A State-Space Modeling Approach for Localization of Focal Current Sources From MEG

Makoto Fukushima, Okito Yamashita*, Atsunori Kanemura, *Member, IEEE*, Shin Ishii, Mitsuo Kawato, and Masa-aki Sato

***Abstract*—State-space modeling is a promising approach for current source reconstruction from magnetoencephalography (MEG) because it constrains the spatiotemporal behavior of inverse solutions in a flexible manner. However, state-space model-based source localization research remains underdeveloped; extraction of spatially focal current sources and handling of the high dimensionality of the distributed source model remain problematic. In this study, we propose a novel state-space model-based method that resolves these problems, extending our previous source localization method to include a temporal constraint by state-space modeling. To enable focal current reconstruction, we account for spatially inhomogeneous temporal dynamics by introducing dynamics model parameters that differ for each cortical position. The model parameters and the intensity of the current sources are jointly estimated according to a Bayesian framework. We circumvent the high dimensionality of the problem by assuming prior distributions of the model parameters to reduce the sensitivity to unmodeled components, and by adopting variational Bayesian inference to reduce the computational cost. Through simulation experiments and application to real MEG data, we have confirmed that our proposed method successfully reconstructs focal current activities, which evolve with their temporal dynamics.**

***Index Terms*—Bayesian inference, magnetoencephalography (MEG), source localization, spatial focality, state-space model.**

I. Introduction

MAGNETOENCEPHALOGRAPHY (MEG) is a noninvasive brain imaging modality that has been widely used in neuroscience studies and clinical applications. One advantage of MEG is its high temporal resolution, which allows measurement of neural activity changes to the millisecond order. On the other hand, a main disadvantage of MEG is its ill-posedness, i.e., MEG cannot uniquely determine the locations of neural activities (current sources or dipoles) [1].

To elucidate the locations and intensities of neural current sources from magnetic fields external to the head, the source localization problem must be solved [2]. This poses an inverse problem requiring some *a priori* information or constraints on the current sources. Examples of such constraints on solutions are minimization of the total power [3] and maximization of spatial smoothness [4] and spatial sparseness [5].

Recently, the sparseness constraint that leads to spatially focal solutions has received attention and has been adopted in norm regularization methods [5]–[7], subspace methods [8]–[10], recursive weighed methods [11], [12], and Bayesian methods [13]–[16]. Spatially focal solutions are important because they are consistent with the well-accepted notion of *functional specialization*, which states that cortical activities specialized to specific stimulus attributes or experimental contexts are restricted to small regions [17]. In particular, the sparseness constraint provides good matches to brain activities evoked by external triggers such as visual, auditory, and somatosensory stimuli.

In addition to spatial constraints, several studies have extended source localization methods by incorporating temporal constraints [18]–[27]. A natural way to combine spatial and temporal constraints is to formulate the source localization problem by a state-space model. According to this framework, the measurement process is represented by the observation equation, while spatiotemporal dynamics of the current sources are represented by the state equation. Random walk (RW) like models [21]–[24] and autoregressive models, whose parameters are determined from measurements [25]–[27], are based upon this framework.

Although several source localization methods based on a state-space model have been proposed, none of them provide spatially focal estimates. This arises from a simplifying assumption that all current sources follow the same temporal evolution over the entire brain [21]–[24] or over a given cortical area [25]–[27]. Taking spatially inhomogeneous temporal evolution into account, the spatially focal currents with temporal constraints could be reconstructed.

In this paper, we extend our previous source localization method, the hierarchical variational Bayesian method (VB) [13], [28], to incorporate the evolution of focal current activities under their temporal dynamics based on the state-space modeling. We focus on linear temporal dynamics, as in existing

Manuscript received November 13, 2011; revised January 24, 2012; accepted February 17, 2012. Date of publication March 1, 2012; date of current version May 18, 2012. This work was supported by a contract with the National Institute of Information and Communications Technology, Japan, entitled, "Multimodal integration for brain imaging measurements," and a contract with the Ministry of Internal Affairs and Communications entitled, 'Novel and innovative R&D making use of brain structures' and by the Japan Society for the Promotion of Science (JSPS) through Grant-in-Aid for JSPS Fellows (23-3907) and Grant-in-Aid for Young Scientists (B) (23700281). *Asterisk indicates corresponding author.*

M. Fukushima is with the Graduate School of Information Science, Nara Institute of Science and Technology, Nara 630-0192, Japan, and also with ATR Neural Information Analysis Laboratories, Kyoto 619-0288, Japan (e-mail: mfukushima@atr.jp).

*O. Yamashita is with the ATR Neural Information Analysis Laboratories, Kyoto 619-0288, Japan (e-mail: oyamashi@atr.jp).

A. Kanemura and M. Sato are with ATR Neural Information Analysis Laboratories, Kyoto 619-0288, Japan (e-mail: atsu-kan@atr.jp; masa-aki@atr.jp).

S. Ishii is with the Graduate School of Informatics, Kyoto University, Kyoto 611-0011, Japan and also with ATR Neural Information Analysis Laboratories, Kyoto 619-0288, Japan (e-mail: ishii@i.kyoto-u.ac.jp).

M. Kawato is with ATR Computational Neuroscience Laboratories, Kyoto 619-0288, Japan, and also with the Graduate School of Information Science, Nara Institute of Science and Technology, Nara 630-0192, Japan (e-mail: kawato@atr.jp).

Digital Object Identifier 10.1109/TBME.2012.2189713

state-space methods, but introduce the dynamics model parameters in a spatially nonuniform manner, i.e., a current source at every cortical location is assumed to evolve under its individual temporal dynamics. The spatially focal current is reconstructed by estimating both the dynamics parameters and the intensities of current sources from data.

Since the assumed number of the current sources on the cortical surface is large (of order 10^3), our state-space model contains many degrees of freedom. As a result, the currents and the parameters are not readily determinable because of high sensitivity to unmodeled data components and high computational costs [29]. The sensitivity is reduced by introducing prior distributions of parameters that are common to all current sources. These prior distributions control the overall adaptability of each model parameter to the data. The computational cost is reduced by employing variational Bayesian inference [30]–[32], which yields an approximate posterior distribution of variables of interest. This approximation enables the complexity of the current source computation to grow linearly with the number of sources. By solving these two problems, we have developed the first state-space model-based source localization method that can estimate spatially focal activities.

II. Methods

A. Notation

The following notations are used throughout this paper. The probability distribution of a random variable x and the conditional probability distribution of a random variable x given y are represented by $P(x)$ and $P(x|\,y)$, respectively. $\mathcal{N}(\mathbf{x}|\,\bar{\mathbf{x}},\mathbf{\Sigma})$ denotes a vector random variable $\mathbf{x}$ following the multivariate Gaussian distribution with mean vector $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{\Sigma}$. This representation is also used for the univariate Gaussian distribution as $\mathcal{N}(x|\,\bar{x},\sigma)$. $\Gamma(x|\,\bar{x},\gamma)$ denotes a scalar random variable x following the gamma distribution with mean $\bar{x}$ and shape parameter γ (note that the scale parameter corresponds to $\bar{x}\,\gamma^{-1}$). As in [13], we refer to the shape parameter as the confidence parameter. $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$, respectively, denote the sets of real n-dimensional vectors and real $m\times n$ matrices. A diagonal matrix with diagonal elements $\mathbf{x}$ is denoted $\mathrm{diag}(\mathbf{x})$. The trace and the nth diagonal element of a matrix $\mathbf{X}$ are denoted $\mathrm{tr}(\mathbf{X})$ and $(\mathbf{X})_{n,n}$, respectively. For notational simplicity, $1:T$ is used in place of $1,\ldots,T$.

B. State-Space Modeling

In our state-space model, the process of observing magnetic fields from unobserved currents is described by the observation equation, and the temporal evolution of the currents is described by the state equation. These two equations act as spatiotemporal constraints upon the model.

1) Observation Equation: Let M, N, and T be the number of MEG sensors, current sources, and time samples, respectively. The observed magnetic fields $\mathbf{B}_t\in\mathbb{R}^M$ are generated from the current sources $\mathbf{J}_t\in\mathbb{R}^N$ on the basis of Maxwell's equations [33]:

$$\mathbf{B}_t=\mathbf{G}\mathbf{J}_t+\boldsymbol{\epsilon}_t \tag{1}$$

where $\mathbf{G}\in\mathbb{R}^{M\times N}$ is the lead-field matrix, $\boldsymbol{\epsilon}_t\in\mathbb{R}^M$ is an additive disturbance, and the subscript t ($1\le t\le T$) is the time index. In this paper, we set $M=400$, the number of sensors of the MEG system used in experiments (see Section III-E), and $N=2000$, yielding an average intersource distance of 6.51 mm. The lead-field matrix $\mathbf{G}$ is obtained by solving the so-called forward problem [34]. The disturbance $\boldsymbol{\epsilon}_t$ is assumed to follow a zero mean Gaussian distribution with covariance $\beta^{-1}\mathbf{S}$, where scalar β is a scaling parameter and $\mathbf{S}\in\mathbb{R}^{M\times M}$ is a scaled covariance matrix. $\mathbf{S}$ is normalized to satisfy $\mathrm{tr}(\mathbf{S})=M$ and is determined from prestimulus rest period measurements. Thus, this covariance structure represents summation of the measurement noise and background brain activities.

The observation equation is represented by the following likelihood function in our probabilistic formulation:

$$P(\mathbf{B}_t|\,\mathbf{J}_t,\beta)=\mathcal{N}(\mathbf{B}_t|\,\mathbf{G}\mathbf{J}_t,\beta^{-1}\mathbf{S}). \tag{2}$$

2) State-Equation: We model the temporal evolution of a current source by a first-order autoregressive [AR(1)] process, which differs at each cortical location. In particular, we consider the following special class of linear dynamics:

$$\mathbf{J}_t=\mathrm{diag}(\mathbf{a})\mathbf{J}_{t-1}+\boldsymbol{\omega}_t \tag{3}$$

where the vector $\mathbf{a}\in\mathbb{R}^N$ is the AR(1) parameter and $\boldsymbol{\omega}_t\in\mathbb{R}^N$ is system noise. We assume that each current source on each cortical location evolves independently of the others, i.e., $J_{n,t}=a_nJ_{n,t-1}+\omega_{n,t}$ where $J_{n,t}$, a_n, and $\omega_{n,t}$ are the nth elements of $\mathbf{J}_t$, $\mathbf{a}$, and $\boldsymbol{\omega}_t$, respectively. Although the linear model (3) is too simple to model complex brain dynamics, it is a significant improvement upon previous state-space methods, in which the temporal evolution is assumed uniform over the entire brain [21]–[24] or within a cortical area [25]–[27]. The system noise $\boldsymbol{\omega}_t$ is assumed to follow a Gaussian distribution $\mathcal{N}\big(\boldsymbol{\omega}_t|\,\mathbf{0},(\beta\mathrm{diag}(\mathbf{q}))^{-1}\big)$, whose variance also differs at each cortical location. Here, $\mathbf{q}\in\mathbb{R}^N$ is the parameter vector and the variance for the nth cortical site, denoted σ_n, is equal to $(\beta q_n)^{-1}$ [q_n is the nth element of $\mathbf{q}$]. The scaling parameter β is introduced to the system noise for technical convenience [13], [28], [35] during application of the variational Bayesian inference.

From a Bayesian point of view, specifying the state equation corresponds to setting the following prior distribution on the current sources:

$$P(\mathbf{J}_t|\,\mathbf{J}_{t-1},\beta,\mathbf{a},\mathbf{q})=\mathcal{N}\big(\mathbf{J}_t|\,\mathrm{diag}(\mathbf{a})\mathbf{J}_{t-1},(\beta\mathrm{diag}(\mathbf{q}))^{-1}\big). \tag{4}$$

C. Prior Distributions

To approach the source localization problem from Bayesian perspectives, we introduce prior distributions of the model parameters β, $\mathbf{a}$, and $\mathbf{q}$. For these parameters, the following conjugate priors [32] are used to simplify the derivation of their (approximate) posterior distributions. The prior distribution of the scaling parameter β is set as

$$P(\beta)\propto\beta^{-1} \tag{5}$$

which corresponds to a noninformative prior [32]. For prior distribution of the AR(1) parameter $\mathbf{a}$, we select a zero mean Gaussian distribution:

$$P(\mathbf{a}) = \prod_{n=1}^{N} \mathcal{N}(a_n \,|\, 0, \eta_0^{-1}) \tag{6}$$

where a_n is the nth element of $\mathbf{a}$, and η_0 is a hyperparameter that controls the reliability of this prior (the higher the η_0, the more reliable the prior). For the parameter $\mathbf{q}$, the prior distribution is assumed as

$$P(\mathbf{q}) = \prod_{n=1}^{N} \Gamma(q_n \,|\, q_0, \gamma_0) \tag{7}$$

where both q_0 and γ_0 are hyperparameters (the mean and confidence parameter of Gamma distribution, respectively). High γ_0 means high reliability of this prior. The hyperparameters η_0, q_0, and γ_0 are shared by all current sources. Their values are determined as described in Section II-F. It should be noted that these prior distributions are dynamical extensions of the automatic relevance determination (ARD) prior [36] that gives nondynamical focal estimation [13], [28]. Indeed, when all entries of $\mathbf{a}$ are set to 0 (i.e., $\eta_0 \to \infty$), (4) and (7) reduce to the hierarchical prior distributions of ARD.

The noninformative priors for $\mathbf{a}$ and $\mathbf{q}$, derived by setting both η_0 and γ_0 to 0, were used in the simulations in Sections III-A, III-B, and III-C, whereas the informative priors for these parameters (nonzero η_0 and γ_0) were used in the more realistic simulations in Section III-D and the real data analysis in Section III-E. Choosing informative priors in the realistic data setting reduces the effects of modeling errors on parameter estimation. Since the dimension of $\mathbf{a}$ and $\mathbf{q}$ is large, these parameters are highly sensitive to modeling errors, such as discrepancies between the assumed model and the data generation process. Thus, to render our estimation robust to these errors, we reduce the sensitivity by constraining the parameter space through the informative priors.

D. Joint Probability Distribution

The likelihood function and all of the prior distributions comprise the joint probability distribution:

$$P(\mathbf{B}_{1:T}, \mathbf{J}_{1:T}, \beta, \mathbf{a}, \mathbf{q}) = \left\{ \prod_{t=1}^{T} P(\mathbf{B}_t \,|\, \mathbf{J}_t, \beta) \right\} P(\mathbf{J}_1 \,|\, \bar{\mathbf{J}}_0, \beta, \mathbf{a}, \mathbf{q}) \times \left\{ \prod_{t=2}^{T} P(\mathbf{J}_t \,|\, \mathbf{J}_{t-1}, \beta, \mathbf{a}, \mathbf{q}) \right\} P(\beta) P(\mathbf{a}) P(\mathbf{q}) \tag{8}$$

whose graphical representation is shown in Fig. 1. The initial current $\bar{\mathbf{J}}_0$ is set as a zero vector rather than a vector of random variables, as appropriate for analysis of data whose starting point immediately follows a rest period. This setting was used for all analyses in Section III. A minor alteration to the estimation algorithm would allow us to estimate $\bar{\mathbf{J}}_0$ also.

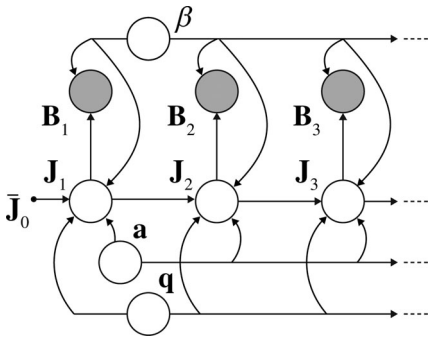

Fig. 1. Graphical representation of the joint probability distribution. Each circle node represents a random variable (gray: observed, white: unobserved) and each arrow indicates a dependence between variables. $\bar{\mathbf{J}}_0$ (black dot) is constant.

E. Variational Bayesian Inference

Estimates of the currents $\mathbf{J}_{1:T}$ and the parameters β, $\mathbf{a}$, and $\mathbf{q}$ are obtained from Bayesian inference. The objective of the inference is to derive the joint posterior distribution $P(\mathbf{J}_{1:T}, \beta, \mathbf{a}, \mathbf{q} \,|\, \mathbf{B}_{1:T})$. However, the exact calculation of the posterior is often intractable when a model structure is complicated, as in [37], and this is also the case for our model. To solve this problem, we employ variational Bayesian inference [30]–[32], which yields an analytically tractable but approximate posterior distribution, and which is widely used in such situations. For notational simplicity, the set of currents and parameters is denoted by $X = \{\mathbf{J}_{1:T}, \beta, \mathbf{a}, \mathbf{q}\}$ and one of its elements by $x \in X$. Within the variational Bayesian framework, the approximate form of joint posterior distribution $Q(X)$ makes the following independence assumption among the elements of X:

$$Q\left(X = \{\mathbf{J}_{1:T}, \beta, \mathbf{a}, \mathbf{q}\}\right) = \left\{ \prod_{t=1}^{T} Q(\mathbf{J}_t) \right\} Q(\beta) Q(\mathbf{a}) Q(\mathbf{q}). \tag{9}$$

Here, $Q(x)$ [a factor of $Q(X)$] is an approximate posterior distribution for each variable of interest. This independence assumption allows computation of the joint approximate posterior $Q(X)$ with a computational complexity that grows linearly with number of sources (see Appendix for details). Note that our independence assumption is stronger than that in the variational Kalman smoother (VKS) algorithm [37], [38], which does not assume independence within $Q(\mathbf{J}_{1:T})$.

All of the approximate posteriors $Q(x)$ are alternately and iteratively updated to decrease the Kullback–Leibler divergence (KLD) between the approximate and the true posteriors $\mathrm{KL}(Q(X) \| P(X \,|\, \mathbf{B}_{1:T}))$. Minimizing KLD is equivalent to maximizing the free energy, i.e.,

$$F(Q(X)) = \log P(\mathbf{B}_{1:T}) - \mathrm{KL}(Q(X) \| P(X \,|\, \mathbf{B}_{1:T})). \tag{10}$$

The approximate posteriors are obtained by iteratively maximizing the free energy with respect to $Q(x)$ using calculus of variations

$$\frac{\partial F(Q(X))}{\partial Q(x)} = 0$$
$$\Rightarrow \log Q(x) = \langle \log P(\mathbf{B}_{1:T}, X) \rangle_{Q(X \setminus x)} + \text{const} \quad (11)$$

where $\langle \cdot \rangle_{Q(X \setminus x)}$ is the expectation of the approximate posterior of X other than x, and const is a term independent of x. The iterative maximization continues until the free energy converges. Through these iterations, we obtain the $Q(X)$ closest to $P(X|\,\mathbf{B}_{1:T})$ in terms of KLD.

The estimation process of the proposed method is summarized as follows.

1) Set the initial values required for updating $Q(\mathbf{J}_1)$, and also set the hyperparameters.
2) $\mathbf{J}$-step: For $t = 1{:}T$, update $Q(\mathbf{J}_t)$.
3) β-step: Update $Q(\beta)$, and compute the free energy.
4) $\mathbf{a}$-step: Update $Q(\mathbf{a})$.
5) $\mathbf{q}$-step: Update $Q(\mathbf{q})$.
6) If the change of the free energy relative to the value of the free energy at the previous step is smaller than 10^{-9}, then exit; otherwise, go to 2.

Each update rule is derived from (11) by substituting the corresponding variable of interest into x. Detailed derivation of the update rules is presented in Appendix.

F. Initial Values and Hyperparameters

Since the proposed method adopts an iterative estimation algorithm, it requires initial values for the first update of $Q(\mathbf{J}_1)$ in the $\mathbf{J}$-step. In our algorithm, the mean and variance of $Q(a_n)$ and the mean of $Q(q_n)$ must be initialized [see (15)–(17) in Appendix]. To avoid the need to initialize the means of $Q(\mathbf{J}_{2:T})$, the mean and variance of $Q(a_n)$ are initialized to 0 for all n. The initial mean of $Q(q_n)$ is set as the estimated inverse variance parameter of VB described in [28], which is equivalent to the mean of $Q(q_n)$ when no temporal evolution is assumed (i.e., when all a_n are fixed at 0).

In addition, three hyperparameters must be specified. Two hyperparameters η_0 and γ_0, both of which control reliability of the priors, are selected from an appropriate range of values explored in a grid-search manner (see Section III-D). As noted previously, η_0 and γ_0 are set to 0 during the estimations of $\mathbf{a}$ and $\mathbf{q}$ with the noninformative priors (see Sections III-A–III-C). The remaining hyperparameter q_0 is determined such that the product $(\beta q_0)^{-1}$, corresponding to the prior for the system noise variance σ_n for any n, becomes the average variance of all current sources over the cortex. When specifying this variance, we use the currents estimated by the minimum norm estimation (MN) [3] and we set β^{-1} as the variance of the observation noise estimated by MN, with the optimal regularization parameter of MN derived from variational Bayesian inference.

G. Required Computation Time

All analyses were performed in MATLAB 7.5.0. Approximately 15 min were required for convergence of the iterative process using a standard multicore PC (2 cores, 2.66-GHz CPU, and 8-GB RAM) for real data analysis in Section III-E ($M = 400$, $N = 2000$, $T = 300$, and $\eta_0 = \gamma_0 = 10^2$). In this case, 89 iterations were required to maximize the free energy. When the free energy calculation was bypassed, the computation time was reduced to around 3 min only. Note that the number of iterations required to converge depends on the values of the hyperparameters. For smaller η_0 and γ_0, more iterations are needed. For the first simulation in Section III-A, which assumes noninformative priors, 926 iterations were required for free energy convergence.

H. Previous Methods for Comparison

We present here two previously proposed methods against which our method is compared in Sections III-A and III-E. This comparison reveals the effects of temporal constraints on the estimated currents, and also the effects of spatially inhomogeneous and data-adaptive dynamics parameters.

To meet the first purpose, our model is compared with VB [13], [28]. The model in VB can be viewed as a specialized form of our model with all entries of $\mathbf{a}$ set to 0. Under this simplification, the prior distribution of the currents is not dependent on past and future currents, thereby removing temporal constraints.

For the second purpose, our model is compared with that in [22], (one of RW models). Our model reduces to the RW model when all entries of $\mathbf{a}$ are set to 1 (constant) and the vector $\mathbf{q}$ is replaced by a scalar q. In this paper, all entries of $\mathbf{a}$ are set to 0.9 (i.e., less than 1), to avoid gradual RW deviation of the amplitudes from the baseline (slight changing the value of $\mathbf{a}$ does not substantially alter the results).

I. Local Spatial Smoothing

For all methods, a local spatial smoothness constraint is imposed on the current sources by applying a spatial smoothing filter. The relationship between the smoothed currents $\mathbf{J}_t$ and the unsmoothed currents $\mathbf{Z}_t$ is assumed to be

$$\mathbf{J}_t = \mathbf{W}\mathbf{Z}_t \quad (12)$$

where $\mathbf{W}$ represents a Gaussian smoothing filter with 8-mm full-width at half-maximum as used in [13] and [28]. By substituting (12) into (1), the observation process is rewritten as

$$\mathbf{B}_t = \hat{\mathbf{G}}\mathbf{Z}_t + \boldsymbol{\epsilon}_t \quad (13)$$

where $\hat{\mathbf{G}} = \mathbf{G}\mathbf{W}$ is a smoothed lead-field matrix. To obtain the solutions $\mathbf{J}_{1:T}$, we first solve the source localization problem of $\mathbf{Z}_{1:T}$ with the smoothed lead-field matrix $\hat{\mathbf{G}}$ and, then, convert $\mathbf{Z}_{1:T}$ to $\mathbf{J}_{1:T}$ by computing (12).

III. Results

A. Simulation Study 1: Method Comparison

First, we conducted simulation experiments to compare the performance of the three source localization methods (the proposed method, VB, and RW). Time courses and cortical positions of the simulated active current sources are summarized in the leftmost column of Fig. 2. Three active sources were assumed, following time courses of nonsmooth with high amplitude, smooth with high amplitude, and smooth with low amplitude (hereafter denoted Sources 1, 2, and 3, respectively). Chosen positions of the three sources are indicated by the three

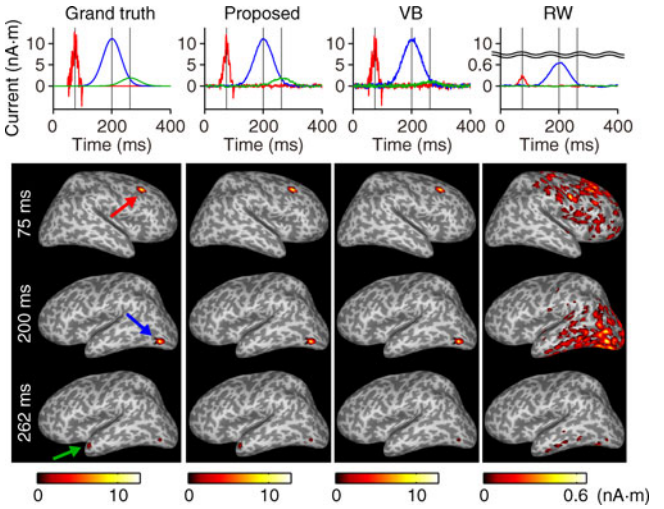

Fig. 2. Spatiotemporal profiles of the true and estimated currents in simulation study 1. The leftmost column illustrates the grand truth. The other columns (left to right) show the estimates from the proposed method, VB, and RW, respectively. Upper: time courses of the currents at the three active positions (Sources 1, 2 and 3 plotted in red, blue and green, respectively). Lower: spatial maps of the currents at peak time points 75, 200, and 262 ms. The estimated currents are mapped onto an inflated cortical surface, and the currents exceeding one-tenth of the maximum amplitude are shown.

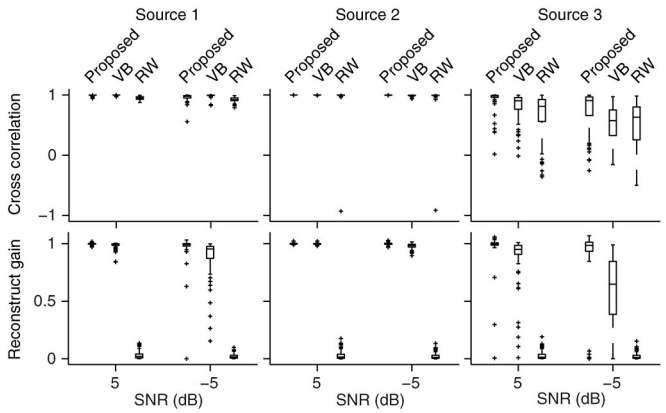

Fig. 3. Boxplot representation of cross correlation (upper) and reconstruction gain (lower) of the currents estimated by the proposed method, VB, and RW at SNRs ± 5 dB.

arrows in Fig. 2. Each current source was blurred with a local spatial smoothing filter (see Section II-I). Although the time course of Source 1 appears physiologically unrealistic, it usefully illustrates how nonsmooth temporal profiles affect our estimation results.

The simulated MEG, for which sampling frequency was assumed to be 1 kHz, was generated by multiplying the simulated currents with the lead-field matrix and then adding white Gaussian noise. The lead-field matrix was based on the anatomical structure of the cortex extracted from a T1 image of magnetic resonance imaging (MRI). The covariance matrix of Gaussian noise was determined from real MEG data recorded without a subject and its scale was adjusted to a signal-to-noise ratio (SNR)[1] -5 dB, similar to that of single-trial measurements of stimulus-evoked responses. In concordance with the real data analysis in Section III-E, the structure of this covariance was estimated from rest period measurements, which contained only Gaussian noise in this simulation.

The proposed method and VB successfully reconstructed all sources as spatially focal, whereas RW reconstructed them as spatially diffuse (see Fig. 2). The waveform of Source 1 was best estimated by VB. In contrast, it was slightly oversmoothed by the proposed method. Both methods reconstructed Source 2 similarly. For Source 3, the proposed method generated better amplitude estimates than VB. The amplitudes of all activities estimated by RW were suppressed, being below one-tenth of the assumed amplitudes.

To statistically evaluate the performance of the three methods, we repeated the above simulation 100 times for different positions of the three sources. We also considered two SNR cases: 5 dB and -5 dB. The cross correlation and the reconstruction gain were used as performance measures. The cross correlation is defined as the correlation between the time courses of the estimated and true currents. The reconstruction gain is the ratio of the estimated amplitude of the currents to the true one. Both measures were computed at the cortical locations of the true sources. The results are summarized as boxplots in Fig. 3. For Source 1, the cross correlation was highest in VB but the proposed method yielded the higher reconstruction gain. For Source 2, the estimates from the proposed method and VB were very similar. For Source 3, both cross correlation and reconstruction gain were closest to 1 for the proposed method (detailed investigation reveals that the lead-field norm [13] of the true source position, rather than the distance between sources, influences the estimation accuracy). The cross correlation of RW was comparable to those of the proposed method and VB, but RW yielded low reconstruction gains in all three cases.

To further understand the results of the proposed method, we investigated the relationship between the estimated currents and the estimated dynamics parameters [AR(1) parameter and system noise variance]. These estimates, resulting from the simulation data shown in Fig. 2, are presented in Fig. 4. The currents were nonzero at positions where either the AR(1) parameter or the system noise variance was estimated as nonzero. When the SNR was high (5 dB), the AR(1) parameter was estimated as nonzero only at the three positions of the grand truth. When the SNR was low (-5 dB), negative AR(1) parameter and large system noise variance were estimated at several positions of false positive currents. In both SNR cases, temporally smooth currents (Sources 2 and 3) assessed the AR(1) parameter as close to 1 and produced low system noise variance. For the nonsmooth current (Source 1), the AR(1) parameter was lower than for the previous two cases, while the system noise variance was much higher. Although the estimated AR(1) parameter appears unexpectedly high, this result is reasonable because the low-frequency component in Source 1 generates some autocorrelation.

[1]The SNR is defined as $10\log_{10}\dfrac{\sum_{m=1}^{M}\sum_{t=1}^{T}\hat{B}_{m,t}^{2}}{MT\beta^{-1}}$, where $\hat{B}_{m,t}$ is the noise-free sensor measurements on the mth sensor at time point t.

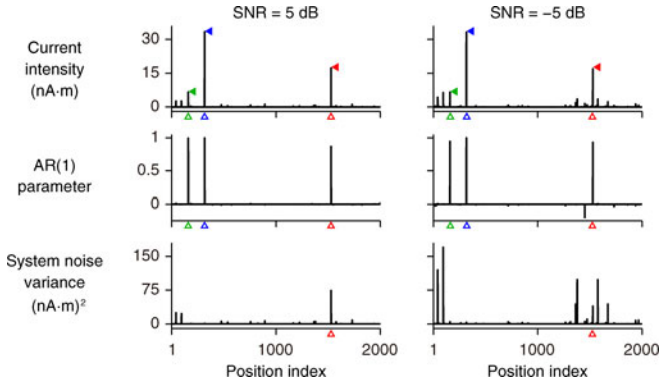

Fig. 4. Estimated currents and dynamics parameters from the proposed method. Upper: current intensity (shown as standard deviation). Center: AR(1) parameter. Lower: system noise variance. Positions of Sources 1, 2, and, 3 are indicated by red, blue, and, green unfilled triangles, respectively. The amplitude of the grand truth currents is indicated by filled triangles. The currents $\mathbf{Z}_{1:T}$ before applying the local spatial filter are presented rather than the smoothed currents $\mathbf{J}_{1:T}$ to clearly show the correspondence of position indices between the currents and the parameters.

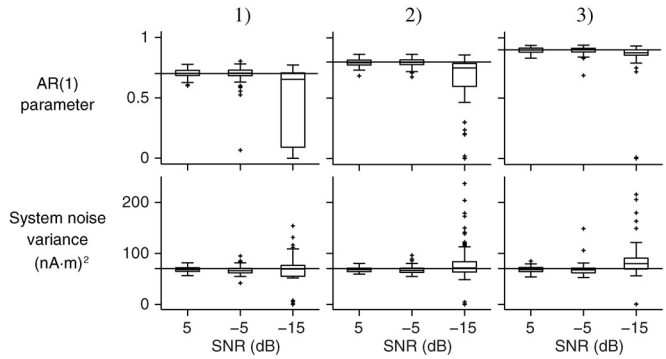

Fig. 5. Boxplot representation of the AR(1) parameter (upper) and system noise variance (lower) estimated by the proposed method at SNRs 5 dB, −5 dB, and −15 dB. The true parameter values of the three active sources are represented by horizontal lines slicing the boxes.

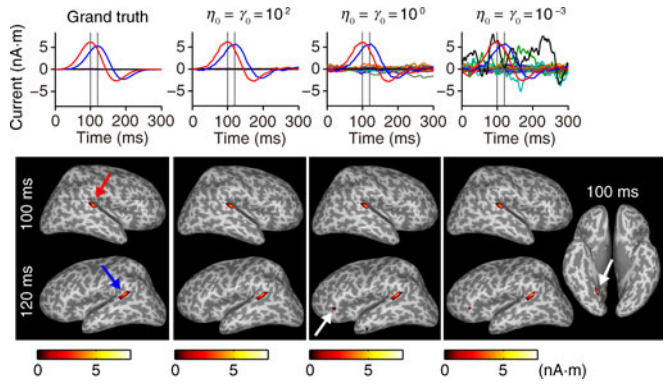

Fig. 6. Spatiotemporal profiles of the true and estimated currents in simulation study 4. The leftmost column illustrates the grand truth. The other columns (left to right) show the estimates from the proposed method for both η_0 and γ_0 equaling 10^2, 10^0, and 10^{-3}, respectively. Upper: time courses of the currents over the whole cortex. Lower: spatial maps of the currents at peak time points 100 and 120 ms (white arrows indicate false positives). The mapping threshold is that used in Fig. 2.

B. Simulation Study 2: Estimation for Correlated Sources

To examine performance of the proposed method for correlated sources, we changed the source configuration from simulation study 1 to a situation in which two nearby sources (10–15 mm apart) have an identical waveform (that of Source 2 in simulation study 1). This simulation was repeated 100 times for different positions of the paired sources.

The proposed method successfully reconstructed the correlated sources for most of the simulations when the SNR was 5 dB. The reconstruction gain averaged over the two true sources was almost 1 for 77 times. In these cases, the false positive sources were not observed around the peak time point (one-tenth of the maximum amplitude was used as a threshold). Even at the low SNR (−5 dB), the reconstruction gain was almost 1 for over half (56 times) of the simulations, while the false positives were appeared in 15 out of 56 times.

C. Simulation Study 3: Validation of Parameter Estimation

We conducted a further set of simulation experiments to check whether the dynamics parameters can be estimated correctly by variational Bayesian inference. To assess the accuracy of the parameter estimation, we generated time courses of the grand truth currents directly from our temporal model shown in (3). Three active sources were assumed, with AR(1) parameter values a_n of 1) 0.7, 2) 0.8, and 3) 0.9, and system noise variance σ_n of 70 $(\mathrm{nA}\cdot\mathrm{m})^2$ for all cases. For the other nonactive sources, AR(1) parameters and system noise variances were both set to 0. Apart from the waveform of the simulated currents and the SNRs, the simulation setting was that of the repeated experiments of simulation study 1.

As shown in Fig. 5, the dynamics parameters were estimated correctly even under the variational Bayesian approximation when the SNR was ±5 dB. At an SNR of −15 dB, far below that of single-trial measurements, the AR(1) parameter was underestimated and the system noise variance was overestimated several times. In addition, the estimated parameters were close to 0 in some cases, as a result of mislocalization of the active sources. Nevertheless, the median of the parameters remained close to the true values.

D. Simulation Study 4: Effects of Hyperparameters

Next, we conducted more realistic simulation experiments to assess the importance of appropriate hyperparameter values. Two active current sources in the left and right temporal lobes were assumed, mimicking auditory evoked responses (see the blue and red arrows of the leftmost column in Fig. 6). MEG measurements were generated as for simulation study 1, but replacing white Gaussian observation noise with more realistic background noise. The background noise was derived from 300 ms prestimulus periods of trial-averaged data from the real auditory experiment described in Section III-E (SNR was around 7 dB). Since such real background brain activities are neither Gaussian nor white, these components could not be properly modeled in our method. Therefore, the hyperparameters η_0 and γ_0 were carefully chosen to minimize the effect of modeling errors on the simulation results.

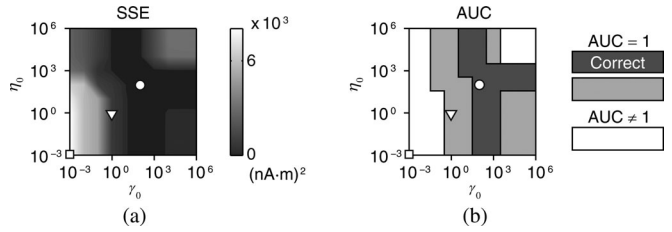

Fig. 7. (a) Interpolated contour map of the SSE between the true and estimated currents. (b) Evaluation of source reconstruction results based on AUC. Sets of hyperparameters resulting in AUC equal to 1 are colored gray (dark and light) while those resulting in AUC not equal to 1 are white. The former region is further divided into two subregions, indicated by dark and light gray, based on quantitative differences (see text). The small circle, triangle, and square in each map correspond to the hyperparameter sets shown in Fig. 6.

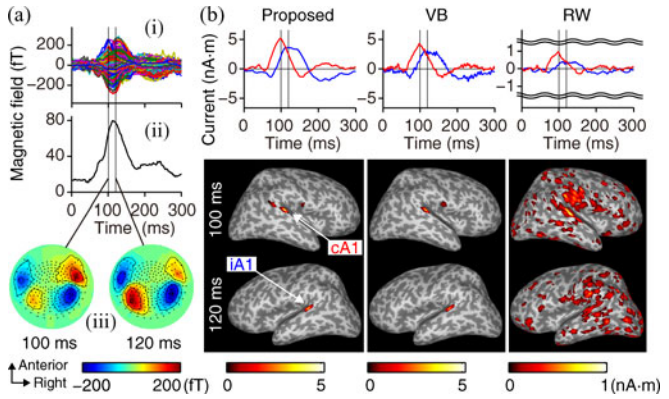

Fig. 8. (a) Preprocessed trial-averaged MEG data from the auditory evoked fields: (i) superposition of the sensor signals; (ii) square root of power averaged over all the sensors; (iii) 2-D scalp maps at two time points, 100 and 120 ms. (b) Spatiotemporal profiles of the estimated currents from the MEG data shown in (a). Upper: time courses of the currents at cA1 (red) and iA1 (blue) (plotted vertices are common to the three methods). Lower: spatial maps of the currents at peak time points 100 and 120 ms. The mapping threshold is the same as for Figs. 2 and 6. Small difference in the prespecified threshold did not affect the basic properties of the mapped results.

Fig. 6 illustrates the estimates from the simulated measurements. With the hyperparameters set to $\eta_0 = \gamma_0 = 10^2$, the two current activities were correctly extracted and the others were suppressed. When η_0 and γ_0 were reduced, i.e., the prior distributions of the parameters were closer to noninformative ones, parameter estimation became more sensitive to disturbance. In this case (both η_0 and γ_0 set to 10^0 or 10^{-3}), the estimated current activities appeared not only at the positions of the grand truth but also at other positions.

To determine the appropriate range of the hyperparameter values, we repeated the estimation by varying η_0 and γ_0 from 10^{-3} to 10^6 (10 degrees). Fig. 7(a) shows an interpolated contour map of the sum of squared error (SSE) between the true and estimated currents. Fig. 7(b) presents the results of receiver operating characteristic (ROC) analysis [39]. In this map, the hyperparameter settings for which the area under the ROC curve (AUC), averaged over 90–130 ms, equals 1, are delineated by dark and light gray regions. The dark gray region is regarded as the "correct" source reconstruction case, in which the current activities are obtained only at the two true positions, using one-tenth of the maximum amplitude as a threshold. The best case of the estimates in Fig. 6 is indicated by a small circle, and the other cases, yielding false positives, are indicated by a small triangle and a small square. The SSE and AUC maps also show that large hyperparameter values result in incorrect source reconstruction. As shown in Fig. 7(b), correct estimates were obtained for a relatively wide range of γ_0 when η_0 was 10^2 or 10^3. On the other hand, setting γ_0 to 10^2 enabled a large range of η_0 to be successfully used. Thus, appropriate ranges for the hyperparameters, at least in the tested simulation setting, are $\eta_0 = 10^2$–10^3 and $\gamma_0 \sim 10^2$.

E. Real Data Analysis: Auditory Evoked Response

Finally, we examined whether the previously specified hyperparameter values provide a physiologically plausible result in a real data analysis. We applied the proposed method to the auditory evoked fields from one right-handed healthy subject (a 27-year-old male). A signed informed consent approved by the ATR Human Subject Review Committee was obtained prior to experiment. Measurement data were acquired from a 400-channel MEG system (MEGvision PQ1400RM, Yokogawa Electric Co.) with a sampling frequency of 1 kHz. The cortical surface model was constructed from a segmented gray matter image of T1-weighted MRI collected by a 3 T scanner (MAGNETOM Trio, A Tim System 3T, Siemens).

An 800-Hz tone burst of duration 500 ms was presented to the left ear of the subject. To keep the subject awake, we asked him to press a button 2–3 s after each stimulus. The trial-averaged data in the time window 0–300 ms with respect to the stimulus onset were used to estimate the currents. The noise covariance matrix was determined in advance from nonaveraged rest period data in a-100-ms prestimulus window.

The following preprocessing operations were applied: baseline correction, high-pass filtering (cutoff 0.5 Hz), low-pass filtering (cutoff 200 Hz), trial segmentation, and trial rejection. Artifactual trials, determined from the amplitudes of MEG and electrooculogram and from subject behavior (button press), were rejected. As a result, 112 out of 128 trials were accepted for further analysis.

Preprocessed trial-averaged MEG data, showing typical time courses and scalp maps of the auditory evoked fields, are displayed in Fig. 8(a). Fig. 8(b) presents the estimated currents for $\eta_0 = \gamma_0 = 10^2$. Focal current activities were estimated at the contra-lateral and ipsilateral primary auditory cortices (cA1 and iA1, respectively), consistent with previous findings from a cat electrophysiological study [40], a human intracranial study [41], and a human MEG study [42]. The estimated time courses show that the latency and amplitude of cA1 activities are shorter and larger respectively, than those from iA1. These results are also consistent with previous physiological findings [43], [44]. When η_0 was increased to 10^3, the number of active positions slightly decreased while the current time courses at cA1 and iA1 did not effectively change.

Similar results were obtained from VB using the same hyperparameter values, but some differences were observed: the amplitude and temporal smoothness of the currents were lower

than those of the proposed method, and the active brain region was more limited. In contrast, RW generated spatially diffuse currents with the lowest amplitudes.

IV. Discussion

We have developed a state-space model-based source localization method that achieves current source reconstruction to high spatial resolution (i.e., localization of focal current source patterns). To our knowledge, the method is the first to reconstruct spatially focal current sources utilizing temporal constraints from the state-space model. The key to our method is to assume different temporal evolution for each current source and to estimate all parameters in the state equation from data. The joint estimation of both the current sources and the model parameters is based on the Bayesian framework. We overcome difficulties presented by the high dimensionality of the problem by 1) introducing prior distributions for the parameters, which reduces the sensitivity of the parameter estimation to data, especially when discrepancies exist between the model assumptions and the data generation process, and 2) employing variational Bayesian inference, which reduces the computational cost of the estimation algorithm.

The proposed method is a dynamical extension of VB [13], [28]. Its advantage over VB is high reconstruction gain, especially for low-amplitude currents (see Fig. 3). This indicates that, by introducing temporal constraints, considerable improvement is achieved over a simple smoothing operation such as low-pass filtering. The enhanced reconstruction gain of our method demonstrates that the amplitude of current source underestimated by VB can be compensated by the effective use of the current sources at adjacent time points, owing to the estimated temporal dynamics. Indeed, VB, along with FOCUSS [11] and empirical Bayesian methods [14], [16] is a special form of covariance component estimation methods [15]. Unlike the proposed method, none of these established methods incorporate temporal constraints. Introducing state-space modeling into FOCUSS and empirical Bayesian methods might enhance their performance in reconstructing weak sources, as for VB.

A key factor in reconstructing focal activities is estimating the temporal dynamics of each source individually, and this is not apparent from the form of our probabilistic model unlike the explicit constraint term of the $l1$ or $l0$ norm-regularization methods [5]–[7]. To understand how our method can localize focal activities, we must consider the prior distribution for any single-current source, given by $\mathcal{N}(J_{n,t}|\, a_n J_{n,t-1}, \sigma_n)$, and the update rules of a_n and q_n [see (24) and (27) in Appendix, respectively]. The prior distribution ensures that when the AR(1) parameter a_n and the system noise variance σ_n are close to 0, the currents are softly constrained to 0. For such near-zero currents, the update rules dictate that a_n and σ_n (which is proportional to q_n^{-1}) are close to 0 also (σ_n is not exactly zero when $\gamma_0 \neq 0$). These two mechanisms provide positive feedback effects on promoting the most of $J_{n,1:T}$, a_n, and σ_n estimated to be near 0, resulting in spatially focal activities (see Fig. 4). We have confirmed that this property also holds under initialization with nonfocal estimates such as MN. In contrast to the proposed method, homogeneous RW dynamics imposes no zero-value constraints on any of the current sources, since all sources fluctuate around their past values in a spatially uniform manner.

There are another type of focal source localization methods in the subspace approach, such as MUSIC [8] and FINES [9], [10]. As reported in [10], these methods have difficulty in reconstructing perfectly correlated and closely located sources; however, simulation study 2 shows that these sources can be reconstructed by the proposed method. This is a possible advantage of our method over MUSIC and FINES, in addition to a desirable property that the number of active sources does not need to be estimated beforehand.

Variational Bayesian inference [30]–[32] has enabled the computationally demanding inference of the high-dimensional state-space model with little detriment to accuracy. A main drawback of our estimation algorithm appears to be the decrease in parameter estimation accuracy due to the independence assumption on the approximate posterior distribution of the current sources $Q(\mathbf{J}_{1:T}) = Q(\mathbf{J}_1) \cdots Q(\mathbf{J}_T)$. This presents the major difference between our algorithm and the VKS algorithm [37], [38]. However, as shown in simulation study 3, parameter estimation of our method was reasonably accurate for the modest SNR case (± 5 dB) (see Fig. 5). In fact, we also observed little difference in accuracy between our algorithm and the VKS algorithm in a small-size simulation. The estimation accuracy of the dynamics would come from the fact that temporal dependence between current $\mathbf{J}_t$ is propagated through the means of the posterior distributions at neighboring time samples $\mathbf{J}_{t-1}$ and $\mathbf{J}_{t+1}$ in our algorithm [see (15)–(17) in Appendix].

To reduce the sensitivity of the parameter estimation to unmodeled components, in particular for real data analysis, appropriate values of the hyperparameters are required. An example of such unmodeled components for our method is background brain activity, which cannot be modeled by white Gaussian observation noise. With no informative prior distributions, our source reconstruction could be affected by these factors. We searched the appropriate ranges of the hyperparameters in simulation study 4, in which the background activities of the real experiment were used and confirmed that those hyperparameter settings derived physiologically reasonable results from real MEG data of the auditory evoked fields [see Fig. 8(b)]. In fact, by applying the same hyperparameters to somatosensory evoked fields, we have also reconstructed plausible current sources that were consistent with literature results (data not shown).

Recently, both spatially focal and temporally smooth current source reconstructions have been achieved in [19] and [20], via group-wise $l1$ regularization and temporal basis functions. Although such methods might yield similar results to ours, the state-space modeling approach can model spatiotemporal properties of the current sources and is thus more flexible in constraining inverse solutions.

The temporal dynamics assumed in this study might appear to be oversimplified; however, our model substantially advances those of previous state-space methods. The model could be further extended by using a matrix form of the AR(1) parameter to account for source-space connectivity. This extension would be much more demanding because the number of parameters

increases with square order of the state dimension. The large degrees of freedom could be reduced by incorporating anatomical connections as prior knowledge. This possibility will be pursued in future study.

V. Conclusion

The proposed state-space model-based method has successfully extracted focal current activities in simulation and in application to real data. The amplitude of current activities, especially for weak components, was preserved in our method owning to the source dynamics model.

Appendix

Algorithm details

We present here the detailed derivation of the update rules for the currents ($\mathbf{J}$-step) and for the parameters (β-step, $\mathbf{a}$-step, and $\mathbf{q}$-step).

In the $\mathbf{J}$-step, the approximate posterior distributions of the currents $Q(\mathbf{J}_t)$ for $t = 1:T$ are updated. The distribution $Q(\mathbf{J}_t)$ is obtained by substituting $\mathbf{J}_t$ into (11):

$$\log Q(\mathbf{J}_t) = \langle \log P(\mathbf{B}_{1:T}, X) \rangle_{Q(X \setminus \mathbf{J}_t)} + \text{const.} \tag{14}$$

By calculating the integral in the right-hand side of (14), we obtain the linear and quadratic terms with respect to $\mathbf{J}_t$. Therefore, $Q(\mathbf{J}_t)$ is a Gaussian distribution:

$$Q(\mathbf{J}_t) = \mathcal{N}(\mathbf{J}_t | \, \bar{\mathbf{J}}_t, \bar{\beta}^{-1} \bar{\mathbf{V}}_t) \tag{15}$$

where $\bar{\beta}$ represents the mean of $Q(\beta)$. The (scaled) covariance matrix $\bar{\mathbf{V}}_t$ and the mean $\bar{\mathbf{J}}_t$ are derived as

$$\begin{cases} \bar{\mathbf{V}}_t = \left(\mathbf{G}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{G} + \hat{\mathbf{V}}_t^{-1} \right)^{-1} \\ \bar{\mathbf{J}}_t = \bar{\mathbf{V}}_t \left(\mathbf{G}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{B}_t + \hat{\mathbf{V}}_t^{-1} \hat{\mathbf{J}}_t \right) \end{cases} \tag{16}$$

where we have introduced auxiliary parameters $\hat{\mathbf{V}}_t$ and $\hat{\mathbf{J}}_t$. These are defined as

$$\hat{\mathbf{V}}_t = \begin{cases} \left(\left(\mathbf{I} + \langle \mathrm{diag}(\mathbf{a}) \mathrm{diag}(\mathbf{a}) \rangle \right) \mathrm{diag}(\bar{\mathbf{q}}) \right)^{-1}, & \text{for } t \neq T \\ \left(\mathrm{diag}(\bar{\mathbf{q}}) \right)^{-1}, & \text{for } t = T \end{cases}$$

$$\hat{\mathbf{J}}_t = \begin{cases} \left(\mathbf{I} + \langle \mathrm{diag}(\mathbf{a}) \mathrm{diag}(\mathbf{a}) \rangle \right)^{-1} \\ \qquad \times \mathrm{diag}(\bar{\mathbf{a}}) (\bar{\mathbf{J}}_{t-1} + \bar{\mathbf{J}}_{t+1}), & \text{for } t \neq T \\ \mathrm{diag}(\bar{\mathbf{a}}) \bar{\mathbf{J}}_{t-1}, & \text{for } t = T \end{cases} \tag{17}$$

with $\bar{\mathbf{a}}$, $\bar{\mathbf{q}}$, $\bar{\mathbf{J}}_{t-1}$, and $\bar{\mathbf{J}}_{t+1}$ being the means of their approximate posteriors, which are updated at the previous iteration, and $\mathbf{I}$ being the identity matrix. Equation (16) can be rewritten in a computationally efficient form using the Woodbury's matrix inversion lemma [32]:

$$\begin{cases} \bar{\mathbf{V}}_t = \left(\mathbf{I} - \mathbf{K} \mathbf{G} \right) \hat{\mathbf{V}}_t \\ \bar{\mathbf{J}}_t = \hat{\mathbf{J}}_t + \mathbf{K} \left(\mathbf{B}_t - \mathbf{G} \hat{\mathbf{J}}_t \right) \end{cases} \tag{18}$$

where $\mathbf{K} = \hat{\mathbf{V}}_t \mathbf{G}^{\mathrm{T}} (\mathbf{G} \hat{\mathbf{V}}_t \mathbf{G}^{\mathrm{T}} + \mathbf{S})^{-1}$ plays a role similar to the Kalman gain. By adopting this form, the matrix inversions of large size (in our application 2000×2000) in (16) are avoided.

It should be noted that since $\hat{\mathbf{V}}_t$ is diagonal owing to the independence approximation on $Q(\mathbf{J}_{1:T})$, the computational complexity required for obtaining $\bar{\mathbf{V}}_t$ and $\bar{\mathbf{J}}_t$ increases linearly with number of current sources.

In the β-step, the approximate posterior distribution of the scaling parameter $Q(\beta)$ is updated. The distribution $Q(\beta)$ is obtained by substituting β into (11):

$$\log Q(\beta) = \langle \log P(\mathbf{B}_{1:T}, X) \rangle_{Q(X \setminus \beta)} + \text{const.} \tag{19}$$

By calculating the integral in the right-hand side of (19), we obtain the linear and logarithmic terms with respect to β. Therefore, $Q(\beta)$ is a gamma distribution:

$$Q(\beta) = \Gamma(\beta | \, \bar{\beta}, \gamma_\beta) \tag{20}$$

where the confidence parameter γ_β and the mean $\bar{\beta}$ are derived as

$$\begin{cases} \gamma_\beta = \frac{1}{2} MT \\ \bar{\beta} = \gamma_\beta \Big\{ \frac{1}{2} \sum_{t=1}^{T} \mathrm{tr} \big(\mathbf{S}^{-1} (\mathbf{B}_t \mathbf{B}_t^{\mathrm{T}} - 2 \, \mathbf{G} \bar{\mathbf{J}}_t \mathbf{B}_t^{\mathrm{T}} + \mathbf{G} \bar{\mathbf{J}}_t \bar{\mathbf{J}}_t^{\mathrm{T}} \mathbf{G}^{\mathrm{T}} \big) \\ \qquad + \frac{1}{2} \sum_{t=1}^{T} \mathrm{tr} \Big(\mathrm{diag}(\mathbf{q}) (\bar{\mathbf{J}}_t \bar{\mathbf{J}}_t^{\mathrm{T}} - 2 \, \mathrm{diag}(\mathbf{a}) \bar{\mathbf{J}}_{t-1} \bar{\mathbf{J}}_t^{\mathrm{T}} \\ \qquad\qquad + \langle \mathrm{diag}(\mathbf{a}) \mathrm{diag}(\mathbf{a}) \rangle \bar{\mathbf{J}}_{t-1} \bar{\mathbf{J}}_{t-1}^{\mathrm{T}}) \Big) \Big\}^{-1}. \end{cases} \tag{21}$$

In the $\mathbf{a}$-step, the approximate posterior distribution of the AR(1) parameter $Q(\mathbf{a})$ is updated. The distribution $Q(\mathbf{a})$ is obtained by substituting $\mathbf{a}$ into (11):

$$\log Q(\mathbf{a}) = \langle \log P(\mathbf{B}_{1:T}, X) \rangle_{Q(X \setminus \mathbf{a})} + \text{const.} \tag{22}$$

By calculating the integral in the right-hand side of (22), we obtain the sum of the linear and quadratic terms with respect to a_n. Therefore, $Q(\mathbf{a})$ is a product of Gaussian distributions:

$$Q(\mathbf{a}) = \prod_{n=1}^{N} \mathcal{N}(a_n | \, \bar{a}_n, \eta_n^{-1}) \tag{23}$$

where the variance η_n^{-1} and the mean $\bar{a}_n$ are derived as

$$\begin{cases} \eta_n^{-1} = \left\{ \eta_0 + \sum_{t=1}^{T} \bar{\beta} \left(\mathrm{diag}(\bar{\mathbf{q}}) \langle \mathbf{J}_{t-1} \mathbf{J}_{t-1}^{\mathrm{T}} \rangle \right)_{n,n} \right\}^{-1} \\ \bar{a}_n = \eta_n^{-1} \sum_{t=1}^{T} \bar{\beta} \left(\mathrm{diag}(\bar{\mathbf{q}}) \bar{\mathbf{J}}_{t-1} \bar{\mathbf{J}}_t^{\mathrm{T}} \right)_{n,n}. \end{cases} \tag{24}$$

In the $\mathbf{q}$-step, the approximate posterior distribution of the vector parameterizing the system noise $Q(\mathbf{q})$ is updated. The distribution $Q(\mathbf{q})$ is obtained by substituting $\mathbf{q}$ into (11):

$$\log Q(\mathbf{q}) = \langle \log P(\mathbf{B}_{1:T}, X) \rangle_{Q(X \setminus \mathbf{q})} + \text{const.} \tag{25}$$

By calculating the integral in the right-hand side of (25), we obtain the sum of the linear and logarithmic terms with respect to q_n. Therefore, $Q(\mathbf{q})$ is a product of gamma distributions:

$$Q(\mathbf{q}) = \prod_{n=1}^{N} \Gamma(q_n | \, \bar{q}_n, \gamma_n) \tag{26}$$

where the confidence parameter γ_n and the mean $\bar{q}_n$ are derived as

$$\begin{cases} \gamma_n = \gamma_0 + \frac{1}{2}\,T \\ \bar{q}_n = \gamma_n \Big\{ \gamma_0\, q_0^{-1} + \frac{1}{2}\sum_{t=1}^{T} \bar{\beta} \Big(\langle \mathbf{J}_t \mathbf{J}_t^{\mathrm{T}} \rangle - 2\mathrm{diag}(\mathbf{a}) \bar{\mathbf{J}}_{t-1} \bar{\mathbf{J}}_t^{\mathrm{T}} \\ \qquad + \langle \mathrm{diag}(\mathbf{a})\mathrm{diag}(\mathbf{a}) \rangle \langle \mathbf{J}_{t-1} \mathbf{J}_{t-1}^{\mathrm{T}} \rangle \Big)_{n,n} \Big\}^{-1}. \end{cases} \tag{27}$$

Remaining expectations in the update equations are evaluated as follows [$\boldsymbol{\eta}$ is a vector whose entries are η_n in (24)]:

$$\langle \mathrm{diag}(\mathbf{a})\mathrm{diag}(\mathbf{a}) \rangle = \mathrm{diag}(\mathbf{a})\mathrm{diag}(\mathbf{a}) + (\mathrm{diag}(\boldsymbol{\eta}))^{-1} \tag{28}$$

$$\sum_{t=1}^{T} \langle \mathbf{J}_{t-1} \mathbf{J}_{t-1}^{\mathrm{T}} \rangle = \bar{\mathbf{J}}_0 \bar{\mathbf{J}}_0^{\mathrm{T}} + \sum_{t=2}^{T} \left(\bar{\mathbf{J}}_{t-1} \bar{\mathbf{J}}_{t-1}^{\mathrm{T}} + \bar{\beta}^{-1} \bar{\mathbf{V}}_t \right) \tag{29}$$

$$\sum_{t=1}^{T} \langle \mathbf{J}_t \mathbf{J}_t^{\mathrm{T}} \rangle = \sum_{t=1}^{T} \left(\bar{\mathbf{J}}_t \bar{\mathbf{J}}_t^{\mathrm{T}} + \bar{\beta}^{-1} \bar{\mathbf{V}}_t \right). \tag{30}$$

ACKNOWLEDGMENT

The authors are grateful to N. Hiroe (ATR) for providing the real experimental data and to T. Yoshioka, R. Hayashi and T. Sako (ATR) for solving the forward problem and for helping with data preprocessing.

REFERENCES

[1] P. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. New York: Oxford Univ. Press, 2006.
[2] S. Baillet, J. Mosher, and R. Leahy, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 14–30, Nov. 2001.
[3] M. Hämäläinen and R. Ilmoniemi, "Interpreting magnetic fields of the brain: Minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, pp. 35–42, Jan. 1994.
[4] R. Pascual-Marqui, C. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain," *Int. J. Psychophysiol.*, vol. 18, pp. 49–65, Oct. 1994.
[5] K. Matsuura and Y. Okabe, "Selective minimum-norm solution of the biomagnetic inverse problem," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 6, pp. 608–615, Jun. 1995.
[6] K. Uutela, M. Hämäläinen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, pp. 173–180, Aug. 1999.
[7] P. Xu, Y. Tian, X. Lei, and D. Yao, "Neuroelectric source imaging using 3SCO: A space coding algorithm based on particle swarm optimization and *l*0 norm constraint," *NeuroImage*, vol. 51, pp. 183–205, May 2010.
[8] J. Mosher, P. Lewis, and R. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 6, pp. 541–557, Jun. 1992.
[9] X.-L. Xu, B. Xu, and B. He, "An alternative subspace approach to EEG dipole source localization," *Phys. Med. Biol.*, vol. 49, pp. 327–343, Jan. 2004.
[10] L. Ding and B. He, "Spatio-temporal EEG source localization using a three-dimensional subspace FINE approach in a realistic geometry inhomogeneous head model," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 9, pp. 1732–1739, Sep. 2006.
[11] I. Gorodnitsky, J. George, and B. Rao, "Neuromagnetic source imaging with FOCUSS: A recursive weighted minimum norm algorithm," *Electroencephalogr. Clin. Neurophysiol.*, vol. 95, pp. 231–251, Oct. 1995.
[12] P. Xu, Y. Tian, H. Chen, and D. Yao, "Lp norm iterative sparse solution for EEG source localization," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 400–409, Mar. 2007.
[13] M. Sato, T. Yoshioka, S. Kajihara, K. Toyama, N. Goda, K. Doya, and M. Kawato, "Hierarchical Bayesian estimation for MEG inverse problem," *NeuroImage*, vol. 23, pp. 806–826, Nov. 2004.
[14] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, pp. 1104–1120, Feb. 2008.
[15] D. Wipf and S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, pp. 947–966, Feb. 2009.
[16] X. Lei, P. Xu, C. Luo, J. Zhao, D. Zhou, and D. Yao, "fMRI functional networks for EEG source imaging," *Hum. Brain Mapping*, vol. 32, pp. 1141–1160, Sep. 2011.
[17] R. Frackowiak, K. Friston, C. Frith, R. Dolan, C. Price, S. Zeki, J. Ashburner, and W. Penny, *Human Brain Function*, 2nd ed. New York: Academic, 2004.
[18] N. Trujillo-Barreto, E. Aubert-Vázquez, and W. Penny, "Bayesian M/EEG source reconstruction with spatio-temporal priors," *NeuroImage*, vol. 39, pp. 318–335, Jan. 2008.
[19] W. Ou, M. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, pp. 932–946, Feb. 2009.
[20] A. Bolstad, B. Van Veen, and R. Nowak, "Space-time event sparse penalization for magneto-/electroencephalography," *NeuroImage*, vol. 46, pp. 1066–1081, Jul. 2009.
[21] S. Baillet and L. Garnero, "A Bayesian approach to introducing anatomofunctional priors in the EEG/MEG inverse problem," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 5, pp. 374–385, May 1997.
[22] F. Darvas, U. Schmitt, A. Louis, M. Fuchs, G. Knoll, and H. Buchner, "Spatio-temporal current density reconstruction (stCDR) from EEG/MEG-data," *Brain Topogr.*, vol. 13, pp. 195–207, Mar. 2001.
[23] U. Schmitt, A. Louis, F. Darvas, H. Buchner, and M. Fuchs, "Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data," *IEEE Trans. Med. Imag.*, vol. 20, no. 4, pp. 314–324, Apr. 2001.
[24] J. Daunizeau, J. Mattout, D. Clonda, B. Goulard, H. Benali, and J. Lina, "Bayesian spatio-temporal approach for EEG source reconstruction: conciliating ECD and distributed models," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 503–516, Mar. 2006.
[25] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, and P. Valdés-Sosa, "A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering," *NeuroImage*, vol. 23, pp. 435–453, Oct. 2004.
[26] O. Yamashita, A. Galka, T. Ozaki, R. Biscay, and P. Valdés-Sosa, "Recursive penalized least squares solution for dynamical inverse problems of EEG generation," *Hum. Brain Mapping*, vol. 21, pp. 221–235, Apr. 2004.
[27] M. Barton, P. Robinson, S. Kumar, A. Galka, H. Durrant-Whyte, J. Guivant, and T. Ozaki, "Evaluating the performance of Kalman-filter-based EEG source localization," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 1, pp. 122–136, Jan. 2009.
[28] T. Yoshioka, K. Toyama, M. Kawato, O. Yamashita, S. Nishina, N. Yamagishi, and M. Sato, "Evaluation of hierarchical Bayesian method through retinotopic brain activities reconstruction from fMRI and MEG signals," *NeuroImage*, vol. 42, pp. 1397–1413, Oct. 2008.
[29] C. Lamus, C. Long, M. Hämäläinen, E. Brown, and P. Purdon, "Parameter estimation and dynamic source localization for the magnetoencephalography (MEG) inverse problem," in *Proc. 4th IEEE Int. Symp. Biomed. Imag.*, 2007, pp. 1092–1095.
[30] H. Attias, "Inferring parameters and structure of latent variable models by variational Bayes," in *Proc. 15th Conf. Uncertainty Artif. Intell.*, 1999, pp. 21–30.
[31] M. Sato, "Online model selection based on the variational Bayes," *Neural Comput.*, vol. 13, pp. 1649–1681, Jul. 2001.
[32] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006.
[33] M. Hämäläinen, R. Hari, R. Ilmoniemi, J. Knuutila, and O. Lounasmaa, "Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Mod. Phys.*, vol. 65, pp. 413–497, Apr. 1993.
[34] J. Mosher, R. Leahy, and P. Lewis, "EEG and MEG: Forward solutions for inverse methods," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 3, pp. 245–259, Mar. 1999.
[35] A. Nummenmaa, T. Auranen, M. Hämäläinen, I. Jääskeläinen, M. Sams, A. Vehtari, and J. Lampinen, "Automatic relevance determination based hierarchical Bayesian MEG inversion in practice," *NeuroImage*, vol. 37, pp. 876–889, Sep. 2007.
[36] R. Neal, *Bayesian Learning for Neural Networks*. New York: Springer-Verlag, 1996.
[37] J. Daunizeau and K. Friston, "A mesostate-space model for EEG and MEG," *NeuroImage*, vol. 38, pp. 67–81, Oct. 2007.

[38] M. Beal, "Variational algorithm for approximate Bayesian inference" Ph.D. dissertation, Univ. College London, London, U.K., 2003.
[39] J. Swets, "Measuring the accuracy of diagnostic systems," *Science*, vol. 240, pp. 1285–1293, Jun. 1988.
[40] T. Imig and H. Adrián, "Binaural columns in the primary field (A1) of cat auditory cortex," *Brain Res.*, vol. 138, pp. 241–257, Dec. 1977.
[41] B. Godey, D. Schwartz, J. de Graaf, P. Chauvel, and C. Liégeois-Chauvel, "Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients," *Clin. Neurophysiol.*, vol. 112, pp. 1850–1859, Oct. 2001.
[42] R. Hari, K. Aittoniemi, M. Järvinen, T. Katila, and T. Varpula, "Auditory evoked transient and sustained magnetic fields of the human brain localization of neural generators," *Exp. Brain Res.*, vol. 40, pp. 237–240, Sep. 1980.
[43] C. Pantev, B. Ross, P. Berg, T. Elbert, and B. Rockstroh, "Study of the human auditory cortices using a whole-head magnetometer: Left vs. right hemisphere and ipsilateral vs. contralateral stimulation," *Audiol. Neurootol.*, vol. 3, pp. 183–190, Mar.–Jun. 1998.
[44] D. Gabriel, E. Veuillet, R. Ragot, D. Schwartz, A. Ducorps, A. Norena, J. Durrant, A. Bonmartin, F. Cotton, and L. Collet, "Effect of stimulus frequency and stimulation site on the N1m response of the human auditory cortex," *Hear. Res.*, vol. 197, pp. 55–64, Nov. 2004.

Authors' photographs and biographies not available at the time of publication.