

Dynamic MEG Imaging of Focal Neuronal Sources¹

†James W. Phillips, †Richard M. Leahy, and ‡John C. Mosher

†Signal & Image Processing Institute
University of Southern California
Los Angeles, California 90089-2564 USA

‡Los Alamos National Laboratory
PS-21 MS D454
Los Alamos, New Mexico 87545 USA

Abstract

We describe inverse methods for using the magnetoencephalogram (MEG) to image neural current sources associated with functional activation in the cerebral cortex. A Bayesian formulation is presented that is based on a Gibbs prior which reflects the sparse, focal nature of neural activation. The model includes a dynamic component so that we can utilize the full spatio-temporal data record to reconstruct a sequence of images reflecting changes in the current source amplitudes during activation. The model consists of the product of a binary field, representing the areas of activation in the cerebral cortex, and a time series at each site which represents the dynamic changes in the source amplitudes at the active sites. Our estimation methods are based on the optimization of three different functions of the posterior density. Each of these methods requires the estimation of a binary field which we compute using a mean field annealing method. We demonstrate and compare our methods in application to computer generated and experimental phantom data.

I. INTRODUCTION

The primary advantage of MEG over other modalities that can image functional activation, such as fMRI and PET, is the ability to follow changes in neural activity on a millisecond time scale. Here we address the problem of producing *dynamic* images of neural activity using a Bayesian formulation that is based on an extension of the model we described in [9] for imaging with a single time-slice of data.

Physiological models for the MEG assume primary sources are constrained to the cortex with current flow oriented normal to the local surface [8]. By tessellating the cortex with N disjoint regions and representing the sources in each region at time t by an equivalent constrained current dipole oriented normal to the surface with amplitude $y_i(t)$, the problem can be expressed in terms of a linear model. The linear forward model relating the L time samples $\{t_j, j = 1 \dots L\}$ for N sources \mathbf{Y} ($N \times L$) and the M recording measurements \mathbf{B} ($M \times L$) can be written,

$$\mathbf{B} = \mathbf{G}\mathbf{Y} + \mathbf{N}. \quad (1)$$

The i 'th row of the $M \times N$ system matrix \mathbf{G} may be viewed as a discrete projection 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 $M \times L$ matrix \mathbf{N} represents noise generated within the sensors and by unwanted electromagnetic sources, and errors resulting from model mismatch.

Most MEG imaging approaches that have been described to date have been based on weighted minimum l_p -norm formulations, e.g. [10]. While methods of this type, or at least regularized versions of them, can perform well in other less ill-posed problems, the limitation of the number of spatial measurements to at best a few hundred and the inherent ill-posedness of the quasi-static inverse problem, result in very poor performance for these methods.

In order to produce a reasonable source image from MEG data, it is essential to introduce a substantial degree of prior information to the problem. For this reason, we propose an alternative approach based on a Bayesian formulation [9]. It is well known that the distribution of functional activation within the cerebral cortex is highly localized [1]. Consequently, we would not expect to see MEG source images showing very large areas of activation, but rather the source images should be sparse and focal. This information is incorporated into the Bayesian formulation using a prior that includes an indicator process, in the form of a binary Gibbs field, that represents the specific sites that are active; the Gibbs energy is chosen to ensure that, with high probability, these fields are sparse and focal. The model for the source image is completed by combining the binary field with a time series at each active site that represents the time varying amplitudes of the current sources.

Once the posterior density for the source image is specified, we choose a cost function to optimize which gives us a point estimate of the source image. The first cost function we investigated was the joint maximum *a posteriori* (MAP) estimator with respect to both the binary indicator field and the amplitude of the current sources. By integrating over all possible source amplitudes, we can form the binary marginal density with respect to the indicator process only. Our second method forms a (marginalized) MAP estimate over this marginalized posterior density. Our third approach minimizes the expected number of incorrectly identified active sites in the indicator process. This result can be found using the maximizer of posterior marginals (MPM) estimate [6]. We describe each of these methods below, apply them to simulated and experimental phantom data.

¹ This work was supported by the National Institute of Mental Health Grant No. R01-MH53213, the National Eye Institute Grant No. R01-EY08610-04, and Los Alamos National Laboratory, operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

II. THE BAYESIAN FORMULATION

We described a Bayesian formulation of the MEG imaging problem in a prior publication [9]. Here we extend this formulation to include a dynamic component. Our expectation that primary current sources are sparse and focal is naturally introduced into the inverse problem using a Bayesian paradigm in which the source is modeled as a random field. Since sources are sparse, a large majority of source pixels will have zero amplitude over the entire time series. We therefore use a binary indicator process \mathbf{x} to model whether each source dipole is *on* ($x_i=1$) or *off* ($x_i=0$). Active sites are assumed to have a temporally white Gaussian amplitude, $z_i(t)$ over the L time samples. We can write the (NxL) source image matrix \mathbf{Y} as,

$$\mathbf{Y} = \mathbf{XZ} \quad (2)$$

where $\mathbf{X} = \text{diag}(\mathbf{x})$ is a diagonal matrix. Assuming independence of the indicator and amplitude processes, the posterior probability for \mathbf{x} and \mathbf{Z} given the MEG data matrix \mathbf{B} is,

$$p(\mathbf{x}, \mathbf{Z}|\mathbf{B}) = \frac{p(\mathbf{B}|\mathbf{x}, \mathbf{Z})p(\mathbf{x})p(\mathbf{Z})}{p(\mathbf{B})}. \quad (3)$$

The joint probability $p(\mathbf{x})$ uses a Markov Random Field (MRF) 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,

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

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

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

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

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

where ∂_{ij} is the absolute distance between pixels i and j . The clustering term is small if neighboring pixels are of the same binary magnitude. The exponential parameter Q determines the strength of the clustering. As Q increases, the size of the clusters tends to increase.

The source amplitude process, \mathbf{Z} , is assumed to be a set of zero mean, temporally white, Gaussian random variables with spatial covariance \mathbf{C}_Z . The data is assumed to be corrupted by zero mean, additive Gaussian noise with covariance \mathbf{C}_N . Using these definitions, we can write the posterior density as,

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

where K_B is the posterior partition function and $U(\mathbf{x}, \mathbf{Z}|\mathbf{B})$ is the overall posterior energy function,

$$\frac{1}{2} \text{Tr}\{[\mathbf{B} - \mathbf{GXZ}]^T \mathbf{C}_N^{-1} [\mathbf{B} - \mathbf{GXZ}] + \mathbf{Z}^T \mathbf{C}_Z^{-1} \mathbf{Z}\} + LV(\mathbf{x}). \quad (8)$$

III. OPTIMIZATION TECHNIQUES

We investigate three cost functions based on the posterior energy defined in the previous section. Optimization of the cost function gives us an estimate of the current density.

A. The Joint MAP Estimate

The joint MAP estimate is found by maximizing over the log-posterior for both \mathbf{x} and \mathbf{Z} , or equivalently,

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

Optimizing jointly over continuous and discrete variables as is proposed in (9) is difficult. We note, however, that the energy function $U(\mathbf{x}, \mathbf{Z}|\mathbf{B})$ is quadratic in \mathbf{Z} . We can derive a closed form expression for the optimal \mathbf{Z} as a function of any particular indicator process \mathbf{x} ,

$$\mathbf{Z}^*(\mathbf{x}) = \arg \min_{\mathbf{Z}} U(\mathbf{x}, \mathbf{Z}|\mathbf{B}). \quad (10)$$

Setting the derivative of (8) to zero and solving yields,

$$\mathbf{Z}^*(\mathbf{x}) = \mathbf{C}_Z \mathbf{X} \mathbf{G}^T (\mathbf{G} \mathbf{X} \mathbf{C}_Z \mathbf{X} \mathbf{G}^T + \mathbf{C}_N)^{-1} \mathbf{B}. \quad (11)$$

Substituting $\mathbf{Z}^*(\mathbf{x})$ into $U(\mathbf{x}, \mathbf{Z}|\mathbf{B})$ results in,

$$\tilde{U}(\mathbf{x}|\mathbf{B}) = U(\mathbf{x}, \mathbf{Z}|\mathbf{B}) \Big|_{\mathbf{Z}=\mathbf{Z}^*(\mathbf{x})} \quad (12)$$

which is a Gibbs energy function for the density $\tilde{p}(\mathbf{x}|\mathbf{B})$, a function of \mathbf{x} only. We discuss minimization of a function of a binary field in Section IV.

B. Marginal MAP Estimation

An alternative solution can be found by considering only the locations of the sources. In this case we can remove the source amplitudes by marginalization of the posterior. To prepare for the marginalization, we rearrange (8) as:

$$\frac{1}{2} \text{Tr}\{(\mathbf{Z} - \mathbf{M}_Z)^T \mathbf{Q}^{-1} (\mathbf{Z} - \mathbf{M}_Z) - \mathbf{A}(\mathbf{x})\} + LV(\mathbf{x}) \quad (13)$$

where the new variables are defined as,

$$\mathbf{Q}^{-1} = \mathbf{X} \mathbf{G}^T \mathbf{C}_N^{-1} \mathbf{G} \mathbf{X} + \mathbf{C}_Z^{-1} \quad (14)$$

$$\mathbf{M}_Z = \mathbf{Q} \mathbf{X} \mathbf{G}^T \mathbf{C}_N^{-1} \mathbf{B} \quad (15)$$

$$\mathbf{A}(\mathbf{x}) = \mathbf{B}^T \mathbf{C}_N^{-1} \mathbf{G} \mathbf{X} \mathbf{Q} \mathbf{X} \mathbf{G}^T \mathbf{C}_N^{-1} \mathbf{B} - \mathbf{B}^T \mathbf{C}_N^{-1} \mathbf{B}. \quad (16)$$

To find the marginalized posterior, we integrate over all \mathbf{Z} :

$$p(\mathbf{x}|\mathbf{B}) = \int_{\mathbf{Z}} p(\mathbf{x}, \mathbf{Z}|\mathbf{B}) d\mathbf{Z}. \quad (17)$$

After removing terms from the integrand that are independent of \mathbf{Z} , we see that the integral is simply equal to the partition function of a temporally white Gaussian random process with spatial covariance \mathbf{Q} . We take the log of the partition function $(2\pi)^{N/2} |\mathbf{Q}|^{1/2}$ and substitute this for the integral to get a Gibbs distribution with marginal energy function,

$$U(\mathbf{x}|\mathbf{B}) = -\frac{L}{2} \ln |\mathbf{Q}| - \frac{1}{2} \text{Tr}\{\mathbf{A}(\mathbf{x})\} + LV(\mathbf{x}). \quad (18)$$

This is the energy function of the marginal posterior distribution, and the marginal MAP estimate is given by:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} U(\mathbf{x}|\mathbf{B}). \quad (19)$$

As with the joint MAP estimate, this requires minimization with respect to a binary field, as described in Section IV.

C. Marginal MPM Estimation

We can disregard the MAP approach entirely and attempt to find the best estimate of the set of pixel sites that are active during a specific response in the sense of minimizing the number of incorrectly identified active and inactive sites. The appropriate cost for this problem for a particular realization \mathbf{x} is:

$$e(\mathbf{x}, \hat{\mathbf{x}}) = \sum_i e(x_i, \hat{x}_i) = \sum_i (x_i - \hat{x}_i)^2 \quad (20)$$

where \hat{x}_i is the estimated pixel indicator value. The optimal Bayesian estimator \mathbf{x}^* with respect to the cost functional e is defined as the global minimizer of the expected value \bar{e} over all possible \mathbf{x} given the data \mathbf{B} :

$$\mathbf{x}^* = \arg \min_{\hat{\mathbf{x}}} \sum_{\mathbf{x}} e(\mathbf{x}, \hat{\mathbf{x}}) p(\mathbf{x}|\mathbf{B}). \quad (21)$$

If the posterior marginal distributions for every element of the field are known, the optimal Bayesian estimator with respect to any positive definite cost function may be found by independently minimizing the marginal expected cost for each element [6]. The optimal estimate of \mathbf{x} with respect to the positive definite cost functional e can therefore be found by minimizing independently the marginal expected cost for each binary element. i.e.,

$$x_i^* = \begin{cases} 0, & \text{if } \sum_{x_i} (x_i - 0)^2 p(x_i|\mathbf{B}) \leq \sum_{x_i} (x_i - 1)^2 p(x_i|\mathbf{B}) \\ 1, & \text{otherwise} \end{cases}. \quad (22)$$

Equivalently, if \bar{x} is the mean of the marginalized posterior $p(\mathbf{x}|\mathbf{B})$, then the MPM solution is given by:

$$x_i^* = \begin{cases} 0 & \text{if } \bar{x}_i \leq \frac{1}{2} \\ 1, & \text{else} \end{cases}. \quad (23)$$

Therefore, we find the MPM estimate by first finding the marginalized posterior mean field of the binary process, and then compare this mean to a threshold of 1/2 for each site.

IV. MEAN FIELD ANNEALING

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

$$x_i^{(n+1)} = E_T\{x_i|x_j^{(n)} \quad \forall j \neq i\} \quad (24)$$

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

$$p_T(\mathbf{x}|\mathbf{B}) = \frac{1}{K_T} \exp\left\{-\frac{1}{T} U(\mathbf{x}|\mathbf{B})\right\}. \quad (25)$$

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 $U(\mathbf{x}|\mathbf{B})$. This energy should be set to (12) for MAP estimation or (18) for marginalized MAP. If we use the same procedure, but iterate with constant temperature $T = 1$, we typically converge to a fixed point which represents the mean of a mean field approximation [3] of the original Gibbs distribution. We use this technique to compute an approximate solution to the marginal MPM problem.

V. SIMULATIONS

We have conducted extensive simulations based on a simplified 2D source model with a 1D array of 80 sensors. All dipolar sources are constrained to the annular segment of the x-y plane shown in Fig. 1. The distance between source locations was set to 1.5 mm providing an isotropically sampled image grid of 1,098 pixels. All dipoles were constrained in orientation perpendicular to the x-y plane. The source images were chosen

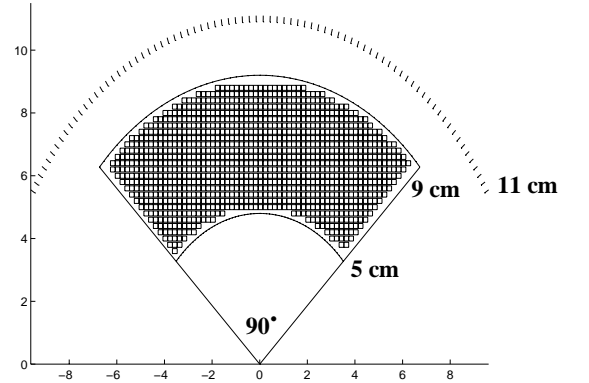


Figure 1. Reconstruction region for 2D simulations. 1,098 pixels, 80 sensors.

using stochastic sampling from $p(\mathbf{x})$ and $p(\mathbf{Z})$ with $Q = 2.5$, $\alpha_i = 0.200$, $\beta_j = 0.050$, and $\sigma_i^2 = 10 \text{ nAm}$ $i = 1 \dots N$. The first three parameters were chosen using the number of clusters and cluster size as the metric in a Markov chain Monte Carlo study [4]. We chose values of α_i , β_j , and Q which produce an average of 2-3 clusters with an area of 0.3 cm^2 for each cluster. The fourth parameter, the standard deviation of the dipole moments, reflects the typical activity seen in an evoked response study [5]. The true solution is shown in Fig. 2. We set \mathbf{C}_n to $v^2 \cdot \mathbf{I}$ where v^2 is the added noise variance which we assume known.

The simulation results for our three methods are shown in Fig. 3. The percent residual error (% R.E.) is shown above each image. For the binary solutions of marginal MAP and marginal MPM, we used (11) to find \mathbf{Z}^* and calculated % R.E. with that

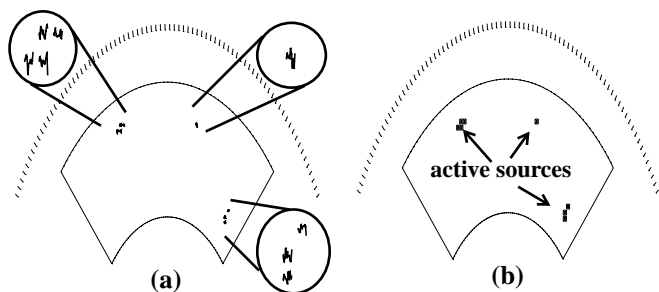


Figure 2. True solution. (a) Time series image, individual plots shown magnified. 8 active sites, 10-point time-series. (b) The same image in binary form with active sites shown as black rectangles.

amplitude process. The % R.E. for the true solution was 1.055 for data with SNR of 20 dB and 14.492 for data with SNR of 8 dB. The three solutions had a better fit than the true solution and achieved it with smaller clusters (lower energy). Note however that in this example all three methods gave very similar results, and placed three clusters of sources close to the correct locations for the three clusters in the true solution.

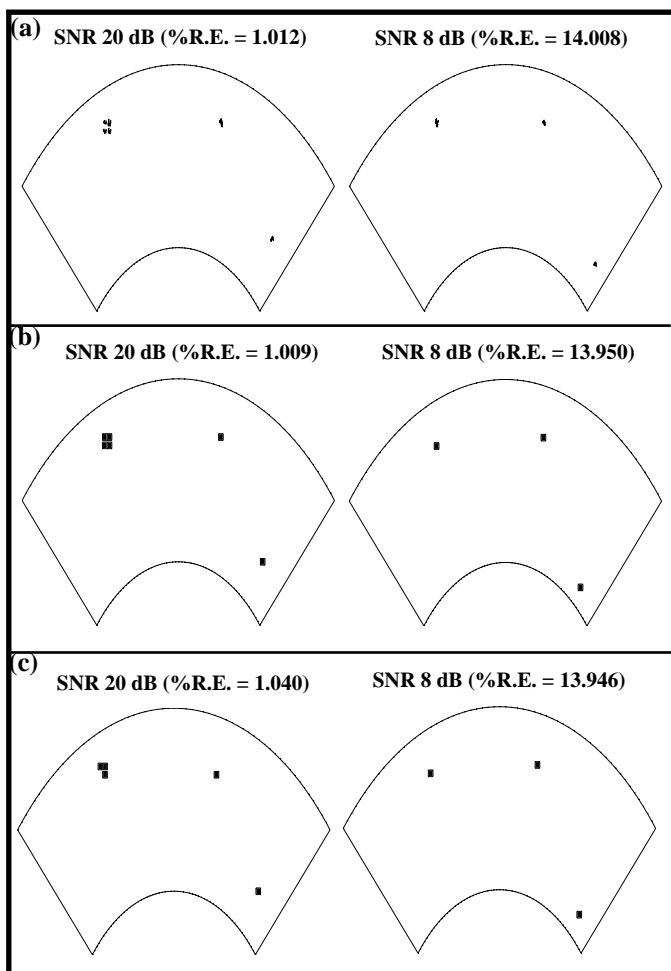


Figure 3. 2D simulation results shown with SNR of 20 dB and 8 dB. (a) Joint MAP estimate. (b) Marginal MAP estimate. (c) Marginal MPM estimate. Percent residual error (%R.E.) is shown above each image.

VI. PHANTOM EXPERIMENTS

We have applied the same techniques used in Section V to experimental phantom data collected with a Neuromag-122 system [5] using the 26 dipole phantom supplied by the manufacturer. The Neuromag-122 system employs 61 dual-channel planar first-order gradiometer units in a helmet-shaped configuration at a radius of 10-11 cm, measuring the magnetic field gradient in two orthogonal tangential directions, for a total of 122 individual sensor measurements.

The phantom consists of two half circles with a 7 cm radius in the x-z plane and y-z plane, with dipoles in fixed positions in these planes oriented tangential to the outer edge. The image grid consisted of 768 locations spaced 4 mm apart on these two half circles, with an inner radius of 3 cm and an outer radius of 7 cm. Fig. 4 shows a picture of the phantom and the reconstruction region used in our phantom experiments.

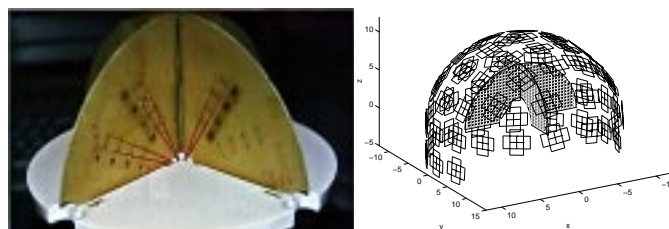


Figure 4. (a) The phantom. (b) Reconstruction region surrounded by the Neuromag-122 sensors.

We used the same parameters in the reconstruction as we did in Section V. These parameters reflect what we expect to see in real data. Since this phantom only activates individual dipoles, it is impossible to generate clustered sources. This phantom is therefore inadequate in testing the clustering properties of our methods. The true solution was generated by activating three dipoles and recording measurements of the magnetic field gradient for 20 time samples. The true solution is shown in Fig. 5.

The phantom data was scaled to reflect a reasonable evoked field response. We then added data collected in the same system from a passive human subject (100 averages, eyes closed, and no external stimulus present). This background was added to the phantom data to obtain a SNR of 15 dB and 7 dB.

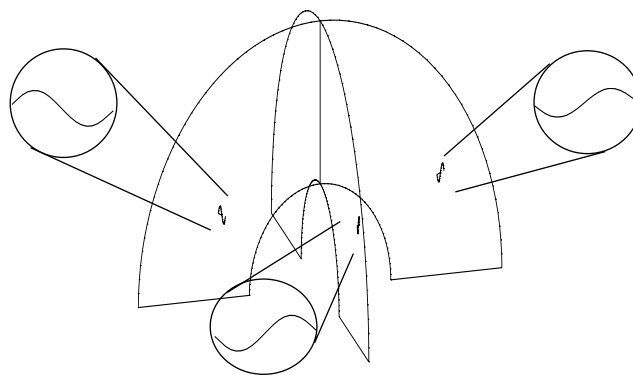


Figure 5. True image showing time series of 3 active dipoles magnified. 50-point time-series.

The results for the phantom experiment are shown in Fig. 6. The % R.E. for the true solution was 3.430 for SNR 15 dB and 19.979 for SNR of 7 dB. The reconstruction in every case found a three-dipole solution which better fit the data than the true solution. These reconstructions would therefore be preferred over the true solution by our optimization methods. However, all reconstructions are again close to the original source distribution.

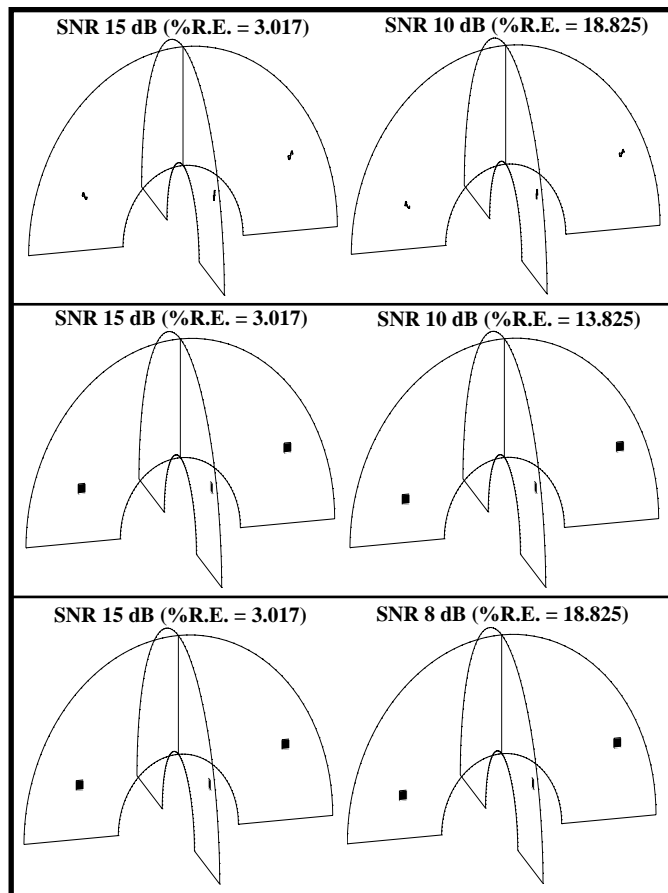


Figure 6. Results of phantom experiments. (a) Joint MAP estimate. (b) Marginal MAP estimate. (c) Marginal MPM estimate. Percent residual error (%R.E.) is shown above each image.

VII. DISCUSSION AND CONCLUSIONS

In the simulations and phantom studies we have conducted, our Bayesian methods generally found results which closely resembled the true solution. All reconstructions closely fit the data and exhibited the sparse clustered property that we expect. We note that in most cases all three of the methods described give similar performance in terms of localization accuracy. Since the computation cost varies considerably, with marginal MPM requiring the least computation time, the results provide support for using this approach over the more time consuming methods. We note however, that the two MAP methods are not guaranteed to give a global maximum and that the marginalized MPM method is based on a mean field approximation. Consequently it remains to be seen whether the similarity

between the three solutions is due to the true nature of the posterior density or whether the high degree of similarity may be partially attributed to the approximations and numerical optimization techniques that we used. In the future, we will perform more exhaustive quantitative comparisons of these methods.

The results that we have presented assume sources are constrained to 2D planes. In the future, the methods described here will be applied to sources constrained to a realistic cortical surface. Furthermore, since EEG and MEG are complimentary modalities, we plan to combine these in an attempt to maximize localization accuracy.

Other viable methods exist for source localization that were not discussed here, such as non-linear multiple dipole techniques [7]. We plan to compare our methods to these techniques. Eventually, we hope to determine realistic limits on the ability of MEG and EEG to usefully image neural activity.

We would like to thank Dr. Jeffrey Lewine of the New Mexico Institute of Neuroimaging, Albuquerque, for providing access to the NeuroMag-122 system and data. We also thank Dr. Cheryl Aine for her pre-stimulus data.

VIII. REFERENCES

- [1] C.J. Aine, "A Conceptual Overview and Critique of Functional Neuroimaging Techniques in Humans: 1. MRI/fMRI and PET," *Crit Rev Neurobiol*, vol. 9, nos. 2 and 3, pp. 229-309, 1995.
- [2] 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.
- [3] D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, pp. 131-138, 1987.
- [4] J. Goutsias, "Markov Random Fields: Interacting Particle Systems for Statistical Image Modeling and Analysis," *IEEE Proc*, submitted, 1996.
- [5] M. Hämäläinen *et al.*, "Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev Mod Phys*, Vol. 65(2), pp. 413-497, 1993.
- [6] J.L. Marroquin, "Probabilistic Solution of Inverse Problems," Tech. Rep TR-860, MIT, Artificial Intell. Lab., 1985.
- [7] J.C. Mosher, P.S. Lewis, R. Leahy, and M. Singh, "Multiple Dipole Modeling and Localization from Spatio-Temporal MEG Data," *IEEE Trans Biomed Eng*, Vol. 39(6), pp. 541-557, 1992.
- [8] Y. Okada, "Neurogenesis of Evoked Magnetic Fields," in *Biomagnetism, an Interdisciplinary Approach*, S.J. Williamson *et al.*, Ed., Plenum Press, N.Y., pp. 399-408, 1983
- [9] W. Phillips, R. M. Leahy, and J. C. Mosher, "MEG-based Imaging of Focal Neuronal Current Sources," *IEEE Trans Med. Imag.*, accepted, 1996.
- [10] 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 Biomed Eng*, Vol. 39(7), pp. 665-675, 1992.