International Journal of Bioelectromagnetism
Vol. 13, No. 4, pp. 167-172, 2011

Evaluating an empirical Bayesian source estimation method with dense array EEG and a realistic head model

J. Song[a,†], H. T. Attias[b], P. Luu[a], and D. Tucker[a]

[a] *EGI, 1600 Millrace Dr. Suite 307, Eugene OR 97403, USA*
[b] *ConvexImaging, 650 California St. San Francisco, CA 94108, USA*
[†] Corresponding author, email: jsong@egi.com, phone +1 541 687 7962, fax +1 541 687 7963

Abstract. Source analysis of the electroencephalography (EEG) involves finding the current distribution within the brain that explains the electric potential recorded on the surface of the head. In a linear inverse formulation, with distributed sources in the brain, the inverse problem is ill-posed because the number of electrodes is typically less than the number of current sources to be estimated. In this paper, we evaluate an empirical Bayesian method described by Wipf et al. (2010) for solving the inverse problem through modeling the temporal variability of the sources and the property of noise in the data. We evaluate this method with both simulations and recordings of 256 channel dEEG, using an accurate high-resolution, finite difference model of the head.

Keywords: Minimum-norm, sLORETA, source localization, empirical Bayesian analysis, EEG

1. Introduction

For source localization of the electroencephalogram (EEG), the *forward problem* consists of calculating the electric fields on the scalp, given the sources distribution within the brain (cortex) and a head model that specifies the geometry and conductivity of head tissues. In contrast, the *inverse problem* consists of finding the magnitude and orientation of the current sources, given the electric field recorded on the head's surface and a head model. For distributed sources in the cortex, the inverse problem involves solving for more unknowns (sources) than the measurement points (electrodes) on the head surface. As a result, there can be no unique solution.

An alternative to the distributed linear inverse formulation is the equivalent current dipole (ECD) approach. The ECD is a method of estimating the location, orientation and amplitude of single or multiple dipoles in the brain. Although ECD fitting provides robust estimates if the number of sources is known, the results are inaccurate when the assumed number of dipoles differs from the true number. There is no general way to determine the true number of dipoles. Furthermore, in practice it is rare that only a few neural sources are active. Thus, it is necessary to include sources for each region of the cortex to prevent contributions of un-modeled regions from distorting localization results. As a result, an ECD model that covers all cortical regions begins to approximate a distributed solution.

In the distributed inverse approach, the source space (i.e., the cortex) is discretized into a certain number of locations, where each location represents a dipole. If the cortical surface is known, the dipoles are oriented perpendicular to the surface; otherwise, orientations vary in 3-dimensional space. The solutions are computed by minimizing a suitable cost function. Minimum norm estimates (MN) are widely used, constraining the solution to the minimum total energy. The MN is based on a regularized linear inverse operator. In practice MN solutions are rather diffuse and are biased toward superficial sources.

Various MN constraints, including Weighted Minimum Norm (WMN), Local Autoregressive Average (LAURA), Low Resolution Electromagnetic Tomography (LORETA) and standardized Low Resolution Electromagnetic Tomography (sLORETA) have been proposed to deal with the shortcomings of the MN technique. Each method has its own advantages. The standardization of dipole moments for the particular head model (as in sLORETA) improves inverse accuracy considerably, but only if the head model used for the standardization matches that used for the inverse. Applied to real EEG, this requires a forward head model that is accurate for the head geometry and conductivity of the person whose EEG sources are to be localized. To date there is little evidence

that forward models achieve the requisite precision to insure that the standardization guarantees accuracy with real data in the way it does with simulations.

These inverse methods all share two limitations: they fail to take advantage of the knowledge of statistical properties of the sources, and they fail to explicitly handle noise contaminating the sensor data. Both limitations lead to inaccurate inverse solutions. In this paper, we evaluate an empirical Bayesian method of distributed source localization (Champagne) described by Wipf et al. (2010) that addresses these two limitations. This method places prior distributions over the sources, with hyperparameters that are learned from the sensor data using an iterative algorithm. In the formulation of Champagne evaluated here, the goal is to represent the data with a few sources, such as might be expected for focal sensory or motor activations. The algorithm also embeds a statistical noise model, leading to a more accurate reconstruction of the event-related response.

2. Methods

2.1. Model

The source localization problem, or inverse problem, involves determining source activations based on EEG signals on the scalp. Distributed inverse methods assume thousands of dipoles are distributed evenly throughout the cortex in fixed locations and orientations.

The voltages measured at the scalp are denoted by $\Phi \in \mathbb{R}^{N_e \times N_t}$, where N_e equals the number of sensors and N_t is the number of time points at which measurements are made. Each unknown activation of the dipole J_i is a N_d-dimensional vector with values at N_t time points. If the dipole orientation is known, $N_d = 1$, otherwise $N_d = 3$. Let $J \equiv [J_1^T, \ldots, J_{N_v}^T]^T \in \mathbb{R}^{N_j}$, where $N_j = N_d N_v$ and N_v is the number of dipole locations. The candidate N_v locations can be obtained by segmenting a structural MR scan of a human subject and tessellating the brain volume with a set of vertices.

There is a linear relationship between the activations of the dipoles and the voltages measured at the scalp. For a given source distribution J and electrical measurement Φ, there exists a so-called lead field matrix $K \in \mathbb{R}^{N_e \times N_j}$ such that

$$\Phi = KJ + \mathcal{E}, \tag{1}$$

where K can be calculated by modeling the flow of current through the head. There is a variety of methods for estimating the lead field matrix, including the spherical model, boundary element model, finite element model, and finite difference model. These models present trade-offs in terms of speed and accuracy of computation. The factors in this calculation include the head geometry, tissue conductivities and sensor positions. K is solved through the use of Poisson′s equation, which is based on Maxwell′s equations. Finally, $\mathcal{E}$ is the additive measurement error or noise and may include interfering signals that are non event-related.

To obtain a reasonable spatial resolution, the number of candidate source locations will necessarily be much larger than the number of sensors ($N_e \ll N_j$). This renders the inverse problem ill-posed, i.e., severely underdetermined. Since the mapping from source distribution J to sensor measurements Φ is many-to-one, all reconstructions of the source distribution are heavily dependent on prior assumptions.

Such prior assumptions can be described in terms of source properties, or prior probability distributions over sources, which are often considered to be fixed and known, as in the case of MN, WMN, LAURA, LORETA and sLORETA. Alternatively, empirical Bayesian approaches have been proposed to estimate model covariance components from data [Friston et al., 2008].

The empirical Bayesian method [Wipf et al., 2010] we evaluate in this paper improves on existing methods in terms of source distribution accuracy and computational robustness and efficiency.

2.2. Modeling assumptions

In this section we present a probability-based model that describes the joint distribution of the sensor data. This model will be used to derive an algorithm for learning the hyperparameters and reconstruct the source activation.

We start by translating Eq.1 into a statistical form. Assuming the noise $\mathcal{E}$ is an i.i.d. Gaussian variable with mean zero and covariance matrix $\Sigma_{\mathcal{E}}$, the probability of the data conditioned on the sources is given by

$$p(\Phi|J) \propto \exp(-\frac{1}{2}\|\Phi - KJ\|^2_{\Sigma_{\mathcal{E}}^{-1}}), \tag{2}$$

where $\|X\|_W$ denotes the weighted matrix norm $\sqrt{trace(X^TWX)}$. The unknown noise covariance $\Sigma_{\mathcal{E}}$ will be estimated from the data. For now, we assume it is fixed and known. Next we adopt the following source prior for J:

$$p(J|\Gamma) \propto \exp(-\frac{1}{2}trace\Big[\sum_{i=1}^{N_v} J_i^T\Gamma_i^{-1}J_i\Big]). \tag{3}$$

This is equivalent to assuming independently at each time point, a zero-mean Gaussian distribution with covariance Γ_i for each dipole J_i. We define Γ to be the $N_j \times N_j$ block diagonal matrix formed by ordering each Γ_i along the diagonal of an otherwise zero-valued matrix.

If Γ is known, then the posterior distribution $p(J|\Phi,\Gamma) \propto p(\Phi|J)p(J|\Gamma)$ is a Gaussian distribution with mean and covariance given by

$$\begin{aligned} &E(J|\Phi,\Gamma) = \Gamma K^T(\Sigma_{\mathcal{E}} + K\Gamma K^T)^{-1}\Phi, \\ &Cov(J|\Phi,\Gamma) = \Gamma - \Gamma K^T(\Sigma_{\mathcal{E}} + K\Gamma K^T)^{-1}K\Gamma. \end{aligned} \tag{4}$$

Since Γ is actually unknown, an approximation $\hat{\Gamma} \approx \Gamma$ must be estimated first. One way to estimate is to integrate with respect to sources J and to maximize

$$\begin{aligned} &p(\Phi|\Gamma) = \int p(\Phi|J)p(J|\Gamma)dJ \propto |\Sigma_\Phi|^{-\frac{1}{2}} \exp(-\frac{1}{2}\Phi^T\Sigma_\Phi^{-1}\Phi), \\ &\Sigma_\Phi \equiv \Sigma_{\mathcal{E}} + K\Gamma K^T. \end{aligned} \tag{5}$$

This is equivalent to minimizing the cost function

$$\mathcal{L}(\Gamma) = -2\log p(\Phi|\Gamma) = \log|\Sigma_\Phi| + trace[C_\Phi\Sigma_\Phi^{-1}], \tag{6}$$

where $C_\Phi = \frac{1}{N_t}\Phi\Phi^T$ is the empirical covariance.

2.3. Learning the hyperparameter Γ

Given $\Sigma_{\mathcal{E}}$ and Γ, the posterior distribution over J has been computed above. We now turn to the task of estimating these unknown quantities. First, we estimate Γ assuming $\Sigma_{\mathcal{E}}$ is known. In the next section, we estimate $\Sigma_{\mathcal{E}}$ from data.

Estimating the hyperparameter Γ is performed by minimizing Eq. 6 with respect to Γ. Several EEG imaging algorithms rely on the form of the expectation-maximization (EM) algorithm [Friston et al., 2008]. However, these algorithms are very slow when N_v is large. Our approach to this task implements a variant of the EM algorithm that converges quickly and handles arbitrary dipole orientations.

We consider several forms for Γ. First, a simple form of Γ_i is $\Gamma_i = \gamma_i I$, where γ_i is a positive scalar. Second, the prior covariance of each source J_i is a diagonal. Third, Γ_i is a general covariance matrix with no restriction. This gives the following parameterizations :

$$\Gamma_i = \begin{bmatrix} \gamma_i & 0 & 0 \\ 0 & \gamma_i & 0 \\ 0 & 0 & \gamma_i \end{bmatrix}, \quad \Gamma_i = \begin{bmatrix} \gamma_{i_1} & 0 & 0 \\ 0 & \gamma_{i_2} & 0 \\ 0 & 0 & \gamma_{i_3} \end{bmatrix}, \quad \Gamma_i = \begin{bmatrix} \gamma_{i_{11}} & \gamma_{i_{12}} & \gamma_{i_{13}} \\ \gamma_{i_{12}} & \gamma_{i_{22}} & \gamma_{i_{23}} \\ \gamma_{i_{13}} & \gamma_{i_{23}} & \gamma_{i_{33}} \end{bmatrix}. \tag{7}$$

For the case of scalar lead field ($N_d = 1$), all 3 forms are identical. The update rule for Γ_i is

$$\Gamma_i^{new} \rightarrow Z_i^{-1/2}(Z_i^{1/2}X_iX_i^TZ_i^{1/2})^{1/2}Z_i^{-1/2}, \tag{8}$$

where $X_i = \Gamma_i K_i^T \Sigma_\Phi^{-1} \Phi$ and $Z_i = K_i^T \Sigma_\Phi^{-1} K_i$. The details can be found in Wipf et al. (2010).

Starting with a suitably chosen initial value for Γ_i, the algorithm iterates between Eq. 4 and Eq. 8 until convergence is achieved. It can be shown that each iteration decreases the cost function $\mathcal{L}(\Gamma)$ or leaves it unchanged. The hyperparameters may be initialized using a commonly used localization method like MN or sLORETA, or simply to $\Gamma_i = I$.

2.4. Learning the interference covariance $\Sigma_\mathcal{E}$

The noise covariance matrix $\Sigma_\mathcal{E}$ used by the above algorithm reflects contributions from all interference sources, including spontaneous brain activity, eye movements, muscle artifacts, and sensor noise. In this paper we assume that data is available from a time period containing only the noise $\mathcal{E}$, where the event-related source activity vanishes. This means that $J = 0$ and Eq. 1 becomes $\Phi = \mathcal{E}$. Such a data segment is usually termed baseline. For example, in data collected using the evoked stimulus experimental paradigm, the baseline consists of pre-stimulus data. We further assume that the noise is stationary, i.e., the noise statistics in the baseline are approximately the same as during event-related activity. In our case, this means the noise covariance remains nearly unchanged.

Under these assumptions, the covariance $\Sigma_\mathcal{E}$ can be obtained directly from the baseline. One simple way to estimate it is using the empirical covariance, $\Sigma_\mathcal{E} = C_\Phi$. However, in cases where the number of time points N_t is not sufficiently large compared to the number of sensors N_e, the empirical covariance may be non-invertible due to its low rank. This is rather common in data collected with a dense EEG array, where typically $N_e = 128$ or 256. Instead, we use a factor analysis (FA) parametrization,

$$\Sigma_\mathcal{E} = FF^T + \Lambda, \tag{9}$$

where F is a $N_e \times N_f$ matrix termed the interference mixing matrix (or factor loading) and Λ is a diagonal matrix. This corresponds to the FA model $\Phi = Fx + u$, where $x \sim N(0, I)$ is a N_f-dimensional vector containing the i.i.d. Gaussian factors, and $u \sim N(0, \Lambda)$ is a N_e-dimensional vector containing the noise. The matrices F, Λ and the number of factors N_f can be estimated using a variant of the EM algorithm [Zumer et al., 2007]. This algorithm differs from usual FA estimation methods by treating the mixing matrix in a Bayesian manner, i.e., computing a full posterior distribution over it rather than a point estimate. This is done in the variational Bayesian framework [Attias, 2000], which is particularly suitable to data with a low-rank empirical covariance, and facilitates estimating the optimal number of factors needed to explain the data.

In this paper we go a step further, and estimate F and Λ using not just the baseline but also the event-related data. To do this, we extend the above FA model to $\Phi = Gv + Fx + u$, where $v \sim N(0, I)$ is a N_g-dimensional vector which is non-zero only during event-related activity, and G is the $N_e \times N_g$ event-related mixing matrix. The data covariance during event-related activity is

$$\Sigma_\Phi = GG^T + FF^T + \Lambda, \tag{10}$$

Notice that the source contribution $K\Gamma K^T$ is modeled here by GG^T. Next, we learn F, G and Λ from both baseline and event-related data, obtaining a noise model (Eq. 9) more robust to non-stationarity.

3. Validation

In this section, we test the performance of Champagne on both simulated and real data sets.

Simulated data

We first construct tests using simulated data with a realistic source distribution. The brain was segmented into 7 mm voxels and a three-orientation ($N_d = 3$) forward lead field was calculated using Finite Difference Model (FDM) in which tissue conductivity was specified for each voxel of the atlas (typical) MRI, segmented into scalp, skull, CSF, and gray and white matter of the brain.

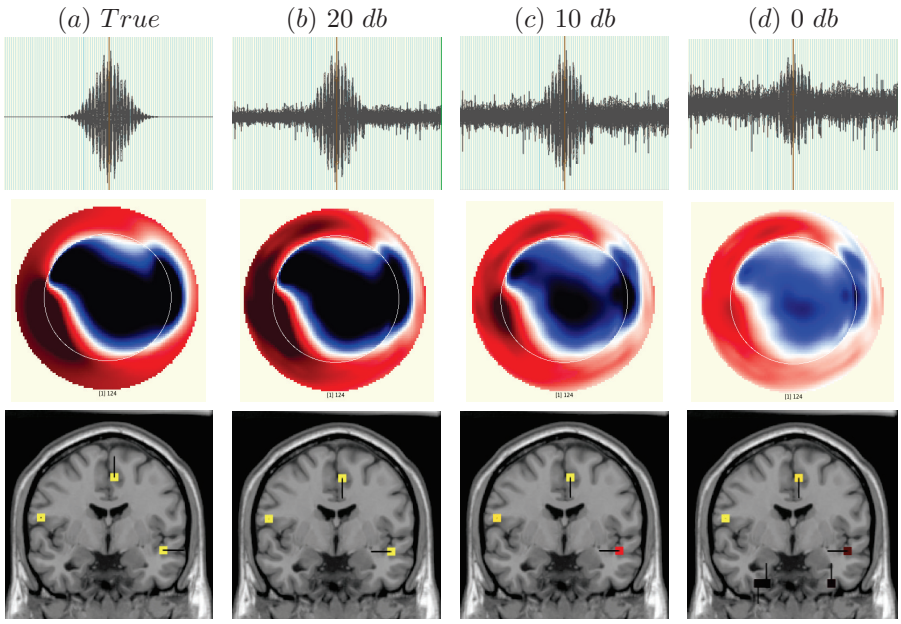

Figure 1. *Source localization performance on simulated 256-channel EEG data at different noise levels. Top row: time courses in butterfly plot. Middle row: scalp potential topographies. Bottom row: source distribution obtained by the Champagne method at sample 65, the time of peak intensity, of the post-stimulus period. Low SNR indicates the data is more contaminated by noise.*

The data time courses were divided into pre- and post-stimulus periods. The pre-stimulus period (50 samples) contains only noise, whereas the post-stimulus periods (200 samples) contains both noise and source activity of interest. The noise consists of resting state sensor recordings collected from a human subject. This activity was on-going and continued into the post stimulus period, to which the stimulated source signals (artificial event-related potentials) were added. Source time courses were seeded at locations in the brain with damped-sinusoidal signals and this voxel activity was projected onto the sensors through the FDM lead-field. Three source locations were randomly chosen. We adjusted the signal-to-noise ratio (SNR) and the correlation among source time courses. The SNR is defined as

$$SNR \equiv 20 \log \frac{\|KJ\|_{\mathcal{F}}}{\|\mathcal{E}\|_{\mathcal{F}}}. \tag{11}$$

We ran simulations of three randomly selected sources at SNR level of 0, 10, 20 dB. Fig. 1 shows the time courses, topomaps, and source distributions obtained by the Champagne method.

Real data

Next, we used the stimulus-evoked data collected from a normal healthy subject with a 256-channel Hydrocel Geodesic Sensor Net. The data is from a motor potential paradigm, which the subject pressed a button with the right thumb for 225 trials. The time of interest is the motor preparation period, which is from 100 ms before button press to the time of the button press. In a parallel fMRI session with the same paradigm and subject, the BOLD response was localized to the hand knob identified in the structural MRI [Yousry et al., 1997]. The oriented lead field matrix of the subject is used for source localization. Both the sLORETA and Champagne methods localized the premotor potential to the hand knob area. The source localization by sLORETA was more diffuse than that by Champagne, as shown in Fig. 2.

4. Conclusion

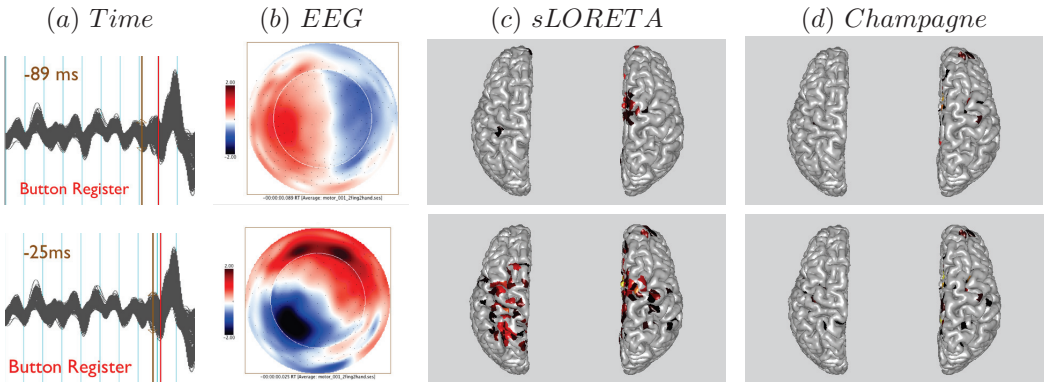

Figure 2. *Motor mapping performance on real data from 256 EEG channels. Column a: time courses. The red vertical lines indicate the moment of button press and the brown lines are the times for topographies and source distributions. Column b: scalp potential topographies at 89 ms and 25 ms before button press. Column c: source distribution obtained by sLORETA. Column d: source distribution by the Champagne method.*

This paper evaluates the Champagne method, which provides for unambiguous reconstruction of multiple sources. In contrast, sLORETA and other localization methods may not disambiguate multiple simultaneous sources. The source localization of Champagne was focal and sparse, consistent with simulated data, and consistent with our understanding of the averaged motor potential.

Statistical modeling of the temporal properties of the EEG data is an important direction for improving source localization. Furthermore, systematic artifacts such as eye movements or muscle activity may require more appropriate noise models. The Champagne method provides an instructive example for how, given adequate spatial sampling of the head surface and a high resolution realistic model of head conductivity, explicit modeling of the statistical properties of the brain signal and noise can improve the accuracy of EEG source localization.

References

Attias HT. A variational Bayesian framework for graphical models. *Advances in Neural Information Processing Systems*, 12:209-215, 2000.

Friston K, Harrison L, Daunizeau J, Phillips C, Trujillo-Barreto N, Henson R, Flandin G, Mattout J. Multiple sparse priors for the MEG/EEG inverse problem. *NeuroImage*, 39(1):1104-1120, 2008.

Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18:49-65, 1994.

Wipf DP, Owen JP, Attias HT, Sekihara K, Nagarajan SS. Robust bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *NeuroImage*, 49:641-655, 2010.

Yousry TA, Schmid UD, Alkadhi H, Schmidt D, Peraud D, Buettner A, Winkler P. Localization of the motor hand area to a knob on the precentral gyrus: A new landmark. *Brain*, 120:141-157, 1997.

Zumer JM, Attias HT, Sekihara K, Nagarajan SS. A probabilistic algorithm for interference suppression and source reconstruction from MEG/EEG data. *Neuroimage*, 37:102-115, 2007.