_{g_{p,q}}| = \max_{p,q} r_{\phi_{p,q}} r_{g_{p,q}} \\ &\leq (\max_{p,q} r_{\phi_{p,q}})(\max_{p,q} r_{g_{p,q}}) = \mathcal{M}(\boldsymbol{G})\mathcal{M}(\boldsymbol{\Phi}). \end{aligned} \quad (32)$$

□.

Observation 2:

$$\begin{aligned} \mathcal{M}(\boldsymbol{A}) &\leq \mathcal{M}_1(k, \boldsymbol{A}) \leq K(N, N_s, N_d, k)(\|\tilde{\boldsymbol{\Phi}}^T \tilde{\boldsymbol{\Phi}} - \boldsymbol{I}_N\|_F^2 \\ &+ \|\tilde{\boldsymbol{G}}^T \tilde{\boldsymbol{G}} - \boldsymbol{I}_N\|_F^2), \text{ for } k = 1, ..., N-1. \end{aligned} \quad (33)$$

where $\|\boldsymbol{E}\|_F = \sqrt{\sum_{i,j} e_{i,j}^2}$ denotes the Frobenius norm of the matrix quantity $\boldsymbol{E}$; $K(N, N_s, N_d, k)$, $i = 1, 2$ is a constant that depends on N, N_s,N_d and k; $\tilde{\boldsymbol{\Phi}}$ and $\tilde{\boldsymbol{G}}$ are $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ matrices whose columns are normalized to unity, and $\boldsymbol{I}_N$ denotes the N-by-N identity matrix.

□.

See Appendix A for the derivation of (33), and an explicit expression for K.

We summarize the implications of the results and observations made in this section in the following list of remarks:

Remarks -

- The normalized inner product $r_{A_{p,q}}$ of the p^{th} and q^{th} columns of the FDOT forward matrix is the product $r_{\phi_{p,q}}$ and $r_{g_{p,q}}$, the normalized inner product of the p^{th} and q^{th} columns of $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ matrices, i.e., $r_{A_{p,q}} = r_{\Phi_{p,q}} r_{G_{p,q}}$.
- To reduce the mutual coherence and cumulative coherence of the FDOT forward matrix $\boldsymbol{A}$, we wish to minimize $r_{A_{p,q}}$, which is equivalent to minimizing $r_{g_{p,q}}$ and $r_{\phi_{p,q}}$.
- The mutual coherence of the FDOT forward matrix is upper bounded by the product of the mutual coherence of $\boldsymbol{\Phi}$ and $\boldsymbol{G}$, i.e., $\mathcal{M}(\boldsymbol{A}) \leq \mathcal{M}(\boldsymbol{\Phi})\mathcal{M}(\boldsymbol{G})$.
- The mutual coherence and the cumulative coherence of $\boldsymbol{A}$ are bounded. Thus, reducing $\|\tilde{\boldsymbol{\Phi}}^T \tilde{\boldsymbol{\Phi}} - \boldsymbol{I}_N\|_F^2$ and $\|\tilde{\boldsymbol{G}}^T \tilde{\boldsymbol{G}} - \boldsymbol{I}_N\|_F^2$, reduces $r_{A_{p,q}}$, which in turn, reduces the cumulative coherence, $\mathcal{M}_1(k, \boldsymbol{A})$, and mutual coherence, $\mathcal{M}(\boldsymbol{A})$, of the FDOT forward matrix.

Without loss of generality for the rest of the paper, we assume that the columns of $\boldsymbol{G}$ and $\boldsymbol{\Phi}$ are normalized to unity and drop tilde from $\tilde{\boldsymbol{G}}$ and $\tilde{\boldsymbol{\Phi}}$ to simplify our notation for rest of the paper.

V. The FDOT Forward Matrix Optimization

In this section, we focus on the optimization of the FDOT forward matrix using the results in the previous section. In particular, we minimize $\|\boldsymbol{\Phi}^T \boldsymbol{\Phi} - \boldsymbol{I}_N\|_F$ and $\|\boldsymbol{G}^T \boldsymbol{G} - \boldsymbol{I}_N\|_F^2$ by designing two preconditioners on $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ to reduce both

[1]Note that since $\boldsymbol{\Phi}$ is composed of the excitation light field and $\boldsymbol{G}$ is composed of the discretized Green's function of the diffusion equation for the emission light field. Therefore, every entry in $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ is non-negative. As a result, $r_{\phi_{p,q}} \geq 0$ and $r_{g_{p,q}} \geq 0$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

the mutual coherence and average coherence of the FDOT forward matrix. Our preconditioner design follows a method similar to the one in [12].

The entries of the matrix $\mathbf{\Phi}$ represent the excitation field at each voxel location due to light sources. They are determined by the source function, the Green's function of (1) and (3) and the endogenous optical properties of the imaging domain. Thus, $\mathbf{\Phi}$ is a source related term. We refer to the preconditioner on $\mathbf{\Phi}$ as the *optical mask*. The optical mask is an $N_s \times N_s$ intensity matrix, with each row being a different intensity pattern applied to N_s point sources for N_s different illumination patterns.

The entries of the matrix $\boldsymbol{G}$ represent the emission field at the detectors due to a point source at each voxel location. They are determined by the Green's function of (2) and (4), and the endogenous optical properties of the imaging domain. Thus, $\boldsymbol{G}$ can be viewed as a measurement-related term. We refer to the preconditioner on $\boldsymbol{G}$ as the *measurement mask*. The measurement mask is an $N_d \times N_d$ linear filter applied to the measured data collected by N_d detectors corresponding to each illumination pattern.

In Subsection V-A we present the design of the optical mask; and in Subsection V-B, we present the design of measurement mask. In Subsection V-C, we describe the practical, algorithmic implementation of the FDOT imaging process with structured illumination and detection patterns.

A. Design of the Illumination Patterns - The Optical Mask

In diffuse optical tomography, the light sources are typically time multiplexed, i.e., operated one at a time, and the scattered light field due to each source is measured at the detectors. In designing the optical mask, we assume that all N_s sources transmit simultaneously into the imaging domain according to some intensity patterns designed to optimize the FDOT forward matrix. Furthermore, we assume that the number of illumination patterns is the same as the number of point sources. Thus, our objective is to determine $N_s \times N_s$ different light intensities to precondition $\mathbf{\Phi}$, and hence the FDOT forward matrix.

Let $\boldsymbol{M}_s$ denote the $N_s \times N_s$ optical mask. Let $\boldsymbol{m}_k$ denote each row of $\boldsymbol{M}_s$, i.e.,

$$\boldsymbol{M}_s = [\boldsymbol{m}_1{}^T, \boldsymbol{m}_2{}^T, \ldots \boldsymbol{m}_{N_s}{}^T]^T. \tag{34}$$

We refer to $\boldsymbol{m}_k$, $k = 1, ..., N_s$ as the k^{th} illumination pattern for reasons that will be made clear below. Recall that the excitation light field due to the i^{th} point source is given by

$$\phi_x^i(\boldsymbol{r}) = \int_\Omega g_x(\boldsymbol{r}, \boldsymbol{r}') s^i(\boldsymbol{r}') d\boldsymbol{r}'. \tag{35}$$

Without loss of generality, we assume that the initial light sources, $s^i(\boldsymbol{r})$, are unit impulse functions,

$$s^i(\boldsymbol{r}) = \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{36}$$

where $\boldsymbol{r}_i$ denotes the location of the i^{th} point source.

Let $m_{k,i}$ be the intensity of the i^{th} source, $i = 1, ..., N_s$, for the k^{th} illumination pattern, i.e.,

$$\boldsymbol{m}_k = [m_{k,1}, m_{k,2}, \ldots m_{k,N_s}]. \tag{37}$$

Then, if all point sources transmit simultaneously, the source illuminating the imaging domain can be expressed as

$$Q^k(\boldsymbol{r}) = \sum_{i=1}^{N_s} m_{k,i} s^i(\boldsymbol{r}). \tag{38}$$

Thus, the resulting excitation light field due to the k^{th} illumination pattern becomes

$$\begin{aligned} \phi_{pre}^k(\boldsymbol{r}) &= \int_\Omega g_x(\boldsymbol{r}, \boldsymbol{r}') Q^k(\boldsymbol{r}') d\boldsymbol{r}' \\ &= \sum_{i=1}^{N_s} m_{k,i} \int_\Omega g_x(\boldsymbol{r}, \boldsymbol{r}') s^i(\boldsymbol{r}') d\boldsymbol{r}' \\ &= \sum_{i=1}^{N_s} m_{k,i} \phi_x^i(\boldsymbol{r}). \end{aligned} \tag{39}$$

Reexpressing (39) in matrix notation introduced in (34) and (37), we write the new matrix composed of the excitation light field due to sources Q^k ($k = 1, ..., N_s$) as follows:

$$\mathbf{\Phi}_{pre} = \boldsymbol{M}_s \mathbf{\Phi}, \tag{40}$$

where $\mathbf{\Phi}_{pre}$ is now the preconditioned $\mathbf{\Phi}$ matrix. From (40), clearly the optical mask $\boldsymbol{M}_s$ can be designed to precondition $\mathbf{\Phi}$. We design $\boldsymbol{M}_s$ to minimize the normalized inner products between different columns of $\mathbf{\Phi}_{pre}$. Or equivalently, we seek to find $\boldsymbol{M}_s$ such that the Gramm matrix $\mathbf{\Phi}_{pre}{}^T \mathbf{\Phi}_{pre}$ is as close to identity as possible,

$$\mathbf{\Phi}_{pre}{}^T \mathbf{\Phi}_{pre} = \mathbf{\Phi}^T \boldsymbol{M}_s^T \boldsymbol{M}_s \mathbf{\Phi} \approx \boldsymbol{I}_N. \tag{41}$$

To find such a matrix $\boldsymbol{M}_s$, we consider the following optimization problem:

$$\min_{\boldsymbol{M}_s} \|(\boldsymbol{M}_s \mathbf{\Phi})^T \boldsymbol{M}_s \mathbf{\Phi} - \boldsymbol{I}_N\|_F. \tag{42}$$

Note that (42) attempts to minimize not only the greatest normalized inner product between two columns of $\mathbf{\Phi}_{pre}$, but all the normalized inner products between different columns of $\mathbf{\Phi}_{pre}$.

To solve (42), we multiply (41) by $\mathbf{\Phi}$ and $\mathbf{\Phi}^T$ on both sides,

$$\mathbf{\Phi}\mathbf{\Phi}^T \boldsymbol{M}_s^T \boldsymbol{M}_s \mathbf{\Phi}\mathbf{\Phi}^T \approx \mathbf{\Phi}\mathbf{\Phi}^T, \tag{43}$$

and consider the singular value decomposition of $\mathbf{\Phi}$,

$$\mathbf{\Phi} = \boldsymbol{U}_s \mathbf{\Sigma}_s \boldsymbol{V}_s, \tag{44}$$

where $\mathbf{\Sigma}_s$ is the $N_s \times N_s$ diagonal matrix containing the singular values of $\mathbf{\Phi}$. Let $\mathbf{\Lambda}_s = \mathbf{\Sigma}_s \mathbf{\Sigma}_s^T$. Substituting (44) into (43), and simplifying (43), we obtain

$$\mathbf{\Lambda}_s \boldsymbol{U}_s^T \boldsymbol{M}_s^T \boldsymbol{M}_s \boldsymbol{U}_s \mathbf{\Lambda}_s \approx \mathbf{\Lambda}_s. \tag{45}$$

Thus, we choose $\boldsymbol{M}_s$ as

$$\boldsymbol{M}_s = \mathbf{\Lambda}_s^{-1/2} \boldsymbol{U}_s^T. \tag{46}$$

Note that if $\mathbf{\Phi}$ has poor condition number, $\boldsymbol{M}_s$ can be stabilized by $(\mathbf{\Lambda}_s + \varepsilon \boldsymbol{I})^{-1/2} \boldsymbol{U}_s^T$ where $\boldsymbol{I}$ is the identity matrix and $\varepsilon > 0$ is a suitably chosen regularization parameter.

The new $\mathbf{\Phi}$ matrix after the application of $\boldsymbol{M}_s$ becomes

$$\mathbf{\Phi}_{pre} = \boldsymbol{M}_s \mathbf{\Phi} = \mathbf{\Lambda}_s^{-1/2} \boldsymbol{U}_s^T \boldsymbol{U}_s \mathbf{\Sigma}_s \boldsymbol{V}_s = \mathbf{\Lambda}_s^{-1/2} \mathbf{\Sigma}_s \boldsymbol{V}_s. \tag{47}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

The matrix $\boldsymbol{M}_s$ may have negative entries. Thus directly applying $\boldsymbol{M}_s$ onto source intensities is not impossible. We decompose $\boldsymbol{M}_s$ into a linear combination of two matrices with non-negative entries as follows:

$$\boldsymbol{M}_s = \boldsymbol{M}_s^{(+)} - (-\boldsymbol{M}_s^{(-)}), \tag{48}$$

where the entries of $\boldsymbol{M}_s^{(+)}$ and $\boldsymbol{M}_s^{(-)}$ are either zero or equal to the non-negative and negative entries of $\boldsymbol{M}_s$, respectively. Thus, the entries of both $\boldsymbol{M}_s^{(+)}$ and $-\boldsymbol{M}_s^{(-)}$ are non-negative. We then sequentially apply $\boldsymbol{M}_s^{(+)}$ and $(-\boldsymbol{M}_s^{(-)})$ onto the source intensities. Let $\boldsymbol{\Gamma}^{(+)}$ and $\boldsymbol{\Gamma}^{(-)}$ denote the resultant measurements, we take the difference of the measurement $\boldsymbol{\Gamma}^{(+)}$ and $\boldsymbol{\Gamma}^{(-)}$, and obtain

$$\begin{aligned} \boldsymbol{\Gamma} &= \left((\boldsymbol{M}_s^{(+)}\boldsymbol{\Phi}) * \boldsymbol{G}_{pre}\right)\boldsymbol{x} - \left((-\boldsymbol{M}_s^{(-)}\boldsymbol{\Phi}) * \boldsymbol{G}_{pre}\right)\boldsymbol{x} \\ &= (\boldsymbol{\Phi}_{pre} * \boldsymbol{G}_{pre})\boldsymbol{x} = \boldsymbol{A}_{pre}\boldsymbol{x}. \end{aligned} \tag{49}$$

Let $m_{k,i}^{(\pm)}$ denote the $(k,i)^{th}$ entry of $\boldsymbol{M}_s^{(\pm)}$, which is the intensity of the i^{th} point source in the k^{th} illumination pattern. Thus, the light source generated by the k^{th} illumination pattern is given by

$$Q^k(\boldsymbol{r}) = \sum_{i=1}^{N_s}(m_{k,i}^{(+)} - (-m_{k,i}^{(-)}))s^i(\boldsymbol{r}), \tag{50}$$

and the resulting excitation light field due to the k^{th} illumination pattern becomes

$$\begin{aligned} \phi_{pre}^{(k)}(\boldsymbol{r}) &= \sum_{i=1}^{N_s} m_{k,i}^{(+)}\phi_x^i(\boldsymbol{r}) - \sum_{i=1}^{N_s}(-m_{k,i}^{(-)})\phi_x^i(\boldsymbol{r}), \\ & \qquad k = 1, ..., N \quad \boldsymbol{r} \in \Omega. \end{aligned} \tag{51}$$

B. Design of Detection Patterns - The Measurement Mask

In this section, our objective is to design a measurement mask, or a linear filter, that will be applied to the measured data to reduce the normalized inner product between the columns of $\boldsymbol{G}$, i.e., $r_{g_{p,q}}$.

Recall from Section 2 that the measurements $\boldsymbol{\Gamma}^i$ due to the i^{th} light source is given as $\boldsymbol{\Gamma}^i = \boldsymbol{B}^i\boldsymbol{x}$ where

$$\boldsymbol{B}^i = [\boldsymbol{g}_1\phi_{ex,1}^i, \boldsymbol{g}_2\phi_{ex,2}^i, ..., \boldsymbol{g}_N\phi_{ex,N}^i].$$

If a linear transform is applied to the measurements, the new set of measurements becomes:

$$\boldsymbol{\Gamma}_{pre}^i = \boldsymbol{M}_d\boldsymbol{\Gamma}^i = \boldsymbol{M}_d\boldsymbol{B}^i\boldsymbol{x} \tag{52}$$

where $\boldsymbol{M}_d \in \mathbb{R}^{N_d \times N_d}$ is the linear (not necessarily Toeplitz) filter, which we refer to as the measurement mask. We note that

$$\boldsymbol{M}_d\boldsymbol{B}^i = [\boldsymbol{M}_d\boldsymbol{g}_1\phi_{ex,1}^i, \boldsymbol{M}_d\boldsymbol{g}_2\phi_{ex,2}^i, ..., \boldsymbol{M}_d\boldsymbol{g}_N\phi_{ex,N}^i]. \tag{53}$$

Let

$$\boldsymbol{G}_{pre} = \boldsymbol{M}_d\boldsymbol{G} \tag{54}$$

be the preconditioned $\boldsymbol{G}$ matrix. We choose to design $\boldsymbol{M}_d$ so that the normalized inner products between different columns of $\boldsymbol{G}$ are minimized.

Designing the measurement mask, $\boldsymbol{M}_d$, is similar to designing the optical mask. Following a step similar to the one in (43), we obtain

$$\boldsymbol{G}\boldsymbol{G}^T\boldsymbol{M}_d^T\boldsymbol{M}_d\boldsymbol{G}\boldsymbol{G}^T \approx \boldsymbol{G}\boldsymbol{G}^T. \tag{55}$$

Let $\boldsymbol{G} = \boldsymbol{U}_d\boldsymbol{\Sigma}_d\boldsymbol{V}_d^T$ be the singular value decomposition of $\boldsymbol{G}$. We choose

$$\boldsymbol{M}_d = \boldsymbol{\Lambda}_d^{-1/2}\boldsymbol{U}_d^T \tag{56}$$

where $\boldsymbol{\Lambda}_d = \boldsymbol{\Sigma}_d\boldsymbol{\Sigma}_d^T$. Again, if $\boldsymbol{G}$ has poor condition number, $\boldsymbol{M}_d$ can be modified to $(\boldsymbol{\Lambda}_d + \varepsilon\boldsymbol{I})^{-1/2}\boldsymbol{U}_d^T$ where $\boldsymbol{I}$ is the identity matrix and $\varepsilon > 0$ is a suitably chosen regularization parameter.

The preconditioned $\boldsymbol{G}$ matrix becomes

$$\boldsymbol{G}_{pre} = \boldsymbol{M}_d\boldsymbol{G} = \boldsymbol{\Lambda}_d^{-1/2}\boldsymbol{U}_d^T\boldsymbol{U}_d\boldsymbol{\Sigma}_d\boldsymbol{V}_d = \boldsymbol{\Lambda}_d^{-1/2}\boldsymbol{\Sigma}_d\boldsymbol{V}_d. \tag{57}$$

An alternative approach in designing a measurement mask is to directly precondition the forward sensing matrix $\boldsymbol{A}$ [37]. In this case, the measurement mask becomes

$$\boldsymbol{M}_A = (\boldsymbol{\Sigma}_A\boldsymbol{\Sigma}_A^T)^{-1/2}\boldsymbol{U}_A^T \tag{58}$$

where $\boldsymbol{U}_A\boldsymbol{\Sigma}_A\boldsymbol{V}_A^T$ is the singular value decomposition of $\boldsymbol{A}$.

In some applications, depending on the source-detector geometry, number of sources, as well as the discretization of the diffusion equations and the imaging domain, it is favorable to precondition $\boldsymbol{G}\boldsymbol{G}^T$ over preconditioning $\boldsymbol{A}\boldsymbol{A}^T$ to avoid noise amplification due to large condition number. It is straightforward to show that $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{\Phi}^T\boldsymbol{\Phi} \circ \boldsymbol{G}^T\boldsymbol{G}$ where $\circ$ denotes the Hadamard product. As a result, $\|\boldsymbol{A}\boldsymbol{A}^T\| \leq \max_{n=1,...,N}[\boldsymbol{\Phi}^T\boldsymbol{\Phi}]_{nn}\|\boldsymbol{G}\boldsymbol{G}^T\|$ where $[\boldsymbol{\Phi}^T\boldsymbol{\Phi}]_{nn}$ denotes the nth diagonal entry of $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$ and $\|\cdot\|$ stands for any unitarily invariant matrix norm [38]. Hence, if $\max_{n=1,...,N}[\boldsymbol{\Phi}^T\boldsymbol{\Phi}]_{nn} \leq 1$, it may be favorable to precondition $\boldsymbol{G}\boldsymbol{G}^T$ over preconditioning $\boldsymbol{A}\boldsymbol{A}^T$. However, $\max_{n=1,...,N}[\boldsymbol{\Phi}^T\boldsymbol{\Phi}]_{nn}$, that is, the maximum power of the excitation photon density field due to all point sources among all voxels depends on the source-detector geometry, number of sources and the discretization of the diffusion equations and imaging domain.

C. Forward Sensing Matrix Construction

The new forward sensing matrix $\boldsymbol{A}_{pre}$ composed of the columnwise Kronocker product of $\boldsymbol{\Phi}_{pre}$ and $\boldsymbol{G}_{pre}$,

$$\boldsymbol{A}_{pre} = [\boldsymbol{\phi}_{pre,1} \otimes \boldsymbol{g}_{pre,1}, ..., \boldsymbol{\phi}_{pre,N} \otimes \boldsymbol{g}_{pre,N}] \tag{59}$$

where $\boldsymbol{\phi}_{pre,i}$ and $\boldsymbol{g}_{pre,i}$, $i = 1, ..., N$ denote the columns of $\boldsymbol{\Phi}_{pre}$ and $\boldsymbol{G}_{pre}$.

VI. Numerical Simulations

In this section, we extensively present numerical simulations to demonstrate the effects of the optical mask $\boldsymbol{M}_s$ and measurement mask $\boldsymbol{M}_d$ on the coherence of the forward sensing matrix as well as on the quality of the reconstructed FDOT images.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

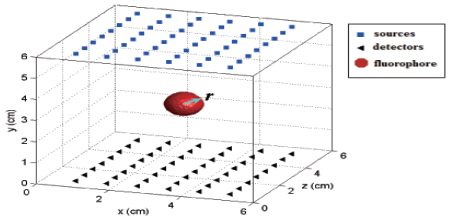


Fig. 1. An illustration of the simulated phantom and source-detector configuration.

A. 3D-Simulation with Single Heterogeneity

We set up a $6 \times 6 \times 6$ cm^3 cubic phantom shown in Fig.1. We set the background absorption coefficient to $\mu_{am} = 0.05\text{cm}^{-1}$, and the diffusion coefficient to $\mu_s' = 8\text{cm}^{-1}$ ($D = 1/3(\mu_a + \mu_s')$) at both the excitation and emission wavelengths. The circular heterogeneity with radius r simulated the fluorophore concentration with fluorophore absorption coefficient $\mu_{axf} = 0.005\text{cm}^{-1}$. 48 sources (shown in squares) and 48 detectors (shown in triangles) are uniformly placed at the top and bottom of the imaging domain. We discretized the imaging domain into $20 \times 20 \times 20$ voxels. Thus, the forward sensing matrix is 2304 by 8000. We simulated data sets with a single heterogeneity corresponding to 3 different values of the radius r : 0.5, 0.75, and 1.0cm.

To simulate the measurement noise, we assume that, when a sufficiently large number of photons are detected, the noise at each detector can be approximated by a Gaussian random variable with its variance proportional to the magnitude of the detector reading. We define the *signal-to-noise-ratio* (SNR) of the measurements as

$$SNR = 20 \log_{10} \frac{\|\mathbf{\Gamma}\|_2}{\|\boldsymbol{\epsilon}\|_2}. \tag{60}$$

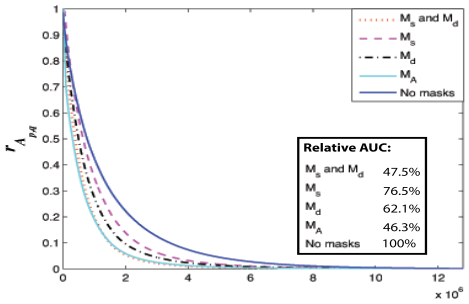
We simulated three sets of noise contaminated measurements with approximately 10%, 3% and 1% noise, corresponding to the SNR value of 20dB, 30dB and 40dB, respectively.

1) Simulation Results - Coherence of the FDOT Forward Matrix: We evaluated the coherence of the forward matrix for five different cases: (1) no masks corresponding to the time-multiplexed point source configuration with unit intensity; (2) only optical mask $\boldsymbol{M}_s$; (3) only measurement mask $\boldsymbol{M}_d$; (4) both optical and measurement masks; and (5) large preconditioning matrix $\boldsymbol{M}_A$. Fig. 2(a) shows the largest 40% of all the normalized inner products between different columns of the forward sensing matrix, arranged in a descending order. The remaining 60% of the normalized inner products between different columns are close to 0. Clearly, application of $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ and $\boldsymbol{M}_A$ all reduces the large correlations between different columns of $\boldsymbol{A}$. When $\boldsymbol{A}$ is preconditioned using $\boldsymbol{M}_A$, the resultant normalized inner product plot is very close to using both $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$. To quantify the improvements, we computed the *area under the curves* (AUC), which we defined as the sum of all the normalized inner products between different pairs of columns:

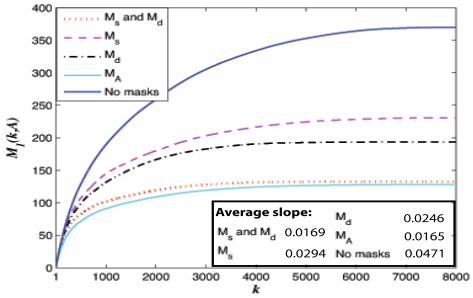
$$\text{AUC} = \sum_{p,1=1(p\neq q)}^{N} |r_{A_{p,q}}|. \tag{61}$$

We define the relative AUC as the ratio of the AUC to that of the no-mask case. The results of relative AUC are given in the box in Fig. 2(a). We see that the optical and measurement masks alone reduce the AUC to 76.5% and 62.1%, respectively as compared to the no-mask case. When both masks are applied, the AUC reduces to 47.5%. The application of $\boldsymbol{M}_A$ reduces AUC to 46.3%, which is slightly smaller than the case that both $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$ are applied.

Fig. 2(b) shows the plot of the cumulative coherence $\mathcal{M}_1(k, \boldsymbol{A})$ as a function of k. As stated before, the cumulative coherence is a nondecreasing function of k. When $\mathcal{M}_1(k, \boldsymbol{A})$ increases slowly, the forward matrix is said to be quasi-incoherent [10]. From Fig. 2(b), the application of $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ and $\boldsymbol{M}_A$ all result in slower increase of cumulative coherence. The average slope of each curve is provided in the box in Fig. 2(b) for a quantitative comparison.



(a) Top 40% values of all the normalized inner products



(b) Cumulative coherence

Fig. 2. The normalized inner products and cumulative coherence of the forward matrix before and after applying optical and measurement masks in numerical simulations.

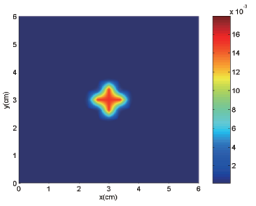
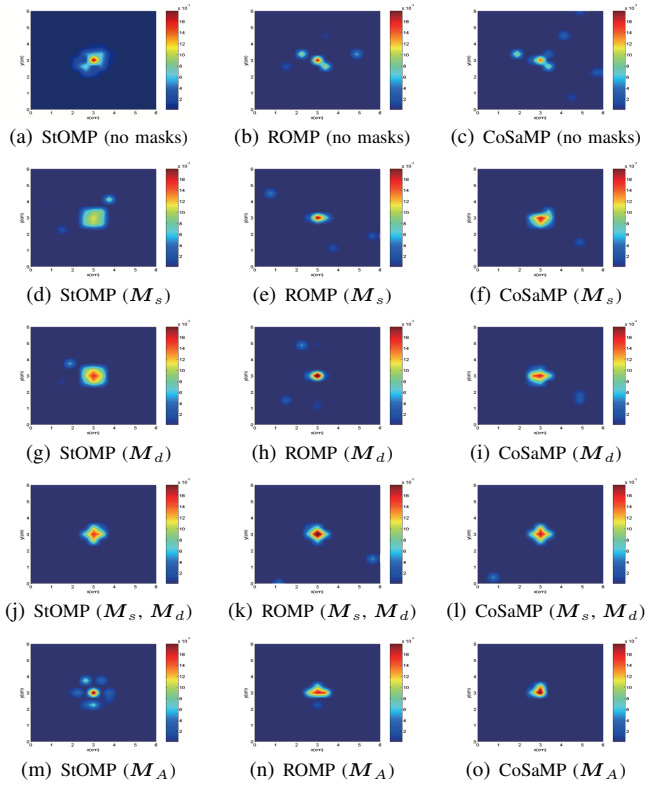


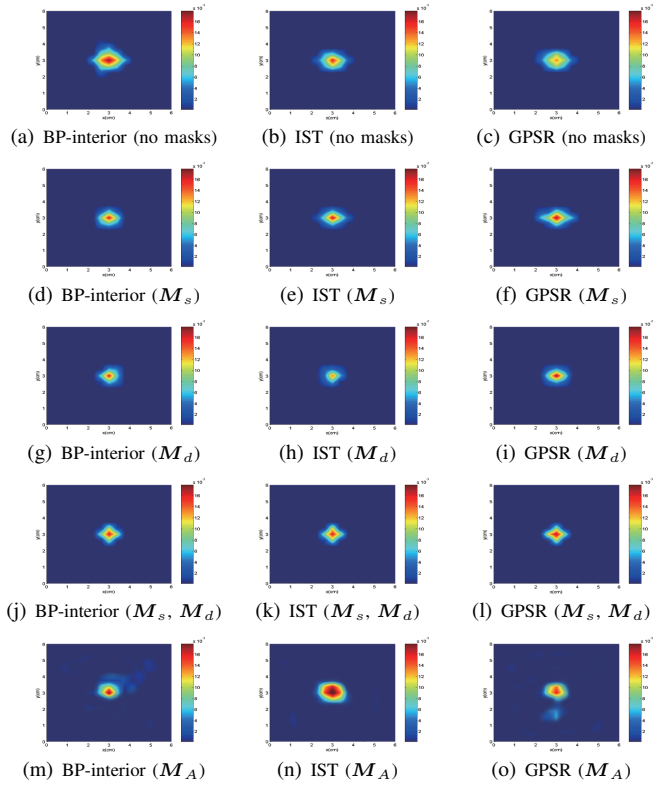
Fig. 3. The cross sections at $z = 3$cm of the simulated phantoms.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing



(a) StOMP (no masks) (b) ROMP (no masks) (c) CoSaMP (no masks)
(d) StOMP ($\boldsymbol{M}_s$) (e) ROMP ($\boldsymbol{M}_s$) (f) CoSaMP ($\boldsymbol{M}_s$)
(g) StOMP ($\boldsymbol{M}_d$) (h) ROMP ($\boldsymbol{M}_d$) (i) CoSaMP ($\boldsymbol{M}_d$)
(j) StOMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) ROMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) CoSaMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(m) StOMP ($\boldsymbol{M}_A$) (n) ROMP ($\boldsymbol{M}_A$) (o) CoSaMP ($\boldsymbol{M}_A$)

Fig. 4. The cross sections of the reconstructed images at $z = 3$cm of the phantom using greedy algorithms in numerical simulations, centered at $(3, 3, 3)$cm, $r = 0.5$cm, 1% noise.



(a) BP-interior (no masks) (b) IST (no masks) (c) GPSR (no masks)
(d) BP-interior ($\boldsymbol{M}_s$) (e) IST ($\boldsymbol{M}_s$) (f) GPSR ($\boldsymbol{M}_s$)
(g) BP-interior ($\boldsymbol{M}_d$) (h) IST ($\boldsymbol{M}_d$) (i) GPSR ($\boldsymbol{M}_d$)
(j) BP-interior ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) IST ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) GPSR ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(m) BP-interior ($\boldsymbol{M}_A$) (n) IST ($\boldsymbol{M}_A$) (o) GPSR ($\boldsymbol{M}_A$)

Fig. 5. The cross sections of the reconstructed images at $z = 3$cm of the phantom using convex relaxation algorithms in numerical simulations, centered at $(3, 3, 3)$cm, $r = 0.5$cm, 1% noise.

2) Simulation Results - Image Reconstruction: To reconstruct the fluorophore concentration map, we used the forward matrix without applying any masks, with either $\boldsymbol{M}_s$ or $\boldsymbol{M}_d$ alone, with both masks applied simultaneously, and with $\boldsymbol{M}_A$ to precondition $\boldsymbol{A}$ directly. We used six different sparsity promoting reconstruction methods available in the CS literature. Specifically, we used: *stagewise orthogonal matching pursuit* (StOMP) [39], *regularized orthogonal matching pursuit* (ROMP) [40], *acrostic compressive sampling matching pursuit* (CoSaMP) [34], BP-interior [35], *iterative shrinkage/thresholding* (IST) [41], and *gradient projection for sparse reconstruction* (GPSR) [42]. The first three are greedy type algorithms, and the last three are convex relaxation based algorithms.

We reconstructed the fluorophore concentration map of a phantom with a single heterogeneity at 3 different sparsity levels corresponding to 3 different radii using measurements at 3 different noise levels. We measure the accuracy of the reconstructed images using the normalized mean square error (NMSE) given by

$$\text{NMSE} = \frac{\|\boldsymbol{x}_{\text{true}} - \boldsymbol{x}_{\text{recon}}\|_2}{\text{\# of voxels}}. \tag{62}$$

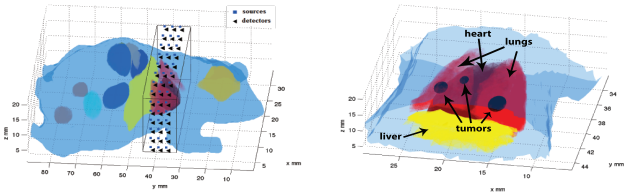
We averaged the NMSE of the reconstructed images over 50 realizations of noise. The results are tabulated in Table I. The results show that, in general, the NMSE of the reconstructed images increases as the sparsity level of the signal increases. By applying $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ or $\boldsymbol{M}_A$, the NMSE can be reduced for all reconstructions. However, applying both $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$ is more effective in reducing the noise than applying $\boldsymbol{M}_A$, since the large condition numbers in $\boldsymbol{M}_A$ is likely to amplify the measurement noise. Table I shows a greater reduction in NMSE for greedy-type algorithms than that of convex relaxation techniques.

Fig. 3 shows the cross sections of the reconstructed images at the middle of the imaging domain when the radius of the heterogeneity is $r = 0.5$cm. The cross sections of the reconstructed fluorophore yield maps using greedy type algorithms are shown in Fig.4, and those using convex relaxation techniques are shown in Fig.5. For each type of algorithm, the application of $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ or $\boldsymbol{M}_A$ results in reconstructed images that are in better agreement with the original fluorophore yield map. Although applying $\boldsymbol{M}_A$ is effective in improving the incoherence of $\boldsymbol{A}$, as indicated in Fig. 2, it is less effective in improve the image reconstruction results. The visual improvements are most obvious for the greedy algorithms, consistent with the NMSE results. In the greedy type algorithms, the support of the signal is determined by selecting the columns of the forward matrix that have the greatest correlation with the measurements. The reduction of the normalized inner products between different columns in the forward matrix has a direct effect on the column selection procedure at each iteration.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

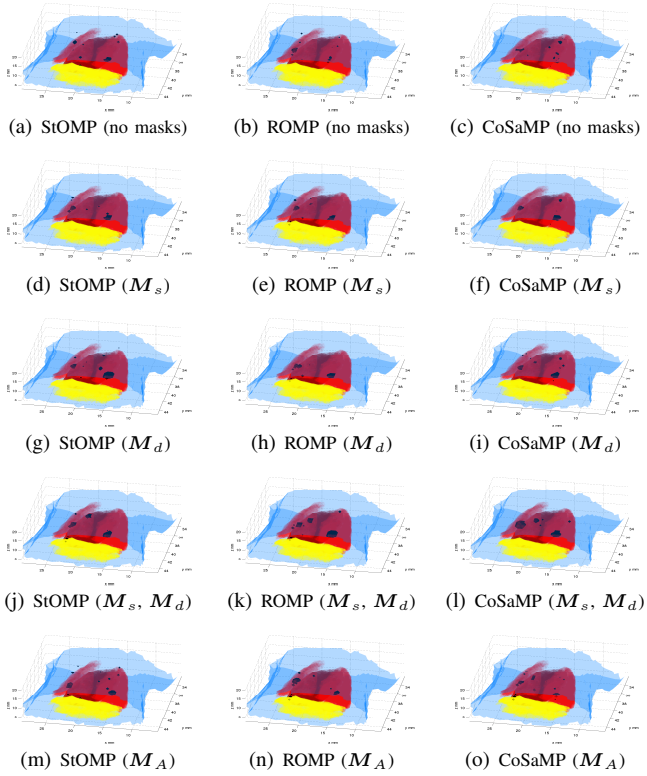
TABLE I
NMSE OF THE RECONSTRUCTED FLUOROPHORE CONCENTRATION MAP IN NUMERICAL SIMULATIONS (10^{-5}).

Algorithm		$r = 0.5$cm			$r = 0.75$cm			$r = 1.0$cm		
StOMP	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	2.05	6.47	12.1	3.87	14.5	28.0	9.44	31.8	63.7
	$\boldsymbol{M}_s$	0.98	2.05	7.04	1.90	6.47	13.9	5.00	13.2	33.3
	$\boldsymbol{M}_d$	0.90	1.95	6.41	1.76	5.77	9.26	4.37	11.7	33.0
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.55	1.31	2.58	0.93	2.63	5.55	2.12	7.67	16.4
	$\boldsymbol{M}_A$	1.40	4.20	8.51	2.74	9.65	20.6	6.94	22.3	45.2
ROMP	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	1.66	5.85	13.2	3.68	13.0	31.6	8.01	29.0	64.2
	$\boldsymbol{M}_s$	0.68	2.45	6.97	2.19	6.39	19.8	3.16	15.4	38.1
	$\boldsymbol{M}_d$	0.54	2.54	6.62	1.89	5.19	16.9	3.71	13.9	35.3
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.35	1.55	2.49	0.80	2.58	5.82	1.76	8.40	19.9
	$\boldsymbol{M}_A$	1.26	3.76	8.13	2.09	11.8	16.4	5.77	22.4	38.5
CoSaMP	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	2.11	6.35	13.0	5.97	14.5	28.9	11.8	34.7	67.8
	$\boldsymbol{M}_s$	1.69	2.11	5.98	3.69	8.96	13.4	4.29	16.1	35.6
	$\boldsymbol{M}_d$	1.13	2.08	6.30	3.08	7.37	11.9	5.35	15.9	32.1
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.61	1.37	3.45	1.31	2.77	6.08	2.66	8.06	18.5
	$\boldsymbol{M}_A$	1.24	3.28	8.18	3.78	9.73	22.0	6.26	23.7	49.1
BP-interior	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	1.56	3.03	7.05	2.44	7.13	15.3	5.03	20.8	39.1
	$\boldsymbol{M}_s$	1.18	2.31	4.85	1.97	5.07	10.2	4.02	10.8	23.4
	$\boldsymbol{M}_d$	0.93	2.03	4.50	2.08	4.37	9.93	3.68	8.91	20.1
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.39	1.28	1.98	1.09	2.06	3.85	1.84	5.18	9.11
	$\boldsymbol{M}_A$	1.04	2.07	5.96	1.87	6.09	13.2	4.31	18.7	37.2
IST	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	1.20	3.55	7.49	3.51	6.97	19.2	6.38	23.3	49.9
	$\boldsymbol{M}_s$	0.89	2.05	3.91	2.31	5.32	11.8	5.06	11.7	26.9
	$\boldsymbol{M}_d$	0.92	1.79	3.61	1.93	4.19	10.0	4.41	9.95	24.3
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.46	1.19	2.38	0.81	1.86	4.36	2.03	6.28	12.5
	$\boldsymbol{M}_A$	1.07	2.98	4.74	2.43	5.50	11.8	5.80	20.7	45.1
GPSR	SNR(dB)	40	30	20	40	30	20	40	30	20
	No masks	1.29	3.28	7.07	3.41	7.13	17.1	5.51	22.2	46.6
	$\boldsymbol{M}_s$	0.97	2.27	4.08	2.03	5.01	13.5	4.92	11.3	29.0
	$\boldsymbol{M}_d$	0.82	1.76	3.26	1.50	3.76	10.0	4.83	8.74	24.8
	$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	0.40	1.23	2.56	0.88	1.68	3.82	2.15	5.45	10.2
	$\boldsymbol{M}_A$	1.06	2.24	6.14	2.10	5.67	13.5	4.40	16.4	39.4



(a) The digital mouse phantom with the source-detector configuration (b) The chest region of the digital mouse

Fig. 6. An illustration of the digital mouse phantom.



(a) StOMP (no masks) (b) ROMP (no masks) (c) CoSaMP (no masks)
(d) StOMP ($\boldsymbol{M}_s$) (e) ROMP ($\boldsymbol{M}_s$) (f) CoSaMP ($\boldsymbol{M}_s$)
(g) StOMP ($\boldsymbol{M}_d$) (h) ROMP ($\boldsymbol{M}_d$) (i) CoSaMP ($\boldsymbol{M}_d$)
(j) StOMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) ROMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) CoSaMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(m) StOMP ($\boldsymbol{M}_A$) (n) ROMP ($\boldsymbol{M}_A$) (o) CoSaMP ($\boldsymbol{M}_A$)

Fig. 7. The reconstructed digital mouse phantom images using greedy algorithms (1% noise, isosurfaces plotted at 50% of the maximum values).

B. 3D-Simulations using the Digital Mouse Phantom [43], [44]

In this subsection, we present numerical simulation results using the digital mouse phantom reported in [43], [44]. The digital mouse phantom was generated using coregistered CT and cryosection images and has segmented internal tissue structures. We assume that the tumors are located in the lungs, and consider a 24mm×13mm×18mm rectangular imaging domain around the chest of the mouse phantom. Three tumors of diameter 1.0mm, 1.5mm, and 2.0mm are located in the lungs with fluorophore absorption coefficient $\mu_{af} = 0.05\text{cm}^{-1}$, as shown in Fig. 6(b). 84 point sources and 84 detectors that are uniformly placed on the four surfaces of the rectangular region as shown in Fig. 6(a). The optical properties of different organ structures are tabulated in Table II, which are estimated at wavelength 750nm [45]. We assume that the rectangular imaging domain is filled with matching fluid having the same optical properties as the mouse skin. The imaging domain is discretized into $33 \times 19 \times 26$ voxels using finite element method. 1%, 3% and 10% Gaussian noise are added to the measurements.

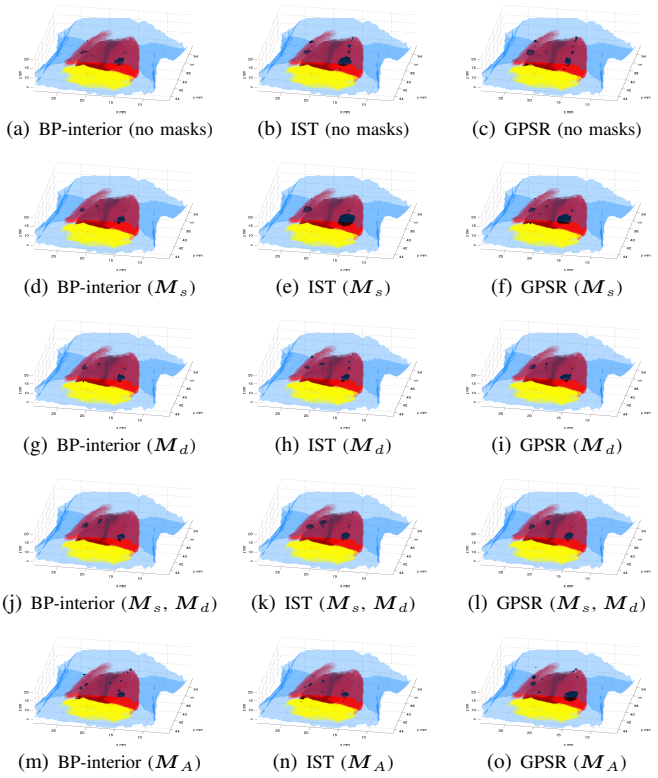
TABLE II
OPTICAL PROPERTIES OF THE MOUSE CHEST AT 750NM (CM^{-1}) [45]

Tissue type	Skin	muscle	bone	Lung	Heart	Liver
μ_s'	19.60	3.122	21.14	20.48	8.203	6.023
μ_a	0.397	0.478	0.329	1.054	0.333	1.936

The NMSE of the reconstructed images are shown in Table III, which are averaged over 30 realizations of noise. The results show that applying the $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ or $\boldsymbol{M}_A$ reduces NMSE when used in conjunction with different sparse signal recovery algorithms. However, applying $\boldsymbol{M}_A$ is not effective in reducing NMSE as compared to using both $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$, which is due to the large condition number of $\boldsymbol{M}_A$. The reconstructed images using greedy type algorithms are shown in Fig. 7, and those using convex relaxation techniques are shown in Fig. 8.

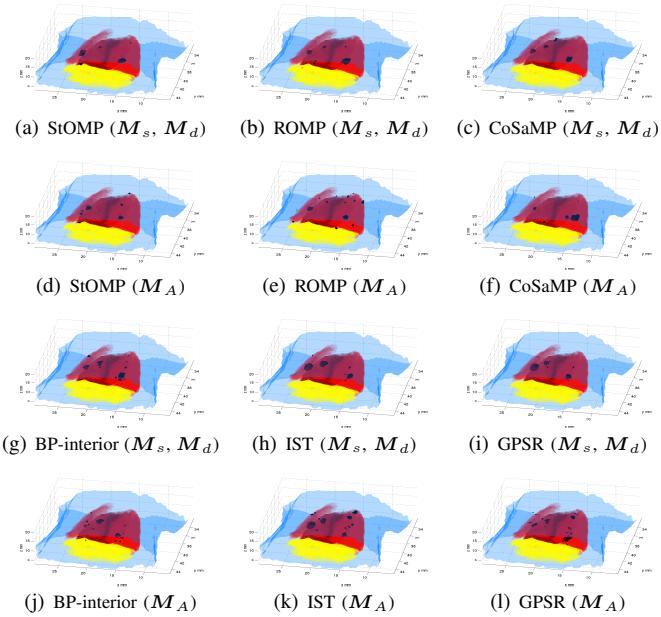
To check the robustness of our method with respect to errors in the endogenous absorption coefficient estimation, we assumed that the endogenous absorption coefficient estimation μ_{am} deviates from the true value by 1% and built $\boldsymbol{G}$, $\boldsymbol{\Phi}$ and $\boldsymbol{A}$ matrices. Fig. 9 shows the reconstruction results when erroneous endogenous absorption coefficients were used along with the application of optical and measurement masks, and with preconditioning matrix $\boldsymbol{A}$. Visual inspection of the images shows that the reconstruction performance

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing



(a) BP-interior (no masks) (b) IST (no masks) (c) GPSR (no masks)
(d) BP-interior ($\boldsymbol{M}_s$) (e) IST ($\boldsymbol{M}_s$) (f) GPSR ($\boldsymbol{M}_s$)
(g) BP-interior ($\boldsymbol{M}_d$) (h) IST ($\boldsymbol{M}_d$) (i) GPSR ($\boldsymbol{M}_d$)
(j) BP-interior ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) IST ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) GPSR ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(m) BP-interior ($\boldsymbol{M}_A$) (n) IST ($\boldsymbol{M}_A$) (o) GPSR ($\boldsymbol{M}_A$)

Fig. 8. The reconstructed digital mouse phantom images using convex relaxation algorithms (1% noise, isosurfaces plotted at 50% of the maximum values).



(a) StOMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (b) ROMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (c) CoSaMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(d) StOMP ($\boldsymbol{M}_A$) (e) ROMP ($\boldsymbol{M}_A$) (f) CoSaMP ($\boldsymbol{M}_A$)
(g) BP-interior ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (h) IST ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (i) GPSR ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)
(j) BP-interior ($\boldsymbol{M}_A$) (k) IST ($\boldsymbol{M}_A$) (l) GPSR ($\boldsymbol{M}_A$)

Fig. 9. The reconstructed digital mouse phantom images with 1% deviation of μ_{am} (1% noise, isosurfaces plotted at 50% of the maximum values).

decreases. Nevertheless, the tumors are still identifiable and the reconstructed images are visually better than the images reconstructed without the masks shown in Fig. 7 and Fig. 8.

TABLE III
NMSE OF THE RECONSTRUCTED FLUOROPHORE CONCENTRATION OF THE DIGITAL MOUSE PHANTOM (10^{-4}).

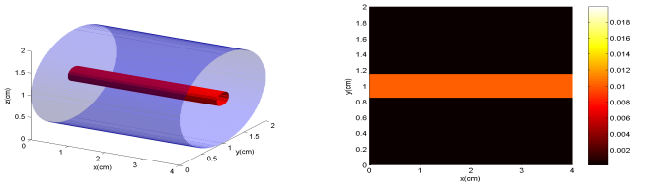
Algorithm	StOMP			ROMP			CoSaMP		
SNR(dB)	40	30	20	40	30	20	40	30	20
No masks	9.95	18.7	57.9	7.32	24.9	70.2	11.4	22.9	81.6
$\boldsymbol{M}_s$	2.65	15.5	36.5	3.74	17.8	37.4	5.50	17.3	41.0
$\boldsymbol{M}_d$	2.87	13.2	28.8	3.13	19.5	36.2	4.41	15.0	45.1
$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	1.33	5.72	17.6	1.99	8.29	20.8	2.32	9.57	18.9
$\boldsymbol{M}_A$	5.88	13.9	30.8	5.23	20.1	35.9	8.17	19.3	39.8
Algorithm	BP-interior			IST			GPSR		
SNR(dB)	40	30	20	40	30	20	40	30	20
No masks	15.4	21.6	35.5	10.3	31.2	44.0	18.7	27.9	39.2
$\boldsymbol{M}_s$	7.18	10.2	28.9	7.86	21.9	23.8	10.6	17.8	29.0
$\boldsymbol{M}_d$	7.96	12.3	27.3	7.35	22.1	29.1	13.9	16.8	27.2
$\boldsymbol{M}_s$, $\boldsymbol{M}_d$	4.01	6.92	10.1	5.27	7.30	13.2	4.07	8.68	13.3
$\boldsymbol{M}_A$	7.07	17.3	26.2	8.01	18.6	37.2	15.2	22.9	33.7

VII. Silicon Phantom Experiment

In this section, we present the effect of using $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$ on the coherence of the forward sensing matrix and the reconstructed images using real data collected in a silicon phantom experiment. The effect of preconditioning the $\boldsymbol{A}$ as described in subsection V-B on the coherence of the forward matrix and the reconstructed images using real data is presented in [37].

A. Real Phantom Configuration

Fig.10(a) shows an illustration of the cylindrical phantom used in the experiment. The phantom was made of silicone rubber with diameter of about 2cm, and length of 4cm. The phantom had homogeneous absorption coefficient $\mu_a = 0.2\text{cm}^{-1}$ and scattering coefficient $\mu_s' = 12\text{cm}^{-1}$ at both the excitation and emission wavelengths (743nm and 767nm). The silicon phantom contained a hollow cylindrical tube in the middle with approximately 3mm in diameter. which was filled with intralipid and ink to mimic the same optical properties as the background. The intralipid and ink contained 1 micromolar of Cy7 as the fluorophore. The cross section of the fluorophore yield at $z = 1$cm is shown in Fig. 10(b).



(a) The silicon phantom (b) The cross section of the fluorophore yield at $z = 1$cm (middle) of the silicon phantom

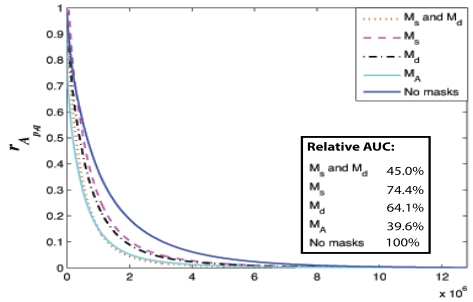
Fig. 10. The configuration of the real silicon phantom and the cross section of the fluorophore yield.

The FDOT measurements were collected using the FDOT imaging system reported in [46]. Specifically, focused collimated laser beams were used as point light sources to excite the fluorophore. We had 60 point sources in total. The fluorescence measurements were collected by an electrically cooled CCD camera. The reading of the detector was recorded as the mean value of a subregion with 5×5 pixels around each detector location. We selected 60 detector locations. The

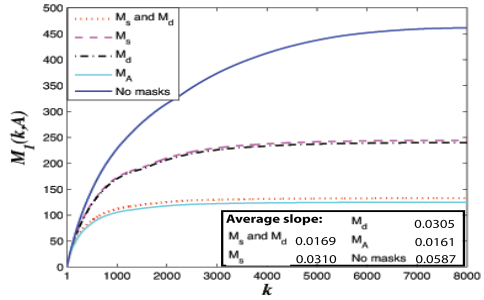
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

measurements due to each point sources were collected at the detectors. Assuming that the measurements obtained for each source-detector pair have high SNR, we approximated the measurements due to each illumination pattern as a linear combination of measurements due to a single source due to the linearity of (11) and (35). We discretized the imaging domain into $20 \times 20 \times 20$ voxels. Thus, the forward sensing matrix is of dimension 3600 by 8000.

B. Coherence of the FDOT Forward Matrix



(a) Top 40% values of all the normalized inner products



(b) Cumulative coherence

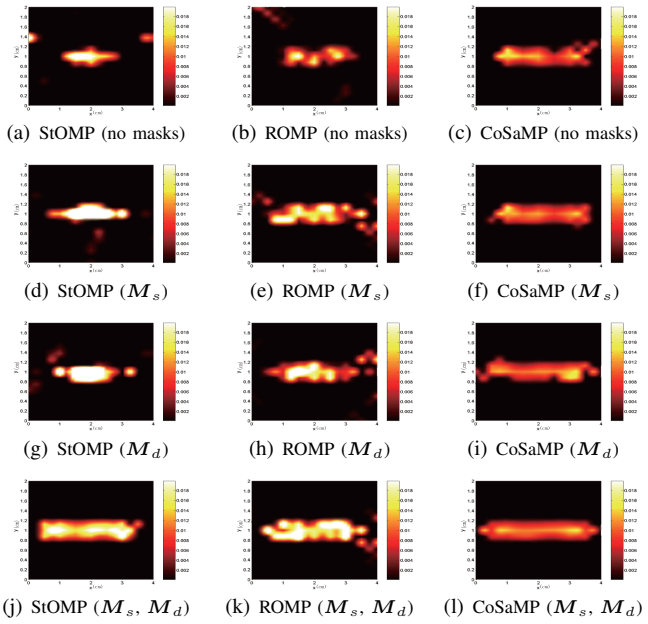
Fig. 11. The normalized inner products and cumulative coherence of the forward matrix before and after applying optical and measurement masks in real silicon phantom experiment.

We evaluated the effect of $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ and $\boldsymbol{M}_A$ on the coherence of the forward matrix and on the quality of reconstructed images. Fig. 11(a) shows the largest 40% of all the normalized inner products between different columns of the forward sensing matrix, arranged in a descending order. Note that the remaining 60% of the normalized inner products between different columns are close to 0. $\boldsymbol{M}_s$, $\boldsymbol{M}_d$ and $\boldsymbol{M}_A$ all decrease the normalized inner products between different columns. The relative percentages of the AUC as compared to the no-masks case are given in the box in Fig. 11(a).

Fig. 11(b) shows the cumulative coherence $\mathcal{M}_1(k, \boldsymbol{A})$ of the forward matrix. We see a reduction in the average slope of each curve as shown in the box in Fig. 11(b).

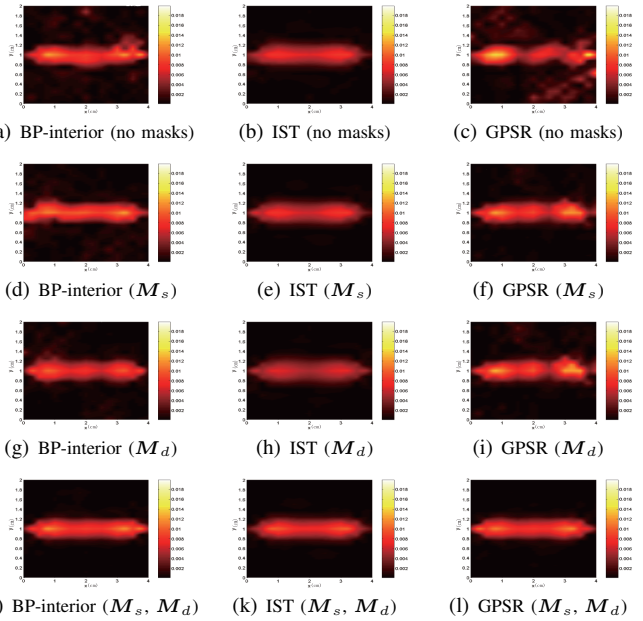
C. Image Reconstruction

The cross section of the original phantom at $z = 1$ (middle) is shown in Fig. 10(b). The cross sections of the reconstructed fluorophore concentration maps of the silicon phantom using different greedy algorithms are shown in Fig.12 and those using convex relaxation techniques are shown in



(a) StOMP (no masks) (b) ROMP (no masks) (c) CoSaMP (no masks)
(d) StOMP ($\boldsymbol{M}_s$) (e) ROMP ($\boldsymbol{M}_s$) (f) CoSaMP ($\boldsymbol{M}_s$)
(g) StOMP ($\boldsymbol{M}_d$) (h) ROMP ($\boldsymbol{M}_d$) (i) CoSaMP ($\boldsymbol{M}_d$)
(j) StOMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) ROMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) CoSaMP ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)

Fig. 12. The cross sections of the reconstructed phantom at $z = 1$cm (middle) using greedy algorithms in real silicon phantom experiment.



(a) BP-interior (no masks) (b) IST (no masks) (c) GPSR (no masks)
(d) BP-interior ($\boldsymbol{M}_s$) (e) IST ($\boldsymbol{M}_s$) (f) GPSR ($\boldsymbol{M}_s$)
(g) BP-interior ($\boldsymbol{M}_d$) (h) IST ($\boldsymbol{M}_d$) (i) GPSR ($\boldsymbol{M}_d$)
(j) BP-interior ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (k) IST ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$) (l) GPSR ($\boldsymbol{M}_s$, $\boldsymbol{M}_d$)

Fig. 13. The cross sections of the phantom at $z = 1$cm (middle) using convex relaxation techniques in real silicon phantom experiment.

Fig.13. We observe that the application of $\boldsymbol{M}_s$ and $\boldsymbol{M}_d$ improves the reconstruction results for all algorithms. Similar to the numerical simulation results, the improvements in the visual quality are most noticeable when the greedy algorithms are used.

VIII. Conclusion

We presented a method of designing illumination and detection patterns for FDOT imaging to optimize the reconstruction of sparse fluorophore concentration maps. Based on the assumption of weak fluorophore, we formulated the FDOT

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

image reconstruction as a linearized inverse problem. In CS theory, accurate recovery of a sparse signal from an underdetermined linear system requires the underlying forward matrix to be incoherent. We showed that the FDOT forward matrix can be represented as the Kronecker product of two matrices. The first matrix is determined by the excitation light field, and the second matrix is determined by the Green's function of the emission light field. We showed that the incoherence of the FDOT forward matrix is related to the incoherence of these two matrices and design preconditioners to reduce the incoherence of these matrices. The preconditioners result in a set of intensity patterns for multiple spatially distributed sources illuminating the imaging domain simultaneously and a linear filter applied to the corresponding measurements collected at multiple spatially distributed detectors. To reconstruct the fluorophore concentration map, we used sparse signal recovery techniques, including both the greedy type and convex relaxation algorithms. We showed that the application of intensity patterns and filtering of the measurements in conjunction with sparse signal recovery techniques improves the visual quality of reconstructed images and reduces the mean square error in both numerical simulations and in a real phantom experiment.

Our approach can be viewed as a method of determining the optimal source and detector locations and source intensities. Given a fluorophore designed to accumulate in certain regions and anatomical a priori information, optimal geometry for source-detector locations and source intensities can be determined via numerical simulations prior to the imaging process. This may eliminate unnecessary illumination, optimize imaging process and result in better image quality than the conventional FDOT imaging process.

We note that the approach introduced in this work can be easily extended to nonlinear FDOT inverse problem within an iterative perturbation approach, and FDOT imaging with multiple modulation frequencies.

Finally, our method is not limited to FDOT imaging, it can be extended to other imaging reconstruction problems involving wave propagation, such as DOT [47], [48], electromagnetic and acoustic imaging whenever sparsity is sought.

APPENDIX A
UPPER BOUNDS ON THE CUMULATIVE COHERENCE

Let

$$R(p,Q_k)=\sum_{|Q_k|=k,p\notin Q_k,q\in Q_k}|r_{A_{p,q}}|, \tag{63}$$

and

$$\begin{aligned} R(p_0,Q_{k_0}) &= \mathcal{M}_1(k,\boldsymbol{A})=\max_{p,Q_k}R(p,Q_k) \\ &\quad \text{such that } |Q_k|=k \text{ and } p\notin Q_k. \end{aligned} \tag{64}$$

Note that

$$R(p_0,Q_{k_0})=\sum_{q\in Q_{k_0}} r_{\phi_{p_0,q}}r_{g_{p_0,q}}\le\frac{1}{2}\sum_{q\in Q_{k_0}}(r^2_{\phi_{p_0,q}}+r^2_{g_{p_0,q}}). \tag{65}$$

Let $\tilde{\boldsymbol{\Phi}}$ and $\tilde{\boldsymbol{G}}$ denote the $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ matrices with their columns normalized to unity; and $\tilde{\boldsymbol{\Phi}}_{Q_{k_0}}$ is the $N_s\times k$ submatrix of $\tilde{\boldsymbol{\Phi}}$ composed the columns in Q_{k_0}, and $\tilde{\boldsymbol{G}}_{Q_{k_0}}$ is the $N_d\times k$ submatrix of $\tilde{\boldsymbol{G}}$ composed the columns in Q_{k_0}. Thus, we have

$$\begin{aligned} &\textstyle\sum_{q\in Q_{k_0}}(r^2_{\phi_{p_0,q}}+r^2_{g_{p_0,q}})=\|\tilde{\boldsymbol{\Phi}}^T_{Q_{k_0}}\tilde{\boldsymbol{\phi}}_{p_0}\|_2^2+\|\tilde{\boldsymbol{G}}^T_{Q_{k_0}}\tilde{\boldsymbol{\phi}}_{p_0}\|_2^2 \\ &\le\lambda_{\max_{\phi_k}}\|\tilde{\boldsymbol{\phi}}_{p_0}\|_2^2+\lambda_{\max_{g_k}}\|\tilde{\boldsymbol{g}}_{p_0}\|_2^2=\lambda_{\max_{\phi_k}}+\lambda_{\max_{g_k}} \end{aligned} \tag{66}$$

where $\lambda_{\max_{\phi_k}}$ is the maximum eigenvalue of the $k\times k$ matrix $\boldsymbol{H}_\phi=\tilde{\boldsymbol{\Phi}}^T_{Q_{k_0}}\tilde{\boldsymbol{\Phi}}_{Q_{k_0}}$, and $\lambda_{\max_{g_k}}$ is the maximum eigenvalue of the $k\times k$ matrix $\boldsymbol{H}_g=\tilde{\boldsymbol{G}}^T_{Q_{k_0}}\tilde{\boldsymbol{G}}_{Q_{k_0}}$. The maximum value of $\lambda_{\max_{\phi_k}}$ and $\lambda_{\max_{g_k}}$ is k. From (64)-(66),

$$\mathcal{M}_1(k,\boldsymbol{A})\le\frac{1}{2}(\lambda_{\max_{\phi_k}}+\lambda_{\max_{g_k}}). \tag{67}$$

Let $\boldsymbol{H}_\phi=\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}$ and $\boldsymbol{H}_g=\tilde{\boldsymbol{G}}^T\tilde{\boldsymbol{G}}$ denote the Gram-type matrix, and let $\lambda_{\phi_1}\ge\lambda_{\phi_2}\ge\ldots\lambda_{\phi_{N_s}}\ge0$ denote the nonzero eigenvalues of $\boldsymbol{H}_\phi$, $\lambda_{g_1}\ge\lambda_{g_2}\ge\ldots\lambda_{g_{N_d}}\ge0$ denote the nonzero eigenvalues of $\boldsymbol{H}_g$. Then, we have

$$\|\boldsymbol{H}_\phi-\boldsymbol{I}_N\|_F^2=\sum_{i=1}^{N_s}\lambda^2_{\phi_i}-N \text{ and } \|\boldsymbol{H}_g-\boldsymbol{I}_N\|_F^2=\sum_{i=1}^{N_d}\lambda^2_{g_i}-N. \tag{68}$$

We derive the ratio, r_1, between $\lambda_{\max_{\phi_k}}$ and $\|\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}-\boldsymbol{I}_N\|_F^2$ as follows:

$$\begin{aligned} r_1 &= \frac{\lambda_{\max_{\phi_k}}}{\|\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}-\boldsymbol{I}_N\|_F^2}=\frac{\lambda_{\max_{\phi_k}}}{\sum_{i=1}^{N_s}\lambda^2_{\phi_i}-N} \\ &\le \frac{\lambda_{\max_{\phi_k}}}{\frac{N_s^2}{N_s}-N}\le\frac{k}{\frac{N_s^2}{N_s}-N}. \end{aligned} \tag{69}$$

The first inequality holds because the squared summation of the eigenvalues is minimum, if all the eigenvalues λ_{ϕ_i} $(i=1,\ldots,N_s)$ are equal to each other $(\lambda_{\phi_i}=\frac{N}{N_s})$. The second inequality holds since $\lambda_{\max_{\phi_k}}$ has the maximum value k.

The upper bound derived in (69) is not tight for large k values. In this case, we note that $\tilde{\boldsymbol{\Phi}}_{Q_{k_0}}$ is a submatrix of $\tilde{\boldsymbol{\Phi}}$. Thus $\lambda_{\phi_1}\ge\lambda_{\max_{\phi_k}}$, we have

$$\begin{aligned} r_1 &= \frac{\lambda_{\max_{\phi_k}}}{\|\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}-\boldsymbol{I}_N\|_F^2}=\frac{\lambda_{\max_{\phi_k}}}{\sum_{i=1}^{N_s}\lambda^2_{\phi_i}-N}\le\frac{\lambda_{\phi_1}}{\sum_{i}^{N_s}\lambda^2_{\phi_i}-N} \\ &\le\frac{\lambda_{\phi_1}}{\lambda^2_{\phi_1}+\frac{(N-\lambda_{\phi_1})^2}{N_s-1}-N}\le\frac{N_s-1}{2(\sqrt{N_sN^2-N_s^2N+N_sN}-N)}. \end{aligned} \tag{70}$$

Using (69) and (70), we obtain

$$r_1=\frac{\lambda_{\max_{\phi_k}}}{\|\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}-\boldsymbol{I}_N\|_F^2}\le\bar{r}_1 \tag{71}$$

where $\bar{r}_1=\min\{\frac{kN_s}{N^2-NN_s},\ \frac{N_s-1}{2(\sqrt{N_sN^2-N_s^2N+N_sN}-N)}\}$.

Following similar steps as above, we derive the ratio, r_2, between $\lambda_{\max_{g_k}}$ and $\|\tilde{\boldsymbol{G}}^T\tilde{\boldsymbol{G}}-\boldsymbol{I}_N\|_F^2$ and obtain

$$r_2=\frac{\lambda_{\max_{g_k}}}{\|\tilde{\boldsymbol{G}}^T\tilde{\boldsymbol{G}}-\boldsymbol{I}_N\|_F^2}\le\bar{r}_2 \tag{72}$$

where $\bar{r}_2=\min\{\frac{kN_d}{N^2-NN_d},\ \frac{N_d-1}{2(\sqrt{N_dN^2-N_d^2N+N_dN}-N)}\}$.

From (67), (71) and (72),

$$\mathcal{M}_1(k,\boldsymbol{A})\le\frac{\max\{\bar{r}_1,\bar{r}_2\}}{2}(\|\tilde{\boldsymbol{\Phi}}^T\tilde{\boldsymbol{\Phi}}-\boldsymbol{I}_N\|_F^2+\|\tilde{\boldsymbol{G}}^T\tilde{\boldsymbol{G}}-\boldsymbol{I}_N\|_F^2). \tag{73}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

REFERENCES

[1] V. Ntziachristos, "Fluorescence molecular imaging," *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 1–33, 2006.

[2] B. Tromberg, B. Pogue, K. Paulsen, A. Yodh, D. Boas, and A. Cerussi, "Assessing the future of diffuse optical imaging technologies for breast cancer management," *Med. Phys.*, vol. 35, no. 6, p. 2443C2451, 2008.

[3] L. Zhou and B. Yazici, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography in the presence of measurement noise," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1094–1111, Apr. 2011.

[4] B. Alacam and B. Yazici, "Direct reconstruction of pharmacokinetic-rate images of optical fluorophores from nir measurements," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1337–1353, Sep. 2009.

[5] B. Alacam, B. Yazici, X. Intes, and B. Chance, "Extended kalman filtering for the modeling and analysis of icg pharmacokinetics in cancerous tumors using nir optical methods," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 1861–1871, 2006.

[6] K. Ren, G. Bal, and A. Hielscher, "Transport- and diffusion-based optical tomography in small domains: a comparative study," *Appl. Optics*, vol. 46, no. 27, pp. 6669–6679, 2007.

[7] M. Guven, L. Zhou, L. Reilly-Raska, and B. Yazici, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography:part i," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 217–229, Feb. 2010.

[8] ——, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography:part ii," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 230–245, Feb. 2010.

[9] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[10] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory.*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[11] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.

[12] J. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.

[13] K.Schnass and P. Vandergheynst;, "Dictionary preconditioning for greedy algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1994–2002, May 2008.

[14] L.Zelnik-Manor, K.Rosenblum, and Y. Eldar, "Sensing matrix optimization for block-sparse decoding," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4300–4312, Sep. 2011.

[15] W. Bajwa, R. Calderbank, and S. Jafarpour, "Why gabor frames? two fundamental measures of coherence and their role in model selection," *J. Commun. Netw.*, vol. 12, no. 4, pp. 289–307, Aug. 2010.

[16] A. Juditsky and A. Nemirovski, "On verifiable sufficient conditions for sparse signal recovery via l_1 minimization," *Mathematical Programming: Series A and B - Special Issue on 'Optimization and Machine learning'*, vol. 127, no. 1, 2011. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:0809.2650

[17] A. Serdaroglu, B. Yazici, and K. Kwon, "Optimum source design for detection of heterogeneities in diffuse optical imaging," in *Proceedings of SPIE Photonic West*, vol. 6139, 2006.

[18] A. Joshi, W. Bangerth, and E. Sevick-Muraca, "Non-contact fluorescence optical tomography with scanning patterned illumination," *Opt. Express.*, vol. 14, no. 14, pp. 6516–6534, Jul. 2006.

[19] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Adaptive finite element modeling of optical fluoresc enceenhanced tomography," *Opt. Express.*, vol. 12, pp. 5402–5417, 2004.

[20] A. Joshi, W. Bangerth, K. Hwang, J. Rasmussen, and E. M. Sevick-Muraca, "Plane wave fluorescence tomography with adaptive finite elements," *Opt. Lett.*, vol. 31, no. 2, pp. 193–195, 2006.

[21] A. Godavarty, A. Thompson, R. Roy, M. Gurfinkel, M. J. Eppstein, C. Zhang, and E. Sevick-Muraca, "Diagnostic imaging of breast cancer using fluorescence-enhanced optical tomography: phantom studies," *J. Biomed. Opt.*, vol. 9, no. 3, pp. 488–496, 2004.

[22] V. Lukic, V. A. Markel, and J. C. Schotland, "Optical tomography with structured illumination," *Opt. lett.*, vol. 34, no. 7, pp. 983–985, 2009.

[23] J. Chen, V. Venugopal, F. Lesage, and X. Intes, "Time-resolved diffuse optical tomography with patternedlight illumination and detection," *Opt. lett.*, vol. 35, no. 13, pp. 188–191, 2010.

[24] O. Lee, J. Kim, Y. Bresler, and J. Ye, "Compressive diffuse optical tomography: Non-iterative exact reconstruction using joint sparsity," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1129–1142, May 2011.

[25] S. Meikle, "An investigation of coded aperture imaging for small animal spect," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 3, pp. 816–821, Jun. 2001.

[26] H. Arguello and G. Arce, "Code aperture optimization for spectrally agile compressive imaging," *JOSA A*, vol. 28, no. 11, pp. 2400–2413, 2011.

[27] R. Marcia and R. W. Z. Harmany, "Compressive coded apertures for high-resolution imaging," in *Proceedings of SPIE In Photonics Europe*, vol. 7723, Apr. 2010, pp. 772 304–77 230 411.

[28] P. Mohajerani, A. Eftekhar, J. Huang, and A. Adibi, "Optimal sparse solution for fluorescent diffuse optical tomography: theory and phantom experimental results," *Appl. Opt.*, vol. 46, no. 10, pp. 1679–1685, 2007.

[29] N. Cao, A. Nehorai1, and M. Jacob, "Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm," *Opt. express.*, vol. 15, no. 21, pp. 13 695–13 708, 2007.

[30] M. Suzen, A. Giannoula, and T. Durduran, "Compressed sensing in diffuse optical tomography," *Opt. Express.*, vol. 18, no. 23, pp. 23 676–23 690, 2010.

[31] M. Suzen, A. Giannoula, P. Zirak, N. Oliverio, U. Weigel, P. Farzam, and T. Durduran, "Sparse image reconstruction in diffuse optical tomography: an application of compressed sensing," in *Biomedical Optics (BIOMED)*, Miami, Florida, Apr. 2010.

[32] J. Ye, S. Lee, and Y. Bresler, "Exact reconstruction formula for diffuse optical tomography using simultaneous sparse representation," in *Biomedical Imaging: From Nano to Macro, ISBI 2008. 5th IEEE International Symposium on*, May 2008, pp. 1621–1624.

[33] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*, fisrt ed. Springer, 2010.

[34] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[35] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientific Computing*, vol. 20, pp. 33–61, 1999.

[36] G. Tang and A. Nehorai, "Verifiable and computable performance analysis of sparsity recovery,," *submitted for publication*.

[37] A. Jin, B. Yazici, A. Ale, and V. Ntziachristos, "Performance evaluation of fluorescence diffuse optical tomography sensing matrix preconditioning using real data," *Opt. lett.*, vol. 37, no. 20, pp. 4326–4328, 2012.

[38] N. Kahale and R. Urbanke, "Matrix theory and its applications," *Lecture notes prepared by Americal Mathematicl Society Short Course*, Jan. 1989.

[39] D. Donoho, Y. Tsaig, I. Drori, and J. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," Tech. Rep., 2006.

[40] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.*, vol. 9, no. 3, pp. 317–334, Apr. 2009.

[41] I. Daubechies, M. Defriese, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[42] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 586–597, Dec. 2007.

[43] B. Dogdas, D. Stout, A. Chatziioannou, and R. Leahy, "A 3-d whole body mouse atlas from ct and cryosection data," *Phys. Med. Biol.*, vol. 52, pp. 577–587, 2007.

[44] D. Stout, P. Chow, R. Silverman, R. Leahy, X. Lewis, S. Gambhir, and A. Chatziioannou, "Creating a whole body digital mouse atlas with pet, ct and cryosection images," *Molecular Imaging and Biology*, vol. 4, no. 4, p. S27, 2002.

[45] G. Alexandrakis, F. Rannou, and A. Chatziioannou, "Tomographic bioluminescence imaging by use of a combined optical-pet (opet) system: a computer simulation feasibility study," *Phys. Med. Biol.*, vol. 50, no. 17, pp. 4225–4241, 2005.

[46] N. Deliolanis, T. Lasser, D. Hyde, A. Soubret, J. Ripoll, and V. Ntziachristos, "Free-space fluorescence molecular tomography utilizing 360 degrees geometry projections," *Opt. Lett.*, vol. 32, no. 4, pp. 382–384, 2007.

[47] K. Kwon, B.Yazici, and M. Guven, "Two-level domain decomposition methods for diffuse optical tomography," *Inverse Probl.*, vol. 22, pp. 1533–1559, 2006.

[48] M. Guven, B. Yazici, and B. C. X. Intes, "An adaptive multigrid algorithm for region of interest diffuse optical tomography," in *Proc. IEEE International Conference on Image Processing (ICIP)*, vol. 2, Sep. 2003, pp. 1533–1559.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIP.2014.2300756, IEEE Transactions on Image Processing

An Jin received the B.S. degree in biomedical engineering from Zhejiang University, China, and the Ph.D. degree in biomedical engineering from Rensselaer Polytechnic Institute, Troy, NY, in 2006 and 2012, respectively. She currently is algorithm engineer III at DRV. Technologies in Bellevue, WA. Dr. Jin's research area includes medical imaging, biomedical optics, signal/image processing and machine learning.

Birsen Yazıcı received BS degrees in Electrical Engineering and Mathematics in 1988 from Bogazici University, Istanbul Turkey and MS and Ph.D. degrees in Mathematics and Electrical Engineering both from Purdue University, W. Lafayette IN, in 1990 and 1994, respectively. From September 1994 until 2000, she was a research engineer at General Electric Company Global Research Center, Schenectady NY. During her tenure in industry, she worked on radar, transportation, industrial and medical imaging systems. From 2001 to June 2003, she was an assistant professor at Drexel University, Electrical and Computer Engineering Department. In Fall 2003, she joined Rensselaer Polytechnic Institute where she is currently a full professor in the Department of Electrical, Computer and Systems Engineering and in the Department of Biomedical Engineering.

Prof. Yazıcıs research interests span the areas of statistical signal processing, inverse problems in imaging, image reconstruction, biomedical optics, radar and X-ray imaging. From 2008 to 2012, she was an associate editor for the IEEE Transactions on Image Processing. She currently serves as an associate editor for SIAM Journal on Imaging Science. She is the recipient of the Rensselaer Polytechnic Institute 2007 and 2013 School of Engineering Research Excellence Awards. She holds 11 US patents.

Vasilis Ntziachristos has research area that encompasses the development of new methods and devices for biological and medical imaging focusing on innovative non-invasive approaches that visualize previously unseen physiological and molecular processes in tissues. His research also aims to translate these methods to advance biological discovery, accelerate drug development and offer new efficient methods for diagnostics and theranostics. Prof. Ntziachristos studied electrical engineering at Aristotle University in Thessaloniki, Greece and received his Masters and Doctorate degrees from the Bioengineering Department of the University of Pennsylvania. He served as assistant professor and director of the Laboratory for Bio-Optics and Molecular Imaging at Har-vard University and Massachusetts General Hospital. His current professorial chair is closely linked with the Institute for Biological and Medical Imaging at the Helmholtz Zentrum Mnchen, of which Prof. Ntziachristos is director.