

Title: MEG-BASED IMAGING OF FOCAL NEURONAL CURRENT SOURCES

RECEIVED  
JUL 19 1996  
OSTI

Author(s): J. W. Phillips, R. M. Leahy, J. C. Mosher

Submitted to: Biomagnetism '96, Santa Fe, NM, February 1996

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER



**Los Alamos**  
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

MEG-Based Imaging of Focal Neuronal Current Sources<sup>†</sup><sup>1</sup>Phillips, J.W., <sup>1</sup>Leahy, R.M., and <sup>2</sup>Mosher, J.C.<sup>1</sup>University of Southern California, Los Angeles, California, USA; <sup>2</sup>Los Alamos  
National Laboratory, Los Alamos, New Mexico, USA**Abstract**

We describe a new approach to imaging neuronal current sources from measurements of the magnetoencephalogram (MEG) associated with sensory, motor, or cognitive brain activation. Previous approaches to this problem have concentrated on the use of weighted minimum norm inverse methods. While these methods ensure a unique solution, they do not introduce information specific to the MEG inverse problem, often producing overly smoothed solutions and exhibiting severe sensitivity to noise. We describe a Bayesian formulation of the inverse problem in which a Gibbs prior is constructed to reflect the sparse focal nature of neuronal current sources associated with evoked response data. The prior involves a binary process indicating active sources and a continuous Gaussian process designating associated amplitudes. An estimate of the primary current source distribution for a specific data set is formed by maximizing over the posterior probability with respect to the binary and continuous variables.

**Introduction**

An array of SQUID biomagnetometers may be used to measure the spatio-temporal MEG signal produced by the brain. We wish to construct an image of neural activity which produced this signal. Physiological models for the MEG assume primary sources are constrained to the cortex with current flow oriented normal to the local surface [7]. The image can therefore be constrained to the cortical surface which can be extracted from a registered volume MR image of the subject's head. By tessellating the cortex with  $N$  disjoint regions and representing the sources in each region by an equivalent constrained current dipole oriented normal to the surface with amplitude  $y_i$ , the MEG inverse problem can be expressed in terms of a linear model. The linear forward model relating the  $N$  sources  $\mathbf{y}$  ( $N \times 1$ ) and the  $M$  MEG measurements  $\mathbf{b}$  ( $M \times 1$ ) can be written,

$$\mathbf{b} = \mathbf{G}\mathbf{y} + \mathbf{n} \quad (1)$$

where the  $i$ 'th row of the  $M \times N$  system matrix  $\mathbf{G}$  specifies a discrete representation of the lead field (sensitivity) of the  $i$ 'th sensor. The  $j$ 'th column of  $\mathbf{G}$  specifies the gain vector for the  $j$ 'th constrained dipole component. The  $M \times 1$  vector  $\mathbf{n}$  represents noise generated within the sensor and by unwanted electromagnetic sources (power lines, the heart, background brain activity, etc.).

The search for an appropriate imaging method is primarily concerned with finding a way to choose within a set of images that produces essentially the same fit to the data. Weighted minimum norm inverse methods [8,9] typically find solutions which match the data while minimizing a weighted  $l_2$ -norm on the solution vector. These techniques tend to smear sources over the entire reconstruction region and are generally unstable due to ill-conditioning of the system matrix. The instability can be overcome using Tikhonov regularization [6] but the reconstructions remain smeared. The iteratively reweighted minimum norm method [5] is a nonlinear approach to overcoming the problem of smeared sources in which the norm weighting is updated at each iteration based on the result of the previous iteration. This method uses a weighting matrix which, as the iterations proceed, reinforces strong sources and reduces weak ones. This results in very sparse solutions, but again the method is extremely sensitive to noise and highly dependent on the initial estimate.

Here we propose an alternative approach to the inverse problem based on a Bayesian formulation. Rather than use an arbitrary weighted  $l_2$ -norm to select the solution, we introduce a prior distribution on the source which is used to resolve the ambiguities inherent in the inverse problem. This prior is constructed to favor the reconstruction of physiologically plausible solutions. Basic studies of functional activation, such as somatotopic or retinotopic map-

<sup>†</sup>This work was supported by the National Institute of Mental Health Grant No. R01-MH53213, and National Eye Institute Grant No. R01-EY08610-04.

ping using fMRI or PET, reveal the sparse and highly localized nature of activation in the cerebral cortex [1]. Our prior is therefore specifically designed to reflect the expectation that current sources tend to be sparse and focal. This prior is combined with a Gaussian likelihood model for the data which is based on the linear model (1) and an assumption of additive white Gaussian noise. Maximization over the resulting posterior probability results in a maximum *a posteriori* (MAP) estimate of the primary current sources.

## Bayesian Technique Description

Since the primary sources of the MEG are widely accepted to be sparse and focal, we suggest that this information be used in reconstructing the image. This information can be naturally introduced into the problem using the Bayesian paradigm in which the source is modeled as a random field. Since we assume that sources are sparse, the large majority of source pixels will have zero amplitude. We therefore use a binary indicator process  $\mathbf{x}$  to model whether each source dipole is *on* ( $x_i=1$ ) or *off* ( $x_i=0$ ). Those sites that are active are assumed to have a Gaussian amplitude,  $z_i$ . We can then write the source image vector  $\mathbf{y}$  as  $\mathbf{y} = \mathbf{x}.*\mathbf{z} = \mathbf{X}\mathbf{z}$ , where  $.*$  signifies the Schur product (element by element pair-wise multiplication) and  $\mathbf{X} = \text{diag}(\mathbf{x})$ . Assuming independence of the indicator and amplitude processes, we can write the posterior probability for  $\mathbf{x}$  and  $\mathbf{z}$  given the MEG data  $\mathbf{b}$  as,

$$p(\mathbf{x}, \mathbf{z}|\mathbf{b}) = \frac{p(\mathbf{b}|\mathbf{x}, \mathbf{z})p(\mathbf{x})p(\mathbf{z})}{p(\mathbf{b})}. \quad (2)$$

We find a MAP estimate  $\mathbf{y}^*$  of  $\mathbf{y}$  as  $\mathbf{y}^* = \mathbf{x}^*.*\mathbf{z}^*$  where,

$$\mathbf{x}^*, \mathbf{z}^* = \arg \max_{\mathbf{x}, \mathbf{z}} p(\mathbf{x}, \mathbf{z}|\mathbf{b}). \quad (3)$$

The joint probability  $p(\mathbf{x})$  is chosen to reflect the expectation that the sources are sparse and focal. To achieve this goal, we use a Markov Random Field (MRF) [4] model for which sparse focal sources have a higher probability of occurring than more distributed sources. We define  $p(\mathbf{x})$  to be a Gibbs distribution,

$$p(\mathbf{x}) = (1/K) \exp \{-V(\mathbf{x})\} \quad (4)$$

where  $K$  is a proportionality constant and the energy function  $V(\mathbf{x})$  is given by,

$$V(\mathbf{x}) = \sum_i \left[ \underbrace{\alpha_i x_i}_{\text{Sparseness Term}} + \underbrace{\beta_i C_i\{x_i, x_j \mid j \in \xi_i\}}_{\text{Clustering Term}} \right] \quad (5)$$

where the parameters  $\alpha_i > 0$  and  $\beta_i > 0$  determine the relative weights of the sparseness and clustering terms. The potential function  $C_i\{x_i, x_j \mid j \in \xi_i\}$  is defined in terms of each pixel and its neighbors  $\xi_i$  as,

$$C_i\{x_i, x_j \mid j \in \xi_i\} = \left[ \sum_{j \in \xi_i} (x_i - x_j)^2 \right]^Q. \quad (6)$$

This clustering term is small if neighboring pixels are of the same magnitude. The exponential parameter  $Q$  determines the strength of the clustering. As  $Q$  increases, the size of the clusters tends to increase. In the results presented below we constrain sources to a plane and use a four nearest-neighbor interaction. Some examples of binary images, produced by sampling from this prior using a Gibbs sampler [4], are shown in Fig. 1. The source amplitude process,  $\mathbf{z}$ , is assumed to be a set of independent zero mean Gaussian random variables with covariance  $\mathbf{C}_z$ .

Using the definitions above and assuming the noise process in (1) is zero mean Gaussian with covariance  $\mathbf{C}_n$ , we can write,

$$p(\mathbf{x}, \mathbf{z}|\mathbf{b}) = \frac{1}{Z} \exp \{-U(\mathbf{x}, \mathbf{z}|\mathbf{b})\} \quad (7)$$

where  $Z$  is the posterior partition function. The posterior energy function is given by:

$$U(\mathbf{x}, \mathbf{z}|\mathbf{b}) = \frac{1}{2} [\mathbf{b} - \mathbf{G}\mathbf{X}\mathbf{z}]^T \mathbf{C}_n^{-1} [\mathbf{b} - \mathbf{G}\mathbf{X}\mathbf{z}] + \frac{1}{2} \mathbf{z}^T \mathbf{C}_z^{-1} \mathbf{z} + V(\mathbf{x}). \quad (8)$$

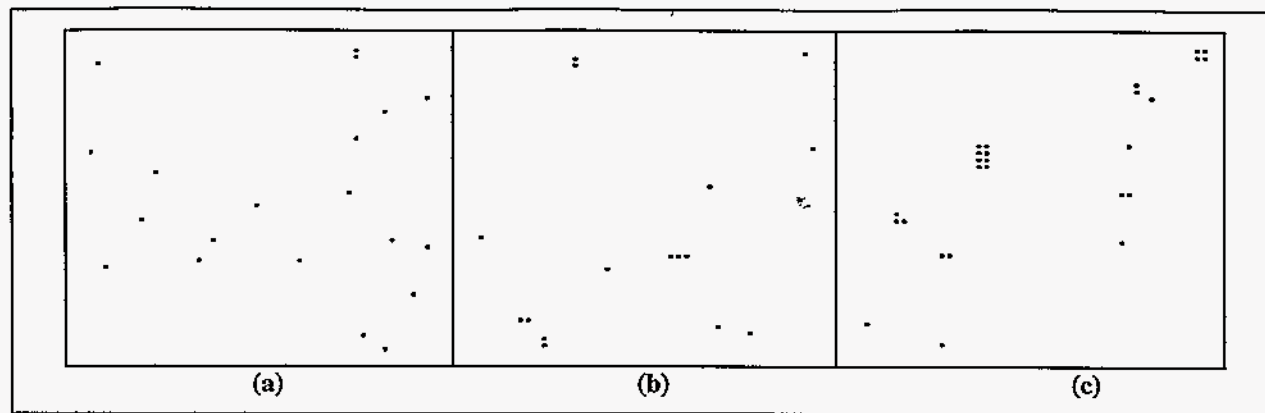


Fig. 1. Three examples of a random sampling from  $p(\mathbf{x})$ . (a)  $Q = 1$ ,  $\alpha = 1.0$ , and  $\beta = 0.4$  (b)  $Q = 2$ ,  $\alpha = 1.0$ , and  $\beta = 0.2$  (c)  $Q = 3$ ,  $\alpha = 1.0$ , and  $\beta = 0.1$ .

The MAP estimate is found by maximizing over the log-posterior, or equivalently,

$$\mathbf{x}^*, \mathbf{z}^* = \arg \min_{\mathbf{x}, \mathbf{z}} U(\mathbf{x}, \mathbf{z} | \mathbf{b}). \quad (9)$$

The solution to the optimization problem (9) provides our estimate of the neural current sources. The solution is clearly a function of the parameters of the likelihood function and the prior probabilities. Methods for selecting these parameters will be addressed in a future publication. However, we note that since the parameters  $\alpha_i$  represent the relative probabilities that each source pixel is active, it is straightforward to include pixel-wise probability weightings determined from fMRI or PET activation studies to influence the formation of these sparse images.

## Mean Field Annealing

Minimization of  $U(\mathbf{x}, \mathbf{z} | \mathbf{b})$  is difficult since the optimization must be performed over a mixture of discrete and continuous variables. Since this function is quadratic in the continuous variables,  $\mathbf{z}$ , we can derive a closed form expression for the optimal  $\mathbf{z}$  as a function of any particular indicator process  $\mathbf{x}$ , i.e. the vector  $\mathbf{z}^*(\mathbf{x})$  which minimizes (8) given the binary vector  $\mathbf{x}$  is given by,

$$\mathbf{z}^*(\mathbf{x}) = \mathbf{C}_z \mathbf{X} \mathbf{G}^T (\mathbf{G} \mathbf{X} \mathbf{C}_z \mathbf{X} \mathbf{G}^T + \mathbf{C}_n)^{-1} \mathbf{b}. \quad (10)$$

Substituting  $\mathbf{z}^*(\mathbf{x})$  into  $U(\mathbf{x}, \mathbf{z} | \mathbf{b})$  results in a new energy function  $U(\mathbf{x})$ , a function of  $\mathbf{x}$  only. We can therefore first find the optimal indicator process by minimizing  $U(\mathbf{x})$ , then substitute this result in (10) to obtain the optimal amplitude process.

Coordinate-wise optimization with respect to a collection of binary variables using, for example, iterated conditional modes (ICM) [2], tends to produce rapid convergence to an undesirable local minimum. Here we use, instead, a continuation method based on mean field annealing (MFA) [3]. We visit each pixel in turn using the following update strategy:

$$x_i^{(n+1)} = E_T \{ x_i | x_j^{(n)} \quad \forall j \neq i \} \quad (11)$$

where the conditional expectation is computed with respect to the modified joint probability,

$$p_T(\mathbf{x}) = \frac{1}{Z_T} \exp \left\{ -\frac{1}{T} U(\mathbf{x}) \right\}. \quad (12)$$

The temperature parameter  $T$  is slowly reduced as the iterations proceed. As  $T \rightarrow 0$  the iteration will converge to a binary solution which is a local minimum of  $U(\mathbf{x})$ . MFA typically results in a better local minimum than is achieved using ICM. We will address the convergence behavior and the basis for the mean field approach in a future extended publication.

## Simulations and Phantom Experiments

We have conducted extensive comparisons of this technique with minimum norm based methods using simulated and experimental phantom data. Some example images of representative comparisons are shown in our sister publication in this conference proceedings ("A comparative study of minimum norm inverse methods for MEG imaging," R.M. Leahy, J.C. Mosher, and J.W. Phillips). We found that all of the linear minimum norm methods produced results exhibiting a relatively large degree of smoothness. For all additive noise scenarios tested, all methods generally gave similar residual errors in the fit to the data. All results can be considered 'correct' in the sense that they are configurations that could have produced the observed data. This observation emphasizes the severely under-determined nature of this problem. Clearly, in order to select between these feasible solutions we must use additional information concerning the expected nature of the source.

## Conclusions

In the simulations and phantom studies we have conducted, our MAP solution was generally superior to those obtained using minimum norm methods. However, this is true only when the sources exhibit the sparse focal characteristics on which our method is based. We stress that all of the methods provide good fits to the data, and hence are physically (if not physiologically) plausible. This ambiguity is inherent in attempting to infer spatial information from on the order of 100 external sensor measurements. Specific prior information is essential if useful spatial information is to be extracted from the data.

In this work we have developed a Bayesian framework for image estimation from MEG data. This approach can be extended to include information from other modalities (fMRI or PET) as well as using anatomical MR images to constrain sources to the cortex. The method can also be used for combined MEG/EEG data. We can also directly extend the model for dynamic imaging by simply replacing each of the amplitude processes,  $z_i$ , in our model with a time series model,  $z_i(t)$ .

The results that we have presented assume sources are constrained to 2D planes. Future research will focus on sources constrained to a realistic cortical surface. Only then can we begin to establish realistic limits on the ability of MEG to usefully image neural activity.

## References

- [1] C.J. Aine, "A Conceptual Overview and Critique of Functional Neuroimaging Techniques in Humans: 1. MRI/fMRI and PET," *Critical Reviews in Neurobiology*, vol. 9, nos. 2 and 3, pp. 229-309, 1995.
- [2] J. Besag, "On the statistical analysis of dirty pictures (with discussion)," *J. of the Royal Statist. Soc., Series B*, vol. 48, pp. 259-302, 1986.
- [3] G. Bilbro., et al. "Optimization by Mean Field Annealing," in: Touretzky D.S., *Advances in Neural Information Processing Systems*, San Mateo, CA: Morgan-Kaufman, pp. 91-98, 1989.
- [4] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans on Patt. Anal. and Mach. Int.*, vol. 6, no. 6, pp. 721-741, 1984.
- [5] I.F. Gorodnitsky, J.S. George, and B.D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *EEG and clinical Neurophysiol.*, 95, pp. 231-251, 1995.
- [6] P.C. Hansen, *Regularization Tools, A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems*, available in Postscript from <http://www.mathworks.com/>, March 1993.
- [7] Y. Okada, "Neurogenesis of Evoked Magnetic Fields," in: S.J. Williamson, G.L. Romani, L. Kaufman, and I. Miedna, *Biomagnetism, an Interdisciplinary Approach*, Plenum Press, N.Y., pp. 399-408, 1983.
- [8] J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Physics in Medicine and Biology*, vol. 32, pp. 11-22, 1987.
- [9] J.Z. Wang, S.J. Williamson, and L. Kaufman, "Magnetic Source Images Determined by a Lead-Field Analysis: The Unique Minimum-Norm Least-Squares Estimation," *IEEE Trans. on Bio. Eng.*, vol. 39, no. 7, pp. 665-675, 1992.