NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 28 (2005) 84 – 98

# Spatiotemporal Bayesian inference dipole analysis for MEG neuroimaging data

Sung C. Jun,* John S. George, Juliana Paré-Blagoev, Sergey M. Plis, Doug M. Ranken, David M. Schmidt, and C.C. Wood

*Biological and Quantum Physics Group, MS D454, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

Received 17 November 2004; revised 24 May 2005; accepted 1 June 2005
Available online 15 July 2005

**Recently, we described a Bayesian inference approach to the MEG/EEG inverse problem that used numerical techniques to estimate the full posterior probability distributions of likely solutions upon which all inferences were based [Schmidt, D.M., George, J.S., Wood, C.C., 1999. Bayesian inference applied to the electromagnetic inverse problem. Human Brain Mapping 7, 195; Schmidt, D.M., George, J.S., Ranken, D.M., Wood, C.C., 2001. Spatial-temporal bayesian inference for MEG/EEG. In: Nenonen, J., Ilmoniemi, R. J., Katila, T. (Eds.), Biomag 2000: 12th International Conference on Biomagnetism. Espoo, Norway, p. 671]. Schmidt et al. (1999) focused on the analysis of data at a single point in time employing an extended region source model. They subsequently extended their work to a spatiotemporal Bayesian inference analysis of the full spatiotemporal MEG/EEG data set. Here, we formulate spatiotemporal Bayesian inference analysis using a multi-dipole model of neural activity. This approach is faster than the extended region model, does not require use of the subject's anatomical information, does not require prior determination of the number of dipoles, and yields quantitative probabilistic inferences. In addition, we have incorporated the ability to handle much more complex and realistic estimates of the background noise, which may be represented as a sum of Kronecker products of temporal and spatial noise covariance components. This reduces the effects of undermodeling noise. In order to reduce the rigidity of the multi-dipole formulation which commonly causes problems due to multiple local minima, we treat the given covariance of the background as uncertain and marginalize over it in the analysis. Markov Chain Monte Carlo (MCMC) was used to sample the many possible likely solutions. The spatiotemporal Bayesian dipole analysis is demonstrated using simulated and empirical whole-head MEG data.**


*Keywords:* Bayesian inference; Source localization; Spatiotemporal analysis; Dipole analysis; Magnetoencephalography; Markov Chain Monte Carlo

## Introduction

Magnetoencephalography (MEG) and electroencephalography (EEG) are non-invasive techniques. These methods measure direct physical consequences of neuronal currents and are capable of resolving temporal patterns of neural activity in the millisecond range. The MEG/EEG source localization problem, which identifies active brain regions from measurements on or outside of the human head, has been important in medical diagnosis of conditions like epilepsy, in surgical planning, and in neuroscience research. However, the MEG/EEG source localization inverse problem is mathematically ill-posed, that is, it has no unique solution.

For several decades, researchers have worked to develop MEG/EEG source localization methods to try to overcome the inherent ill-posed nature of the inverse problem. A number of localization methods which assume a dipolar source or an extended source have been developed (see Hämäläinen et al. (1993) for review). Most existing approaches fall into two broad categories: (1) few-parameter models having $N_p \ll N_s$ and (2) many-parameter models having $N_p \gg N_s$, where $N_p$ is the number of parameters to be estimated in the model and $N_s$ is the number of measurements, typically the number of sensors in MEG/EEG system. In general, few-parameter models are solved by finding a best-fitting solution through various nonlinear optimization techniques (Hämäläinen et al., 1993; Mosher et al., 1992; Huang et al., 1998; Uutela et al., 1998a; Jun et al., 2002). Many-parameter models are usually solved by the minimum norm method, or variants of the same, that selects the one solution minimizing a specified norm from the many solutions that fit the data equally well (Hämäläinen and Ilmoniemi, 1994; Gorodnitsky et al., 1995; Robinson and Vrba, 1999; Pascual-Marqui et al., 1994).

Recently, new probabilistic approaches to the MEG/EEG source localization problem based on Bayesian inference using Markov Chain Monte Carlo (MCMC) have been reported by Schmidt et al. (1999), Bertrand et al. (2001a,b), and Kincses et al. (2003). Unlike other probabilistic approaches (Baillet and Garnero, 1997; Phillips et al., 1997), the Bayesian inference

* Corresponding author.
*E-mail address:* jschan@lanl.gov (S.C. Jun).
**Available online on ScienceDirect (www.sciencedirect.com).**

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.06.003

approach does not result in a single best solution to the problem but produces a large number of likely solutions that fit both the data and any prior information. From the many sampled likely solutions, we can characterize some statistical information on any feature of solutions. This provides an effective means for quantifying uncertainty that is distinct from the other approaches to quantify uncertainty in inverse algorithms (Medvick et al., 1989; Singh and Harding, 2000; Darvas et al., 2005). Schmidt and Bertrand focused on the analysis of data at a single point in time and demonstrated the utility of Bayesian inference both for including pertinent prior information (anatomical location and orientation, sparseness of regions of activity, limitations on current strength, and spatial correlation) and for yielding robust results in spite of the under-determined inverse problem. Schmidt et al. (1999) used an extended region model for neural activity and Reversible Jump (RJ) MCMC method, while Bertrand et al. (2001a,b) used a multi-dipole model and combined RJ-MCMC and Parallel Tempering (PT) MCMC method. Schmidt et al. (2001) extended their work to a Bayesian inference analysis of the full spatiotemporal MEG/EEG data set, using their extended region model for neural activity.

Here, we present a spatiotemporal Bayesian inference technique for multi-dipole analysis. Compared to the full spatiotemporal analysis for extended regions, it is faster and does not require the use of the subject's anatomical information. Furthermore, in distinction to most other dipole analyses, it does not require the prior determination of the number of dipoles.

We begin with an overview of the general techniques of Bayesian inference. Then, we formulate the posterior probability distribution by incorporating the relevant priors into the Bayesian framework. To reduce computation costs and to improve MCMC performance, the posterior probability distribution is simplified by a marginalization technique over current time courses and a noise covariance matrix. A speed-up strategy for computing the posterior probability distribution is proposed, the MCMC sampling technique is briefly introduced, and then noise covariance approximation is discussed. Finally, results from experiments on simulated and empirical data are presented.

## Formulation of Bayesian inference

Bayesian inference is a general procedure for constructing a posterior probability distribution for quantities of interest from the measurements and the given prior probability distributions for all uncertain parameters. The method is conceptually simple and relatively straightforward for even complicated problems.

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

$$P(\theta, \mathbf{B}) = P(\theta|\mathbf{B})P(\mathbf{B}) = P(\mathbf{B}|\theta)P(\theta),$$

If $\theta$ represents parameters of interest and $\mathbf{B}$ represents data depending on $\theta$, then the probability of $\theta$ given $\mathbf{B}$ is

$$P(\theta|\mathbf{B}) = \frac{P(\theta, \mathbf{B})}{P(\mathbf{B})} = \frac{P(\mathbf{B}|\theta)P(\theta)}{P(\mathbf{B})}.$$

Here, $P(\theta, \mathbf{B})$ is the joint probability distribution for $\theta$ and $\mathbf{B}$, $P(\theta|\mathbf{B})$ is the conditional probability distribution of $\theta$ given $\mathbf{B}$, $P(\mathbf{B})$ is the marginal probability distribution of $\mathbf{B}$, and $P(\theta)$ is the prior probability distribution of $\theta$, which represents one's information of $\theta$ before measurement. $P(\mathbf{B}|\theta)$ is the likelihood function which modifies the prior $P(\theta)$ to produce the posterior probability distribution $P(\mathbf{B}|\theta)$. Since $P(\mathbf{B})$ is independent of $\theta$, it is constant and can be omitted from the posterior density:

$$P(\theta|\mathbf{B}) \propto P(\mathbf{B}|\theta)P(\theta).$$

Bayes' rule of probability formulates how prior information and measurements can be combined and encoded in the posterior distribution. Commonly, the obtainable posterior distribution is complex and in such cases is numerically sampled using MCMC techniques (Chen et al., 2000; Gilks et al., 1995).

In this work, we propose a spatiotemporal MEG/EEG dipole analysis based on Bayesian inference. This analysis is formulated in the following way: assuming a localized effective dipole nature of the neuromagnetic sources that can explain the spatiotemporal data, we construct a current model that assumes a variable number of current dipoles of brain activity that are composed of their locations within a sphere of some radius $R_0$, dipole orientations, and current time courses representing dipole magnitudes over time. Furthermore, we assume a fixed dipole model, where dipole locations and orientations are fixed over time, but dipole magnitudes vary over time. There can be any number $N$ of active current dipoles from minimum $N_{\min}$ up to some maximum $N_{\max}$. We used a spherical head model and the Sarvas forward model (Sarvas, 1987), but our analysis could employ other forward models as well.

Given the spatiotemporal measurement set, the Bayesian formulation is as follows:

| $P(N, \boldsymbol{X}, \boldsymbol{\Theta}, \mathbf{J}\|\mathbf{B}) \propto$ | $P(\mathbf{B}\|N, \boldsymbol{X}, \boldsymbol{\Theta}, \mathbf{J})\ P(\boldsymbol{\Theta}\|\boldsymbol{X}, N)\ P(\mathbf{J}\|N)\ P(\boldsymbol{X}\|N)\ P(N)$ |
|---|---|
| $\mathbf{B}$ | $T \times L$ matrix representing observed spatiotemporal data. $L$ and $T$ represent the number of sensors and the number of time samples in measurements, respectively. |
| $N$ | A priori unknown number of dipole sources. |
| $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)$ | Vector of $N$ dipole sources, with each $X_i = (x_i, y_i, z_i)$ representing the location of the $i$-th dipole. |
| $\mathbf{J} = (J_1, J_2, \ldots, J_N)$ | Vector of $N$ current time courses, with each $J_i = (j_i^1, j_i^2, \ldots, j_i^T)$ representing signed dipole moment magnitude over time of $i$-th dipole. Negative sign means that dipole moment orientation is reversed. |
| $\boldsymbol{\Theta} = (\theta_1, \theta_2, \ldots, \theta_N)$ | Vector of $N$ dipole moment orientations, with each $\theta_i$ representing a unit tangential direction of $i$-th dipole. |

The prior distributions are constructed as follows:

- The dipole current time course prior distribution is chosen as a Gaussian distribution:

$$P(\mathbf{J}|N) = \frac{1}{\prod_{\alpha=1}^{N}\left[\left(2\pi\sigma_\alpha^2\right)^{T/2}|\mathbf{C}_{\mathrm{cu}}|^{1/2}\right]} e^{-\frac{1}{2}\sum_{\alpha=1}^{N}\frac{1}{\sigma_\alpha^2}J'_\alpha\mathbf{C}_{\mathrm{cu}}^{-1}J_\alpha}. \quad (1)$$

Here, $|\cdot|$ denotes the determinant. $\mathbf{C}_{\mathrm{cu}}$ is the temporal correlation matrix of one time point with another, which allows us to include the temporal correlation at nearby latencies. $\sigma_\alpha$ represents the prior standard deviation of time varying current magnitudes of each dipole. Both $\mathbf{C}_{\mathrm{cu}}$ and $\sigma_\alpha$ are predetermined based on spatiotem-