

A Graphical Model for Estimating Stimulus-Evoked Brain Responses in Noisy MEG data with Large Background Brain Activity

S.S. Nagarajan¹, H.T. Attias², K.E. Hild¹, K. Sekihara³

¹Department of Radiology, University of California-San Francisco, CA, USA

²Golden Metallic, Inc., San Francisco, CA, USA

³Department of Systems Engineering, Tokyo Metropolitan Institute of Technology, Japan

Abstract — This paper formulates a novel probabilistic graphical model for stimulus-evoked MEG and EEG sensor data obtained in the presence of large background brain activity. The model describes the observed data in terms of unobserved evoked and background sources. We present an expectation maximization algorithm that estimates the model parameters from data. Using the model, the algorithm cleans the stimulus-evoked data by removing interference from background sources and noise artifacts, and separates those data into contributions from independent factors. We demonstrate on real and simulated data that the algorithm outperforms benchmark methods for denoising and separation. We also show that the algorithm improves the performance of existing localization techniques.

Keywords— Bayesian methods, graphical models, ICA

I. INTRODUCTION

Electromagnetic source imaging (ESI), the reconstruction of the spatiotemporal activation of brain sources from MEG and EEG data, is increasingly being used for functional brain imaging of normal and diseased human brain function. A major problem in ESI is that MEG and EEG measurements, which use sensors located outside the brain, generally contain not only signals associated with brain sources of interest, but also signals from other sources such as spontaneous brain activity, eye blinks and other artifacts. Interference signals from those sources overlap spatially and temporally with those from the brain sources of interest, making it difficult to obtain accurate reconstructions. Many approaches have been taken to address the problem of interference from background brain activity and noise artifacts [1]. These approaches have provided some benefits but require a subjective choice of many parameters, such as thresholds for truncation and choice of specific components and usually don't work well for low SNR. Most of these algorithms also provide no principled mechanism for model order selection. Hence, the selection of the number of spatially distinct sources and the number of sources of interest must be based on ad-hoc methods or by expert analysis.

This paper presents a novel and powerful approach for the suppression of interference signals and the separation of signals from individual evoked sources. This approach is formulated in the framework of probabilistic graphical models with latent variables, which has been developed and studied in the fields of machine learning and statistics [2]. In

the graphical modeling framework, the observed data are modeled in terms of a set of latent variables, which are signals that are not directly observable. The dependence of the data on the latent variables is specified by a parameterized probability distribution. The latent variables are modeled by their own probability distribution. The combined distributions define a probabilistic model for the observed data. The model parameters are inferred from data using an expectation-maximization (EM) type algorithm, which is a standard technique for performing maximum likelihood in latent variable models. Problems such as interference suppression and source separation translate to the problem of probabilistic (Bayesian) inference of appropriate latent variables.

In our case, the observed MEG and EEG data are modeled in terms of the latent variables, termed factors, which represent evoked sources and interference sources. This paper focuses on MEG, but the framework is applicable to EEG as well. The model is defined in mathematical terms in section II. Section III presents a VB-EM (a generalization of standard EM) algorithm for this model[3]. The algorithm infers from data not just the model parameters, but also the number of factors required to represent evoked and interference sources (i.e., the dimensionalities of the evoked and interference subspaces). Next, it estimates the separate contribution of each individual evoked factor to the sensor data. This estimator uses the model to remove interferences and separate the evoked response into independent contributions. Section III also presents an automatically regularized estimator of the covariance matrix of the evoked response. Section IV demonstrates, using real and simulated data, that the algorithm provides interference-robust estimates of the time course and dimensionality of the stimulus evoked response. Furthermore, it shows that using the regularized evoked covariance in an existing source localization method improves its performance.

II. A PROBABILISTIC GRAPHICAL MODEL FOR STIMULUS-EVOKED MEG DATA

This section defines the graphical model for stimulus evoked MEG sensor data. Our model uses an independent factor analysis approach[4]. Assume we have K sensors that record N -point time series data. Let y_{in} denote the signal recorded by sensor $i = 1:K$ at time $n = 1:N$. We describe this multivariate time series as a linear combination of L stimulus evoked factors and M interference factors, both

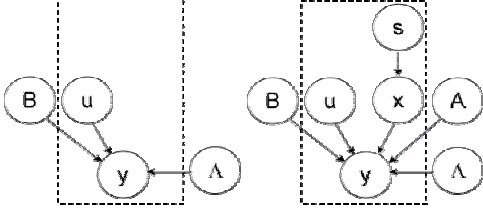


Figure 1: Probabilistic graphical model for MEG sensor data. Left: pre-stimulus period; data arise from interference factors u (modeled by Gaussians) via mixing matrix B . Right: post-stimulus period; data arise from interference factors u , plus independent stimulus-evoked factors x (modeled by MOGs with state vector s) via mixing matrix A . In both periods, the data include additive noise with precision matrix Λ . Nodes inside the dotted box denote model variables (which change with time), and nodes outside the box denote model parameters.

unobserved, contaminated by additive sensor noise. Let x_{jn} denote the activity of evoked factor $j=1:L$, let u_{jn} denote the activity of interference factor $j=1:M$, and let v_{jn} denote the noise in sensor i , all at time n . We also assume an evoked stimulus paradigm with a stimulus onset at time $n = N_0 + 1$. The evoked factors are active only during the post-stimulus period, and vanish during the pre-stimulus period, whereas the interference factors and sensor noise are active throughout. Hence $y_{in} = \sum_{j=1}^L A_{jn} x_{jn} + \sum_{j=1}^M B_{jn} u_{jn} + v_{in}$,

where the matrices A, B , termed mixing (or factor loading) matrices, define the linear combination relating the sensor and factor signals. In vector notation,

$$\begin{aligned} y_n &= Bu_n + v_n, & n &= 1: N_0 \\ y_n &= Ax_n + Bu_n + v_n, & n &= N_0 + 1: N \end{aligned} \quad (1.1)$$

To define the graphical model, we must specify probability distributions for the different signals. We model the interference factors and sensor noise as Gaussian. specifically, the interference factors are modeled by M independent, zero mean Gaussians with unit precision, and the sensor noise is modeled by a K dimensional, zero mean Gaussian with a diagonal precision matrix Λ (precision is defined as the inverse covariance).

Notice that this model describes the non-evoked part of the data as a K dimensional Gaussian with covariance $BB^T + \Lambda^{-1}$, which has $K(M+1)$ parameters. Alternatively, one could have used a general covariance matrix, requiring $K(K+1)/2$ parameters. This number is quadratic in the number of sensors and could become quite large, hence difficult to estimate accurately from finite data, resulting in an ill-conditioned covariance matrix. In contrast, the covariance in our model is always well conditioned, and with an appropriate choice of M the number of parameters is sufficiently small to allow accurate estimation. Moreover, rather than setting M manually, the graphical modeling

framework facilitates model order selection, i.e., inferring M (and L) from data, as described below.

Next, the evoked factors are modeled by independent non-Gaussian distributions. There are two reasons for avoiding Gaussians. First, evoked brain sources are often characterized by spikes or by modulated harmonic functions, leading to non-Gaussian distributions. Second, it is well known from work on ICA that mixed Gaussian sources cannot be separated. Here we model each factor by a 1 dimensional mixture of Gaussians (MOG) with S components (or states). Such a model provides an accurate description of arbitrary distributions, given an appropriate choice of S . It is also convenient to work with as a building block in larger graphical models. Here, for simplicity, we model all evoked factors by the same MOG distribution. For Gaussian state $s=1:S$, let μ_s, ν_s, π_s denote its mean, precision, and mixing proportion. Finally, we specify a prior distribution over the mixing matrices. We use an independent Gaussian for each matrix element A_{ij}, B_{ij} with hyperparameters α_i, β_j , respectively. This completes the definition of our probabilistic model, whose graphical representation is given in Fig. 1.

III. VB-EM ALGORITHM

We now present an algorithm, termed stimulus-evoked independent factor analysis (SEIFA), which infers the model parameters and reconstructs the evoked sources x from data. As usual in latent variable models, this is an iterative algorithm of the EM type. Here, however, we use a generalized version termed variational Bayesian EM (VB-EM) [3]. Whereas standard EM computes point estimates of parameters, VB-EM is formulated in the Bayesian framework and computes a full posterior distribution over the parameters A, B . This is necessary in order to perform model order selection. We also compute point estimates of the noise precision Λ and the hyperparameters α, β . The MOG parameters of the evoked factors are fitted to a particular sparse distribution (details omitted) in advance, and are kept fixed.

VB-EM is guaranteed to converge, since it maximizes an objective function which is upper-bounded by the marginal likelihood [3]. It is iterated from a random initialization to convergence. Each iteration includes an E-step and an M-step. Here we provide only the algorithm; details on the derivation will be given elsewhere.

E-step: Here we update the posterior distribution over the factors x, u , and compute their sufficient statistics (which are moments of the posterior) required for the M-step. For the pre-stimulus period, the posterior over u_n is Gaussian with mean $\bar{u}_n$ and covariance matrix Φ , where

$$\begin{aligned} \bar{u}_n &= \Phi \bar{B}^T \Lambda y_n \\ \Phi^{-1} &= \bar{B}^T \Lambda \bar{B} + I + K \Psi_{BB} \end{aligned} \quad (1.2)$$

For the post-stimulus we first define several quantities. Let $L' = L + M$ be the combined number of evoked and interference factors. Let $A' = (A, B)$ denote the $K \times L'$ matrix containing A and B , and let $x'_n = (x_n^T, u_n^T)^T$ denote the $L' \times 1$ vector containing x_n and u_n . Let s_n denote the state vector of the evoked factors at time n . Then conditioned on $s_n = s$, the posterior over all the factors is Gaussian, with mean $\bar{x}'_{sn}$ and covariance matrix Γ_s given by

$$\begin{aligned}\bar{x}'_{sn} &= \Gamma_s (\bar{A}'^T \Lambda y_n + v_s' \mu_s') \\ \Gamma_s^{-1} &= \bar{A}'^T \Lambda \bar{A}' + v_s' + K \Psi\end{aligned}\quad (1.3)$$

where Ψ is a sufficient statistic of the posterior over parameters computed in the M-step. The state posterior, i.e., the probability that $s_n = s$, is

$$\bar{\pi}_{sn} = \frac{1}{z_n} \pi_s \sqrt{|v_s| |\Gamma_s|} \exp\left(-\frac{1}{2} (y_n^T \Lambda y_n - \mu_s^T v_s \mu_s + \bar{x}_n'^T \Gamma_s^{-1} \bar{x}_n')\right) \quad (1.4)$$

where z_n is a normalization constant ensuring $\sum \bar{\pi}_{sn} = 1$.

M-step: Here we update the posterior distribution over the mixing matrices A, B , and compute their sufficient statistics required for the E-step. We also update Λ, α, β . The posterior over $A' = (A, B)$ is Gaussian, with mean $\bar{A}'$ and covariance matrix Ψ given by

$$\begin{aligned}\bar{A}' &= \begin{pmatrix} R_{yx} & R_{yu} \end{pmatrix} \Psi \\ \Psi &= \begin{pmatrix} R_{xx} + \alpha & R_{xu} \\ R_{xu}^T & R_{uu} + \beta \end{pmatrix}^{-1}\end{aligned}\quad (1.5)$$

where R_{xx}, R_{xu}, R_{uu} are the cross-correlations of the posterior means of the factors $\bar{x}_n, \bar{u}_n$ computed in the E-step. The updates of the hyperparameters and noise precision are given by

$$\begin{aligned}\alpha^{-1} &= \text{diag}\left(\frac{1}{K} \bar{A}'^T \Lambda \bar{A}' + \Psi_{AA}\right) \\ \beta^{-1} &= \text{diag}\left(\frac{1}{K} \bar{B}'^T \Lambda \bar{B}' + \Psi_{BB}\right) \\ \Lambda^{-1} &= \frac{1}{N + L + M} \text{diag}(R_{yy} - \bar{A} R_{yx}^T - \bar{B} R_{yu}^T)\end{aligned}\quad (1.6)$$

where $\Psi_{AA}, \Psi_{AB}, \Psi_{BB}$ are the top-left, top-right, bottom-right parts of Ψ corresponding to xx, xu, uu , respectively.

Model order selection: Inferring the number of evoked and interference factors L, M is done via the hyperparameters α, β . When the number of evoked factors (i.e., columns of A) is larger than needed to model the data, the VB-EM estimated values of α_j corresponding to the irrelevant columns end up approaching ∞ . Since α_j is the precision of the prior over column j , it causes the matrix elements A_{ij} to vanish, effectively canceling factor j . A similar effect occurs for the interference factors. This effect is termed automatic relevance determination and results from the Bayesian treatment (i.e., computing full posteriors over the parameters A, B) of our model [5].

Interference suppression and evoked source separation: Following convergence, one can also compute Bayes optimal estimates for the contribution of each evoked factor to the sensor signals. Furthermore, the algorithm automatically regularizes sensor covariance matrix for each evoked factor, which is often required for further processing such as localization [6]. Let $z'_{in} = A_{ij} x_{jn}$ denote the clean sensor signal i produced by evoked factor j , then its estimator is $\bar{z}'_{in} = \bar{A}_{in} \bar{x}_{jn}$ and its covariance is given by $C_{ik}^j = \bar{A}_{ij} (R_{xx})_{jj} \bar{A}_{jk}^T + \delta_{ik} \frac{1}{\Lambda_i} (R_{xx})_{jj} (\Psi_{AA})_{jj}$. Notice that the first term alone is ill-conditioned. The second, which is a direct

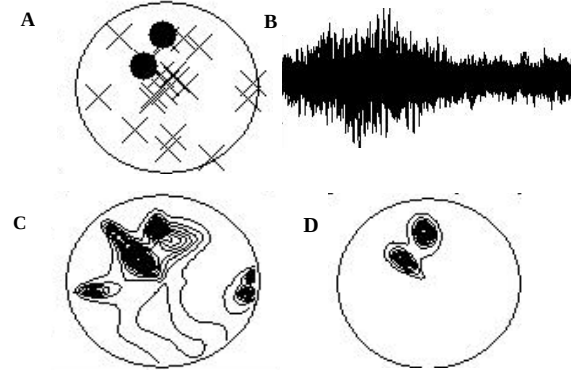


Figure 2. (A) Simulated location of stimulus-evoked sources (dots) and 21 background sources (crosses) assumed to lie on a plane with sensors all around the head. (B) Waveforms showing raw sensor data at a Signal to Interference Ratio (SIR) of 0 dB. (C) Beamformer reconstruction of raw sensor data cannot distinguish between true sources and interferences. (D) Regularized beamformer reconstruction of the denoised sources inferred from SEIFA shows accurate reconstruction of true sources.

result of the Bayesian treatment of our model, is the regularizer for the covariance matrix.

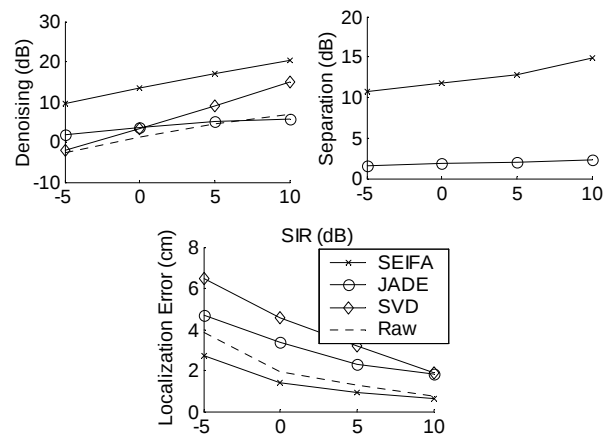


Figure 3: Top left – Mean denoising performance as a function of SIR plotted as a mean of 50 Monte Carlo simulations for different positions of active and interference sources. Error bars on mean are ~ 1 dB. Top right – Mean separation performance for SEIFA compared to JADE. Error bars on mean are ~ 1 dB. Bottom – Localization performance of beamformers for SEIFA, JADE, SVD and Raw data. SEIFA clearly outperforms all other algorithms on all three metrics.

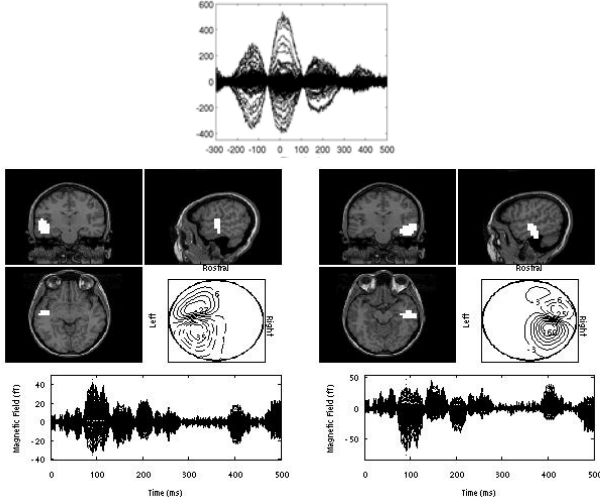


Figure 4: SEIFA denoising of eye-blink data. Top – The averaged magnetic field response to a tone in the presence of large eye-blink artifact is shown. Typical response peaking around 100 ms cannot be seen in the raw data. Bottom – Beamformer localization of responses in the left and right hemisphere computed from a denoised and separated factor. The denoised sensor data shows peak responses in each hemisphere around 100ms that can be well localized to auditory cortex in each hemisphere.

V. RESULTS

A synthetic dataset was constructed using a small number of evoked and interference sources, each of which has a sinusoidal time course of random frequency (chosen uniformly over a finite range). The finite length of each evoked source was enforced by modulating the corresponding sinusoid by a Hanning window having random length and random placement within the post-stimulus window. The subject's head was assumed to be spherical, reconstruction was done only for the $x = 0$ (coronal) plane, and 275 channel MEG data were simulated using a lead field matrix for a random physical location and orientation for each neural signal and then adding white Gaussian noise to each sensor. The power of the evoked sources relative to the power of the interference sources (computed in sensor space and averaged over all channels) is referred to as the signal-to-interference ratio (SIR). Likewise, the signal-to-noise ratio (SNR) is used to quantify the power of the evoked sources relative to the power of the additive sensor noise. SNR for simulations were fixed at 10 dB with $N=1000$ and $N_0=375$. The results below are given for several different values of SIR ranging from -5 dB to +10 dB. Comparisons include the SEIFA algorithm, an ICA method (JADE), results obtained by selecting the most energetic singular-value components (SVD), and the results obtained when no denoising procedure is used (Raw). Although the number of factors could be computed from our algorithm using marginal likelihood estimates, for these simulations we assumed that the numbers of evoked and interference sources were known. ICA was performed using JADE following an SVD truncation, where the number of

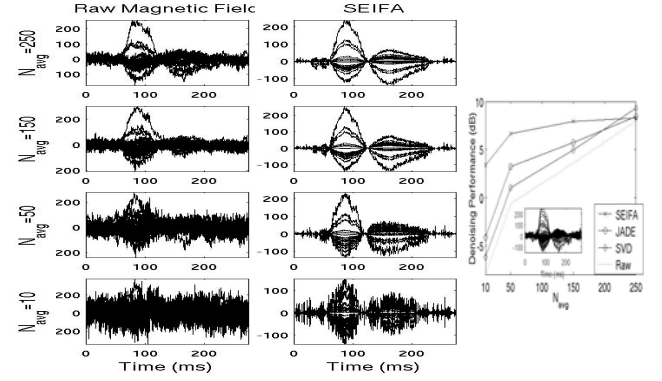


Figure 5: Left – Averaged auditory evoked response for varying number of trials ranging from 250 to 10 (top to bottom). Middle – SEIFA denoised data that shows good denoising occurs even for averages with 10 trials. Right – Performance comparison between SEIFA, JADE, SVD that indicates best performance for SEIFA.

components kept during the truncation is $L+M$. The denoised sensor signals were based on the L ICA components that have the highest ratio of post-stimulus energy to pre-stimulus energy. SVD results were based on keeping the number of components as $L+M$. Figure 2 shows an example from our simulations where noisy MEG data denoised using SEIFA clearly reconstructs the true sources whereas neither JADE, nor truncated SVD methods reconstruct the true sources. We compare performance across 50 Monte Carlo trials with different positions and orientations of the stimulus-evoked and background sources. As seen in figure 3, SEIFA significantly outperforms JADE and SVD methods in the three metrics of performance – denoising, separation and localization. Performance on real MEG data are shown in Figures 4 and 5. In figure 5, we were able to denoise large eye-blink artifacts that contaminated an auditory evoked response and to localize auditory cortex using SEIFA. Figure 5 shows SEIFA performance in extracting auditory responses for varying number of trials in the average. The algorithm was able to extract clean auditory evoked responses from even ten trials and outperforms SVD and JADE for small trial averages.

ACKNOWLEDGMENT

This work was funded by NIH grant (DC004855-01A1).

REFERENCES

- [1] S. Makeig, S. Debener, J. Onton, and A. Delorme, "Mining event-related brain dynamics," *Trends Cogn Sci*, vol. 8, pp. 204-10, 2004.
- [2] M. I. Jordan, "Learning in Graphical Models." Cambridge, Massachusetts: The MIT Press, 1998.
- [3] H. Attias, "A Variational Bayes framework for Graphical Models," *Advances in Neural Information Processing Systems*, vol. 12, pp. 209-215, 2000.
- [4] H. Attias, "Independent factor analysis," *Neural Comput*, vol. 11, pp. 803-51, 1999.
- [5] R. M. Neal, *Bayesian Learning for Neural Networks*: Springer, 1996.
- [6] K. Sekihara, S. S. Nagarajan, D. Poeppel, A. Marantz, and Y. Miyashita, "Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique," *IEEE Trans Biomed Eng*, vol. 48, pp. 760-71, 2001.