A probabilistic algorithm integrating source localization and noise suppression for MEG and EEG data

Johanna M. Zumer,[a,b] Hagai T. Attias,[c] Kensuke Sekihara,[d] and Srikantan S. Nagarajan[a,b,*]

[a]Biomagnetic Imaging Lab., Department of Radiology, University of California, San Francisco, San Francisco, CA 94143-0628, USA
[b]UCSF/UC Berkeley Joint Graduate Group in Bioengineering, University of California, San Francisco, San Francisco, CA 94143-0628, USA
[c]Golden Metallic Inc., PO Box 475608, San Francisco, CA 94147, USA
[d]Department of Systems Design and Engineering, Tokyo Metropolitan University, Tokyo 191-0065, Japan

We have developed a novel probabilistic model that estimates neural source activity measured by MEG and EEG data while suppressing the effect of interference and noise sources. The model estimates contributions to sensor data from evoked sources, interference sources and sensor noise using Bayesian methods and by exploiting knowledge about their timing and spatial covariance properties. Full posterior distributions are computed rather than just the MAP estimates. In simulation, the algorithm can accurately localize and estimate the time courses of several simultaneously active dipoles, with rotating or fixed orientation, at noise levels typical for averaged MEG data. The algorithm even performs reasonably at noise levels typical of an average of just a few trials. The algorithm is superior to beamforming techniques, which we show to be an approximation to our graphical model, in estimation of temporally correlated sources. Success of this algorithm using MEG data for localizing bilateral auditory cortex, low-SNR somatosensory activations, and for localizing an epileptic spike source are also demonstrated.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Magnetoencephalography; Electroencephalography; Inverse methods; Bayesian inference; Denoising

Introduction

Mapping the spatiotemporal neural activity of the entire brain is an important problem in basic neuroscience research. It is also clinically important for patients with brain tumors and epilepsy, both in localizing regions important for cognitive function and for identifying epileptogenic brain regions. Such brain mapping procedures are useful to guide neurosurgical planning, navigation, and resection.

Many noninvasive techniques have emerged for functional brain mapping, such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG). Although fMRI is the most popular method for functional brain imaging with high spatial resolution, it suffers from poor temporal resolution since it measures blood oxygenation level-dependent (BOLD) signals with fluctuations in the order of seconds. These BOLD signals are also indirect measures of neural activity which might not accurately reflect neural activity, especially in regions of altered vasculature. However, dynamic neuronal activity has fluctuations in the submillisecond time scale that can only be directly measured with magnetoencephalography (MEG) and/or electroencephalography (EEG). MEG data are the measurements of tiny magnetic fields surrounding the head while EEG data are the measurements of voltage potentials using an electrode array placed on the scalp.

The past decade has shown rapid development of whole-head MEG/EEG sensor arrays and of algorithms for reconstruction of brain source activity from MEG and EEG data, termed source localization. All existing methods for brain source localization are hampered by the many sources of noise present in MEG/EEG data. The magnitude of the stimulus-evoked neural sources is on the order of noise on a single trial, and so typically 50–200 averaged trials are needed in order to clearly distinguish the sources above noise. This limits the type of cognitive questions that can be asked and is prohibitive for examining processes such as learning that can occur over just one or several trials. Needing to average trials is time consuming and therefore difficult for a subject or patient to hold still or pay attention through the duration of the experiment. Gaussian thermal noise or Gaussian electrical noise is present at the MEG or EEG sensors themselves. Background room interference such as from powerlines and electronic equipment can be problematic. Biological noise such as heartbeat, eyeblink or other muscle artifact can also be present. Ongoing brain activity itself, including the drowsy-state alpha (∼10 Hz) rhythm can drown out evoked brain sources. Finally, many localization algorithms have

* Corresponding author. Biomagnetic Imaging Lab., Department of Radiology, University of California, San Francisco, San Francisco, CA 94143-0628, USA. Fax: +1 415 502 4302.
E-mail address: sri@radiology.ucsf.edu (S.S. Nagarajan).
URL: http://bil.ucsf.edu (S.S. Nagarajan).

Available online on ScienceDirect (www.sciencedirect.com).

difficulty in separating neural sources of interest that have temporally overlapping activity.

Noise in MEG and EEG data is typically reduced by a variety of preprocessing algorithms before being used by source localization algorithms. Simple forms of preprocessing include filtering out frequency bands not containing a brain signal of interest. Additionally and more recently, ICA algorithms have been used to remove artefactual components, such as eyeblinks (Jung et al., 2000). More sophisticated techniques have also recently been developed using graphical models for preprocessing prior to source localization (Nagarajan et al., 2005, 2006). Therefore, current algorithms for source localization from MEG and EEG data typically use a two-stage procedure — the first for noise/interference removal and the second for source localization.

This paper presents a probabilistic modeling framework for MEG/EEG source localization that estimates Source Activity using Knowledge of Event Timing and Independence from Noise and Interference (SAKETINI). The framework uses a probabilistic hidden variable model that describes the observed sensor data in terms of activity from unobserved brain and interference sources. The unobserved source activities and model parameters are inferred from the data by a variational Bayesian Expectation Maximization algorithm. The algorithm then creates a spatiotemporal image of brain activity by scanning the brain, inferring the model parameters and variables from sensor data, and using them to compute the likelihood of a dipole at each grid location in the brain.

We first describe the generative model for the data. We complete specification of the model for post-stimulus sources by including prior distributions and compute the unknown quantities learned from the data. We then describe the model for learning interference and noise sources from pre-stimulus data. We finish the Methods by showing that an established source localization method, the minimum variance adaptive beamformer (MVAB) (Sekihara et al., 2001), is an approximation of our framework. In the Results section, we show performance of SAKETINI relative to MVAB and sLORETA (Pascual-Marqui, 2002) both in localization ability and in time course estimation for MEG data. We show the effect of number of sensors and time points for all three methods. We further show the proposed method's performance applied to a real auditory-evoked MEG dataset, a low-SNR somatosensory MEG dataset, and an epileptic spike MEG dataset. We conclude with a discussion on the model order and inputs to the algorithm, extensions of the model, and its relationship to other methods in the literature. A preliminary report of this work was shown in Zumer et al. (2007).

Methods

Probabilistic model integrating source localization and noise suppression

This section describes the generative model for the data. We assume that the MEG/EEG data have been collected such that stimulus onset or some other experimental marker indicated the "zero" time point. Ongoing brain activity, biological noise, background environmental noise, and sensor noise are present in both pre-stimulus and post-stimulus periods; however, the evoked neural sources of interest are only present in the post-stimulus time period. We therefore assume that the sensor data can be described

as coming from four types of sources: (I) evoked source at a particular voxel (grid point), (II) all other evoked sources not at that voxel, (III) all background noise sources with spatial covariance at the sensors (including brain, biological, or environmental sources), and (IV) sensor noise. We first infer the model describing source types (III) and (IV) from the pre-stimulus data, then fix certain quantities (described below in next subsection) and infer the full model describing the remaining source types (I) and (II) from the post-stimulus data (described below in the third subsection). After inference of the model, a map of the source activity is created, as well as a map of the likelihood of activity across voxels.

Let y_n denote the $K \times 1$ vector of data from sensor data for time point n, where K is the number of sensors (typically ~200). Time ranges from $-N_{\text{pre}}:0:N_{\text{post}}-1$ where $N_{\text{pre}}(N_{\text{post}})$ indicates the number of time samples in the pre-(post-)stimulus period. The generative model for data y_n is

$$y_n = \begin{cases} Bu_n + v_n & n = -N_{\text{pre}}, \ldots, -1 \\ F^r s_n^r + A^{\backslash r} x_n^{\backslash r} + Bu_n + v_n & n = 0, \ldots, N_{\text{post}} - 1 \end{cases} \tag{1}$$

The $K \times q$ forward lead field matrix F^r represents the physical (and linear) relationship between a dipole source at voxel r for each dipole orientation q along a coordinate basis and its influence on sensor $k = 1:K$ (Sarvas, 1987). The lead field F^r is calculated from knowing the geometry of the source location to the sensor location, as well as assuming the type of conducting medium in which the source lies. The human head is most commonly approximated as a single-shell sphere volume conductor for MEG data; for EEG data, a more detailed model taking into account the tissue conductance boundaries is usually needed. In the most general case, $q = 3$ for all three possible directions of coordinate bases of a source dipole. In the case of the single-shell sphere, the radial component of source dipoles contributes nothing to MEG sensors, thus $q = 2$. If there is knowledge of subject-specific cortical anatomy, the source may be constrained to be perpendicular to the gray matter surface, thus $q = 1$. Throughout the rest of this paper, we use the single-shell model with $q = 2$ for both simulations and real data from MEG.

The source activity s_n^r is a $q \times 1$ vector of dipole strength at time n for the voxel r. The $K \times L$ matrix A^r and the $L \times 1$ vector x_n^r represent the post-stimulus mixing matrix and evoked non-localized factors, respectively, corresponding to source type (II) discussed above, where the $\backslash r$ superscript indicates for all voxels not at voxel r. We leave these superscripts off in the rest of this paper for clarity. The $K \times M$ matrix B and the $M \times 1$ vector u_n represent the background mixing matrix and background factors, respectively, corresponding to source type (III) discussed above. The $K \times 1$ vector v_n represents the sensor-level noise, which is assumed to be drawn from a Gaussian distribution with zero-mean and precision (inverse covariance) defined by the diagonal matrix λ. All quantities depend on r in the post-stimulus period except for B, u_n and λ, which will be learned from the pre-stimulus data, and fixed as the other quantities are learned for each voxel. Note, however, that the posterior update for u_n (represented by $\bar{u}_n$) does depend on the voxel r. The graphical model is shown in Fig. 1. This generative model becomes a probabilistic model when we specify prior distributions, as described in the next two sections (see Jordan (1999) and Beal et al. (2003) for additional information on related graphical models).
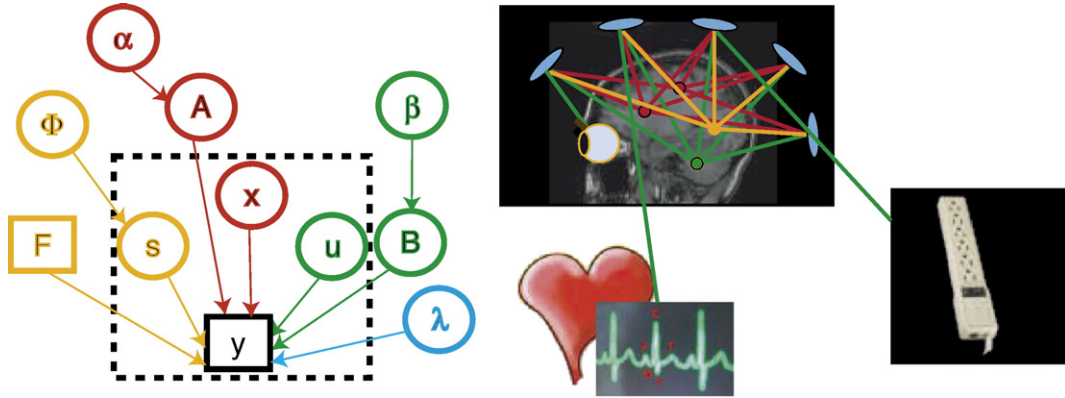
Fig. 1. (Left) SAKETINI graphical model. Variables dependent on time are inside dotted box; parameters independent of time outside dotted box. Values in circles are unknown and learned from the model, and values in squares are known. After applying a simpler model to the pre-stimulus data to learn the sensor noise precision and interference mixing matrix, these quantities are then known and placed in squares. (Right) Representation of factors influencing the data recorded at the sensors. In orange, a post-stimulus source at the voxel of interest, focused on by the lead field *F*. In red, other post-stimulus sources not at that particular voxel. In green, all background sources, including on-going brain activity, eyeblinks, heartbeat, and electrical noise. In blue, thermal noise present in each sensor.

Separation of background sources learned from pre-stimulus data

We learn the background mixing matrix B and sensor noise precision $\lambda = \mathrm{diag}(\lambda_1,...,\lambda_K)$ from the pre-stimulus data using a variational Bayesian factor analysis model. All variables, parameters and hyperparameters are hidden and are learned from the pre-stimulus data. We assume Gaussian prior distributions on the background factors and sensor noise; we assume a flat prior on the sensor noise precision. We further assume that the background factors are independent and identically distributed (i.i.d.) across time. All normal distributions specified in this paper are defined by their mean and precision (inverse covariance). The following is assumed for $n = -N_{\mathrm{pre}}: -1$:

$$p(u) = \prod_n p(u_n); \quad p(u_n) = \mathcal{N}(u_n|0,I), \tag{2}$$

$$p(v) = \prod_n p(v_n); \quad p(v_n) = \mathcal{N}(v_n|0,\lambda), \quad p(\lambda) = \mathrm{const}. \tag{3}$$

$$p(y|u, B, \lambda) = \prod_n p(y_n|u_n, B, \lambda); \quad p(y_n|u_n, B, \lambda) = \mathcal{N}(y_n|Bu_n, \lambda) \tag{4}$$

We define the notation $p(u_n) = \mathcal{N}(u_n|\mu,\Delta)$ to mean that the probability distribution of u_n is a normal Gaussian distribution with mean μ and precision (inverse covariance) Δ. To complete the specification of this model, we need prior distributions on the model parameters. We use a conjugate prior for the background mixing matrix B, as follows,

$$p(B) = \prod_{km} p(B_{km}); \quad p(B_{km}) = \mathcal{N}(B_{km}|0,\lambda_k\beta_m) \tag{5}$$

where β_m is a hyperparameter over the *m*th column of B and λ_k is the precision of the *k*th sensor. The matrix $\beta = \mathrm{diag}(\beta_1,...,\beta_M)$ provides a robust mechanism for automatic model order selection, so that the optimal size of B is inferred from the data through β (for general reference, see Neal (1996)).

In general, for any "free" probability density function (pdf) q, the "free energy" $\mathcal{F}$ (so-called in statistical physics) is written as

$$
\begin{aligned}
\mathcal{F}(q,\Theta) &= \int du\, dB\, q(u,B|y,\Theta)[\log p(y,u,B|\Theta) - \log q(u,B|y,\Theta)] \\
&= E_{q(u,B|y,\Theta)}[l(y,u,B|\Theta)] + H_q \\
&= \int du\, dB\, q(u,B|y,\Theta)\left[\log\frac{p(u,B|y,\Theta)}{q(u,B|y,\Theta)} + \log p(y|\Theta)\right] \\
&= \log p(y|\Theta) - KL[q(u,B|y,\Theta)||p(u,B|y,\Theta)] \\
&= \mathcal{L}(\Theta) - KL[q(u,B|y,\Theta)||p(u,B|y,\Theta)]
\end{aligned} \tag{6}
$$

where Θ is the set of all model parameters, in this case $\Theta = \{\beta, \lambda\}$, and H_q is the entropy of $q(u,B|y,\Theta)$. The Kullback–Leibler divergence between distributions q and p, defined $KL(q||p) = \int dx\, q(x)[\log q(x) - \log p(x)]$, is always non-negative and can equal zero if the two distributions are equal; thus, $\mathcal{F}$ is a lower bound to the log likelihood function $\mathcal{L}$. Exact inference on this model is intractable using the true joint posterior over the background factors u and mixing matrix B. We desire the marginal conditional posterior estimates of the background factors and mixing matrix, given the data. We could choose to estimate just the maximum a posteriori (MAP) estimates, but this ignores the full distribution including its precision. We could also choose to estimate points of the distribution extensively, using, for example, Markov Chain Monte Carlo methods (Jun et al., 2005; Gelman and Rubin, 1996); however this is computationally costly as well as dependent on the sampled points. Instead, we choose to factorize the distribution assuming conditional independence of the factors and parameters, also termed the mean field approximation.

$$p(u,B|y) \approx q(u,B|y) = q(u|y)q(B|y) \tag{7}$$

This approximation restricts the posterior to a product of factor distributions but allows the solution to be computed analytically. $\mathcal{F}$ is equal to $\log p(y|\Theta)$ when the approximation in Eq. (7) is true, thus making the *KL*-divergence zero. The variational Bayesian Expectation Maximization (VB-EM) algorithm (Attias, 1999; Ghahramani and Beal, 2001; Beal, 2003) iteratively maximizes $\mathcal{F}$ with respect to (w.r.t.) each factorized distribution to, at least, a local maximum of $\mathcal{F}$, alternating w.r.t. the posteriors $q(u|y)$ and $q(B|y)$. In

the E-step, $\mathcal{F}$ is maximized w.r.t. $q(u|y)$, keeping $q(B|y)$ constant, and the sufficient statistics of the hidden variables are computed. In the M-step, $\mathcal{F}$ is maximized w.r.t. $q(B|y)$, keeping $q(u|y)$ constant, and the MAP estimate of the parameters and hyperparameters is computed. The $q(u,B|y)=q(u|y)q(B|y)$ that maximizes $\mathcal{F}$ is the distribution that minimizes the *KL*-divergence, thus making the mean field approximation as close to equivalence as possible under the constraint implied by Eq. (7). A derivation of a similar model is described in the Appendix of Nagarajan et al. (2006) and the variational approximation has been shown to be a powerful and flexible tool (Jordan et al., 1999). It can further be shown that, given our previous assumptions of Gaussian prior distributions, the form of the posterior $q(u|y)$ must also be Gaussian, by showing that the function $\log q$ is quadratic in u.

In the E-step, maximizing $\mathcal{F}$ w.r.t. $q(u|y)$ holding $q(B|y)$ fixed gives the following posterior estimate update for the factors.

$$q(u|y) = \prod_n q(u_n|y_n); \quad q(u_n|y_n) = \mathcal{N}(u_n|\bar{u}_n, \gamma) \tag{8}$$
$$\bar{u}_n = \gamma^{-1}\bar{B}^T\lambda y_n; \quad \gamma = \bar{B}^T\lambda\bar{B} + K\psi^{-1} + I$$

For all variables in this paper, a bar over the variable indicates its posterior mean. In the M-step, we compute the full posterior distribution of the background mixing matrix B, including its precision matrix ψ, and the MAP estimates of the noise precision λ and the hyperparameter β.

$$q(B|y) = \mathcal{N}(B|\bar{B},\psi)$$
$$\bar{B} = R_{yu}\psi; \qquad \psi = (R_{uu} + \beta)^{-1} \tag{9}$$
$$\beta^{-1} = \mathrm{diag}\left(\frac{1}{K}\bar{B}^T\lambda\bar{B} + \psi\right); \quad \lambda^{-1} = \frac{1}{N}\mathrm{diag}(R_{yy} - \bar{B}R_{yu}^T)$$

The matrices, such as R_{yu}, represent the posterior covariance between the two subscripts.

$$R_{yy} = \sum_{n=1}^{N} y_n y_n^T, \quad R_{yu} = \sum_{n=1}^{N} y_n \bar{u}_n^T \tag{10}$$

Localization of evoked sources learned from post-stimulus data

In the stimulus-evoked paradigm, the source strength at each voxel is learned from the post-stimulus data. The background mixing matrix B and sensor noise precision λ are fixed, after having been learned from the pre-stimulus data, as described in the previous section. We assume that those quantities remain constant through the post-stimulus period and are independent of source location. We assume Gaussian prior distributions on the source factors and interference factors. We further make the assumption that both of these factors are i.i.d. across time.

The source factors have prior precision given by the $q \times q$ matrix Φ, which relates to the strength of the dipole in each of q directions. We have used the superscript r here as a reminder that s is for only one voxel at a time, although in general we will leave off this superscript.

$$p(s^r) = \prod_n p(s_n^r); \quad p(s_n^r) = \mathcal{N}(s_n^r|0, \Phi^r) \tag{11}$$

The interference and background factors are assumed to have identity precision (uninformed prior).

$$p(x) = \prod_n p(x_n); \quad p(x_n) = \mathcal{N}(x_n|0, I);$$
$$p(u) = \prod_n p(u_n); \quad p(u_n) = \mathcal{N}(u_n|0, I) \tag{12}$$

We also use a conjugate prior for the interference mixing matrix A, where $\alpha = \mathrm{diag}(\alpha_1,...,\alpha_L)$ is a hyperparameter, similar to the expression for the background mixing matrix, which helps in model order selection for determining the size of A from the data.

$$p(A) = \prod_{kl} p(A_{kl}); \quad p(A_{kl}) = \mathcal{N}(A_{kl}|0, \lambda_k\alpha_l) \tag{13}$$

We now specify the full model:

$$p(y|s, x, u, A, B, \lambda) = \prod_n p(y_n|s_n, x_n, u_n, A, B, \lambda);$$
$$p(y_n|s_n, x_n, u_n, A, B, \lambda) = \mathcal{N}(y_n|Fs_n + Ax_n + Bu_n, \lambda) \tag{14}$$

Exact inference on this model is intractable using the joint posterior over the evoked non-localized factors x and mixing matrix A; thus a variational Bayesian approximation for the posteriors is again used.

$$p(s, x, u, A|y) \approx q(s, x, u, A|y) = q(s, x, u|y)q(A|y) \tag{15}$$

All variables, parameters and hyperparameters are unknown and are learned from the data. We learn the hidden variables and parameters from the post-stimulus data, iterating through each voxel across the brain, using the VB-EM algorithm. Since maximizing the log $p(y|\Theta)$, for $\Theta = \{\alpha, \Phi\}$ would be mathematically intractable, we maximize a lower bound to $\log p(y|\Theta)$ defined by the free energy $\mathcal{F}$ in the following equation.

$$\begin{aligned}
\mathcal{F}(q, \Theta) &= \int du\, dx\, ds\, dA\, q(s,x,u,A|y, \Theta)[\log p(y,s,x,u,A|\Theta) \\
&\quad - \log q(s, x, u, A|y, \Theta)] \\
&= \log p(y|\Theta) \\
&\quad - KL[q(s, x, u, A|y, \Theta)||p(s, x, u, A|y, \Theta)]
\end{aligned} \tag{16}$$

In the E-step, the posterior distribution of the factors given the data is computed:

$$q(x_n'|y_n) = \mathcal{N}(x_n'|\bar{x}_n',\Gamma); \quad \bar{x}_n = \Gamma^{-1}\bar{A}'^T\lambda y_n;$$
$$\Gamma = \bar{A}'^T\lambda\bar{A}' + K\Psi + I' \tag{17}$$

where we define:

$$\bar{x}_n' = \begin{pmatrix} \bar{s}_n \\ \bar{x}_n \\ \bar{u}_n \end{pmatrix}; \quad \bar{A}' = (F\ \bar{A}\ \bar{B});$$
$$I' = \begin{pmatrix} \Phi & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}; \quad \Psi = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Psi_{AA} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{18}$$

In the M-step, we update the posterior distribution of the interference mixing matrix A including its precision Ψ_{AA}. Note that the lead field $\mathcal{F}$ is fixed and known based on the geometry of the sensors relative to the head (same quantity assumed known by beamforming and other approaches) and B was learned and fixed from the pre-stimulus data. The sensor noise precision λ is also kept fixed from the pre-stimulus period. The MAP values of the hyperparameter α and source factor precision Φ are learned here from the post-stimulus data.

$$\bar{A} = (R_{yx} - FR_{sx} - \bar{B}R_{ux})\Psi_{AA};$$
$$\Psi_{AA} = (R_{xx} + \alpha)^{-1}; \quad \Phi^{-1} = \frac{1}{N}R_{ss};$$
$$\alpha^{-1} = \mathrm{diag}\left(\frac{1}{K}\bar{A}^T\lambda\bar{A} + \Psi_{AA}\right) \tag{19}$$

Explicit definitions for all combinations of posterior covariances are omitted, but some are shown here for example:

$$R_{ss} = \sum_{n=1}^{N} \bar{s}_n \bar{s}_n^T + N\Sigma_{ss}, \; R_{sx} = \sum_{n=1}^{N} \bar{s}_n \bar{x}_n^T + N\Sigma_{sx} \qquad (20)$$

where $\Sigma = \Gamma^{-1}$ is specified as:

$$\Sigma = \begin{pmatrix} \Sigma_{ss} & \Sigma_{sx} & \Sigma_{su} \\ \Sigma_{sx}^T & \Sigma_{xx} & \Sigma_{xu} \\ \Sigma_{su}^T & \Sigma_{xu}^T & \Sigma_{uu} \end{pmatrix} \qquad (21)$$

In each iteration of VB-EM, the marginal likelihood is increased. The variational likelihood function (the lower bound on the exact marginal likelihood) is given as follows:

$$\mathcal{L}^r = \frac{N}{2} \log \frac{|\lambda||\Phi^r|}{|\Gamma^r|} - \frac{1}{2} \sum_{n=1}^{N} \left(y_n^T \lambda y_n - \bar{x}_n^{\prime rT} \Gamma^r \bar{x}_n^{\prime r} \right)$$
$$+ \frac{K}{2} \log|\alpha^r||\Psi^r| \qquad (22)$$

This likelihood function is dependent on the source voxel r, since we update the variables for each voxel, and thus a map of the likelihood across the brain can be displayed. Furthermore, an image of the source power estimates and the time course of activity at each voxel could be plotted.

We note that the computational complexity of the proposed algorithm is on the order $O(KLNS)$, roughly equivalent to a single dipole scan, which is on the order $O(N(K^2+S))$. These are much smaller than the complexity of a multi-dipole scan which is on the order $O(NS^P)$ where P is the number of dipoles, and if S represents roughly several thousand voxels.

We further note that the number of hidden variables to be estimated is less than the number of data points observed, thus not posing problems for estimation accuracy. See Fig. 7 for performance as a function of data points.

Relationship of SAKETINI to minimum variance adaptive beamforming

Minimum variance adaptive beamforming (MVAB) is one of the best performing source localization techniques (Sekihara et al., 2001, 2002, 2005; Küçükaltun-Yildirim et al., 2006). MVAB estimates the dipole source time series by $\hat{s}_n = W_{\text{MVAB}} y_n$, where

$$W_{\text{MVAB}} = (F^T R_{yy}^{-1} F)^{-1} F^T R_{yy}^{-1} \qquad (23)$$

Thus, MVAB also has computational complexity equivalent to a single-dipole scan, on the order $O(N(K^2+S))$. MVAB is known to fail for two sources that are highly correlated in time (Sekihara et al., 2002), although methods exist to correct this issue (Dalal et al., 2006; Brookes et al., 2007). In this section, we derive that MVAB is an approximation to SAKETINI.

We start by rewriting Eq. (1) as $y_n = Fs_n + z_n$, where z_n is termed the total noise and is given by $z_n = Ax_n + Bu_n + v_n$. It has mean zero and precision matrix $\Upsilon = (AA^T + BB^T + \lambda^{-1})^{-1}$. Assuming we have

estimated the model parameters A, B, and λ, the MAP estimate of the dipole source time series is $\bar{s}_n = Wy_n$, where

$$W = \Omega^{-1} F^T \Upsilon; \quad \Omega = F^T \Upsilon F + \Phi \qquad (24)$$

The Φ term exists from the prior distribution of s_n. It can be shown that Eq. (24) is equivalent to Eq. (23).

In the infinite data limit, the data covariance satisfies $R_{yy} = F\Phi^{-1}F^T + \Upsilon^{-1}$. Its inverse is found, using the matrix inversion lemma, to be $R_{yy}^{-1} = \Upsilon - \Upsilon F \Omega^{-1} F^T \Upsilon$. Hence, we obtain

$$F^T R_{yy}^{-1} = (I - F^T \Upsilon F \Omega^{-1}) F^T \Upsilon = \Phi \Omega^{-1} F^T \Upsilon \qquad (25)$$

where the last step used the expression for Ω. Next, we approximate $\Omega \approx F^T \Upsilon F$, which is equivalent to the prior term on s_n having zero precision. We then obtain

$$W \approx (F^T \Upsilon F)^{-1} F^T \Upsilon = (F^T \Upsilon F)^{-1} \Omega \Phi^{-1} \cdot \Phi \Omega^{-1} F^T \Upsilon$$
$$= (F^T R_{yy}^{-1} F)^{-1} F^T R_{yy}^{-1} = W_{\text{MVAB}} \qquad (26)$$

where the last step uses Eq. (25).

Simulations: generation and analysis

SAKETINI was tested in a variety of scenarios, including many simulations with real brain noise and real MEG datasets. Performance of the method is compared to the MVAB and sLORETA (Pascual-Marqui, 2002). Simulations and real data were analyzed using NUTMEG (Neurodynamic Utility Toolbox for MEG) (Dalal et al., 2004), a toolbox developed using MATLAB (MathWorks, Natick, MA, USA), obtainable from http://bil.ucsf.edu. NUTMEG is useful for coregistration of fiducial points to a structural MRI, selection of volume-of-interest, computation of forward field, filtering and other denoising preprocessing methods, as well as a variety of source reconstruction methods, including MVAB, sLORETA and now SAKETINI.

Simulations were created using a variety of realistic source configurations reconstructed on a 5-mm voxel grid. A single-shell spherical volume conductor model was used to calculate the forward lead field (Sarvas, 1987). Simulated datasets were constructed by placing Gaussian-damped sinusoidal time courses at specific locations inside a voxel grid based on realistic head geometry. Sources were set to be active only during a post-stimulus period, which always comprised 62.5% of the total data available, while the remaining 37.5% was pre-stimulus data. Typically 700 total data points were used, unless specified otherwise.

In one set of simulations, termed "simulated interference" cases, we created background activity drawn from the statistical distributions assumed by the model. The simulated sources mentioned above were projected to the sensors using the computed forward field. Noise drawn from a Gaussian distribution was added to each sensor across all time points. Furthermore, to simulate "ongoing brain activity," time courses were drawn from a Gaussian noise distribution and were placed in 30 random locations throughout the brain voxel grid, occurring in both pre- and post-stimulus periods. Their activity was projected onto the sensors and added to both Gaussian sensor noise and source activity. These simulated background brain sources add noise to the sensors in a spatially correlated manner.

In order to test simulation performance using data with more realistic (and unknown) statistical distributions, we created a second set of simulations termed "real brain noise." Real MEG sensor data were collected from a CTF MEG System with 275 axial gradio-

meters while a human subject was alert but not performing tasks or receiving stimuli. This real background data thus include real sensor noise plus real ongoing brain activity that could interfere with evoked sources and adds spatial correlation to the sensor data. Since throughout this paper we work with averaged data, these real data were binned into 100 trials of 700 data points each and averaged. We varied the signal to noise ratio (SNR) and the corresponding signal to noise-plus-interference ratio (SNIR). SNIR is calculated from the ratio of the sensor data resulting from sources only to the sensor data from noise plus interference.

Simulation methods: average performance

In addition to some illustrative examples, we examine average results over many realizations of simulations in which certain source parameters were varied. Two different source configurations were used: one with sources more near the surface as depicted in Fig. 3 and the other configuration with deeper sources. The orientation of the source was fixed in half the simulations and allowed to rotate over time in the other half. The correlation of two of the three sources with each other was set to be 0%, 95%, or 100%; the third source was always uncorrelated with the other two sources. Either simulated interference or real brain noise was used as background noise. SNIR was set at 5 dB, 0 dB, or −5 dB. Each combination of parameters was tested for 10 different randomly generated source time courses and source orientations. Thus, a total of 720 simulations were run using all combinations of simulation parameters.

Performance was measured in two ways: localization ability and estimation of time course. To assess localization ability, it is important to take into account source strength, source localization error, and presence of false positives. Thus we used the ROC (receiver-operator characteristic) method modified for brain imaging results as suggested by Darvas et al. (2004), which is a measure of hit rate versus false positive rate. The free-response ROC (FROC) curve in particular allows for multiple "hits" per image (Bunch et al., 1978).

We define a "local peak" as a voxel that is greater in value than its 26 three-dimensional neighbors. We define a "hit" as a local peak that is within a specified distance of the true location and above a certain threshold. A "miss" is defined as a true source location that has no "hit" within the specified distance. A "false positive" is a local peak above a certain threshold but further than the specified distance from a true source location. A "true negative" is any voxel that is none of the above.

FROC curves are generated by varying the threshold and allowable distance error, thus varying the tradeoff of sensitivity and specificity. We used $5 * \sqrt{3}mm$, $10 * \sqrt{3}mm$, or $15 * \sqrt{3}mm$ for the allowable localization error from a true source. We varied the threshold to be 30%, 50%, 70% or 90% of the maximum value in the whole image. We thus recorded a hit rate (HR) and false positive rate (FR) for each of 12 combinations of threshold/error for each of the 720 simulations.

Since these HR versus FR points do not increase monotonically, as they would if the threshold was the only criterion varied, we chose to use the measure of A'. A' is a way to approximate the area under the ROC curve for one HR/FR point (Snodgrass and Corwin, 1988). The larger the area under the ROC, the better the method is performing, since this means a higher HR relative to FR for specified thresholds/localization errors. For each simulation, the twelve computed A' values were averaged to give one A' value per

simulation. The SAKETINI likelihood map was used as the spatial map to test localization. For MVAB and sLORETA, the power map (sum of squares of post-stimulus time points) was used.

Simulation methods: effects of amount of data

Previous studies have shown advantage of sensor arrays with larger number of channels (Hamalainen et al., 1993). We sought to determine how many sensors were required to prevent performance degradation. Likewise, increased amount of data points usually leads to improved estimation of unknown quantities and therefore we tested performance with less data available.

To test the effect of the number of sensors, simulations were created similarly to those discussed above with three uncorrelated sources. Two values of SNIR were created using real brain noise: 0 dB and −10 dB. Ten different realizations of source time course and orientation were tested for each case. All simulations discussed previously were created using the full 275 channel array from the CTF system. We tested the method's performance when only a random subset of sensors were selected, using 150, 74, 37, 25 or 15 sensors.

To test the effect of the number of data points, the full set of 275 channels were used but reduced the available amount of data points. All previous simulations have used 700 total data points, where 62.5% were in the post-stimulus period. We kept the ratio of data points in the post-stimulus period the same but reduced the total number to 300, 200, 150, 100, or 50 points.

Collection of real data

Several real datasets were analyzed with the proposed method and compared to existing methods. For all data, the 275-channel CTF MEG System in a magnetically shielded room was used to collect data. All healthy subjects and one patient with epilepsy gave written, informed consent to participate in each study, according to UCSF institutional review board approval.

Auditory

The first dataset was obtained by presenting 120 repetitions of a 1-kHz tone binaurally to a healthy subject, at an intertrial interval of 1.4 s. The trials were averaged locked to stimulus onset. This auditory stimulus is known to invoke bilateral auditory cortex to be active simultaneously, causing problems for the MVAB's ability to localize the auditory sources.

Somatosensory

We next examine a somatosensory dataset in which the localization of primary somatosensory cortex is relatively easy for all methods when many trials are available to average. A small diaphragm was placed on the subject's right index finger and was driven by compressed air. The stimulus was given 256 times every 500 ms. However, if we limit the available data to only a small subset of trials, the lower SNR can become limiting for all methods. We first applied SAKETINI, MVAB and sLORETA to the average of all 256 trials to assess performance for the standard (high) SNR case. We then applied all three methods to the average of only the first 5 trials. To further test if the performance was consistent across other sets of just 5-trial averages, we applied the three methods to the 5-trial average of trials 6–10, 11–15, and 16–

20. We then averaged the results of these four different results. Any location found consistently will show up in the average.

Epileptiform spikes

We next tested the proposed method on its ability to localize interictal spikes obtained from a patient with epilepsy. No sensory stimuli were presented to this patient in this dataset, which was collected in the same MEG device described above. A Registered EEG/Evoked Potential Technologist marked segments of the continuously collected dataset which contained spontaneous spikes, as well as segments that clearly contained no spikes. One segment of data with a spike marked at 400 ms, as well as three additional spikes in the 800 ms segment, was used here as the "post-stimulus" period and a separate, spike-free, segment of equal length was used as the "pre-stimulus" period.

Results

We first report results from an example simulation. Then we describe performance for the averages of the simulations discussed above, including varying the amount of sensors or time points available. We finish by demonstrating performance in the real datasets discussed above.

Simulation results

Individual example simulation results

We first examine localization error for a single active source placed randomly within the voxel grid space. Fig. 2 shows localization error of SAKETINI relative to the MVAB. The largest peak in the SAKETINI likelihood map was found and the distance from this point to the true source was recorded. Each data point is an average of 20 realizations of the source configuration, with error bars showing standard error. This simulation was performed for a variety of SNIR's and for both cases of simulated and real brain
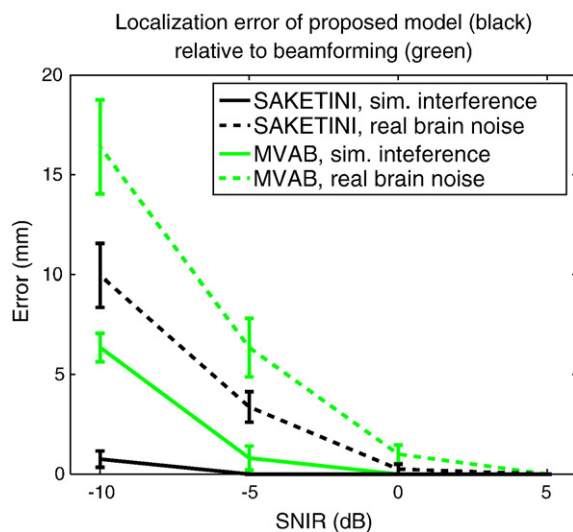

Fig. 2. Localization error of SAKETINI compared to MVAB. A single source was placed randomly in a realistic brain volume and its SNIR was varied. Error was measured as the three-dimensional distance from the maximal peak in the reconstruction to the true location.

noise. This figure clearly shows that the error in localization is smaller for SAKETINI (black) than for MVAB (green).

We next examine the performance of the proposed model when three active sources are present. Fig. 3 shows performance in one example with three uncorrelated sources. On the left is the likelihood map and on the right are the estimated time courses (green) of the three sources compared to the true time courses (blue). The likelihood map correctly locates all three sources and the time courses are estimated well.

We further test performance when two of the three sources are highly correlated in time. In one example simulation, Fig. 4 shows the SAKETINI likelihood map on the left compared to the MVAB power map on the right. The likelihood map detects the presence of all three sources, whereas the MVAB fails to find the left source and instead finds a false source in the middle between the two correlated sources.

Simulation results: average performance

We now show performance of the proposed methods according to the metrics of A' (area under ROC curve) and time course estimation, as defined in the Methods. Fig. 5(a) plots A' for each method, for each value of source correlation and SNIR, and for both types of interference (simulated and real). SAKETINI (black lines) has a larger average A' than MVAB (green) and sLORETA (red) for nearly all values of SNIR and source correlation, indicating overall better localization ability.

The other main test of performance was the ability to estimate the source time course. The estimated time courses for all three methods were obtained from the true source locations, regardless of whether their respective localization maps found that source as a "hit." The correlation of the true time course with the estimated time course was computed for each simulation and the averages are plotted in Fig. 5(b). SAKETINI estimates the time course better than MVAB for all simulation types and better than sLORETA in nearly all simulation types.

Simulation results: effects of amount of data

Fig. 6 shows simulation performance resulting from reduced number of sensors. Fig. 6(a) shows the A' metric for localization performance for SAKETINI relative to MVAB and sLORETA, where Fig. 6(b) shows estimation accuracy of the source time course. The top row in both (a) and (b) is for SNIR = 0 dB and the bottom row is for SNIR = −10 dB. SAKETINI's performance is not degraded for as few as 74 sensors for both values of SNIR and can even perform well using as few as 25 sensors for SNIR = 0 dB.

Fig. 7 tests the proposed method against the other methods with decreased number of time points available. Fig. 7(a) shows that SAKETINI has a larger A' value relative to MVAB and sLORETA for all amounts of data tested and for both SNIRs. However, SAKETINI's ability to estimate the source time course (Fig. 7(b)) decreases with less than roughly 150 data points total, even though MVAB and sLORETA do not show this sensitivity. Like other minimum-norm techniques, sLORETA is not data-dependent, and thus is not expected to show sensitivity to number of data points. The MVAB is dependent on the data to provide an estimate of the data covariance matrix. Since the simulations in both top and bottom rows are with relatively high noise (SNIR = 0 dB and −10 dB, respectively), the data covariance estimate might not change much with decreased data, since it is already noisy (note the time course correlation does not reach above 0.5 for any
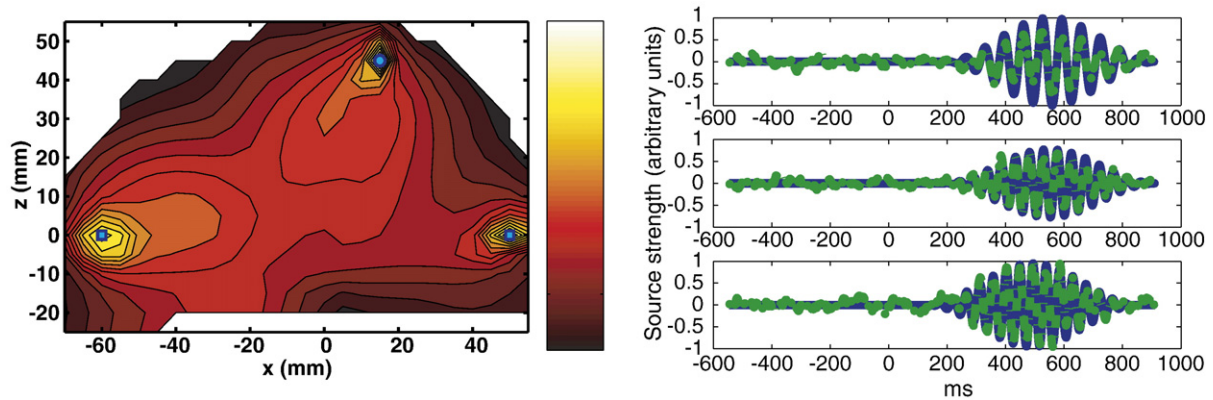
Fig. 3. SAKETINI's performance in simulated example with three uncorrelated sources. (Left) Blue circles and squares indicate location of true sources. Intensity of map corresponds to normalized likelihood map. (Right) Blue lines indicate simulated time series for each of the three source locations. Green lines indicate proposed method's estimates of the source time series at those three locations.

number of data points tested at SNIR = − 10 dB). However, SAKETINI relies on an initial estimate of interference sources from the pre-stimulus data; perhaps this estimate is not as robust when only 30 or less pre-stimulus data points are present to estimate the 30 interference sources projected to 275 sensors, allowing for some cross-talk of interference with "evoked" time estimates.

Results from real data

Auditory

The well-characterized auditory-evoked M100 peak was observed and the sensor map for this latency is shown in Fig. 8(a). Bilateral auditory cortex is clearly activated. As mentioned above, MVAB performance deteriorates when neural sources are highly correlated in time and, in general, bilateral auditory cortex is known to be correlated (Dalal et al., 2006). Fig. 8(b) shows the MVAB performance on this dataset: the MVAB fails to find either left or right auditory cortex and instead finds a false source in the center. However, both SAKETINI and sLORETA are able to find bilateral auditory cortex (Figs. 8(c) and (d)). Additionally shown in Fig. 8(d) are the time courses of the SAKETINI source estimates extracted from the peak voxel on each side. The correlation of the left with the right source estimates was found to be 0.93.

Somatosensory

The left panel of Fig. 9(a) shows typical somatosensory-evoked MEG data with the largest peak at 50 ms, expected to be coming from primary somatosensory cortex in the posterior wall of the central sulcus. The next three panels of Fig. 9(a) show localization performance of SAKETINI, MVAB and sLORETA. All three methods accurately localize activity to the contralateral primary somatosensory cortex. However, performance changes when only 5 trials are used in the average. The left panel of Fig. 9(b) shows the sensor data averaged over trials 1–5 of the same somatosensory dataset. The next three panels of Fig. 9(b) show errors in localization in all methods. SAKETINI shows the least error, relative to the peak location found using all 256 trials. We note that other averages of 5 trials showed varied performance, but that, when averaging four different sets of 5-trial averages together, SAKETINI showed localization closest to primary somatosensory cortex, as shown in Fig. 9(c), whereas MVAB and sLORETA mislocalize this source.

Epileptiform spikes

Fig. 10 shows the proposed method's performance on estimating single spikes relative to a spike-free "pre-stim" period. Fig. 10(a) shows the raw sensor data for the segment containing the marked spike. Fig. 10(b) shows the location of the equivalent-current dipole (ECD) fit to 20 spikes from this patient; the centroid of this cluster of
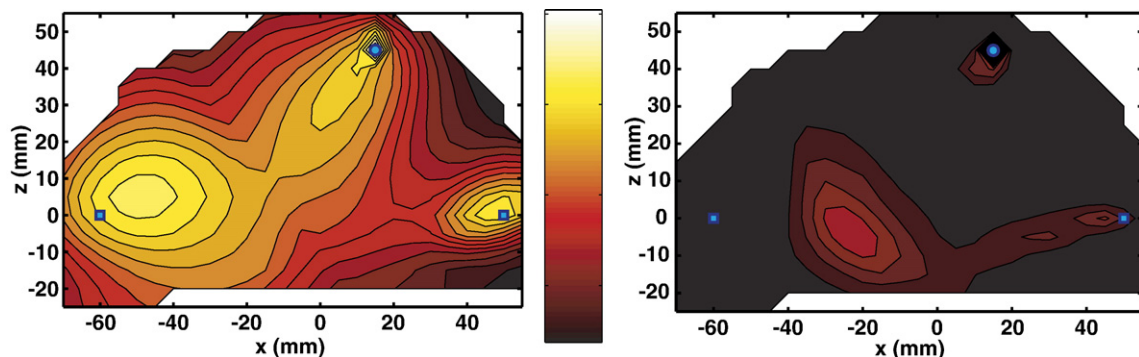

Fig. 4. Example of SAKETINI and MVAB for correlated source simulation. Blue squares indicate source locations that are highly temporally correlated. Blue circle indicates simulated source location that was not correlated with the other two sources. (Left) SAKETINI likelihood map. (Right) MVAB power map.
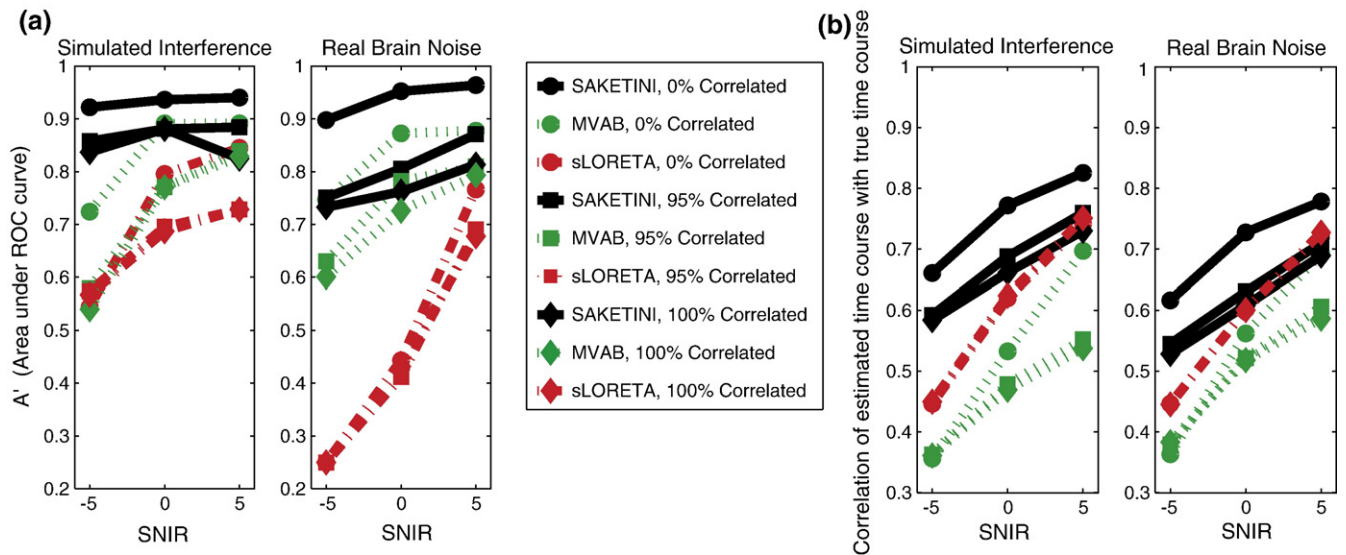
Fig. 5. Performance of SAKETINI relative to MVAB and sLORETA for variety of simulated datasets. Each data point is an average of 40 simulations, consisting of two different source locations and either a fixed or rotating source orientation. Standard errors were less than 0.05 for all points (not shown). (a) A measure of area under ROC curve A' is plotted for the three methods as a function of SNIR for both simulated and real brain interference. See text for discussion of the A' metric. (b) The correlation of the estimated with the true time course is plotted for each method.

ECD fits would normally be used clinically. Fig. 10(d) shows the SAKETINI likelihood map based on the data in Fig. 10(a); the peak is in clear agreement with the standard ECD localization. The maximum voxel of the SAKETINI likelihood map is 6.3 mm from the ECD fit of that particular spike and is 3.2 mm from the ECD cluster centroid. Fig. 10(c) shows the time course estimated for the likelihood spatial peak. The spike at 400 ms is clearly seen; this cleaned waveform could be of use to the clinician in analyzing peak shape (information usually not available from the ECD). Finally, Figs. 10(e) and (f) show a source time course from a randomly selected location far from the epileptic spike source (shown with cross-hairs on bottom right plot), in order to show the low noise level

and to show lack of cross-talk onto source estimates elsewhere. In addition to this single spike analysis with SAKETINI, the sensor data of all 20 spikes were first averaged, then examined by both ECD and SAKETINI.

Typically this average is only used clinically if the ECD of single spikes gives large parameter fit errors. The difference between the ECD and SAKETINI localizations based on this average spike data was only 3.1 mm. Both of these also corresponded well to the 20-spike ECD cluster centroid mentioned above, with differences of only 4.7 mm (ECD) and 3.1 mm (SAKETINI).

Furthermore, the cluster of 20 ECD localizations corresponded well to the electrocorticography mapping performed in this patient.
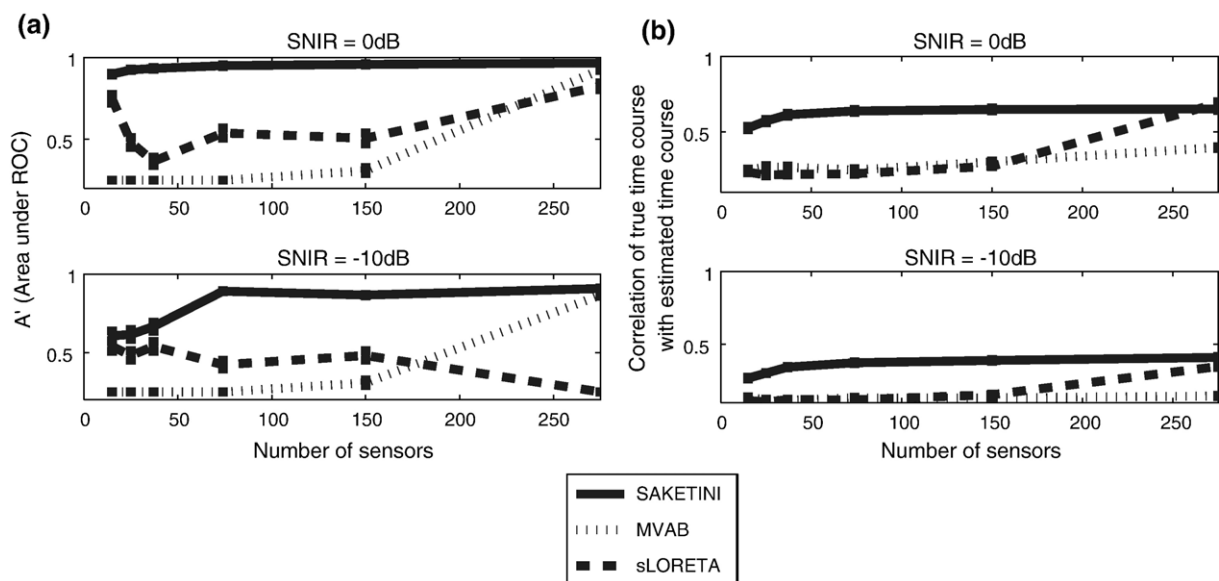

Fig. 6. (a) A' and (b) time course estimation as a function of the number of MEG sensors for simulated data with 3 uncorrelated sources. The top row shows SNIR=0 dB and bottom row shows SNIR=−10 dB using real brain noise. Error bars represent standard error.
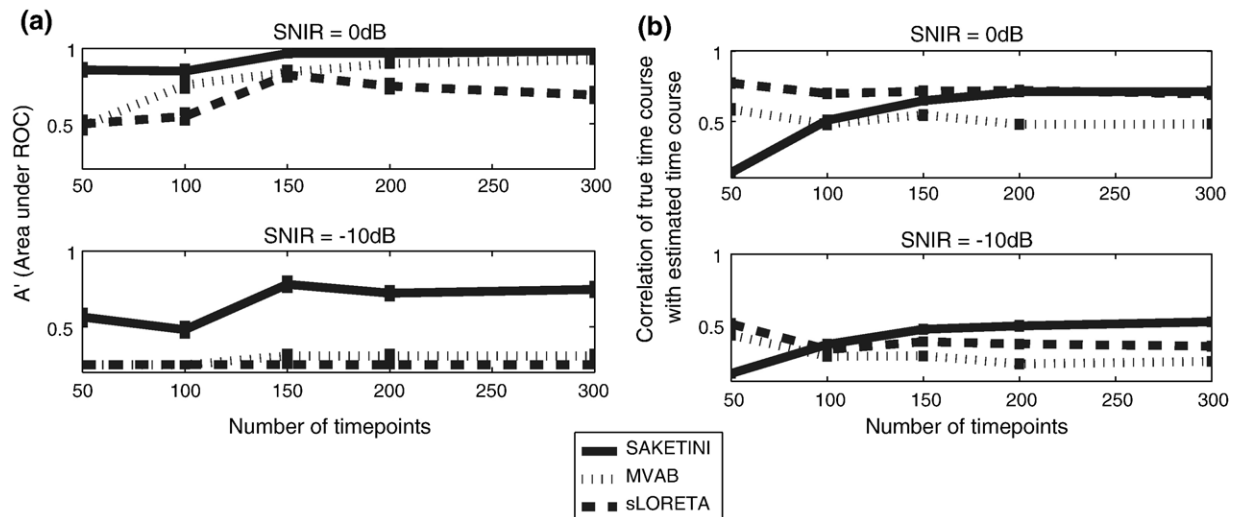
Fig. 7. (a) *A′* and (b) time course estimation as a function of the number of total data points for simulated data with 3 uncorrelated sources. The top row of each shows SNIR = 0 dB and bottom row shows SNIR = −10 dB using real brain noise. Error bars represent standard error.

Intraoperative electrodes were spaced 1 cm apart and the cluster resided within this 1 cm error. This brain region was chosen for resection and this patient was seizure-free after the operation. This indicates that the brain region localized pre-surgically and noninvasively by the ECD cluster (and consistent with the SAKETINI localization) correctly specified the epileptogenic zone.

Discussion

We have described the proposed method in a graphical model framework, which is a powerful and flexible technique for describing probabilistic dependencies between observed and unobserved quantities. We have chosen a scanning-based method
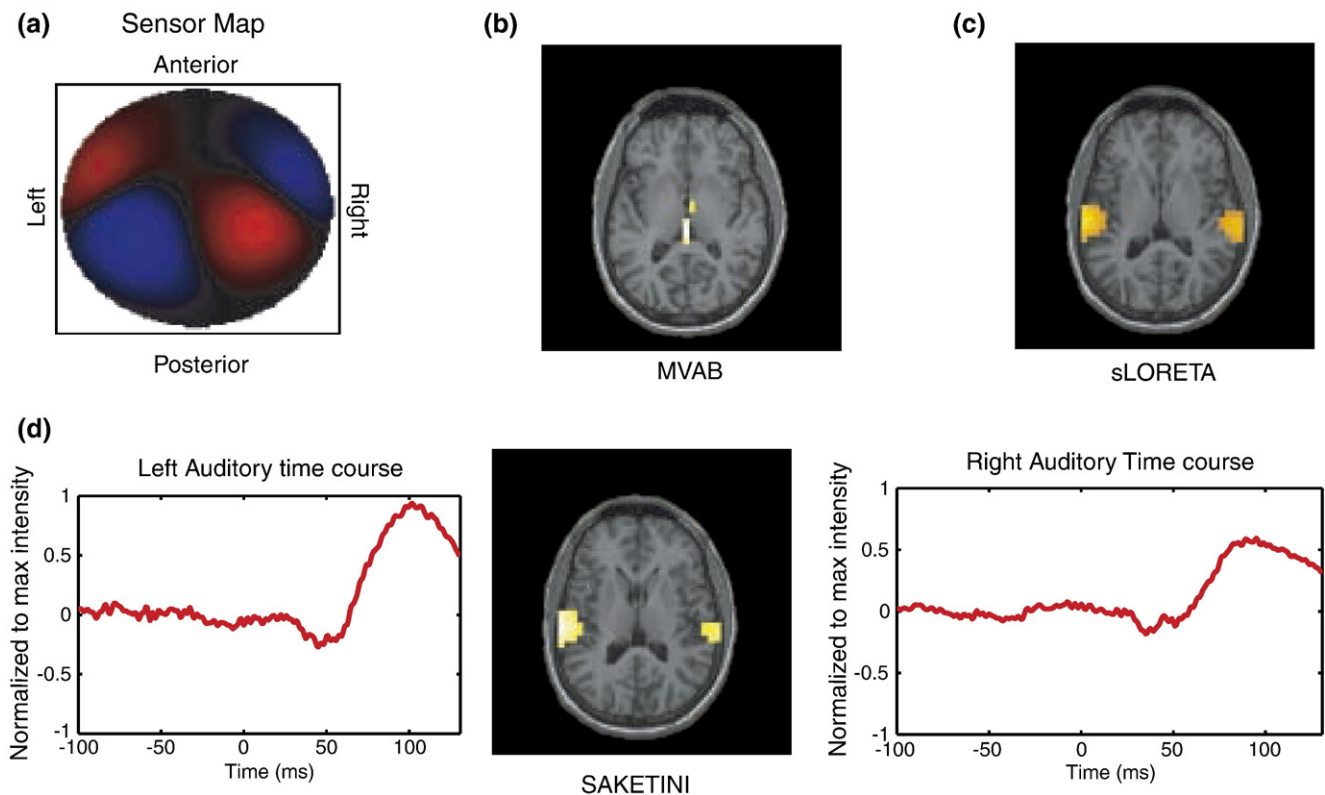

Fig. 8. Performance of methods using a real auditory-evoked MEG dataset. (a) Sensor map showing bilateral activation at M100 latency. (b) MVAB power map, showing failure to localize sources to auditory cortex. (c) sLORETA power map localizes bilateral auditory cortex. (d) SAKETINI likelihood map localizes activity in bilateral auditory cortex. Time series from left and right peaks are also shown, where intensity is normalized to maximum value across all voxels and time points.
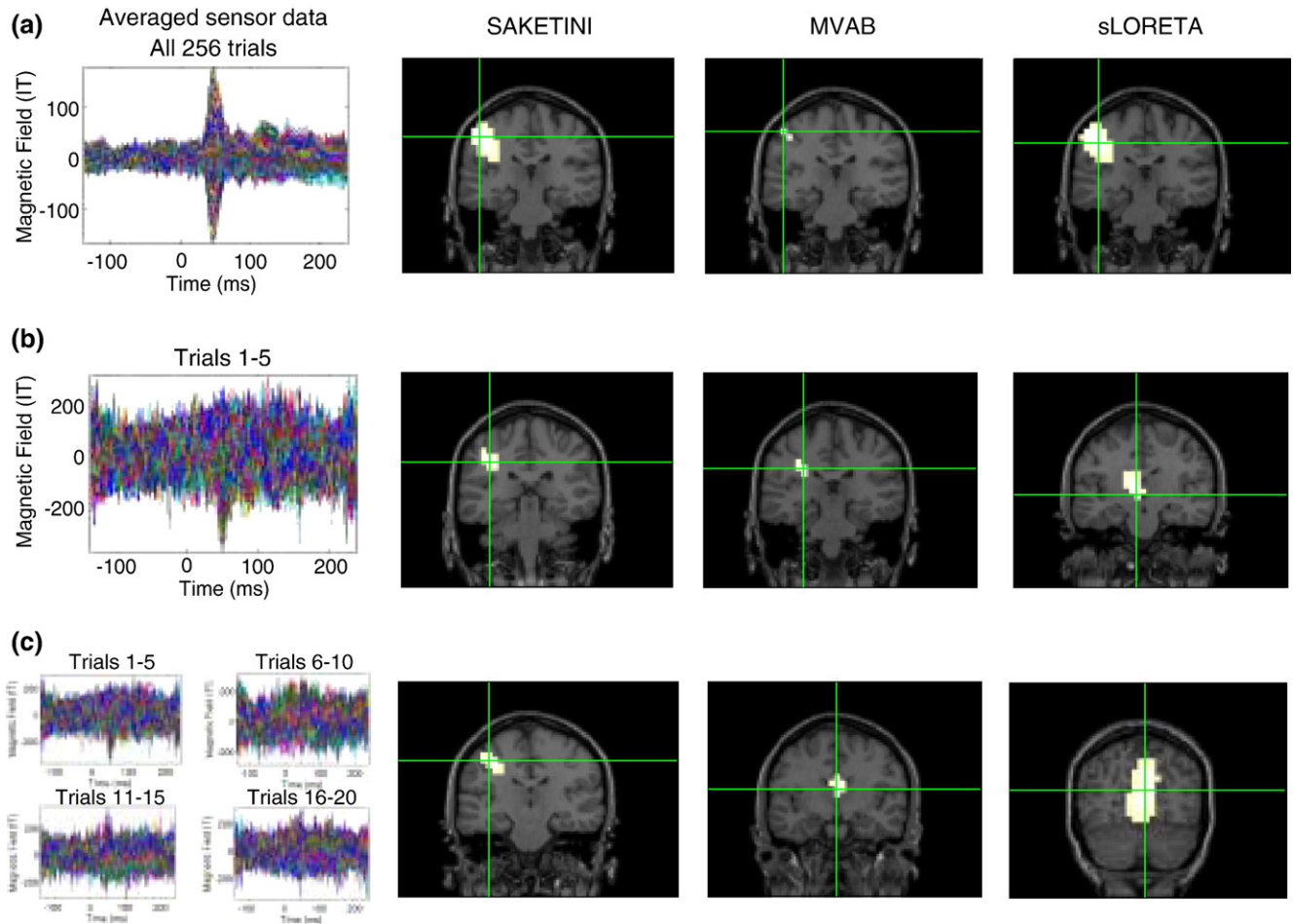
Fig. 9. Performance of methods using real somatosensory data as a function of the number of trials. Left column shows sensor data averaged over varied number of trials, while remaining columns show localization performance of SAKETINI, MVAB and sLORETA. Row (a) shows performance of the three methods applied to the average of all 256 trials. Row (b) shows the localization performance to the average of only the first 5 trials. In order to show performance over other subsets of 5-trial averages, the spatial maps in row (c) are spatial averages of the localization of 4 different 5-trial averages. See Methods for details. Crosshairs in localization maps show peak location within "active" voxels at the slice of peak location, where the threshold for "active" was defined at 90% of the maximum for all maps.

to formulate the problem, rather than to solve the full tomographic problem which is very ill parameterized. The variational Bayesian Expectation Maximization algorithm is used to solve for the values of the unknown quantities that maximize the marginal log likelihood. A map of the maximized likelihood is used to depict localized sources and estimates of their time course are extracted from those spatial peaks.

We have shown that the proposed model is not as sensitive to temporally correlated sources as the standard formulation of MVAB. However, it is possible to reduce the MVAB's dependence on correlated sources through a modified weight matrix computed subject to additional constraints, if a rough idea of the location of sources is known (Dalal et al., 2006).

Model order and initialization

As the number of MEG and EEG channels has increased in recent years, the ability to accurately localize sources throughout the brain has increased (Vrba et al., 2004). However, performing calculations of high-dimensional data, such as inverting a data

covariance matrix, becomes more difficult and can lead to errors. Meanwhile, the dimensionality of the underlying neural activity remains the same. Thus, many variations of PCA and ICA have been used on MEG/EEG data for removal of noise/artefactual components as well as for data dimension reduction (Jung et al., 2000; Ikeda and Toyama, 2000). Factor analysis also aims to reduce the dimensionality of the data to a linear mixture of factors that best account for the data while accounting for noise at the sensor level. We have used an extended version of factor analysis to partition which factors are event-related activity and which factors are background interference.

All methods which perform dimension reduction need a criterion for choosing the reduction number. Using PCA, a plot of eigenvalues can often give a reasonable intuition for the dimension of "signal" in the data. ICA has no ordering of components. In the method proposed here, there are two variables affecting model dimension: number of background interference factors (u) and number of non-localized evoked factors (x).

The model itself has two built-in ways of assisting in the dimension choice. First, the hyperparameters over the columns of
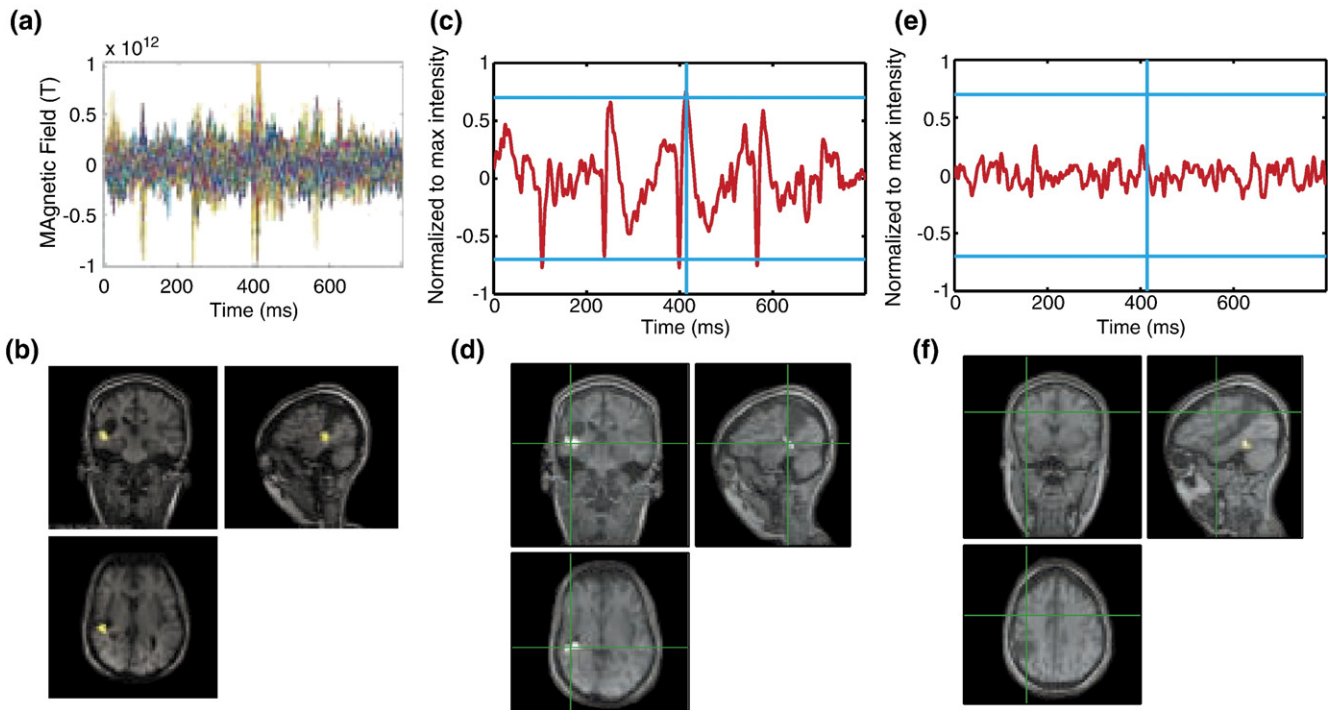
Fig. 10. SAKETINI applied to interictal spike data from a patient with epilepsy. (a) Segment of data contains several spikes marked by a technician. No stimulus was presented to the patient in this data. (b) Cluster of Equivalent Current Dipole (ECD) localizations to 20 spikes are shown from this patient. The middle column shows SAKETINI time course estimation (c) and likelihood map (d) for this segment of data, showing spatial agreement with ECD. The right column shows SAKETINI time course (e) from a voxel (f) randomly selected far away from the location of the spike localization, indicating lack of crosstalk of the spikes into normal tissue. The time course intensities in panels c and e are normalized to maximum value across all voxels and time points.

the mixing matrices A and B are updated in the M-step; thus their final values can be plotted and used as indication of which columns (and therefore which factors) contribute more, similar to a plot of eigenvalues for PCA. As long as the number of dimensions chosen is larger than the true dimension, the extra dimensions can be "zeroed out" by the model as those corresponding hyperparameters decrease.

A second way to examine model order is by learning the model for several values of model order and then comparing the final maximized likelihood value. In general, the greater the model order, the easier to explain the data, thus the likelihood will increase for greater model order. However, there is usually an inflection point in the plot of likelihood versus model order, again similar to a plot of eigenvalues for PCA, around which the ideal model order resides.

All methods which do not have a closed form solution require initialization of the values to be iteratively updated. We have found that choice of initialization can change the final results somewhat but not largely, and so we did not extensively examine these effects. After finding one method of initialization that worked well in a few test simulations, that set was used for all results shown. We note that in testing data with only sensor noise (no simulated or real brain interference at all), we do see a greater dependence on initialization values. This is worth mentioning since many source reconstruction algorithms are only tested on simulated sources plus sensor noise, thus comparison of this method with other methods on sensor-noise-only data should only be done with this in mind. However, since this type of data does not follow the assumptions of the model that the pre-stimulus

data has spatially covarying noise, and is therefore not realistic, we chose not to focus the results of this paper on this type of data.

A consideration of algorithms that iteratively update values is the longer computation time. SAKETINI infers the model values for each voxel and convergence typically occurs after about 10–20 VB-EM iterations, which takes roughly one second per voxel on a standard Linux personal computer with 2.0-GHz processor. Changing the number of channels or time points also affects computation time.

Extensions of model

The graphical model proposed here is one variation of a class of models that could be used to perform source reconstruction for MEG and EEG data. Certainly many aspects of this model can be modified in order to relax assumptions made here. One assumption is that the background interference sources are stationary across the pre-stimulus and post-stimulus periods. If background is non-stationary, the model could be modified to include a term similar to Φ which would influence the posterior update for $\bar{u}$. This would occur, for instance, if ongoing brain activity decreased as a result of the stimulus/event onset.

If the background is stationary, but the source activations vary drastically over the post-stimulus period, it might help to break up the post-stimulus data into partitions of relatively similar activity. Specifically, the data could be decomposed into time–frequency windows where inference of this model on each window would result in a separate likelihood map.

In all results shown, we used a single-shell spherical conductor model for generation of the forward field F. Certainly more sophisticated forward fields could be computed using multisphere methods or BEM models, which could further improve performance for all methods. Distributed sources could be modeled explicitly using a forward field computed from patch bases (Limpiti et al., 2006). However, errors in coregistration of MEG or EEG data to the structural MRI can affect any computation of forward field. If one wishes to restrict the source activity to one orientation based on cortical surface, errors in this segmentation can lead to errors in orientation estimation. Localization of EEG data has larger errors since tissue conductance values cause uncertainty in the forward field calculation. Thus, we can take advantage of the probabilistic framework of the proposed model by including a hyperparameter over the forward field F that would enable the algorithm to make small changes to the computed F, using the data to "learn" an improved estimate of the forward field.

One could further relax assumptions about the source and interference factors. They could be modeled by non-Gaussian distributions, although care should be taken in choice of distribution so that they could still be solved for analytically. The temporal progression of the factors could be explicitly modeled, such as using an auto-regressive model or Kalman filter.

The scanning framework of solving for one voxel's activity at a time can be relaxed. Spatial smoothness or spatial priors from other modalities, such as structural or functional MRI, could be incorporated. Furthermore, one is not limited to s_n in a single voxel; the above formulation holds for any P arbitrarily chosen dipole components, no matter which voxels they belong to, and for any value of P. Of course, as P increases the inferred value of Φ becomes less accurate, and one might choose to restrict it to a diagonal or block-diagonal form.

Related methods for MEG and EEG source reconstruction

Source localization algorithms can be broadly classified as parametric or tomographic, making different assumptions to overcome the ill-posed inverse problem. Parametric methods, including equivalent current dipole (ECD) fitting techniques, assume knowledge about the number of sources and their approximate locations. A single dipolar source can be localized well, but ECD techniques poorly describe multiple sources or sources with large spatial extent. Parametric methods could also involve multipole, line source and/or patch reconstructions (Schmidt et al., 1999; Yetik et al., 2005; Jerbi et al., 2002). Many parametric methods involve computation of noise and or signal covariances from data where several Bayesian approaches have been reported (Mosher and Leahy, 1998; Jun et al., 2005, 2006; Huizenga et al., 2002; Waldorp et al., 2002; de Munck et al., 2004; Bijma et al., 2005).

Tomographic methods reconstruct an estimate of source activity at grid points across the whole brain, either simultaneously or through scanning. Tomographic methods could also be classified as adaptive or non-adaptive. Of many tomographic algorithms, the adaptive beamformer has been shown to have the best spatial resolution and zero localization bias (Sekihara et al., 2002, 2005; Küçükaltun-Yildirim et al., 2006); however, it is sensitive to temporally correlated sources. Several tomographic algorithms use a Bayesian framework to explicitly incorporate prior assumptions about source configuration in the model (Sato et al., 2004; Phillips

et al., 2005; Mattout et al., 2006). Wipf et al. (2007) provide a unifying framework relating methods through their choice in specification of a neural source prior, whether fixed or learned from the data.

In contrast to many of these methods which reconstruct whole-brain source activity as MAP estimates, SAKETINI is a scanning based method that reconstructs a full posterior distribution of the source activity at each scanning voxel. The closest algorithm related to this approach is Dogandzic and Nehorai (2000), however, their algorithm does not make use of latent variable modeling and assumes that the temporal basis functions of the source activity at a scan location are known. SAKETINI does not make this assumption. Furthermore, instead of estimating a full-rank noise covariance, SAKETINI uses variational Bayesian factor analysis to estimate a dimension-reducing mixing matrix with number of columns much less than number of rows. Thus, more robust estimates of noise covariances can be made with fewer parameters to estimate.

Jun et al. (2005) create a multi-dipole model that explicitly computes the posterior for number of dipoles, location and orientation. They use prior terms for the temporal correlations of the sources, chosen based on the data. A prior on location or orientation can include information from fMRI or structural MRI, although they just use a uniform prior. The prior on number of dipoles roughly scales with the inverse of the number. They acknowledge the problem that the noise covariance has many parameters to estimate, so they use a Wishart prior distribution and estimate the noise posterior. They also integrate over the time courses to obtain the posterior for location, orientation and number of dipoles.

In this paper, we have not included any specific spatial prior information, although it certainly can be incorporated. SAKETINI only estimates one dipole at a time, by scanning through the voxel grid, thus estimation of the number of dipoles is not explicitly performed. Thresholding of the likelihood map can be viewed as a posterior probability map thresholding procedure.

Conclusion

We have described a novel probabilistic algorithm which performs source localization while robust to interference and demonstrated its superior performance over standard methods in a variety of simulations and real datasets. The model takes advantage of knowledge of when sources of interest are not occurring (such as in the pre-stimulus period of a evoked response paradigm). It learns the statistical structure of the interference sources from the pre-stimulus period and then can suppress these signals in the post-stimulus period. Set in a probabilistic framework, the use of prior probabilities and hyperparameters assists in the computation of unknown quantities. The flexibility of the model allows for many possible extensions.

Acknowledgments

References

Attias, H., 1999. Inferring parameters and structure of latent variable models by variational Bayes. Proc. 15th Conf. Uncertainty Artif. Intell., pp. 21–30.

Beal, M.J., May 2003. Variational algorithms for approximate Bayesian inference. PhD thesis, University of London.

Beal, M.J., Jojic, N., Attias, H., 2003. A graphical model for audiovixual object tracking. IEEE Trans. Pattern Anal. Mach. Intell. 25, 828–836.

Bijma, F., de Munck, J.C., Heethaar, R.M., 2005. The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. Neuro-Image 27, 402–415.

Brookes, M.J., Stevenson, C.M., Barnes, G.R., Hillebrand, A., Simpson, M.I.G., Francis, S.T., Morris, P.G., 2007. Beamformer reconstruction of correlated sources using a modified source model. NeuroImage 34, 1454–1465.

Bunch, P., Hamilton, J., Sanderson, G., Simmons, A., 1978. A free response approach to the measurement and characterization of radiographic-observer performance. J. Appl. Photogr. Eng. 4, 166–172.

Dalal, S.S., Zumer, J.M., Agrawal, V., Hild, K.E., Sekihara, K., Nagarajan, S.S., 2004. NUTMEG: a neuromagnetic source reconstruction toolbox. Neurol. Clin. Neurophysiol 52.

Dalal, S.S., Sekihara, K., Nagarajan, S.S., 2006. Modified beamformers for coherent source region suppression. IEEE Trans. Biomed. Eng. 53, 1357–1363.

Darvas, F., Pantazis, D., Kucukaltun-Yildirim, E., Leahy, R.M., 2004. Mapping human brain function with MEG and EEG: methods and validation. NeuroImage 23 (Suppl. 1), S289–S299.

de Munck, J.C., Bijma, F., Gaura, P., Sieluzycki, C.A., Branco, M.I., Heethaar, R.M., 2004. A maximum-likelihood estimator for trial-to-trial variations in noisy MEG/EEG data sets. IEEE Trans. Biomed. Eng. 51, 2123–2128.

Dogandzic, A., Nehorai, A., 2000. Estimating evoked dipole responses in unknown spatially correlated noise with EEG/MEG arrays. IEEE Trans. Signal 13–25.

Gelman, A., Rubin, D.B., 1996. Markov chain Monte Carlo methods in biostatistics. Stat. Methods Med. Res. 5, 339–355.

Ghahramani, Z., Beal, M., 2001. Graphical models and variational methods. In: Opper, M., Saad, D. (Eds.), Advanced Mean Field Methods—Theory and Practice. MIT Press.

Hamalainen, M., Hari, R., IImoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65, 413–497.

Huizenga, H.M., de Munck, J.C., Waldorp, L.J., Grasman, R.P.P.P., 2002. Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. IEEE Trans. Biomed. Eng. 49, 533–539.

Ikeda, S., Toyama, K., 2000. Independent component analysis for noisy data–MEG data analysis. Neural Netw. 13, 1063–1074.

Jerbi, K., Mosher, J.C., Baillet, S., Leahy, R.M., 2002. On MEG forward modelling using multipolar expansions. Phys. Med. Biol. 47, 523–555.

Jordan, M.I. (Ed.), 1999. Learning in Graphical Models. MIT Press, Cambridge, MA, USA.

Jordan, M., Ghaharamani, Z., Jaakkola, T., Saul, L., 1999. An introduction to variational methods for graphical models. In: Jordan, M. (Ed.), Learning in Graphical Models. MIT Press, Cambridge, MA, USA.

Jun, S.C., George, J.S., Paré-Blagoev, J., Plis, S.M., Ranken, D.M., Schmidt, D.M., Wood, C.C., 2005. Spatiotemporal Bayesian inference dipole analysis for MEG neuroimaging data. NeuroImage 28, 84–98.

Jun, S.C., George, J.S., Plis, S.M., Ranken, D.M., Schmidt, D.M., Wood, C.C., 2006. Improving source detection and separation in a spatio-temporal Bayesian inference dipole analysis. Phys. Med. Biol. 51, 2395–2414.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178.

Küçükaltun-Yildirim, E., Pantazis, D., Leahy, R.M., 2006. Task-based comparison of inverse methods in magnetoencephalography. IEEE Trans. Biomed. Eng. 53, 1783–1793.

Limpiti, T., Van Veen, B.D., Wakai, R.T., 2006. Cortical patch basis model for spatially extended neural activity. IEEE Trans. Biomed. Eng. 53, 1740–1754.

Mattout, J., Phillips, C., Penny, W.D., Rugg, M.D., Friston, K.J., 2006. MEG source localization under multiple constraints: an extended Bayesian framework. NeuroImage 30, 753–767.

Mosher, J.C., Leahy, R.M., 1998. Recursive MUSIC: a framework for EEG and MEG source localization. IEEE Trans. Biomed. Eng. 45, 1342–1354.

Nagarajan, S., Attias, H., Hild, K., Sekihara, K., 2005. Stimulus evoked independent factor analysis of MEG data with large background activity. Adv. Neural Inf. Process. 18.

Nagarajan, S.S., Attias, H.T., Hild, K.E., Sekihara, K., 2006. A graphical model for estimating stimulus-evoked brain responses from magnetoen-cephalography data with large background brain activity. NeuroImage 30, 400–416.

Neal, R.M. (Ed.), 1996. Bayesian Learning for Neural Networks. Springer, New York.

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electro-magnetic tomography (sLORETA): technical details. Meth. Find. Exp. Clin. Pharmacol. 24 (Suppl D), 5–12.

Phillips, C., Mattout, J., Rugg, M.D., Maquet, P., Friston, K.J., 2005. An empirical Bayesian solution to the source reconstruction problem in EEG. NeuroImage 24, 997–991011.

Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22.

Sato, M.-A., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage 23, 806–826.

Schmidt, D.M., George, J.S., Wood, C.C., 1999. Bayesian inference applied to the electromagnetic inverse problem. Hum. Brain Mapp. 7, 195–212.

Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., Miyashita, Y., 2001. Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. IEEE Trans. Biomed. Eng. 48, 760–771.

Sekihara, K., Nagarajan, S., Poeppel, D., Marantz, A., 2002. Performance of an MEG adaptive-beamformer technique in the presence of correlated neural activities: effects on signal intensity and time-course estimates. IEEE Trans. Biomed. Eng. 49, 1534–1546.

Sekihara, K., Sahani, M., Nagarajan, S., 2005. Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. NeuroImage 25, 1056–1067.

Snodgrass, J.G., Corwin, J., 1988. Pragmatics of measuring recognition memory: applications to dementia and amnesia. J. Exp. Psychol. Gen. 117, 34–50.

Vrba, J., Robinson, S.E., McCubbin, J., 2004. How many channels are needed for MEG? Neurol. Clin. Neurophysiol. 2004, 99.

Waldorp, L.J., Huizenga, H.M., Grasman, R.P.P.P., Böcker, K.B.E., de Munck, J.C., Molenaar, P.C.M., 2002. Model selection in electromag-netic source analysis with an application to VEFs. IEEE Trans. Biomed. Eng. 49, 1121–1129.

Wipf, D., Ramirez, R., Palmer, J., Makeig, S., Rao, B., 2007. Analysis of empirical Bayesian methods for neuroelectromagnetic source localiza-tion. In: Schölkopf, B., Platt, J., Hoffman, T. (Eds.), Adv. Neural Inf. Process. Syst., vol. 19. MIT Press, Cambridge, MA.

Yetik, I.S., Nehorai, A., Muravchik, C.H., Haueisen, J., 2005. Line-source modeling and estimation with magnetoencephalography. IEEE Trans. Biomed. Eng. 52, 839–851.

Zumer, J., Attias, H., Sekihara, K., Nagarajan, S., 2007. A probabilistic algorithm integrating source localization and noise suppression for MEG and EEG data. In: Schölkopf, B., Platt, J., Hoffman, T. (Eds.), Adv. Neural Inf. Process. Syst., vol. 19. MIT Press, Cambridge, MA.