Implementation of a computationally efficient least-squares algorithm for highly under-determined three-dimensional diffuse optical tomography problems

Phaneendra K. Yalavarthy,[a)] Daniel R. Lynch, and Brian W. Pogue[b)]
Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755

Hamid Dehghani
Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755
and School of Physics, University of Exeter, Stocker Road, Exeter, EX4 4QL, United Kingdom

Keith D. Paulsen
Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755

Three-dimensional (3D) diffuse optical tomography is known to be a nonlinear, ill-posed and sometimes under-determined problem, where regularization is added to the minimization to allow convergence to a unique solution. In this work, a generalized least-squares (GLS) minimization method was implemented, which employs weight matrices for both data-model misfit and optical properties to include their variances and covariances, using a computationally efficient scheme. This allows inversion of a matrix that is of a dimension dictated by the number of measurements, instead of by the number of imaging parameters. This increases the computation speed up to four times per iteration in most of the under-determined 3D imaging problems. An analytic derivation, using the Sherman–Morrison–Woodbury identity, is shown for this efficient alternative form and it is proven to be equivalent, not only analytically, but also numerically. Equivalent alternative forms for other minimization methods, like Levenberg–Marquardt (LM) and Tikhonov, are also derived. Three-dimensional reconstruction results indicate that the poor recovery of quantitatively accurate values in 3D optical images can also be a characteristic of the reconstruction algorithm, along with the target size. Interestingly, usage of GLS reconstruction methods reduces error in the periphery of the image, as expected, and improves by 20% the ability to quantify local interior regions in terms of the recovered optical contrast, as compared to LM methods. Characterization of detector photo-multiplier tubes noise has enabled the use of the GLS method for reconstructing experimental data and showed a promise for better quantification of target in 3D optical imaging. Use of these new alternative forms becomes effective when the ratio of the number of imaging property parameters exceeds the number of measurements by a factor greater than 2. © *2008 American Association of Physicists in Medicine*. [DOI: 10.1118/1.2889778]

Key words: near infrared, diffuse optical tomography, three-dimensional imaging, image reconstruction, inverse problems, least-squares minimization

I. INTRODUCTION

Diffuse optical tomography (DOT) uses near infrared wavelengths (600–1000 nm) to obtain optical absorption and scattering images for characterizing functional properties of the tissue under investigation.[1–4] The most important step in forming these images is solving the inverse problem, i.e., estimating the optical properties by matching the experimental data with modeled results in the least-squares sense.[3,5,6] This problem is typically ill posed and ill determined depending on the noise in the data, the number of measurements, and the dimensions of the parameter space.[6] Even though light travels in three dimensions (3D), most of the numerical models reported in the literature have been two dimensional (2D) because of computational considerations. Moreover, the 3D DOT imaging problem is more under-determined relative to the 2D case and has been found to generate poor quantitative estimates of the optical properties when compared to 2D results.[7–15] Several methods have appeared in the literature describing efficient 3D computations,[11,14,16] but no unified approach to the problem has been discussed. Recently, a generalized least-squares (GLS) minimization scheme was presented for 2D DOT image reconstruction.[17] This article reports a computationally efficient approach for implementing GLS minimization in

3D which shows an improvement in the quantification of optical properties relative to earlier studies. Even though the focus is on GLS implementation, equivalent forms for other methods are also presented in light of the GLS framework.

The inverse problem in DOT is solved by minimizing the objective function (Ω) over the range of optical properties (μ). Methods based on gradient optimization, which do not require an explicit inversion of the Hessian matrix (here, Hessian approximates the second derivative; this can be some form equivalent to $\mathbf{J}^T\mathbf{J}$ or $\mathbf{J}\mathbf{J}^T$, in general some form of $\mathbf{J}^T\mathbf{J}$, $\mathbf{J}$ being the Jacobian) are known to be computationally efficient.[18,19] But these methods require an optimization scheme, which can be thought of as an inner iteration, for choosing the step size, and are not as straightforward as a direct inversion of the matrix. Alternatively, the full-Newton methods require calculation of the Jacobian ($\mathbf{J}$), the forward data, and inversion of the dense Hessian matrix at each iteration. Because full-Newton methods are relatively easy to implement, they are widely used for DOT image reconstruction even though they require large matrix inversions at every iteration. Thus, while the full-Newton method is ideal for small problems, it rapidly becomes intractable for larger domains, such as those encountered in 3D imaging problems. This manuscript presents a formal approach, using the Sherman–Morrison–Woodbury identity, to construct a more efficient alternative form of update equations of GLS and Levenberg–Marquardt (LM) minimization schemes, which generates a Hessian matrix to invert. The dimension of this Hessian is dictated by the number of measurements, rather than the number of parameter estimates which can be considerably lower for highly under-determined problems, and therefore, much more efficient computationally. This equivalent form is also shown for other common minimization methods, namely Tikhonov.

The later part of this article describes a way to characterize the systematic noise using a simple analytical formula, when the photomultiplier tube (PMT) is used as a detector. Characterizing noise behavior of the experimental data leads to the use of the GLS technique in the experimental data case and it is also shown that usage of noise characteristics will lead to better quality and quantification of target in a experimental test case.

II. DIFFUSE OPTICAL TOMOGRAPHY: FORWARD PROBLEM

Near-infrared (NIR) light propagation in a biological tissue like breast can be modeled using the diffusion equation (DE)[6,20] which in the frequency domain becomes

$$-\nabla \cdot D(r) \nabla \Phi(r,\omega) + (\mu_a(r) + i\omega/c)\Phi(r,\omega) = Q_o(r,\omega), \tag{1}$$

where the optical diffusion and absorption coefficients are given by $D(r)$ and $\mu_a(r)$, respectively. The light source, represented by $Q_o(r,\omega)$, is modeled as isotropic. $\Phi(r,\omega)$ is the photon fluence rate at a given position r. The light modulation frequency is denoted by ω, where $\omega = 2\pi f$, (here $f = 100$ MHz). The velocity of light in tissue is represented by c, which is assumed to be constant. Note that

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

where $\mu_s'(r)$ is the reduced scattering coefficient which is defined as $\mu_s' = \mu_s(1-g)$; μ_s is the scattering coefficient and g is the anisotropy factor. The finite element method (FEM) is used to solve Eq. (1) to generate modeled data ($G(\mu)$) for a given distribution of optical properties (μ),[11,13,21] where $\mu = [\mu_s'(r); \mu_a(r)]$. A Type-III boundary condition is employed to account for the refractive-index mismatch at the boundary.[22] Under the Rytov approximation, the data (y) becomes the natural logarithm of the amplitude (A) and phase (θ) of the frequency-domain signal; $y = [lnA; \theta]$.

III. DIFFUSE OPTICAL TOMOGRAPHY: INVERSE PROBLEM

III.A. Levenberg–Marquardt (LM) minimization

The most-common approach for solving the inverse problem in DOT is LM minimization.[1,6,11,13,17,20,23,24] A detailed discussion of this method is available in Ref. 17 and it is only briefly reviewed here.

The objective function[25,26] for this approach is defined as

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

where y is the experimental data and $G(\mu)$ is the modeled response. Minimization of this objective function with respect to μ is achieved by setting the first-order derivative equal to zero

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

where δ is the data-model misfit, $\delta = y - G(\mu)$, and $\mathbf{J}$ represents the Jacobian ($\mathbf{J} = \partial G(\mu)/\partial \mu$). Due to the ill-conditioned nature of the problem, the update equation for the optical properties at iteration "i" is regularized to be

$$\Delta\mu_i = [\mathbf{J}^T\mathbf{J} + \alpha\mathbf{I}]^{-1}\mathbf{J}^T\delta_{i-1} \tag{5}$$

or equivalently (See Appendix 2)

$$\Delta\mu_i = \mathbf{J}^T[\mathbf{J}\mathbf{J}^T + \alpha\mathbf{I}]^{-1}\delta_{i-1}, \tag{6}$$

where $\Delta\mu_i$ represents the update of the optical property parameters at the ith step; α is the regularization parameter, which monotonically decreases with increasing iteration

(always > 0).[26] In this approach (Eq. (5)), the Jacobian is normalized by its optical properties. Moreover, α is chosen empirically (it typically starts at 10 and reduced by a factor of $10^{0.25}$ at every following iteration after being multiplied by the maximum of the diagonal values of $\mathbf{JJ}^T$.[17,27]). The iterative procedure is stopped when the L2 norm of the data-model misfit (δ) does not improve (in our experience, by more than 1% because beyond these values the LM procedure can become unstable[17]).

Even though LM minimization or its modified versions have been used successfully for DOT image reconstruction,[1,6,11,13,17,20,23,24,27] the final image depends on the choice of α due to the ill-conditioned nature of the problem. Moreover, the approach ignores the noise characteristics of the data and optical properties. A more systematic and generalized method for image reconstruction can be based on GLS minimization. The GLS scheme is discussed extensively in Ref. 17 and is only briefly reviewed here.

III.B. Generalized least squares (GLS) minimization

In GLS, the objective function is given by[17,28]

$$\Omega = (y - G(\mu))^T \mathbf{W}_\delta (y - G(\mu)) + (\mu + \mu_0)^T \mathbf{W}_{\mu-\mu_0} (\mu - \mu_0), \tag{7}$$

where $\mathbf{W}_\delta$ and $\mathbf{W}_{\mu-\mu_0}$ are the weight matrices for the data-model misfit (δ) and optical properties ($\mu - \mu_0$), respectively. Note that $\mathbf{W}_\delta = (\mathbf{C}_\delta)^{-1}$, where $\mathbf{C}$ represents the covariance matrix and similarly $\mathbf{W}_{\mu-\mu_0} = (\mathbf{C}_{\mu-\mu_0})^{-1}$ (see Appendix A-4 of Ref. 28). These weight matrices; are symmetric and positive definite (because they are, inverses of covariance matrices). No regularization parameter is involved because the weight matrices include the noise characteristics of the experimental data and optical properties.[17] Similarly to the LM approach, minimization of Ω (Eq. (7)) is accomplished by setting the first derivative of Ω with respect to μ equal to zero:

$$\frac{\partial \Omega}{\partial \mu} = \mathbf{J}^T \mathbf{W}_\delta \delta - \mathbf{W}_{\mu-\mu_0} (\mu - \mu_0) = 0. \tag{8}$$

Linearizing the problem leads to the iterative update equation (for ith iteration)[17]

$$\Delta \mu_i = [\mathbf{J}^T \mathbf{W}_\delta \mathbf{J} + \mathbf{W}_{\mu-\mu_0}]^{-1} (\mathbf{J}^T \mathbf{W}_\delta \delta_{i-1} - \mathbf{W}_{\mu-\mu_0} (\mu_{i-1} - \mu_0)). \tag{9}$$

Explicit definitions of the weight matrices ($\mathbf{W}_\delta$ and $\mathbf{W}_{\mu-\mu_0}$) are given in Ref. 17. Although any number of forms for $\mathbf{W}_{\mu-\mu_0}$ can exist, only one is considered here, specifically, where the covariance matrix is defined as[17]

$$[\mathbf{C}_{\mu-\mu_0}]_{ij} = (\sigma_{\mu-\mu_0})^2 \left(1 + \frac{r_{ij}}{l}\right) e^{-(r_{ij}/l)} \tag{10}$$

with l being the correlation length (here $l=15$ mm) and r_{ij} being the distance between the FEM nodes i and j ($\sigma_{\mu-\mu_0})^2$ is the expected variance of $\mu - \mu_0$. Strategies for calculating the expected variances are given in Ref. 17. In this work, the expected variance is determined from the prior knowledge

that the expected contrast between tumor and normal tissue is about 50%–400%. To demonstrate the robustness of the GLS reconstruction procedure, for the results discussed here, the variance was chosen to be $(4 * \mu)^2$. Both weight matrices, $\mathbf{W}_\delta$ and $\mathbf{W}_{\mu-\mu_0}$, are computed before the reconstruction procedure begins, whereas the Jacobian ($\mathbf{J}$), and modeled data, $G(\mu)$, are calculated at each iteration. The iterative procedure is stopped when the L2 norm of the data-model misfit (δ) does not improve by more than 0.001%. Beyond these values, the round-off error dominates.

III.B.1. GLS implementation

The parameters recovered in the case of this GLS scheme are $(\mu_s'; \mu_a)$, which is different from some previous approaches that estimate $(D; \mu_a)$. The later case has a mismatch because the units of D are mm whereas those of μ_a are mm^{-1}. In its implementation typically the whole equation is normalized by the optical properties (outlined in Ref. 17) which becomes computationally intensive especially for GLS in 3D because the update equation must be left and right multiplied by the optical properties at every iteration. Here, the GLS problem was reformulated in terms of $(\mu_s'; \mu_a)$, so that both parameters have the same units (mm^{-1}). While this is a relatively minor alteration in the form of the algorithm, it has important implications for the computational time required for matrix preconditioning.

A simple transformation converts the diffusion part of the Jacobian $(\partial G(\mu)/\partial D)$ to its scattering component $(\partial G(\mu)/\partial \mu_s')$:

$$\frac{\partial G(\mu)}{\partial \mu_s'} = \frac{\partial G(\mu)}{\partial D} \frac{\partial D}{\partial \mu_s'}. \tag{11}$$

Using Eq. (2)

$$\frac{\partial D}{\partial \mu_s'} = \frac{1}{3} \left(\frac{-1}{[(\mu_a + \mu_s')]^2}\right) = -3D^2 \tag{12}$$

and substituting Eq. (12) in Eq. (11), leads to

$$\frac{\partial G(\mu)}{\partial \mu_s'} = \frac{\partial G(\mu)}{\partial D} (-3D^2). \tag{13}$$

After this transformation (Eq. (13)), the Jacobian ($\mathbf{J}$) has the form

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial lnA_1}{\partial \mu'_{s1}} & \dfrac{\partial lnA_1}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial lnA_1}{\partial \mu'_{sNN}} & \dfrac{\partial lnA_1}{\partial \mu_{a1}} & \dfrac{\partial lnA_1}{\partial \mu_{a2}} & \cdots & \dfrac{\partial lnA_1}{\partial \mu_{aNN}} \\[2mm] \dfrac{\partial \theta_1}{\partial \mu'_{s1}} & \dfrac{\partial \theta_1}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial \theta_1}{\partial \mu'_{sNN}} & \dfrac{\partial \theta_1}{\partial \mu_{a1}} & \dfrac{\partial \theta_1}{\partial \mu_{a2}} & \cdots & \dfrac{\partial \theta_1}{\partial \mu_{aNN}} \\[2mm] \dfrac{\partial lnA_2}{\partial \mu'_{s1}} & \dfrac{\partial lnA_2}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial lnA_2}{\partial \mu'_{sNN}} & \dfrac{\partial lnA_2}{\partial \mu_{a1}} & \dfrac{\partial lnA_2}{\partial \mu_{a2}} & \cdots & \dfrac{\partial lnA_2}{\partial \mu_{aNN}} \\[2mm] \dfrac{\partial \theta_2}{\partial \mu'_{s1}} & \dfrac{\partial \theta_2}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial \theta_2}{\partial \mu'_{sNN}} & \dfrac{\partial \theta_2}{\partial \mu_{a1}} & \dfrac{\partial \theta_2}{\partial \mu_{a2}} & \cdots & \dfrac{\partial \theta_2}{\partial \mu_{aNN}} \\[2mm] \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\[2mm] \dfrac{\partial lnA_{NM}}{\partial \mu'_{s1}} & \dfrac{\partial lnA_{NM}}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial lnA_{NM}}{\partial \mu'_{sNN}} & \dfrac{\partial lnA_{NM}}{\partial \mu_{a1}} & \dfrac{\partial lnA_{NM}}{\partial \mu_{a2}} & \cdots & \dfrac{\partial lnA_{NM}}{\partial \mu_{aNN}} \\[2mm] \dfrac{\partial \theta_{NM}}{\partial \mu'_{s1}} & \dfrac{\partial \theta_{NM}}{\partial \mu'_{s2}} & \cdots & \dfrac{\partial \theta_{NM}}{\partial \mu'_{sNN}} & \dfrac{\partial \theta_{NM}}{\partial \mu_{a1}} & \dfrac{\partial \theta_{NM}}{\partial \mu_{a2}} & \cdots & \dfrac{\partial \theta_{NM}}{\partial \mu_{aNN}} \end{bmatrix}, \tag{14}$$

where NM and NN represent the number of measurements and number of property parameters associated with the FEM mesh, respectively. The implementation of the GLS update equation (Eq. (9)) requires assembly of the weight matrix ($\mathbf{W}_{\mu-\mu_0}$) for simultaneous reconstruction of μ'_s and μ_a and is accomplished by writing Eq. (9) in block matrix form

$$\left\{ \begin{bmatrix} H_{\mu'_s 2} & H_{\mu'_s \mu_a} \\ H_{\mu'_s \mu_a} & H_{\mu_a^2} \end{bmatrix} + \begin{bmatrix} W_{\mu'_s - \mu'_{s0}} & 0 \\ 0 & W_{\mu_a - \mu_{a0}} \end{bmatrix} \right\} \begin{bmatrix} \Delta \mu'_{si} \\ \Delta \mu_{ai} \end{bmatrix}$$
$$= \begin{bmatrix} (J^T)_{\mu'_s} W_\delta \delta_{i-1} \\ (J^T)_{\mu_a} W_\delta \delta_{i-1} \end{bmatrix} - \begin{bmatrix} W_{\mu'_s - \mu'_{s0}}(\mu_{s'_{i-1}} - \mu'_{s0}) \\ W_{\mu_a - \mu_{a0}}(\mu_{a_{i-1}} - \mu_{a0}) \end{bmatrix}, \tag{15}$$

where H represents the Hessian matrix ($\mathbf{J}^T \mathbf{W}_\delta \mathbf{J}$). Here, the cross terms in the weight matrix ($\mathbf{W}_{\mu-\mu_0}$) are zero because μ'_s and μ_a are independent parameters in the estimation procedure. Note that the dimensions of the matrices in Eq. (9) are: $\mathbf{J}: 2NM \times 2NN$, $\mathbf{W}_{\mu-\mu_0}: 2NN \times 2NN$, $\mathbf{W}_\delta: 2NM \times 2NM$, $\delta = 2NM \times 1$, and $\Delta \mu = 2NN \times 1$. Most 3D-D0T problems are ill determined, i.e., $NM \ll NN$.

Computing an update of the optical properties ($\Delta \mu_i$, from Eq. (9)) requires an inversion (or its equivalent) of a large matrix with dimensions $2NN \times 2NN$. Inverting a matrix of dimension $N \times N$ typically requires an order of N^3 operations and N^2 memory.[28] Hence, any gain in reducing the dimensionality of the matrix to be inverted will reduce the computation time cubically and the memory requirement quadratically. An alternative form of Eq. (9), which requires few operations, is

$$\Delta \mu_i = [\mathbf{I} - \mathbf{C}_{\mu-\mu_0} \mathbf{J}^T (\mathbf{J} \mathbf{C}_{\mu-\mu_0} \mathbf{J}^T + \mathbf{C}_\delta)^{-1} \mathbf{J}]$$
$$\times \{ \mathbf{C}_{\mu-\mu_0} \mathbf{J}^T \mathbf{W}_\delta \delta_{i-1} - (\mu_{i-1} - \mu_0) \}. \tag{16}$$

Full derivation of this alternative form is given in the Appendix 1, along with the equivalent expressions for other minimization methods. Equation (16) requires an inversion of a matrix with dimensions $2NM \times 2NM$ (same is true for Eq. (6)).

Note that the covariance/weight matrices are calculated before the start of the iterative procedure and are used throughout the iteration. For nodes where the sensitivity (column sum of Jacobian) fell below 1% of the maximum sensitivity, the expected variance of the optical properties were chosen to be 1% of background μ (in Eq. (10)).

IV. SIMULATION STUDIES: THREE-DIMENSIONAL TEST PROBLEM

For all numerical experiments discussed here, the imaging domain was chosen to be cylindrical (as shown in Fig. 1) with a diameter of 84 mm and height of 109 mm. The background optical properties were $\mu_a = 0.01$ mm^{-1} and $\mu'_s = 1.0$ mm^{-1}. Two meshes were used: (1) a cylinder consisting of 21 440 nodes corresponding to 110 483 linear tetrahedral
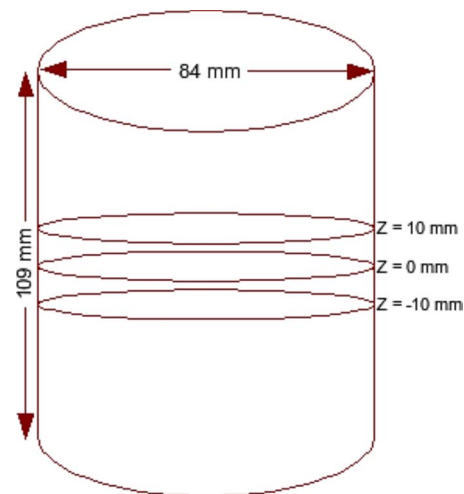

FIG. 1. Schematic diagram of the three-dimensional cylindrical imaging domain.

elements for the forward model and (2) a cylinder having 9211 nodes corresponding to 45 980 linear tetrahedral elements for the reconstruction. The data-collection geometry consisted of 48 fibers that were arranged in a circular, equally spaced fashion in three layers spaced 10 mm apart (Fig. 1), with 16 fibers per plane. One fiber was used at a time as the source while the fibers in the same "source fiber plane" were used as detectors to generate 720 $(3 \times 16 \times 15)$ measurement locations or a total of 1440 values (720 *lnA* data points and 720 θ data points). The sources were modeled as a Gaussian profile with a full width at half maximum of 3 mm to represent the distribution used in an experimental setup.[29] The source was also placed one mean transport length inside the boundary.

Both spherical and cylindrical objects were considered as targets. The cylindrical target had a contrast of 2:1 with the background in both μ_a and μ_s' and a diameter of 15 mm. It extended in Z direction throughout the domain (height of 109 mm) and was placed at the center (at (0,0), first row of Fig. 5) and near the boundary (at (30,0)). Two-dimensional cross sections of both reconstructed and actual 3D volumes, displayed in increments of 5 mm spanning from $z=$ −25 mm to $z=25$ mm (from left-hand side to right-hand side), are shown in Figs. 3, 5, and 6. Cross section below z =−25 mm and beyond $z=25$ mm did not show any deviation from the starting values of iterations as the sensitivity in this region is almost negligible compared to the rest of the domain, so these cross sections are omitted for display purposes. Measurements with a noise level of 1% were assumed as the experimental data (y) in most of the cases discussed here. The noise variance also used with GLS reconstruction algorithm.[17] The background optical properties were selected as starting values for the iterative image reconstruction procedures discussed in Sec. III. All computations were carried out on a Linux work station with an AMD Dual Core Opteron 280 processor (2.2 GHz) with 8 GB of random access memory (RAM).

V. PHANTOM STUDIES

V.A. Data variance estimation

Use of weight matrices in the GLS scheme ($\mathbf{W}_\delta$ in Eq. (9)) requires an estimation of data variance, which requires experimental characterization of the expected values, prior to patient/phantom imaging. This was achieved by tracking the detected voltage measured at the photomultiplier tube (PMT). PMTs are used as a detectors in the experimental system at Dartmouth, details of the experimental system are given in Ref. 29. Note that this characterization includes only systematic errors associated with low signal levels, but errors due to poor fiber-tissue coupling are not accounted for in this model.

Starting from the assumption that the detected signal at the PMT in diffuse optical imaging is shot noise limited leads to

$$\sigma(N) = \sqrt{N}, \tag{17}$$

where σ is the standard deviation of N and N is the number of photons reaching the PMT. The voltage (V; representing the detected ac intensity signal[29]) measured at the PMT is directly proportional to N, which also implies that measured amplitude (A) of the detected frequency domain signal (y) is proportional to this voltage (V). This is written as

$$A \propto V \Rightarrow A = kV$$

$$\sigma(A) = k\sigma(V). \tag{18}$$

Here, k acts as a proportionality constant. In the reconstruction procedure, the Rytov approximation is used, leading to data being represented as *lnA* rather than A. If $f(x)$ is a function of x and is continuous and differentiate, then

$$\sigma(f(x)) = \frac{\partial f(x)}{\partial x}\sigma(x). \tag{19}$$

Similar to the previous equation (Eq. (19)), writing the standard deviation of *InA* leads to

$$\sigma(lnA) = \frac{1}{A}\sigma(A). \tag{20}$$

now using Eq. (18) leads to

$$\sigma(lnA) = \frac{\sigma(V)}{V}. \tag{21}$$

From the above equation (Eq. (21)), it can be concluded that the variance in data (σ^2) can be known by measuring the deviation in the PMT voltage (V).

To measure the deviation in the measured signal, a series of light signal measurements were taken through homogeneous intra-lipid phantom experiments which were conducted with increasing levels of blood (HbT) concentration, varying from 7.3 to 36 μM, leading to a decrease in the measured PMT voltage. To achieve this, the gain of the PMT was kept at 0.9. A concise discussion of the PMT gain setting in the system characterization is given in Ref. 29. A single source and the farthest detector was used for these transmission measurements. For every concentration, 200 data points were collected to estimate the deviation in the measured voltage using the same gain settings. The approach for the characterization is similar to the one described in Ref. 29, except the raw detected voltage was used here for estimation of the error (or deviation σ). Note also that two sets of diameters, 56 and 84 mm, were used to get the voltage in the range of 0–1 V. This was repeated for all the wavelengths to ensure uniformity of performance in the signal, and to ensure that the observed trend was independent of wavelength and gain setting.

Figure 9 gives a plot of error ($\sigma(V)/V$) as a function of measured PMT voltage for 785 nm wavelength. A similar trend was observed for other wavelengths. This plot also gives a deviation in phase ($\sigma(\theta)$) in degrees for the same voltage. Each of these points represents a sample size of 200. The lowest measured voltage of PMT was 0.001 V. The
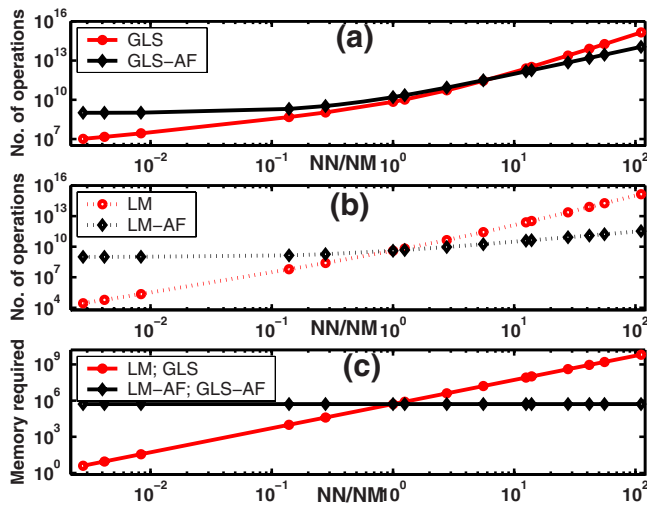
FIG. 2. Comparison of the number of operations required for the original update equation and its alternative form (a); for GLS (Eq. (9), represented by GLS) and its alternative form (Eq. (16), represented by GLS-AF) (b); for LM (Eq. (5), represented by LM) and its alternative form (Eq. (6), represented by LM-AF) as a function of the ratio of number of estimation parameters to number of measurements (represented by NN/NM). Memory required for implementing the inversion procedure is plotted in (c).

measured deviations were 1% for *lnA* and 0.5° for θ for PMT voltages above 0.005 V. These values are similar to the ones reported in the literature (Mcbride *et al.*[29] reported 0.32% for the PMT voltage and 0.48° in phase). A solid line in Fig. 9 shows these average deviation using $1/V^2$ fitting (following the shot-noise model). From this plot, the weight matrix for the data-model misfit ($\mathbf{W}_\delta$ in Eq. (9)) can be written as

$$[\mathbf{W}_\delta]_{ij}$$

$$
= \begin{cases}
0 & \text{if } i \neq j \\[2mm]
\dfrac{1}{[(5.7/V^2) + 0.6]^2} \text{for } \ln A \text{ with } V \geq 0.001 & \text{if } i = j \\[2mm]
\dfrac{1}{[(1.3/V^2) + 0.4]^2} \text{for } \theta \text{ with } V \geq 0.001 & \text{if } i = j . \\[2mm]
1 \text{ for } lnA \text{ with } V < 0.001 & \text{if } i = j \\[2mm]
\dfrac{1}{(\theta)^2} \text{ for } \theta \text{ with } V < 0.001 & \text{if } i = j
\end{cases}
$$

$$(22)$$

This implies that if the PMT voltage is below 0.001 V, the signal is considered to be in the noise floor. For the signals above 0.001 V, the variance can be estimated using $1/V^2$ fitted curve (Eq. (22)). This characterization enables the usage of the GLS scheme in the case of experimental data, with a requirement that the voltage measured at the PMT is available.

V.B. Gelatin phantom

A multilayer cylindrical gelatin phantom of diameter 86 mm, height 25 mm was fabricated using different mixtures of India ink for absorption and Titanium oxide (TiO₂)
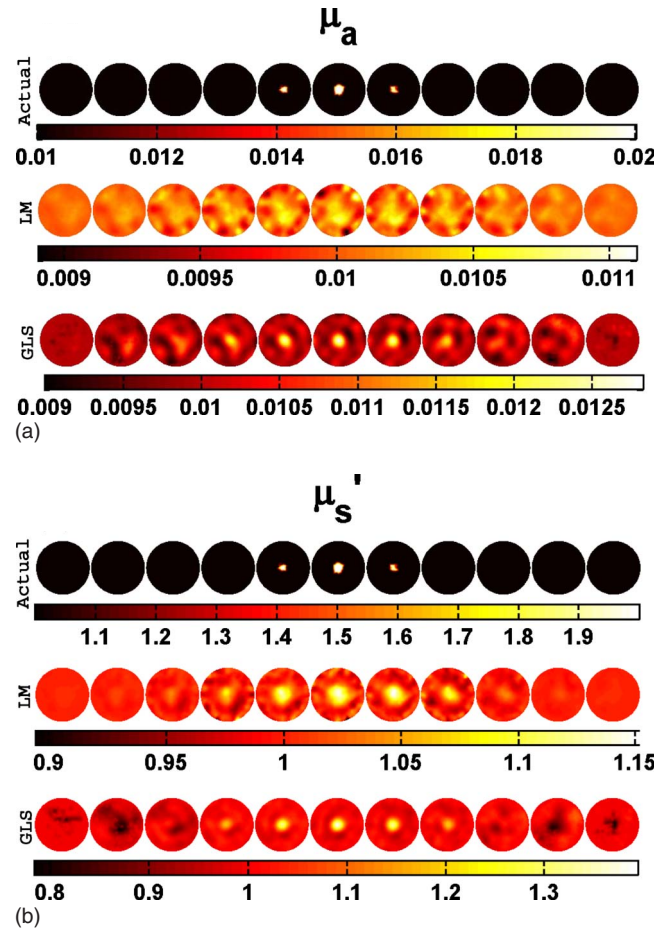

FIG. 3. Actual and reconstructed (a) μ_a and (b) μ_s' distributions of a spherical target having a diameter of 15 mm located at the center (at (0,0,0)) using 1% noisy data. Two-dimensional cross sections of the 3D volume in 5 mm increments spanning from $z=-25$ mm to $z=25$ mm (from left to right) are shown. Actual distributions are given in the first row. Reconstructed distribution using the LM minimization scheme and GLS minimization scheme are presented in the middle and last rows, respectively.

for scattering. These different layers of gelatin were fabricated by successively hardening heated gelatin solutions (typically 80% deionized water and 20% gelatin (G2625, Sigma Inc)) along with different amounts of ink and TiO₂ (Sigma Inc). A cylindrical hole extending in Z direction (diameter of 16 mm and height of 24 mm) filled with intra-lipid mixed with India ink acted as a target having the optical properties $\mu_a=0.02$ mm⁻¹ and $\mu_s'=1.2$ mm⁻¹. The outer layer with optical properties $\mu_a=0.0065$ mm⁻¹ and $\mu_s' =0.65$ mm⁻¹ had a thickness of 10 mm mimicking the typical fatty layer of the breast.[30] The middle layer with 76 mm diameter, mimicking fibro-glandular layer, had optical properties $\mu_a=0.01$ mm⁻¹ and $\mu_s'=1.0$ mm⁻¹. Validation of individual layer optical properties was performed by the data collected on large cylindrical samples of each layer using 785 nm wavelength laser diode as the source. Two-dimensional cross sections of this gelatin phantom optical properties are displayed in increments of 2.5 mm spanning from $z=-12.5$ mm to $z=12.5$ mm (from left-hand side to right-hand side) in top rows of Figs. 10(a) and 10(b). In this
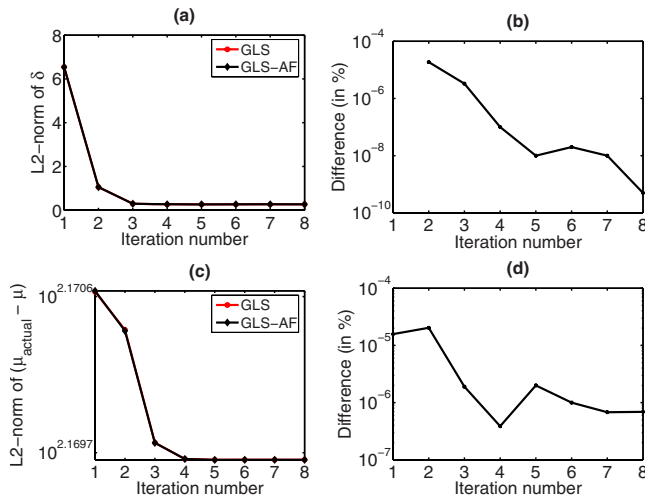
FIG. 4. Comparison of results from the GLS update equation (Eq. (9), represented by GLS) and its alternative form (Eq. (16), represented by GLS-AF). (a) L2 norm of data-model misfit (δ) as a function of iteration, (b) difference (in %) in the curves in (a), (c) L2-norm of the solution space ($\mu_{actual} - \mu$) as a function of iteration, (d) difference (in %) in the curves in (c).

phantom case, dater were collected using only one layer of fibers (at $z=0$ mm) leading to 240 $ln A$ data points and 240 θ data points. A cylindrical mesh consisting 8990 nodes corresponding to 44 803 linear tetrahedral elements was used and the experimental data were also calibrated using a reference homogenous phantom data. The outer layer (mimicking fatty layer) optical properties were used as initial guess for reconstruction procedures. A second mesh with the same geometry containing 3718 nodes (16 627 linear tetrahedral elements) was used as a reconstruction mesh.

VI. RESULTS

The number of operations required to produce an optical property update ($\Delta\mu_i$) for both the original and alternative GLS update equations (Eqs. (9) and (16), respectively) was compared as a function of the ratio of the number of estimation parameters (NN) to the number of measurements (NM).

A similar comparison for LM update equations (Eqs. (5) and (6)) was also performed. The results are plotted in Figs. 2(a) and 2(b), respectively. The expressions used for calculating the number of operations are given in Appendix 4. Memory required for implementing these inversions is presented in Fig. 2(c). To relate the analysis to existing experiments, NM was chosen to be 720. The number of nodes (equivalently estimation parameters) was varied from 2 to 80 000. The number of computations increases for both LM and GLS cases as NN increases, but the alternative form for GLS (GLS-AF) has a lower computational cost when NN/NM is greater than 6. In the case of alternative form LM (LM-AF), this is when NN/NM is greater than 2. In terms of memory, as soon as $NN > NM$, both GLS-AF and LM-AF requirement is less than GLS and LM counterparts.

In order to assess algorithm performance, a series of test reconstructions were evaluated. Data with 1% noise from the cylinder containing a 15-mm-diam spherical target (first rows of Figs. 3(a) and 3(b)) was reconstructed for the optical properties using LM and GLS techniques (middle and last rows and Figs. 3(a) and 3(b)). In the case of the GLS scheme, the reformulated update equation (alternative form - Eq. (16)) was used (last rows of Figs. 3(a) and 3(b)). However, it was also important to confirm that the two forms of the update equation (Eqs. (9) and (16)) produced numerically equivalent solutions. Figure 4 shows a comparison of results generated with the original GLS update equation (Eq. (9)) and its alternative form (Eq. (16)) in terms of data-model misfit (δ) and reconstructed optical properties. The difference plots (Figs. 4(b) and 4(d)) demonstrate that Eqs. (9) and (16) are equal within the limits of the numerical precision to be expected ($<10^{-8}$ of the L2-norm value, after the first few iterations). A similar analysis between the original LM update equation (Eq. (5)) and its alternative form (Eq. (6)) was performed and gave similar results (not shown here). These results indicate that there were no unexpected numerical difficulties in proving that these alternative forms of update equations (GLS: Eq. (16); LM: Eq. (6)) are equivalent to its original forms (GLS: Eq. (9); LM: Eq. (5)). Reconstruction results with different target shapes and positions are summa-

TABLE I. Mean and standard deviation of the reconstructed μ_a and μ_s' values in mm^{-1} for the background and target with LM and GLS techniques using data with 1% noise. The spherical target had a diameter of 15 mm. One set of reconstructed images for the target in the center is presented in Fig. 3. The cylindrical target diameter was 15 mm and extended throughout the imaging domain in the z direction.

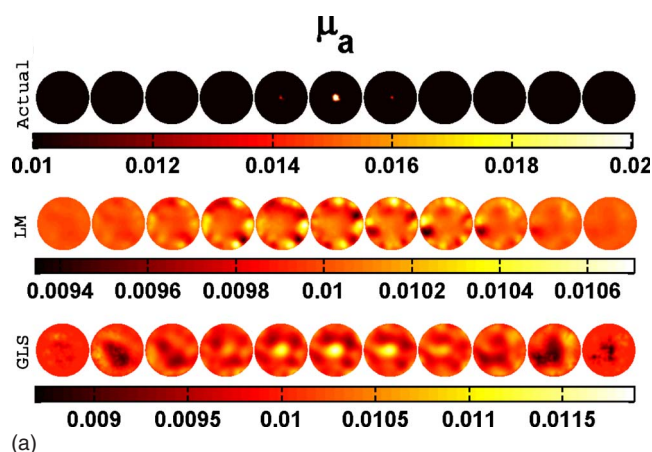| Methods	Shape	Target Position	μ_a	Target μ_s'	μ_a	Background μ_s'
Actual	$\cdots$	$\cdots$	0.01	1.0	0.02	2.0
LM	Spherical	(0,0,0)	0.0101 ± 0.003	1.0079 ± 0.0322	0.0104 ± 0.0002	1.1259 ± 0.0160
		(30,0,0)	0.0101 ± 0.0006	1.0063 ± 0.0500	0.0126 ± 0.0009	1.4514 ± 0.1152
	Cylindrical	(0,0)	0.0102 ± 0.0010	1.0120 ± 0.0874	0.0151 ± 0.0012	1.4308 ± 0.0854
		(30,0)	0.0101 ± 0.006	1.0030 ± 0.0663	0.0148 ± 0.0015	1.8406 ± 0.2470
GLS	Spherical	(0,0,0)	0.0101 ± 0.0001	1.0100 ± 0.0212	0.0122 ± 0.0004	1.2903 ± 0.0326
		(30,0,0)	0.0100 ± 0.0004	1.0108 ± 0.0250	0.0241 ± 0.0006	1.4498 ± 0.0853
	Cylindrical	(0,0)	0.0102 ± 0.0008	1.0055 ± 0.0500	0.0159 ± 0.0009	1.4750 ± 0.0941
		(30,0)	0.0101 ± 0.0008	1.0043 ± 0.0588	0.0170 ± 0.0012	1.6793 ± 0.1848
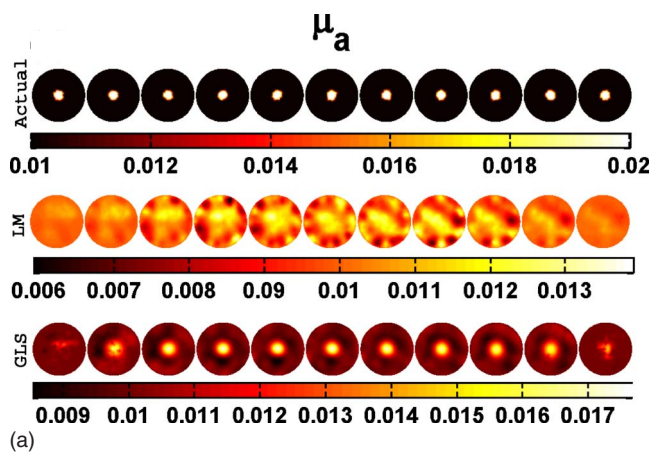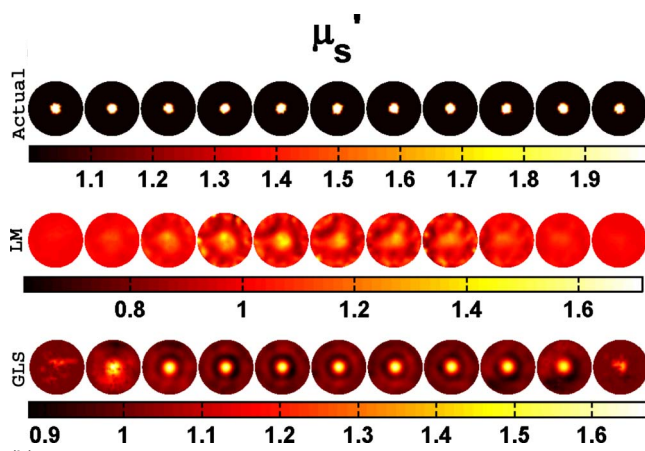
FIG. 5. Actual and reconstructed (a) μ_a and (b) μ_s' distributions of a cylindrical target located at the center (diameter — 15 mm) using 5% noisy data. Two-dimensional cross sections of the 3D volume in 5 mm increments spanning from $z=-25$ mm to $z=25$ mm (from left to right) are shown. Actual distributions are given in the first row. Reconstructed distribution using the LM minimization scheme and GLS minimization scheme are presented in the middle and last rows, respectively.
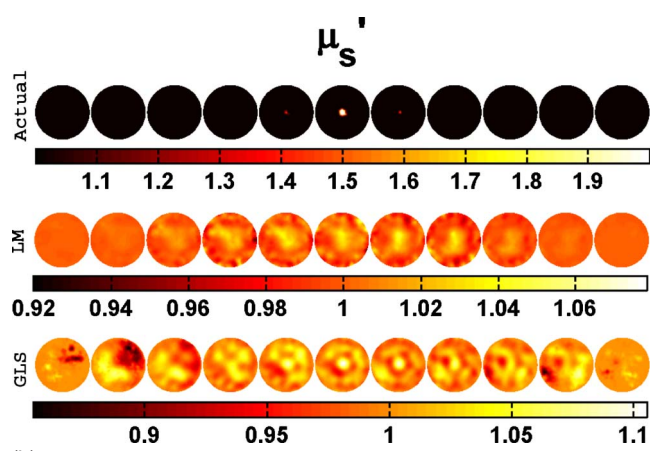
FIG. 6. Actual and reconstructed (a) μ_a and (b) μ_s' distributions of a spherical target having a diameter of 10 mm located at the center (at (0,0,0)) using 1% noisy data. Two-dimensional cross sections of the 3D volume in 5 mm increments spanning from $z=-25$ mm to $z=25$ mm (from left to right) are shown. Actual distributions are given in the first row. Reconstructed distribution using the LM minimization scheme and GLS minimization scheme are presented in the middle and last rows, respectively.

rized in Table I which reports the mean and standard deviation of recovered μ_a and μ_s' values in the background and target areas. Note that the recovered optical properties between $Z=15$ mm and $Z=-15$ mm were used because the reconstructed optical properties were equal to the actual background value and the standard deviation was zero (within round-off error limits) above/below these Z values.

To show the robustness of the GLS procedure, data with 5% noise was used in the reconstruction of a cylindrical target located in the center (as shown in first rows of Figs. 5(a) and 5(b)). The reconstruction results using the LM and GLS schemes are presented in middle and last rows of Figs. 5(a) and 5(b), respectively. The GLS minimization technique was able to localize the target more clearly than the LM method.

To see the effect of target size on the recovery of its contrast using these reconstruction techniques, a series of simulations were performed where the diameter of the spherical target located at the center was varied (from 10 to 35 mm). One set of results is presented for the 15-mm-diam target in Fig. 3. Another sample set for a target with diameter

of 10 mm is shown in Fig. 6. A comparison plot is given in Fig. 7. The data noise level for the cases considered here is 1%. Increase in the diameter of the spherical target increases the contrast recovery.

A performance comparison of these reconstruction techniques with increases in target (spherical object with diameter of 25 mm) contrast (from 2 to 10 with respect to background optical properties) located at the center (0,0,0) and (20,0,0) is presented in Fig. 8. Again, the noise level in the data was 1%. The recovery of contrast is much lower in the case of the centered target compared to the off-centered location in both LM and GLS techniques. Between the LM and GLS methods, the latter performs better in terms of recovery of contrast.

A study was conducted to evaluate estimation parameter independence (cross-talk) in these reconstruction procedures. The spherical target having a contrast of 2:1 only in μ_a was considered at the center and near the boundary of the imaging domain. Synthetic data with 1% noise were used in the reconstructions, and the results are presented in Table II in
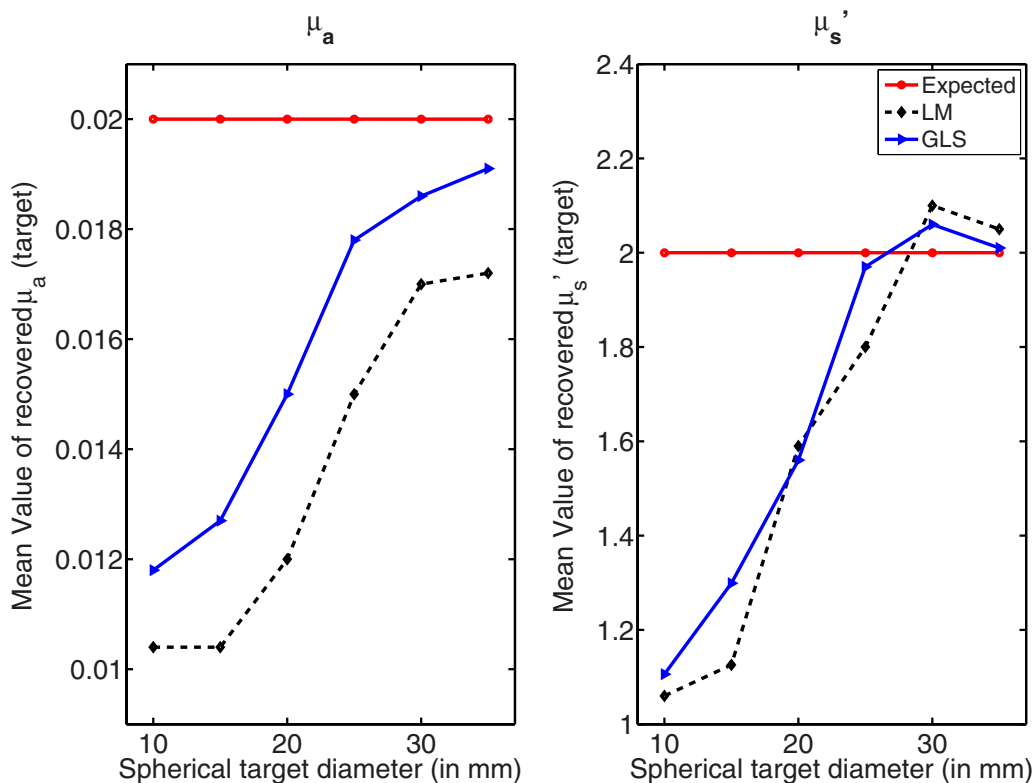
FIG. 7. Comparison of recovered optical properties as a function of inclusion size (diameter) for the spherical target located at (0,0,0) using both LM and GLS techniques. Standard deviations were of the same order as given in Table I.

terms of recovered mean and standard deviation of the optical properties. The recovery of contrast was higher in the GLS case and the amount of cross-talk was less (roughly 50% in the LM case compared to 30% in the GLS case). A similar study with a contrast of only in μ_s' was also conducted (not shown) and it also showed a similar trend in terms of cross-talk.

Finally, using the experimental data (Sec. V B) collected using a multilayered gelatin phantom, reconstructions using both LM and GLS techniques were performed. For the GLS technique, experimental data variance was estimated (Sec. V A) using an analytical equation, given in Eq. (22). Two-dimensional cross sections of the actual and reconstructed optical properties distribution are plotted in Figs. 10(a) and 10(b). The middle and bottom rows correspond to the reconstruction results obtained by LM and GLS techniques. The variance in the data (noise charters tics) is embedded in the GLS reconstruction procedure, resulting in optimal weighting (highly noisy data points get less weightage and vice versa for less noisy data points), leading to better quantification of tumor region.

VII. DISCUSSION

Appendix 1 presents a computationally efficient form for implementing an iterative GLS reconstruction scheme which reduces the dimensionality of the matrix to be inverted. The alternative forms for other minimization methods are also developed in the appendices (Sec. II and III). Appendix 4

presents expressions for estimating the operations count of both forms of the GLS update equations (Eqs. (9) and (16)) for a single iteration. Appendix 4 also gives operations count for both LM and its alternative form (Eqs. (5) and (6)) as well. Figure 2 shows a log-log plot of operations count as a function of the ratio of number of optical property parameters (*NN*) to number of measurements (*NM*) which determines the form of the GLS and LM update equation to be preferred (given that both produce numerically equivalent results as reported in Fig. 4). For example, when spatial priors are available, the number of optical unknowns can be reduced to the number of regions that can be segmented[31] which in the case of breast tissue is typically NN=3 (assuming fatty, fibro-glandular and tumor regions).[31] Here, since NM≫NN, the original LM and GLS update equation (Eqs. (5) and (9), respectively) is effective. In under-determined problems, such as the cases considered in this article, where NM≪NN (NN/NM ratio of 12 in the test problems), the GLS alternative form reduces the number of operations (by up to two times in Fig. 2(a) when NN/NM=12). In fact, the alternative form of the GLS update equation (Eq. (16)) becomes effective once NN/NM>6 and the number of operations decreases by an order of magnitude when NN/NM reaches 100. For the LM minimization scheme, alternative form reduces the number of operations by a factor of 6 in Fig. 2(b) when NN/NM=12. The memory required for inverting such matrices is plotted in Fig. 2(c) as a function of NN/NM. It is also important to recognize that the memory
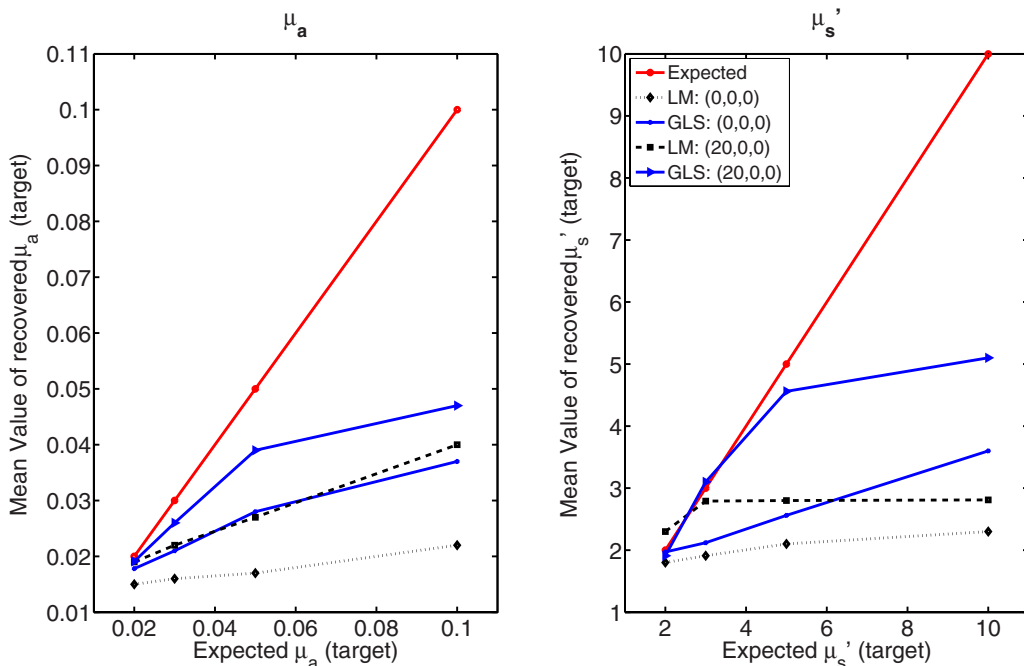
FIG. 8. Comparison of recovered optical properties with respect to expected values in a spherical target located at (0,0,0) and (20,0,0) having a diameter of 25 mm using both LM and GLS reconstruction techniques. Observed standard deviations are of similar order to that reported in Table I.

required to complete these operations can become critical because the cache sizes and RAM available on different architectures is variable but influences the efficiency of the computational processes executing on a given platform.

Under-determinedness of the imaging problem (i.e., NN/NM > 1) leads to nonuniqueness in the solution space, but regularization helps to give a unique solution in these cases. Typically NN/NM values are between 2 and 10 for a typical two-dimensional (2D) problem, as the choice depends on the expected resolution in the reconstructed image, imaging domain size, shape, data-collection geometry, and prior information available. For a 3D problem, this choice (NN/NM values) also depends on these factors, but one expects this ratio to be higher than the 2D case as the imaging domain size is bigger reflecting in the number of imaging parameters (NN) to be larger [Typical example: NN=600 (in 2D) and 6000 (in 3D)]. Ideally, one would like to have this ratio (NN/NM) constant between 2D and 3D; this implies that the number of measurements has to increase by the same

factor (typical case requires a factor of 10), which might not be feasible due to instrumentation constrains.[24] This leads to choice of NN/NM greater than 6 at least, where the derived forms are effective (even though in the case of LM, the alternative form is effective when NN/NM > 2). It is also important to note that, ideally one would like mesh the volume where the sensitivity is greater (for the imaging domain discussed here *z* between −25 and 25 mm) finer and rest of the domain coarser, to keep NN/NM in the same range as in 2D (typically 3–8). This adaptive meshing gets trickier when the patients are imaged, in our experience. We could not find appropriate tools that could be used in real time for this meshing problem, even though efforts to solve this patient-specific adaptive meshing are being pursued.[32,33] Even when NN/NM=3 (lowest ratio one expects in a 3D imaging problem), from Fig. 2(b), the alternative form (LM-AF, Eq. 6) in the case of LM minimization scheme becomes effective.

Figure 4 demonstrates that the two GLS update forms (Eqs. (9) and (16)) are equivalent numerically (within the

TABLE II. Mean and standard deviation of reconstructed μ_a and μ_s' in mm^{-1} values for the background and a spherical target with no scattering contrast using the LM and GLS techniques. The diameter of the spherical inclusion was 15 mm. Data with 1% noise were used.

Methods	Target position	Background		Target	
		μ_a	μ_s'	μ_a	μ_s'
Actual	$\cdots$	0.01	1.0	0.02	1.0
LM	(0,0,0)	0.0101 ± 0.0003	1.0025 ± 0.0211	0.0109 ± 0.0001	1.0500 ± 0.0071
	(30,0,0)	0.0101 ± 0.0004	1.0009 ± 0.0205	0.0119 ± 0.0004	1.0934 ± 0.0269
GLS	(0,0,0)	0.0101 ± 0.0001	1.0016 ± 0.0269	0.0126 ± 0.0002	1.0924 ± 0.0253
	(30,0,0)	0.0100 ± 0.0004	1.0029 ± 0.0189	0.0136 ± 0.0003	1.1002 ± 0.0294

numerical precision of the L2 norm value). For the cases considered here, the computation time for each iterative update using Eq. (16) was approximately 46 min, which was three times faster than with Eq. (9) (computation time ~126 min per iteration). In terms of operations count, a factor of 2 reduction would be estimated from Fig. 2(a) for the NN/NM ratio involved. The deviation in run time that occurs in practice is likely due to the cost of memory management alluded to above in the case of Eq. (9). It is also important to note that implementation of Eq. (9) requires an inversion of the covariance matrix (Eq. (10)), whereas this matrix can be used directly in Eq. (16). In the case of LM update equations, computation time for an iteration using Eq. (6) was approximately 21 min and for Eq. (5) was 91 min. The deviation from factor of 6 (from Fig. 2(b)) is mainly due to the memory required to perform these operations; this affects run time in turn.

The middle and last rows of Figs. 3(a) and 3(b) indicate that with 1% noise in the data, LM has failed to recover μ_a in a spherical target with a diameter of 15 mm located at the center, whereas GLS was able to identify the target very well. The failure of LM minimization is indicative of a lack of sensitivity at the center of the domain[24] which is improved through the GLS approach by including the noise characteristics and covariances associated with the problem. When the same target is located near the boundary (at (30,0,0)), both techniques were able to recover the contrast approximately 20% better relative to the center position (Table I).

When 1% noisy data generated from a centered cylindrical target with a diameter of 15 mm (first rows of Figs. 5(a) and 5(b)) were used in the reconstruction, both LM and GLS techniques were able to recover approximately 50% of the expected contrast (Table I). For the same type of target located near the boundary (at (30,0)), the recovery of contrast was approximately 70% in the case of GLS. For LM, the recovery was only 50% for μ_a and 85% for μ_s' under these conditions. The reconstruction results also show that recovery of the centered target is always poor relative to an object near the boundary. This is primarily due to the hypersensitivity at the boundary in these cases.[24] The extended cylindrical target is essentially equivalent to the two-dimensional case of a circular inclusion and the trend observed in 3D of recovering more contrast for a target near the boundary is similar to the behavior found in 2D.[24]

When 5% noisy data were used, LM reconstruction (middle rows of Figs. 5(a) and 5(b)) performs poorly in terms of localization of the target, whereas GLS was able to reconstruct optical images with better quality and quantitation (up to 70%). Even though the reconstructed results using very noisy data were presented here from only one type of target, similar trends were also observed in other cases that mimic the 2D reconstructions reported in Ref. 17. These results show that GLS outperforms LM even though the data noise level is high because stability is retained by including the noise characteristics into the weight matrices used for normalization.

Accuracy in contrast recovery of local targets increases as the size of the target increases, as shown in Fig. 7. For ex-
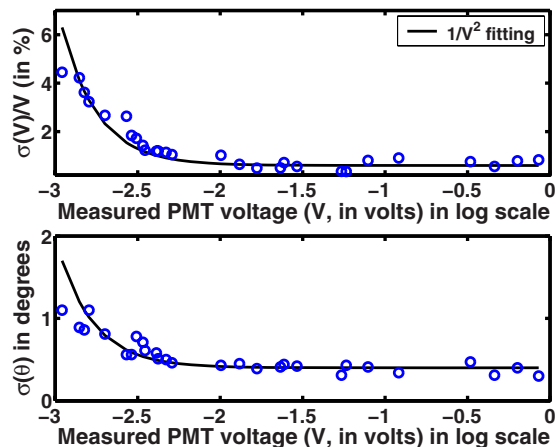

FIG. 9. An error (deviation, σ) plot of the measured voltage and phase (θ) as a function of mean of measured PMT voltage. The legend of the figure represents the fitting model used. Each data point corresponds to a sample size of 200.

ample, the contrast recovered for a centered target below 20 mm in diameter is only about 30% of the true value and as low as 10% for the LM algorithm (Fig. 7), whereas increasing the size of target to 30 mm leads to quantitative accuracy near 100%. The GLS approach provides maximal contrast recovery and superior image quality at all sizes relative to LM (example: Fig. 6). Even when the target size is as low as 5 mm (Fig. 6), the object was well localized in the GLS case, but not with LM (Fig. 6).

The performance comparison of the algorithms in terms of contrast recovery (Fig. 8) confirms that the position of the target dictates the response. When the target had 10:1 contrast in comparison to the background, the maximum recovery of contrast was ~5:1. GLS outperformed LM in this regard but there is a plateau in recovery of contrast at 400% of the background value.

Table II shows that estimation parameter dependence (cross-talk) is lower (by 20%) for GLS compared to LM, by reinforcing the independence of μ_a and μ_s' through elimination of any cross terms in the weight matrix $\mathbf{W}_{\mu-\mu_0}$ (Eq. (15)). The inter-parameter dependence is complex because of the non-unique relationship between the optical property distribution and the incomplete boundary data, indicating that different formulations of the inversion tend to perform differently. Nonetheless, the estimation parameter dependence is substantially higher in 3D data-limited situations, relative to when the ratio of data to number of estimates is less skewed.

In the case of phantom data (Fig. 10), as expected, GLS reconstructions showed more promising results in this test case. Characterization of the data collection system, leading to variance estimation depending on the voltage measure at PMT (Fig. 9) enabled the employment of the GLS technique for experimental data reconstructions. Both techniques (LM and GLS) were able to give qualitative information about the target, in terms of quantification the GLS technique overtakes the LM technique (Fig. 10). It should be noted that this
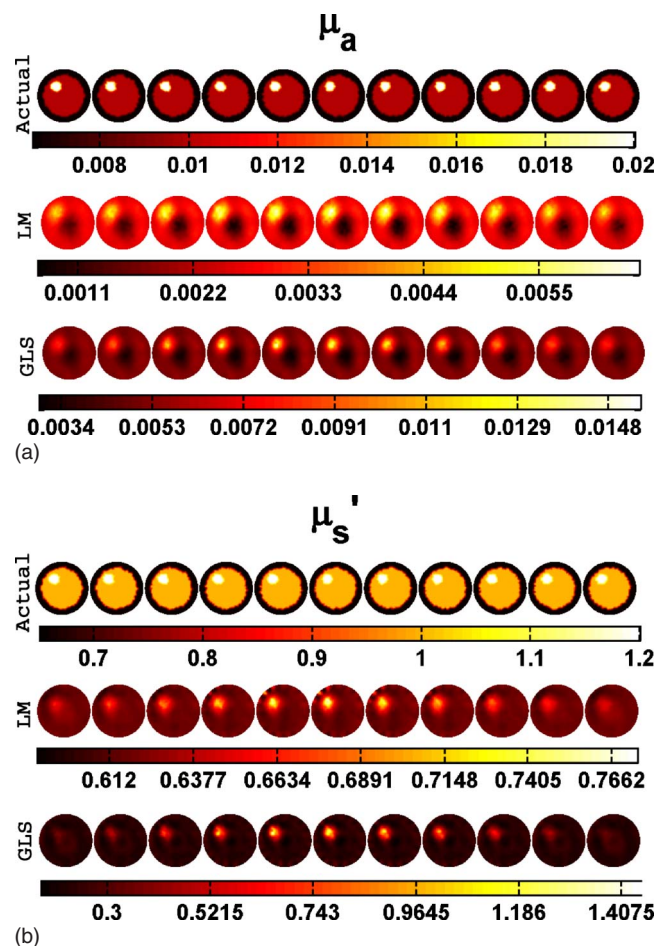
FIG. 10. Actual and reconstructed (a) μ_a and (b) μ_s' distributions of a cylindrical target using experimental multilayered phantom data. Two-dimensional cross sections of the 3D volume in 2.5 mm increments spanning from $z=-12.5$ mm to $z=12.5$ mm (from left to right) are shown. Actual distributions are given in the first row. Reconstructed distribution using the LM minimization scheme and GLS minimization scheme are presented in the middle and last rows, respectively.

type of characterization of the experimental system does not take into account coupling errors between the light collection/delivery fiber and tissue surface. These kinds of unsystematic errors are difficult to estimate as it depends on many parameters, such as tissue surface roughness, tissue elastic properties, design of tissue-fiber coupling interface, repeatability, and alignment of fibers. Recent advances, such as inclusion of coupling errors as a part of reconstruction scheme by making them as unknowns and iteratively solving for them along with optical parameters, are possible.[34–37] But this makes the reconstruction problem more computationally complex in nature as the number of unknowns of this procedure are larger than the techniques considered here. As the main focus of this article is about presenting a computationally efficient algorithm for reconstructing optical parameters, this issue was not considered as a part of this work. Development of methods, which were presented as a part of this work, that can include systematic noise characteristics in the reconstruction procedure, will be taking a step in the right direction. Moreover, as can be seen from Eq. (22), covari-

ances among the data points were ignored, making the data weight matrix ($\mathbf{W}_\delta$) a simple diagonal matrix. Inclusion of covariances can offer a better weighting in case of experimental data; this is under investigation as of now. Even though this test case showed very promising results, in our experience, in cases where coupling errors are dominant in the data, the GLS scheme did not yield any meaningful results. In these cases, the LM technique was able to give reasonable results.

Even though the data used here are generated by using in-plane data, it was collected only from the source fiber plane. Previous investigations indicate that the use of out-of-plane data (when the data were collected from rest of the fibers in all three planes when one fiber was used as source) may not give enough advantage in terms of reconstructed image quality given an increase in the data-acquisition time and computational cost compared to in-plane data (in this case the data were collected only from those fibers which lie in the same plane as the source fiber).[24] So for all the simulation studies conducted here, in-plane data collection scheme is used. The experimental phantom study uses only one plane of data, as the phantom optical properties were not varying in Z direction. Even though considering a experimental phantom where the optical properties are varying in Z direction might have been ideal to demonstrate the depth resolution of reconstruction procedures, in here only a simple case was considered as the main aim is to prove that the techniques developed here can be used in experimental data reconstruction. It is also important to note that the studies conducted here are generic in nature, especially in terms of proving the computational efficiency of alternative forms (Fig. 2), as NN/NM was changed over a range of 0.0028–100 (spanning from well-determined to highly under-determined problems).

Partial volume effects can be observed in the recovery of contrast as a function of target size. The recovery of contrast was much higher for the extended cylinder target compared to the spherical inclusion (Table I). The quantitative accuracy of reconstructed images increases with an increase in target size (Fig. 7). GLS reconstruction results of the data from a centered cylindrical object are encouraging, demonstrating recovery of more than 30% contrast in this case (other Newton-type algorithms have reported a maximum of 10% contrast recovery[7,11,13,14,24]).

VIII. CONCLUSIONS

Three-dimensional diffuse optical tomography is more computationally intensive because of the size of the parameter space to be reconstructed. Newton-based inversion methods that operate on a Hessian matrix, which has dimensions of the number of measurements rather than the number of parameters, can be derived using the Sherman–Morrison–Woodbury identity and become computationally more efficient once the number of estimation parameters exceeds two times the number of measurements. Representative examples demonstrate that this form of update equation can be at least six times faster in practice in the highly under-determined

problems which commonly occur in 3D. Three-dimensional diffuse optical tomographic reconstruction algorithms also suffer from partial volume effects that significantly degrade the accuracy with which optical properties can be quantified. The GLS approach which incorporates structured weight matrices consisting of the variance and covariance of the data-model misfit and the optical properties, improves the quantification of optical properties by at least 20% in 3D. The GLS estimate is also robust to data noise as high as 5%—conditions under which other algorithms fail when the problem is highly under-determined. By characterizing the detector noise for systematic errors, using a multi-layered gelatin phantom data, the GLS technique can be easily employed for reconstructing experimental data and can yield better quantification of targets compared to conventional reconstruction methods. Future investigations will include thorough examination of the GLS technique when applied to phantom and clinical data and extension of the technique to direct-spectral reconstruction. The test data used in this article, along with computer algorithms, are available on a web page.[38]

APPENDIX

1. Alternative form for GLS update equation

Before deriving the alternative form, it is useful to catalog several properties of the weight matrices and their inverses[28]

$$\mathbf{W}_\delta = (\mathbf{C}_\delta)^{-1}; \mathbf{W}_{\mu-\mu_0} = (\mathbf{C}_{\mu-\mu_0})^{-1}$$

$$(\mathbf{W}_\delta)^T = \mathbf{W}_\delta; (\mathbf{W}_{\mu-\mu_0})^T = \mathbf{W}_{\mu-\mu_0}$$

$$(\mathbf{C}_\delta)^T = \mathbf{C}_\delta; (\mathbf{C}_{\mu-\mu_0})^T = \mathbf{C}_{\mu-\mu_0}. \tag{A1}$$

If a square matrix $\mathbb{A}$ has block form

$$\mathbb{A} = \begin{bmatrix} -\mathbf{C}_\delta & \mathbf{J} \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0} \end{bmatrix} \tag{A2}$$

with dimensions $(2NM+2NN) \times (2NM+2NN)$, it is readily shown to be symmetric by invoking the relationships in Eq. (A1):

$$(\mathbb{A}^T) = \begin{bmatrix} -\mathbf{C}_\delta & \mathbf{J} \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0} \end{bmatrix}^T = \begin{bmatrix} -\mathbf{C}_\delta^T & (\mathbf{J}^T)^T \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0}^T \end{bmatrix}$$

$$= \begin{bmatrix} -\mathbf{C}_\delta & \mathbf{J} \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0} \end{bmatrix} = \mathbb{A}. \tag{A3}$$

Since inverses of both $\mathbf{C}_\delta$ and $\mathbf{W}_{\mu-\mu_0}$ exist, then $\mathbb{A}^{-1}$ also exists and can be expressed in block form as well

$$\mathbb{A}^{-1} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix} \tag{A4}$$

in which case

$$\mathbb{A}\mathbb{A}^{-1} = \begin{bmatrix} -\mathbf{C}_\delta & \mathbf{J} \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0} \end{bmatrix} \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{A5}$$

requires that

$$-\mathbf{C}_\delta\mathbf{P} + \mathbf{J}\mathbf{R} = \mathbf{I} \tag{A6}$$

$$-\mathbf{C}_\delta\mathbf{Q} + \mathbf{J}\mathbf{S} = \mathbf{0} \tag{A7}$$

$$\mathbf{J}^T\mathbf{P} + \mathbf{W}_{\mu-\mu_0}\mathbf{R} = \mathbf{0} \tag{A8}$$

$$\mathbf{J}^T\mathbf{Q} + \mathbf{W}_{\mu-\mu_0}\mathbf{S} = \mathbf{I}. \tag{A9}$$

These relationships can be manipulated through a series of substitutions to express the blocks of $\mathbb{A}^{-1}$ in terms of combinations of the block components of $\mathbb{A}$. Specifically, Eqs. (A7) and (A8) along with the weight matrix properties in Eqs. (A1) imply that

$$\mathbf{Q} = W_\delta\mathbf{J}\mathbf{S} \tag{A10}$$

and

$$\mathbf{R} = -\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T\mathbf{P}. \tag{A11}$$

Substituting Eq. (A11) into Eq. (A6) to form the expression

$$\mathbf{P} = -(\mathbf{C}_\delta + \mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T)^{-1}, \tag{A12}$$

which is put back into Eq. (A11) produces

$$\mathbf{R} = \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{C}_\delta + \mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T)^{-1}. \tag{A13}$$

A similar series of steps starting with Eqs. (A10) and (A9) to write

$$\mathbf{S} = (\mathbf{W}_{\mu-\mu_0} + \mathbf{J}^T\mathbf{W}_\delta\mathbf{J})^{-1}, \tag{A14}$$

which is combined again with Eq. (A10) yields

$$\mathbf{Q} = \mathbf{W}_\delta\mathbf{J}(\mathbf{W}_{\mu-\mu_0} + \mathbf{J}^T\mathbf{W}_\delta\mathbf{J})^{-1}. \tag{A15}$$

Since $\mathbb{A}$ is symmetric and invertible, $\mathbb{A}^{-1}$ is symmetric as well.

$$(\mathbb{A}^{-1})^T = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix}^T = \begin{bmatrix} \mathbf{P}^T & \mathbf{R}^T \\ \mathbf{Q}^T & \mathbf{S}^T \end{bmatrix} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix} = \mathbb{A}^{-1} \tag{A16}$$

in which case

$$\mathbf{Q}^T = \mathbf{R}. \tag{A17}$$

Substituting the forms of $\mathbf{Q}$ and $\mathbf{R}$ (Eqs. (A15) and (A13), respectively) into Eq. (A17) results in

$$[\mathbf{W}_\delta\mathbf{J}(\mathbf{W}_{\mu-\mu_0} + \mathbf{J}^T\mathbf{W}_\delta\mathbf{J})^{-1}]^T = \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{C}_\delta + \mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T)^{-1} \tag{A18}$$

Equation (A16) also requires $\mathbf{S}^T = \mathbf{S}$, where $\mathbf{S}$ is given by Eq. (A15), which when identified in the term on the left side of Eq. (A18) allows it to be rewritten as

$$(\mathbf{W}_{\mu-\mu_0} + \mathbf{J}^T\mathbf{W}_\delta\mathbf{J})^{-1}\mathbf{J}^T\mathbf{W}_\delta = \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{C}_\delta + \mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T)^{-1}.$$

(A19)

Alternately,

$$\mathbb{A}^{-1}\mathbb{A} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix} = \begin{bmatrix} -\mathbf{C}_\delta & \mathbf{J} \\ \mathbf{J}^T & \mathbf{W}_{\mu-\mu_0} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

(A20)

or

$$\mathbf{RJ} + \mathbf{SW}_{\mu-\mu_0} = \mathbf{I}.$$

(A21)

Solving for $\mathbf{S}$

$$\mathbf{S} = \mathbf{C}_{\mu-\mu_0} - \mathbf{RJC}_{\mu-\mu_0}$$

(A22)

and substituting Eqs. (A13) and (A14) for $\mathbf{R}$ and $\mathbf{S}$ in Eq. (A22) produces

$$(\mathbf{W}_{\mu-\mu_0} + \mathbf{J}^T\mathbf{W}_\delta\mathbf{J})^{-1}$$
$$= \mathbf{C}_{\mu-\mu_0} - \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{C}_\delta + \mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T)^{-1}\mathbf{J}\mathbf{C}_{\mu-\mu_0}.$$

(A23)

Note that this derivation was adapted from Liebelt.[39] A variant of Eq. (A23) exists in the literature with many names, the most common being the **Sherm–Morrison–Woodbury identity**.[40–47] It is also known as the **matrix inversion lemma**.[48,49] Even though one can start from this equation and derive the alternative forms, the complete derivation is presented here for completeness.

Substituting Eq. (A23) back into Eq. (9) yields

$$\Delta\mu_i = [\mathbf{C}_{\mu-\mu_0} - \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T + \mathbf{C}_\delta)^{-1}\mathbf{J}\mathbf{C}_{\mu-\mu_0}]$$
$$\times \{\mathbf{J}^T\mathbf{W}_\delta\delta_{i-1} - \mathbf{W}_{\mu-\mu_0}(\mu_{i-1} - \mu_0)\}$$

(A24)

or

$$\Delta\mu_i = [\mathbf{I} - \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T + \mathbf{C}_\delta)^{-1}\mathbf{J}]$$
$$\times \{\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T\mathbf{W}_\delta\delta_{i-1} - \mathbf{C}_{\mu-\mu_0}\mathbf{W}_{\mu-\mu_0}(\mu_{i-1} - \mu_0)\}$$

(A25)

which results in

$$\Delta\mu_i = [\mathbf{I} - \mathbf{C}_{\mu-\mu_0}\mathbf{J}^T(\mathbf{J}\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T + \mathbf{C}_\delta)^{-1}\mathbf{J}]$$
$$\times \{\mathbf{C}_{\mu-\mu_0}\mathbf{J}^T\mathbf{W}_\delta\delta_{i-1} - (\mu_{i-1} - \mu_0)\}$$

(A26)

as the alternative form for the update equation (Eq. (16)).

The next two subsections show the alternative forms of other least-squares minimization techniques, namely LM and Tikhonov minimizations.

2. Alternative form for LM update equation

The LM update equation (Eq. (5)) is a special case of the GLS update equation (Eq. (9)) when $\mathbf{W}_{\mu-\mu_0} = \alpha\mathbf{I}$ and $\mathbf{W}_\delta = \mathbf{I}$ (see Sec. III.V.4 in Ref. 17). Using these forms in Eq. (A19) leads to an alternative form to Eq. (5)

$$\Delta\mu_i = (\alpha\mathbf{I})^{-1}\mathbf{J}^T(\mathbf{J}(\alpha\mathbf{I})^{-1}\mathbf{J}^T + \mathbf{I}^{-1})^{-1}\delta_{i-1}.$$

(A27)

Rearranging the terms in Eq. (A27) leads to

$$\Delta\mu_i = \frac{\mathbf{J}^T}{\alpha}\left(\frac{\mathbf{J}\mathbf{J}^T + \alpha\mathbf{I}}{\alpha}\right)^{-1}\delta_{i-1}$$

(A28)

which can be simplified to produce

$$\Delta\mu_i = \mathbf{J}^T(\mathbf{J}\mathbf{J}^T + \alpha\mathbf{I})^{-1}\delta_{i-1}.$$

(A29)

Equation (A29) is also known as under-determined form in the literature.[11,14]

3. Alternative form for Tikhonov update equation

The objective function[17,50] for the Tikhonov scheme is

$$\Omega = \|\mathbf{y} - G(\mu)\|^2 + \lambda\|\mathbf{L}(\mu - \mu_0)\|^2.$$

(A30)

Minimization of Eq. (A30) and linearizing the problem leads to update equation[17,31]

$$\Delta\mu_i = [\mathbf{J}^T\mathbf{J} + \lambda\mathbf{L}^T\mathbf{L}]^{-1}(\mathbf{J}^T\delta_{i-1} - \lambda\mathbf{L}^T\mathbf{L}(\mu_{i-1} - \mu_0)).$$

(A31)

Equation (A31) is a special case of the GLS update equation (Eq. (9)) with weight matrices (see Sec. III.B.4 in Ref. 17)

$$\mathbf{W}_\delta = \mathbf{I}; \mathbf{W}_{\mu-\mu_0} = \lambda\mathbf{L}^T\mathbf{L}.$$

(A32)

From Eq. (A26) one can write

$$\Delta\mu_i = [\mathbf{I} - (\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T(\mathbf{J}(\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T + \mathbf{I}^{-1})^{-1}\mathbf{J}]$$
$$\times \{(\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T\mathbf{I}\delta_{i-1} - (\mu_{i-1} - \mu_0)\}.$$

(A33)

This leads to

$$\Delta\mu_i = [\mathbf{I} - (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T(\mathbf{J}(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T + \lambda\mathbf{I})^{-1}\mathbf{J}]$$
$$\times \{(\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T\delta_{i-1} - (\mu_{i-1} - \mu_0)\}.$$

(A34)

Assuming $\mu_{i-1} = \mu_0$, the single-step Tikhonov update equation (or its equivalent)[14,30,51] becomes

$$\Delta\mu_i = (\mathbf{J}^T\mathbf{J} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T\delta_{i-1}.$$

(A35)

Using Eqs. (A19) and (A32) leads to

$$\Delta\mu_i = (\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T(\mathbf{J}(\lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T + \mathbf{I}^{-1})^{-1}\delta_{i-1},$$

(A36)

which can be rearranged to

$$\Delta\mu_i = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T(\mathbf{J}(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{J}^T + \lambda\mathbf{I})^{-1}\delta_{i-1}.$$

(A37)

Equation (A37) is also known as the under-determined Tikhonov single-step update equation.[11,14,16]

4. Calculation of number of operations for LM and GLS update equations

The number of operations was estimated by assuming that divisions/multiplications consume most of the processor cycles. Note that Gaussian elimination was used in computing matrix inversion. Typically, Gaussian elimination for an $N \times N$ matrix requires $((N^3/3) + N^2 - (N/3))$ operations.[47] The memory required to invert a matrix of dimension $N \times N$ is N^2.[47] The number of operations only includes solution of the update equation and does not account for the number of operations required to form the matrices/vectors used in these equations.

For the GLS update equation (Eq. (9)), the number of operations required for iteration i is (using the dimensions defined after Eq. (15))

Number of operations

$$= \left[(2NM * 2NM * 2NN) + (2NN * 2NM * 2NN) \right.$$
$$+ \left. \left(\frac{(2NN)}{3} + (2NN)^2 - \frac{2NN}{3} \right) \right] + (2NN * 2NN * 1)$$
$$+ \left[(2NN * 2NM * 1) + (2NN * 2NN * 1) \right.$$
$$+ \left. (2NN * 2NN * 1) \right]. \tag{A38}$$

For the alternative form for the GLS update equation (Eq. (16)), the number of operations is

Number of operations

$$= \left[(2NN * 2NN * 2NM) \right.$$
$$+ (2NM * 2NN * 2NM) + (2NN * 2NN * 2NM)$$
$$+ \left(\frac{(2NM)^3}{3} + (2NM)^2 - \frac{2NM}{3} \right)$$
$$+ \left. (2NM * 2NM * 2NN) \right] + (2NN * 2NN * 1)$$
$$+ \left[(2NN * 2NN * 2NM) + (2NN * 2NM * 1) \right.$$
$$+ \left. (2NM * 2NM * 1) \right]. \tag{A39}$$

Similarly, for the LM update equation (Eq. (5)), the number of operations required for iteration i is

Number of operations

$$= \left[(2NN * 2NM * 2NN) \right.$$
$$+ \left. \left(\frac{(2NN)^3}{3} + (2NN)^2 - \frac{2NN}{3} \right) \right]$$
$$+ (2NN * 2NM * 1). \tag{A40}$$

The number of operations for the alternative form for the LM update equation (Eq. (6)) is

Number of operations

$$= (2NN * 2NM * 1)$$
$$+ \left[(2NM * 2NN * 2NM) \right.$$
$$+ \left. \left(\frac{(2NM)^3}{3} + (2NM)^2 - \frac{2NM}{3} \right) \right]$$
$$+ (2NM * 2NM * 1). \tag{A41}$$

Note that the computation time for these update equations is not only dependent on the number of operations needed to be performed, but also on the memory required for implementing the operations.

[a)] Present address: Department of Radiation Oncology, Washington University School of Medicine, St. Louis, Missouri 63110.

[b)] Author to whom correspondence should be addressed. Electronic mail: Pogue@dartmouth.edu

[1] 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).

[2] 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).

[3] A. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical tomography," Prog. Nucl. Energy **50**, R1–R43 (2005).

[4] D. R. Leff, O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Pattenl, J. Hebden, G. Z. Yang, and A. Darzil, "Diffuse optical imaging of the healthy and diseased breast: A systematic review," Breast Cancer Res. Treat. **108**, 9–22 (2008).

[5] S. R. Arridge and J. C. Hebden, "Optical imaging in medicine: II. Modeling and reconstruction," Phys. Med. Biol. **42**, 841–853 (1997).

[6] R. Arridge, "Optical tomography in medical imaging," Inverse Probl. **15**, R41–R93 (1999).

[7] M. Schweiger and S. R. Arridge, "Comparison of two- and three-dimensional reconstruction methods in optical tomography," Appl. Opt. **37**, 7419–7428 (1998).

[8] H. Jiang, Y. Xu, N. Iftimia, J. Eggert, K. Klove, L. Baron, and L. Fajardo, "Three-dimensional optical tomographic imaging of breast in a human subject," IEEE Trans. Med. Imaging **20**, 1334–1340 (2001).

[9] B. W. Pogue, S. Geimer, T. O. McBride, S. Jiang, U. L. Sterberg, and K. D. Paulsen, "Three-dimensional simulation of near-infrared diffusion in tissue: Boundary condition and geometry analysis for finite element image reconstruction," Appl. Opt. **40**, 588–600 (2001).

[10] J. C. Hebden, H. Veenstra, H. Dehghani, E. M. C. Hillman, M. Schweiger, S. R. Arridge, and D. T. Delpy, "Three-dimensional time-resolved optical tomography of a conical breast phantom," Appl. Opt. **40**, 3278–3287 (2001).

[11] H. Dehghani, B. W. Pogue, S. P. Poplack, and K. D. Paulsen, "Multiwavelength three-dimensional near-infrared tomography of the breast: Initial simulation, phantom, and clinical results," Appl. Opt. **42**, 135–145 (2003).

[12] A. Gibson, R. M. Yusof, E. M. C. Hillman, H. Dehghani, J. Riley, N. Everdale, R. Richards, J. C. Hebden, M. Schweiger, S. R. Arridge, and D. T. Delpy, "Optical tomography of a realistic neonatal head phantom," Appl. Opt. **42**, 3109–3116 (2003).

[13] H. Dehghani, B. W. Pogue, J. Shudong, B. Brooksby, and K. D. Paulsen, "Three-dimensional optical tomography: Resolution in small-object imaging," Appl. Opt. **42**, 3117–3128 (2003).

[14] M. Schweiger, S. R. Arridge, and I. Nissila, "Gauss-Newton method for image reconstruction in diffuse optical tomography," Phys. Med. Biol. **50**, 2365–2386 (2005).

[15] J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemp, V. Ntziachristos, D. N. Pattanayak, B. Chance, and A. G. Yodh, "Three-dimensional diffuse optical tomography in the plane parallel transmission geometry: Evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging," Med. Phys. **30**, 235–247 (2003).

[16] M. J. Eppstein, D. E. Dougherty, D. J. Hawysz, and E. M. Sevick-Muraca, "Three-dimensional Bayesian optical image reconstruction with domain decomposition," IEEE Trans. Med. Imaging **20**, 147–163 (2001).

[17] P. K. Yalavarthy, B. W. Pogue, H. Dehghani, and K. D. Paulsen, "Weight-matrix structured regularization provides optimal generalized least-squares estimate in diffuse optical tomography," Med. Phys. **34**(6), 2085–2098 (2007).

[18] S. R. Arridge and M. Schweiger, "A gradient based optimization scheme for optical tomography," Opt. Express **2**, 213–226 (1998).

[19] A. H. Hielscher, A. D. Klose, and K. M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography,"

IEEE Trans. Med. Imaging **18**, 262–271 (1999).

[20] 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).

[21] S. R. Arridge and M. Schweiger, "Photon-measurement density functions. Part 2: Finite-element-method calculations," Appl. Opt. **34**, 8026–8037 (1995).

[22] 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).

[23] M. Schweiger, S. R. Arridge, and D. T. Delpy, "Application of the finite element method for the forward and inverse models in optical tomography," J. Math. Imaging Vision **3**, 263–283 (1993).

[24] 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).

[25] K. Levenberg, "A method for the solution of certain nonlinear problems in least squares," Q. Appl. Math. **2**, 164–168 (1944).

[26] D. W. Marquardt, "An algorithm for least squares estimation of nonlinear parameters," J. Soc. Ind. Appl. Math. **11**, 431–441 (1963).

[27] S. Srinivasan, B. W. Pogue, H. Dehghani, S. Jiang, X. Song, and K. D. Paulsen, "Improved quantification of small objects in near-infrared diffuse optical tomography," J. Biomed. Opt. **9**, 1161–1171 (2004).

[28] L. R. Lynch, *Numerical Partial Differential Equations for Environmental Scientists and Engineers—A First Practical Course* (Springer, New York, 2005).

[29] 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).

[30] B. Brooksby, S. Jiang, H. Dehghani, B. W. Pogue, K. D. Paulsen, J. Weaver, C. Kogel, and S. P. Poplack, "Combining near infrared tomography and magnetic resonance imaging to study in vivo breast tissue: Implementation of a Laplacian-type regularization to incorporate magnetic resonance structure," J. Biomed. Opt. **10**, 051504:1–10 (2005).

[31] P. K. Yalavarthy, B. W. Pogue, H. Dehghani, C. M. Carpenter, S. Jiang, and K. D. Paulsen, "Structural information within regularization matrices improves near infrared diffuse optical tomography," Opt. Express **15**, 8043–8058 (2007).

[32] 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).

[33] M. E. Eames, B. W. Pogue, P. K. Yalavarthy, and H. Dehghani, "An efficient Jacobian reduction method for diffuse optical image reconstruction," Opt. Express **15**, 15908–15919 (2007).

[34] D. A. Boas, T. J. Gaudette, and S. R. Arridge, "Simultaneous imaging and optode calibration with diffusive optical tomography," Surv. Math. Ind. **8**, 263–270 (2001).

[35] J. J. Stott, J. P. Culver, S. R. Arridge, and D. A. Boas, "Optode positional calibration in diffuse optical tomography," Appl. Opt. **42**, 3154–3162 (2003).

[36] M. Schweiger, I. Nissil, D. A. Boas, and S. R. Arridge, "Image reconstruction in optical tomography in the presence of coupling errors," Appl. Opt. **46**, 2743–2756 (2007).

[37] S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector calibration in three-dimensional Bayesian optical diffusion tomography," J. Opt. Soc. Am. A **19**, 1983–1993 (2002).

[38] http://nir.thayer.dartmouth.edu/gls3d.html.

[39] P. B. Liebelt, *An Introduction to Optimal Estimation* (Addison-Wesley, Reading, MA, 1967).

[40] W. J. Duncan, "Some devices for the solution of large sets of simultaneous equations (with an appendix on the reciprocation of partitioned matrices)," London, Edinburgh Dublin Philos. Mag. J. Sci. **35**, 660–670 (1944).

[41] 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).

[42] 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).

[43] M. Wooodbury, "Inverting Modified Matrices," Memorandum Report No. 42, Statistical Research Group, Princeton (1950).

[44] A. S. Householder, *The Theory of Matrices in Numerical Analysis* (Dover, New York, 1964), p. 123.

[45] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," SIAM Rev. **23**, 53–60 (1981).

[46] H. Golub and C. F. Van Loan, *Matrix Computations* 3rd ed. (The John Hopkins University Press, Baltimore and London, 1996), p. 50.

[47] J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations* (Wiley, New York, 1968), p. 132.

[48] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communication, and Control* (Prentice-Hall, Eaglewood Cliffe, NJ), p. 62.

[49] C. Wunsch, *The Ocean Circulation InverseProblem* (Cambridge University Press, Cambridge, 1996), p. 99.

[50] A. N. Tikhonov and V. A. Arsenin, *Solution of Ill-Posed Problems* (Winston and Sons, New York, NY, 1977).

[51] 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," Surv. Math. Ind. **15**, 4066–4082 (2007).