LA-UR- 99-554

Title: BAYESIAN INFERENCE FOR NEURAL ELECTROMAGNETIC SOURCE LOCALIZATION: ANALYSIS OF MEG VISUAL EVOKED ACTIVITY

Author(s): David M. Schmidt, John S. George, Charles C. Wood, P-21 Biophysics Group

Submitted to: Proceedings of SPIE Medical Imaging Conference, San Diego, CA, February 20-26, 1999.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

LA-UR- 99-554

Bayesian Inference for Neural Electromagnetic Source Localization: Analysis of MEG Visual Evoked Activity

D.M. Schmidt, J.S. George, and C.C. Wood

Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

ABSTRACT

We have developed a Bayesian approach to the analysis of neural electromagnetic (MEG/EEG) data that can incorporate or fuse information from other imaging modalities and addresses the ill-posed inverse problem by sampling the many different solutions which could have produced the given data. From these samples one can draw probabilistic inferences about regions of activation. Our source model assumes a variable number of variable size cortical regions of stimulus-correlated activity. An active region consists of locations on the cortical surf ace, within a sphere centered on some location in cortex. The number and radii of active regions can vary to defined maximum values. The goal of the analysis is to determine the posterior probability distribution for the set of parameters that govern the number, location, and extent of active regions. Markov Chain Monte Carlo is used to generate a large sample of sets of parameters distributed according to the posterior distribution. This sample is representative of the many different source distributions that could account for given data, and allow s identification of probable (i.e. consistent) features across solutions. Examples of the use of this analysis technique with both simulated and empirical MEG data are presented.

Keywords: MEG, Inverse Problem, MCMC

1. INTRODUCTION

Under suitable conditions of spatial and temporal synchronization, neuronal currents are accompanied by electric potentials and magnetic fields that are sufficiently large to be recorded non-invasively from the surface of the head. These are known as the electroencephalogram (EEG) and magnetoencephalogram (MEG), respectively. In contrast to PET and fMRI, which measure cerebral vascular changes secondary to changes in neuronal activity, EEG and MEG are direct physical consequences of neuronal currents and are capable of resolving temporal patterns of neural activity in the millisecond range.[1–3] Unlike PET and fMRI, however, the problem of estimating the current distribution in the brain from surface EEG and MEG measurements (the so-called electromagnetic inverse problem) is mathematically ill-posed; that is, it has no unique solution in the most general, unconstrained case.[4,5]

Existing approaches to the electromagnetic inverse problem fall into two broad categories: (1) "few-parameter models" (i.e., those in which $M \ll N$, where M is the number of parameters to be estimated in the model and N is the number of recording sites); (2) and "many-parameter models" (i.e., those in which $M \geq N$). A well-known example of the "few parameter" approach is the single- or multiple-dipole model,[6–8] in which the current is assumed to be represented by a few point-dipoles, the "order" of the model is estimated using Chi-square or related statistical techniques, and the best-fitting values of the dipole parameters (locations, orientations, and magnitudes) are estimated using non-linear numerical minimization techniques. A well-known example of the "many-parameter" approach is the "minimum-norm linear inverse",[9,10,1] in which the problem is under-determined (because $M \geq N$) and a strictly mathematical criterion is used to select among the many solutions that fit the data equally well; in the case of the minimum-norm approach the mathematical criterion is the solution that minimizes the sum of squared current strengths.

We have developed a new probabilistic approach to the electromagnetic inverse problem,[11] based on Bayesian inference.[12,13] Unlike other approaches to this problem, including other recent applications of Bayesian methods,[14,15] our approach does not result in a single "best" solution to the problem. Rather, we estimate a probability distribution of solutions upon which all subsequent inferences are based. This distribution provides a means of identifying and estimating the likelihood of features of current sources from surface measurements that explicitly emphasizes the multiple solutions that can account for any set of surface EEG/MEG measurements.

Correspondence should be addressed to D.M. Schmidt, Email: DSchmidt@LANL.GOV

In addition to emphasizing the inherent probabilistic character of the electromagnetic inverse problem, Bayesian methods provide a formal, quantitative means of incorporating additional relevant information, independent of the EEG/MEG measurements themselves, into the resulting probability distribution of inverse solutions. Such information can include constraints derived from anatomy on the likely location and/or orientation of current,[16,17,14,18] maximum current strength, spatial and/or temporal smoothness of current, etc.

2. BAYESIAN INFERENCE

Bayesian inference (BI) is a general procedure for constructing a (posterior) probability distribution for quantities of interest from the measurements given (prior) probability distributions for all of the uncertain parameters—both those that relate the quantities of interest to the measurements and the quantities of interest themselves. The method is conceptually simple, using basic laws of probability, making its application even to complicated problems relatively straightforward. The posterior probability distribution is often too complicated to be calculated analytically, but can usually be adequately sampled using modern computer techniques, even in problems with many parameters. The method is outlined here, more detailed presentations can be found elsewhere.[13]

The starting point for Bayesian inference is Bayes' rule of probability:

$$P(\theta, y) = P(\theta \mid y)P(y), \tag{1}$$

where $P(\theta, y)$ is the joint probability distribution for the quantities, θ and y, $P(\theta \mid y)$ is the conditional probability distribution of θ given y, and $P(y)$ is the marginalized probability distribution of y; $P(y) = \sum_\theta P(\theta, y)$ (or $P(y) = \int P(\theta, y)\, d\theta$ for continuous θ). If θ represents parameters about which we wish to learn and y represents data bearing upon θ, then the probability of θ given y can be constructed from Bayes' rule as:

$$P(\theta \mid y) = \frac{P(\theta, y)}{P(y)} = \frac{P(y \mid \theta)P(\theta)}{P(y)}. \tag{2}$$

Here $P(\theta)$ is the prior probability distribution of θ which represents one's knowledge of θ prior to the measurement. This is modified by the data through the likelihood function, $P(y \mid \theta)$, to produce the posterior probability distribution, $P(\theta \mid y)$. Since $P(y)$ is independent of θ it can be considered a normalizing constant and can be omitted from the unnormalized posterior density:

$$P(\theta \mid y) \propto P(y \mid \theta)P(\theta). \tag{3}$$

As summarized in,[13] "These simple expressions encapsulate the technical core of Bayesian inference: the primary task of any specific application is to develop the model $P(\theta, y)$ and perform the necessary computations to summarize $P(\theta \mid y)$ in appropriate ways."

3. BAYESIAN INFERENCE APPLIED TO THE EEG/MEG INVERSE PROBLEM

3.1. Activity Model

In applying the methods of Bayesian inference to the EEG/MEG inverse problem we constructed a model for regions of activation which is intended to be applicable in evoked response experiments. There is both theoretical and experimental evidence that EEG and MEG recorded outside the head arise primarily from neocortex, in particular from apical dendrites of pyramidal cells.[19,10,1] We therefore constructed a model that assumes a variable number of variable size cortical regions of stimulus-correlated activity in which current may be present. Specifically, an active region is assumed to consist of those locations which are identified as being part of cortex and are located within a sphere of some radius r centered on some location w, also in cortex. There can be any number n of these active regions up to some maximum n_{max} and the radius can have any value up to some maximum, r_{max}. The goal, in our approach, is to determine the posterior probability values for the set of activity parameters $\alpha = \{n, w, r\}$ which govern the number, location, and extent of active regions.

3.2. Probability Model for Activity Parameters

The first step in Bayesian inference is to construct a probability model that relates the activity parameters to the measurements. The probability model we constructed is outlined here, a more detailed description may be found elsewhere.[11] Let the N measurements at one instant in time be denoted by $\mathbf{b} = \{b_1, \ldots, b_N\}$. The conditional probability of the activity parameters given the observed data, $P(\alpha \mid \mathbf{b})$, can be expressed using Bayes' rule of probability as

$$P(\alpha \mid \mathbf{b}) \propto P(\mathbf{b} \mid \alpha) P(\alpha) \tag{4}$$

where $P(\alpha)$ is the prior probability for the activity parameters and $P(\mathbf{b} \mid \alpha)$ is the probability of the data given a particular set of values for the activity parameters. The prior probability for the activity parameters will be set by the experimenter using physiological information about the particular experiment being analyzed. Because the data do not depend on the activity parameters directly, but rather on a given current distribution, the function $P(\mathbf{b} \mid \alpha)$ needs to be expanded to reflect this fact. This may be accomplished by marginalizing out the current in the joint probability of the data and current such that

$$\begin{aligned} P(\mathbf{b} \mid \alpha) &= \int P(\mathbf{b}, \mathbf{j} \mid \alpha)\, \mathcal{D}\mathbf{j} \\ &= \int P(\mathbf{b} \mid \mathbf{j}, \alpha) P(\mathbf{j} \mid \alpha)\, \mathcal{D}\mathbf{j} \end{aligned} \tag{5}$$

where the integral is a functional integral over all current distributions. This integral is present because there are many different current distributions that could be present within a given set of active regions.

The function, $P(\mathbf{b} \mid \mathbf{j}, \alpha)$, is the likelihood function of the data. It is taken to be a Gaussian whose covariance reflects the uncertainty in the data and whose mean incorporates the forward model, which models what would be measured if the given current distribution were present. Specifically,

$$P(\mathbf{b} \mid \mathbf{j}, \alpha) \propto \exp\left\{ -\frac{1}{2} \sum_{k=1}^{N} \left(\tilde{b}_k - \langle \tilde{\mathbf{a}}_k, \mathbf{j} \rangle \right)^2 \right\}. \tag{6}$$

where $\mathbf{a}$ are the forward fields or measurement kernel such that if there were no noise or background the measurements would be related to the current by the inner product:

$$b_k = \langle \mathbf{a}_k, \mathbf{j} \rangle = \int \mathbf{a}_k(\mathbf{x}) \cdot \mathbf{j}(\mathbf{x})\, d^3x. \tag{7}$$

The symbols $\tilde{b}$ and $\tilde{\mathbf{a}}$ in Eq. 6 indicate that any noise covariance has been absorbed into these measurement and forward field vectors.

The function $P(\mathbf{j} \mid \alpha)$, which gives the probability of any current given a particular set of activity parameters, was constructed to incorporate a good deal of prior information. We used a Gaussian distribution such that

$$P(\mathbf{j} \mid \alpha) \propto |\mathbf{V}_\alpha|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \langle \mathbf{j}, \mathbf{V}_\alpha^{-1} \mathbf{j} \rangle \right\} \tag{8}$$

where $\mathbf{V}_\alpha^{-1}$ is the inverse of the covariance operator (matrix) of the current. The diagonal elements, or the variances, serve to limit the current strength, and the off-diagonal elements, which are related to the correlation coefficients, can serve to restrict the smoothness and orientation of the current distribution. The variance at locations which are not part of any active region for a given α is set to zero. The experimenter needs to set the values of the covariance matrix, based on knowledge of the experiment to be analyzed, using prior information about the strength, orientation and spatial variability of current within active regions.

The full probability model for the activity parameters is

$$P(\alpha \mid \mathbf{b}) \propto P(\alpha) |\mathbf{V}_\alpha|^{-\frac{1}{2}} \int \exp\left[-\frac{1}{2} \left\{ \sum_{k=1}^{N} \left(\tilde{b}_k - \langle \tilde{\mathbf{a}}_k, \mathbf{j} \rangle \right)^2 + \langle \mathbf{j}, \mathbf{V}_\alpha^{-1} \mathbf{j} \rangle \right\} \right] \mathcal{D}\mathbf{j} \tag{9}$$

Figure 1. Gray matter regions are tagged from anatomical MRI data. These tagged voxels constitute the anatomical model used to implement the cortical location and orientation prior information.

and $P(\alpha)$ is set by the experimenter. In practice, the integral over the current is constructed using the eigenvalues, $\{\lambda_\theta(\alpha)\}$, and normalized eigenvectors, $\{\psi_\theta(\alpha)\}$ $(\theta = 1, \ldots, N)$, of the matrix $G_{k,l}(\alpha) = \langle \tilde{\mathbf{a}}_k, \mathbf{V}_\alpha \tilde{\mathbf{a}}_l \rangle$, using standard numerical techniques.

Using these eigenvalues and eigenvectors the formula for the posterior probability distribution becomes

$$P(\alpha \mid \mathbf{b}) \propto P(\alpha) \exp\left[-\frac{1}{2}\left\{\sum_{k,\theta,l} \tilde{b}_k \frac{\psi_{k,\theta}(\alpha)\psi_{l,\theta}(\alpha)}{1+\lambda_\theta(\alpha)} \tilde{b}_l + \sum_\theta \ln(1+\lambda_\theta(\alpha))\right\}\right]. \tag{10}$$

This formula is well-behaved and is not overly sensitive to very small eigenvalues. Moreover, it is relatively simple to compute because it only depends on the N by N matrix, $G_{k,l}(\alpha)$.

3.3. Sampling the Posterior

The next step in Bayesian inference is to use the posterior probability distribution in order to answer questions related to the activity parameters in terms of probability. Examples of such questions include: what is the probability that there were m regions of activity? What are the locations for these active regions at a 95% probability level? In cases where the number of different possible sets of activation parameters is small, one can evaluate the complete posterior distribution. Generally, however, the number of different possible sets of activation parameters is large. In such cases the method of Markov Chain Monte Carlo (MCMC) can be used to generate a sample of sets of activity parameters which are distributed according to the posterior distribution. This is known as sampling the posterior, the techniques for which are described in detail elsewhere.[13]

4. EXAMPLES

While the methods just described apply to models for both EEG and MEG data, in the remainder of this paper we will use MEG data to illustrate the properties of the approach. Both simulated and empirical MEG data for a Neuromag-122 whole-head system were used.[20] The physical setup of the actual MEG experiment was used to determine the location of the subject's head relative to the sensors in the simulated data examples. In addition, an anatomical MRI data set acquired from the subject in the MEG experiment was used to determine the location of cortex (actually gray matter) using MRIVIEW (Fig 1), a software tool developed in our laboratory.[21] About 50,000 voxels were tagged and the normal directions for each of these voxels was then determined by examining the curvature of the local tagged region.

A spherically symmetric conductivity model was used to calculate the expected measurements given a current source both for the simulated data and in the likelihood calculations.[22] The same prior assumptions were used with the simulated data sets, with only minor changes for the real data example. Specifically the prior probability function $P(\alpha)$ was uniform so that each set of activation parameters had the same prior probability. The number of

active regions was allowed to range from 0 to 8 and the radius of any region of activity was allowed to range from 0 to 10 mm.

The covariance matrix was factored such that

$$ {}^{\beta\gamma}V_{\alpha}(i,j) = \sigma_{\alpha}(i)\sigma_{\alpha}(j)\rho(i-j)\,{}^{\beta\gamma}\Omega(i,j) \tag{11} $$

where β and γ are orientation indices, i and j are location indices, $\sigma_{\alpha}(i)$ is the standard deviation at location i, $\rho(i-j)$ is the spatial correlation function, and ${}^{\beta\gamma}\Omega(i,j)$ is the orientation covariance. The correlation function was chosen to be a Gaussian with zero mean and 7 mm standard deviation which imposes spatial smoothness on scales of about 7 mm or less. Because of this prior information concerning spatial correlation, the continuous current distributions and integrals of the previous section may be well-approximated by discrete distributions and sums over the volume elements (voxels) that were tagged from the anatomical MRI data. For example, in evaluating the posterior probability value using Eq. 10 the matrix $\mathbf{G}$ is calculated in the following examples by approximating the continuous integral with a sum over tagged voxels. This is a good approximation because the covariance operator has a correlation length of 7 mm which is larger than the voxel dimensions of 2 mm on a side.

To complete the specification of the covariance operator, a value of 2 nAm was used for $\sigma_{\alpha}(i)$ at all locations in active regions and 0 nAm elsewhere. The orientation covariance was chosen such that there was no correlation between the orientations at different locations and the orientation distribution at any given location was symmetric with respect to the direction normal to the cortical surface at that location and had a mean equal to the cortical norm direction and a standard deviation of 30°. Unlike other recent implementations of cortical constraints in distributed inverse solutions,[10,14] this procedure results in a distribution of orientations around the perpendicular, not a fixed normal orientation.

Specifically, the orientation covariance was constructed as follows:

$$ {}^{\alpha\beta}\Omega(i,j) = \begin{cases} \frac{S+1}{S+2}\varepsilon^{\alpha}(i)\varepsilon^{\beta}(j) & \text{if } i \neq j, \\ {}^{\alpha\beta}E(i;S) & \text{if } i = j. \end{cases} \tag{12} $$

Here $\varepsilon^{\alpha}(i)$ is the unit vector normal to cortex at voxel i and S is a parameter controlling how tight the prior distribution is about the normals to cortex. The $\mathbf{E(i;S)}$ matrix is constructed from

$$ \mathbf{E(i;S)} = \mathbf{R^{T}(i)E_{n}(S)R(i)} \tag{13} $$

which is a rotation of the matrix

$$ \mathbf{E_{n}(S)} = \begin{pmatrix} \frac{1}{2}(1-\frac{S+1}{S+3}) & 0 & 0 \\ 0 & \frac{1}{2}(1-\frac{S+1}{S+3}) & 0 \\ 0 & 0 & \frac{S+1}{S+3} \end{pmatrix} \tag{14} $$

using the standard rotation matrix $\mathbf{R(i)}$ that rotates from a coordinate system in which positive z is along the cortical norm direction at voxel i (i.e. along $\varepsilon(i)$) to the standard, global $\{x,y,z\}$ coordinate system being used. This formulation was derived from calculating the covariance of unit norm orientation vectors using the distribution

$$ P(\theta) = \frac{S+1}{2\pi}\cos^{S}(\theta) \tag{15} $$

where θ is the angle from the cortical norm direction at each voxel. From the form of this distribution, $S = 0$ yields a uniform distribution (no preference towards the cortical norm), while the larger S is the more restrictive is the distribution towards the cortical norm direction. A value for $S = 40$ was used for the following examples.

Finally, the same noise was added to all simulated data sets, which was Normal with a standard deviation of 10 fT. The values used here in the prior probability distribution are meant to be an example of what one might choose for a MEG analysis and should be chosen for each particular MEG data set.

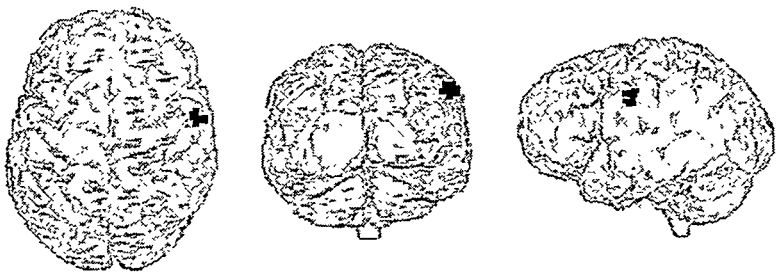

Figure 2. Maximum intensity projection of the location and extent (in black) of the active region used to generate simulated MEG data for Example 1.

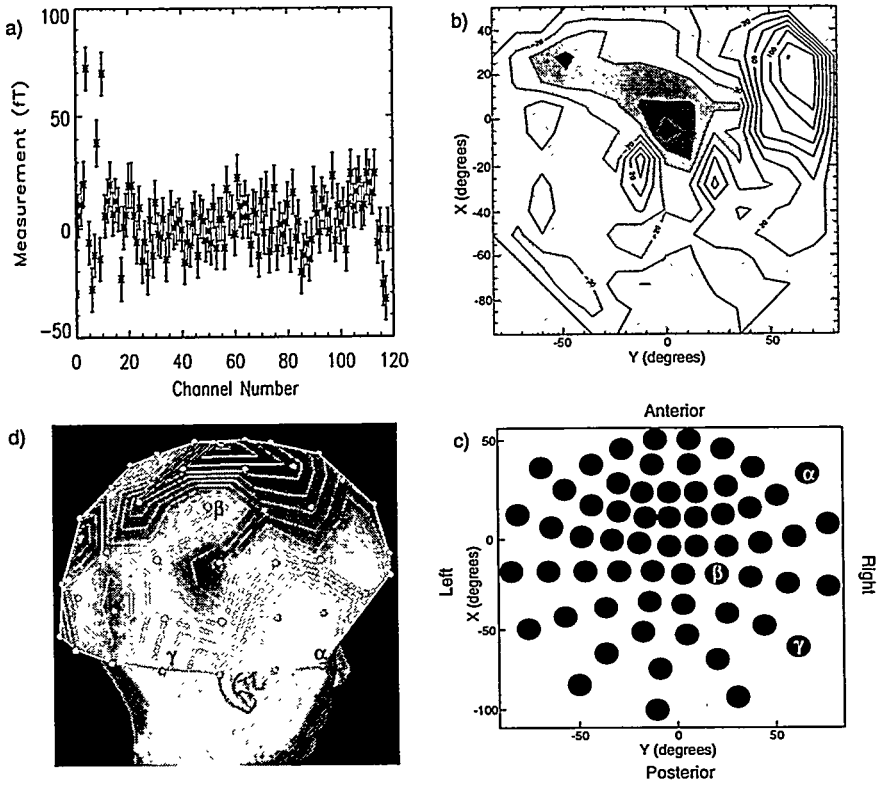

Figure 3. The simulated data used in Example 1 as a function of channel number, a). To aid in visualizing the field pattern associated with this data panels b) and d) display iso-amplitude contours of a field on the sensor helmet surface that is consistent with the data shown in a). Positive fields (flux emerging from the head) are displayed in light shades and negative fields (reentering flux) are shown in dark shades. b) is a polar projection, viewing the helmet surface from above, for which the relative position of the sensors and orientation of the subject's head are shown in c). The sensors labeled α, β and γ in c) correspond to those with the same labels in d).

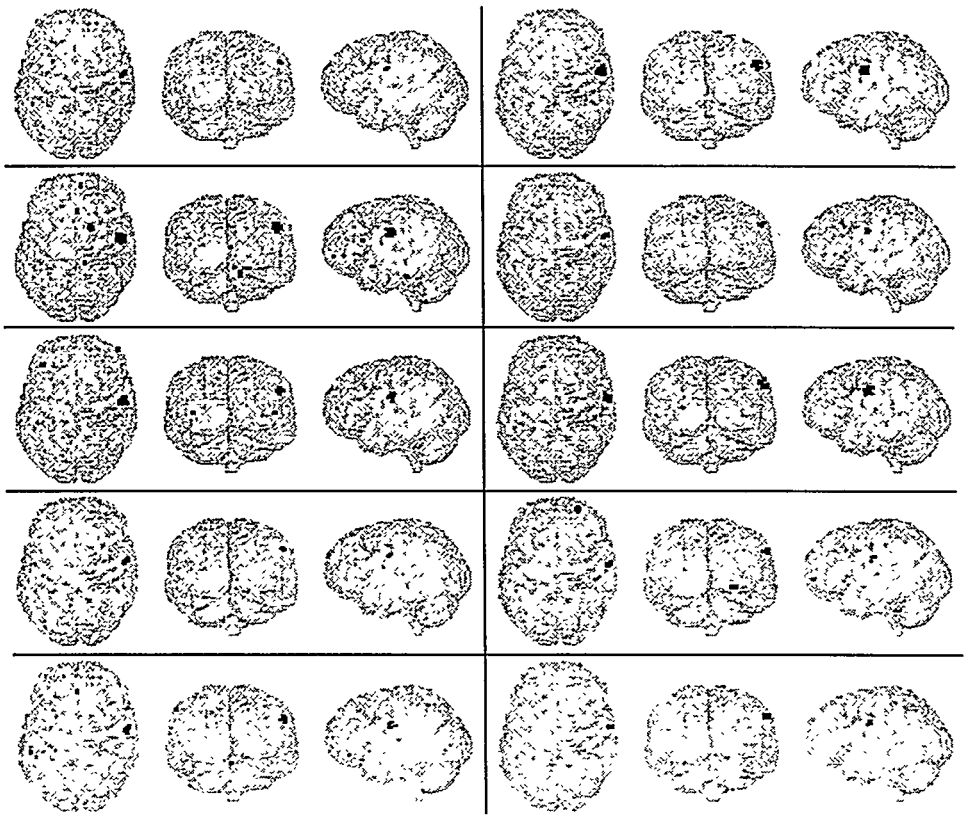

Figure 4. A few of the 9,000 samples drawn from the posterior probability distribution of Example 1. Each panel shows 3 views of the maximum intensity projection of all of the active regions from a single sample. All of these samples could have produced the same MEG data set.

4.1. Example 1

The location and extent of the active region used to generate the simulated MEG data is shown in Fig. 2. The bounding radius of the active region was 5 mm and the current dipole strength at each voxel was 2 nAm oriented in the cortical normal direction. A plot of the simulated data and some views of its field pattern are shown in Fig. 3. Ten thousand samples were drawn from the posterior distribution using a MCMC algorithm. The program ran overnight, not having been optimized, on a current personal computer. It took about 600 samples to progress from the starting point which had a low probability to one that had a high probability and was therefore representative of the posterior distribution. Only the final 9,000 samples, a few of which are shown in Fig. 4, were used in making probabilistic inferences as discussed below. All of the samples shown in Fig. 4 are among the 95% most probable and therefore fit both the data and the prior expectations quite well. Any of these could have produced the given MEG data, yet there are clearly vast differences among the samples. The number of active regions ranges from 1 to 5, the sizes of the regions vary greatly and the locations of the active regions vary nearly across the entire tagged region of the brain (when considering all 9,000 samples). This variability is a representation of the degree of the ambiguity of the inverse problem for these MEG data, even with the prior information present.

Despite the degree of variability among the samples in Fig. 4 a property common to all is apparent; namely an active region in the dorsal, lateral region of the right hemisphere. A feature, such as this, common to all or most of the samples, is associated with a high degree of probability. This probability can be quantified because the MCMC samples are distributed according to the posterior probability distribution. The smallest set of voxels which contains the center of the active region in the dorsal, lateral region in 95% of the samples was identified and is shown in Fig. 5. This region, which contains a center of activity with a probability of 95%, in fact encompasses the region of activity which was used to produce the simulated data set (Fig. 2). Although it is nice to see this agreement, it is not

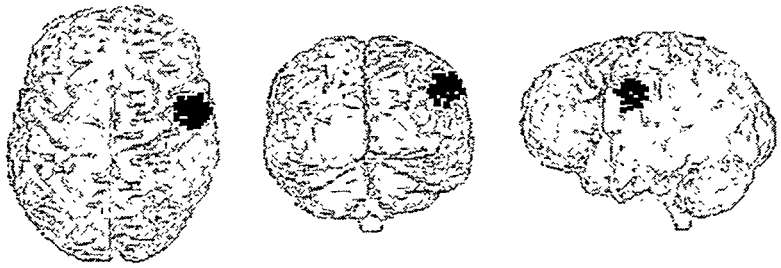

Figure 5. Maximum intensity projections of the location and extent of a region containing a center of activity at a 95% probability level in Example 1.

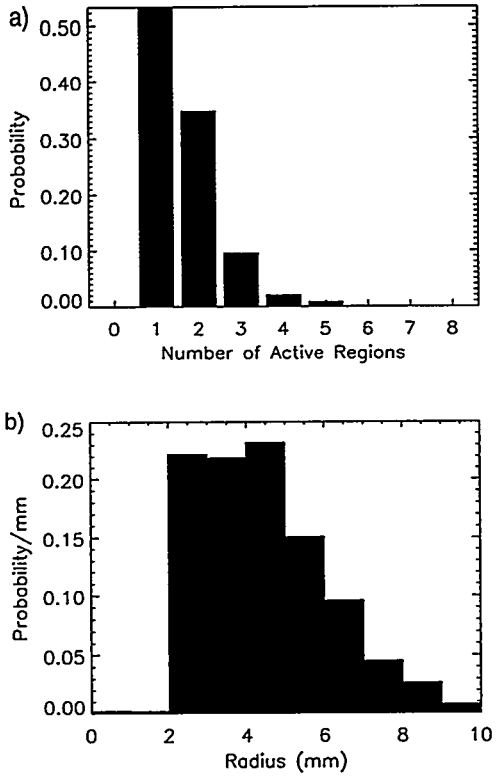

Figure 6. The posterior probability for a) the number of active regions present in Example 1 and b) the radius of the sphere bounding activity in Example 1, assuming there was only one active region present..

sufficient to justify this or any MEG inverse method based solely on whether it produces results consistent with the true active regions because any of the sets of active regions shown in Fig. 4 could have also been used to generate the same MEG data. Any robust and highly probable result or inference therefore should be consistent with the wide range of possible sets of active regions, as is the result in Fig. 5 by construction. This is a very important feature of BI which is necessarily missing from any other analysis method that only considers just one possible result, even if it happens to be the most likely result within a given model.

In addition to the information about the locations of probable regions of activity, the Bayesian approach combined with this activity model also provides probabilistic information about the number and size of active regions. The posterior distribution for the number of active regions was constructed by histograming the number of regions across the MCMC samples. This histogram is shown in Fig. 6a. One active region is the most probable; however, two active regions are quite likely as well. Although the location of one active region was identified in the last paragraph, the

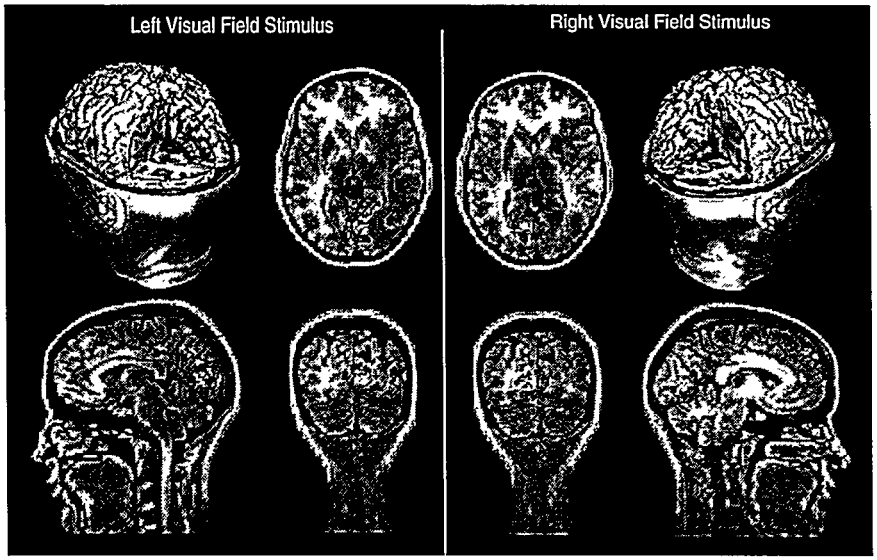

Figure 7. Four views of a region that was found to contain activity at a 95% probability level for both a left and a right visual field stimulus, at 110 ms latency. The two-dimensional views show the probable regions of activity within the anatomical MRI data. The horizontal and coronal views are from the top and from the back of the subject, respectively; the sagital views are from the left for the left visual field stimulus panel and from the right for the right visual field stimulus panel. The three-dimensional views are useful for showing the location of the regions relative to other brain structures. These results indicate that the probability of activity is maximal in the calcarine region of the hemisphere contralateral to the visual field stimulated.

location of a second could not be well localized because it occurred in a wide range of locations across the MCMC samples. Assuming there was only one active region present, we can obtain information about its size by histograming the size of active regions in the MCMC samples that had only one region present. This histogram is shown in Fig. 6b and represents the posterior probability for the bounding radius of the active region, assuming that there was only one region active. Regions smaller than 2 mm and larger than 8 mm in radius are very unlikely whereas regions that are around 5 mm in radius are likely. The size of the region used to produce the simulated data was 5 mm. We believe that much of the information on size derives from prior information about location, orientation and strength of neural current.

Other inferences could be drawn using the MCMC samples in a similar manner. For example, one could construct the probability for the size of the active region, assuming there was one centered within the 95% probability region shown in Fig. 5, rather than assuming there was only one active region present throughout the entire head as was done above.

4.2. Example 2

In order to examine the sensitivity of the Bayesian approach to detect known features of human visual cortex organization, we compared Bayesian analyses of MEG responses to visual stimuli in the left and right visual fields.[23] Based on the crossed anatomical projections of the visual fields to the brain and on previous lesion, MEG, and fMRI studies in humans,[24–26] initial cortical activation for stimuli in the left and right visual fields should occur near the calcarine fissure in the occipital region of the contralateral hemispheres.

The MEG data were acquired from a Neuromag-122 whole-head system.[20] An anatomical MRI data set acquired from the subject in the MEG experiment was used to determine the location of cortex (actually gray matter) using MRIVIEW (Fig 1), a software tool developed in our laboratory.[21] About 50,000 voxels were tagged and the normal directions for each of these voxels was then determined by examining the curvature of the local tagged region.

Our Bayesian inference analysis was applied separately to the data for each visual field stimulus at 110 ms post-stimulus latency; a latency that should include robust activation of the calcarine region.[26] Ten thousand samples

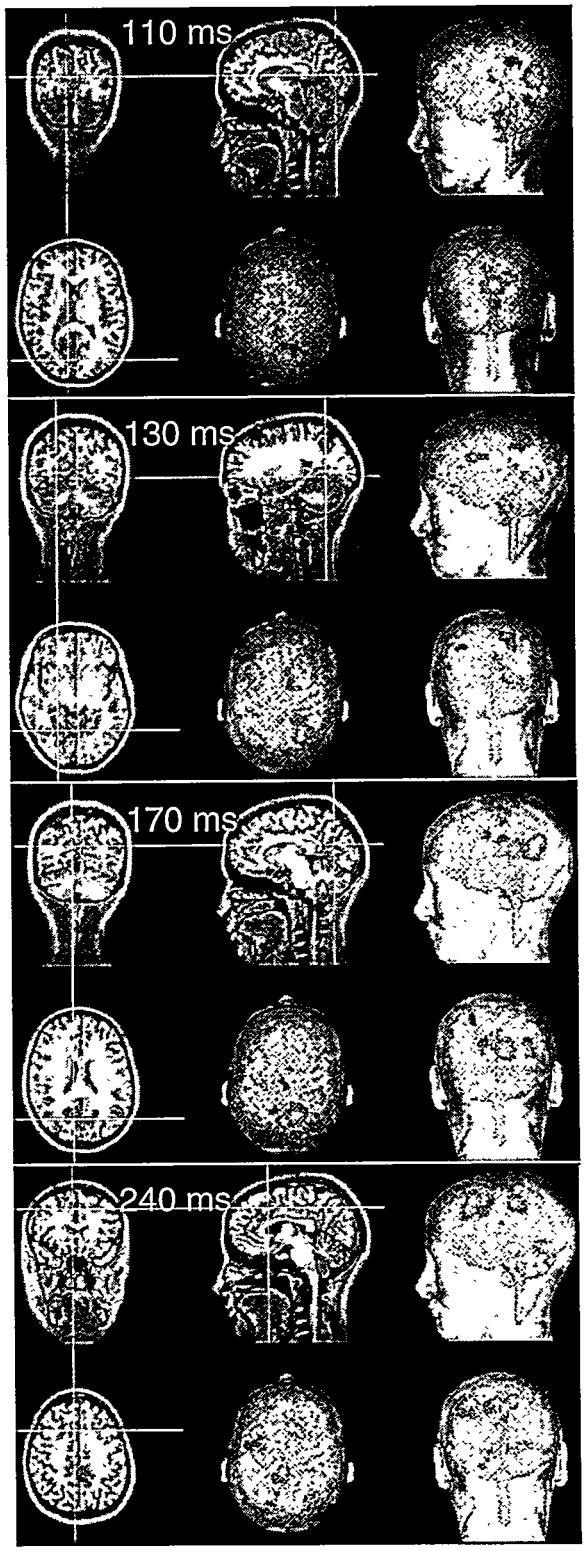

Figure 8. Results from four latencies of the left visual field data.

of the posterior probability were generated for each data set. The smallest sets of voxels which contained the center of the active region in the occipital region in 95% of the samples was identified for both data sets and are shown in Fig. 7. This figure depicts relative probability of activation within these regions on a color scale in three orthogonal slices through the calcarine region and a three-dimensional rendering of the occipital region. For the left visual field stimulus, maximal probability of activation at 110 ms was located in the right (contralateral) hemisphere, centered upon the calcarine region. This pattern was reversed for the right visual field stimulus at 110 ms, consistent with the predictions from anatomy, and from the lesion, fMRI, and previous MEG studies cited above.

Two additional features of the results should be noted. First, although maximal probability of activation at the 110 ms latency was indeed located in the opposite hemisphere, there exists sizable probability for activity in the ipsilateral hemisphere near the mid-line. The extent of the 95% probability regions shown in Fig. 7 is indicative of both the extent of estimated activation and the degree of error or uncertainty in that estimate even allowing for the possibility of different numbers of active regions of variable extent.

Second, as shown in Fig. 8, analyses at other latencies suggest a progressively increasing number of probable regions of activation, in both the ipsilateral and contralateral hemispheres, over the latency region from 110 to 240 ms following stimulus onset. There was also evidence for the contralateral calcarine region reactivating near 170 *ms*. It will be of considerable interest to explore the time dependence of the Bayesian inference analyses in relation to evidence for multiple, functionally organized areas of striate and extra-striate visual cortex and to examine the value of temporal prior information (not included in the current activation model) in the form of, for example, temporal covariance constraints. In addition, the Bayesian approach provides a natural means for incorporating information from other functional imaging modalities such as PET or fMRI.[17,27,18] The latter can be readily achieved with the Bayesian framework and with this activity model by assigning prior probabilities to possible locations of active regions based on results from the other modality or modalities. Such a Bayesian formulation of multi-modality integration would yield an inherently probabilistic result in which the quantity estimated would be the probability of activation as a function of both space and time.

REFERENCES

1. M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Mod. Phys.* **65**, pp. 413–497, apr 1993.
2. C. J. Aine, "A conceptual overview and critique of functional neuroimaging techniques in humans: I. MRI/fMRI and PET," *Critical Reviews in Neurobiology* **9**, pp. 229–309, 1995.
3. A. W. Toga and J. C. Mazziotta, *Brain Mapping: The Methods*, Academic Press, New York, 1996.
4. H. von Helmholtz, "Ueber einige gesetze der vertheilung elektrischer strome in korperlichen leitern, mit anwendung auf die thierisch-elektrischen versuche," *Ann. Phys. Chem* **89**, pp. 211–233 and 353–377, 1853.
5. P. Nunez, *Electrical fields of the brain: neurophysics of the EEG*, Oxford University Press, Oxford, 1981.
6. R. N. Kavanaugh, T. M. Darcey, D. Lehmanh, and D. H. Fender, "Evaluation of methods for three-dimensional localization of electrical sources in the human brain," *IEEE Trans. Biomed. Eng.* **25**, pp. 421–429, 1978.
7. M. Scherg and D. von Cramon, "Evoked dipole source potentials of the human auditory cortex," *Electroenceph. Clin. Neurophysiol.* **65**, pp. 344–360, 1986.
8. J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomed. Eng.* **39**(6), pp. 541–557, 1992.
9. M. Hämäläinen and R. Ilmoniemi, "Interpreting measured magnetic fields of the brain: estimates of current distributions," Tech. Rep. TKK-F-A559, Helsinki University of Technology, Finland, 1984.
10. A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach," *J. Cognitive Neurosci.* **5**, pp. 162–176, 1993.
11. D. M. Schmidt, J. S. George, and C. C. Wood, "Bayesian inference applied to the electromagnet inverse problem (To be published in Human Brain Mapping)," Tech. Rep. LA-UR-97-4813, Los Alamos National Laboratory, 1997.
12. J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, Wiley, New York, 1994.
13. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, 1995.

14. S. Baillet and L. Garnero, "A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem," *IEEE Trans. Biomed. Eng.* **44**, pp. 374–385, may 1997.
15. J. W. Phillips, R. M. Leahy, and J. C. Mosher, "MEG-based imaging of focal neuronal current sources," *IEEE Trans. Med. Imag.* **16**, pp. 338–348, jun 1997.
16. J. Z. Wang, S. J. Williamson, and L. Kaufman, "Magnetic source images determined by a lead-field analysis: the unique minimum-norm," *IEEE Trans. Biomed. Eng.* **39**, pp. 665–675, jul 1992.
17. J. S. George, C. J. Aine, J. C. Mosher, D. M. Schmidt, D. M. Ranken, H. A. Schlitt, C. C. Wood, J. D. Lewine, J. A. Sanders, and J. W. Belliveau, "Mapping function in the human brain with MEG, anatomical MRI, and functional MRI," *J. Clin. Neurophys.* **12**, pp. 406–431, 1995.
18. A. M. Dale, "Strategies and limitations in integrating brain imaging and electromagnetic recording," *Soc. Neurosci. Abst.* **23**, p. 1, 1997.
19. T. Allison, C. C. Wood, and G. McCarthy, "The central nervous system," in *Psychophysiology: Systems, Processes and Applications: A Handbook*, E. Donchin, S. Porges, and M. Coles, eds., pp. 5–25, Guilford Press, New York, 1986.
20. A. I. Ahonen, M. S. Hämäläinen, M. J. Kajola, J. E. T. Knuutila, P. P. Laine, O. V. Lounasmaa, L. T. Parkkonen, J. T. Simola, and C. D. Tesche, "122-channel SQUID instrument for investigating the magnetic signals from the human brain," *Physica Scripta* **T49A**, pp. 198–205, 1993.
21. D. M. Ranken and J. S. George, "MRIVIEW: An interactive computational tool for investigation of brain structure and function," in *Visualization '93*, pp. 324–331, IEEE Computer Society, 1993.
22. J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Phys. Med. Biol.* **32**(1), pp. 11–22, 1987.
23. C. Aine, H.-W. Chen, D. Ranken, J. Mosher, E. Best, J. George, J. Lewine, and K. Paulson, "An examination of chromatic/achromatic stimuli presented to central/peripheral visual fields: An MEG study," *Neuroimage* **5**, p. 153, 1997.
24. J. C. Horton and W. F. Hoyt, "The representation of the visual field in human striate cortex," *Proc. Roy. Soc. London* **132**, pp. 348–361, 1991.
25. M. I. Sereno, A. M. Dale, J. B. Reppas, K. K. Kwong, J. W. Belliveau, T. J. Brady, B. R. Rosen, and R. B. H. Tootell, "Borders of multiple visual areas in human revealed by functional magnetic resonance imaging," *Science* **268**, pp. 889–893, 1995.
26. C. Aine, S. Supek, J. George, D. Ranken, J. Lewine, J. Sanders, E. Best, W. Tiee, E. Flynn, and C. Wood, "Retinotopic organization of human visual cortex: Departures from the classical model," *Cerebral Cortex* **6**, pp. 354–361, 1996.
27. J. Belliveau, "Dynamic human brain mapping using combined fMRI, EEG and MEG," in *A presentation to the symposium on approaches to cognitive neuroscience by means of functional brain imaging*, (Caen, France), jun 1997.