Bayesian estimation of optical properties of the human head via 3D structural MRI

Alex H. Barnett[a], Joseph P. Culver[b], A. Gregory Sorensen[b], Anders M. Dale[b] and David A. Boas[b]

[a]Courant Institute, New York University, 251 Mercer Street, New York, NY 10012
[b]NMR Center, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129

ABSTRACT

Knowledge of the baseline optical properties of the tissues of the human head is essential for absolute cerebral oximetry, and for quantitative studies of brain activation. In this work we numerically model the utility of signals from a small 6-optode time-resolved diffuse optical tomographic apparatus for inferring baseline scattering and absorption coefficients of the scalp, skull and brain, when complete geometric information is available from magnetic resonance imaging (MRI). We use an optical model where MRI-segmented tissues are assumed homogeneous. We introduce a noise model capturing both photon shot noise and forward model numerical accuracy, and use Bayesian inference to predict errorbars and correlations on the measurments. We also sample from the full posterior distribution using Markov chain Monte Carlo. We conclude that $\sim 10^6$ detected photons are sufficient to measure the brain's scattering and absorption to a few percent. We present preliminary results using a fast multi-layer slab model, comparing the case when layer thicknesses are known versus unknown.

Keywords: Diffuse optical tomography, Bayesian, posterior, magnetic resonance imaging

1. INTRODUCTION

Diffuse optical tomography (DOT) is arousing growing interest as a tool for noninvasive imaging of optically active materials within living tissue. In particular, in the last decade DOT has been used for in vivo studies of brain activation,[1,2] including detection of the fast signal.[3] DOT has many advantages over function magnetic resonance imaging (fMRI) for brain activation studies, and has been validated against fMRI in the rat model.[4] The use of near-infrared light at multiple wavelengths allows independent measurement of oxy-hemoglobin (HbO_2) and deoxy-hemoglobin (HbR) concentrations (fMRI is only sensitive to HbR). DOT is portable, an order of magnitude cheaper than MRI, has an order of magnitude faster temporal resolution, and could be used for longer term bedside monitoring. Because of their different strengths, we view DOT and MRI as complementary modalities.

In DOT, tissues of interest are typically many times thicker than the ballistic mean free paths of photon transport, so the passage of light from a source to a tissue region of interest, and finally to a detector, is essentially diffusive. This removes the possibility of focussing imaging. Instead an inverse problem must be solved to create a spatial map of absorption coefficient μ_a and reduced scattering coefficient μ_s', given surface intensity measurements. This is mathematically and computationally nontrivial, and has hindered the widespread use of DOT as an imaging modality.

In this work we introduce two techniques to tackle the inverse problem:

1. We use structural information from a MRI in order to build an optical model for DOT. The optical model assumes homogenous scattering and absorption coefficients assigned to each of several tissue types: scalp, skull, cerebro-spinal fluid (CSF), and brain. The three-dimensional geometrical shapes of these tissue types are extracted with high resolution (1 mm) from structural MRI. Thus we reduce the inverse problem to one with a relatively small number (10 or less) of unknowns, namely the absorption and scattering coefficients of the tissue types. This is often referred to as a 'fitting' problem.

Further author information: Send correspondence to A.H.B., E-mail: barnett@cims.nyu.edu, Telephone: +1 212 998 3296

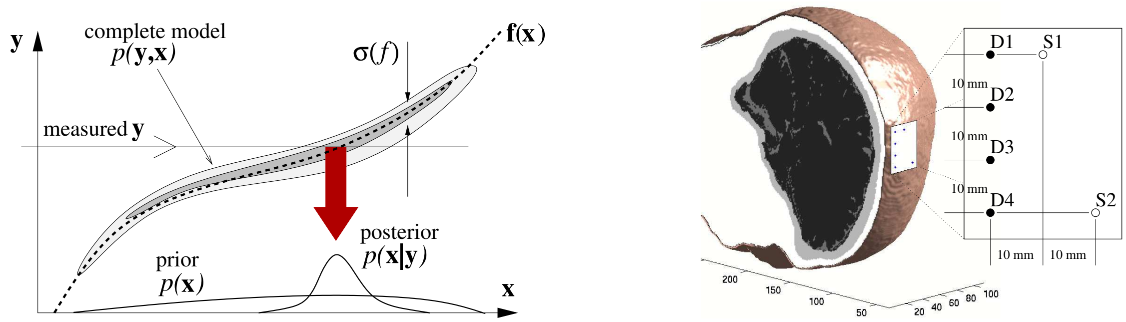

Figure 1. Left (a): Illustration of Bayesian inference of unknown parameters $\mathbf{x}$ given the measured signal vector $\mathbf{y}$. The many-dimensional spaces of both $\mathbf{x}$ and $\mathbf{y}$ are each represented as one-dimensional for clarity. The posterior is simply the slice across the joint distribution at the detected signal value $\mathbf{y}$. **Right (b):** slice through MRI-segmented head model, showing scalp, skull, CSF, and brain, and our surface optode arrangement.

2. We use Bayesian inference to keep track of exactly what our signal measurements are telling us about these unknowns. This is a *model-based* approach, which forces us to be clear about our noise model assuptions, and allows us to predict errorbars on our inferred optical properties. Our Bayesian approach goes further than those in the existing DOT literature (see Section 2.2). We believe our approach will be invaluable for the noisy multi-parameter fitting problems which arise in layered DOT (and other areas of applied science), especially when unknown calibration parameters are included.

The utility of accurate baseline optical measurements of the human head is manyfold, however obtaining baseline accuracy has proved a harder problem than obtaining changes. Absolute cerebral oximetry and blood volume measurement, which require baseline measurement of μ_a of the brain, have clinial importance for stroke,[6] head trauma, migraine, neonatal ischemia[7] and brain development.[8] Knowledge of accurate baseline μ_a and μ_s' for each of the scalp, skull, and brain itself, is essential to proceed with quantitative brain activation studies.[9,10] Knowledge of depth information (the local thickness of the scalp and skull can vary significantly across the cranium of a single subject) is also needed for accurate activation studies.[9,11] In Section 4.2 we show that with a multi-layer slab model, unknown layer thicknesses can be inferred as part of the inverse problem, but that prior knowledge of the thickness improves the accuracy of optical measurements.

2. METHODS

The main question under study is: Given an optode arrangement, and a structural MRI, how accurately can time-resolved DOT measure the baseline μ_a and μ_s' of the scalp, skull, and brain? We will use numerical simulations of all data to answer this question. We consider a 2-source, 4-detector apparatus (Fig. 1b), which collects single photons, recording which timegate they arrive in, thus recording the temporal point spread function (TPSF). The timegates are equally spaced in time, with width 0.2 ns, starting from 0.2 ns and ending at 2 ns. We have found that for our parameter values the signals are rarely needed beyond 2 ns.

We use the notation $\mathbf{x}$ to represent the length-N vector of unknown optical parameters, and $\mathbf{y}$ to represent the signal vector (of length M). $\mathbf{y}$ contains the histogram of photon counts in each timegate, *i.e.* the TPSF, for each of the 8 source-detector pairs. Fig 2b shows a typical simulated signal. In this example we have not adjusted (numerical) attenuation factors: it is possible to set these so each detector receives the same number of photons (as would be usually be done in an experimental setting).

2.1. Segmented Diffusion Forward Model

The forward model $\mathbf{f}(\mathbf{x})$ takes as input the parameters $\mathbf{x}$ and returns the corresponding detected signal vector $\mathbf{f}$. The diffusion approximation to photon transport[12] describes the time-evolution of the fluence $\phi(\mathbf{r}, t)$ (proportional to photon density) according to

$$\frac{1}{c}\frac{\partial \phi}{\partial t} \;=\; \nabla \cdot \left(\frac{1}{3\mu_s'(\mathbf{r})} \nabla \phi \right) - \mu_a(\mathbf{r})\phi + q(\mathbf{r}, t), \tag{1}$$

Table 1. Default homogeneous tissue optical properties used in the study, with approximate geometry

tissue	μ_a (mm^{-1})	μ_s' (mm^{-1})	shape
scalp	0.015	0.8	$\sim$ 5 mm layer
skull	0.01	1.0	$\sim$ 7 mm layer
CSF	0.0004	0.4	folded 1–3 mm sheet
brain	0.018	1.3	$\sim$ 1 cm folds (sulci)

where we have used the good approximation[12] that the diffusion constant is $1/(3\mu_s')$. Diffusion is valid when $\mu_a \ll \mu_s'$ and length-scales of interest are $\gg 1/\mu_s'$, and if appropriate extrapolated boundary conditions are used. The source term q is chosen as an impulsive point source a distance $1/\mu_s'$ below the location of the source fiber.

Eq. 1 is solved numerically using a finite-difference scheme based on a cubic lattice of cells. Each cell is assigned a μ_a and μ_s' according to which tissue type it predominantly falls inside: scalp, skull, CSF, brain (grey and white matter are lumped together, since only grey matter is optically relevant). The segmentation map comes from an automatic procedure for segmenting 3d structure MRI images due to Dale, Fischl and Sereno.[13] Our default optical properties, shown in Table 1, are literature values (for instance Ref. 15), however there is much uncertainty about their possible range. The parameter vector is $\mathbf{x} = \{\mu_{a,\text{scalp}},\ \mu_{a,\text{skull}},\ \mu_{a,\text{brain}},\ \mu'_{s,\text{scalp}},\ \mu'_{s,\text{skull}},\ \mu'_{s,\text{brain}}\}$.

We use an effective μ_s' for the CSF of 0.4 mm^{-1}, as explained in our previous work,[16] even though the μ_s' of the bare CSF clear liquid may be as low as 0.01 mm^{-1}. The issue of the CSF will not be discussed further here. It is an important issue whose impact on the validity of the diffusion approximation to light transport in the human head has yet to be resolved. Recent fitted figures for an effective μ_s' of the CSF using Monte Carlo models[14] give the figures 0.16-0.32mm^{-1}, similar to our choice. The consequence is that diffusion may be an accurate option.

The finite difference scheme simulates the $\phi(\mathbf{r}, t)$ evolution in a sufficiently-large 3D chunk of the head surrounding the optodes. It can handle arbitrary tissue geometry. At a typical $h = 2$ mm cell size this comprises about 4×10^4 cells. Timestepping is performed using a simple forward Euler (explicit) scheme, giving a CPU time on a 1 GHz P3 or about 8 seconds per source. Currently the extrapolated boundary conditions a the scalp-air interface are handled using a crude voxellated definition of the interface. The whole scheme can be improved upon greatly, for instance using implicit timestepping methods (which would improve upon the current $O(h^{-5})$ scaling of numerical effort) and using sub-voxel (smoother) representations of the tissue-tissue and tissue-air interfaces. Currently the forward model error is a few percent

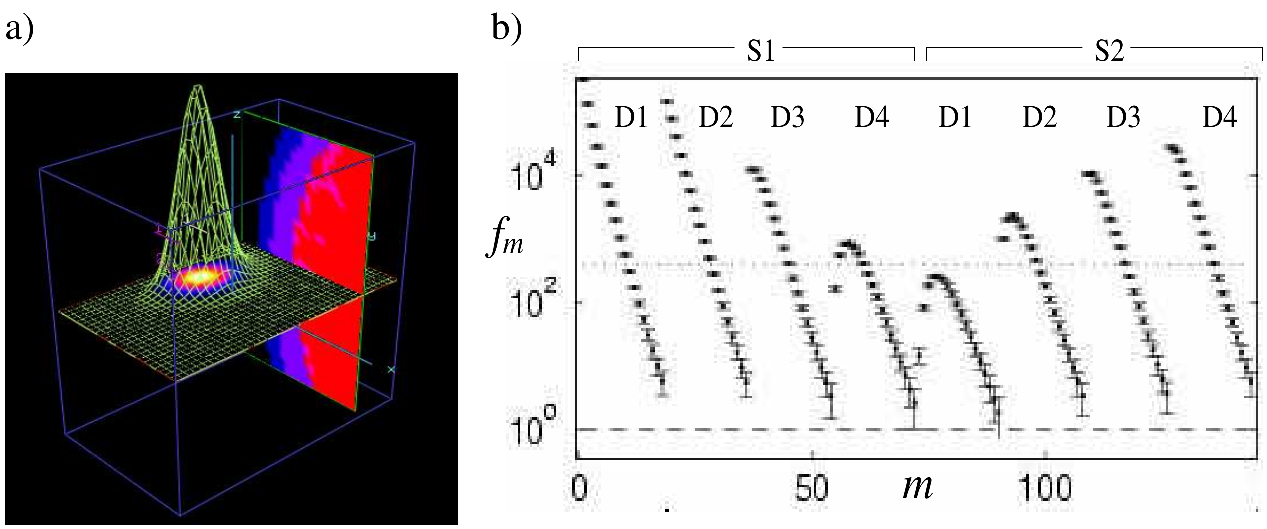

Figure 2. Left (a): Horizontal slice of fluence ϕ at a particular instant of time shortly after fluence injection, inside the head, as calculated by the forward model. The tissue segmentation (2 mm voxels) is shown by a vertical slice. **Right (b):** Example signal vector $\mathbf{y}$ extracted from $\phi(t)$ (and its normal derivative) at the detector locations, due to each source.

at later times, and over 10% before the first 200 ps. The error scales at best linearly with h due to the crude voxellated representation of the geometry. More numerical details can be found in Ref 16.

2.2. Bayesian Inference

Bayesian inference[17] is a statistically rigorous procedure for handling uncertainty in model-based fitting. Since the DOT inverse problem is ill-posed,[12] our segmented fitting problem retains much of this ill-posedness in the form of well- and badly-constrained directions in $\mathbf{x}$-space. It is necessary therefore to handle the full probability density function (PDF) on $\mathbf{x}$. Bayesian methods embrace the ill-posedness, allowing you to assess the constraints placed on the unknowns by the data, before any regularization is applied.

As sketched in Fig. 1a, the complete model is a joint distribution $p(\mathbf{y}, \mathbf{x})$. Assumptions about $\mathbf{x}$ before the signal is measured are encoded in the *prior* $p(\mathbf{x})$; when the detected signal vector $\mathbf{y}$ arrives our new complete knowledge about $\mathbf{x}$ is encoded in the *posterior*,

$$p(\mathbf{x}|\mathbf{y}) \;\propto\; p(\mathbf{y}, \mathbf{x}) \;=\; p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x}). \tag{2}$$

The *likelihood* $p(\mathbf{y}|\mathbf{x})$ describes how likely we believe a hypothetical signal $\mathbf{y}$ to be given known parameters $\mathbf{x}$. Generally regularization in other methods is interpreted as a choice of prior; here we use a flat (uniform) prior only.

Studying properties of the full posterior is an essential part of the Bayesian inverse method, because distributions on some unknowns can be very wide. Previous Bayesian DOT (for instance, the work of Oh[18]) has elected only to find the *best fit* solution, corresponding to the *maximum a-priori* (MAP) value $\mathbf{x}_{\text{MAP}}$. It is quite easy to show that the location of $\mathbf{x}_{\text{MAP}}$ can be moved around by reparametrization (for instance using diffusion coefficient $1/(3\mu_s')$ rather than μ_s' to parametrize the degree of scattering), therefore that $\mathbf{x}_{\text{MAP}}$ has a certain arbitrariness. This effect becomes significant only when the posterior PDF is wide in the corresponding parameter. This indicates that consideration of the full PDF can be important. The errorbars and correlations on fitted parameters also follow from the posterior PDF; it is our view that it is essential to present errorbars along with any fitted results.

We explore the posterior on $\mathbf{x}$ using two methods:

1. Gaussian approximation of the posterior. This assumption allows us to characterize the posterior $p(\mathbf{x}|\mathbf{y})$ by $\mathbf{x}_{\text{MAP}}$ and the second derivative (Hessian) matrix $H = -\nabla\nabla \log p(\mathbf{x}|\mathbf{y})$ alone. The search for $\mathbf{x}_{\text{MAP}}$ is done via Levenburg-Marquardt optimization, using an approximation to H (details in Ref. 16).

2. Sampling the exact posterior via Markov chain Monte Carlo[19] (MCMC). This statistical technique requires many $\mathbf{f}(\mathbf{x})$ evaluations (we find about 40 per independent sample are needed for $N = 6$ dimensions, using the Metropolis algorithm using a hyperellipsoid step distribution given by H). However, there are very few other ways to examine a high-dimensional distribution, therefore we consider this CPU-intensive method important.

The latter method allows us to see when the former approximation is valid.

2.3. Noise Model

The choice of noise model is crucial since the size of the noise is responsible for determining the final measurement errorbars. We assume independent Gaussian noise with signal-dependent width $\sigma(f)$. The assumption of independence is valid since our timegates The likelihood function is then $p(\mathbf{y}|\mathbf{x}) \propto e^{-L(\mathbf{y};\mathbf{x})}$, with

$$L(\mathbf{y}; \mathbf{x}) \;=\; \frac{1}{2}\sum_{m=1}^{M} \ln \sigma_m^2 + \frac{1}{2}\sum_{m=1}^{M} \frac{(f_m(\mathbf{x}) - y_m)^2}{\sigma_m^2}, \tag{3}$$

where σ_m abbreviates $\sigma(f_m)$. Thus our Bayesian approach corresponds (approximately) to a least-squares fit.

Detector shot noise is Poissonian, and would imply a noise function choice $\sigma(f) = f^{1/2}$ for $f \gg 1$. For instance, if 10^6 photons were detected in a signal component f_m, the corresponding relative error would be 0.1%. However, it is unrealistic to trust our forward model to this high accuracy! Therefore we use a noise function which transitions to constant relative error ϵ at large signal values:

$$\sigma(f) \;=\; 1 + f^{1/2} + \epsilon f. \tag{4}$$

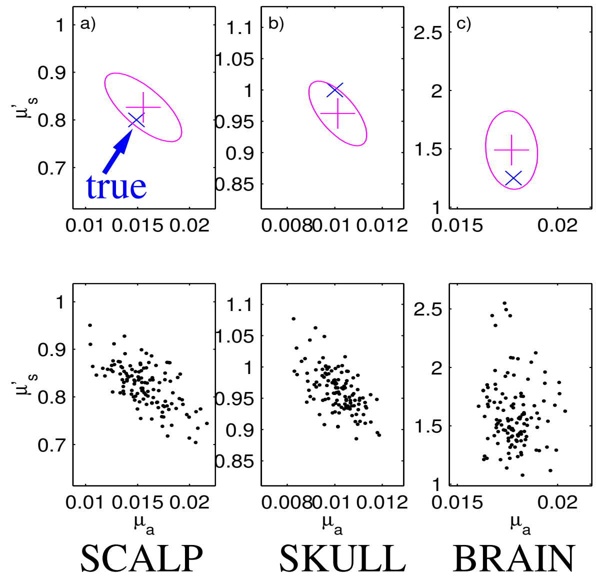

Figure 3. Views of the posterior distribution $p(\mathbf{x}|\mathbf{y})$ with signals $\mathbf{y}$ from the same forward model ($h = 2$ mm), using standard optical parameters in Table 1. The three columns a,b,c show the μ_a–μ'_s plane, separately for each tissue type. The top row shows true $\mathbf{x}$ (cross), $\mathbf{x}_{\text{MAP}}$ (crosshairs), and the *marginal* PDF as an elliptical contour enclosing 63% of the probability mass in the Gaussian posterior approximation. The bottom row shows 123 independent samples from the posterior found using MCMC, on the same pairs of axes. This figure is adapted from Ref 16.

Typically ϵ is a few percent, and includes errors in the physical assumptions and numerical forward model, and systematic multiplicative measurement errors. For now we assume uncorrelated forward model errors (the timegate width 0.2 ns is not too much smaller than any typical correlation times that this is unreasonable). Typical $\sigma(f)$ sizes are shown as errorbars in the signal Fig. 2b.

3. SEGMENTED RESULTS

All results are performed using simulated signals which include shot noise. Fig. 3 shows the posterior distribution on the six unknown optical parameters $\mathbf{x}$ which results when $N_p = 3.2 \times 10^5$ photons are detected, and an assumption of 10% model errors is used ($\epsilon = 0.1$). The PDF has been projected (marginalized) onto 3 pairs of axes (in the Gaussian PDF approximation this requires matrix manipulations described in Ref. 16). The full 6-dimensional PDF is harder to visualize (it turns out that it has a minor to major axis ratio of about 10^2, in the Gaussian approximation, implying it is pancake-like but happens not to be oriented along the axes). Notice that the true values fall within the bulk of the inferred posterior PDF, indicating a correct inference.

Fig. 4 shows corresponding one-dimensional projections(for convenience these have actually been generated using the forward model of Section 4.1, but the results are very similar). It is clear that even with this relatively small number of photons then all parameters except $\mu'_{s,\text{brain}}$ have acceptable widths of their distribution, that is, less than 10% measurement errorbars. The width of the marginal PDF is interpreted simply as the measurement errorbar on each parameter. Note that for $\mu'_{s,\text{brain}}$, which has not been very well measured (over 20% errors), that the 68% confidence interval generated from MCMC sampling is poorly approximated by the Gaussian approximation errorbar (68% confidence intervals correspond to $\pm 1\sigma$ for a Gaussian).

By varying N_p and ϵ we can find out what values are needed to reduce the errors on $\mu'_{s,\text{brain}}$ to 10%. Typically we find $N_p = 2 \times 10^6$ and $\epsilon = 5\%$ is sufficient. Of course, these values will change for other choices of scalp and skull thicknesses and optical properties; a study of more subject data is needed. Fig. 5 shows the tracking of measured versus true optical properties of the brain, using these values. The tracking is very linear, and absorption-scattering crosstalk is negligible. Crosstalk with the other four scalp and skull parameters is also negligible (not shown).

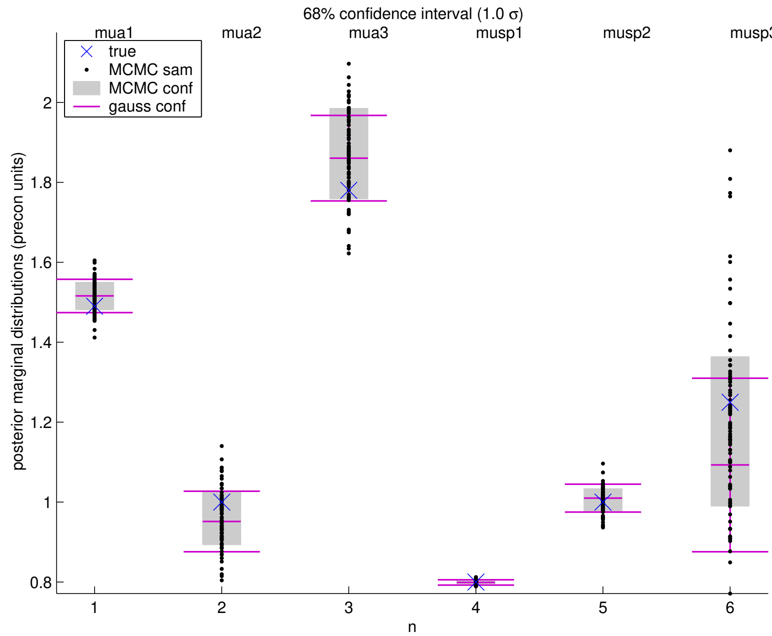

Figure 4. Posterior PDF marginalized onto each of the 6 parameters x_n separately. True values are shown by the crosses. Confidence intervals for the Gaussian approximation are shown as errorbars. Confidence intervals from MCMC are shown as a grey bar (MCMC samples are shown as dots, as before). $n = 1, 2, 3$ are μ_a for scalp, skull, brain (multiplied by factor 10^2), and $n = 4, 5, 6$ are corresponding μ'_s. Units are mm^{-1}. Note the poor Gaussian approximation for $n = 6$ ($\mu'_{s,\text{brain}}$), due to the long tail in the PDF.

4. PARTIAL GEOMETRIC INFORMATION

Here we present preliminary results using a different forward model which approximates the human head geometry in the region of the optodes: a multi-layer slab geometry. This forward model is about 10^2 times faster than the 3D segmented model, allowing us to tackle new questions. We will give evidence for the improvement that knowledge of layer thickness can provide, hence for the value of MRI structural information.

4.1. Layered Slab Forward Model

The layered diffusion problem is separable, allowing us to Fourier decompose the fluence $\phi(R, z, t)$ where z is depth into the medium and R is radial distance from the source in the xy-plane, as follows,

$$\phi(R, z, t) \;=\; \sum_{m=1}^{M} B_m \psi_m(R) u_m(z, t). \tag{5}$$

We have truncated the system size to a cylindrical volume of maximum radius R_{max} (hence the discrete rather than continuous sum in m), and maximum depth z_{max}. The ψ_m are Bessel functions (eigenmodes of the 2D laplacian in the circle of radius R_{max}, with Dirichlet boundary conditions). The B_m are constants which depend on the initial condition. The exact solution requires $M \to \infty$, however, we can truncate M to a finite value (typically of order 50), by using the facts that high Fourier modes die rapidly in time, and that signals before $t_{\text{min}} = 0.2$ ns are not needed. Each $u_m(z, t)$ obeys a 1D diffusion equation with modified $\mu_a(z)$. These 1D equations are discretized in z via a finite difference lattice (typically $h = 0.25$ mm), and Crank-Nicholson timestepping is used. Because of the lack of a source term beyond $t = 0$, the timesteps can be made proprotional to t. In this way the total solution time scales like $\log(t_{\text{max}}/t_{\text{min}})$. This forward model requires a CPU time of 0.2 sec on a 1 GHz P3, and has errors less than 1% (relative to the exact diffusion solution). Essentially arbitrary continuous or discontinous μ_a and μ'_s as a function of z require little extra effort. Details will appear in a forthcoming publication.

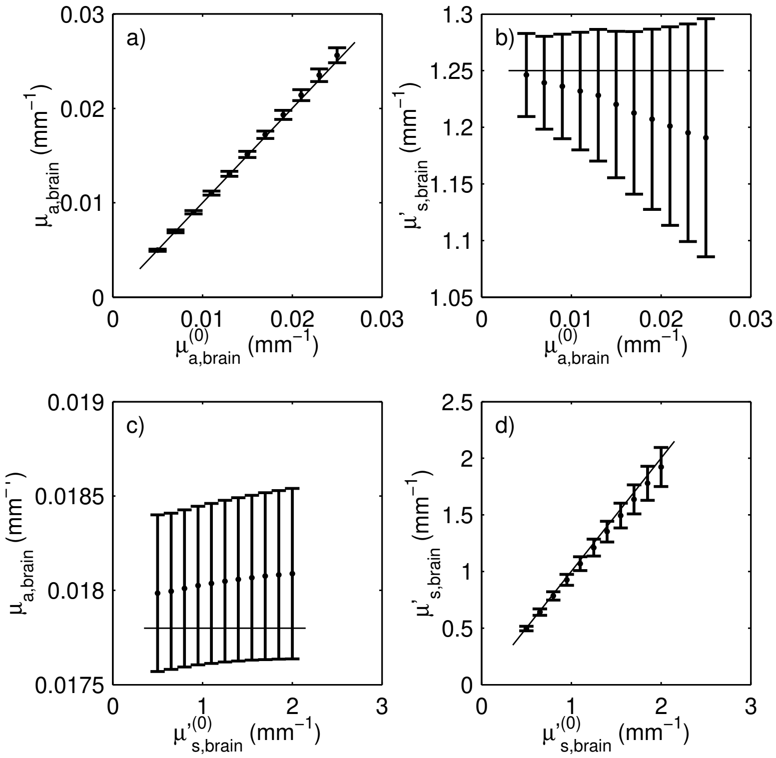

Figure 5. Tracking of inferred vs true optical parameters of the brain. a) and b) show the effect that changing the true absorption $\mu_{a,\text{brain}}^{(0)}$ (horizontal axis) has on inferred $\mu_{a,\text{brain}}$ and $\mu'_{s,\text{brain}}$ respectively (vertical axes). c) and d) show the same except changing the true reduced scattering $\mu'^{(0)}_{s,\text{brain}}$. (The scalp and skull are fixed at the standard parameters, and inferred scalp and skull parameters are not shown). This figure is taken from Ref 16.

4.2. Results: Knowledge of Layer Thickness

We apply the above forward model to a 2-layer system: a top layer representing the scalp and skull, the lower layer representing the brain (as before, optically dominated by grey matter). We have verified that a layered system can well approximate the true head geometry, as far as the diffusion approximation is concerned. (Verification against Monte Carlo simulation of the transport equation is a matter for future investigation).

Fig. 6 shows the relative errorbars from the Bayesian fitting procedure, as a function of top layer thickness Δz. The solid lines show the errorbar size with unknown parameters $\mathbf{x} = \{\mu_{a,1},\ \mu_{a,2},\ \mu'_{s,1},\ \mu'_{s,2},\ \Delta z\}$, for each parameter. The dotted lines show the errorbar size when Δz is known exactly and $\mathbf{x}$ comprises only the remaining four optical parameters. 4 detectors at distances from 10 mm to 30 mm from the source are used, with $N_p = 4 \times 10^5$ and $\epsilon = 2\%$. For $\Delta z > 10$ mm, knowledge of Δz improved errorbars on $\mu_{a,2}$ by about a factor of 3. In order words, we predict a factor 3 improvement in baseline accuracy of the absorption of the brain. This provides an indication of the utility that MRI structural information will provide in future DOT applications.

5. CONCLUSIONS

We have shown that baseline optical measurement of the tissue types of the human head is possible when full geometric information is available. The method we demonstrate allows the scalp, skull and brain optical properties to be measured independently without crosstalk. The detection of $\sim 10^6$ photons in a time-resolved single-photon-counting apparatus with a small number of optodes is predicted to allow measurement of the brain's absorption to a few percent and scattering to about ten percent. This type of error analysis allows optimal design of DOT apparatus. Because complete geometric information, and no calibration parameters, are assumed, these results are in effect the 'best-case scenario'. Hence they give lower bounds on the errors that could be achieved in practice. Introduction of calibration parameters and other sources

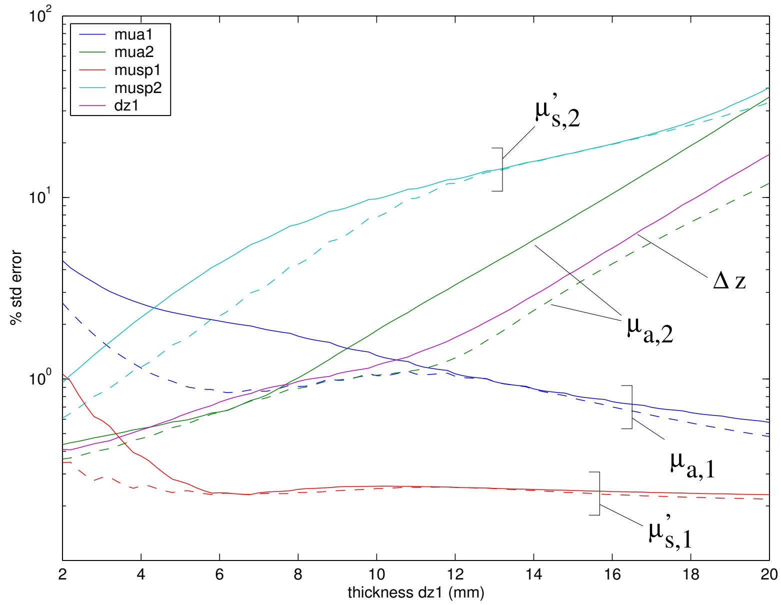

Figure 6. Relative errorbar size (in %) for fitted parameters in the vector $\mathbf{x}$ for a 2-layer model. Solid: the 4 optical parameters and thickness Δz are fitted for simultaneously. Dotted: only the 4 optical parameters are fitted for, with Δz taken as known. See text for details.

of uncertainty can only increase the errors. However we have not yet tried to optimize optode locations or numbers; this should give significant improvements to the errorbar figures quoted for the one particular opdote arrangement chosen above.

We have also shown that exact knowledge of layer thickness in a 2-layer approximate model of the head increases our measurement accuracy of μ_a of the lower layer (*i.e.* the brain) by about a factor of 3, compared to the incomplete knowledge case where the layer thickness was also fit for. This demonstrates the improvement that complete MRI structural information can provide over "blind" DOT. Clearly this 2-layer model is just a first step to investigating a series of approximate optical models for the human head. We believe that combined MRI and DOT will provide unprecedented information about brain function.

The small $N = 6$ dimensional parameter space allowed us to explore the full posterior on $\mathbf{x}$ using Markov chain Monte Carlo: this required many forward model calcualations. This would be much harder for the 10^3 or more unknowns that are used in imaging DOT inverse methods. However, we believe the Gaussian approximation could be tractable, and that the resulting knowledge of errorbars is essential. Since CPUs are advancing faster than DOT instrumentation, it will become increasingly advantageous to use CPU-intensive methods such as Bayesian inference which can extract the most information from the DOT signals.

As a final note, it is clear that the community is sorely in need of fast forward models (diffusion and ultimately transport equation) for complex geometries, of known accuracy. In this work we have not presented validation of diffusion against Monte Carlo simulated transport-equation signals. (However, reasonable agreement with Monte Carlo has been found[20]). This is a subtle issue which requires further investigation.

ACKNOWLEDGMENTS

We thank Anna Custo for her work on validating our finite-difference forward model. AHB is grateful for support from the Courant Institute of Mathematical Sciences. We also acknowledge funding from grants NIH 2P01-NS35611-06, NIH R29-NS38842, NIH P41-RR14075 and from the Center for Innovative Minimally Invasive Therapies. This research was funded in part by the US Army, under Cooperative Agreement No. DAMD17-99-2-9001. The material presented does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

1. D. A. Boas, D. H. Brooks, E. L. Miller, C. .A. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang, “Imaging the body with diffuse optical tomography”, IEEE Signal Processing Magazine **18**(6), 57–75 (2001).
2. A. Villringer and B. Chance, “Non-invasive optical spectroscopy and imaging of human brain function”, Trends Neurosci. **20**, 435–442 (1997).
3. J. Steinbrink, M. Kohl, H. Obrig, G. Curio, F. Syre, F. Thomas, H. Wabnitz, H. Rinneberg, and A. Villringer, “Somatosensory evoked fast optical intensity changes detected non-invasively in the adult human head”, Neuroscience Letters **291**, 105–108 (2000).
4. A. M. Siegel, J. P. Culver, J. B. Mandeville, and D. A. Boas, “Temporal comparison of functional brain imaging with diffuse optical tomography and fMRI during rat forepaw stimulation”, Phys. Med. Biol. **48** 1391–1403 (2003).
5. G. Strangman, J. .P. Culver, J. H. Thompson, and D. A. Boas, “A quantitative comparison of simlutaneous BOLD fMRI and NIRS recordings during functional brain activation”, NeuroImage **17**, 719–731 (2002).
6. W. D. Heiss, “Ischemic penumbra: Evidence from functional imaging in man”, Journal of Cerebral Blood Flow and Metabolism **20** 1276–93 (2000).
7. S. R. Hintz, W.-F. Cheong, J. P. van Houten, D. K. Stevenson, D. A. Benaron “Bedside imaging of intracranial hemorrhage in the neonate using light: comparison with ultrasound, computed tomography, and magnetic resonance imaging”, Pediatr. Res. **45**, 54–59 (1999).
8. S. R. Hintz, D. A. Benaron, A. M. Siegel, A. Zourabian, D. K. Stevenson and D. A. Boas, “Bedside functional imaging of the premature infant brain during passive motor activation”, J. Perinat. Med. **29** 335–343 (2001).
9. X. Cheng and D. A. Boas, “Systematic diffuse optical image errors resulting from uncertainty in the background optical properties”, Opt. Express **4**, 299–307 (1999).
10. A. Liebert, H. Wabnitz, J. Steinbrink, H. Obrig, M. Möller, R. Macdonald, H. Rinneberg, “Intra- and extracerebral changes of hemoglobin concentration by analysis of moments of distributions of times of flight of photons”, *this issue*.
11. E. Okada, “Effect of thickness of the superficial tissues on the sensitivity profile for optical imaging”, *this issue*.
12. S. R. Arridge, “Optical tomography in medical imaging”, Inverse Problems **15** R41–R93 (1999).
13. A. M. Dale, B. Fischl, and M. I. Sereno, “Cortical surface-based analysis I. Segmentation and surface reconstruction”, NeuroImage **9**, 179–194 (1999).
14. E. Okada and D. T. Delpy, “Near-infrared light propagation in an adult head model. I. Modeling of low-level scattering in the cerebrospinal fluid layer”, Appl. Opt. **42**, 2906–14 (2003).
15. F. Bevilacqua, D. Piguet, P. Marquet, J. D. Gross, B. J. Tromberg, C. Depeursinge, “In vivo local determination of tissue optical properties: applications to human brain”, Appl. Opt. **38**, 4939–50 (1999).
16. A. H. Barnett, J. P. Culver, A. G. Sorensen, A. Dale, and D. A. Boas, “Robust inference of baseline optical properties of the human head with 3D segmentation from magnetic resonance imaging”, Appl. Opt. **42**, 3095–3108 (2003).
17. D. S. Sivia, *Data Analysis: A Bayesian Tutorial*, Oxford, 1996.
18. S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, “Source-detector calibration in three-dimensional Bayesian optical diffusion tomography”, J. Opt. Soc. Am. A, to appear.
19. Neal, R. M., *Probabilistic inference using Markov chain Monte Carlo methods*, Technical Report CRG-TR-93-1, Dept. of Computer Science, University of Toronto (1993). Available at http://www.cs.toronto.edu/∼radford/review.abstract.html
20. Anna Custo, Masters thesis, MIT, *expected* Sept 2003.