

Nonquadratic penalization improves near-infrared diffuse optical tomography

Ravi Prasad K. Jagannath and Phaneendra K. Yalavarthy*

Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560 012, India

*Corresponding author: phani@serc.iisc.ernet.in

Received March 20, 2013; revised June 8, 2013; accepted June 10, 2013;
posted June 10, 2013 (Doc. ID 187491); published July 15, 2013

A new approach that can easily incorporate any generic penalty function into the diffuse optical tomographic image reconstruction is introduced to show the utility of nonquadratic penalty functions. The penalty functions that were used include quadratic (ℓ_2), absolute (ℓ_1), Cauchy, and Geman–McClure. The regularization parameter in each of these cases was obtained automatically by using the generalized cross-validation method. The reconstruction results were systematically compared with each other via utilization of quantitative metrics, such as relative error and Pearson correlation. The reconstruction results indicate that, while the quadratic penalty may be able to provide better separation between two closely spaced targets, its contrast recovery capability is limited, and the sparseness promoting penalties, such as ℓ_1 , Cauchy, and Geman–McClure have better utility in reconstructing high-contrast and complex-shaped targets, with the Geman–McClure penalty being the most optimal one. © 2013 Optical Society of America

OCIS codes: (170.0170) Medical optics and biotechnology; (100.3010) Image reconstruction techniques; (100.3190) Inverse problems; (110.6960) Tomography; (110.6955) Tomographic imaging; (110.0113) Imaging through turbid media.

<http://dx.doi.org/10.1364/JOSAA.30.001516>

1. INTRODUCTION

Near-infrared (NIR) diffuse optical tomography [1–4] is an emerging medical imaging modality that is capable of providing functional information of soft-tissue. Nonionizing NIR (600–1000 nm) light is delivered by using optical fibers on the surface of the tissue under investigation, making it an attractive noninvasive imaging modality, as tissues may be exposed several times without any risk of radiation-related hazards [3]. Because of multiple scattering, the light diffuses through the medium and is collected by the fibers placed at the boundary. From the limited measured intensities on the boundary of the imaging domain, the distribution of the internal optical properties of the tissue are estimated [5,6]. When multiple-wavelengths data is available, then clinically relevant functional properties of the tissues, such as oxy-hemoglobin, deoxy-hemoglobin, and water concentration can be determined by using this data [4].

Diffuse optical tomography recovers the internal distribution of the optical properties of tissue, and, in practice, it requires solving a nonlinear inverse problem [6] involving a numerical model for light propagation. The inverse problem associated with diffuse optical imaging is ill-posed and often underdetermined, as scattering is the dominant mechanism experienced by the NIR light in the tissue [4]. Solving this inverse problem, also known as the image reconstruction problem, requires powerful and advanced computational methods that can recover accurate solutions using limited noisy boundary measurements [4,7]. Moreover, to overcome the ill-posedness of the problem and to obtain a unique solution, a penalty term is deployed in the estimation process. The penalty term, also known as a regularization scheme, not only stabilizes the solution, but also promotes faster

convergence to the iterative solutions. More precisely, instead of only trying to minimize the data-model misfit, the objective function to be minimized for diffuse optical tomographic image reconstruction incorporates additional information about the image to be reconstructed by including a penalty term. Several penalty terms are proposed in the literature that can effectively be used to model the unknown optical property distribution [8–11]. In multimodal diffuse optical imaging systems, the structural information provided by the traditional imaging modalities has been incorporated into the penalty terms to improve diffuse optical image reconstruction [10]. The standard and the simplest method is Tikhonov-type (quadratic) regularization, in which the image quality is determined by the regularization parameter (scalar value). A higher value over smoothes the reconstructed image, which results in the loss of image quality due to poor spatial contrast and resolution, whereas a lower value amplifies the high-frequency noise [9,12].

Even though often maligned [12], quadratic penalties provide a useful reduction in image noise over a large range of spatial resolutions. Nonquadratic potential functions as penalty terms may induce additional nonuniformities beyond the resolution effects. In this case, the model of the imaging domain has to be consistent with the actual image to be reconstructed [13–17]. In this paper, a simple, yet effective, framework to incorporate nonquadratic penalty terms is introduced for the inverse problem encountered in diffuse optical tomography, and implementation of the same is achieved through MATLAB-based open-source NIRFAST [7]. Using this framework, three such nonquadratic penalty terms are deployed to show their usefulness in some special cases. Mainly the resolving power, robustness against noise levels, and

high-contrast recovery of the proposed penalty functions are compared with the standard uniform quadratic regularization parameter. Moreover, the generalized cross-validation (GCV) method is incorporated into this framework for automated choice of the regularization parameter for obtaining unbiased reconstruction results. The reconstruction results using different penalization schemes deployed in this work have been compared quantitatively via use of both relative error (RE) and Pearson correlation (PC) figures of merit. As the emphasis in this work is on investigating the effect of nonquadratic penalty functions on image reconstruction, the discussion is limited to the continuous-wave (CW) case. In this case, the amplitude of the light is the boundary data collected, from which the unknown optical absorption coefficient distribution is recovered.

2. METHODS

A. Diffuse Optical Tomography: Forward Problem

The modeling of CW NIR light propagation through a thick biological tissue is achieved by solving a diffusion equation (DE) [5,18], which is given as

$$-\nabla \cdot [D(r)\nabla\Phi(r)] + \mu_a(r)\Phi(r) = S_o(r), \quad (1)$$

where the optical diffusion coefficient is defined as $D(r)$, the optical absorption coefficient as $\mu_a(r)$, and the spatially resolved photon fluency density is denoted by $\Phi(r)$. The right-hand side of the DE, $S_o(r)$, represents the CW light source, and is modeled as isotropic. The diffusion coefficient is defined as

$$D(r) = \frac{1}{3[\mu_a(r) + \mu'_s(r)]}, \quad (2)$$

where $\mu_a(r)$ is absorption coefficient defined above and $\mu'_s(r)$ is the reduced scattering coefficient, which is defined as $\mu'_s = \mu_s(1-g)$ with μ_s as the scattering coefficient (and g is the anisotropy factor of the medium). Note that $\mu'_s(r)$ is assumed to be known and remains constant in the CW case. In order to account for the refractive-index mismatch at the boundary [19], a Type-III boundary condition (Robin-type) is employed for the DE. The solution of the DE [Eq. (1)] for the given distribution of the optical properties [$\mu_a(r)$] of the imaging domain is obtained by applying the finite element method (FEM) [20], and thus the forward problem solution gives the modeled data at the detector positions. The modeled data (y), under the Rytov approximation, is the natural logarithm of the intensity (I), i.e., $y = \ln(I)$, and this forward model is repeatedly used in solving the inverse problem [5].

B. Diffuse Optical Tomography: Inverse Problem

The inverse problem solving involves estimation of the unknown optical property distribution using the limited number of CW boundary measurements via a model-based approach. This is typically an iterative procedure, where the experimental measurements are matched iteratively with the modeled data in the least-squares sense over the range of μ_a . The inverse problem is popularly solved as an optimization problem [1,5–7,10,19,21] and involves direct calculation of the Jacobian (J , also known as the sensitivity matrix) with the help of

a forward model, followed by solving a linear system of equations to estimate the update in the optical properties.

1. Quadratic Penalization

A typical (standard) objective function (Ω) to be minimized with respect to μ_a can be written as follows:

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

where y denotes the experimental data, i.e., $y = \ln(I)$ ^{measured} and $G(\mu_a)$ denotes the model data for the given distribution of μ_a . Because of the ill-posedness of the problem, a regularization or penalty term is added; here it is the generalized Tikhonov regularization, one of the most popular penalty terms. The initial guess for the background absorption coefficient is μ_{a0} . Taylor's series expansion of the $G(\mu_a)$ around μ_{a0} gives

$$G(\mu_a) = G(\mu_{a0}) + G'|_{\mu_{a0}}(\mu_a - \mu_{a0}) + \dots, \quad (4)$$

where $G' = J = \frac{dG(\mu_a)}{d\mu_a}$ is the Jacobian. The Jacobian matrix has the dimensions of $NM \times NN$, where NM represents the number of measurements and NN represents the number of nodes in the finite element mesh. Ignoring the higher-order terms by assuming that the initial guess μ_{a0} is very close to the solution leads to the linearized inverse problem [20,22]; using Eq. (4) in Eq. (3),

$$\Omega = \|\delta - J\Delta\mu_a\|^2 + \lambda\|\Delta\mu_a\|^2, \quad (5)$$

where δ is the data model misfit, $\delta = [y - G(\mu_a)]$, and $\Delta\mu_a = (\mu_a - \mu_{a0})$. Minimizing the above objective function with respect to $\Delta\mu_a$ by applying the first-order derivative condition yields

$$[J^T J + \lambda I]\Delta\mu_a = J^T \delta. \quad (6)$$

This leads to the iterative update equation [22]

$$[J_{i-1}^T J_{i-1} + \lambda_i I]\Delta\mu_a^i = J_{i-1}^T \delta_{i-1}, \quad (7)$$

where $\Delta\mu_a^i$ represent the update of the optical absorption coefficient (μ_a) at the i th step, the Jacobian J_{i-1} is calculated from the distribution of the absorption coefficient from the $(i-1)$ th step (μ_a^{i-1}), which is obtained from the previous update $\Delta\mu_a^{i-1}$, and I is the identity matrix of order equal to number of FEM nodes (NN). The regularization parameter (λ_i), in Eq. (7) can be shown to be the ratio of the variance in the data and the estimated optical properties [10]. For a given problem, finding a suitable λ requires the prior information about the noise level in the data and also the noise characteristics of the image to be reconstructed, which may not be possible in a realistic scenario. In the present work using simulated data, 1% normally distributed Gaussian noise was added to the amplitude, which is a typical noise observed in experimental data [23]. The iterative procedure is stopped when the ℓ_2 -norm of the data-model misfit (δ) does not improve by more than 2%.

2. Nonquadratic Penalization

In the linearized objection function [Eq. (5)] instead of the Tikhonov regularization term, which is quadratic

Table 1. Different Penalty Functions and Their Derivatives Used for Estimation of Optical Properties from Eq. (10)

Penalty function	$\rho(\Delta\mu_a)$	$\rho'(\Delta\mu_a) = \frac{\partial\rho(\Delta\mu_a)}{\partial\Delta\mu_a}$	$D_{\Delta\mu_a} = \frac{\rho'(\Delta\mu_a)}{\Delta\mu_a}$
Quadratic (ℓ_2)	$\frac{1}{2} \left\ \frac{\Delta\mu_a}{\sigma_{\Delta\mu_a}} \right\ _2^2$	$\left(\frac{\Delta\mu_a}{\sigma_{\Delta\mu_a}^2} \right)$	$\frac{1}{\sigma_{\Delta\mu_a}^2}$
Absolute value (ℓ_1)	$\left\ \left(\frac{\Delta\mu_a}{\sigma_{\Delta\mu_a}} \right) \right\ _1$	$\frac{\text{sgn}(\Delta\mu_a)}{\sigma_{\Delta\mu_a}}$	$\frac{1}{\sigma_{\Delta\mu_a} \Delta\mu_a }$
Cauchy	$\frac{1}{2} \ln \left[1 + \left(\frac{\Delta\mu_a}{\sigma_{\Delta\mu_a}} \right)^2 \right]$	$\frac{\Delta\mu_a}{\sigma_{\Delta\mu_a}^2 + \Delta\mu_a^2}$	$\frac{1}{\sigma_{\Delta\mu_a}^2 + \Delta\mu_a^2}$
Geman–McClure	$\frac{1}{2} \left[\frac{\Delta\mu_a^2}{\sigma_{\Delta\mu_a}^2 + \Delta\mu_a^2} \right]$	$\frac{\Delta\mu_a \sigma_{\Delta\mu_a}^2}{\sigma_{\Delta\mu_a}^2 + \Delta\mu_a^2}$	$\frac{1}{(\sigma_{\Delta\mu_a}^2 + \Delta\mu_a^2)^2}$

(ℓ_2 -norm-based), one can deploy nonquadratic penalty terms [13–15,17]. Some of these penalties are given in Table 1 (including the ℓ_2 -norm-based one) with $\sigma_{\Delta\mu_a}^2$ representing the variance in $\Delta\mu_a$. Typically, these penalty terms are added to stabilize the solution, thus ensuring convergence [16,24] and removing high-frequency components [9,11]. Different penalty functions model the distribution of the update, in turn modeling the absorption coefficient of the domain. The inclusion of $\sigma_{\Delta\mu_a}^2$ in the update equation assists in controlling the reconstructed image quality at each iterative step. The generalized objective function in this case can be written as

$$\Omega = \frac{1}{2} \|\delta - J\Delta\mu_a\|^2 + \lambda\rho(\Delta\mu_a), \quad (8)$$

where $\rho(\Delta\mu_a)$ is one of the penalty function given in Table 1. Minimizing the above objective function with respect to $\Delta\mu_a$ and equating the first-order derivative to zero leads to

$$-J^T(\delta - J\Delta\mu_a) + \lambda\rho'(\Delta\mu_a) = 0. \quad (9)$$

Similar to Eq. (7), the update equation for the i th iteration becomes

$$[J_{i-1}^T J_{i-1} + \lambda_i D_{\Delta\mu_a^{i-1}}] \Delta\mu_a^i = J_{i-1}^T \delta_{i-1}, \quad (10)$$

where $(D_{\Delta\mu_a^{i-1}})$ is a diagonal penalty matrix consisting of values of $[\rho'(\Delta\mu_a^{i-1})/\Delta\mu_a^{i-1}]$ corresponding to each FEM node. The details of generalized inversion along with the comparison with the standard (quadratic penalty) are given in Appendix A. The value of λ_i is chosen optimally in an automated fashion using the GCV method (explained below). Note that the first iteration of the reconstruction scheme is same for all cases, where we use a standard quadratic penalty and the regularization parameter used is 0.01. From the second iteration onward the penalty terms given in Table 1 are deployed, and the regularization parameter (λ_i) is automatically chosen by using the GCV method.

The GCV method is the most popular method for estimating the optimal regularization parameter in ill-posed inverse problems [25,26]. The main advantage of this method is that it does not require accurate estimation of the noise-level (σ^2) present in the system for finding the optimal regularization parameter. The L-curve method is another popular scheme for estimation

of the optimal regularization parameter for a linear inverse problem, which typically assumes that the penalty term is quadratic [27,28]. Moreover, in diffuse optical tomography, L-curve based regularization parameter choice was found to result in overly smooth solutions [27,28]. Using the GCV method, obtaining an estimate for the regularization parameter used in Eq. (10) is achieved by minimizing a continuous function $G(\lambda)$, with $\lambda > 0$, which is given by

$$G(\lambda) = \frac{\frac{1}{NN} \| (I - A(\lambda)) \delta \|^2}{\left[\frac{1}{NN} \text{trace}(I - A(\lambda)) \right]^2}, \quad (11)$$

where

$$A(\lambda) = [J_{i-1} (J_{i-1}^T J_{i-1} + NN \cdot \lambda \cdot D_{\Delta\mu_a^{i-1}})^{-1} J_{i-1}^T]. \quad (12)$$

Here *trace* represents the trace of the matrix (sum of the diagonal entries). Since there exists an optimal regularization [29] or at least a suboptimal regularization for noisy data, a direct line search methods can be applied to find the optimal λ that minimizes $G(\lambda)$. For finding such an optimal λ , here a gradient-free simplex-method-type algorithm is used owing to its computational compactness [30,31].

3. NUMERICAL EXPERIMENTS

In order to investigate the different characteristics of the reconstructed images obtained as a result of using different penalty functions, five numerical experiments were conducted. In the initial four experiments, the imaging domain considered is circular in shape, having a radius of 43 mm, where the source–detector fibers are placed on the boundary. The optical properties of the background are at 0.01 mm^{-1} for the absorption coefficient and 1.00 mm^{-1} for the reduced scattering coefficient. For the generation of the experimental amplitude measurements [$y = \ln(I)$], a fine FEM mesh having 10,249 nodes (corresponding to 20,160 triangular elements) was used, whereas for the generation of modeled data $G(\mu_a)$, a coarser FEM mesh with 1785 nodes (corresponding to 3148 triangular elements) was used. The measurement setup consists of 16 equi-spaced fibers arranged along the circumference of the circular domain. When one fiber is used at a time as the source of NIR light, then the remaining 15 fibers act as detectors, resulting in 240 (16 \times 15) measurements. To mimic the experimental conditions, Gaussian-type modeling

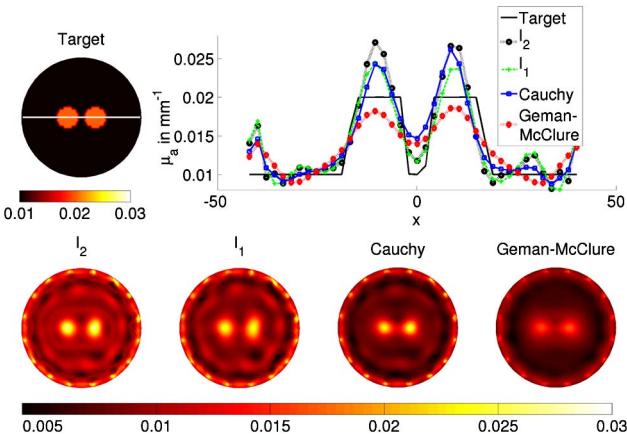


Fig. 1. Reconstructed images using the penalization functions discussed here (Table 1) in the numerical experiment with 1% noisy data where target distribution is given as in the top left-hand corner. The penalization functions used are given at the top of each corresponding reconstructed image. The one-dimensional cross-sectional profile along the solid line of the target distribution is given at the top right-hand corner.

of the source with a full-width at half-maximum of 3 mm was used [23].

Initially, two circular absorbing targets are placed along the x axis with a 5 mm separation and equidistant from the center of the circular domain with contrast of 2:1 compared to the background μ_a value. The target (expected) μ_a distribution is shown in Fig. 1 (top left-hand corner). The numerically generated data using a fine mesh were added with 1% (Fig. 1) and 3% (Fig. 2) white Gaussian noise to mimic a typical experimental scenario. Next, similar to Fig. 1, the circular absorbing targets are placed in an asymmetric position as shown in the top left-hand corner of Fig. 3. Here, to mimic experimental data, numerically generated data using the fine mesh were added with 1% white Gaussian noise. One more case, with a centrally absorbing target with a high contrast of 4:1 when compared with the background μ_a , was considered as the next test problem. The expected target distribution is given in the top left-hand corner of Fig. 4. In this case, numerically generated data was added with only 1% white Gaussian noise. The next test problem had two rectangular absorbers placed horizontally (centered at $[0, -14 \text{ mm}]$) and vertically (centered at

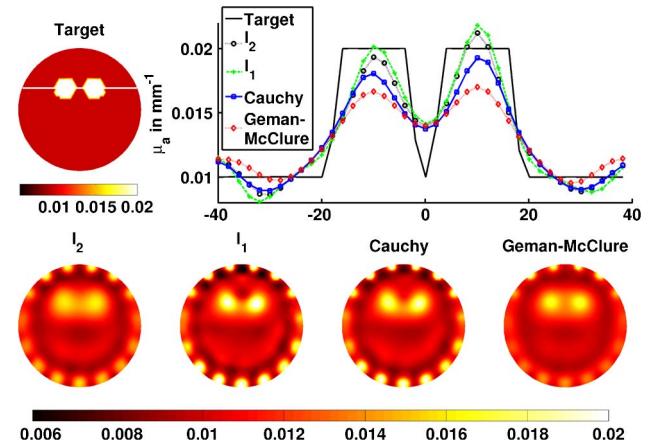


Fig. 3. Similar to Fig. 1 except that the absorption targets are placed asymmetrically close to the boundary (top left-hand corner).

$[-10 \text{ mm}, 0]$) to obtain an L-shaped target (shown in Fig. 5, top left-hand corner) with a uniform thickness of 7 mm. The absorption contrast of the target is 2:1. Similar to the previous case, 1% white Gaussian noise was added to the numerical data. In all cases, data calibration was performed to remove biases introduced due to irregularities in modeling source-detectors [32].

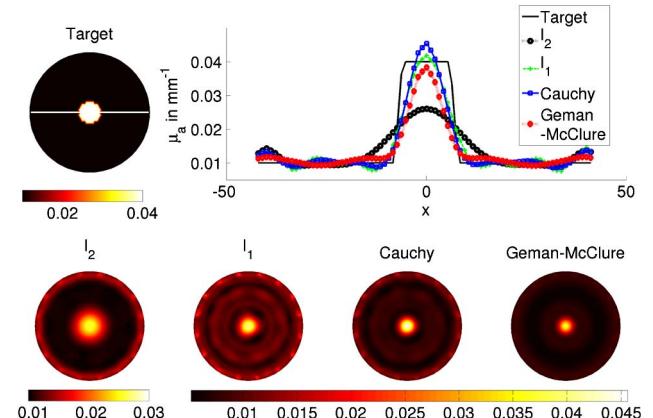


Fig. 4. Similar to Fig. 1 except the target is located at the center and has a contrast of 4:1.

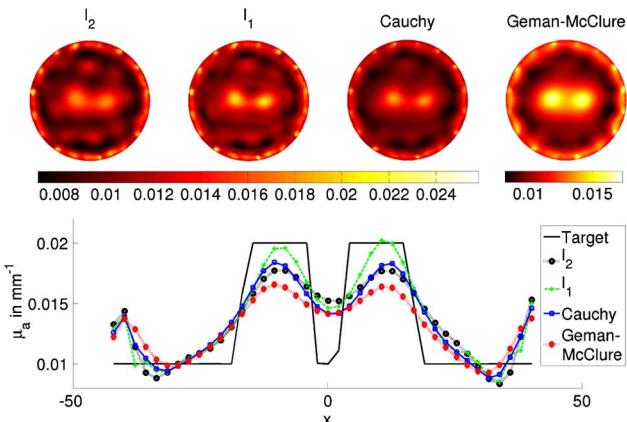


Fig. 2. Similar to Fig. 1 except that the data noise level is 3%; the target distribution is the same as in Fig. 1.

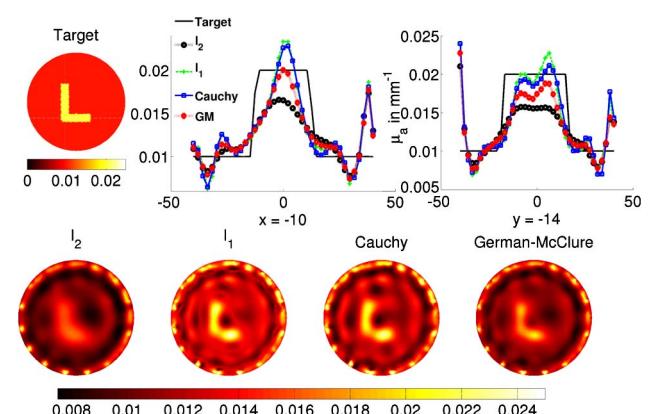


Fig. 5. Similar to Fig. 1 except for an L-shaped target.

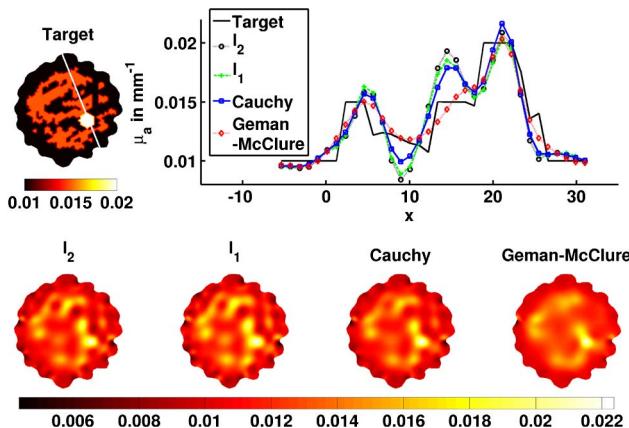


Fig. 6. Reconstructed optical images with various penalty functions listed in Table 1, using a realistic MRI derived patient mesh. The one-dimensional cross-sectional profile along the solid line of target distribution is given at the top right-hand corner.

In order to effectively assess the reconstruction performance of different penalizations, a numerical experiment involving a typical magnetic resonance imaging (MRI) NIR human breast mesh is considered. The segmented high-resolution MRI images clearly show three breast regions, and the corresponding optical properties of the tissue types used in this simulation case were $\mu_a = 0.01 \text{ mm}^{-1}$ for the fatty region, $\mu_a = 0.015 \text{ mm}^{-1}$ for fibro-glandular tissue, and tumor with $\mu_a = 0.02 \text{ mm}^{-1}$. The target distribution of the same is shown in the left-hand top corner of Fig. 6. The value of μ'_s was assumed to be known and was kept constant at 1 mm^{-1} throughout this imaging domain, also. The mesh used for data collection was a fine mesh with 4876 nodes (corresponding to 9567 triangular elements), and 1% Gaussian noise was added to mimic the real-life condition. For the reconstruction a coarse mesh with 1969 nodes (corresponding to 3753 triangular elements) was used.

A. Quantitative Analysis

To better assess the performance of different penalty functions deployed in the test cases presented here, a quantitative analysis was undertaken using the metrics RE and PC. As the studies involved here are numerical in nature, these metrics are easily computable.

The RE is a quantitative metric to measure the deviation of the reconstructed image from the target image and is defined as

$$\text{RE} = \left[\frac{\|(\mu_a)_{\text{TRUE}} - (\mu_a)_{\text{RECON}}\|_2}{\|(\mu_a)_{\text{TRUE}}\|_2} \right] \times 100, \quad (13)$$

where $(\mu_a)_{\text{TRUE}}$ and $(\mu_a)_{\text{RECON}}$, respectively, represent the target and the reconstructed images. The less RE, the better is the reconstruction performance of the penalty function.

PC quantitatively [33] measures the detectability in the reconstructed images by measuring the linear correlation between the target and the reconstructed image and is defined as

$$\text{PC}((\mu_a)_{\text{TRUE}}, (\mu_a)_{\text{RECON}}) = \frac{\text{COV}((\mu_a)_{\text{TRUE}}, (\mu_a)_{\text{RECON}})}{\sigma((\mu_a)_{\text{TRUE}})}, \quad (14)$$

where COV is the covariance and σ indicates the standard deviation. The higher the PC value, the better is the detectability of the anomalies in the imaging domain.

All reconstruction results presented in this work were performed on a Linux workstation with dual quad-core Intel Xeon processor 2.33 GHz with 64 GB of RAM.

4. RESULTS

The reconstruction results in the numerical experiment conducted using the target distribution having two circular targets (top left-hand corner of Fig. 1) are given in Figs. 1 and 2 for 1% and 3% noisy data cases, respectively. For the case of the absorption target's being close to the boundary (top left-hand corner of Fig. 3), the results are given in Fig. 3. The penalty functions used are given at the top of each reconstructed image, and the one-dimensional cross-sectional profile along the solid line of the target distribution is also given in the same figures. The corresponding penalty function forms are given in Table 1. The RE and PC quantitative metrics for the analysis of these results are given in Tables 2 and 3, respectively. These results indicate that the ℓ_2 , ℓ_1 , and Cauchy penalty function performance in resolving the anomalies is comparable, while the Geman–McClure has the best performance compared to the others. But even though the reconstruction is not visually showing good performance for the Geman–McClure, the RE value is the lowest, as it is evident from the profile plot that the variation from the target distribution is far less when compared with the other penalty functions.

The numerical experimental results (1% noisy data) of the high-contrast target placed at the center of the imaging domain are given in Fig. 4 along with the one-dimensional cross-sectional profile. Similar to earlier cases, the quantitative metrics that analyze the results are given in Tables 2 and 3. Here, it is clearly seen how the nonquadratic penalization gives better reconstruction than the traditional ℓ_2 -based

Table 2. Quantitative Comparison of RE [Eq. (13)] of Reconstruction Results Using the Discussed Penalty Functions

Figures	ℓ_2	ℓ_1	Cauchy	Geman–McClure
Fig. 1	30.3253	29.8520	26.7255	20.6825
Fig. 2	25.6591	24.9072	22.6244	20.0364
Fig. 3	29.1088	29.7643	27.4685	19.4516
Fig. 4	32.5844	30.0740	28.3723	29.2519
Fig. 5	35.6552	35.4859	31.5659	25.5824
Fig. 6	16.9519	16.9585	15.8667	15.2391

Table 3. Quantitative Comparison of Pearson Correlation (PC) [Eq. (14)] of the Reconstruction Results Using the Discussed Penalty Functions

Figures	ℓ_2	ℓ_1	Cauchy	Geman–McClure
Fig. 1	0.4794	0.4744	0.4825	0.5270
Fig. 2	0.4258	0.4599	0.4781	0.5283
Fig. 3	0.3884	0.4045	0.3907	0.5373
Fig. 4	0.6762	0.7193	0.7552	0.7944
Fig. 5	0.3516	0.3506	0.3321	0.3744
Fig. 6	0.6495	0.6502	0.6847	0.7046

penalization, both visually and also by one-dimensional line-profile plot shown in the same figure.

In Fig. 5, the reconstruction results for the L-shaped target is provided, using penalty functions described here. Similar to the earlier results, the corresponding figures of merit of the reconstructed results are given in Tables 2 and 3. These results indicate, in terms of quantitative metrics, that ℓ_2 - and ℓ_1 -based penalty performance is inferior compared to the others and that the nonconvex penalty functions (Cauchy and Geman–McClure) performance is superior.

The reconstruction results with a patient mesh obtained from MRI images using various penalty functions are given in the Fig. 6 along with the one-dimensional cross-sectional profiles. Tables 2 and 3 also have the quantitative comparison of these reconstruction results. Quantitatively the performance of all penalty functions is similar; the Geman–McClure result shows that it is effective in suppressing the background artifacts (also true for the results presented in Fig. 5). Further, the shape recovery more closely follows the expected distribution (line profile of Fig. 6).

5. DISCUSSION

The inverse problem of estimating the internal distribution of tissue optical properties in NIR diffuse optical tomography is a challenging task and involves using advanced numerical techniques to get reasonable estimates. The ill-posedness encountered in the inverse problem is compensated by using penalty or regularization terms to get a unique solution. Often, these penalties dictate the reconstructed image characteristics. Many penalties have been proposed in the literature; more often they are based on prior information available to the user about the imaging domain and data-noise level [34–37]. Here, the framework was more tailored toward use of nonquadratic penalties, especially in cases where no prior information, including the structural information, about the tissue under investigation and data-noise level is available. The proposed generalized framework can easily incorporate any penalty function, including the standard quadratic penalty, for reconstruction of optical properties. Also, the proposed scheme is independent of any heuristics imposed by the choice of the regularization parameter (λ_i), which is automatically estimated in this work by using the GCV method.

Even though there are only three other penalties, other than the standard quadratic penalty, discussed in this paper, the proposed framework can be extended to include any generic penalty function. The quantitative analysis of the results via the figures of merit has provided some insights into the deployment of appropriate penalties for a given problem at hand. For example, if the imaging problem demands suppression of boundary artifacts in the reconstructed images (Fig. 5), use of traditional quadratic (ℓ_2 norm) penalties should be encouraged. On the other hand, if one is interested in the recovery of the shapes of the target, the use of a nonquadratic penalty should be discouraged (Fig. 4). For the case of a high-contrast target (Fig. 3), the traditional method (ℓ_2 -norm based) may not be optimal, as it encourages smooth solutions compared with other nonquadratic penalties. Also, it is evident from other results that the nonconvex penalties (Cauchy and Geman–McClure) may be the optimal ones for the cases discussed here, as they promote sparse solutions. The Geman–McClure penalty function yields a sparser result

compared to Cauchy [14]. From the figures of merit, one could see that Geman–McClure performance is superior compared to the others; visually, the Geman–McClure also produces more appealing results by reducing the background artifacts (Figs. 4 and 6).

Even though the cases considered here are limited in nature, our aim in this work is to show that nonquadratic penalization can improve the diffuse optical tomographic image reconstruction by providing additional information or feedback into the minimization scheme. Note that as the quantitative comparison of the results requires adequate information about the region of interest, the discussion was limited to the numerical experiments. Future work includes extending this work to real-time phantom and patient data cases.

6. CONCLUSIONS

For diffuse optical tomographic reconstruction, a new framework that can easily incorporate nonquadratic penalty terms was developed, and through numerical cases these penalization schemes were systematically compared with one another. The nonquadratic penalty terms that were used in this work include, ℓ_1 , Cauchy, and Geman–McClure. The GCV method was also deployed in this work for automatic estimation of the regularization parameter in each of these penalization schemes to remove any heuristics involved in the optical property estimation process. The results indicate that nonquadratic penalization has advantage compared to traditional quadratic penalty in terms of improving the reconstructed image characteristics, especially in cases where high-contrast and complex shapes need to be reconstructed.

APPENDIX A: IMAGE RECONSTRUCTION THROUGH GENERIC PENALTY FUNCTIONS

Diffuse optical tomographic image reconstruction is an iterative procedure in which modeled data is matched with experimental data in the least-squared sense as explained in Subsection 2.B. So the objective function to be minimized is posed as follows:

$$\Omega = \|y - G(\mu_a)\|^2. \quad (15)$$

Since the inverse problem is ill posed, it requires regularization to be added to the cost function, which becomes [34]

$$\Omega = \frac{1}{2} \|y - G(\mu_a)\|^2 + \lambda \rho(\mu_a - \mu_{a0}), \quad (16)$$

where the penalty function is denoted $\rho(\mu_a - \mu_{a0})$, with μ_{a0} being the initial guess (typically assumed to be a uniform scalar value). Taking the first-order approximation to $G(\mu_a)$ results in a linear inverse problem, whose objective function is written as

$$\Omega = \frac{1}{2} \|(\delta - J\Delta\mu_a)\|^2 + \lambda \rho(\Delta\mu_a), \quad (17)$$

where $\Delta\mu_a = \mu_a - \mu_{a0}$. To minimize the above objective function, we have to differentiate the $\rho(\Delta\mu_a)$ function. If it is possible, then we can find the first-order derivative condition for the function minimization as

$$-J^T(\delta - J\Delta\mu_a) + \lambda\rho'(\Delta\mu_a) = 0. \quad (18)$$

Rearranging the terms above, we get the condition to be satisfied,

$$-J^T\delta + J^TJ\Delta\mu_a + \lambda\rho'(\Delta\mu_a) = 0. \quad (19)$$

This can be further simplified to

$$[J^TJ + \lambda D_{\Delta\mu_a}]\Delta\mu_a = J^T\delta, \quad (20)$$

where $D_{\Delta\mu_a}$ is a diagonal matrix determined by the nonquadratic penalty function $\rho'(\Delta\mu_a)$,

$$(D_{\Delta\mu_a})_{ii} = \left(\frac{\rho'(\Delta\mu_a)}{\Delta\mu_a} \right)_i \quad \text{for } i = 1, 2, \dots, \text{NN}, \quad (21)$$

where NN represents the number of FEM nodes reconstructed. Actual functional forms of $D_{\Delta\mu_a}$ of some of the penalty functions are listed in the fourth column of Table 1. An iterative scheme similar to the standard image reconstruction procedure [Eq. (7)] can be formulated in a similar way and is given in Subsection 2.B.1 [Eq. (10)]. The first row of Table 1 gives the ℓ_2 -norm-based penalization, and the last column of the first row gives the diagonal matrix $D_{\Delta\mu_a}$,

$$(D_{\Delta\mu_a})_{ii} = \frac{1}{\sigma_{\Delta\mu_a}^2} \quad \text{for } i = 1, 2, \dots, \text{NN}, \quad (22)$$

where $\sigma_{\Delta\mu_a}^2$ represents the variance of $\Delta\mu_a$, and this makes $D_{\Delta\mu_a}$ a scalar matrix. Now this gives the updated equation, which is same as the traditional update [refer to Eq. (7)] wherein the regularization parameter scales $D_{\Delta\mu_a}$. On careful observation, for the ℓ_2 -norm-based penalization, finding the update using Eq. (20) requires only the knowledge of the variance of the update, and it is part of the regularization parameter, making $D_{\Delta\mu_a}$ a scalar multiple of the identity matrix.

Different functional forms of $\rho'(\Delta\mu_a)$ along with the traditional ℓ_2 -norm-based functions is listed in the second column of Table 1, and, in this table, note that the standard deviation of the $\Delta\mu_a$ is represented as $\sigma_{\Delta\mu_a}$.

Similarly if one takes different nonquadratic penalty functions as given in Table 1, we can get different forms of the diagonal matrix $D_{\Delta\mu_a}$ in the update Eq. (20). For any generic penalty function it is not possible to find the diagonal matrix $D_{\Delta\mu_a}$, as $\Delta\mu_a$ itself is the unknown; here we are using information obtained from the previous iteration to find the diagonal matrix, and for starting the iteration we are using a single scalar value for all nodes, which is obtained from the GCV method explained in the Subsection 2.B. So, the modeling parameters for defining all the penalty functions in this approach are progressively generated from the previous update and are independent of any heuristics.

ACKNOWLEDGMENTS

This work is supported by the Department of Atomic Energy, Government of India, through DAE young scientist research award (Sanction No. 2010/20/34/6/BRNS). The authors are also thankful to the NIR imaging group of Dartmouth for

providing necessary finite element meshes that were needed to carry out this work. Ravi Prasad K. Jagannath acknowledges a University Grants Commission—Senior Research Fellowship (UGC-SRF). The authors also acknowledge the IISc Mathematic Initiative program.

REFERENCES

- D. A. Boas, D. H. Brooks, E. L. Miller, C. A. DiMarzio, M. Kilmel, R. J. Gaudette, and Q. Zhang, "Imaging the body with diffuse optical tomography," *IEEE Signal Process. Mag.* **18**, 57–75 (2001).
- S. L. Jacques and B. W. Pogue, "Tutorial on diffuse light transport," *J. Biomed. Opt.* **13**, 041302 (2008).
- A. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical tomography," *Phys. Med. Biol.* **50**, R1–R43 (2005).
- A. Gibson and H. Dehghani, "Diffuse optical imaging," *Phil. Trans. R. Soc. A* **367**, 3055–3072 (2009).
- S. R. Arridge, "Optical tomography in medical imaging," *Inverse Probl.* **15**, R41–R93 (1999).
- S. R. Arridge and J. C. Hebden, "Optical imaging in medicine: II. Modeling and reconstruction," *Phys. Med. Biol.* **42**, 841–853 (1997).
- H. Dehghani, S. Srinivasan, B. W. Pogue, and A. Gibson, "Numerical modelling and image reconstruction in diffuse optical tomography," *Phil. Trans. R. Soc. A* **367**, 3073–3093 (2009).
- X. Intes, C. Maloux, M. Guven, B. Yazici, and B. Chance, "Diffuse optical tomography with physiological and spatial a priori constraints," *Phys. Med. Biol.* **49**, N155–N163 (2004).
- B. W. Pogue, T. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, "Spatially variant regularization improves diffuse optical tomography," *Appl. Opt.* **38**, 2950–2961 (1999).
- 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).
- 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).
- J. W. Stayman and J. A. Fessler, "Spatially-variant roughness penalty design for uniform resolution in penalized-likelihood image reconstruction," in *Proc. of International Conference on Image Processing (ICIP)*, Chicago, IL, (4–7, October 1998, Vol. 2, IEEE), pp. 685–689.
- A. H. Hielscher and S. Bartel, "Use of penalty terms in gradient-based iterative reconstruction schemes for optical tomography," *J. Biomed. Opt.* **6**, 183–192 (2001).
- P. Zwartjes and A. Gisolf, "Fourier reconstruction with sparse inversion," *Geophys. Prospect.* **55**, 199–221 (2007).
- P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.* **6**, 298–311 (1997).
- R. Acar and C. R. Vogel, "Analysis of bounded variation penalty methods for ill-posed problems," *Inverse Probl.* **10**, 1217–1229 (1994).
- M. D. Sacchi, T. J. Ulrych, and C. Walker, "Interpolation and extrapolation using a high resolution discrete Fourier transform," *IEEE Trans. Signal Process.* **46**, 31–38 (1998).
- 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).
- 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).
- S. R. Arridge and M. Schweiger, "Photon-measurement density functions. Part 2: finite-element-method calculations," *Appl. Opt.* **34**, 8026–8037 (1995).
- 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).

22. 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).
23. 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).
24. R. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems* (Academic, 2005).
25. G. Golub and U. von Matt, "Generalized cross-validation for large-scale problems," *J. Comput. Graph. Stat.* **6**, 1–34 (1997).
26. N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Trans. Image Process.* **10**, 573–583 (2001).
27. J. P. Culver, R. Choe, M. J. Holboke, L. Zubkov, T. Durduran, A. Slemp, V. Ntziachristos, 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).
28. 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).
29. R. P. K. Jagannath and P. K. Yalavarthy, "Minimal residual method provides optimal regularization parameter for diffuse optical tomography," *J. Biomed. Opt.* **17**, 106015 (2012).
30. 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).
31. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).
32. B. W. Pogue, C. Abele, H. Kaufman, and K. D. Paulsen, "Calibration of near infrared frequency-domain tissue spectroscopy for absolute absorption coefficient quantitation in neonatal head-simulating phantoms," *J. Biomed. Opt.* **5**, 185–193 (2000).
33. K. B. Flach, R. Kueres, W. Semmler, M. Kachelrie, and S. Bartling, "Constrained reconstructions for 4D intervention guidance," *Phys. Med. Biol.* **58**, 3283–3300 (2013).
34. 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).
35. H. Niu, P. Guo, L. Ji, Q. Zhao, and T. Jiang, "Improving image quality of diffuse optical tomography with a projection-error-based adaptive regularization method," *Opt. Express* **16**, 12423–12434 (2008).
36. N. Cao, A. Nehorai, and M. Jacob, "Image reconstruction for diffuse optical tomography using sparsity regularization and expectation–maximization algorithm," *Opt. Express* **15**, 13695–13708 (2007).
37. M. E. Eames and H. Dehghani, "Wavelength dependence of sensitivity in spectral diffuse optical imaging: effect of normalization on image reconstruction," *Opt. Express* **16**, 17780–17791 (2008).