Cramér-Rao estimation of error limits for diffuse optical tomography with spatial prior information

Phaneendra Yalavarthy,[a] Rajesh Langoju,[b] Brian Pogue,[a] Hamid Dehghani,[c] Abhijit Patil,[b] and Keith Paulsen[a]

[a]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755 USA;
[b]Applied Computing and Mechanics Laboratory, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland;
[c]School of Physics, University of Exeter, Stocker Road, Exeter, EX4 4QL, UK

ABSTRACT

Cramér-Rao Bounds (CRB) for the expected variance in the parameter space were examined for Diffuse Optical Tomography (DOT), to define the lower bound (CRLB) of an ideal system. The results show that the relative standard deviation in the optical parameter estimate follows an inverse quadratic function with respect to signal to noise ratio (SNR). The CRLB was estimated for three methods of including spatial constraints. The CRLB estimate decreased by a factor of 10 when parameter reduction using spatial constraints (hard-priors) was enforced whereas, inclusion of spatial-priors in the regularization matrix (soft-priors) decreased the CRLB estimate only by a factor of 4. The maximum reduction in variance from the use of spatial-priors, occurred in the background of the imaging domain as opposed to localized target regions. As expected, the variance in the recovered properties increased as the number of parameters to be estimated increased. Additionally, increasing SNR beyond a certain point did not influence the outcome of the optical property estimation when prior information was available.

Keywords: Diffuse Optical Tomography, Structural Priors, Cramér-Rao Lower Bounds

1. INTRODUCTION

Diffuse optical tomography (DOT) involves estimating the interior optical properties of tissue from measurements recorded non-invasively from the surface. The main DOT applications include breast and brain imaging.[1,2] Typically, light is delivered and collected through a set of optical fiber bundles which contact the tissue. Wavelengths in the near infrared (NIR) range (650-1000nm) are used to take advantage of the lower absorption of water and hemoglobin (Hb), and therefore the higher penetration of light through tissue. The spatial resolution of NIR imaging is typically low, near 3-5 mm, due to the dominance of scattering. Yet, DOT has significant potential to become a clinical imaging modality because of the physiological information it provides through non-ionizing radiation. The procedure is also non-invasive.[1]

DOT image formation involves two important steps.[3] The first is simulation of the way light propagates in tissue, or the so called *forward problem*, which gives the distribution of the fluence of the light due to a source under a known or estimated optical property map. Sampling this distribution at the boundaries, where the fiber probes reside, provides the modeled data ($\mathbf{Z}^c$). The forward problem is typically solved by finite-element modeling (FEM) of the diffusion equation (DE), which is an approximation to the radiative transport equation (RTE).[3] The second and perhaps more critical step is the parameter estimation process, or the so-called *inverse problem* which involves minimization of the data-model misfit in a least-squares sense (Eq. 1). Here, the *data* refers to the experimental measurements ($\mathbf{Z}$) and the *model* refers to $\mathbf{Z}^c$. The objective function Ω has form:

$$\Omega = \min_{\Theta} \{\|\mathbf{Z} - \Gamma(\Theta)\|^2 + \lambda \|L(\Theta - \Theta_0)\|^2\} \quad (1)$$

Further author information: (Send correspondence to P.K.Y.)
P.K.Y: E-mail: Phani@dartmouth.edu, Telephone: (603)-646-2685

where the parameter vector is represented by Θ and Θ_0 is its prior estimate. Γ is the forward model ($\Gamma(\Theta) = \mathbf{Z}^c$). λ and L are the regularization parameter and regularization matrix, respectively.[4] In DOT, typically this minimization process is iterative and involves re-computing the first-order derivatives of the forward model (Γ) with respect to the parameters. The DOT inverse problem is inherently ill-posed,[3] meaning that small changes in data can give rise to large changes in optical properties. The ill-posedness arises from the fact that there is a large range in sensitivity (of boundary data to the parameter distribution) within the imaging domain, which makes it prone to noise.[3–5] Moreover, the inverse problem is often under-determined, because the number of optical property parameters required to complete the forward modeling is often very high, whereas the amount of data that can be recorded in practice is far less.[3] To reduce the under-determinedness, dual mesh strategies have been developed[7] where the inverse problem is executed on fewer degrees of freedom than the forward problem. However, to ensure that the imaging problem is solvable even in the face of being ill-posed and under-determined, regularization is added. Typically, L is an identity matrix (I) when structural-priors are not available. Other forms of L are often used to incorporate structural-priors.[6] Some studies have indicated that the error in the reconstructed optical property distribution is decreased by factor of 2 with structural-priors.[8,9]

This work is focused on estimating a theoretical limit on the optical parameter precision in DOT using the Cramér-Rao Lower Bound (CRLB)[10,11] which is based on known results from information theory. CRLB provides theoretical performance limits associated with the problem complexity, and the CRLB approach to estimate these limits is generically applicable to many estimation procedures. Specifically, any algorithm that produces an unbiased parameter estimate cannot attain a lower error bound than that defined by the CRLB under certain regularity conditions.[11] In iterative methods like DOT, it is assumed that bias in the parameter estimation asymptotically goes to zero.[11] This is true for DOT, as shown in Ref.,[12] the bias in the problem dominates only in the first few iterations. For regularized problems, such bounds have been obtained in the past.[13–15] However, these bounds are useful in evaluating the effect of SNR on the performance of the estimation technique, yet, they have not been determined in the past for DOT. In this work, the effect of the regularization parameter (λ) and regularization matrix (L) on the estimation of optical parameters is examined. These studies provide an understanding of the precision limits on optical parameter estimation in DOT. For completeness of results, the estimation techniques considered here are compared as a function of SNR. Note that the discussion in this paper is limited for SNR values above 10 dB, below which the DOT parameter estimation problem becomes severely biased. This is especially true in the cases of considering magnitude only data.[15]

2. METHODS

2.1. Parameter estimation in DOT

Because scattering dominates absorption at NIR wavelengths, light transport is modeled using the DE in thick tissues. In the frequency domain, the DE is given by[16]

$$-\nabla.k(\mathbf{r})\nabla\Phi(\mathbf{r},\omega) + (\mu_a(\mathbf{r}) + i\omega/c)\Phi(\mathbf{r},\omega) = q_o(\mathbf{r},\omega) \tag{2}$$

where $\Phi(\mathbf{r},\omega)$ is the photon density at position $\mathbf{r}$. The frequency of light modulation is represented by ω, c gives the speed of light in tissue and $q_o(\mathbf{r},\omega)$ represents the isotropic light source at $\mathbf{r}$, $\mu_a(\mathbf{r})$ is the optical absorption coefficient and $k(\mathbf{r})$ is the optical diffusion coefficient, which is defined as

$$k(\mathbf{r}) = \frac{1}{3[\mu_a(\mathbf{r}) + \mu_s'(\mathbf{r})]} \tag{3}$$

where $\mu_s'(\mathbf{r})$ represents the reduced scattering coefficient, which is defined as $\mu_s' = \mu_s(1-g)$. Here, μ_s is the scattering coefficient and g is the anisotropy factor. A Robin (Type-III) boundary condition is generally applied to model the refractive-index mismatch at the boundary.

A finite element implementation of Eq. 2 will generate simulated data, $\Gamma(\Theta)$, where $\Theta \equiv [k;\mu_a]$. Note that the data in the frequency domain case is the logarithm of amplitude ($\ln A$) and phase (ϕ); $\mathbf{Z}$ *or* $\mathbf{Z}^c \equiv [\ln A;\phi]$.

Minimizing the objective function (Ω given by Eq. 1) by setting first-order derivatives with respect to Θ to zero leads to update equation

$$(J^T J + \lambda L^T L)\triangle\Theta = J^T(\mathbf{Z} - \Gamma(\Theta)) - \lambda L^T L(\Theta - \Theta_0) \tag{4}$$

where J is the Jacobian matrix used in the inverse solution. This procedure is repeated until $\Gamma(\Theta)$ iteratively matches $\mathbf{Z}$ within a preset tolerance $\epsilon(\approx$ data noise). Note that for the discussion here, only amplitude data ($\mathbf{Z} = log(A)$) is considered and the parameter that is being estimated is the absorption coefficient ($\Theta = \mu_a$) in order to simplify the analysis.

2.1.1. Forms of Regularization Matrix (L)

The regularization matrix (L) in Eq. 1 and 4 can take on many forms depending upon the application. Two formulations were examined in this work. When $L = I$, spatial-priors are not available and generally a Levenberg-Marquardt (LM)[17]-[18] type of regularization procedure has been applied leading to[16]

$$(J^T J + 2\lambda L^T L)\triangle\Theta = J^T(\mathbf{Z} - \Gamma(\Theta)) \tag{5}$$

where $\triangle\Theta = \Theta - \Theta_0$. Note that, in the literature, $\lambda^* \equiv 2\lambda$,[3,16] for Levenberg-Marquardt (LM) minimization, which does not involve the parameter field in the objective function (Eq. 1).[17,18] In the LM approach, λ begins as the ratio of variances and is reduced at each iteration by some factor (here, it is $\sqrt{10}$ multiplied by the maximum of the diagonal values of $\mathbf{J^T J}$[19]). In general, the prior values, (Θ_0), are obtained from a pre-reconstruction step where the data is calibrated by a homogeneous fitting procedure.[20,21]

When spatial-priors are encoded, Brooksby et al.[8] have developed a Laplacian-type of regularization matrix, often referred to as *soft-priors*, depending on the region information which is available (for breast tissue, fibroglandular and tumor tissue are typically involved)

$$L(i,j) = \begin{cases} 0 & \text{if i and j do not belong to same region} \\ -1/N & \text{if i and j belong to same region} \\ 1 & \text{if i = j} \end{cases} \tag{6}$$

where N is the number of nodes in the region. A parameter reduction technique, often referred to as *hard-priors*, is also possible when spatial-priors are available and the different regions of tissue are expected to be homogeneous. In this case, the parameter space is reduced to the number of regions indicated by the spatial-priors, making the inverse problem well-determined and making regularization unnecessary. Typically, it is also assumed that the spatial-priors are exact, i.e. uncorrupted by spatial uncertainty.

2.2. Cramér-Rao Lower Bound (CRLB) on the standard deviation of μ_a

A detailed derivation of CRLB is given in the appendix. The CRLB defines the minimum variance of the parameters in a given estimation problem for a specified SNR in the data. Here, standard deviation of μ_a, square root of this CRLB variance, is considered to be the theoretical limit. A test object (as shown in Figure 1), which mimics the typical optical properties of breast at 785 nm,[22] was used for the studies reported in this paper. The imaging domain was centered at (0,0) and had a diameter of 86 mm with an internal target (located at (30,0)) having a diameter of 15 mm (Figure 1(a)). The optical properties were: $\mu_a = 0.01$ mm^{-1} and $\mu_s' = 1$ mm^{-1} for the background and $\mu_a = 0.02$ mm^{-1} and $\mu_s' = 1$ mm^{-1} for the target. The frequency of modulation was 100 MHz, the refractive index of the tissue was 1.33. Data was collected through 16 equi-distant fibers that are placed on the boundary of the domain, where one fiber acts as a source while the other 15 fibers are used as detectors, to give 240 total measurement points. The source is a Gaussian distribution with a Full Width Half Maximum (FWHM) of 3 mm which was placed one reduced scattering distance ($= 1/\mu_s'$) within the external boundary. The source and detector positions are illustrated in Figure 1(b). The FEM mesh (plotted in Figure 1(b)) used for the problem had 1785 nodes (corresponding to 3418 linear triangular elements) making the parameter (μ_a) space of dimension 1785x1. In the estimation process, a reconstruction basis of 30x30 pixels was used.[23]

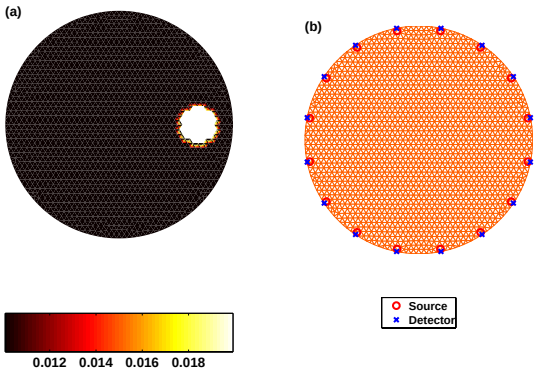

Figure 1. Original chosen domain diameter is 86 mm with a target diameter of 15 mm (a). The true μ_a image is shown with target contrast: 2:1. The source and detector locations along with the mesh are shown in (b).

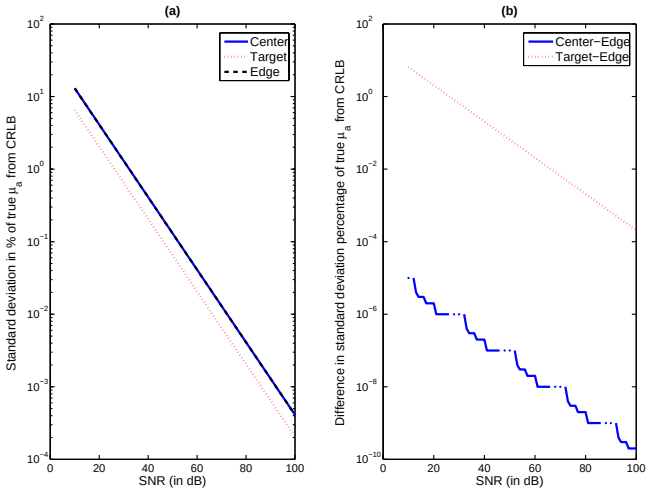

Figure 2. (a) Plot of standard deviation in μ_a for different parts of the domain using the Cramer-Rao Lower Bound (CRLB) on a 1785 node mesh (shown in Figure 1(b)) for λ=10 (b).

At first, Gaussian noise with a given SNR was added to the synthetic data generated from the model to form an experimental data set. For SNR ranging from 10 to 100 dB, standard deviation and variance of μ_a are calculated using CRLB (see Appendix). The calculated standard deviation (sd) is plotted as a percentage of true μ_a for a given SNR:

$$\text{standard deviation in } \% = \left(\frac{sd}{true\ \mu_a}\right) * 100\% \tag{7}$$

Note that an average of the standard deviation % over 40 nodes in the given part of the domain was calculated for plotting purposes. This averaging was performed to overcome the spatial irregularities (spacing between the nodes) due to finite element meshing.

To study the effect of the regularization parameter (λ) on μ_a estimation for SNR values of 20, 40, 60 and 80 dB, CRLB calculations were completed when λ was varied from 0.01 to 100. The standard deviation was estimated for both the no-prior and soft-prior regularizations. The hard-prior case was also explored with 2 and 4 parameter estimates and the standard deviation was plotted as a function of the SNR.

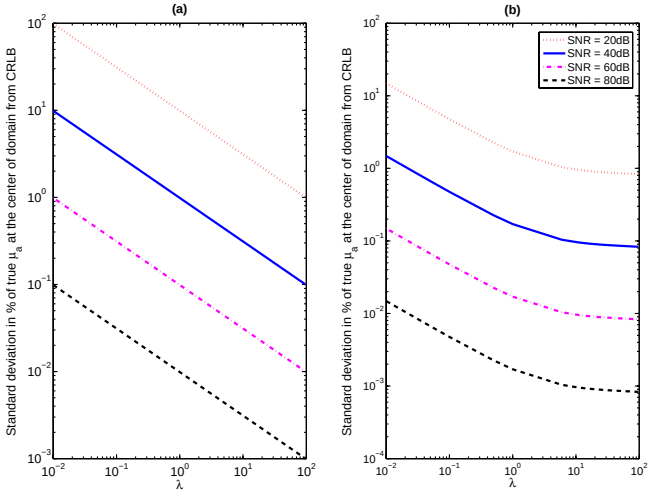

Figure 3. The effect of regularization parameter λ on the CRLB for different SNR values is shown for (a) no-priors ($L = I$) and (b) soft-priors (L in Eq. 6). SNR values are given in the legend. Standard deviations are from the center of the domain.

To compare the performance of the estimation processes presented here (no-priors, soft-priors and hard-priors), estimation of optical properties using an ensemble of noise added data sets and the standard deviation among these estimates was calculated. The iterative estimation process was repeated until the synthetic data was matched with modeled data within the noise level. The data was generated on the same test object described in Figure 1 and Gaussian noise was added repeatedly, for a given SNR, to form 60 synthetic data sets. For the cases involving no-priors and soft-priors, $\lambda = 10$ was used. Note that the stopping criterion in these procedures is changed adoptively with the noise level.

3. RESULTS

Figure 2(a) shows the standard deviation in percentage of the true μ_a as a function of SNR for the test object described in Sec. 2.1. The standard deviation predicted by the CRLB is inversely proportional to the square of the SNR as given by Eq. 27. The difference in the standard deviation % between different parts of the domain is plotted in Figure 2(b).

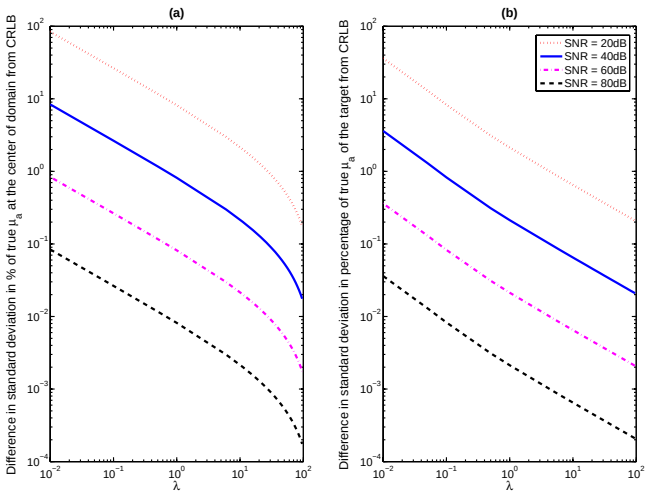

Figure 4. Difference in standard deviation (Figure 3) of μ_a between the no-prior and soft-prior regularizations for (a) the center of the domain and (b) the target

.

For SNR values of 20, 40, 60 and 80 dB, the standard deviation % is plotted as a function of λ for the no-prior and soft-prior cases in Figure 3(a) and (b), respectively. The CRLB estimates depend inversely on λ, such that larger regularization leads to lower variances as expected. Interestingly, in the case of soft-priors, CRLB estimates become largely independent of the regularization parameter above a value of 6. The difference in standard deviation between the no-prior and soft-prior regularizations is shown in Figure 4 for 4 different SNR values.

The hard-prior results for 2 and 4 parameter estimates (3 targets and background) are plotted in Figure 5(a) and (b), respectively. The positions (center) of the targets are given in the legend of the Figure 5. Results from the different estimation processes on the ensemble of 60 noisy data sets (as described in Sec. 2.2) are plotted in Figure 6 as a function of SNR. The CRLB of the no-priors case ($\lambda = 10$) is included for comparison. Using the ensemble of estimated parameter distributions, the standard deviation over this ensemble is plotted in Figure 6. The legend of the Figure 6 gives the estimation technique.

4. DISCUSSION

For a given regularization parameter, the standard deviation (equivalently precision) in the estimated μ_a follows an inverse quadratic relationship with SNR (Figure 2(a)). This functional form implies that the limit on precision of the μ_a estimation increases as SNR increases. The current experimental-system at Dartmouth has an SNR $\sim$ 20 dB.[24] Interestingly, simulations and experimental studies have suggested that the lower limit on detectable contrast is around 16%.[25] For SNR = 20 dB, the CRLB on standard deviation percentage is approximately 1-3% depending upon location within the imaging domain (Figure 2(a)) which indicates that it is not possible to create an imaging system with 20 dB SNR which would be able to detect a target within 1-3% contrast. Conversely, if the noise level could be lowered, the minimum detectable contrast could be lowered according to the line in Figure 2(a). Figure 2(b) also suggests that the target would always have a lower standard deviation relative to the rest of the domain, which is consistent with the results presented by Pogue et al.[12] However, it is important to note that the standard deviations reported in previous work are significantly higher than the ones presented here because the Cramér-Rao formulation provides a theoretical lower limit on the standard deviation that can be achieved.

Diffuse Optical Tomography is inherently ill-posed and the effect of regularization on the precision of μ_a estimation has been investigated extensively,[26,27] but not in the context of the CRLB. Figure 3 clearly shows

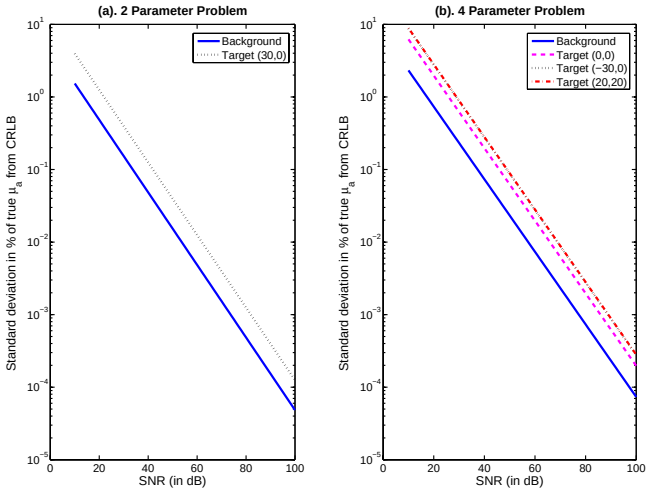

Figure 5. Standard deviation in μ_a using CRLB as a function of SNR (in dB) for hard-priors: (a) 2 parameter problem and (b) 4 parameter problem. Target location as given in figure legend.

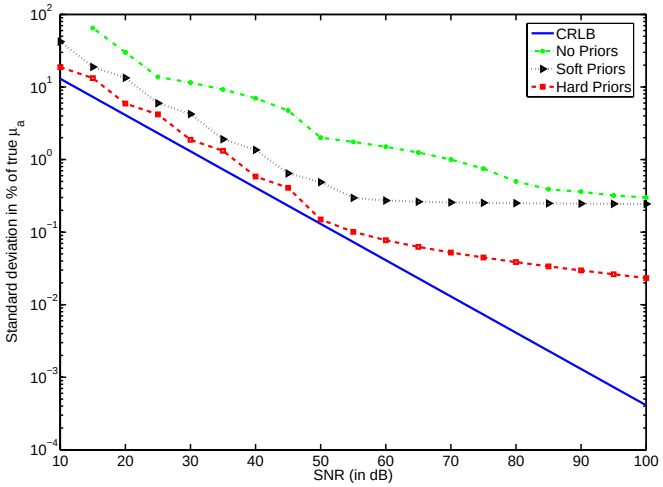

Figure 6. Estimation procedures performance as a function of SNR in comparison with CRLB for no-priors. $\lambda = 10$ was used in the no-priors and soft-priors estimation procedures. Each point in this graph represents the mean of reconstruction results at the center of the domain from an ensemble (N = 60) of noise added data sets for a given SNR.

that as the regularization parameter λ is increased, the precision in μ_a also increases. This in turn, indicates that the term $\|L(\Theta - \Theta_0)\|^2$ in Eq. 1 is given more importance than the data-model misfit, which reduces the variance in the parameter ($\Theta = \mu_a$) estimation. Another important observation is that the precision in μ_a is much higher in the case of soft-priors because the regularization matrix L gives more weight to the parameter field and thus reduces the standard deviation in the estimation process. Increases in λ in the case of soft-priors do not reduce the standard deviation beyond $\lambda = 6$ (Figure 3(b)), because the estimation process has reached the minimum in the parameter error. The same behavior is observed with the L-curve.[28] In the case of no-priors the same minimum was achieved when the maximum value of λ was extended from 100 to 10,000. It is also important to note that the standard deviation values were reduced by a factor of approximately 4 through the inclusion of spatial-priors. The change is more evident in the center of the domain than in the target (Figure 4), which can be explained by the fact that the target always has less standard deviation compared to the rest of the domain. The improvement in the standard deviation limit with soft-priors relative to no-priors, is greater (factor of 2) for the background than the target.

Interestingly, the well-determined problem (hard-priors case) follows the same exponential decay function (Figure 5) with the same decay rate in the precision of μ_a, as was observed in the no-priors case (Figure 2(a)). However, the precision was improved by an order of magnitude and the background had a lower standard deviation compared to the target, which is contrary to the results obtained with no-priors. Further, as the number of parameters are doubled the precision decreases by the same factor (comparing Figures 5(a) and (b)). The lines in Figure 5(b) indicate that the standard deviation is lower for a target located at the center of the domain compared to the edge, because DOT has increased sensitivity near the boundary source and detector locations.

Finally, the CRLB analysis indicates that hard-priors have better performance than the other techniques (Figure 6). The center of the domain was chosen to be examined in detail because it has the highest errors due to low sensitivity. Even though the errors in magnitude were different, all points in the domain followed the same trend presented in Figure 6 where it is seen that after an SNR of 55 dB, the standard deviation does not decrease significantly ($\leq$ 5 %) for the case of soft-priors. At SNR = 20 dB, it can be clearly seen that the standard deviation is a factor of 7.3, 3.3 and 1.44 of the CRLB for no-priors, soft-priors and hard-priors, respectively.

Given prior information on the parameters that are being estimated, it is possible to use the CRB estimate to determine the best possible performance of an imaging system. Rewriting Eq. 1 by assuming that the variance in each data point/parameter is the same and that they are independent of each other, leads to

$$\Omega = \min_{\Theta} \left\{ \frac{\|\mathbf{Z} - \Gamma(\Theta)\|^2}{{\sigma_Z}^2} + \frac{\|L(\Theta - \Theta_0)\|^2}{{\sigma_\theta}^2} \right\} \tag{8}$$

where ${\sigma_Z}^2$ and ${\sigma_\theta}^2$ are the variances of the data and the parameter space, respectively. Comparing this equation with Eq. 1 leads to the observation that $\lambda = (\sigma_Z/\sigma_\theta)^2$ which can be used to estimate λ through CRLB, to calculate ${\sigma_\theta}^2$ (characterizing the data collection system provides ${\sigma_Z}^2$).

Note that in the analysis presented here, the estimation process was assumed to be unbiased, but the DOT problem is known to be biased.[12,29] Estimation of image bias requires the true parameter distribution to be known, which is not possible in most practical imaging situations. The effect of bias error in DOT and upon the CRB estimate is left for future investigation. However, studies of the tradeoff between bias and variance in the recovery of DOT images has shown that while the bias error dominates at early iterations, an optimized algorithm can converge to a solution which is no longer biased, if the data variance is low.[12,29] More over, any practical estimator will have certain bias, as stated before, in all CRB calculations it is assumed that asymptotically bias goes to zero.[11] Although the calculations were performed for a single test problem, the results should be generalizable, at least in terms of the trends observed. It is also important to note that while only light amplitude was used as data and μ_a was the parameter recovered, including the phase (ϕ) and diffusion coefficient (k) produced similar trends. Nonetheless, future studies are needed, especially for complex inverse problems, such as those that use multiple wavelengths or estimate multiple chromophore property images.[30,31]

5. CONCLUSIONS

The Cramér-Rao Lower Bound (CRLB) for diffuse optical tomography was formulated and a closed-form expression was derived to estimate the lower limit in the precision of the estimation of optical properties. The effect of regularization upon the CRB was studied and the results are shown to be in reasonable agreement with those observed in previous publications. Specifically, hard-priors are closer to the CRLB, indicating this is the best estimator. The variance bounds given by the CRLB, were observed to be lower limits that can be used to choose an appropriate regularization parameter for a problem. The choices of regularization procedure and parameter values are clearly the most important factors that influence the outcome of an image-reconstruction procedure, but this analysis approach provides a quantitative interpretation of how regularization affects the variance in the property images.

6. APPENDIX

6.1. Cramér-Rao Lower Bound (CRLB)

This section discusses the Cramér-Rao Lower Bound (CRLB)[10,11] inequality, which places a lower bound on the error variance, a bound that does not depend on a particular estimator. Initially, the CRLB is discussed for the single parameter case, then it is extended to a parameter vector, and finally it is derived for data noise that is Gaussian distributed.

6.2. CRLB for single parameter

Let $\mathbf{Z}$ denote a set of data [i.e.,$\mathbf{Z} = col(z_0, z_1, \ldots, z_{N-1})$] with N discrete observations, which is characterized by the probability density function $p(\mathbf{Z};\theta) \equiv p(\mathbf{Z})$, where θ is the parameter. If $\hat{\theta}$ is the unbiased estimate of θ then the variance of the unbiased estimator $E\{\tilde{\theta}^2\}$ is bounded as $E\{\tilde{\theta}^2\} \geq \frac{1}{\mathbf{E}\left\{\left[\frac{\partial}{\partial\theta}\ln p(\mathbf{Z})\right]^2\right\}}$, where $\tilde{\theta} = \hat{\theta} - \theta$. The denominator is called Fisher's information, which contains the relative rate (derivative) at which the probability density function changes with respect to the data.

Proof: If $\hat{\theta}$ is an unbiased estimate of θ then the variance of $\hat{\theta}$ is given by

$$E\{\tilde{\theta}\} = \int_{\infty}^{\infty} (\hat{\theta} - \theta) p(\mathbf{Z}) d\mathbf{Z} = 0 \tag{9}$$

Differentiating Eq. (9) with respect to θ gives

$$\int_{-\infty}^{\infty} (\hat{\theta} - \theta) \frac{\partial p(\mathbf{Z})}{\partial\theta} d\mathbf{Z} - \int_{\infty}^{\infty} p(\mathbf{Z}) d\mathbf{Z} = 0 \tag{10}$$

Simplification of Eq. (10) leads to

$$\int_{-\infty}^{\infty} (\hat{\theta} - \theta) \frac{\partial p(\mathbf{Z})}{\partial\theta} d\mathbf{Z} = 1 \tag{11}$$

Note that

$$\frac{\partial}{\partial\theta}\ln p(\mathbf{Z}) = \frac{\partial p(\mathbf{Z})}{\partial\theta}\frac{1}{p(\mathbf{Z})} \tag{12}$$

Substituting Eq. (12) into Eq. (11) and rearranging terms leads to

$$\int_{-\infty}^{\infty} \left[(\hat{\theta} - \theta)\sqrt{p(\mathbf{Z})}\right] \left[\sqrt{p(\mathbf{Z})}\frac{\partial}{\partial\theta}\ln p(\mathbf{Z})\right] d\mathbf{Z} = 1 \tag{13}$$

Recall the Schwartz inequality

$$\left[\int_{-\infty}^{\infty} a(\mathbf{Z})b(\mathbf{Z})d\mathbf{Z}\right]^2 \leq \left[\int_{-\infty}^{\infty} a^2(\mathbf{Z})d\mathbf{Z}\right]\left[\int_{-\infty}^{\infty} b^2(\mathbf{Z})d\mathbf{Z}\right] \tag{14}$$

where equality is achieved when $b(\mathbf{Z}) = ca(\mathbf{z})$ in which c is a constant that is not dependent upon $\mathbf{Z}$. Squaring Eq. (13) and applying Eq. (14) to the left hand side gives

$$\begin{gathered}\left[\int_{-\infty}^{\infty}(\hat{\theta}-\theta)^2 p(\mathbf{Z})d\mathbf{Z}\right]\left[\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\ln p(\mathbf{Z})\right]^2 p(\mathbf{Z})d\mathbf{Z}\right] \geq 1 \\ or \\ \mathbf{E}\{\tilde{\theta}^2\}\mathbf{E}\left\{\left[\frac{\partial}{\partial\theta}\ln p(\mathbf{Z})\right]^2\right\} \geq 1\end{gathered} \tag{15}$$

which leads to CRLB

$$\mathbf{E}\{\tilde{\theta}^2\} \geq \frac{1}{\mathbf{E}\left\{\left[\frac{\partial}{\partial\theta}\ln p(\mathbf{Z})\right]^2\right\}} \tag{16}$$

6.3. CRLB for parameter vector

Let $\mathbf{Z}$ denote a set of data [i.e.,$\mathbf{Z} = col(z_1, z_2, \ldots, z_k)$ which is characterized by the probability density function $p(\mathbf{Z};\Theta) \equiv p(\mathbf{Z})$. If $\hat{\Theta}$ is an unbiased estimate of deterministic Θ, then the covariance matrix of the unbiased estimator $\mathbf{E}\{\tilde{\Theta}\tilde{\Theta}^T\}$ is bounded as $\mathbf{E}\{\tilde{\Theta}\tilde{\Theta}^T\} \geq \mathbf{H}^{-1}$ where T is the transpose operator, and $\mathbf{H}$ is the Fisher information matrix given by, $\mathbf{H} = E\{\left[\frac{\partial}{\partial\Theta}\ln p(\mathbf{Z})\right]\left[\frac{\partial}{\partial\Theta}\ln p(\mathbf{Z})\right]^T\}$

Proof: Following the same steps that were taken to reach Eq. (13), we get

$$\int_{-\infty}^{\infty}(\hat{\Theta}-\Theta)\frac{\partial \ln p(\mathbf{Z})}{\partial\Theta}p(\mathbf{Z})d\mathbf{Z} = \mathbf{I} \tag{17}$$

At this point the proof departs from the single parameter case. Multiply Eq. (17) by $\mathbf{a}^T$ from left side and $\mathbf{H}^{-1}\mathbf{a}$ from right side, where $\mathbf{a}$ is an arbitrary vector of length of Θ, gives

$$\int_{-\infty}^{\infty}\mathbf{a}^T(\hat{\Theta}-\Theta)\frac{\partial \ln p(\mathbf{Z})}{\partial\Theta}\mathbf{H}^{-1}\mathbf{a}p(\mathbf{Z})d\mathbf{Z} = \mathbf{a}^T\mathbf{H}^{-1}\mathbf{a} \tag{18}$$

Squaring both sides of this equation, using the fact that the square of a scalar can be treated as the product of the scalar and its transpose, or vice versa, and applying the Schwartz inequality to the result leads us to

$$\begin{aligned}\left(\mathbf{a}^T\mathbf{H}^{-1}\mathbf{a}\right)^2 &\leq \int_{-\infty}^{\infty}\left[\mathbf{a}(\hat{\Theta}-\Theta)\right]^2 p(\mathbf{Z})d\mathbf{Z}\mathbf{a}^T\mathbf{H}^{-1}\int_{-\infty}^{\infty}\left[\frac{\ln p(\mathbf{Z})}{\partial\Theta}\right]^2 p(\mathbf{Z})d\mathbf{Z}\mathbf{H}^{-1}\mathbf{a} \\ &\leq \mathbf{a}^T\mathbf{E}\left\{\tilde{\Theta}\tilde{\Theta}^T\right\}\mathbf{a}\mathbf{a}^T\mathbf{H}^{-1}\int_{-\infty}^{\infty}\left[\frac{\partial \ln p(\mathbf{Z})}{\partial\Theta}\right]\left[\frac{\partial \ln p(\mathbf{Z})}{\partial\Theta}\right]^T p(\mathbf{Z})d\mathbf{Z}\mathbf{H}^{-1}\mathbf{a} \\ &\leq \mathbf{a}^T\mathbf{E}\left\{\tilde{\Theta}\tilde{\Theta}^T\right\}\mathbf{a}\mathbf{a}^T\mathbf{H}^{-1}\mathbf{H}\mathbf{H}^{-1}\mathbf{a} \\ &\leq \mathbf{a}^T\mathbf{E}\left\{\tilde{\Theta}\tilde{\Theta}^T\right\}\mathbf{a}\left(\mathbf{a}^T\mathbf{H}^{-1}\mathbf{a}\right)\end{aligned} \tag{19}$$

Therefore

$$\mathbf{E}\left\{\tilde{\Theta}\tilde{\Theta}^T\right\} \geq \mathbf{H}^{-1} \tag{20}$$

Note that for an under-determined case like Diffuse Optical Tomography, $\mathbf{H}$ is regularized to find its inverse.

6.4. CRLB for real-time data with Gaussian noise

Estimation of parameters from noisy data is of interest in many applications like data set testing, radar,[14] and other measurement situations.[13,15,32] The data model in presence of noise is given by

$$\mathbf{Z} = \Gamma(\Theta) + \Psi \tag{21}$$

where $\mathbf{Z}$ (i.e.,$\mathbf{Z} = col(z_0, z_1, \ldots, z_{N-1})$) is a finite set of measurements and Γ (i.e.,$\Gamma = col(\Gamma_0, \Gamma_2, \ldots, \Gamma_{N-1})$) is the model vector with parameter set Θ (i.e.,$\Theta = col(\Theta_1, \Theta_2, \ldots, \Theta_p)$). Ψ is the noise vector with mean zero and variance σ^2. The joint probability density function of the elements of the sample vector $\mathbf{Z}$, when the unknown parameter vector is Θ, is given by

$$p(\mathbf{Z};\Theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(\mathbf{Z}_n - \Gamma_n)^2\right] \tag{22}$$

The CRLB for parameter set Θ for noisy data where the noise follows a Gaussian distribution can be found by developing the Fisher information matrix $\mathbf{H}$ (definition is given in the previous section).

In other words, if $\hat{\Theta}_r$ (any r^{th} element of $\hat{\Theta}$) is an unbiased estimate of deterministic Θ_r, the lower bound on variance is given by

$$E\{\hat{\Theta}_r^2\} \geq \mathbf{H}_{r,r}^{-1}, \text{ for } r = 1, 2, 3, ..., p \tag{23}$$

where, $\mathbf{H}_{r,r}^{-1}$ is the (r,r) element in matrix $\mathbf{H}^{-1}$. It is well understood that the CRB for each unknown parameter can be determined by observing the diagonal elements of the inverse of the Fisher Information matrix, $\mathbf{H}^{-1}$. Taking the logarithmic function (note that $\log p$ is asymptotic to $p(\mathbf{Z_n};\Theta)$), results in

$$\log p = \text{constant} - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(\mathbf{Z}_n - \Gamma_n)^2 \tag{24}$$

Differentiating Eq. (24) with the r^{th} element in Θ, say Θ_r, we obtain,

$$\frac{\partial \log p}{\partial \Theta_r} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\frac{\partial \Gamma_n}{\partial \Theta_r}(\mathbf{Z}_n - \Gamma_n) \tag{25}$$

Similarly, Eq. (25) can be differentiated with respect to any other s^{th} element Θ_s. The r^{th} and s^{th} element of $\mathbf{H}$ can be written as

$$\mathbf{H}_{r,s} = \mathbf{E}\left\{\frac{1}{\sigma^4}\sum_{n=0}^{N-1}\frac{\partial \Gamma_n}{\partial \Theta_r}(\mathbf{Z}_n - \Gamma_n)\sum_{l=0}^{N-1}\frac{\partial \Gamma_l}{\partial \Theta_s}(\mathbf{Z}_l - \Gamma_l)\right\} \tag{26}$$

Equation (26) can be simplified into the following compact form

$$\mathbf{H}_{r,s} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\frac{\partial \Gamma_n}{\partial \Theta_r}\frac{\partial \Gamma_n}{\partial \Theta_s} \tag{27}$$

Finally, the lower bounds are given by variance, $\mathbf{E}\left\{\tilde{\Theta}_r^2\right\} \geq \mathbf{H}_{r,r}^{-1}$. This is known as *Cramér-Rao Lower Bound*(CRLB). The closer the value of $\mathbf{E}\left\{\tilde{\Theta}_r^2\right\}$ is to $\mathbf{H}_{r,r}^{-1}$, the more efficient the estimator becomes. The standard deviation is the square root of the variance calculated from CRLB.

7. ACKNOWLEDGMENTS

P.K.Y. acknowledges the DOD Breast Cancer pre-doctoral fellowship (BC050309). This work has been sponsored by the National Cancer Institute through grants RO1CA78734, PO1CA80139, and DAMD17-03-1-0405.

REFERENCES

1. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack and K. D. Paulsen, "Interpreting Hemoglobin and Water Concentration, Oxygen Saturation and Scattering Measured In Vivo by Near-Infrared Breast Tomography," Proc. Nat. Acad. Sci. **100**, 12349–12354 (2003).
2. D. A. Boas, A. M. Dale and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution and accuracy," NeuroImage **23**, S275–S288 (2004).
3. S. R. Arridge, "Optical tomography in medical imaging," Inv. Problems **15**, R41–R93 (1999).
4. A. N. Tikhonov, "Regularization of mathematically incorrectly posed problems," Soviet Math. Dokl. **4**, 1624-1627 (1963).
5. X. Cheng and D. Boas, "Systematic diffuse optical image errors resulting from uncertainty in the background optical properties," Opt. Express **4**, 299–307 (1999). http://www.opticsinfobase.org/abstract.cfm?URI=oe-4-8-299
6. P. K. Yalavarthy, H. Dehghani, B. W. Pogue, C. M. Carpenter, S. Jiang and K. D. Paulsen, "Structural information within regularization matrices improves near infrared diffuse optical tomography," IEEE Trans. Med. Imag. (submitted) 2006.
7. K. D. Paulsen, P. M. Meaney, M. J. Moskowitz and J. M. Sullivan Jr., "A dual mesh scheme for finite element based reconstruction algorithms," IEEE Trans. Med. Imag. **14**, 504–514 (1995).
8. B. Brooksby, S. Jiang, H. Dehghani, B. W. Pogue, K. D. Paulsen, J. Weaver, C. Kogel and S. P. Poplack, "Combining near infrared tomography and magnetic resonance imaging to study in vivo breast tissue: implementation of a Laplacian-type regularization to incorporate magnetic resonance structure," J. Biomed. Opt. **10**, 051504:1–10 (2005).
9. B. Brooksby, B. W. Pogue, S. Jiang, H. Dehghani, S. Srinivasan, C. Kogel, T. D. Tosteson, J. Weaver, S. P. Poplack and K. D. Paulsen, "Imaging Breast Adipose and Fibroglandular Tissue Molecular Signatures using Hybrid MRI-Guided Near-Infrared Spectral Tomography," Proc. Nat. Acad. Sci. U.S.A. **103**, 8828–8833 (2006).
10. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, 1993).
11. J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control* (Englewood Cliffs, NJ, USA: Prentice-Hall, 1995)
12. B. W. Pogue, X. Song, T. D. Tosteson, T. O. McBride, S. Jiang and K. D. Paulsen, "Statistical Analysis of Non-linearly Reconstructed Near-Infrared Tomographic Images: Part I - Theory and Simulations," IEEE Trans. Med. Imag. **21**, 755-763 (2002).
13. M. Gustafsson and S. Nordebo, "Cramér-Rao lower bounds for inverse scattering problems of multilayer structures," Inv. Problems **22**, 1359–1380 (2006).
14. J. Chul Ye, Y. Bresler and P. Moulin, "Cramér-Rao bounds for 2-D target shape estimation in nonlinear inverse scattering problems with application to passive radar," IEEE Trans. Anten. Prop. **49**, 771–783 (2001).
15. D. B. Rowe, "Parameter estimation in the magnitude-only and complex-valued fMRI data models," NeuroImage **25**, 1124–1132 (2005).
16. H. Jiang, K. D. Paulsen, U. Osterberg, B. W. Pogue, and M. S. Patterson, "Simultaneous reconstruction of optical absorption and scattering maps in turbid media from near-infrared frequency-domain data," Opt. Lett. **20**, 2128-2130 (1995).
17. K. Levenberg, "A method for the solution of certain nonlinear problems in least squares," Q. Appl. Math. **2**, 164-168 (1944).
18. D. W. Marquardt, "An algorithm for least squares estimation of nonlinear parameters," J. Soc. Ind. Appl. Math. **11**, 431-441 (1963).
19. S. Srinivasan, B. W. Pogue, H. Dehghani, S. Jiang, X. Song, and K. D. Paulsen, "Improved quantification of small objects in near-infrared diffuse optical tomography," J. Biomed. Opt. **9**, 1161–1171 (2004).
20. 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**, 185-193 (2000).

21. S. Jiang, B. W. Pogue, T. O. McBride, M. M. Doyley, S. P. Poplack, and K. D. Paulsen, "Near-infrared breast tomography calibration with optoelastic tissue simulating phantoms," J. Electron. Imaging **12**, 613–620 (2003).
22. B. Brooksby, H. Dehghani, B. W. Pogue and K. D. Paulsen, "Near infrared (NIR) tomography breast image reconstruction with *a priori* structural information from MRI: algorithm development for reconstructing heterogeneities," IEEE J. Sel. Top. Quant. Elec. **9**, 199–209 (2003).
23. P. K. Yalavarthy, H. Dehghani, B. W. Pogue, and K. D. Paulsen, "Critical computational aspects of near infrared circular tomographic imaging: Analysis of measurement number, mesh resolution and reconstruction basis," Opt. Express **14**, 6113–6127 (2006).
24. 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. Instr. **72**, 1817–1824 (2001).
25. B. W. Pogue, S. C. Davis, X. Song, B. Brooksby, H. Dehghani and K. D. Paulsen, "Image analysis methods for Diffuse Optical Tomography," J. Biomed. Opt. 11(3), 033001:1–16 (2006).
26. B. W. Pogue, T. Mcbride, J. Prewit, U. L. Osterberg, and K. D. Paulsen, "Spatially variant regularization improves diffuse optical tomography," App. Opt. **38**, 2950–2961 (1999).
27. A. Douiri, M. Schweiger, J. Riley, and S. Arridge, "Local diffusion regularization method for optical tomography reconstruction by using robust statistics," Opt. Lett. **30**, 2439–2441 (2005).
28. P. C. Hansen, "Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems," Numerical Algorithms **6**, 1–35 (1994).
29. X. Song, B. W. Pogue, T. D. Tosteson, T. O. McBride, S. Jiang and K. D. Paulsen, "Statistical Analysis of Non-linearly Reconstructed Near-Infrared Tomographic Images: Part II Experimental Interpretation," IEEE Trans. Med. Imag. **21**, 764–772 (2002).
30. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani and K. D. Paulsen, "Spectrally Constrained Chromophore and Scattering NIR Tomography Provides Quantitative and Robustness of Reconstruction," App. Opt. **44**, 1858–1869 (2005).
31. 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, "In Vivo Hemoglobin and Water Concentration, Oxygen Saturation, and Scattering Estimates from Near-Infrared Breast Tomography Using Spectral Reconstruction," Acad. Radiology **13**, 195–202 (2006).
32. S. Yazdanfar, C. Yang, M. Sarunic, and J. Izatt, "Frequency estimation precision in Doppler optical coherence tomography using the Cramer-Rao lower bound," Opt. Express **13**, 410–416 (2005).