A LSQR-type method provides a computationally efficient automated optimal choice of regularization parameter in diffuse optical tomography

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

(Received 11 September 2012; revised 4 December 2012; accepted for publication 27 January 2013; published 28 February 2013)

Purpose: Developing a computationally efficient automated method for the optimal choice of regularization parameter in diffuse optical tomography.
Methods: The least-squares QR (LSQR)-type method that uses Lanczos bidiagonalization is known to be computationally efficient in performing the reconstruction procedure in diffuse optical tomography. The same is effectively deployed via an optimization procedure that uses the simplex method to find the optimal regularization parameter. The proposed LSQR-type method is compared with the traditional methods such as L-curve, generalized cross-validation (GCV), and recently proposed minimal residual method (MRM)-based choice of regularization parameter using numerical and experimental phantom data.
Results: The results indicate that the proposed LSQR-type and MRM-based methods performance in terms of reconstructed image quality is similar and superior compared to L-curve and GCV-based methods. The proposed method computational complexity is at least five times lower compared to MRM-based method, making it an optimal technique.
Conclusions: The LSQR-type method was able to overcome the inherent limitation of computationally expensive nature of MRM-based automated way finding the optimal regularization parameter in diffuse optical tomographic imaging, making this method more suitable to be deployed in real-time.
 [http://dx.doi.org/10.1118/1.4792459]

Key words: near infrared imaging, diffuse optical tomography, image reconstruction, inverse problems, regularization parameter

I. INTRODUCTION

Near infrared (NIR) diffuse optical tomography (DOT) uses NIR light (600–1000 nm) as a probing media to obtain optical absorption and scattering images of the tissue under investigation.[1–4] Typically light is collected/delivered on tissue boundary through fiber optical bundles to increase the maximum throughput of transmitted light. As the NIR light is nonionizing and can penetrate thicker tissues, the main applications of NIR tomography are for soft tissue imaging, including breast and brain imaging.[2,3]

The NIR image reconstruction procedure, also called as the inverse problem, is known to be a nonlinear, ill-posed, and underdetermined problem, due to the dominance of scattering in the light propagation and limited boundary data available.[1] This procedure is iterative in nature, requiring regularization to result in an unique solution. This regularization plays an important role in determining the quality of reconstructed images. Typical selection of regularization parameter is either through the empirical choice or prior experience of the user. Recent works have shown that this regularization parameter could be automatically estimated using techniques such as generalized cross validation (GCV) and minimal residual method (MRM) in diffuse optical tomography.[5,6]

In diffuse optical tomography usage of MRM for automated selection of regularization parameter is shown to improve the spatial resolution when compared to traditional GCV-based method.[6] The limitation of that work was MRM-based algorithms are computationally expensive due to repeated solving of forward problem to estimate the optimal regularization parameter.[6] This makes the usage of the MRM-based automated choice of regularization in real-time prohibitive. It is highly desirable to develop a methodology, in principle, that uses the MRM-based technique, which can be deployed in real-time. The aim of this work is to show that conjugate gradient type LSQR (Refs. 7–9) (least squares QR) can provide the computational efficiency needed for automated selection of regularization parameter. Usage of LSQR method for image reconstruction has been widely prevalent,[9,10] specifically it was used in electrical impedance tomography to perform a computationally efficient reconstruction procedure.[9] In Ref. 9, the regularization parameter was estimated using L-curve based approach. In the proposed work, keeping the computational efficiency of LSQR-based reconstruction intact, we have also deployed it to find the optimal regularization parameter. Note that in reconstruction problems, such as diffuse optical tomographic imaging, another automated way of finding the regularization parameter based on L-curve method was shown to provide overly smooth solutions.[11] As LSQR algorithms typically require only matrix-vector computations (as oppose to matrix–matrix computations), it provides better computational efficiency as opposed to traditional methods that solve the ill-conditioned inverse problems.[7,8]

The LSQR method is analytically equivalent to conjugate-gradient least squares (CGLS), requiring lesser number of

operations compared to the traditional CGLS.[7,8] Moreover, the LSQR method exhibits better numerical properties compared to conjugate gradient type procedures when the problem is ill-conditioned (which is the case for diffuse optical tomographic image reconstruction).[7,8] The CGLS method (or its variant) has been proven to be effective for solving the inverse problem of NIR tomography in terms of work required at every iteration.[12,13] These CGLS methods are not as straightforward as the direct inversion (also known as full-Newton methods) due to the ambiguity in choosing the step size,[13,14] typically requiring an optimization procedure in choosing the same. For LSQR-type methods, the choice of number of iterations becomes critical, as it is known to have the same effect as regularization parameter in typical ill-conditioned inverse problem.[7] In this work, we have obtained the required optimal number of iterations based on an optimization procedure. The same optimal number of iterations has been used for estimating the optimal regularization parameter via a simplex method based optimization procedure. Despite deploying two optimization procedures, one for finding the optimal number of iterations and another for optimal regularization parameter, this method is computationally efficient compared to the MRM method (shown later). It will be shown that the obtained reconstructed images using LSQR-method is similar to the one obtained using MRM in both numerical and experimental cases.

For completeness, traditional methods such as L-curve[15] and generalized cross validation,[16] which are capable of finding the optimal regularization parameter in an automated way, were also discussed here. Similar to earlier work,[6] the discussion here is limited to two-dimensional (2D) continuous wave (CW) diffuse optical tomography, where the unknown imaging parameter is the optical absorption coefficient.

II. METHODS

II.A. CW diffuse optical tomographic imaging

Light propagation in thick soft biological tissues, such as breast, in the NIR wavelengths can be described using a diffusion equation (DE). For the CW case, the DE is given by[17]

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

where $Q_o(r)$ and $\Phi(r)$ represent the isotropic light source and photon density (real values), respectively, at position r. The diffusion coefficient $D(r)$ is represented as

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

where $\mu_a(r)$ and $\mu_s'(r)$ represent the absorption coefficient and reduced scattering coefficient, respectively. The Φ is found by solving this partial differential equation [Eq. (1)] using finite element method (popular numerical method for modeling irregular geometries) representing the forward model.[17–19] Modeled data ($G(\mu_a)$) are found by sampling photon density ($\Phi(r)$) at the measurements locations for each source position.[14] The experimental data y are considered to be the natural logarithm of the amplitude ($\ln(A)$).

II.B. Regularized minimal residual method (MRM)

Solving the inverse problem in diffuse optical tomography is a process that matches the modeled data with the experimentally measured boundary data in the least-square sense.[1,14] Penalty term is added in the minimization scheme for stabilizing the solution, as inverse problem is ill-posed, underdetermined, and nonlinear in this modality.[1] Hence the objective function is given by,

$$\Omega = \|y - G(\mu_a)\|^2 + \lambda\|\mu_a - \mu_{a0}\|^2, \tag{3}$$

where λ is the regularization parameter, which affects the resolution characteristics of image reconstruction.[18,20] The initial guess of absorption coefficient is represented as μ_{a0} (typically obtained using a calibration procedure[21]). One of the methods for finding the minima for the objective function [Eq. (3)] is by making the first order condition with respect to μ_a equal to zero (full-Newton method[14]), this results in the following update equation:

$$\triangle\mu_a = [J^T J + \lambda I]^{-1} J^T (y - G(\mu_a)), \tag{4}$$

where J represents the Jacobian matrix [$J = \frac{\partial G(\mu_a)}{\partial \mu_a}$ and of dimension $M \times N$ with M representing the number of measurement and N the number of finite element nodes], $\triangle\mu_a$ represents the update of μ_a, and I is the identity matrix.

The minimization problem given in Eq. (3) could be solved using minimal residual method.[6,22] This method description is given as Algorithm 1 of Ref. 23. This method is briefly reviewed here. The MRM method is equivalent of regularized steepest-descent method, which is an iterative method for solving normal equations. This flow of the method is given in Algorithm 1. The residual (r^i) provides the search direction and the step length ($k^i{}_\lambda$) is based on the positive definite operator defined by the regularization parameter (λ) here. This process is repeated until the norm of residual reaches a small value (ϵ_o), here, 10^{-6}. As this requires only matrix-vector multiplications, resulting in computational burden per step as $O(N^2)$, making it suitable for deployment of finding optimal regularization parameter. As $\triangle\mu_a$ is dependent on λ, the optimal λ should minimize the data-model misfit ($\|y - G(\mu_a)\|^2$). This principle makes estimation of optimal λ as an optimization problem where the minimization function is given by $\|y - G(\mu_a + \triangle\mu_a^\lambda)\|^2$ and $\triangle\mu_a^\lambda$ is obtained using MRM method (Algorithm 1).[6] This optimization problem is solved using simplex method, which offers a computationally

ALGORITHM I. Regularized Minimal Residual Method (MRM)

Calculation of update ($\triangle\mu_a$)
INPUT: J and δ; OUTPUT: $\triangle\mu_a$.
Initialize $\triangle\mu_a^0$ (initial guess), λ.
for $i = 0, 1, \cdots$ (representing inner iteration number)
1. $r^i = J\ \triangle\mu_a^i - \delta$
2. $l_\lambda^i = l_\lambda(\triangle\mu_a^i) = J^T r^i + \lambda\triangle\mu_a^i$
3. $k_\lambda^i = \frac{\|l_\lambda(\triangle\mu_a^i)\|^2}{\|Jl_\lambda(\triangle\mu_a^i)\|^2 + \lambda\|l_\lambda(\triangle\mu_a^i)\|^2}$
4. Update equation: $\triangle\mu_a^{i+1} = \triangle\mu_a^i - k_\lambda^i l_\lambda^i$
5. The iterative process, steps: 1-4, is terminated when the misfit reaches the given stopping criterion(ϵ_o):$\|r^i\|^2 \le \epsilon_o$

efficient gradient-free approach. The existence of optimal λ has been shown in the Appendix of Ref. 6.

One of the bottlenecks of this algorithms in finding optimal λ is the computational complexity associated with it, making it less preferred in the real-time. The equivalence of this approach is achieved through via proposed LSQR-type method, which is described in Sec. II.C.

II.C. LSQR-type method

Conjugate gradient type LSQR algorithm was previously used in diffuse optical tomography with early photons[10] for solving the linearized version of the objective function [Eq. (3)]. The LSQR-type algorithm deployment for estimating the optimal regularization parameter in diffuse optical tomography is the main contribution of this work. This is achieved by using Lanczos bidiagonalization of J as given in Ref. 7. The left and right Lanczos matrices and the bidiagonal matrix related to Jacobian matrix (J) are shown below:[9,24]

$$U_{k+1}(\beta_0 e_1) = \delta, \tag{5}$$

$$J V_k = U_{k+1} B_k, \tag{6}$$

$$J^T U_{k+1} = V_k B_k^T + \alpha_{k+1} v_{k+1} e_{k+1}^T. \tag{7}$$

Here B represents the lower bidiagonal matrix, U and V represent the left and right orthogonal Lanczos matrices, respectively. The unit vector of dimension $k \times 1$ is represented by e_k (=1 at the kth row and 0 elsewhere). The dimensions of U_k and V_k are $(M \times k)$ and $(N \times k)$, with k representing the number of iterations the bidiagonalization is performed. The δ is the data-model misfit $(=y - G(\mu_a))$ and u_i, v_i represent the left and right Lanczos vectors. The structures of U, V are given by[9,24]

$$U_k = [u_1, u_2, \ldots, u_k]; \quad V_k = [v_1, v_2, \ldots, v_k], \tag{8}$$

and B_k is a bidiagonal matrix having $\alpha_1, \ldots \alpha_k$ in the main diagonal and $\beta_1, \ldots, \beta_k$ is the lower subdiagonal of the matrix having a dimension of $((k+1) \times k)$.

As the aim of the least square problem is to match y with $G(\mu_a)$, one can Taylor expand the $G(\mu_a)$ around μ_{a0} (which could be a guess or approximation to original μ_a), leading to[20]

$$G(\mu_a) = G(\mu_{a0}) + J\triangle\mu_a + (\triangle\mu_a)^T H \triangle\mu_a + \cdots, \tag{9}$$

where H and J are the Hessian and Jacobian evaluated at μ_{a0}, respectively, and $\triangle\mu_a = \mu_a - \mu_{a0}$. Linearizing the above equation and assuming $\delta = y - G(\mu_{a0})$ produces a new objective function (linearized inversion) (Ref. 20):

$$\tilde{\Omega} = \|\delta - J\triangle\mu_a\|^2 \tag{10}$$

using Eqs. (5)–(7) in the argument of Eq. (10), resulting in[9,24]

$$\delta - J\triangle\mu_a = U_{k+1}(\beta_0 e_1 - B_k x^{(k)}), \tag{11}$$

where $\triangle\mu_a = V_k x^{(k)}$. Substitution of Eq. (11) in Eq. (10) results in

$$\tilde{\Omega} = \|\beta_0 e_1 - B_k x^{(k)}\|^2. \tag{12}$$

Considering the first order condition of Eq. (12), we get a new update equation as[9,24]

$$x^{(k)} = \left(B_k^T B_k + \lambda I\right)^{-1} \beta_0 B_k^T e_1, \tag{13}$$

where β_0 is the L2-norm of the data-model misfit (δ). Once $x^{(k)}$ is calculated, $\triangle\mu_a^{(k)}$ can be evaluated using $\triangle\mu_a^{(k)} = V_k x^{(k)}$. Note that the number of operations for obtaining $\triangle\mu_a^{(k)}$ in this approach is $O(N^2)$, whereas traditional one [Eq. (4)] requires $O(N^3)$ operations. As stated earlier, the number of iterations (k) plays an important role in determining the reconstructed image quality.

II.D. Estimation of optimal λ using a LSQR-type method

The advantage of LSQR-type method in finding the update ($\triangle\mu_a$) lies in its dimensionality reduction capability, which makes the update as $x^{(k)}$ [Eq. (13)] with $k \ll N$. This kind of evaluation of update turns out to be computationally more efficient when compared to traditional way of finding update using Eq. (4). Here taking the advantage of computational efficiency, we propose a new method of estimating the optimal regularization parameter (λ) using the simplex method based optimization scheme.[25] The objective function in Eq. (10) is minimized with respect to the regularization parameter λ. This is feasible as the update $\triangle\mu_a = V_k x^{(k)}$, where $x^{(k)}$ is a function of λ as given in Eq. (13).

The major role in this entire optimization scheme is played by k (number of iterations of the Lanczos bidiagonalization). This determines the size of the bidiagonal matrix [B_k having dimension of $(k+1) \times k$] and is an important factor in estimation of optimal regularization parameter. Increasing the number of Lanczos iterations is inversely proportional to the estimated optimal regularization parameter. The higher the k, the more is the ill-posedness of the problem, the lower λ will become more optimal (similar to L-curve[15]). The algorithm for determining the optimal number of iterations is given in Algorithm 2. The first 50 iterations of Lanczos iterations are only considered as for $k = 50$ the value of λ turns out to be

ALGORITHM II. Algorithm for determining the optimal number of Lanczos Iterations (k^{opt}) and optimal regularization parameter λ^{opt}

Input: Lanczos Bidiagonal Matrix B_k; V_k (k = 1,2, ..., 50); δ, β_0; J, μ_a, λ_{lim}.

Output: Optimal number of Lanczos iterations: k^{opt} and Optimal regularization parameter: λ^{opt}

Initialize e_1.

for k = 1,2, ..., 50

1. Estimate the optimal λ for the given k (λ_k^{opt}).
   –Use simplex method to find λ_k^{opt} in the range of [0 λ_{lim}.] which minimizes Eq. (10) with $\triangle\mu_a = V_k * x^{(k)}$, where $x^{(k)}$ is found using Eq. (13).
2. Compute $\triangle\mu_a$ with $\lambda = \lambda_k^{opt}$ via Eq. (13).
3. Make $\mu_a^k = \mu_a + \triangle\mu_a$ and compute G(μ_a^k) (modeled data).
4. Estimate $\Omega^k = \|y - G(\mu_a^k)\|^2$.

end

k^{opt} = index of minimal value of Ω^k and $\lambda^{opt} = \lambda_{k^{opt}}^{opt}$

$O(10^{-6})$, which is in the single precision limit. It is important to note that the optimal regularization parameter (λ^{opt}) is estimated at every Gauss–Newton iteration in both MRM- and LSQR-based methods, wherein the optimal λ is searched within the bound of 0 and previous Gauss–Newton iterations optimal λ (similar to the procedure adapted in Ref. 6). For the first iteration the bound (λ_{lim}) for the optimal λ estimation is 1000. The optimal regularization parameter estimated in both these methods reduces with each iteration.

The LSQR method based estimation of optimal regularization parameter is computationally more efficient than MRM-based estimation of regularization parameter, because it performs repeated computation of update using a sparse bidiagonal matrix [Eq. (13)]. MRM kind of estimation requires $O(P\times N^2)$ (where P is the number of inner iterations) computation while LSQR type of estimation of optimal regularization parameter requires $O(2\times Q\times k^2)$ (where Q is the number of function evaluations). Since the value of k is very small, it can be intuitively seen that LSQR kind of approach is more efficient compared to MRM. The same has been verified using numerical experiments and experimental phantom data. Moreover, the LSQR-type method is equivalent of traditional direct method of solving the minimization problem via Tikhonov regularization scheme and the equivalence is shown in the Appendix for a given λ.

In Secs. II.E and II.F, established methods for finding an optimal regularization parameter in cases where there is no prior information is available about the data noise level or the parameter distribution is discussed.

II.E. L-curve method

The L-curve method is one of the popular schemes for estimation of optimal regularization parameter for a linear inverse problem[15,16] and the same was used previously in diffuse optical tomography and was found to result in overly smooth solutions.[11] In adoption of this method to the diffuse optical tomographic inverse problem, it is applied at every linear step of the reconstruction problem. At every iteration (global) of the image reconstruction scheme, a graph is plotted between the residual norms of the misfit ($||J\triangle\mu_a - \delta||^2$) and the solution norm $||\triangle\mu_a||^2$ as function of regularization parameter (λ), as $\triangle\mu_a$ is a function of λ. This curve is suppose to be in L-shape, where the corner of L-curve will represent the optimal choice.[16] Note that this corner point will be at least distance from the origin.

In diffuse optical imaging, the L-curve does not exhibit a clear corner,[11] making the choice of optimal λ purely based on the least distance from the origin. Note also that, for nonlinear inverse problems the L-curve method results only an suboptimal regularization parameter.[14,16]

II.F. Generalized cross-validation

Another popular method for estimation of optimal regularization parameter is GCV.[5,26,27] The regularization parameter in case of GCV is estimated by minimizing a function $G(\lambda)$, where $G(\lambda)$ is defined as

$$G(\lambda) = \frac{||(JJ^T + \lambda I)^{-1}\delta||^2}{(\text{trace}(JJ^T + \lambda I)^{-1})^2}. \tag{14}$$

This method works on the principle that an omitted data point could be easily estimated using the regularized solution obtained with reduced data set. The regularized solution needs to be the optimal for such an estimation.

Both L-curve and GCV methods require repeated estimation of $\triangle\mu_a$ (or its equivalent), efficient way of achieving the same is using singular value decomposition (SVD) of Jacobian[6] and the same is utilized here.

All computations were carried out using open-source MATLAB-based NIRFAST package,[28] which uses finite element based numerical model for light propagation in tissue. The algorithm along with necessary MATLAB code is provided as an open-source for enthusiastic users to use this algorithm.[29] The Lanczos bidiagonalization was performed using the MATLAB-based regularization tools, which is also an open source.[16] A Linux workstation with dual six-core Intel Xeon processor 2.66 GHz with 64 GB RAM has been used to perform the computations in this work.

II.G. Quantitative analysis of algorithms

The estimation of optimal regularization parameter is linked with the optimal solution of the inverse problem. To better assess these solutions (reconstructed images) using the regularization parameters estimated through the above proposed method, quantitative metrics are necessary. Moreover, it will be highly desirable that these metrics do not require the expected/target absorption coefficient values, which are not feasible to obtain in the clinical case. In this work, two such metrics were used, namely, contrast-to-noise ratio (CNR) and contrast resolution.[30]

The CNR indicates how well the region of interest has been reconstructed and is given by[30,31]

$$\text{CNR} = \frac{\mu_{a_{\text{roi}}} - \mu_{a_{\text{back}}}}{\left(w_{\text{roi}}\sigma^2_{a_{\text{roi}}} + w_{\text{back}}\sigma^2_{a_{\text{back}}}\right)^{\frac{1}{2}}}, \tag{15}$$

where $\mu_{a_{\text{roi}}}$ and $\mu_{a_{\text{back}}}$ represent the mean of the reconstructed absorption coefficient, respectively, in the region of interest (ROI) and the background. The $\sigma_{a_{\text{roi}}}$ and $\sigma_{a_{\text{back}}}$ represent the standard deviation of the reconstructed absorption coefficient in the ROI and background, respectively. Here the weights $w_{\text{back}} = \frac{A_{\text{back}}}{A_{\text{tot}}}$ and $w_{\text{roi}} = \frac{A_{\text{roi}}}{A_{\text{tot}}}$ represent the ratio of areas between the background and total area as well as ROI and total area, respectively. Note that the theoretical limit on CNR for the detection of ROI by the human eye is 4 and above.[31] The higher the CNR value, the higher is the detectability of the ROI.

The contrast resolution C is a figure of merit that gives the distinguishability of the ROI with the background and is defined as

$$C = \frac{\mu_{a_{\text{roi}}} - \mu_{a_{\text{back}}}}{\mu_{a_{\text{roi}}} + \mu_{a_{\text{back}}}}. \tag{16}$$

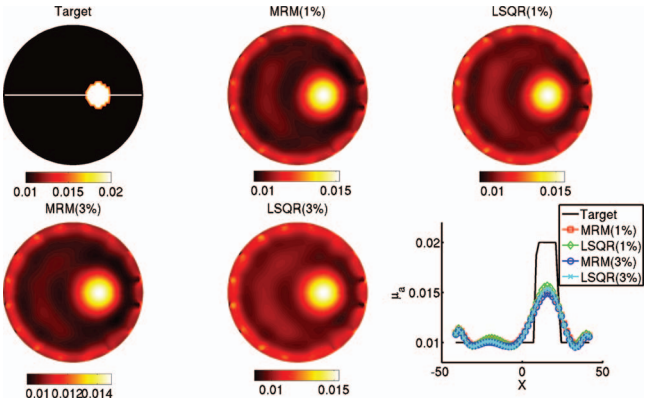

FIG. 1. The target μ_a distribution is shown on the top-left corner. The method that was used to find the optimal regularization parameter along with the data noise level is given on top of each reconstructed μ_a distribution. The one-dimensional cross-sectional plot of target and reconstructed μ_a distributions along the line shown in the target image are given in the bottom-right corner.

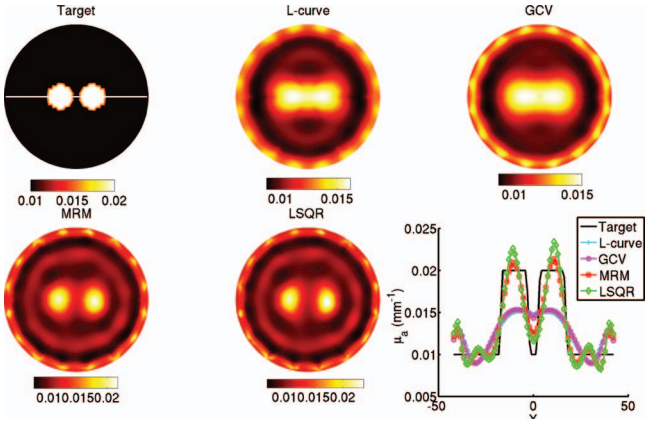

FIG. 2. Similar effort as Fig. 1 except the target μ_a distribution has two anomalies and the noise level in the data is kept at 1% and the estimation of optimal regularization parameter was found using L-curve, GCV, MRM, and the proposed LSQR methods.

The higher the C, better is the algorithm performance.

III. NUMERICAL AND EXPERIMENTAL EVALUATIONS

III.A. Numerical experiments

To asses the effectiveness of optimal estimation of the regularization parameter of LSQR-based proposed method, a circular mesh with background optical properties as $\mu_a = 0.01$ mm^{-1}, $\mu_s' = 1$ mm^{-1}, and uniform refractive index of 1.33 is considered. The diameter of the circular mesh is 86 mm. It had two regions mimicking fatty (with background optical properties) and tumor regions. The tumor (target) is a circular region of radius 7.5 mm centered at (15,0) having optical properties as $\mu_a = 0.02$ mm^{-1} and $\mu_s' = 1$ mm^{-1}. The target μ_a distribution is given in the Fig. 1 (top-left corner). To generate numerical experimental data, a fine mesh having 10 249 nodes (corresponding to 20 160 linear triangular elements) is used. These data are added with 1% and 3% normally distributed Gaussian noise to test the robustness of the method with increasing noise levels. The reconstructions were carried on a coarser mesh consisting of 1933 nodes (corresponding to 3726 linear triangular elements) after calibration of the data.[21] The data-collection system had 16 fibers arranged on the boundary of the circular domain, where when one fiber acts as a source, rest act as detectors. This set up results in $16\times15 = 240$ number of measurements (M). Each source was modeled as Gaussian source having full width at half maximum of 3 mm to mimic the experimental case[32] and is placed at one mean transport length inside the boundary. The calibrated data were reconstructed using the λ obtained by MRM and the proposed LSQR methods to show their equivalence.

The proposed method is also evaluated for multiple target case. The same meshes that were used as in the earlier case for generating the data and performing the reconstruction were used for this experiment also. The anomalies having a radius of 7.5 mm were placed at (10,0) and (−10,0) (hence the anomalies are separated by a distance of 5 mm, which is the typical resolution limit of diffuse optical tomography). The mesh had the optical properties for the two regions similar as explained for circular mesh (whose target distribution is given in the top-left most corner of Fig. 2). The images were reconstructed using the synthetic data with 1% Gaussian distributed noise. The data collection strategy was same as earlier case. Here, the λ is obtained in an automated fashion using each method discussed in this work and corresponding reconstruction was performed to compare the results.

III.B. Gelatin phantom experiment

The proposed scheme was also evaluated with experimental gelatin phantom data.[6] This gelatin phantom, having a radius 43 mm and height 25 mm, was made using a mixture of India ink for producing absorption effect and titanium oxide (TiO_2) for scattering. Layers of gelatin were fabricated by hardening heated gelatin solution (having a concentration of 20% of gelatin [G2625, Sigma Inc] and 80% of deionized water) successively to produce a breast mimicking phantom. The outer layer mimicking the adipose region has a thickness of 10 mm, and having $\mu_a = 0.0065$ mm^{-1} and $\mu_s' = 0.65$ mm^{-1}. The middle layer has a 76 mm diameter, which mimics the fibroglandular layer, having the optical properties at $\mu_a = 0.01$ mm^{-1} and $\mu_s' = 1.0$ mm^{-1}. A cylindrical hole extending in Z direction was filled with intralipid mixed with India ink acting as a tumor having optical properties of $\mu_a = 0.02$ mm^{-1} and $\mu_s' = 1.2$ mm^{-1} with a radius of 8 mm and height of 24 mm. The target laser diode with 785 nm wavelength was used as source for validation of individual layers optical properties from the data obtained using large cylindrical samples of each layer. The data are collected using a single layer of fibers (at $z = 0$ mm) leading to 240 data points. The collected data were calibrated using the coarser mesh with 1933 FEM nodes (corresponding to 3726 linear triangular elements). Similar to earlier multitarget cases, all methods that were discussed in this work were deployed for finding λ at

TABLE I. Comparison of total computational time (in seconds) including the overhead time using the methods discussed in this work for the results presented in Figs. 2 and 4. The total number of iterations taken to converge are given in the parentheses. The corresponding regularization parameter values are given in Table II.

Method	Multiple targets (Fig. 2.)	Experimental phantom (Fig. 4.)
L-curve (Sec. II.E)	16.34 (5)	8.94 (2)
GCV (Sec. II.F)	7.19 (3)	9.19 (4)
MRM (Sec. II.B)	133.42 (3)	179.98 (3)
LSQR (Sec. II.D)	25.92 (7)	15.17 (4)

every iteration and the corresponding reconstruction results were compiled for comparison.

IV. RESULTS

The reconstruction results using MRM and LSQR-type methods for the case of numerical experiment with single target and varying noise levels are shown in Fig. 1 along with the one-dimensional profile along the line of the target image. The method along with percentage of noise level in the data (in parentheses) is given on top of corresponding reconstructed μ_a distribution in Fig. 1. The total computation time for the L-curve, GCV, MRM, and the proposed method to converge has been reported in first and second columns of Table I, respectively. The results show that the proposed method performance is similar to MRM method, leading to conclusion that the proposed methods are in spirit similar to the MRM method. Figure 2 shows the reconstructed results obtained using the multiple target data with the regularization parameter obtained using L-curve, GCV, MRM, and the proposed LSQR methods. The method used for obtaining the regularization parameter is given correspondingly on top of each reconstructed μ_a distribution. These results also indicate that the LSQR type of optimal regularization estimation may be better than the traditional L-curve, GCV, and MRM-based estimation of optimal λ. For the results presented here, an example L-curve is plotted in Fig. 3 to show that typical diffuse optical image reconstruction scheme does not exhibit a clear

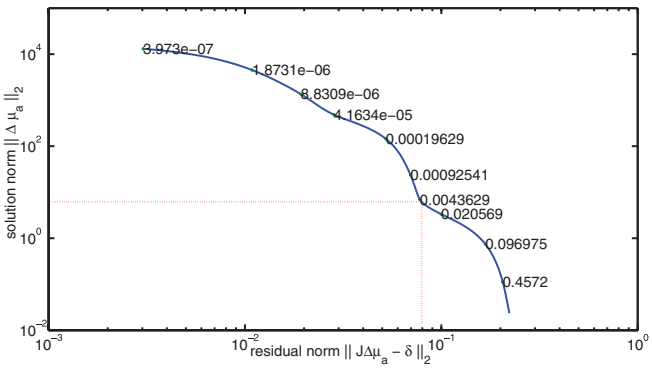

FIG. 3. An example L-curve for the last iteration of the result is presented in Fig. 2, which is a function of regularization parameter (given as text on the curve). The optimal regularization parameter is the corner of the curve (0.0043629), indicated by the dotted line.

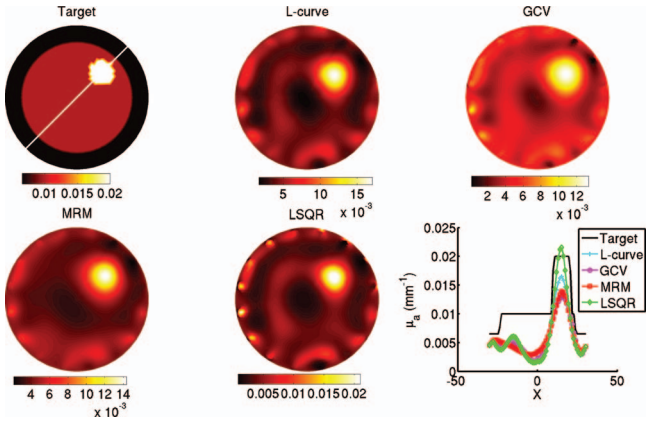

FIG. 4. Performance comparison of the proposed LSQR scheme with the existing techniques L-curve, GCV, and MRM for selection of optimal regularization parameter (λ) using the experimental gelatin phantom data, where the technique used for finding the optimal λ is given correspondingly on top of each distribution. The one-dimensional cross-sectional plot along the solid line indicated on the target image for the reconstructed images is given at the bottom right-hand corner of the figure.

corner, making the L-curve based automated estimation of optimal λ not effective.

The experimental results that were obtained using gelatin phantom data are given in Fig. 4 along with one-dimensional cross-sectional profiles on the right hand side of the same figure. Even here, the LSQR method performs better compared to other traditional methods. The total computational time and estimated λ's were given in Tables I and II correspondingly for the results presented in this work. The result in Fig. 4 also indicates that the contrast recovery of the proposed method was found to be higher than the standard methods resulting in quantitatively better reconstruction distributions using the proposed method. For an objective comparison of the results, the figures of merits that were discussed in this work, namely, CNR and C were compiled in Table III. This quantitative results affirm that performance of the proposed method is at least 16% higher in contrast resolution, with an added advantage of the proposed method being more computationally efficient to MRM method.

TABLE II. Comparison of regularization parameters obtained using the L-curve, GCV, MRM, and LSQR methods for the results presented in this work. The λ values are decreasing with increasing number of iterations for L-curve, MRM, and LSQR methods.

Method	Multiple targets (Fig. 2.)	Experimental phantom (Fig. 4.)
L-curve (Sec. II.E)	$\lambda = 0.0148, 0.0126, 0.0067, 0.0062, 0.0044$	$\lambda = 0.0034, 0.0035$
GCV (Sec. II.F)	$\lambda = 0.0181, 0.00696, 999.9$	$\lambda = 0.00325, 0.377, 0.888, 999.9$
MRM (Sec. II.B)	$\lambda = 0.0139, 0.0010, 5.6 \times 10^{-5}$	$\lambda = 0.0126, 0.0029, 4.7\times 10^{-4}$
LSQR (Sec. II.D)	$\lambda = 0.0048, 5.7\times 10^{-4}, 2.99\times 10^{-4}, 2.93\times 10^{-4}, 2.02\times 10^{-4}, 1.91\times 10^{-4}, 1.91\times 10^{-4}$	$\lambda = 0.0188, 0.0042, 0.0014, 4.4\times 10^{-4}$

TABLE III. Comparison of figures of merit CNR and contrast resolution (C) obtained for the reconstruction results using the methods discussed in this work.

Method	Target (Fig. 1.)	Target (Fig. 2.)	Experimental phantom (Fig. 4.)
L-curve (Sec. II.E)	...	CNR = 2.37 $C = 0.1459$	CNR = 5.41 $C = 0.4816$
GCV (Sec. II.F)	...	CNR = 2.46 $C = 0.1385$	CNR = 4.69 $C = 0.4470$
MRM (Sec. II.B)	CNR = 5.30 $C = 0.1536$	CNR = 2.62 $C = 0.2175$	CNR = 5.49 $C = 0.4507$
LSQR (Sec. II.D)	CNR = 5.27 $C = 0.1639$	CNR = 2.67 $C = 0.2279$	CNR = 5.87 $C = 0.5803$

In all these methods the value of optimal regularization parameter was estimated at each iteration, it is observed that the optimally chosen regularization parameter reduces with each Gauss–Newton iteration (refer to Table II). The computation time reported in Table I is the total time taken for performing reconstruction (including estimation of optimal λ and solving the inverse problem using this optimal λ). Even though the computational time for the proposed LSQR method is two times higher than that of the L-curve and GCV methods, the reconstruction results (Table III) prove that the proposed method indeed provides the optimal λ.

V. DISCUSSION AND CONCLUSIONS

The main advantage of the LSQR-based optimal estimation of regularization parameter is that it eliminates any heuristic (empirical) choice of reconstruction parameters, such as regularization parameter (λ). The proposed algorithm is compared with existing methods of automated estimation of optimal regularization parameter, such as L-curve, GCV, and MRM, it could be seen that the performance of the proposed method is better than the existing methods (Table III). The proposed algorithm is also efficient in terms of performing the reconstruction, compared to traditional matrix inversion. For the results presented here, the total time taken by traditional reconstruction [Eq. (4)] is 62.29 s for two anomalies case, 28.13 s for single target case, and 14.55 s for experimental case. The corresponding computation time taken by LSQR method is reported in Table I. The number of iterations taken for convergence in traditional reconstruction for two anomalies target case is 35, single anomaly target case is 16, and for experimental case is 7. For LSQR-type method the corresponding number of iterations were 7, 2, and 4, leading to a computationally efficient reconstruction procedure.

The proposed algorithm is also compared with traditional schemes such as L-curve and GCV scheme of optimal regularization parameter estimation. It could be seen from Fig. 3 that estimation of L-curve corner for obtaining a optimal regularization parameter is difficult as the plot of residual norm against solution norm does not result in an exact L-shape and the obtained solutions are not optimal as shown in Fig. 2. The inability of GCV-based scheme of estimation of optimal regularization parameter is explained briefly in Ref. 11, this may be mainly because the function may have multiple minima or the minimization function being too flat. The quantitative comparison of the results (Table III) also indicate that the proposed method performance is superior compared to its counter parts.

Even though it is possible for obtaining a optimal regularization parameter using generalized objective function in which the regularization term is explicit (rather than being identity as in this work), the proposed algorithm will loose its computational efficiency. If the explicit regularization term and Jacobian are not bidiagonalizable by the same orthogonal Lacozs matrices (may not be the case for all regularization matrices), the dimensionality reduction may not be explicitly possible for the update equation.

The proposed method requires solving of the forward model for estimating the optimal number of Lanczos iterations required, but the solving of forward model is performed only for fixed number of times ($k = 1, 2, \ldots, 50$) which is not the case with the MRM-based scheme (as it is solved repeatedly in a simplex based method[6]). Here, taking the advantage of computational efficiency of LSQR method, we have estimated the optimal regularization parameter using a simplex method. Among the existing methods for automated choice of regularization parameter, the LSQR-type method yielded optimal reconstructions compared to other existing methods. Even though, in principle, LSQR method is similar to MRM method, the better performance of the LSQR method in the results presented here is primarily due to the additional penalty in terms of Lanczos iterations, which reduces the dimensionality as well as ill-posedness of the problem. Development of these type of methods will remove the unwanted bias induced by the heuristic choice of regularization parameter, making diffuse optical tomographic image quality more objective in nature.

ACKNOWLEDGMENTS

The authors are thankful to NIR imaging group of Dartmouth for providing experimental data that were needed to carry out this work. This work is supported by the Department of Atomic Energy Young Scientist Research Award (Award No. 2010/20/34/6/BRNS) by Government of India.

APPENDIX: EQUIVALENCE OF TIKHONOV-TYPE AND LSQR-TYPE UPDATE EQUATIONS

For proving the equivalence, let the update from Tikhonov-type method be $\triangle\mu_a$ [Eq. (4)], which is obtained by solving

$$[J^T J + \lambda I_n]\triangle\mu_a = J^T\delta, \tag{A1}$$

where I_n is the identity matrix of dimensions $N \times N$. Let us consider the update from LSQR-type method to be $\triangle\tilde{\mu}_a$ with $\triangle\tilde{\mu}_a = V_k x^{(k)}$, the update equation [Eq. (13)] is given by

$$\left[B_k^T B_k + \lambda I_k\right]x^{(k)} = \beta_0 B_k^T e_1, \tag{A2}$$

where I_k represents the identity matrix having a dimension $k \times k$. Using Eq. (6), we get $B_k = U_{k+1}^T J V_k$, which makes Eq. (A2) as

$$\left[V_k^T J^T U_{k+1} U_{k+1}^T J V_k + \lambda I_k\right] x^{(k)} = \beta_0 V_k^T J^T U_{k+1} e_1. \quad \text{(A3)}$$

Now, from Lanczos algorithm, using Eq. (5), $U_{k+1} U_{k+1}^T = I_m$ and $V_k^T V_k = I_k$ in the above equation results in

$$V_k^T [J^T J + \lambda I_n] V_k x^{(k)} = V_k^T J^T \delta. \quad \text{(A4)}$$

Substituting $\triangle\tilde{\mu}_a = V_k x^{(k)}$ in Eq. (A4) and left multiplying both side by V_k result in

$$[J^T J + \lambda I_n] \triangle\tilde{\mu}_a = J^T \delta. \quad \text{(A5)}$$

Comparing Eq. (A5) with Eq. (A1) proves that $\triangle\mu_a = \triangle\tilde{\mu}_a$. Hence both these update equations result in same solution for a given regularization parameter (λ).

[a]Author to whom correspondence should be addressed. Electronic mail: phani@serc.iisc.in; Telephone: +91-80-2293 2496; Fax: +91-80-2360 2648.

[1]S. R. Arridge, "Optical tomography in medical imaging," Inverse Probl. **15**, R41–R93 (1999).
[2]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).
[3]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).
[4]D. R. Leff, O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Pattenl, 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).
[5]G. Golub and U. von Matt, "A Generalized cross-validation for large-scale problems," J. Comput. Graph. Stat. **6**, 1–34 (1997).
[6]R. P. K. Jagannath and P. K. Yalavarthy, "Minimum residual method provides optimal regularization parameter for diffuse optical tomography," J. Biomed. Opt. **17**, 106015 (2012).
[7]C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," ACM Trans. Math. Softw. **8**, 43–71 (1982).
[8]C. C. Paige and M. A. Saunders, "Algorithm-583: LSQR sparse linear equations and least-squares problems," ACM Trans. Math. Softw. **8**, 195–209 (1982).
[9]L. Hu, H. Wang, B. Zhao, and W. Yang, "A hybrid reconstruction algorithm for electrical impedance tomography," Meas. Sci. Technol. **18**, 813–818 (2007).
[10]G. M. Turner, G. Zacharakis, A. Soubret, J. Ripoll, and V. Ntziachristos, "Complete-angle projection diffuse optical tomography by use of early photons," Opt. Lett. **30**, 409–412 (2005).
[11]J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemp, V. Ntzi- achristos, B. Chance, and A. G. Yodh, "Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: Evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging," Med. Phys. **30**, 235–247 (2003).
[12]S. R. Arridge and M. Schweiger, "Gradient-based optimisation scheme for optical tomography," Opt. Exp. **2**, 212–226 (1998).
[13]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).
[14]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**, 2085–2098 (2007).
[15]P. C. Hansen and D. P. O. Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," SIAM J. Sci. Comput. **14**, 1487–1503 (1993).
[16]P. C. Hansen, "Regularization Tools Version 4.0 for Matlab 7.3," Numer. Algorithms **46**, 189–194 (2007).
[17]S. R. Arridge and M. Schweiger, "Photon-measurement density functions. Part 2: Finite-element-method calculations," Appl. Opt. **34**, 8026–8037 (1995).
[18]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).
[19]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).
[20]S. H. Katamreddy and P. K. Yalavarthy, "Model-resolution based regularization improves near infrared diffuse optical tomography," J. Opt. Soc. Am. A **29**, 649–656 (2012).
[21]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).
[22]M. S. Zhdanov, *Geophysical Inverse Theory and Regularization Problems*, Methods in Geophysics and Geochemistry, Vol. 36, 1st ed. (Elsevier, 2002).
[23]C. B. Shaw and P. K. Yalavarthy, "Effective contrast recovery in rapid dynamic near-infrared diffuse optical tomography using L1-norm-based linear image reconstruction method," J. Biomed. Opt. **17**, 086009 (2012).
[24]M. E. Kilmer and D. P. O'Leary, "Choosing regularization parameters in iterative methods for ill-posed problems," SIAM J. Matrix Anal. Appl. **22**, 1204–1221 (2001).
[25]J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," SIAM J. Optim. **9**, 112–147 (1998).
[26]N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," IEEE Trans. Image Process. **10**, 573–583 (2001).
[27]P. C. Hansen, J. G. Nagy, and D. P. O. Leary, *Deblurring Images: Matrices, Spectra, and Filtering*, 1st ed. (SIAM, Philadelphia, 2006).
[28]H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, and K. D. Paulsen, "Near infrared optical tomography using NIRFAST: Algorithms for numerical model and image reconstruction algorithms," Commun. Numer. Methods Eng. **25**, 711–732 (2009).
[29]https://sites.google.com/site/sercmig/home/lsqrreg.
[30]N. Ducros, C. D'Andrea, A. Bassi, G. Valentini, and S. Arridge, "A virtual source pattern method for fluorescence tomography with structured light," Phys. med. Biol. **57**, 3811–3832 (2012).
[31]X. Song, B. W. Pogue, S. Jiang, M. M. Doyley, H. Dehghani, T. D. Tosteson, and K. D. Paulsen, "Automated region detection based on the contrast-to-noise ratio in near-infrared tomography," Appl. Opt. **43**, 1053–1062 (2004).
[32]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).