Comparing Two Regularization Techniques for Diffuse Optical Tomography

Scott C. Davis\textsuperscript{1}, Hamid Dehghani\textsuperscript{1,2}, Phaneendra K. Yalavarthy\textsuperscript{1}, Brian W. Pogue\textsuperscript{1}, and Keith D. Paulsen\textsuperscript{1}

\textsuperscript{1}Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA
\textsuperscript{2}School of Physics, University of Exeter, UK
email: scott.c.davis@dartmouth.edu

ABSTRACT
Two techniques to regularize the diffuse optical tomography inverse problem were compared for a variety of simulated test domains. One method repeats the single-step Tikhonov approach until a stopping criteria is reached, regularizing the inverse problem by scaling the maximum of the diagonal of the inversion matrix with a factor held constant throughout the iterative reconstruction. The second method, a modified Levenberg-Marquardt formulation, uses an identical implementation but reduces the factor at each iteration. Four test geometries of increasing complexity were used to test the performance of the two techniques under a variety of conditions including varying amounts of data noise, different initial parameter estimates, and different initial values of the regularization factor. It was found that for most cases tested, holding the scaling factor constant provided images that were more robust to both data noise and initial homogeneous parameter estimates. However, the results for a complex test domain that most resembled realistic tissue geometries were less conclusive.

Keywords: Diffuse optical tomography, image reconstruction

1. INTRODUCTION
Diffuse optical tomography (DOT) for the detection and characterization of diseased tissue has been studied extensively over the past several years. The initial promise of the technique has been tempered by the inherently low image resolution and frequent occurrence of image artifacts when a-priori information is not incorporated into the imaging algorithm. The imaging problem is under-determined and ill-posed requiring a regularized inversion technique for image recovery. Appropriate choice of the regularization technique is critical for accurately estimating image parameters. However, despite years of research with highly evolved imaging systems, regularization techniques vary widely throughout the DOT community. Formal methods to optimize the regularization parameter, such as L-curve analysis [1-4], have been employed, though approaches derived empirically from experimental phantom data are often preferred. This study compares two regularization approaches on simulated data in a variety of test geometries and under a wide range of reconstruction conditions.

2. METHODS
Photon propagation through tissue is commonly approximate as a diffusive field due to the high degree of scattering in the media. The photon fluence rate $\Phi(r, \omega)$ at position $r$ can be described by the diffusion equation, presented here in the frequency domain

$$-\nabla \cdot \kappa(r) \nabla \Phi(r, \omega) + \left(\mu_a(r) + \frac{i\omega}{c(r)}\right) \Phi(r, \omega) = q_0(r, \omega)$$

(1)
where parameters $\mu_a$ and $\mu_s$ are the absorption and reduced scattering coefficients respectively, $q_0(r, \omega)$ is an isotropic source and the diffusion coefficient is given by $\kappa = \frac{1}{3(\mu_a + \mu_s)}$. The speed of light in the medium, $c(r)$ is defined by $c_0/n(r)$, where $n(r)$ is the index of refraction at and $c_0$ is the speed of light in vacuum.

A Finite Element Method (FEM) implementation on is used to solve the forward problem for fluence rate in any given domain. The inverse problem involves minimizing the objective function defined as to simultaneously recover optical properties $\mu = (\mu_a, \kappa)$:

$$\chi^2 = \sum_{i=1}^{NM} (\Phi_i^C - \Phi_i^{Meas})^2 + \lambda \sum_{j=1}^{NN} f(\mu_j - \mu_0)^2$$

where $NM$ is the total number of measurements, $NN$ the number of nodes in the FEM mesh, $\Phi_i^C$ and $\Phi_i^{Meas}$ are the computed and measured fluence rates at the detectors, respectively, $\lambda$ the regularization parameter, $I$ the identity matrix, and $\mu_0$ an initial estimate of parameters $\mu = (\mu_a, \kappa)$. Minimization of this function provides the update equation for each iteration:

$$\Delta \mu = J^T [JJ^T + \lambda I]^{-1} (\Phi^{Meas} - \Phi^C)$$

where $J$ is the Jacobian matrix divided into sub-matrices accounting for absorption and diffusion coefficients simultaneously and can be expressed as

$$J = \begin{bmatrix}
\delta \ln I_1 & \delta \ln I_1 & \cdots & \delta \ln I_1 & \delta \ln I_1 & \delta \ln I_1 & \delta \ln I_1 & \delta \ln I_1 & \cdots & \delta \ln I_1 \\
\delta \kappa_1 & \delta \kappa_2 & \cdots & \delta \kappa_j & \delta \mu_{a1} & \delta \mu_{a2} & \cdots & \delta \mu_{aj} \\
\delta \theta_1 & \delta \theta_1 & \cdots & \delta \theta_1 & \delta \theta_1 & \delta \theta_1 & \cdots & \delta \theta_1 \\
\delta \ln I_2 & \delta \ln I_2 & \cdots & \delta \ln I_2 & \delta \ln I_2 & \delta \ln I_2 & \delta \ln I_2 & \delta \ln I_2 & \cdots & \delta \ln I_2 \\
\delta \kappa_1 & \delta \kappa_2 & \cdots & \delta \kappa_j & \delta \mu_{a1} & \delta \mu_{a2} & \cdots & \delta \mu_{aj} \\
\delta \theta_2 & \delta \theta_2 & \cdots & \delta \theta_2 & \delta \theta_2 & \delta \theta_2 & \cdots & \delta \theta_2 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
\delta \ln I_S & \delta \ln I_S & \cdots & \delta \ln I_S & \delta \ln I_S & \delta \ln I_S & \delta \ln I_S & \delta \ln I_S & \cdots & \delta \ln I_S \\
\delta \kappa_1 & \delta \kappa_2 & \cdots & \delta \kappa_j & \delta \mu_{a1} & \delta \mu_{a2} & \cdots & \delta \mu_{aj} \\
\delta \theta_S & \delta \theta_S & \cdots & \delta \theta_S & \delta \theta_S & \delta \theta_S & \cdots & \delta \theta_S \\
\delta \kappa_1 & \delta \kappa_2 & \cdots & \delta \kappa_j & \delta \mu_{a1} & \delta \mu_{a2} & \cdots & \delta \mu_{aj}
\end{bmatrix}$$

where $\frac{\delta \ln I_j}{\delta \kappa_j}$ and $\frac{\delta \ln I_j}{\delta \mu_{aj}}$ are the sub-matrices that define the relation between the log of the amplitude of the $i^{th}$ measurement with respect to $\kappa$ and $\mu_a$ at the $j^{th}$ reconstructed nodes respectively; $\frac{\delta \theta_i}{\delta \kappa_j}$ and $\frac{\delta \theta_i}{\delta \mu_{aj}}$ are the sub-matrices...
that define the relation between the phase of the \( i^{th} \) measurement with respect to \( \kappa \) and \( \mu_\alpha \) at the \( j^{th} \) node respectively. \( S \) is the total number of measurements. Since the formulation used in this inversion is \( JJ^T \) rather than \( J^TJ \), the components on the diagonal of this product alternate dependence between amplitude and phase. A regularization technique which follows this alternating pattern is used to regularize the phase and amplitude data separately. In the implementation used here, the regularization parameter is based on a constant which is scaled to the maximum of the amplitude and phase sub-diagonals of \( JJ^T \) and can be expressed as

\[
\lambda_i^{\text{amplitude}} = \beta(\max(\text{diag}(JJ^T(\ln I))))
\]

\[
\lambda_i^{\text{phase}} = \beta(\max(\text{diag}(JJ^T(\theta))))
\]

where \( \beta \) is a starting constant determined from formal methods of numerical analysis or chosen empirically from experimental data. Scaling the constant with the maximum of the diagonal ensures that the regularization is proportional to the values in the Jacobian product. In equation 2, \( \lambda I \) becomes

\[
\lambda I = \begin{bmatrix}
\lambda_1^{\text{amplitude}} & 0 & 0 & 0 & \cdots & 0 \\
0 & \lambda_1^{\text{phase}} & 0 & 0 & \cdots & 0 \\
0 & 0 & \lambda_2^{\text{amplitude}} & 0 & \cdots & 0 \\
0 & 0 & 0 & \lambda_2^{\text{phase}} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \lambda_M^{\text{phase}}
\end{bmatrix}
\]

The two techniques investigated here both employ this formulation, however, one implementation is a modified Levenburg-Marquardt (LM) formulation which reduces \( \beta \) at each iteration by a set amount while the other holds \( \beta \) constant throughout the iterative process, essentially repeating the single-step Tikhonov method until a stopping criteria is reached:

Technique 1 (Modified Levenberg-Marquardt): \( \beta_i = \beta_{i-1} / 10^{0.25} \)

Technique 2 (Single-step Tikhonov): \( \beta_i = \beta_0 \)

where \( i \) is iteration number. It should be noted that the LM approach assumes the inversion matrix is positive-definite and the data contains no noise, which is not the case here. However, a modified version of the LM formulation which does not require the regularization to go to zero as the iterative process proceeds has been used frequently for DOT reconstructions. The reducing factor used was determined empirically from series of phantom and patient studies. This procedure forces the algorithm to trust the data more as the algorithm iterates.

The reconstruction algorithm used a 1785 node mesh to calculate the forward fields at each iteration and interpolated these fields to a 30 by 30 pixel basis for parameter updates. This approach reduces the number of unknowns making the problem less ill-posed. The algorithm was stopped when the change between successive iterations was less than 2%.

Figure 1 depicts the four circular test fields used in this study. This series of 86 mm diameter geometries represents increasing levels of complexity to challenge the imaging algorithm. Each test field was surrounded by 16 source/detector fibers resulting in a total of 240 data points of phase and amplitude for an imaging session. Simulated data was generated by solving the forward problem and extracting boundary data. Normally distributed Gaussian noise of 1%, 2%, 5% and 10% amplitude and 1, 2, 5, and 10 degrees phase was added to each data set. This resulted in a total of four data sets for each of the four test domains.
Images of $\mu_a$ and $\mu_s'$ were reconstructed simultaneously with phase and amplitude data using both regularization techniques described in the Theory section under a variety of conditions summarized below.

1. Initial estimates of optical parameters ($\mu_a = [0.01, 0.02, 0.03]$, $\mu_s' = [1.0, 1.6]$).
2. Initial estimate of the regularization parameter varied $[0.1, 1, 5, 10, 100]$.

Each data set was reconstructed with every combination of the above conditions representing 30 reconstructions for each data set and therefore 120 reconstructions for each test geometry. Table 1 lists the reconstruction conditions for a single test geometry.

Table 1. Summary of simulation conditions for a single test field. Four data sets with varying levels of noise were used for each test geometry.

<table>
<thead>
<tr>
<th>Test field</th>
<th>Data</th>
<th>1% amplitude, 1 degree phase noise</th>
<th>2% amplitude, 2 degrees phase noise</th>
<th>5% amplitude, 5 degrees phase noise</th>
<th>10% amplitude, 10 degrees phase noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test field 1</td>
<td>Initial estimate for $\mu_a$</td>
<td>0.01, 0.02, 0.03</td>
<td>0.01, 0.02, 0.03</td>
<td>0.01, 0.02, 0.03</td>
<td>0.01, 0.02, 0.03</td>
</tr>
<tr>
<td>Test field 2</td>
<td>Initial estimate for $\mu_s'$</td>
<td>1.0, 1.6</td>
<td>1.0, 1.6</td>
<td>1.0, 1.6</td>
<td>1.0, 1.6</td>
</tr>
<tr>
<td>Test field 3</td>
<td>Initial value of $\beta$</td>
<td>0.1, 1, 5, 10, 100</td>
<td>0.1, 1, 5, 10, 100</td>
<td>0.1, 1, 5, 10, 100</td>
<td>0.1, 1, 5, 10, 100</td>
</tr>
<tr>
<td>Test field 4</td>
<td>Layered domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. RESULTS
A subset of the images is presented in figures 2-4 for both regularization techniques under investigation. Figure 2 shows examples of images for each test domain reconstructed from data containing 1% log amplitude and 1° phase noise. In general, the technique which holds $\beta$ constant throughout the iterative reconstruction produces a smoother image with fewer image artifacts. This was a consistent result for many of the images generated in this study.
A series of images demonstrating the effects of varying the amount of data noise as well as changing the initial value of β on the three anomaly test domain are provided in figure 3. Again, the constant β images are smoother with fewer artifacts. It is also clear that data noise and initial values of β have a dramatic effect on image quality. When the data noise was increased from 1% log amplitude and 1° phase to 5% log amplitude and 5° phase, image quality for both regularization techniques decreased dramatically. In this higher noise case, image artifacts dominate images reconstructed with the decreasing β algorithm and it is impossible to discern the true anomalies. Alternatively, it is fairly easy to discern anomalies in images generated using the constant β implementation, though parameter cross-talk yields a false positive in the scattering coefficient. Using a higher β in for the noisier data produced a more qualitatively accurate image. This is expected since a higher initial β allows for a larger data-model misfit and therefore is more appropriate for reconstructing noisy data.

Figure 4 shows images for the same test field (three anomaly) reconstructed using homogeneous initial estimates of μₐ = 0.01, 0.02, and 0.03. Varying the initial parameter estimates has a dramatic impact on the algorithm which decreases β at each iteration. An initial estimate of μₐ that is three times the actual value of the background results in a catastrophic failure of this technique. The algorithm holding β constant is more robust to the initial parameter estimate and the images generated using this technique are more consistent, though cross-talk is still a problem in the scattering images.
Fig. 2. Reconstructed images from noisy (1% amplitude and 1° phase) data for all test domains. The initial value of $\beta$ was 10 for all images shown above and the initial homogeneous parameter estimates to start the iterative algorithm were $\mu_a = 0.01$ (1/mm) and $\mu_{s'} = 1.0$ (1/mm) for all images except for the layered domain, which was $\mu_a = 0.02$ (1/mm) and $\mu_{s'} = 1.6$ (1/mm).
Fig. 3. Reconstructed images of the three anomaly test field with varying data noise and initial $\beta$ values. Initial homogeneous parameter estimates to start the iterative algorithm were $\mu_a = 0.01$ (1/mm) and $\mu_s' = 1.0$ (1/mm) for all images. Qualitatively, the constant $\beta$ algorithm is more effective at handling noise and suppressing image artifacts found in images produced by the technique which decreases regularization.
<table>
<thead>
<tr>
<th>Initial Estimate: $\mu_a = 0.01$ $\mu_s' = 1.0$ (mm$^{-1}$)</th>
<th>$\beta$ decreased at each iteration</th>
<th>$\beta$ constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_a$</td>
<td>$\mu_s'$</td>
<td>$\mu_a$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_a$</td>
<td>$\mu_s'$</td>
<td>$\mu_a$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_a$</td>
<td>$\mu_s'$</td>
<td>$\mu_a$</td>
</tr>
</tbody>
</table>

Fig. 4. Images for the three anomaly test domain with different initial estimates of $\mu_a$. These images contained 1% log amplitude and 1° phase noise and $\beta = 10$ initially. The constant $\beta$ algorithm is shown to be less sensitive to initial parameter estimates in this case.

The root-mean-squared error (RMSE) of the image area was used as a simple metric to compile results from the 960 images generated during this study. Figure 5 is an example of the raw RMSE data for images of $\mu_a$ in the three anomaly test field. The Tikhonov (constant $\beta$) approach produced lower RMSE values for almost all reconstruction conditions and data noise levels. Similar results were found for both $\mu_a$ and $\mu_s'$ parameters for the edge, center, and three anomaly test fields. However, results for the complex layered domain were less convincing.
Fig. 5. RMSE data for $\mu_i$ images of the three anomaly domain for a variety of reconstruction conditions. Data noise levels are indicated by letters: A – 1% log amplitude and 1° phase, B – 2% log amplitude and 2° phase, C – 5% log amplitude and 5° phase, and D – 10% log amplitude and 10° phase.

Figures 6 and 7 report the average RMSE values for the three anomaly test domain images as a function of noise in the boundary data and initial estimate of $\mu_i$ respectively. RMSE values above 0.015 were considered failed reconstruction attempts and excluded from the plot. It is apparent that on average, holding $\beta$ constant results in lower RMSE for both absorption and scattering parameters. Similar graphs are reported for the layered domain test field in figures 8 and 9 and are less compelling. In this case, the constant $\beta$ routine was more robust to initial parameter estimates as well as high noise in the data, but resulted in higher RMSE values for data with 1% in log amplitude, 1° in phase and 2% in log amplitude and 2° in phase noise.
Fig. 6. Averaged RMSE data for images of the three anomaly domain covering data noise levels indicated by letters: A – 1% log amplitude and 1° phase, B - 2% log amplitude and 2° phase, C - 5% log amplitude and 5° phase, and D - 10% log amplitude and 10° phase.

Fig. 7. Averaged RMSE data for images of the three anomaly test domain covering initial $\mu_a$ estimates indicated by letters: A – $\mu_a = 0.01$ (1/mm), B - $\mu_a = 0.02$ (1/mm), C - $\mu_a = 0.03$ (1/mm).
Fig. 8. Averaged RMSE data for images of the layered domain covering data noise levels indicated by letters: A – 1% log amplitude and 1° phase, B - 2% log amplitude and 2° phase, C - 5% log amplitude and 5° phase, and D - 10% log amplitude and 10° phase.

Fig. 9. Averaged RMSE data for images of the layered test domain covering initial $\mu_a$ estimates indicated by letters: A – $\mu_a = 0.01$ (1/mm), B - $\mu_a = 0.02$ (1/mm), C - $\mu_a = 0.03$ (1/mm).
4. DISCUSSION AND CONCLUSION

Two regularization techniques for simultaneous recovery of absorption and scattering parameters in tissue were compared in terms of image quality and RMSE of the image area. In many of the images examined visually, the repeated single-step Tikhonov iterative technique suppressed image artifacts and was more robust to data noise and initial homogeneous parameter estimates. Examining the RMSE of the test fields confirmed that this technique outperformed a method which reduces the regularization at each iteration in almost all test cases. However, both imaging algorithms had difficulty with the layered domain, which was also the most realistic test field, and RMSE errors were relatively high for these images. This is expected given the large heterogeneity of optical properties and techniques that incorporate spatial or spectral prior information in the imaging algorithm have been shown to dramatically improve imaging in complex tissue volumes. [5-8]

Acknowledgements
This work was funded by the National Institutes of Health grants RO1 CA109558, RO1 CA69544, U54 CA105480, as well as Department of Defense Breast Cancer pre-doctoral fellowship BC051058.

5. REFERENCES