Data-resolution based optimization of the data-collection strategy for near infrared diffuse optical tomography

Deepak Karkala and Phaneendra K. Yalavarthy[a)]
Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560012, India

(Received 18 January 2012; revised 16 June 2012; accepted for publication 19 June 2012; published 18 July 2012)

Purpose: To optimize the data-collection strategy for diffuse optical tomography and to obtain a set of independent measurements among the total measurements using the model based data-resolution matrix characteristics.
Methods: The data-resolution matrix is computed based on the sensitivity matrix and the regularization scheme used in the reconstruction procedure by matching the predicted data with the actual one. The diagonal values of data-resolution matrix show the importance of a particular measurement and the magnitude of off-diagonal entries shows the dependence among measurements. Based on the closeness of diagonal value magnitude to off-diagonal entries, the independent measurements choice is made. The reconstruction results obtained using all measurements were compared to the ones obtained using only independent measurements in both numerical and experimental phantom cases. The traditional singular value analysis was also performed to compare the results obtained using the proposed method.
Results: The results indicate that choosing only independent measurements based on data-resolution matrix characteristics for the image reconstruction does not compromise the reconstructed image quality significantly, in turn reduces the data-collection time associated with the procedure. When the same number of measurements (equivalent to independent ones) are chosen at random, the reconstruction results were having poor quality with major boundary artifacts. The number of independent measurements obtained using data-resolution matrix analysis is much higher compared to that obtained using the singular value analysis.
Conclusions: The data-resolution matrix analysis is able to provide the high level of optimization needed for effective data-collection in diffuse optical imaging. The analysis itself is independent of noise characteristics in the data, resulting in an universal framework to characterize and optimize a given data-collection strategy. © *2012 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4736820]

Key words: near infrared, diffuse optical tomography, image reconstruction, inverse problems, data resolution

I. INTRODUCTION

Near infrared (NIR) diffuse optical tomography is an emerging imaging modality that has been widely investigated for breast, brain, and small animal imaging.[1–4] Diffuse optical tomography uses a finite set of boundary measurements using light in the spectral range of 600–1000 nm to reconstruct/estimate the internal distribution of optical properties.[5] The probing media, which is the NIR light, in these studies is delivered through optical fibers and the transmitted light is also collected through the same fibers which are in contact with the external surface of the tissue. Using these measurements, distributions of absorption and scattering coefficients of the tissue are reconstructed via a model-based iterative algorithm.[6] As NIR light is nonionizing, it has the advantage of providing functional changes in a tissue over a prolonged time.

The dominance of light scattering in soft-tissue at NIR wavelengths makes the reconstruction (also known as the inverse problem) a nonlinear, ill-posed, and sometimes underdetermined problem.[5] This leads to infinitely many solutions possible for the inverse problem. To overcome this difficulty, typically a regularization term is added to provide a unique solution to the image reconstruction problem. Choice of regularization influences the outcome of the reconstruction procedure and is known to be biasing the solution space. This regularization is typically chosen based on the prior information (or experience) available about the imaging problem, including the noise characteristics of the data-collection system.

The data-collection systems or strategies play an important role in determining the reconstructed image characteristics and there have been techniques that were proposed to evaluate them. These methods quantitatively evaluate the data-collection strategies in terms of the number of useful measurements as well as the reconstructed image quality. Culver *et al.*[7] have performed the singular value analysis of the weight/sensitivity matrix to determine the optimal data-collection strategy for a simple slab medium. Xu *et al.*[4] have extended this approach for heterogeneous and irregular imaging domains to determine the best data-collection strategy for small animal imaging studies. The same analysis was also used by Yalavarthy *et al.*[8] to determine the critical

computational parameters in a circular tomographic imaging, with an emphasis on number of measurements. Chen and Chen[9] have used Cramer–Rao lower bounds to perform the same optimization with an advantage that their analysis provided direct estimations of parameters without solving the inverse problem. Also Dehghani *et al.*[10] studied the total sensitivity to make a choice of the data-collection strategy in brain imaging with dense imaging arrays.

These kind of optimization techniques are critical in designing a data-collection system for diffuse optical imaging studies, where building and testing of these data-collection arrays becomes expensive, time-consuming, and sometime not feasible due to limited resources. The optimization techniques discussed in the literature were able to provide the optimal data-collection strategy, which is typically based on the number of useful measurements. But none of these techniques were able to assess the independent measurements for a given data-collection strategy, specifically they are unable to provide the information whether a particular measurement is independent or not. This specific information about the independent measurements rather than the number alone is of particular importance for the instrumentation aspect of the diffuse optical tomographic imaging system. Knowing this information reduces the data acquisition time, in turn reducing the overall protocol time. This work aims to introduce a novel way of optimizing the data-collection strategies, with a capability of providing specific information about the independent measurements, based on the data-resolution matrix. This data-resolution matrix is primarily based on the sensitivity/weight matrix and the regularization scheme used in the reconstruction.

In the present work, three-dimensional (3D) cylindrical and patient breast imaging geometries were chosen to assess the developed methodology to provide independent measurements. As the emphasis is on presenting a novel optimization method for evaluating a data-collection strategy, continuous wave (CW) case alone is considered here, where only the amplitude data are considered with optical absorption coefficient being the unknown imaging parameter. For completeness, the results in here also include the comparison with traditional singular value based analysis. As presented analysis provides the information about independent measurements, the effect of omission of dependent measurements on the reconstructed absorption image quality is studied systematically using both numerical and phantom experiments. It will be proved that the omission of dependent measurements has no significant effect on the reconstructed image quality.

II. DIFFUSE OPTICAL TOMOGRAPHY: FORWARD PROBLEM

Continuous wave NIR light propagation in a biological tissue like breast can be modeled using diffusion equation (DE),[5,11] given as

$$-\nabla.D(r)\nabla\Phi(r)+\mu_a(r)\Phi(r)=Q_o(r), \tag{1}$$

where the optical diffusion and absorption coefficients are given by $D(r)$ and $\mu_a(r)$, respectively. The continuous wave light source, represented by $Q_o(r)$, is modeled as isotropic. $\Phi(r)$ is the photon fluency density at a given position r. The diffusion coefficient is defined as

$$D(r)=\frac{1}{3[\mu_a(r)+\mu_s'(r)]}, \tag{2}$$

where $\mu_s'(r)$ is the reduced scattering coefficient, which is defined as $\mu_s'=\mu_s(1-g)$ with μ_s as the scattering coefficient and g as the anisotropy factor. In the present work, μ_s' is assumed to be known and remains constant throughout the domain. The finite element method (FEM) is used to solve Eq. (1) to generate modeled data for a given distribution of the absorption coefficient $\mu_a(r)$. Under the Rytov approximation, the modeled data become the natural logarithm of the intensity (A). A Type-III boundary condition is employed to account for the refractive index mismatch at the boundary.[12] This forward model is used repeatedly in an iterative manner to estimate the optical property of the tissue under investigation.[5]

III. DIFFUSE OPTICAL TOMOGRAPHY: INVERSE PROBLEM

The inverse problem primarily involves the estimation of optical absorption coefficients from the CW boundary measurements using a model-based approach. This is achieved by matching the experimental measurements with model-based ones iteratively in the least-squares sense over the range of μ_a. This minimization problem can be solved using several approaches with most common one involving computing of repeated solutions of the forward model [including Jacobian ($\mathbf{J}$)] and solving linear system of equations.

The Levenberg–Marquardt (LM) optimization scheme is most popular in solving the inverse problem,[13] where the objective function is defined as

$$\Omega=\|y-G(\mu_a)\|^2, \tag{3}$$

where y is the natural logarithm of the amplitude of experimental data and $G(\mu_a)$ represents the modeled data. Minimization of this objective function is achieved by setting the first-order derivative of Ω (Ref. 13) with respect to μ_a equal to zero

$$\frac{\partial\Omega}{\partial\mu_a}=\mathbf{J}^T\delta=0, \tag{4}$$

where δ is the data-model misfit, $\delta=(y-G(\mu_a))$, and $\mathbf{J}$ represents the Jacobian [$\mathbf{J}=\partial G(\mu_a)/\partial\mu_a$]. The Jacobian matrix has the dimensions of NM $\times$ NN, where NM represents the number of measurements and NN represents the number of nodes in the finite element mesh. The Jacobian (also known as sensitivity matrix) maps the changes in logarithm of amplitude to the absorption changes at each node of the FEM model.

Due to the ill-conditioned nature of the problem, the update for the optical properties at iteration "i" is written as[13]

$$\triangle\mu_a^i=\left[\mathbf{J}^T\mathbf{J}+\lambda\mathbf{I}\right]^{-1}\mathbf{J}^T\delta_{i-1} \tag{5}$$

or equivalently[14] from the Appendix

$$\triangle\mu_a^i = \mathbf{J}^T\left[\mathbf{J}\mathbf{J}^T + \lambda\mathbf{I}\right]^{-1}\delta_{i-1}, \tag{6}$$

where $\triangle\mu_a^i$ represents the update of the optical property μ_a at the ith step with λ as the regularization parameter, which stabilizes the solution and ensures the convergence.[15] The λ value here is decreased by a small factor with every iteration as given in Ref. 13.

In Eqs. (5) and (6), the regularization parameter λ (strictly positive) is chosen empirically and is multiplied by the maximum of the diagonal values of $\mathbf{J}\mathbf{J}^T$.[13] In the presented work using simulated data, 1% normally distributed Gaussian noise was added to the amplitude, which is a typical noise observed in the experimental data.[16] The iterative procedure is stopped when the L2 norm of the data-model misfit (δ) does not improve, by more than 2%.

III.A. Data-resolution matrix

The resolution characteristics of the inverse problem are primarily dependent on the model used in solving the same. The model here refers to both the forward model and stabilization scheme. This approach primarily relies on matching the inverse solution using model with the expected distribution of the optical properties.[15,17]

Expanding $G(\mu_a)$ using Taylor series around μ_{a0} (where μ_{a0} represents some initial guess for μ_a) gives

$$G(\mu_a) = G(\mu_{a0}) + G'(\mu_a)(\mu_a - \mu_{a0}) + \cdots(\mu_a - \mu_{a0})^T G''(\mu_a)(\mu_a - \mu_{a0}) + \cdots, \tag{7}$$

where $G'(\mu_a) = \mathbf{J}$ is the Jacobian and $G''(\mu_a)$ is the Hessian.

Neglecting the higher order terms results in

$$G(\mu_a) = G(\mu_{a0}) + \mathbf{J}(\mu_a - \mu_{a0}), \tag{8}$$

where $\triangle\mu_a = \mu_a - \mu_{a0}$ is the update and assuming that $G(\mu_a) = y$ leads to

$$\mathbf{J}\triangle\mu_a = \delta' \tag{9}$$

with $\delta' = y - G(\mu_{a0})$, representing the data-model misfit for using a perfect model [arising when $G(\mu_a) = y$]. Substituting for $\triangle\mu_a$ from Eq. (6) in Eq. (9) leads to

$$\left[\mathbf{J}\mathbf{J}^T\left[\mathbf{J}\mathbf{J}^T + \lambda\mathbf{I}\right]^{-1}\right]\delta = \delta' \tag{10}$$

define

$$\mathbf{N} = \mathbf{J}\mathbf{J}^T\left[\mathbf{J}\mathbf{J}^T + \lambda\mathbf{I}\right]^{-1}, \tag{11}$$

where "$\mathbf{N}$" is known as the data-resolution matrix (dimension: NM × NM), ideally giving the relation between using a perfect model to its linearized version. If $\mathbf{N} = \mathbf{I}$, then $\delta' = \delta$.

In Eq. (11), if λ is equal to zero, then the data-resolution matrix is an identity matrix resulting in maximum resolution. The case of $\lambda = 0$ does not arise as $\mathbf{J}\mathbf{J}^T$ is always ill-conditioned and requiring stabilization ($\lambda > 0$) to obtain $\triangle\mu_a$. In which case ($\lambda > 0$), the data-resolution matrix instead describes how the inverse solution modifies the original model into a recovered model.

It is to be noted that the data-resolution matrix does not depend on specific data (y) or error in it but are exclusively the properties of $\mathbf{J}$ and the regularization (λ) used. The closer it is to the identity matrix, the smaller are the prediction errors for δ. Also, note that each predicted data point is a weighted average of the observed data. So the rows of $\mathbf{N}$ show how well the data can be independently predicted or resolved, with magnitude of values being between 0 and 1. The diagonal entries of $\mathbf{N}$ indicate the weight of a data point on its own prediction, in turn indicating the importance of the data. The higher magnitude of off-diagonal entries in a particular row i indicate their dependency on measurement i, resulting in a methodology to predict the dependent measurements. This is described algorithmically in Subsection III.B.

III.B. Determining the independent measurements

The process of determining independent measurements is given in Algorithm 1. As with the iterative procedure of reconstructing optical properties start with an initial guess typically obtained using calibration procedure, the same uniform initial guess is used in calculating $\mathbf{J}$. The calibration procedure typically assumes that the imaging domain is either infinite or semi-infinite for which analytical solutions are readily available. Using these analytical solutions, experimental/numerical data are calibrated to remove the biases caused by variation in detector sensitivities as well as source strength.[18] This procedure also results in bulk optical properties estimated through the analytical solution (similar to least-squares data fit), which tend to be close to background optical properties of the tissue under investigation.[18] These bulk optical properties are used as initial guess to the iterative image reconstruction procedure and the same properties are also used in computing the $\mathbf{J}$ for determining the independent measurements.

A threshold value indicates the dependency level of the off-diagonal elements of a particular row to its diagonal entry, typical value chosen is between 0.8 and 0.99. Choice of 1 for threshold (th) assumes that all measurements are truly independent. After which, the computation of $\mathbf{N}$ is performed and a vector containing ones is initialized for keeping track of independent measurements. Row by row analysis to determine the dependent measurements is performed. The dependent "i" is made zero in the initialized vector (in here a). If a measurement is determined as dependent on earlier one, the analysis is not repeated (ensured by the if condition in step 4 of Algorithm 1). After running through the loop, the independent measurement indices (represented by "ind") are determined by finding the ones in the vector a.

After finding independent indices in the Jacobian, the dependent rows are removed and the reduced Jacobian becomes

$$\mathbf{J_n} = \mathbf{J}(\mathbf{ind}, :). \tag{12}$$

If NK represents the number of independent measurements obtained, the reduced Jacobian $\mathbf{J_n}$ has the dimensions of NK × NN. Accordingly the new update equation will be

$$\triangle\mu_i = \mathbf{J_n}^T\left[\mathbf{J_n}\mathbf{J_n}^T + \lambda\mathbf{I}\right]^{-1}\delta_{i-1}. \tag{13}$$

In our experience, we need to perform this operation only at the first iteration, the set of independent measurements does

ALGORITHM I. Algorithm for determining independent measurements.

1. Calculate Jacobian (**J**) using uniform initial guess and set threshold (th) to be between 0.8 and 1.
2. Determine the data-resolution matrix for a chosen value of λ
   $\mathbf{N} = \mathbf{J}\mathbf{J}^T\left[\mathbf{J}\mathbf{J}^T + \lambda\mathbf{I}\right]^{-1}$
3. Initialize a = ones(NM,1).
4. for i = 1,2,.....NM
   if $a(i) = 1$ then
   a. $x = \mathbf{N}(:, i)$
   b. Dependent measurements (dep) = those columns in x (excluding the i^{th} column) for which $x_j > \text{th} * x(i)$
   c. $a(\text{dep}) = 0$
   else
   go back to step: 4
   end.
5. Indices of independent measurements (ind) = indices for which $a(:) = 1$

not vary significantly with iterations. Also this determination requires one explicit inversion of a matrix having dimension of NM × NM and the computational cost does not exceed more than 8% of the total computational cost of one iteration.

IV. THREE-DIMENSIONAL TEST PROBLEMS

In order to effectively evaluate the measurement selection based on the data-resolution matrix, a series of simulation studies were performed. Initially, a cylindrical imaging domain was chosen for numerical experiments. The cylinder had a diameter of 86 mm and a height of 100 mm. The mesh used for the forward model as well as the reconstruction had 24 161 nodes corresponding to 116 757 linear tetrahedral elements. The data collection strategy consisted of 48 fibers that were arranged in a circular, equally spaced fashion in three layers spaced 10 mm apart with 16 fibers per plane. One fiber was used at a time as the source while all the other fibers were used as detectors to generate 2256 (48 × 47) measurements. This data-collection strategy is discussed in Ref. 8 as three layers out of plane.

A uniform breast mesh was also used which consisted of 18 723 nodes corresponding to 94 936 linear tetrahedral elements. Sixteen optical source detector fibers were placed equidistant in a circular manner (single plane) for data collection. The absorption coefficient distribution was reconstructed using the reduced Jacobian with the threshold values of 0.9, 0.8, and 0.7. It should be noted that a threshold of 1 corresponds to using the full Jacobian (in turn implying all measurements are independent) and will serve as a reference for comparison.

In all cases, the background optical properties were $\mu_a = 0.01$ mm^{-1} and $\mu_s' = 1.0$ mm^{-1}. A spherical object was used as a target in these simulation studies. The optical properties of the target were $\mu_a = 0.03$ mm^{-1} and $\mu_s' = 1.0$ mm^{-1}. The spherical target had a diameter of 20 mm and was placed in the x-plane at a depth of 10 mm from the surface. The target distributions are shown in the first columns of Figs. 2 and 3 for cylindrical and breast geometries, respectively.

Measurements with a noise level of 1% (normally distributed gaussian noise) was added to the computed/modeled data to mimic the experimental data in all cases discussed here. The choice of 1% noise level is on par with the expected noise level in a typical diffuse optical imaging system.[16] The background optical properties were selected as initial guess for the image reconstruction. All computations were carried out on a Linux workstation with an Intel Xeon Dual Quad Core 2.33 GHz processor with 64 GB of RAM.

IV.A. Experimental gelatin phantom

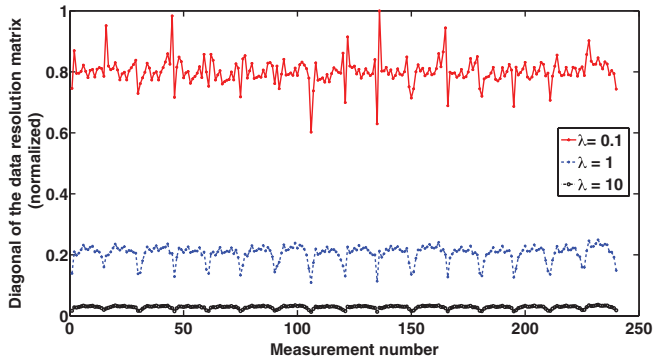
In order to validate the proposed algorithm using experimental data, a cylindrical gelatin based phantom was used. The phantom was fabricated[19] by hardening heated gelatin solution consisting of 80% deionized water and 20% gelatin (G2625, Sigma Inc.). It consisted of different amounts of India ink and titanium oxide (TiO_2) to mimic the absorption and scattering properties of a real tissue. The cylindrical phantom had a height of 60 mm and its diameter was 86 mm. A cylindrical hole (diameter = 16 mm) representing the tumor was placed close to the boundary stretching in the entire Z-direction. The background optical properties were $\mu_a = 0.008$ mm^{-1} and $\mu_s' = 0.9$ mm^{-1} at 785 nm wavelength. The cylindrical hole was filled with intralipid mixed with India ink to result in optical properties of $\mu_a = 0.02$ mm^{-1} and $\mu_s' = 1.0$ mm^{-1}. Near infrared data were collected at the phantom boundary using fibers located in a single plane along the midheight of the phantom.

A cylindrical mesh consisting of 12 695 nodes corresponding to 63 810 linear tetrahedral elements was used in computing the distributions of μ_a and the collected experimental data were calibrated using homogeneous phantom data as described in Ref. 18. The target absorption coefficient distribution is shown in the first column of Fig. 4.

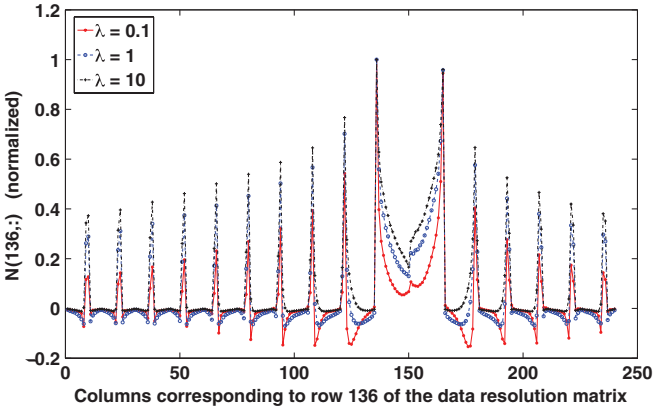
V. RESULTS

The data-resolution matrix can be used to determine how closely the inverse solution matches to a given model. The closer the data-resolution matrix is to an identity matrix, better is the resolution. Accordingly the data-resolution matrix was investigated and the results obtained are presented. The normalized diagonal values of the data-resolution matrix for different values of regularization are shown in Fig. 1(a) for the case of cylindrical imaging domain. Note that the diagonal values are inversely proportional to regularization parameter value. To understand the dependence on other measurements, normalized values of row 136 (where the diagonal of the data-resolution matrix is maximum for λ = 0.1) of the data-resolution matrix for different values of λ are shown in Fig. 1(b). Even though the maximum is at 136, the dependence of it on other measurements is clearly visible.

The number of independent measurements obtained using the proposed method (refer to Algorithm 1) is compared with the number of useful measurements obtained using the singular value analysis[8] in Table I. It can be inferred from the table



(a)



(b)

FIG. 1. Characteristics of data resolution matrix using cylindrical imaging domain with one plane of data-collection. (a) Normalized diagonal values for varying regularization parameter (λ). (b) Plot of magnitude of normalized (to 1) values in row 136 of the data resolution matrix for $\lambda = 0.1$, 1, and 10.

that as the threshold is decreased, the number of independent measurements will decrease.

In order to assess the algorithm's performance, a series of reconstructions were performed using the imaging domains discussed in Sec. IV. The reconstructed images using cylindrical mesh and 3D breast mesh, along with the target distribution (first column) are shown in Figs. 2 and 3, respectively, using thresholds of 1 (corresponding to considering all measurements), 0.9, 0.8, and 0.7. The corresponding threshold levels used to obtain the same are given on top of each column of figures. The number of independent measurements are also listed in Table I. To effectively assess the results, a one-dimensional cross section in the source-detector fiber plane is plotted for varying thresholds and the same is given as Figs. 2(b) and 3(b) along with the independent measurements given in the parenthesis of the legend of these figures.

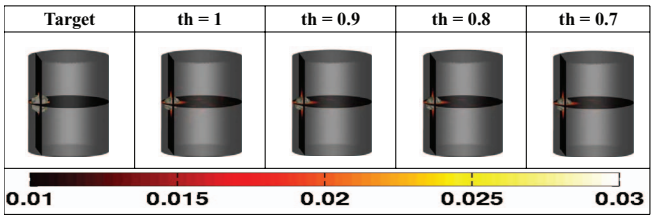
The absorption coefficient (μ_a) distribution for the case of experimental gelatin phantom data was reconstructed using a cylindrical mesh as outlined in Sec. IV.A. The target and the reconstructed distributions are shown in Fig. 4(a). The corresponding cross-sectional plots are given in Fig. 4(b). In order to effectively evaluate the usage of data resolution matrix for the choice of independent measurements, a random set of measurements (same number as that obtained by the analysis) were selected. The reconstructed μ_a distribution is shown in the last column of Fig. 4(a). The reconstruction based on random selection of 162 measurements has led to boundary artifacts and the contrast recovery was also poor [Fig. 4(b)]. Note that in this case, threshold of 0.9 alone is used, as lesser thresholds have resulted in unmeaningful results.

For the results presented in this work, the computation time per iteration for the reconstructed distributions of μ_a given in Figs. 2, 3, and 4(a) is reported in Table II as a function of threshold. The corresponding number of independent measurements for the chosen thresholds are given in Table I. The overhead time for the computation of data-resolution matrix and the associated analysis is also reported in the last column of Table II. The number of iterations taken to reach the stopping criterion is given in parenthesis below the computation time.

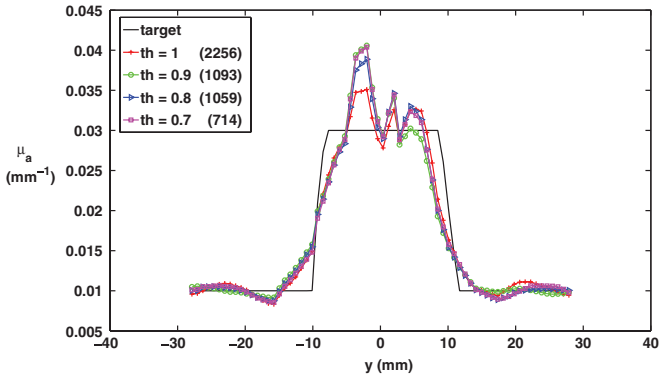
To assess the behavior of the iterative image reconstruction procedure by choosing only the independent measurements in contrast to all measurements for the case of Fig. 3, the L2-norm of the data-model misfit with the iteration number is plotted for different values of threshold in Fig. 5(a) as a function of iteration number. The difference in the mean between the target and reconstructed μ_a distribution is given in Fig. 5(c). The difference plot (in %) between considering only the independent measurements for a given threshold and all measurements (th = 1) for the plots in Figs. 5(a) and 5(c) are shown in Figs. 5(b) and 5(d), respectively. As the threshold is reduced, the total number of iterations required to converge were smaller but the data-model misfit and error in reconstructed μ_a is higher for lesser threshold compared to using all measurements (th = 1). Note that similar trends were observed in other cases considered here.

TABLE I. Comparison of the number of independent measurements obtained using the data resolution matrix [for different values of threshold (th)] and the number of useful measurements obtained using singular value analysis. The regularization parameter λ was 1 in all cases except for the experimental phantom case where it was 100. The imaging domain and the data-collection strategies are discussed in Ref. 8.

			No. of independent measurements		
Imaging domain	No. of measurements	No. of useful singular values	th = 0.9	th = 0.8	th = 0.7
Cylinder: 1 layer	240	89	111	106	57
Cylinder: 3 layers, in plane	720	231	331	320	198
Cylinder: 3 layers, out of plane	2256	282	1093	1059	714
Patient mesh	240	101	121	106	97
Experimental phantom	240	94	162	120	93



(a)



(b)

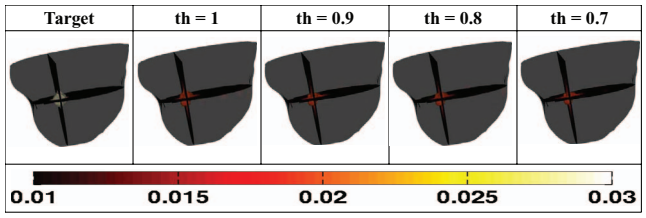
FIG. 2. (a) Reconstructed distributions of μ_a (in mm^{-1}) using cylindrical mesh (with three layers of data collection) along with the target distribution (given in the first column). The different values of threshold (th) used are specified on top of each reconstructed image. (b) The one-dimensional cross-sectional plots of the target and reconstructed μ_a distributions at the source plane for x = 30.

As the choice of threshold plays an important role in terms of number of independent measurements as well as the reconstructed image quality, the same is assessed for the patient mesh case (Fig. 3). In here, the λ was chosen to be 1 and threshold of 1 acted as the reference for computing the relative error percentage. The threshold values of 0.95, 0.9, 0.8, 0.7, and 0.6 were used in this study and the corresponding independent measurements were 128, 121, 106, 97, and 65, respectively. The relative error percentage in terms of reconstructed μ_a in the region of interest is plotted in Fig. 6 for the chosen threshold values. As threshold value is increased, the number of independent measurements increased in turn resulting in decreased error values (Fig. 6). The parameterization of this curve (also shown in Fig. 6) lead to the best fit as fourth order polynomial, indicating that increase in the error in the reconstructed absorption coefficient with respect to decrease in the threshold will follow fourth order polynomial decay.

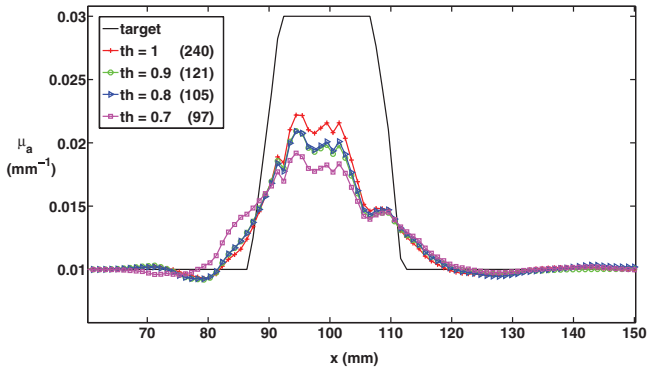
As the data-resolution matrix is computed using the regularization parameter (λ), the effect of the same on determining the number of independent measurements for threshold of 0.9 is studied for the same case as above (Fig. 3). The corresponding independent measurements for varying values of λ (0.01, 0.1, 1, 10, and 100) are plotted in Fig. 7. This plot shows that the number of independent measurements increase with decrease in λ value asserting the same trend as observed in Fig. 1(b).

VI. DISCUSSION

The optimization of data-collection strategies has been a topic of interest in diffuse optical imaging, where the emphasis was in knowing how many useful measurements can be obtained[4,7–9] among the total number of measurements for a given imaging problem. This analysis/optimization played an important role in instrumentation and lead to the existing imaging systems in this area.[20] This analysis has primarily relied on singular value analysis in providing the number of useful measurement information, but lacked the ability to provide specific information about whether a particular measurement is independent or not. This kind of



(a)



(b)

FIG. 3. (a) Reconstructed distributions of μ_a (in mm^{-1}) using patient mesh along with the target distribution (given in the first column). The different values of threshold (th) used are specified on top of each reconstructed image. (b) The one-dimensional cross-sectional plots of the target and reconstructed μ_a distributions at the source plane for y = −100.

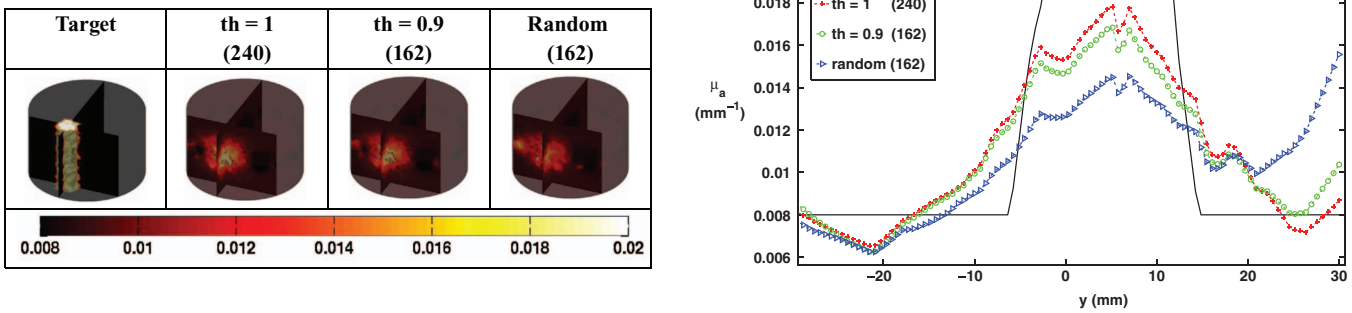


FIG. 4. (a) Reconstructed distributions of μ_a (in mm^{-1}) using cylindrical mesh along with the target distribution (given in the first column) using experimental gelatin phantom data. The different values of threshold (th) used are specified on top of each reconstructed image. (b) The one-dimensional cross-sectional plots of the target and reconstructed μ_a distributions at the source plane for x = −30.

specific information is extremely useful in not only optimizing or choosing a data-collection strategy, but even in measurement selection for a given data-collection strategy. Also, the specific information about the independent measurements avoids collecting redundant data in turn improving the throughput of the data-collection system. In this work, a data-resolution matrix based analysis to obtain the specific independent measurements is presented and has been shown that such selection does not compromise the reconstructed image quality.

The data-resolution matrix relies on the model being used, where the model includes both Jacobian and the regularization associated with the reconstruction scheme.[15,17] Even though the discussion here was limited to Levenberg–Marquardt regularization scheme, the trends and conclusions made in this work should hold for other regularization schemes as well. Also, the data-resolution matrix does not take into account the noise characteristics of the data and is only dependent on the model being used, making this scheme universal irrespective of the signal-to-noise ratio (SNR) in the data, where the SNR is highly dependent on the source strength and sensitivity of detectors.

The dependence of the data-resolution characteristics on the regularization is clearly evident in Fig. 1, where the importance of each data point is decreased as regularization parameter value is increased. Note that the regularization also known as fidelity parameter for the importance of data,[5,21] the higher regularization only meant higher the error associated with that data, in turn decreasing the importance of the data. Also, as the selection of independent measurements was based on how close (in magnitude) are the off-diagonal values in comparison to the diagonal value (refer to Algorithm 1), the parameter λ plays an important role in such selection. The same is clearly shown in Fig. 1(b), where the normalized (to the diagonal value) value of row number 136 is plotted as a function of regularization parameter (λ). The off-diagonal entries in comparison to diagonal entry (located at 136 with magnitude of 1) take higher value for a higher regularization parameter, showing that choice of higher regularization parameter will only yield lesser independent measurements. It should be noted that choosing a low value for regularization parameter in some cases may not improve the condition number of the problem, in turn, leading to nonunique solution [Eqs. (5) and (6)].

It is important to note that the regularization parameter (λ) for the LM scheme used in here stabilizes and improves the convergence of the solution obtained by either Eq. (5) or (6) and does not regularize the nonlinear least-squares as it is not part of the minimization scheme. Even though the

TABLE II. Computation time for the results presented in this work (Figs. 2–4) with varying values of threshold (th). The number of iterations to reach the stopping criterion is given in parenthesis below the computation time. The last column gives the overhead time for choosing the independent measurements based on the data-resolution matrix characteristics.

	Computation time per iteration (in s)				
Imaging domain	th = 1	th = 0.9	th = 0.8	th = 0.7	Overhead time (in s)
Cylinder: 3 layers, out of plane	476.04 (5)	462.12 (5)	461.90 (5)	459.45 (5)	37.98
Patient mesh	40.93 (19)	40.63 (15)	40.55 (11)	40.59 (10)	0.53
Experimental phantom	29.68 (13)	29.28 (12)	29.25 (8)	29.34 (8)	0.35

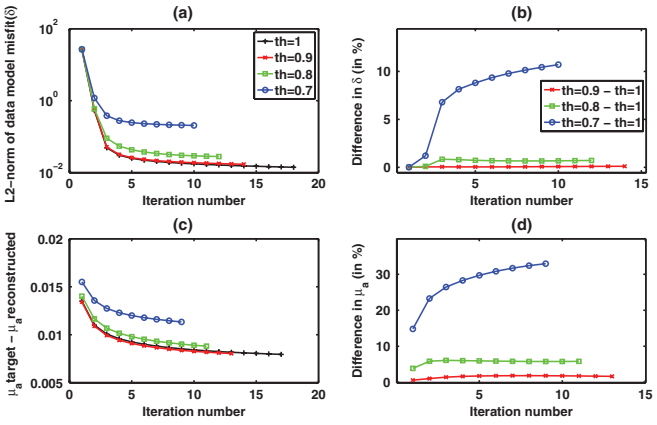


FIG. 5. Quantitative assessment of the data-model misfit and reconstruction parameter error versus the iteration number for the results corresponding to Fig. 3. (a) L2-norm of the data-model misfit for threshold values of 1, 0.9, 0.8, and 0.7. (b) Difference (in %) between the results given in (a) and th = 1. (c) The difference between mean of the target and reconstructed absorption coefficient for threshold values of 1, 0.9, 0.8, 0.7 in the region of interest. (d) Difference (in %) between the results given in (c) and th = 1. The legends of (a) and (c) and (b) and (d) are the same.

update equation [Eq. (5) or (6)] looks more close to the one obtained by Tikhonov minimization scheme,[13,15] it is equivalent to Gauss–Newton minimization scheme for small values of λ. The detailed discussion about LM minimization along with comparison to other schemes (including Tikhonov) is presented in Ref. 13. The initial choice of λ is dependent on the data-model misfit, with a typical choice being greater than or equal to L2 norm of the data-model misfit ($\geq||\delta||^2$),[13] the same has been utilized here. Also note that the optimization of data-collection strategies in the literature has been dealt using the sensitivity matrix obtained at the initial iteration (equivalently assuming the reconstruction procedure is linear),[4,7–9] the same has been adopted here. In our experience, the change in the number of independent measurements with the

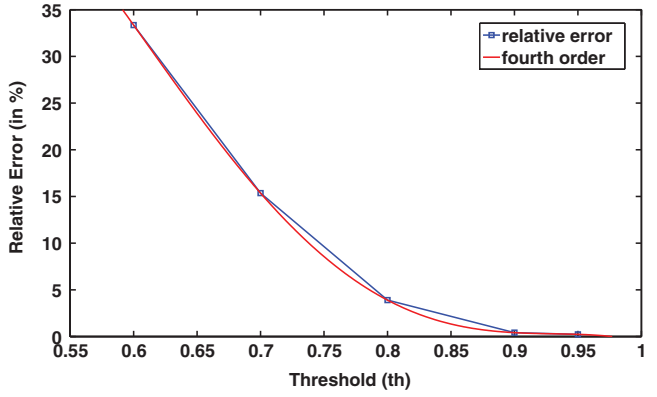


FIG. 6. Quantitative assessment in terms of relative error (RE) observed for the reconstructed absorption coefficient in patient mesh case (Fig. 3) using different threshold (th) values 0.95, 0.9, 0.8, 0.7, and 0.6 (corresponding independent measurements are given in Table I) with respect to the result obtained using threshold of 1 (all measurements). Parameterization of the observed trend follows fourth order polynomial decay, i.e., RE (in %) = $-2681.9\ \text{th}^4 + 8239\ \text{th}^3 - 9053.4\ \text{th}^2 + 4088\ \text{th} - 592.23$.

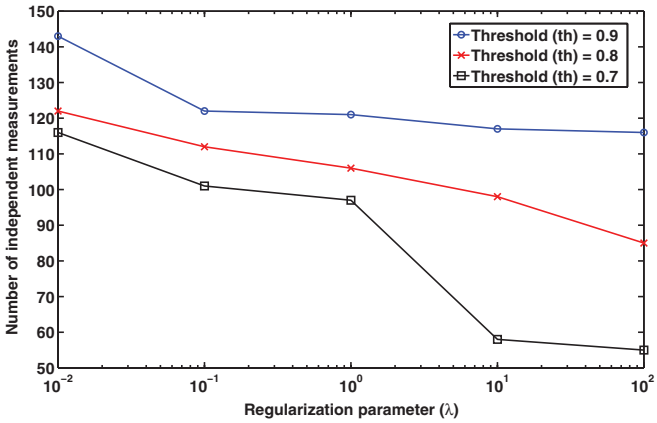


FIG. 7. Plot showing the effect of regularization parameter (λ) on the number of independent measurements for varying threshold values (0.9, 0.8, and 0.7) for the patient mesh case (Fig. 3).

iterations is very minimal (less than 1%) as the inherent model does not change with λ.

The inherent theme of this work is to show that choosing only these independent measurements based on data-resolution matrix does not compromise the reconstructed image quality and the same is shown in Figs. 2–4. As the choice of independent measurements is based on the threshold, the results indicate that a threshold of 0.9 has always lead to reconstructed image quality being on par with using all measurements (threshold of 1). The iterative behavior of reconstruction procedure for the choice of different thresholds for results presented in Fig. 3 are reported in Fig. 5. The threshold of 0.9 always resulted in error both in data-model misfit and reconstructed image to be less than 3% among all iterations, indicating that the selection of independent measurements for threshold of 0.9 did not compromise the reconstructed image quality in comparison to using all the measurements (threshold of 1). The same choice was implemented in the experimental phantom case, which lead to 162 independent measurements among a total of 240 measurements. On using 162 randomly selected measurements, the reconstructed image quality was much poorer compared to their counterparts including lot of artifacts that were formed at the boundary and also resulted in poor contrast recovery as evident in Fig. 4. There is a minimal difference (less than 5%) in the reconstructed μ_a values between using all measurements and threshold of 0.9, where the minimal observed difference could be because of improved SNR provided by using all measurements (where each independent measurement is a weighted average of all its dependent measurements).

A comparison with the traditional singular value analysis in assessing the usefulness of measurements has only revealed that the number of independent measurements are at least 20% higher. Also, it should be noted that the singular value analysis relies on the noise characteristics of data, where the normalized singular values above the noise floor (typically 1%) are chosen as useful measurements by analyzing the singular value spectrum of sensitivity matrix (**J**). Even though the singular value analysis was able to

qualitatively assess a particular data-collection strategy, the assessment is highly dependent on the noise characteristics of the imaging system and does not take into account the effect of regularization that is used in the reconstruction of optical images.

The reduction in total number of measurements that were used in the reconstruction procedure based on the threshold reduced the Jacobian dimensions, in turn reducing the memory used for storing the same. As only underdetermined problems were considered in this work, leading to usage of Eq. (6) for finding an update of μ_a, the associated computation time per iteration is dominated by the matrix multiplication [$\mathbf{JJ^T}$, O($NK * NN^2$)].[14] This lead to very little to no change in the experimentally observed computation time per iteration as given in Table II for the reconstruction cases discussed in this work. Moreover, after the choice of independent measurements using the uniform initial guess and initial value of regularization parameter, in the subsequent iterations, the direct computation of reduced Jacobian ($\mathbf{J_n}$) is performed.

The effect of threshold on the relative error in reconstructed μ_a (at final iteration) revealed that error percentage could be as high as 34% (Fig. 6) by choosing a low threshold value (also resulting in lowest number of independent measurements). The tradeoff between threshold value and error in reconstructed μ_a could be observed from Fig. 6, the lesser the threshold more is the error. The same trend is also observed in other cases. One of the desirable aim of this work is estimate the μ_a with lowest error possible, as the error percentage in the data is 1%, the desirable error should be in the same range. This lead to the natural choice of threshold being 0.9, for which case the relative error was 0.41%. Note that the choice of threshold 0.95 resulted in error percentage of 0.23%, but the number of independent measurements were higher in this case among all thresholds. The error percentage for threshold of 0.8 is 3.89%. As the decrease in error percentage was negligible (comparing 0.95 with 0.9 for threshold) and well below 1%, the choice of threshold 0.9 was used in the experimental phantom case. It should be noted that decrease in the threshold value (in turn decrease in the number of independent measurements) will always reduce the sensitivity of the imaging domain and is well studied in Ref. 8. The reduction in the total sensitivity of the imaging domain with having less number of independent measurements leads to amplification of error present in the data-space[8] and the same trend is observed here. The trend observed here may not be generalized to other system designs, as it is highly dependent on the system/sensitivity matrix ($\mathbf{J}$) characteristics defined by the imaging domain, data-collection strategy, and detector sensitivity. But optimization of the data-collection strategy could be easily achieved by the proposed method.

As the regularization parameter λ dictates the data-resolution characteristics (evident from Fig. 1), the lesser the value of λ, more were the resulting number of independent measurements (Fig. 7). There is a monotonical decrease in the number of independent measurements with increase in λ, for example, threshold of 0.9 the decrease was from 143 ($\lambda = 0.01$) to 116 ($\lambda = 100$), where as for threshold of 0.7, it was from 116 to 55. Note that threshold values of 0.8 and 0.7 resulted in error percentages higher than 1%, where as for threshold of 0.9 it was 0.41% (Fig. 6). Choice of λ lesser than 1 leads to higher number of independent measurements, but results in poor stability in getting the update [Eq. (6)] in subsequent iterations. Values higher than 1 for λ results in highly smoothened images. As LM optimization is known to be converging to local minima in case of nonconvex problems (like diffuse optical image reconstruction) depending on the initial guess and in turn λ value, the choice of λ always plays a pivotal role in determining the final solution to the reconstruction problem.[15] Given a choice of λ, the solution obtained by usage of all measurements compared to using only the independent measurements with threshold of 0.9 in the reconstruction procedure yielded error percentage below 1% (also evident from Fig. 6). Such determination of independent measurements could be obtained using Algorithm 1 and is computationally inexpensive. As the aim of this work is to provide those independent measurements for a given choice of λ, the effect of λ on the reconstructed image quality is beyond the scope of this work.

It should be also noted that even though in here only the results pertaining to transmission type data-collection geometries are considered, the same analysis and methodology are deployable in any data-collection geometry including reflection geometries[10] as it is only dependent on the model that is used in the image reconstruction procedure. The same could also be extended to methods that use boundary elements and other type of advanced techniques that solve diffusion equation.[22] The analysis is also universal with respect to reconstruction techniques that are in spirit similar to diffuse optical imaging, namely, electrical impedance tomography,[23] electrical capacitance tomography,[24] fluorescence optical tomography,[25] and bioluminescence tomography.[26]

VII. CONCLUSIONS

The diffuse optical tomographic reconstruction procedure uses advanced computational models to estimate the optical properties for a given set of boundary measurements.[20] The choice and optimization of a data-collection strategy has been shown to affect the reconstructed image quality and thereby the quantitation of physiological parameters associated with it. Even though traditional singular value analysis was able to provide qualitatively the number of useful measurements for a given data-collection strategy, its inability to provide the information about whether specific measurement is useful or not, makes the traditional data-collection strategy optimization schemes incapable of providing the next level of information. A model-based data-resolution matrix approach is introduced in this work, where the specific information to choose only the independent measurements among the measurements obtained from a data-collection strategy is evaluated using both numerical and experimental data. It is shown that such a choice of using only independent measurements among all measurements has not compromised the image quality, but in turn can reduce the data-collection time and total procedure time needed for the optical data collection. These type of

advanced optimization techniques will surely pave a way in designing highly optimized data collection systems, resulting in most useful data and avoiding the collection of redundant data.

ACKNOWLEDGMENTS

This work is supported by the Department of Atomic Energy, Government of India, through DAE young scientist research award (Sanction No. 2010/20/34/6/BRNS). The authors are also thankful to NIR imaging group of Dartmouth for providing necessary experimental data and finite element meshes that were needed to carry out this work.

APPENDIX: ALTERNATIVE FORM FOR THE UPDATE EQUATION

The alternative form for the update equation [Eq. (5)] can be derived starting from the *Sherman–Morrison–Woodbury identity*.[27–31] This identity is given by

$$(\mathbf{A}+\mathbf{C}^{\mathbf{T}}\mathbf{B}\mathbf{C})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{C}^{\mathbf{T}}(\mathbf{B}^{-1}+\mathbf{C}\mathbf{A}^{-1}\mathbf{C}^{\mathbf{T}})^{-1}\mathbf{C}\mathbf{A}^{-1}, \tag{A1}$$

where $\mathbf{A}$ and $\mathbf{B}$ are positive definite and square matrices. The matrix $\mathbf{C}$ could be a nonsquare and rank-deficient matrix.

Choosing $\mathbf{A}=\lambda\mathbf{I}$, $\mathbf{B}=\mathbf{I}$, and $\mathbf{C}=\mathbf{J}$ leads to

$$(\lambda\mathbf{I}+\mathbf{J}^{\mathbf{T}}\mathbf{J})^{-1}=(\lambda\mathbf{I})^{-1}-(\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}}(\mathbf{I}+\mathbf{J}(\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}})^{-1}\mathbf{J}(\lambda\mathbf{I})^{-1}. \tag{A2}$$

Simplification on the right-hand side (as λ is a scalar) leads to

$$(\lambda\mathbf{I}+\mathbf{J}^{\mathbf{T}}\mathbf{J})^{-1}=(1/\lambda)\mathbf{I}-(1/\lambda)\mathbf{J}^{\mathbf{T}}(\lambda\mathbf{I}+\mathbf{J}\mathbf{J}^{\mathbf{T}})^{-1}\mathbf{J}. \tag{A3}$$

Multiplying by $\mathbf{J}^{\mathbf{T}}$ on both sides and regrouping the terms gives rise to

$$(\mathbf{J}^{\mathbf{T}}\mathbf{J}+\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}}=(1/\lambda)\mathbf{J}^{\mathbf{T}}\left(\mathbf{I}-(\lambda\mathbf{I}+\mathbf{J}\mathbf{J}^{\mathbf{T}})^{-1}\mathbf{J}\mathbf{J}^{\mathbf{T}}\right). \tag{A4}$$

Substitution of $\mathbf{I}=(\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})^{-1}(\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})$ and simplifying leads to

$$(\mathbf{J}^{\mathbf{T}}\mathbf{J}+\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}}=(1/\lambda)\mathbf{J}^{\mathbf{T}}(\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})^{-1}\left((\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})-\mathbf{J}\mathbf{J}^{\mathbf{T}}\right), \tag{A5}$$

simplifying further leads to

$$(\mathbf{J}^{\mathbf{T}}\mathbf{J}+\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}}=(1/\lambda)\mathbf{J}^{\mathbf{T}}(\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})^{-1}(\lambda)\mathbf{I} \tag{A6}$$

finally resulting in

$$(\mathbf{J}^{\mathbf{T}}\mathbf{J}+\lambda\mathbf{I})^{-1}\mathbf{J}^{\mathbf{T}}=\mathbf{J}^{\mathbf{T}}(\mathbf{J}\mathbf{J}^{\mathbf{T}}+\lambda\mathbf{I})^{-1}. \tag{A7}$$

This proves that Eqs. (5) and (6) are equivalent. Note that the same relation is derived starting from the generalized least-squares update equation alternative form in the appendix of Ref. 14, in here it is derived starting from the *Sherman–Morrison–Woodbury identity*.

[a]Electronic mail: phani@serc.iisc.in; Telephone: +91-80-2293 2496; Fax: +91-80-2360 2648.

[1]S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, "Interpreting hemoglobin and water concentration, oxygen saturation and scattering measured *in vivo* by near-infrared breast tomography," Proc. Natl. Acad. Sci. U.S.A. **100**, 12349–12354 (2003).
[2]D. R. Leff, O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Patten, J. C. Hebden, G. Z. Yang, and A. Darzi, "Diffuse optical imaging of the healthy and diseased breast: A systematic review," Breast Cancer Res. Treat. **108**, 9–22 (2008).
[3]B. W. Zeff, B. R. White, H. Dehghani, B. L. Schlaggar, and J. P. Culver, "Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography," Proc. Natl. Acad. Sci. U.S.A. **104**, 12169–12174 (2007).
[4]H. Xu, H. Dehghani, B. W. Pogue, R. Springett, K. D. Paulsen, and J. F. Dunn, "Near-infrared imaging in the small animal brain: Optimization of fiber positions," J. Biomed. Opt. **8**, 102–110 (2003).
[5]S. R. Arridge, "Optical tomography in medical imaging," Inv. Problems **15**, R41–R93 (1999).
[6]D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," IEEE Signal Process. Mag. **18**, 57–75 (2001).
[7]J. P. Culver, V. Ntziachristos, M. J. Holboke, and A. G. Yodh, "Optimization of optode arrangements for diffuse optical tomography: A singular-value analysis," Opt. Lett. **26**, 701–703 (2001).
[8]P. K. Yalavarthy, H. Dehghani, B. W. Pogue, and K. D. Paulsen, "Critical computational aspects of near infrared circular tomographic imaging: Analysis of measurement number, mesh resolution and reconstruction basis," Opt. Express **14**, 6113–6127 (2006).
[9]L. Chen and N. Chen, "Optimization of source and detector configurations based on Cramer-Rao lower bound analysis," J. Biomed. Opt. **16**, 035001 (2011).
[10]H. Dehghani, Brian R. White, Benjamin W. Zeff, Andrew Tizzard, and Joseph P. Culver, "Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography," Appl. Opt. **48**, D137–D143 (2009).
[11]H. Jiang, K. D. Paulsen, U. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency domain data: simulations and experiments," J. Opt. Soc. Am. A **13**, 253–266 (1996).
[12]M. Schweiger, S. R. Arridge, M. Hiroaka, and D. T. Delpy, "The finite element model for the propagation of light in scattering media: Boundary and source conditions," Med. Phys. **22**, 1779–1792 (1995).
[13]P. K. Yalavarthy, B. W. Pogue, H. Dehghani, and K. D. Paulsen, "Weight-matrix structured regularization provides optimal generalized least-squares estimate in dffuse optical tomography," Med. Phys. **34**, 2085–2098 (2007).
[14]P. K. Yalavarthy, D. R. Lynch, B. W. Pogue, H. Dehghani, and K. D. Paulsen, "Implementation of a computationally efficient least-squares algorithm for highly under-determined three-dimensional diffuse optical tomography problems," Med. Phys. **35**, 1682–1697 (2008).
[15]R. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems* (Elsevier, 2005).
[16]T. O. Mcbride, B. W. Pogue, S. Jiang, U. L. Osterberg, and K. D. Paulsen, "A parallel-detection frequency-domain near-infrared tomography system for hemoglobin imaging of the breast in vivo," Rev. Sci. Instrum. **72**, 1817–1824 (2001).
[17]M. S. Zhdanov, *Geophysical Inverse Theory and Regularization Problems*, 1st ed. (Elsevier Science, New York, 2002).
[18]B. W. Pogue, K. D. Paulsen, H. Kaufman, and C. Abele, "Calibration of near infrared frequency-domain tissue spectroscopy for absolute absorption coefficient quantitation in neonatal head-simulating phantoms," J. Biomed. Opt. **5**, 182–193 (2000).
[19]B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," J. Biomed. Opt. **11**, 041102 (2006).
[20]A. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical tomography," Phys. Med. Biol. **50**, R1–R43 (2005).
[21]H. Dehghani, S. Srinivasan, B. W. Pogue, and A. Gibson, "Numerical modelling and image reconstruction in diffuse optical tomography," Philos. Trans. R. Soc. London, Ser. A **367**, 3073–3093 (2009).
[22]S. Srinivasan, B. W. Pogue, C. M. Carpenter, P. K. Yalavarthy, and K. D. Paulsen, "A boundary element approach for image-guided near-infrared absorption and scatter estimation," Med. Phys. **34**, 4545–4557 (2007).

[23]H. Dehghani and M. Soleimani, "Numerical modelling errors in electrical impedance tomography," Phys. Meas. **28**, S45–S55 (2007).
[24]M. Soleimani, P. K. Yalavarthy, and H. Dehghani, "Helmholtz-type regularization method for permittivity reconstruction using experimental phantom data of electrical capacitance tomography," IEEE Trans. Instrum. Meas. **59**, 78–83 (2010).
[25]S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen, "Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization," Opt. Express **15**, 4066–4082 (2007).
[26]H. Dehghani, S. Davis, S. Jiang, B. Pogue, K. Paulsen, and M. Patterson, "Spectrally resolved bioluminescence optical tomography," Opt. Lett. **31**, 365–367 (2006).
[27]J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," Ann. Math. Stat. **20**, 621 (1949).
[28]J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," Ann. Math. Stat. **21**, 124–127 (1950).
[29]M. Wooodbury, *Inverting Modified Matrices*, Memorandum Report 42, Statistical Research Group, Princeton, 1950.
[30]G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (The John Hopkins University Press, Baltimore, 1996), p. 50.
[31]J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations* (Wiley, New York, 1968), p. 132.