NeuroImage 105 (2015) 408–427

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# MEG source reconstruction based on identification of directed source interactions on whole-brain anatomical networks

Makoto Fukushima [a,b], Okito Yamashita [b,c,*], Thomas R. Knösche [d], Masa-aki Sato [b]

[a] Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan
[b] ATR Neural Information Analysis Laboratories, Kyoto, Japan
[c] Brain Functional Imaging Technologies Group, CiNet, Osaka, Japan
[d] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

*Article history:*
Accepted 26 September 2014
Available online 5 October 2014

*Keywords:*
MEG source reconstruction
Multivariate autoregressive model
Effective connectivity
Anatomical connectivity
Prior knowledge
Variational Bayes

ABSTRACT

We present an MEG source reconstruction method that simultaneously reconstructs source amplitudes and identifies source interactions across the whole brain. In the proposed method, a full multivariate autoregressive (MAR) model formulates directed interactions (i.e., effective connectivity) between sources. The MAR coefficients (the entries of the MAR matrix) are constrained by the prior knowledge of whole-brain anatomical networks inferred from diffusion MRI. Moreover, to increase the accuracy and robustness of our method, we apply an fMRI prior on the spatial activity patterns and a sparse prior on the MAR coefficients. The observation process of MEG data, the source dynamics, and a series of the priors are combined into a Bayesian framework using a state-space representation. The parameters, such as the source amplitudes and the MAR coefficients, are jointly estimated from a variational Bayesian learning algorithm. By formulating the source dynamics in the context of MEG source reconstruction, and unifying the estimations of source amplitudes and interactions, we can identify the effective connectivity without requiring the selection of regions of interest. Our method is quantitatively and qualitatively evaluated on simulated and experimental data, respectively. Compared with non-dynamic methods, in which the interactions are estimated after source reconstruction with no dynamic constraints, the proposed dynamic method improves most of the performance measures in simulations, and provides better physiological interpretation and inter-subject consistency in real data applications.



## Introduction

There are two fundamental functional principles of the brain: functional specialization and functional integration (Tononi et al., 1994; Friston, 1994). Identifying functionally specialized brain regions (e.g., for sensory processing, motor control, and cognitive processing) has been a long-term focus of neuroimaging studies. However, for a true understanding of the mechanisms underlying brain function, elucidating the scheme of dynamic integration between these functionally specialized brain regions is indispensable. This topic has received growing interest in recent years (Hutchison et al., 2013).

Magnetoencephalography (MEG) and electroencephalography (EEG) provide ways to investigate such dynamic integration of brain functions (Schoffelen and Gross, 2009; Palva and Palva, 2012), because of their high temporal resolution and large reflection of neuronal electrical activity (Hämäläinen et al., 1993; Nunez and Srinivasan, 2006). The richness of the temporal information in MEG/EEG allows capturing temporal propagation, or event-related dynamics, of neuronal activity occurring over millisecond time scales, which cannot be easily achieved by functional magnetic resonance imaging (fMRI). In contrast to the excellent temporal resolution, the spatial resolutions of MEG and EEG are limited; the spatial distribution of neuronal current sources cannot be uniquely determined from the measurements, unless a priori knowledge or assumptions are imposed as constraints on current sources (Baillet et al., 2001).

Numerous source reconstruction methods have been developed over the past three decades. These methods can be categorized into three approaches; the equivalent current dipole approach, the linear distributed source approach, and the spatial filtering approach. In the equivalent current dipole approach, a small number of focal sources are pre-determined and their locations and amplitudes are estimated by non-linear optimization algorithms (Scherg and Von Cramon, 1985; Mosher et al., 1992). The linear distributed source approach allocates a large number of sources to grid points over the whole brain volume or surface. The amplitude of all sources is simultaneously estimated by solving a system of linear equations. Since the linear equations are underdetermined, additional constraints or prior information are necessary to obtain a unique solution. Prior assumptions used in linear distributed solvers include a spatial prior

* Corresponding author at: ATR Neural Information Analysis Laboratories, 2-2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto 619-0288, Japan. Fax: +81 774 95 1259.
*E-mail address:* oyamashi@atr.jp (O. Yamashita).

http://dx.doi.org/10.1016/j.neuroimage.2014.09.066

forming minimum *l*2 norm regularization (MNE; Hämäläinen and Ilmoniemi, 1994), spatial smoothness priors (LORETA and its variant; Pascual-Marqui et al., 1994; Pascual-Marqui, 2002), spatial sparseness priors (Matsuura and Okabe, 1995; Uutela et al., 1999; Sato et al., 2004; Friston et al., 2008; Wipf et al., 2010), temporal smoothness priors (Baillet and Garnero, 1997; Schmitt et al., 2001; Daunizeau et al., 2006), temporal basis function priors (Trujillo-Barreto et al., 2008; Ou et al., 2009; Bolstad et al., 2009), and fMRI-based spatial priors (Dale et al., 2000; Sato et al., 2004; Daunizeau et al., 2007; Henson et al., 2010; Ou et al., 2010). In the spatial filtering approach, an optimal spatial filter, which maps the sensor measurements to the current source amplitude at each single grid point in the brain, is computed. A popular method for this purpose is the linear constrained minimum variance (LCMV) beamformer (Van Veen et al., 1997). LCMV is used to identify resting-state MEG functional connectivity for neuroscience research (Brookes et al., 2011; Hipp et al., 2012). Wipf and Nagarajan (2009) have recently proposed a framework unifying the beamformer method and some distributed source methods.

In source reconstruction from the linear distributed source approach, introducing prior constraints on the spatiotemporal dynamics of source activities is of particular interest; this type of constraint complements other commonly used constraints (typically spatial) and introduces additional knowledge into the source reconstruction process, for example, on dynamic properties of neuronal populations, anatomical connections between brain areas, and transmission delays of neuronal activities. This knowledge potentially facilitates the extraction of information on directed interactions (i.e., effective connectivity) between sources, while reconstructing spatial source distributions from MEG/EEG data. The spatiotemporal dynamics reflects the generative nature of neuronal current sources, and is readily incorporated into a state-space representation. To formulate such dynamics, previous state-space methods have adopted linear autoregressive models with spatially local interactions (Galka et al., 2004; Lamus et al., 2012) and self-interactions (Yamashita et al., 2004; Daunizeau and Friston, 2007; Fukushima et al., 2012). These methods extend an approach that imposes a simple prior assumption (such as a temporal smoothness prior in Schmitt et al., 2001) on the source dynamics (the effectiveness of imposing simple temporal smoothness is critically evaluated by Dannhauer et al., 2013). Nevertheless, these methods still cannot elucidate the long-range interactions across brain areas. This problem was first solved by Olier et al. (2013), who represented these interactions using the full multivariate autoregressive (MAR) model. However, in this model, the spatiotemporal dynamics was formulated in a low-dimensional latent space rather than in the source space.

To allow the long-range interactions to be directly estimated in the source space, we extend the previous state-space methods into a new MEG source reconstruction method. To achieve this goal, the full MAR model is implemented in the high-dimensional source space. The structure of the MAR model is informed by whole-brain anatomical networks inferred from diffusion MRI (dMRI). More specifically, the MAR coefficients (entries of the MAR matrix) associated with pairs of anatomically connected sources according to dMRI, are estimated from the data, while the others are fixed at zero. The time lags of the MAR model are determined from the mean fiber lengths between pairs of source locations. The anatomical long-range connectivity has been used as a constraint in forward modeling of neuronal dynamics (Honey et al., 2007; Ghosh et al., 2008; Deco et al., 2009), and in estimating the effective connectivity from fMRI data (Stephan et al., 2009; Woolrich and Stephan, 2013). The a priori knowledge of anatomical connectivity also reduces the prohibitively large number of model parameters (in our scenario, from order $10^6$ to order $10^5$ at minimum), thereby improving the feasibility of the estimation. Using this prior information, we can simultaneously estimate the current sources and the source-space effective connectivity. This joint estimation framework distinguishes our method from existing approaches (David et al., 2006; Owen et al., 2009; Hui et al., 2010; Brookes et al., 2011; Hipp et al., 2012; de Pasquale et al., 2012) in which the source time courses and the source connectivity are sequentially estimated. With a low-dimensional MAR model, it was demonstrated that the joint approach yielded better connectivity estimates than the sequential approach (Cheung et al., 2010).

To further improve the reliability of source reconstruction, we apply an fMRI prior on the spatial patterns of source activity. While the fMRI prior is used as a spatial constraint frequently in *non-dynamic* (or not temporally constrained) reconstruction methods (Dale et al., 2000; Sato et al., 2004; Daunizeau et al., 2007; Henson et al., 2010; Ou et al., 2010), it has yet to be applied in the above-mentioned *dynamic* (or state-space) methods. The fMRI prior in the proposed method is implemented similarly to the hierarchical variational Bayesian (hVB) method (Sato et al., 2004; Yoshioka et al., 2008). In forming this prior, the variance of the current noise (an input term driving the spatiotemporal dynamics of the MAR model) is weighted by the fMRI *t*-values. If all MAR coefficients are fixed at zero, this prior becomes identical to the fMRI prior proposed in Sato et al. (2004) and Yoshioka et al. (2008).

The present study unifies the MAR model, prior knowledge on the model parameters, and the measurement process of the current sources into a Bayesian framework. To improve stability of the estimated source dynamics, this framework also includes a sparse prior on the MAR coefficients. All hidden parameters in the unified probabilistic model (such as source amplitudes and the MAR coefficients) are jointly estimated by a variational Bayesian algorithm (Attias, 1999; Sato, 2001). The update rules are similar to those proposed in Fukushima et al. (2012), enabling inference of a high-dimensional dynamic model within a reasonable computation time.

Our method estimates the effective connectivity in the source space without requiring the selection of regions of interest (ROIs). To this end, the source dynamics are formulated using the full MAR model, and the source amplitudes and interactions are estimated simultaneously over the whole brain. These extensions allow exploratory analysis of the integration of brain functions, which complements the confirmatory approach of dynamic causal modeling (DCM; Friston et al., 2003; David et al., 2006). In contrast to our method, DCM initially assigns a small number of ROIs as network nodes, and then examines the validity of the network solutions by post hoc comparison of the model evidence.

The proposed method is quantitatively and qualitatively evaluated on simulation and experimental data, respectively. The results are compared with those of the hVB method, and of MNE and LCMV as benchmark methods. First, we examine the identification accuracy of the MAR model, using data generated from the adopted dynamic source model. We then investigate the estimation performance under more realistic conditions by mimicking stimulus-evoked responses by a network of neural mass models (Jansen and Rit, 1995; David and Friston, 2003; David et al., 2005). Finally, we examine the physiological plausibility of the estimates by application to a publicly available experimental dataset on face perception (Henson et al., 2011). Since the proposed method is a dynamic extension of the hVB method, we refer to it as the dynamic hVB method when comparing the methods.

This paper is organized as follows. The Theory section explains the model formulation and the adopted estimation algorithm. Model construction from the data and schemes for evaluating the estimation performance are described in the Methods section. The next two sections present the settings and results of the evaluation studies. Next, we investigate whether the free energy can be used for model comparison. Finally, we summarize the significance of the present study and discuss the advantages and limitations of the proposed method.

## Theory

### Notation

The following notations are used throughout this paper. $P(x)$ denotes the probability distribution of $x$ and $P(x|y)$ denotes the

conditional probability distribution of $x$ given $y$. $\mathcal{N}(\mathbf{x}|\bar{\mathbf{x}}, \boldsymbol{\Sigma})$ represents the multivariate Gaussian distribution of $\mathbf{x}$ with mean $\bar{\mathbf{x}}$ and covariance matrix $\boldsymbol{\Sigma}$, and $\Gamma(x|\bar{x}, \gamma)$ is the gamma distribution of $x$ with mean $\bar{x}$ and shape parameter $\gamma$ (the scale parameter corresponds to $\bar{x}\gamma^{-1}$). $\langle x \rangle_{P(x)}$ denotes an expectation value of $x$ with respect to $P(x)$. The vector $(\mathbf{x})_C$ contains a subset of $\mathbf{x}$ whose indices are included in a set $C$ in ascending order. Similarly, the matrix $(\mathbf{X})_{C,C}$ contains a subset of $\mathbf{X}$ whose row and column indices are included in $C$. As a special case of this notation, the subscript $k$, $k$ represents the $k$, $k$-th matrix entry. A diagonal matrix with diagonal entries $\mathbf{x}$ is denoted by $\mathrm{diag}(\mathbf{x})$, and the trace of a matrix $\mathbf{X}$ is designated $\mathrm{tr}(\mathbf{X})$. For notational simplicity, 1, …, $N$ and $x_1$, …, $x_N$ are replaced by 1:$N$ and $x_{1:N}$, respectively.

*Observation model*

The magnetic fields observed by MEG sensors and the source distribution over the cortex are linearly related as follows (Hämäläinen et al., 1993):

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

where $\mathbf{B}_t$ and $\mathbf{J}_t$ denote the measurement data and current source activities, respectively, $\mathbf{G}$ is the lead field matrix, $\boldsymbol{\varepsilon}_t$ is the observation noise, and $t$ is an index of time samples (where $t \le T$). $\mathbf{B}_t$ and $\mathbf{J}_t$ are $M$- and $N$-dimensional vectors that vertically concatenate the single channel data $B_{m,t}$ with $m \in \{1{:}M\}$ and the single source activity $J_{n,t}$ with $n \in \{1{:}N\}$, respectively. $\mathbf{G}$ is obtained by solving the forward problem (Mosher et al., 1999), accounting for the structure of the human head (details are described in the Methods section). Here the source orientations are assumed as fixed and perpendicular to the cortical surface. The noise $\boldsymbol{\varepsilon}_t$ is assumed to follow a Gaussian distribution $\mathcal{N}\left(\boldsymbol{\varepsilon}_t|\mathbf{0}, \beta^{-1}\mathbf{S}\right)$, where $\beta$ is a scaling parameter and $\mathbf{S}$ is the noise covariance matrix scaled by $\beta$ ($\mathbf{S}$ is normalized to satisfy $\mathrm{tr}(\mathbf{S}) = M$ and is typically determined from pre-stimulus rest period measurements). The likelihood function for the assumed Gaussian observation noise is written as

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

*Dynamic source model*

The spatiotemporal dynamics of the current sources is directly formulated in the source space by a full MAR model, in which the MAR coefficients represent the source interactions across the whole brain. The MAR coefficients to be estimated, and the time lags between pairs of sources, are based on anatomical brain connectivity inferred from dMRI.

We begin by modeling the dynamics of a single source $J_{n,t}$, which linearly interacts with its anatomically connected sources as follows:

$$J_{n,t} = \sum_{k \in C_n} \left(a_{n,k} J_{k,t-\delta_{n,k}}\right) + \omega_{n,t}, \tag{3}$$

where $C_n$ is an index set of source locations that are anatomically connected to the $n$-th source (including self-connection), $a_{n,k}$ and $\delta_{n,k}$ are the MAR coefficient and time lag, respectively, for interconnection between the $n$, $k$-th source pair, and $\omega_{n,t}$ is the current noise of the $n$-th source. The time lag of the $n$, $k$-th source pair is determined from the length of the anatomical connection $l_{n,k}$, the axonal conduction velocity $v$, and the local delay constant $\tau$:

$$\delta_{n,k} = \frac{l_{n,k}}{v} + \tau. \tag{4}$$

Ways of determining the anatomical connections and the delay parameters are explained in the Methods section. Since a single time lag based on the fiber length is assumed for each pair of anatomically connected sources, the MAR coefficient $a_{n,k}$ itself can be interpreted as a measure of the effective connectivity from the $k$-th to the $n$-th sources.

The dynamic source model (Eq. (3)) of all current sources over the whole brain can be expressed as a single equation:

$$\mathbf{J}_t = \sum_{l=1}^{L} \left(\mathbf{A}_l \, \mathbf{J}_{t-\Delta_l}\right) + \boldsymbol{\omega}_t, \tag{5}$$

where $\Delta_{1:L}$ contains all time lags in Eq. (3) for all $n$ (i.e., the time lags of all anatomically connected pairs of sources) in ascending order. Since a single time lag is assigned to each pair of anatomically connected sources, most entries of $\mathbf{A}_l$ for $l \in \{1{:}L\}$ are zero. The estimated parameters in $\mathbf{A}_l$ are only the entries corresponding to anatomically connected source pairs with the specific time lag $\Delta_l$. Therefore, the locations of non-zero MAR coefficients among the $\mathbf{A}_l$ (where $l \in \{1{:}L\}$) never overlap. Thus, for simplicity, we sum all the $L$ $\mathbf{A}_l$ matrices to define the MAR matrix $\mathbf{A}$:

$$\mathbf{A} = \sum_{l=1}^{L} \mathbf{A}_l, \tag{6}$$

where the original values of the MAR coefficients are not superimposed. The current noise for all current sources is denoted by $\boldsymbol{\omega}_t$ and is assumed to follow a Gaussian distribution $\mathcal{N}\left(\boldsymbol{\omega}_t|\mathbf{0}, (\beta\,\mathrm{diag}(\mathbf{q}))^{-1}\right)$. Here the parameter $\mathbf{q}$ in the variance term is multiplied by $\beta$. We use this parameterization as an extension of the previous source reconstruction methods (Sato et al., 2004; Nummenmaa et al., 2007; Yoshioka et al., 2008; Fukushima et al., 2012). The assumed model of source dynamics and current noise is equivalent to imposing the following prior distribution on the current sources:

$$P\left(\mathbf{J}_t \mid \mathbf{J}_{t-\Delta_1}, \ldots, \mathbf{J}_{t-\Delta_L}, \beta, \mathbf{A}, \mathbf{q}\right) = \mathcal{N}\left(\mathbf{J}_t|\sum_{l=1}^{L}\left(\mathbf{A}_l\mathbf{J}_{t-\Delta_l}\right), (\beta\,\mathrm{diag}(\mathbf{q}))^{-1}\right). \tag{7}$$

The current noise $\boldsymbol{\omega}_t$ in Eq. (5) can be regarded as an input to the dynamic source model. Thus, regions with higher current noise yield higher source amplitude, although activity depends not only on the input term but also on the source activity predicted by the dynamic model. Therefore, the variance of the current noise can be used to specify the spatial prior from fMRI data. In our method, the fMRI prior is constructed by assigning higher-magnitude input terms (i.e., higher current noise variance) to regions of higher $t$-value. This is achieved by basing the prior distribution of $\mathbf{q}$ on the magnitude of fMRI $t$-values (as described in the next subsection).

*Prior distributions on model parameters*

The unknown parameters estimated from the data in the observation and dynamic source models (Eqs. (2) and (7), respectively) are $\beta$, $\mathbf{A}$, and $\mathbf{q}$. For these parameters, the following prior distributions are assumed. The prior distribution of the scaling parameter $\beta$ is a non-informative Jeffreys prior:

$$P(\beta) \propto \beta^{-1}. \tag{8}$$

For the MAR matrix $\mathbf{A}$, the following Gaussian prior distribution is imposed on the MAR coefficients in an element-wise manner:

$$\begin{aligned} P(\mathbf{A}|\boldsymbol{\eta}_{1:N}) &= \prod_{n=1}^{N} P(\mathbf{a}_n|\boldsymbol{\eta}_n) = \prod_{n=1}^{N} \mathcal{N}\left(\mathbf{a}_n|\mathbf{0}, \mathrm{diag}(\boldsymbol{\eta}_n)^{-1}\right) \\ &= \prod_{n=1}^{N}\prod_{k=1}^{K_n} \mathcal{N}\left(\mathrm{a}_{n,k}|0, \eta_{n,k}^{-1}\right), \end{aligned} \tag{9}$$

where $\mathbf{a}_n$ is a column vector containing all non-zero entries of the $n$-th row of $\mathbf{A}$ in ascending order, and $K_n$ is the number of sources anatomically connected to the $n$-th source. In addition, we impose the Gamma prior distribution on the inverse variance $\eta_{n,k}$ in Eq. (9), and thereby apply the automatic relevance determination (ARD) sparse prior (Neal, 1996) to the MAR coefficients:

$$P(\boldsymbol{\eta}_{1:N}) = \prod_{n=1}^{N} P(\boldsymbol{\eta}_n) = \prod_{n=1}^{N}\prod_{k=1}^{K_n} P\left(\eta_{n,k}\right) = \prod_{n=1}^{N}\prod_{k=1}^{K_n} \Gamma\left(\eta_{n,k}|\eta_0, g_0\right), \tag{10}$$

where $\eta_0$ is the mean and $g_0$ is the shape parameter of the Gamma prior distribution (common to all entries of $\boldsymbol{\eta}_{1:N}$). The ARD sparse prior, comprising Eqs. (9) and (10), effectively prunes the coefficients associated with very small source amplitudes, ensuring that solutions are robust toward noise and modeling errors. Finally, we set the prior distribution of the parameter $\mathbf{q}$ as

$$P(\mathbf{q}) = \prod_{n=1}^{N} P(q_n) = \prod_{n=1}^{N} \Gamma\left(q_n|\overline{\nu}_n^{-1}, \gamma_0\right), \tag{11}$$

where $\overline{\nu}_n$ is the prior mean of the current noise variance (scaled by $\beta$), and $\gamma_0$ quantifies the reliability of this prior. To incorporate the fMRI prior into Eq. (11), $\overline{\nu}_n$ is weighted by the fMRI $t$-value $w_n$, normalized from zero to one:

$$\overline{\nu}_n = \nu_0 + (m_0 - 1)\,\nu_0 \cdot w_n^2. \tag{12}$$

In Eq. (12), $\nu_0$ denotes the variance of the source amplitude averaged over the whole brain during a pre-stimulus period, obtained by Bayesian minimum norm estimation (Yoshioka et al., 2008), and $m_0(>1)$ specifies the relative difference of variance intensity, ranked from the lowest (or statistically thresholded) to the highest $t$-value region. To properly represent the fMRI prior as good location information of MEG task-related changes, the experimental contrast of fMRI $t$-values should match that of MEG (i.e., both the baseline condition and the task condition of interest in fMRI and MEG should be identical). This setting is consistently used in the following analyses with the fMRI prior. Under this weighting scheme, regions of higher $t$-value receive larger task-related input activity; up to $m_0$ times larger than the changes associated with averaged resting activity.

*Joint probability distribution*

The likelihood function and all prior distributions (Eqs. (2), (7), (8), (9), (10), and (11)) constitute the joint probability distribution:

$$P(\mathbf{B}_{1:T}, \mathbf{J}_{1:T}, \beta, \mathbf{A}, \boldsymbol{\eta}_{1:N}, \mathbf{q}) = \prod_{t=1}^{T}\left\{P(\mathbf{B}_t|\mathbf{J}_t, \beta)\, P\left(\mathbf{J}_t|\mathbf{J}_{t-\Delta_1}, \ldots, \mathbf{J}_{t-\Delta_L}, \beta, \mathbf{A}, \mathbf{q}\right)\right\} \times P(\beta)\, P(\mathbf{A}|\boldsymbol{\eta}_{1:N})\, P(\boldsymbol{\eta}_{1:N})\, P(\mathbf{q}). \tag{13}$$

The following analyses assume that the analyzed data are trial-averaged evoked responses (apart from *Simulation 1*, where no fMRI prior is used) and that the time samples $-\Delta_L \le t \le 0$ ($\Delta_L \approx 70$ ms) constitute part of the rest period. Therefore, the current sources $\mathbf{J}_t$ (where $t \le 0$), are set to zero. By slightly modifying the dynamic source model, $\mathbf{J}_t$ ($t \le 0$) could be set as model parameters estimated from the data.

*Estimation algorithm*

All of the unknown variables in Eq. (13) are simultaneously estimated by calculating their approximate posterior distributions, such as those of the current sources and the MAR matrix. In real situations, the number of data is limited and non-observable sources would exist even with the mixing effect of the MAR matrix. Therefore, full identification is not necessarily achieved in the dynamic source model. Nevertheless, this estimation scheme enables an exploratory analysis of the effective connectivity in the whole brain, based on observable current sources.

If the probabilistic model is complicated, as in Eq. (13), exact calculation of the posterior distribution is generally intractable. Thus, we adopt the variational Bayesian algorithm (Attias, 1999; Sato, 2001) to analytically compute the approximate poster distribution $Q$. To this end, we first assume that the set of variables of interest contained in $X$ are independent

$$Q(X = \{\mathbf{J}_{1:T}, \beta, \mathbf{A}, \boldsymbol{\eta}_{1:N}, \mathbf{q}\}) = \prod_{t=1}^{T}\{Q(\mathbf{J}_t)\} Q(\beta)\, Q(\mathbf{A})\, Q(\boldsymbol{\eta}_{1:N})\, Q(\mathbf{q}), \tag{14}$$

and then minimize the Kullback–Leibler (KL) divergence of the true posterior distributions relative to the approximate distributions. Note that the additional independence approximation among $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$, applied to the variational Kalman smoother (VKS; Beal, 2003; Barber and Chiappa, 2007), reduces the computation time of deriving $Q(\mathbf{J}_{1:T})$ in high-dimensional (of order 1000) dynamic source models (Fukushima et al., 2012). Unlike VKS, where the time-variant source covariance matrices are computed during Kalman recursions, the source covariance in the proposed method becomes time invariant, except around the end of the time window for analysis (see Eqs. (A.9) and (A.10)). Despite this approximation, the means of the source posteriors are dependent across time points within the short time window of forward model predictions (again, see Eqs. (A.9) and (A.10)). Minimizing the KL divergence is identical to maximizing the free energy $F$, defined as

$$F = \langle \log P(\mathbf{B}_{1:T}, X) - \log Q(X)\rangle_{Q(X)}. \tag{15}$$

Since this equation can be converted to the log model evidence $\log P(\mathbf{B}_{1:T})$ minus the KL divergence, the free energy is a lower bound of the log model evidence. The factorization of $Q(\mathbf{J}_{1:T})$ in Eq. (14) is an additional approximation, but is exact in the non-dynamic equivalent case with $\mathbf{A} = \mathbf{0}$ (the hVB method; Sato et al., 2004; Yoshioka et al., 2008). Therefore, the free energy is a looser lower bound of the log model evidence in our method than in the hVB method. The free energy is monotonically increased by sequentially and iteratively updating the approximate distribution of $x$ (where $x$ is one of $\mathbf{J}_1$, …, $\mathbf{J}_T$, $\beta$, $\mathbf{A}$, $\boldsymbol{\eta}_{1:N}$, or $\mathbf{q}$) using the following update rule:

$$\log Q(x) = \langle \log P(\mathbf{B}_{1:T}, X)\rangle_{Q(X\setminus x)} + \text{const.} \tag{16}$$

$X \setminus x$ denotes the set of $X$ excluding $x$ (for instance, $X \setminus x = \{\mathbf{J}_{1:t-1}, \mathbf{J}_{t+1:T}, \mathbf{A}, \boldsymbol{\eta}_{1:N}, \mathbf{q}\}$ when $x = \mathbf{J}_t$), and const is independent of $x$. After convergence, the approximate posterior distributions maximally approach the true distributions in terms of the KL divergence, under the independence assumption of Eq. (14).

The pseudocode of the estimation algorithm is given below:

1. Initialize $Q(\mathbf{J}_1), \ldots, Q(\mathbf{J}_T)$ and the means of $Q(\boldsymbol{\eta}_{1:N})$, $Q(\mathbf{q})$ and $Q(\beta)$, and set the hyperparameters $\eta_0$, $g_0$, $m_0$, and $\gamma_0$ in Eqs. (10), (11), and (12).
2. **A**-step: Update $Q(\mathbf{A})$ in Eq. (A.1) by computing Eq. (A.2).
3. $\boldsymbol{\eta}$-step: Update $Q(\boldsymbol{\eta}_{1:N})$ in Eq. (A.3) by computing Eq. (A.4).
4. **q**-step: Update $Q(\mathbf{q})$ in Eq. (A.5) by computing Eq. (A.6).
5. **J**-step: Sequentially update $Q(\mathbf{J}_1), \ldots, Q(\mathbf{J}_T)$ in Eq. (A.7) by computing Eqs. (A.9) and (A.10).
6. $\beta$-step: Update $Q(\beta)$ in Eq. (A.11) by computing Eq. (A.12), and compute the free energy by Eq. (A.18).
7. If the maximization of the free energy converges, then exit; otherwise, return to the **A**-step.

Each update rule for computing the approximate posterior distribution is derived from Eq. (16) by substituting a variable of interest

for $x$ (the specific forms of the update equations are presented in Appendix A).

We iterated the algorithm 100 times in simulations and 500 times in real data applications. In the final iterations, the relative change in the free energy was below $10^{-5}$, the level at which the free energy is essentially converged. The computation time for 100 iterations ranged from approximately 12 to 48 h (iterations were performed in MATLAB 7.5.0, running on a multi-core PC: 8 cores, 3.60 GHz CPU, and 96 GB RAM), mainly depending on the number of anatomical connections in the individual brain.

### *Initial values and hyperparameters*

At the beginning of the iterative algorithm, where $Q(\mathbf{A})$ is first updated, $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$ and the means of $Q(\boldsymbol{\eta}_{1:N})$, $Q(\mathbf{q})$, and $Q(\beta)$ must be specified beforehand. As the optimization problem solved in the variational Bayesian algorithm has local maxima, bad initialization degrades the estimation performance. Therefore, setting appropriate initial values is important for obtaining reasonable estimates. In this study, the initial values of $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$ and the means of $Q(\mathbf{q})$ and $Q(\beta)$ were the non-dynamic solutions of the hVB method (the initialization of the hVB method is explained in the methods section below). The probabilistic model of the hVB method in Yoshioka et al. (2008) is identical to the model adopted in the proposed method if all entries of $\mathbf{A}$ are fixed at zero. The means of all elements of $Q(\boldsymbol{\eta}_{1:N})$ are initialized to their prior means $\eta_0$.

In addition, the hyperparameters $\eta_0$, $g_0$, $m_0$, and $\gamma_0$ that control the prior distributions of $\boldsymbol{\eta}_{1:N}$ and $\mathbf{q}$ must be pre-specified. The default values used in the simulations are listed below:

- The hyperparameter $\eta_0$, also used as the initial values for $\boldsymbol{\eta}_{1:N}$, suppresses instabilities in the estimated dynamics. If $\eta_0$ is very small, the MAR coefficients can amplify, which consequently leads to divergence of the source amplitude. On the other hand, if $\eta_0$ is extremely large, all estimated MAR coefficients are close to zero. Thus, we set the default $\eta_0$ to a moderate value (such as 100).
- The hyperparameter $m_0$ in the fMRI spatial prior regulates the maximum strength of the current noise variance relative to the baseline. We set the default $m_0$ to 100, a value recommended by Yoshioka et al. (2008) in the hVB method.
- In the simulations section, the hyperparameters $g_0$ and $\gamma_0$ were set by one of two strategies. When the data were generated from the MAR model identical to the adopted dynamic source model, both $g_0$ and $\gamma_0$ were set to zero, yielding non-informative priors. When discrepancies were found between our model formulation and the data generation process, $g_0$ and $\gamma_0$ were set to 100 to realize informative priors. Such priors enhance the robustness of the estimates to modeling errors (Fukushima et al., 2012).

The optimal hyperparameter values in the event of modeling errors were searched by replacing the values specified above (all 100) in the simulations section, and these values were applied in the real data section.

## Methods

This section briefly introduces the analyzed dataset, and presents the data processing procedures from which the constituents of the model are derived, namely, the anatomical connections, the time lags, and the lead field matrix. We then explain how the estimation performance is evaluated in simulations. Finally, we describe the procedures for estimating the source amplitudes and interactions, which are used in the hVB method and both benchmark methods (MNE and LCMV).

### *Dataset*

Simulation and real data analyses were conducted on a publicly available experimental dataset[1] (the data acquisition and task conditions are detailed in Henson et al., 2011). The dataset contains the anatomical images (T1- and diffusion-weighted MRI) and functional data (MEG and fMRI data collected during face perception) required for our method. Within this dataset, the entire set of MRI, dMRI, MEG, and fMRI data was archived for 11 out of 16 subjects (Sub01–Sub06, Sub09, Sub12–15); these data were used in the real data analysis. The incomplete data of the remaining six subjects were not used; dMRI data are absent for Sub07, Sub08, Sub10, Sub11, and Sub16. Simulation data were generated from the brain data of a single subject (Sub12). In the following, the eleven original subject codes (in ascending order) are renamed Subject I–Subject XI.

### *Diffusion MRI analysis*

#### *Data processing*

The anatomical connections and time lags for the dynamic source model were inferred from fiber tracking with structural and diffusion MRI data (T1 weighted and diffusion weighted images). All MRI data (including fMRI) were acquired on a 3T Trio (Siemens, Erlangen, Germany). The spatial resolutions of the T1-weighted and the diffusion weighted images were 1 mm and 2 mm isotropic, respectively. During dMRI acquisition, the number of gradient directions was 64 and the *b*-value was 1000 s/mm$^2$. To pre-process the dMRI data, subject motion during the image acquisition was corrected by FSL[2]; the consequent bias on the gradient directions was compensated afterwards by rotating the *b*-vector. The fractional anisotropy image was then calculated from the corrected images and was used for registering the diffusion-space to the T1-space by the non-linear registration tool (FNIRT) in FSL.

The seed and target ROIs used for fiber tracking were obtained by parcellating the cortical surfaces extracted by FreeSurfer.[3] In the cortical parcellation, 2000 vertices were first specified by applying the Matlab function *reducepatch* to the high resolution white/gray matter boundary surfaces provided by FreeSurfer. All of the highly-resolved surface vertices were then clustered into 2000 parcels in a nearest-neighbor manner. The vertices in subcortical regions were discarded, and the remaining parcels (numbering around 1840) constituted the final ROIs used for fiber tracking. These surface ROIs were converted to volume ROIs, and then transformed to the diffusion-space.

Fibers were tracked using MRtrix.[4] The local model of fiber orientations was the fiber orientation distribution (FOD), reconstructed at each voxel by constrained spherical deconvolution (Tournier et al., 2007) with six-dimensional spherical harmonics for the response function. Based on the reconstructed FOD, fibers were probabilistically tracked with a step length of 0.2 mm within a mask of white matter volume. The fiber tracks were generated $10^5$ times from each ROI and unidirectionally stepped until they entered another ROI or until the total fiber length reached 300 mm. The fibers were also terminated at voxels of FOD amplitude less than 0.1. The minimum fiber length and radius of curvature were set to 10 mm and 1 mm, respectively.

#### *Anatomical connections*

The anatomical connectivity was quantified from the fiber counts. The strength of the connectivity was computed as the number of fibers within each ROI pair $f_t$ divided by the total number of fibers generated from its seed ROI $f_s$ with voxel size normalization; that is, $(f_t / v_t) / (f_s / v_s)$, where $v_t$ and $v_s$ are the number of voxels in the target and the seed ROI, respectively. Since anatomical connectivity conveys no directional

[1] ftp://ftp.mrc-cbu.cam.ac.uk/personal/rik.henson/wakemandg_hensonrn/.
[2] http://www.fmrib.ox.ac.uk/fsl.
[3] http://surfer.nmr.mgh.harvard.edu/.
[4] http://www.nitrc.org/projects/mrtrix/.

information, the connectivity strengths were symmetrized by assigning the higher strength to both directions.

To determine the anatomically connected ROI pairs, the anatomical connectivity strengths were binarized. The binarized connections were used for specifying pairs of anatomically connected *sources* (and hence for determining the MAR coefficients to be estimated from the data), since we placed a single current source at each ROI when computing the lead field matrix (as explained later in this methods section). The binarization threshold was set as low as possible (in our case, $10^{-4}$). While a lower threshold prolongs the computation time, it reduces the risk of missing true anatomical connections. Indeed, a lower threshold is important for estimating the source interactions in our method. The false negative anatomical connections set the MAR coefficients to zero, even when the corresponding pair of sources is effectively connected. On the other hand, MAR coefficients with false positive anatomical connections are not necessarily increased; the sparse prior on the MAR matrix can make these coefficients go to zero. After the thresholding, between 5 and 9% of all ROI pairs were labeled as anatomically connected (the precise percentage depended on the individual subject). The final number of non-zero entries in the MAR matrix **A** in Eq. (6) was $(2.48 \pm 0.37) \times 10^5$ (mean ± s.d. of the eleven subjects).

*Time lags*

The time lags in the MAR model were computed from the inter-ROI fiber length, the conduction velocity, and the local delay constant (see Eq. (4)):

- The fiber length was obtained by first summing the tracking steps of each fiber track. The lengths of fibers included in each pair of ROIs were then averaged. The mean fiber lengths were finally symmetrized by weighted averaging of both directional fiber lengths by their (pre-symmetrized) connectivity strengths.
- The conduction velocity was set to 6 m/s, as assumed for myelinated axons in a previous simulation study (Ghosh et al., 2008).
- The local delay constant was defined as the time between an input signal and peak population-level neuronal activity evoked by this input. It was set to 27 ms, estimated from the peak to peak time difference between the source activities reconstructed in the primary and secondary visual areas (about 30 ms; Portin et al., 1999). Here we accounted for the time delay introduced by propagation of the source activity along the anatomical connection. This delay was estimated at about 3 ms, based on the distance between the primary and secondary visual areas and the above velocity setting.

The number of time lags L in Eqs. (5) and (6) was 35.5 ± 2.1 (mean ± s.d. of the eleven subjects). The time lags in the self-connections (the diagonal entries in the MAR matrix) were set to the local delay constant. The impact of conduction velocity and local delay constant on the estimation performance was investigated in the simulations section.

*Lead field matrix*

To calculate the lead field for each MEG sensor location, current dipoles of unit strength were successively placed at all cortical vertex positions, and the MEG forward solution was computed. The sensor positions were matched to the positions of 102 magnetometers in a VectorView system (Elekta Neuromag, Helsinki, Finland), and were linearly registered to the T1-space using the fiducials and head points provided in the real dataset. The dipole was oriented perpendicular to the white/gray matter boundary surface. The forward solution was computed by a boundary element method (BEM) using a single-shell head model, based on the inner skull surface obtained from FreeSurfer.

To estimate locally clustered source activities whose minimum cluster size is that of current activities generating measurable MEG signals outside the head (around 40 $mm^2$; Hämäläinen et al., 1993), we modeled locally distributed sources: $\mathbf{J}_t = \mathbf{W}\mathbf{Z}_t$, where **W** is a Gaussian smoothing filter and $\mathbf{Z}_t$ denotes the unsmoothed focal sources, whose vertex positions were specified in the ROI generation. The full width at half maximum (FWHM) of **W** was set to 8 mm, as in Yoshioka et al. (2008). In estimating $\mathbf{Z}_t$, we operationally replaced the original lead field matrix **G** with the smoothed one: $\hat{\mathbf{G}} = \mathbf{GW}$. Smoothing was performed in a vertex space spatially finer than the source space (i.e., the row number in **W** is larger than the column number). By this procedure, we can detect current source activities located between the vertex positions of the $\mathbf{Z}_t$ entries. Using $\mathbf{Z}_t$ and $\hat{\mathbf{G}}$, we estimated the MAR matrix **A** for $\mathbf{Z}_{1:T}$ rather than $\mathbf{J}_{1:T}$. To clearly demonstrate the correspondence between the source dynamics and the MAR coefficients, all results presented in the following sections are based on the unsmoothed sources $\mathbf{Z}_{1:T}$.

*Performance evaluation*

The estimation performance in simulations was quantified by two threshold-free evaluation scores; the area under the receiver operating characteristic (ROC) curve and the normalized root mean square (RMS) error.

The area under the ROC curve (AUC) measures the detection accuracy. The AUC scores were derived by a procedure originally proposed by Grova et al. (2006), which compensates the severe imbalance between a small number of actives and a large number of inactives. In this procedure, the numbers of inactives and actives were equalized by selecting a subset of inactives, subject to the rules presented in Appendix B. Following the selection step, the AUC was computed in the usual manner. Hits and misses of sources were defined for the RMS of the estimated current amplitudes (maximum RMS over the time window), and those of connections were determined for the estimated absolute MAR coefficients. Both quantities were normalized by their maximum values over all sources or connections. By comparing these normalized estimates to the ground truth, we computed the numbers of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) events for thresholds ranging from zero to one. We then calculated the sensitivity and specificity for each threshold, defined as

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \text{ and specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \tag{17}$$

A plot of sensitivity versus (1−specificity) yields the ROC curve. The AUC was determined by the trapezoidal rule.

The normalized root mean square error (nRMSE) quantifies the differences between the estimates and the ground truth. The nRMSE scores of the current sources (or the MAR matrix if the ground truth existed) were computed as the RMS error between the simulated and estimated source amplitudes over all sources and time instances (or MAR coefficients over all connections), divided by the RMS of the simulated ones.

When reconstructing the simulated MAR time series, we evaluated the estimation accuracy by AUC, nRMSE, and two additional evaluation measures that provide an intuitive understanding (see *Simulation 1* in the simulations section). One additional measure is the reconstruction gain, defined as the ratio of the time-averaged estimated source amplitude to the true source amplitude. The other is the non-normalized RMSE of the MAR coefficients.

We also applied thresholds on the estimated sources and MAR coefficients when mapping the results onto the cortical surfaces and when distinguishing between active and inactive sources and connections. In particular, these thresholds were used for qualitative evaluation of the estimation performance when the method is applied to experimental data. Other than when calculating AUC, a source was considered active if its magnitude once exceeded one-tenth of the maximum over all sources. A connection was regarded as effectively connected if its absolute MAR coefficients exceeded 0.1. Although these thresholds were determined in an ad-hoc manner, small threshold differences did not alter

the relative characteristics of the estimates in the following method comparison.

*Method comparison*

To demonstrate the advantages of dynamic modeling and simultaneous estimation, the estimates obtained by our method (the dynamic hVB method) were compared with those of the hVB method, MNE, and LCMV. In all of these methods, source reconstruction with no dynamic constraint was followed by source interaction estimation via the MAR model (Eq. (3)).

*Hierarchical variational Bayesian (hVB) method*

In the first step, the current sources were estimated by the dynamic hVB method with all MAR matrix entries set to zero. The MAR coefficients were then separately estimated from the reconstructed current sources. These coefficients were computed only for anatomically connected pairs of active sources; the coefficients of unconnected pairs and pairs including any inactive sources were set to zero.

In the source reconstruction, the approximate posterior distributions were updated in the following order: $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$, $Q(\beta)$, and $Q(\mathbf{q})$. This ordering requires that only the mean of $Q(\mathbf{q})$ be prespecified. Here, we initialized the $n$-th entry of the mean of $Q(\mathbf{q})$ to the prior mean $\overline{\nu}_n^{-1}$ in Eq. (11). When estimating the source interactions from Eq. (3), the reconstructed active sources were first extracted, and the approximate posteriors of $Q(\mathbf{A})$, $Q(\boldsymbol{\eta}_{1:N})$, and $Q(\mathbf{q})$ were sequentially updated. At the beginning of the iterative updates, $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$ and the means of $Q(\beta)$ and $Q(\mathbf{q})$ were set to the estimates of the hVB method computed in source reconstruction, and the means of all elements of $Q(\boldsymbol{\eta}_{1:N})$ were set to $\eta_0$.

*Benchmark methods*

MNE (Hämäläinen and Ilmoniemi, 1994) and LCMV (Van Veen et al., 1997) were used as benchmarks for method comparison. Both methods have been recently used as inverse solvers for obtaining MEG-based functional brain networks (Brookes et al., 2011; Hipp et al., 2012; de Pasquale et al., 2012).

As an MNE algorithm, we used a Bayesian type of MNE, which derives the optimal regularization parameter from the variational Bayesian inference. The algorithm was run by the dynamic hVB method with all entries of $\mathbf{A}$ set to zero, $q_n$ replaced with $q\,\overline{\nu}_n^{-1}$, and $q$ estimated from the data. In this approach, the current sources are estimated by a weighted MNE algorithm (wMNE) if the fMRI prior is provided. In the source reconstruction, $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$, $Q(\beta)$, and $Q(q)$ were sequentially updated with the mean of $Q(q)$ initialized to 1. To estimate the source interactions, the iterative algorithm in the hVB method was used, fixing $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$ and the mean of $Q(\beta)$ at the MNE (or wMNE) estimates. The mean of the $n$-th element of $Q(\mathbf{q})$ was initialized to $\overline{q}\,\overline{\nu}_n^{-1}$, where $\overline{q}$ is the mean of $Q(q)$ estimated by MNE (or wMNE). $Q(\boldsymbol{\eta}_{1:N})$ was initialized as implemented in the hVB method.

The source reconstruction steps of LCMV are detailed in Appendix C. When the fMRI prior was provided, the source space was restricted to the fMRI-positive regions. Source interactions for LCMV were estimated by the same iterative algorithm applied in the hVB method and MNE. $Q(\mathbf{J}_1)$, …, $Q(\mathbf{J}_T)$ were replaced with the point estimates of current sources reconstructed by LCMV, and the mean of $Q(\beta)$ was the inverse variance of the observation noise during a resting period. The mean of the $n$-th element of $Q(\mathbf{q})$ was initialized by setting $(\overline{\beta}\overline{q}_n)^{-1}$ to the $n$-th source power reconstructed by LCMV, where $\overline{\beta}$ is the above-specified mean of $Q(\beta)$. $Q(\boldsymbol{\eta}_{1:N})$ was initialized by the procedure common to both the hVB method and MNE.

## Simulations

*Simulation 1: MAR time series*

We first examined the estimation performance on a simulated MAR time series. The data were generated from the dynamic source model assumed in the proposed method. The purpose of this simulation was to evaluate the identification accuracy of the MAR model with varying numbers of active sources and effective connections.

*Settings*

The number of active sources was set to 5, 10, 20, or 30 out of 1841 cortical sources. The time courses of the active sources were generated from the dynamic source model. The diagonal MAR matrix entries of active sources were set to 0.4, while anatomically-connected non-diagonal entries were assigned random values uniformly distributed between −0.4 and 0.4. All other entries of the MAR matrix were set to zero. The current noise variance of active and inactive sources was set to 20 $nAm^2$ and zero, respectively. Gaussian simulated observation noise was assumed; the normalized noise covariance matrix was computed from empty room data obtained from the real dataset. The scaling parameter was adjusted to yield a signal to noise ratio[5] of 5 dB in the simulated sensor measurements, a typical value in trial-averaged measurements of simple stimulus-evoked responses.

For each number of active sources specified above, we conducted 20 Monte Carlo simulation trials. In each simulation, the active source positions were sequentially and randomly selected from source positions that were anatomically connected to one of the selected source positions. The current sources and the MAR matrix were estimated from the simulated measurements (duration 400 ms and sampling frequency 1 kHz) with known noise covariance structure and time lags. Throughout this simulation study, non-informative priors were imposed on the model parameters (i.e., $g_0 = \gamma_0 = 0$) and no fMRI information was provided.

*Results*

The estimation accuracies were first examined by evaluating the AUC and nRMSE of the estimated current sources and MAR matrices (see Fig. 1). When calculating the AUC scores, the simulated source positions and non-zero MAR coefficients were assumed as the true sources and connections. The hVB and dynamic hVB methods yielded almost identical AUCs for the current sources. However, the AUC of the MAR matrix and the nRMSE of the current sources and the MAR matrix were improved in the dynamic hVB method. This demonstrates the advantage of explicitly accounting for the source dynamics, and simultaneously reconstructing the current sources and MAR coefficients. As also shown in Fig. 1, AUCs of both the current sources and MAR matrix were also higher in the dynamic hVB method than in MNE and LCMV. This indicates that the dynamic hVB method outperforms existing approaches in terms of detection accuracy. The nRMSE scores of MNE were almost identical to unity for both the current sources and MAR matrix. In LCMV, the nRMSE is not a useful metric because the weight normalization (see Appendix C) alters the magnitudes of the reconstructed sources.

To investigate the estimates of the dynamic hVB method in detail, the reconstruction gain of the current sources and the RMSE of the MAR matrix were computed. The results are listed in Table 1. The reconstruction gains were obtained by averaging the gains across the active source positions (upper line) and the positions at which the true positive sources were successfully estimated (lower line). The RMSE scores were calculated from a subset of the MAR matrix, restricting the row and column indices to the locations of the

[5] The signal to noise ratio is defined as $10\log_{10}\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\hat{B}_{m,t}^2/MT\beta^{-1}\right)$, where $\hat{B}_{m,t}$ denotes the noise-free simulated measurements on the $m$-th sensor at time point $t$.

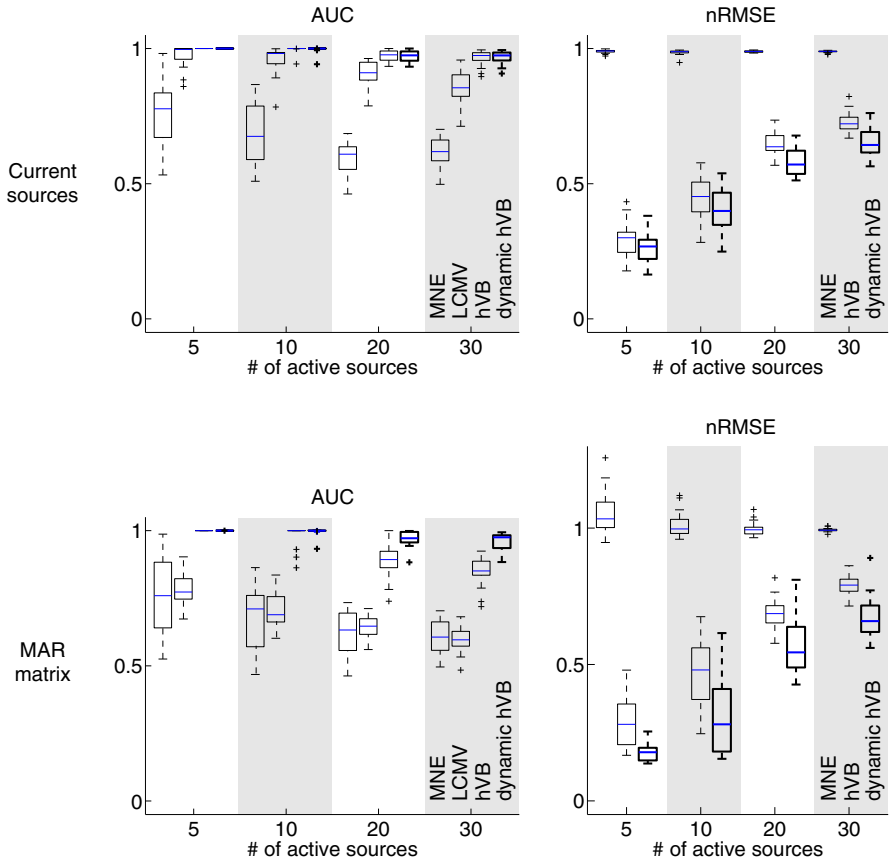


**Fig. 1.** Evaluation of the estimation accuracy in *Simulation 1*. AUC and nRMSE of the current sources (upper) and the MAR matrix (lower) are presented as boxplots. The AUC and nRMSE distributions are shown in the left and right panels, respectively. The blue bar of each box indicates the median and the edges show the upper and the lower quartiles. The length of the whisker is the default length set in the Matlab function *boxplot*.

originally active sources (upper line) and the true positive sources (lower line). This table reveals near-perfect source reconstruction and model identification for 5 active sources. As the number of active sources increased, the estimation performance deteriorated. In addition, sources at less observable cortical positions (with small lead field norms) failed to be identified (data not shown). Similarly, connections with small products of the corresponding lead field norms tended to be missed. Moreover, the estimation accuracy decreased for smaller-magnitude current sources and MAR coefficients among the simulation data. Despite these anomalies, Table 1 shows that even for source numbers of 20 and 30, more than 75% (true positive rate) of the originally active sources were reconstructed with 1) reconstruction gains greater than 0.80 and 2) RMSE of the MAR matrix less than 0.15.

**Table 1**
Reconstruction gains of the current sources and RMSEs of the MAR matrix, estimated from the originally active and true positive sources in *Simulation 1*.

| Number of active sources | 5 | 10 | 20 | 30 |
|---|---|---|---|---|
| Reconstruction gain | | | | |
| of the active sources | 0.95 | 0.88 | 0.73 | 0.66 |
| of the true positive sources | 0.95 | 0.91 | 0.85 | 0.84 |
| RMSE of the MAR matrix | | | | |
| between the active sources | 0.05 | 0.09 | 0.15 | 0.18 |
| between the true positive sources | 0.05 | 0.07 | 0.10 | 0.12 |
| True positive rate | 1.00 | 0.97 | 0.85 | 0.78 |

The displayed values were computed from a subset of the estimates obtained by the dynamic hVB method. Each entry displays the mean of 20 simulations.

*Simulation 2: Stimulus-evoked responses*

Next, to examine the estimation accuracy in a more realistic simulation setting, we applied the methods to synthesized stimulus-evoked responses. The event-related dynamics inherent in the simulation data was generated from a functional network of the neural mass models with external inputs (for details of the neural mass model simulations, see Appendix D).

*Settings*

A schematic of the simulation setting is shown in Fig. 2. The time courses of stimulus-evoked responses, with sampling frequency 1 kHz, were generated from a network of non-linear neural mass models (the constituents of this network model are detailed in Appendix D). When computing the exact time courses of the active sources, the waveforms generated from the network were multiplied by a constant to ensure that the resulting sensor measurements and the real measurements in the dataset were of comparable order of magnitude. We adopted a simplified network model of a face perception task comprising six regions of interest; the bilateral lower visual cortices around the calcarine sulcus (CAL) and the higher visual cortices located in the core regions of face perception (Haxby et al., 2000), the inferior occipital gyrus (IOG) and the fusiform gyrus (FG). The remaining core region, the superior temporal sulcus (STS), was excluded from the network model because direct anatomical connections between STS and IOG/FG were not found in previous dMRI studies (Gschwind et al., 2012; Pyles et al., 2013).

We examined the estimation performance under two simulation scenarios (see Figs. 2A and B) with different numbers of active sources

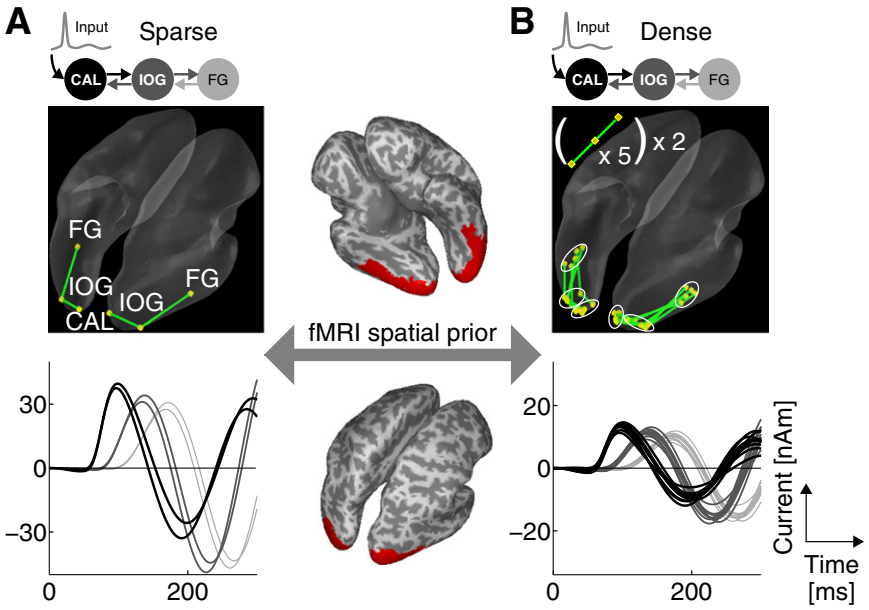


**Fig. 2.** Data generation settings in *Simulation 2*. (A) sparse setting (B) dense setting. The time courses of the active sources in CAL, IOG, and FG are plotted in black, dark gray, and light gray, respectively. The red regions on the cortical surfaces are the fMRI positive regions assumed in this simulation study.

in each ROI; one (the sparse setting) and five (the dense setting). In the sparse setting, each active source in the CAL of both hemispheres received an input signal. The CAL activity was sequentially transmitted to the active sources in IOG and FG in a one-to-one manner via the bidirectional effective connectivity between CAL–IOG and IOG–FG (the input signal and connection parameters are also detailed in Appendix D). A similar connectivity profile was applied in the dense setting, where five sets of CAL–IOG–FG source interactions were assumed in each hemisphere (see Fig. 2B). Note that all forward and backward effective connections were consistent with their anatomical connections. For simplicity, the lateral effective connectivity within each ROI was ignored in this simulation.

An identical fMRI prior was assumed in both sparse and dense settings. To obtain this spatial prior, we first binarized the $t$-value map of the real fMRI data in the dataset, and then masked the binarized areas by extracting regions close to CAL, IOG, and FG (the $t$-value map was computed as described in the real data section). The fMRI positive regions, spread across 69 cortical source positions, include all active sources placed in both simulation settings. Since the $t$-values were binarized, the fMRI positive and negative regions were assigned $w_n = 1$ and $w_n = 0$, respectively, in Eq. (12). The advantages of applying the fMRI prior were assessed by comparing the results with those obtained without the fMRI prior (assigning $w_n = 1$ to the whole cortex).

The observation noise was real background MEG data in the dataset, comprising trial-averaged sensor measurements taken during a 400 ms rest period. During the trial averaging, potential event-related components in the resting data were suppressed by randomly forward-shifting the time course of each trial, then translating the fragmented data projected from the time window to the end of the shifted time series. The magnitudes of the resulting trial-averaged measurements were comparable to those of the actual resting data. The first 100 ms duration of the averaged background MEG data were used to fit the normalized noise covariance matrix; the remaining 300 ms were added to the noise-free simulated measurements.

Various modeling errors typically encountered in real data applications, such as non-linear source dynamics and non-Gaussian observation noise, were introduced to this simulation study. To prevent the entry of possible estimation errors caused by such model discrepancies, we applied informative priors (including the fMRI prior) on the model parameters; that is, we set $g_0$ and $\gamma_0$ to non-zero values. Optimal values of $g_0$ and $\gamma_0$, as well as $\eta_0$ and $m_0$, were searched by varying the values of the hyperparameter sets $(\eta_0, g_0)$ in $P(\boldsymbol{\eta}_{1:N})$ and $(m_0, \gamma_0)$ in $P(\mathbf{q})$ from their default settings (100, 100).

Besides varying the hyperparameters, we also examined the impact of changing the delay parameters in the MAR model, namely, the conduction velocity $v$ and the local delay constant $\tau$, from their default values ($v = 6$ m/s and $\tau = 27$ ms).

*Results*

The current sources and the MAR coefficients estimated with the fMRI prior are shown in Fig. 3. In the sparse setting, both the hVB and dynamic hVB methods successfully reconstructed the active sources and identified their effective connectivity (see Fig. 3A), although the hVB method detected a single false positive source. The benchmark methods with the fMRI prior, namely, wMNE and LCMV, did not correctly estimate the sparse source distribution. Furthermore, the MAR coefficients estimated by these methods were close to zero for several true source connections. The magnitude of the current sources was underestimated in wMNE, and LCMV generated incorrect waveforms in the presence of mutually correlated active sources and real background noise. In the dense setting, none of the methods correctly estimated dense cortical activity and effective connectivity (see Fig. 3B). False positive sources were estimated within the regions covered by the fMRI prior; and false positive connections emerged due to the limitations of the linear autoregressive modeling. In fact, such errors in the estimated connections occurred even when the MAR coefficients were fitted to the true current sources (data not shown). Despite these shortcomings, the dynamic hVB method showed higher sensitivity of current source reconstruction and identification of effective connectivity than the hVB method (see Fig. 3B). In the dynamic hVB method, 23 out of 30 active sources were reconstructed and at least one single effective connection was discovered at each functional ROI pair. In contrast, the hVB method missed half of the active sources (15 out of 30), and detected no effective connection for several functional ROI pairs. The current source amplitude was properly estimated in wMNE and the effective connections were largely detected by wMNE and LCMV. However, both benchmark methods yielded a number of false positive sources and connections.

To clarify the differences among the four methods in the dense setting, we computed ROI-wise source time courses and source

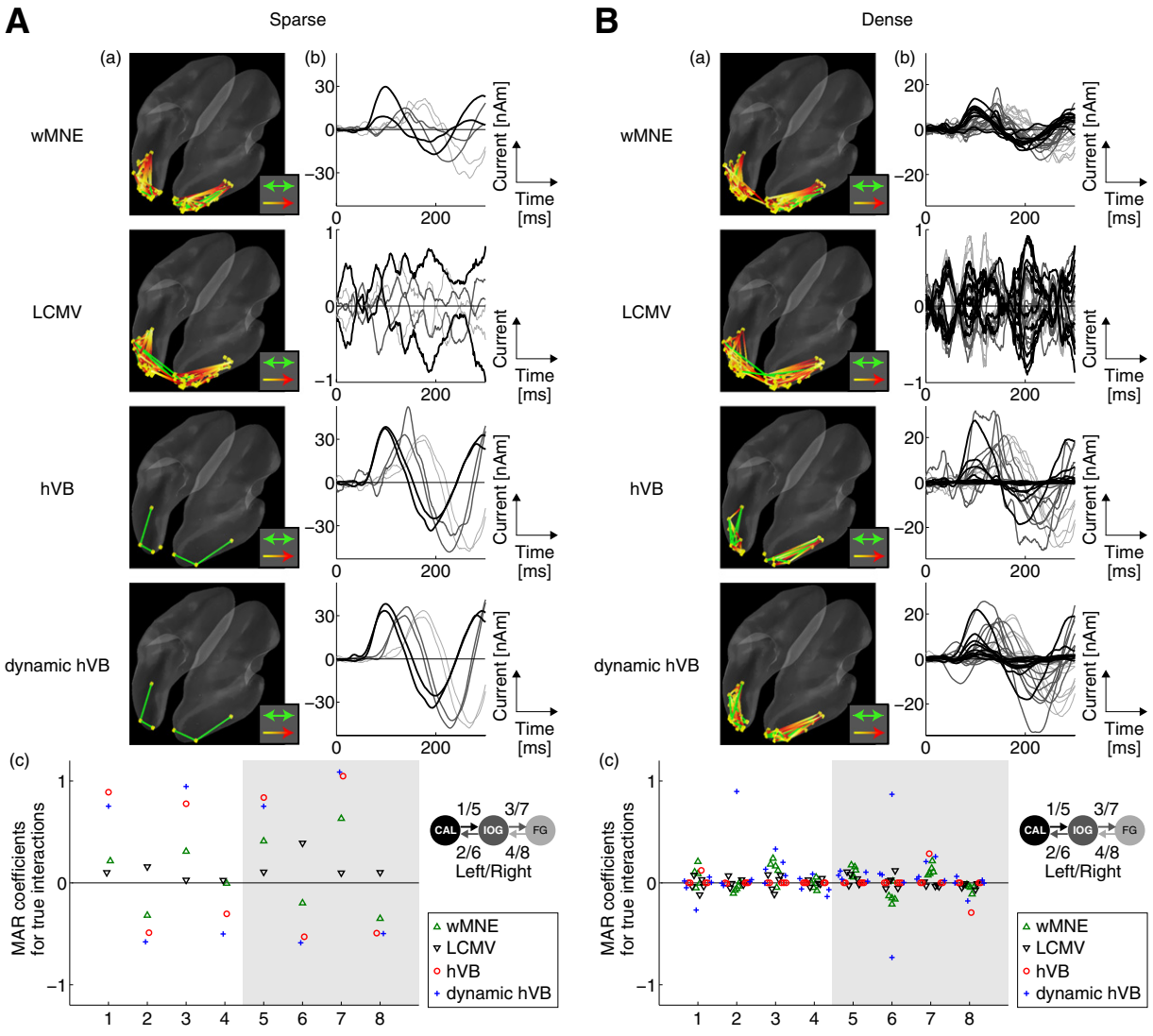


**Fig. 3.** Results obtained from wMNE, LCMV, the hVB method, and the dynamic hVB method with the fMRI prior in *Simulation 2*. (A) sparse setting and (B) dense setting. (a) Estimated effective connectivity among the estimated active sources indicated by yellow dots. Green lines, and lines graduated from yellow to red, denote bi- and uni-directional effective connectivity, respectively. Note that line color is not related to the magnitude of the MAR coefficients. (b) Estimated source time courses on the cortical vertices, on which the true active sources were placed. Colors of time courses (gray scale) are explained in the caption of Fig. 2. The current sources of LCMV are normalized by their maximum in the plotted time courses. (c) Estimated MAR coefficients for pairs of sources between which true interactions were assumed. The ground truth of the MAR coefficients cannot be defined because the data were not generated from MAR models but neural mass models. The horizontal axis indicates the locations of inter-ROI connections (1–4: connections in the left hemisphere; 5–8: connections in the right hemisphere).

interactions. Fig. 4A presents the time courses of the five active sources averaged within each ROI. The ROI-averaged source time courses were reasonably reconstructed by all methods except LCMV. Fig. 4B shows the source interactions, quantified by paring the ROIs and averaging the absolute MAR coefficients within each pair. Averaged source interactions are defined only for ROI pairs linked by at least one anatomical connection. For such a pair, the sum of the absolute MAR coefficients was divided by twice the number of anatomical connections. In the hVB method and LCMV, the overall magnitudes of the averaged MAR coefficients were low. The dynamic hVB method showed higher sensitivity and specificity than wMNE at the plotting threshold of Fig. 4B (AUC scores for these graphs are presented in Table 2).

The AUC and nRMSE scores with and without the fMRI prior are displayed in Table 2. Clearly, the estimation performance was improved by applying the fMRI prior. In the sparse setting, the best score was consistently achieved by the dynamic hVB method with the fMRI prior. The AUCs in wMNE and LCMV with the fMRI prior were close to 1, despite the appreciable number of false positives (see Fig. 3A). This result is attributed to the low detectability of the AUC scores to the false positives near the true source positions (see Appendix B for the computation procedure). In the dense setting, the nRMSE score in the dynamic hVB method with the fMRI prior was approximately unity (0.86), reflecting the difficulty of reconstructing a large number of closely-spaced active sources. These activities were better estimated by wMNE (nRMSE = 0.72). Nevertheless, the dynamic hVB method adequately reconstructed the mean activity across the sources within an ROI (see Fig. 4A). The nRMSE score computed from the ROI-averaged time courses (denoted by ROI-nRMSE in Table 2) was improved from 0.49 in wMNE to 0.39. Similarly, the AUCs of the MAR matrix were higher in wMNE and LCMV than in the dynamic hVB method, whereas the dynamic hVB method yielded the highest AUC score for the ROI-averaged MAR coefficients (denoted by ROI-AUC in Table 2). Here, the ROI-AUC was computed by a standard AUC procedure of the eight active and six inactive interactions among the inter-regional anatomical connections shown in Fig. 4B.

The evaluation scores obtained under various hyperparameter settings are presented in Fig. 5A. The hyperparameters $\eta_0$ and $g_0$

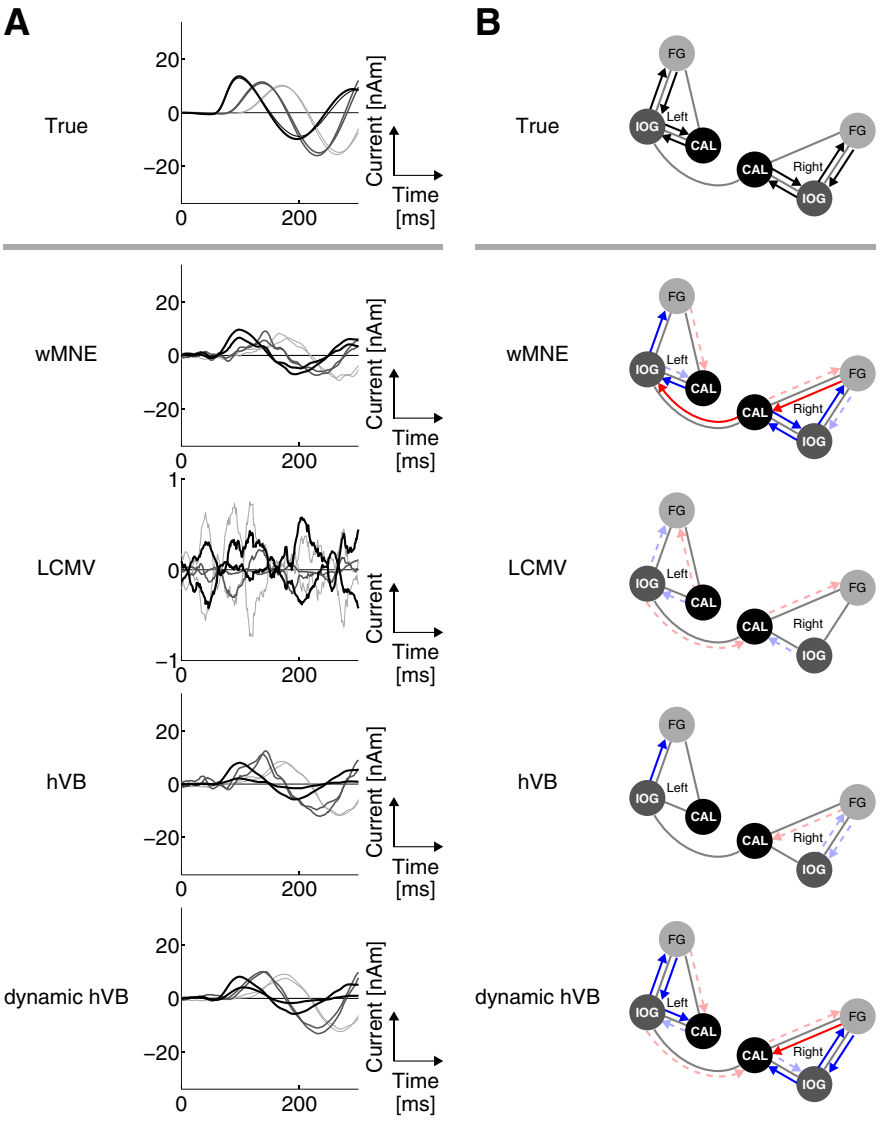


**Fig. 4.** ROI-averaged results in the dense setting of *Simulation 2*. The ground truth is shown in the top row. (A) ROI-averaged source time courses. Colors of the time courses are explained in Fig. 2. The time courses in LCMV were normalized by the maximum current. (B) ROI-averaged source interactions. Anatomically connected ROI pairs are linked by thick gray lines. True and false positive interactions are indicated by blue and red arrows, respectively. Solid and broken arrows indicate that the averaged absolute MAR coefficient of the ROI pair exceeds 0.1 and 0.05, respectively. The ROI-averaged source interactions of the ground truth are illustrated by black solid arrows.

exerted little effect on the AUC and nRMSE of the current sources. However, increasing $\eta_0$ and $g_0$ in the dense setting enlarged the AUC of the MAR matrix. Nevertheless, when $\eta_0$ was increased from 100 to 1000, the sensitivity of the effective connectivity decreased about from 0.3 to 0.1. Therefore, a moderate value of $\eta_0$ (say 100) is recommended. In determining $g_0$, we observed that the perfect source reconstruction in the sparse setting (no false positives or negatives) was achieved only for $g_0 = 100$ or 1000 (and $\eta_0 = 100$). Thus, our recommended

**Table 2**
AUC and nRMSE of the current sources and AUC of the MAR matrix with (and without) applying the fMRI prior in *Simulation 2*.

| | wMNE (MNE) | | LCMV | | hVB | | Dynamic hVB | |
|---|---|---|---|---|---|---|---|---|
| | Sparse | Dense | Sparse | Dense | Sparse | Dense | Sparse | Dense |
| Current sources | | | | | | | | |
| AUC | 1.00 (0.97) | **1.00** (0.95) | 0.98 (0.55) | 0.99 (0.49) | **1** (0.77) | 0.99 (0.51) | **1** (0.99) | **1.00** (0.94) |
| nRMSE | 0.72 (0.94) | **0.70** (0.92) | – | – | 0.18 (0.66) | 0.97 (1.19) | **0.15** (0.64) | 0.86 (1.13) |
| ROI-nRMSE | | 0.49 (0.85) | | – | | 0.45 (0.80) | | **0.39** (0.77) |
| MAR matrix | | | | | | | | |
| AUC | 1.00 (0.87) | **1.00** (0.88) | 1.00 (0.72) | **1.00** (0.54) | **1** (0.75) | 0.55 (0.50) | **1** (0.87) | 0.91 (0.76) |
| ROI-AUC | | 0.57 (0.36) | | 0.46 (0.58) | | 0.80 (0.63) | | **0.81** (0.77) |

The scores with no fMRI prior applied are listed in parentheses. The best scores in each simulation scenario across the methods, applied with and without the prior, are indicated in bold font. The nRMSE scores are not defined in LCMV because of the weight normalization. ROI-nRMSE displays the nRMSE of the ROI-averaged source time courses. ROI-AUC is the AUC score of the ROI-averaged source interactions, computed by a normal AUC calculation procedure. Scores below 1 and exceeding (or equal to) 0.995 are indicated by 1.00.

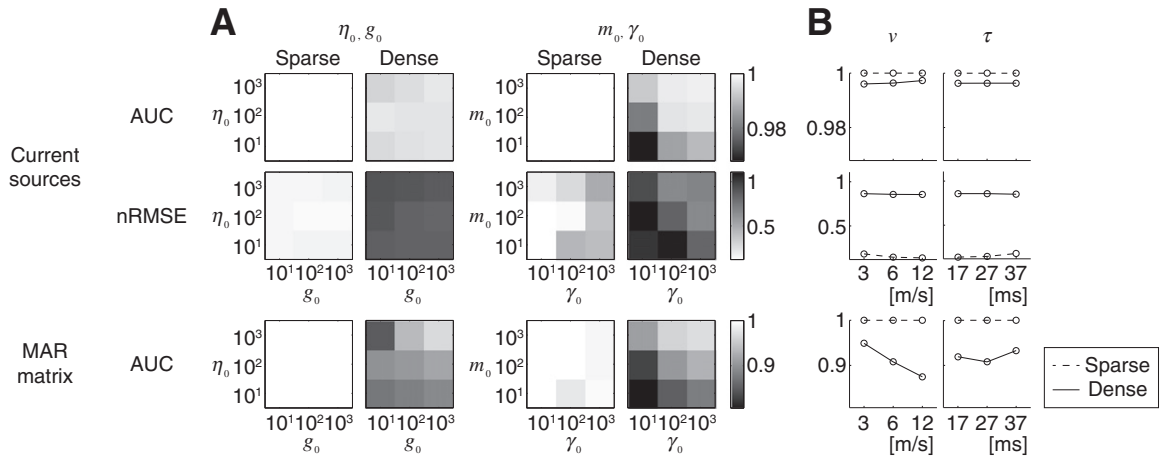


**Fig. 5.** Dependence of the hyperparameters and the delay parameters on the estimation accuracy in *Simulation 2*. (A) Variation in AUC and nRMSE scores as the hyperparameters $\eta_0$, $g_0$, $m_0$, and $\gamma_0$ are varied from 10 to 1000. Scores with the default setting (all hyperparameters set to 100) are shown in the center of each panel. The gray scale is common to each evaluation measure. (B) The AUC and nRMSE scores obtained by varying the conduction velocity ($v = 3, 6, 12$ m/s) and local delay constant ($\tau = 17, 27, 37$ ms). The ranges of the vertical axes correspond to the gray scales in (A).

hyperparameter setting in Eq. (10) is $(\eta_0, g_0) = (100, 100)$ or (100, 1000). The right hand side of Fig. 5A shows that high values of the hyperparameters $m_0$ and $\gamma_0$ yielded higher estimation accuracy in the dense setting and the reverse was true in the sparse setting. Since the central value of 100 achieved relatively high scores in both settings, our recommended hyperparameter setting in Eqs. (11) and (12) is $(m_0, \gamma_0) = (100, 100)$.

Fig. 5B shows the effect of changing the conduction velocity $v$ and the local delay constant $\tau$ on the evaluation measures. Although the AUC of the MAR matrix was noticeably altered in the dense setting, the AUC and nRMSE scores were relatively insensitive to perturbations of $v$ and $\tau$. This indicates that misspecification of the delay parameters does not severely degrade the estimation accuracy.

## Application to real data

Finally, the estimated source distributions and interactions were evaluated on experimental data. To this end, the methods were tested on a publicly available multimodal dataset (Henson et al., 2011). The task data were acquired while pictures of human faces were presented to subjects. The physiological plausibility of the estimated current sources and effective connectivity were qualitatively evaluated from previous literature reports.

### *Settings*

In evaluating the estimation performance, we focused on the temporal propagation of the stimulus-evoked responses. For this purpose, the MEG measurements were trial averaged and the *t*-values for the fMRI prior were computed with a task versus baseline contrast. Here, we used magnetometer measurements recorded in the dataset; the reconstructed sources from these measurements are presented as MEG results in Henson et al. (2011). All hyperparamters for the informative priors were set to 100. As demonstrated in *Simulation 2*, these hyperparamter settings optimized the reconstruction accuracy in both the sparse and dense settings.

The continuous MEG data in the dataset had been pre-processed by temporal signal-space separation (Taulu and Simola, 2006) to remove external noise. These data were further pre-processed by baseline correction, high-pass filtering (cutoff 0.25 Hz), low-pass filtering (cutoff 40 Hz), down sampling (from 1.1 kHz to 1 kHz), trial segmentation (0–300 ms after the stimulus onset), trial rejection, and trial averaging (concatenating the entire cohort of face stimulus conditions; see Fig. 6A). When segmenting the data, the last 100 ms prior to stimulus onset were also extracted to estimate the noise covariance matrix from the trial-averaged data. The trial rejection step eliminated trials with amplitude of MEG exceeding 3000 fT and that of electrooculogram (EOG) exceeding 100 μV. 548 trials (averaged across subjects) remained after the trial rejection step.

The fMRI data (3 mm isotropic) were pre-processed by running a batch Matlab script accompanying the dataset. This script executed a standard data pre-processing pipeline of SPM8.[6] The fMRI prior was computed from a *t*-value map ($p < 0.001$, uncorrected) in which the face stimulus conditions are contrasted against the baseline (the group-level *t*-value map is illustrated in Fig. 6B). Note that the *t*-value maps used to construct the fMRI prior were not group-averaged, but were separately computed from each subject.

### *Results*

Fig. 7A displays the reconstructed cortical sources and their estimated effective connectivity. The spatially clustered active sources estimated from the dynamic hVB method were more consistent among subjects and more interpretable than the activities estimated from the non-dynamic methods. The active sources were mainly estimated in the occipitotemporal cortex containing the face-selective regions IOG and FG (Haxby et al., 2000). In contrast, no sources were estimated in another face-selective region STS; source activity in this region is not easily reconstructed from MEG data because of its insensitivity to radial source components (Hillebrand and Barnes, 2002). The effective connectivity was estimated across the occipitotemporal cortex linking the lower and higher visual cortices. The reconstructed connections anatomically and functionally overlapped with the ventral visual pathways (Goodale and Milner, 1992) responsible for processing visual objects and faces (Haxby et al., 1994).

Fig. 7B shows a representative example of the event-related dynamics of the stimulus-evoked responses. Activity propagation was similarly estimated by wMNE, the hVB method, and the dynamic hVB method. In this example, the effective connectivity between IOG-FG was dominant for the direction from IOG to FG, as inferred from the MAR coefficients. The MAR coefficients of this extracted connection, estimated by the three methods, reflected temporally lagged correlations between the IOG and FG source time courses.

[6] http://www.fil.ion.ucl.ac.uk/spm/.

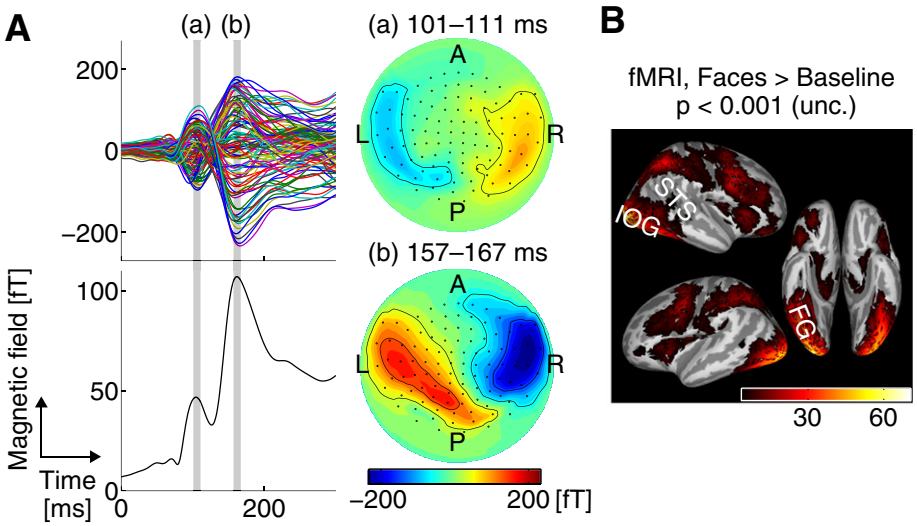


**Fig. 6.** Real MEG and fMRI data acquired during a face perception task. (A) Evoked responses to visual face stimuli (averaged across trials and subjects). The top-left plot shows the time courses of sensor measurements, and the bottom-left plot shows the square root of the measured power (averaged over sensors). The right figures display 2D topography maps at the two times of peak activity, indicated by the gray lines (a) and (b) in the left plots. (B) Group-averaged fMRI *t*-value maps with a contrast of face stimulus conditions against the baseline.

## Free energy

The free energy is popularly used as a model comparison criterion in a number of neuroimaging methods (Friston et al., 2003, 2008; David et al., 2006; Daunizeau and Friston, 2007; Wipf and Nagarajan, 2009; Stephan et al., 2009; Henson et al., 2010, 2011; Olier et al., 2013). In this section, the potential utility of the free energy in model comparisons is briefly investigated in two cases; a comparison between the hVB and dynamic hVB methods, and comparisons within the dynamic hVB method.

### *Comparison between the hVB and dynamic hVB methods*

We compared the free energy of the dynamic hVB method with that of the hVB method in *Simulation 2* and in the real data application. As shown in Table 3, the free energy was lower in the dynamic hVB method than in the hVB method for all evaluated cases. Nevertheless, the dynamic hVB method yielded higher estimation accuracy than the hVB method in *Simulation 2* (see Table 2). A close investigation of the six constituents of the free energy (see Eq. (A.18) in Appendix A) reveals that the slightly improved log likelihood $L$ was masked by the heavily penalized complexity terms $H_J$ of the current sources and $H_A$ of the MAR matrix. The $H_J$ term was much smaller in the dynamic hVB method than in the hVB method, because the factorization of $Q(\mathbf{J}_{1:T})$ in Eq. (14) is an additional approximation in the dynamic hVB method but is exact in the (non-dynamic) hVB method.

### *Comparison within the dynamic hVB method*

For the same number of non-zero MAR coefficients, we compared the dMRI-based connectivity constraint (default) with a random connectivity constraint. The comparison was conducted on the synthesized data of *Simulation 2*. The dMRI-based connectivity consistently yielded higher free energy and a higher estimation accuracy than the random connectivity (see Table 4). Moreover, the dMRI-based constraints yielded higher free energy values in the real data application (dMRI: $F = (8.13 \pm 0.22) \times 10^5$; Random: $F = (8.00 \pm 0.24) \times 10^5$).

We also compared the dynamic models by varying the local delay constant $\tau$ ($\tau = 7$, 17, 27 [default], 37, 47, and 57 ms; see Fig. 8A). The highest free energy was achieved at $\tau = 27$ ms (i.e., the ground truth) in *Simulation 1*. However, in *Simulation 2*, the free energy was monotonically increased, and failed to identify the best model yielding the smallest prediction error of forward simulation (the estimated dynamic model predicted the latter half of the source time courses from the given former half; see Fig. 8B and its caption for more details). These results indicate that comparisons by the free energy criterion are sensitive to the discrepancy between the data-generating model and our estimation model.

We also used the free energy approach to determine a suitable threshold for binarizing the anatomical connectivity strengths. However, this attempt failed because the free energy monotonically increased with increasing threshold (see the supplementary material).

## Discussion

In the present study, we propose a new method of MEG source reconstruction that simultaneously estimates the source amplitudes and interactions across the whole brain. Directed interactions between sources are represented by the full MAR model in the source space. The unknown network structure of this MAR model is informed by prior knowledge of the existence and lengths of anatomical connections, inferred from dMRI data. To improve the reliability and robustness of the estimates, we incorporate the following priors into the model; a spatial prior derived from fMRI activity patterns, and a sparse prior on the MAR coefficients. Source reconstruction using the MAR model with the above prior assumptions is formulated by a state-space representation within a Bayesian framework. The source activity and model parameters are jointly estimated by a variational Bayesian algorithm. The estimation performance of the proposed method was quantitatively and qualitatively evaluated on simulation and experimental data, respectively.

As demonstrated in *Simulation 1*, the source amplitudes and the MAR coefficients were reasonably estimated by the proposed method, even when the number of active sources was as high as 20 or 30. For over 75% of these originally active sources, the reconstruction gains exceeded 0.80 and the RMSEs of the MAR coefficients were below 0.15. However, the estimates yielded a non-negligible number of missing sources and connections with higher number of active sources. Close inspection of the estimates revealed that these misspecifications occurred more frequently when the source locations were less observable from the MEG sensors (data not shown). In general, a less observable source (i.e., a source with a smaller lead field norm) is difficult to correctly estimate, especially when other active sources with larger lead field norms produce a similar spatial pattern of the sensor data. Since

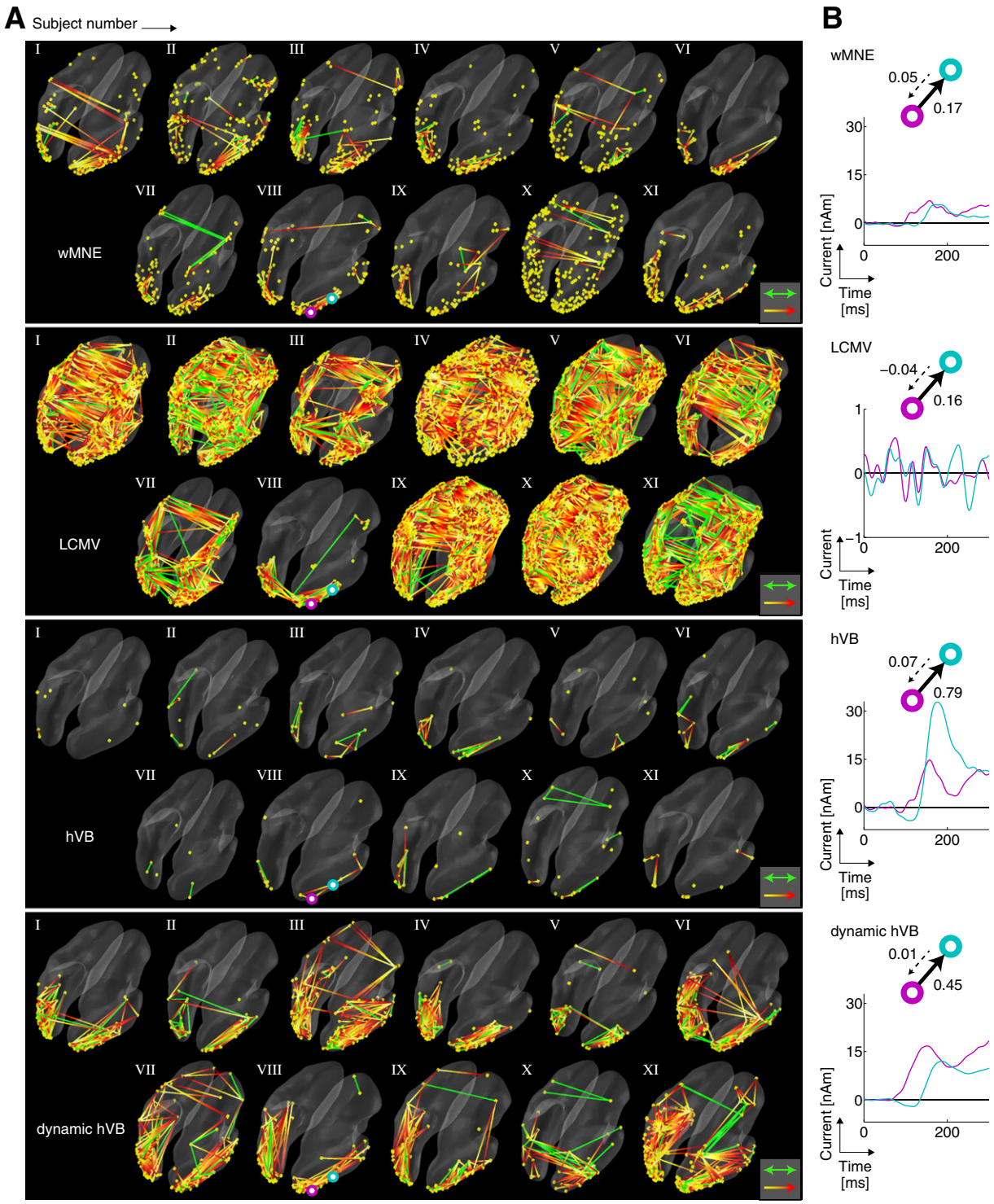


**Fig. 7.** Results of wMNE, LCMV, the hVB method, and the dynamic hVB method on real experimental data. (A) Reconstructed current sources (shown by yellow dots) and their effective connectivity during a face perception task undertaken by eleven subjects. Color legend of the effective connectivity is described in the caption of Fig. 3. (B) Representative example of the event-related dynamics of stimulus-evoked responses (IOG-FG in Subject VIII). The source time courses of LCMV are normalized by the maximum current. Cortical positions of the plotted current sources are indicated by small circles on the cortical surfaces in (A). The numbers beside the arrows connecting the two circles denote the corresponding estimated MAR coefficients. The time courses of each source (magenta and cyan) are plotted in the respective colors of each circle.

such source placements increase with increasing number of active sources in the Monte Carlo simulations, this phenomenon would explain the decay in the source reconstruction accuracy, reported also in Owen et al. (2012) using standard methods, with higher number of active sources. This phenomenon might also explain the declining performance of the connectivity estimation. Overall, however, our results suggest that the source amplitudes and the model parameters were adequately estimated by our proposed method.

In *Simulation 2*, we examined the estimation accuracy of the methods using the synthesized evoked responses. This simulation demonstrated that: 1) our method can correctly estimate the active sources and their effective connectivity in spatially sparse source

**Table 3**
Free energy $F$ and its six constituents defined in Eq. (A.18) ($\times 10^5$).

| | $L$ | $H_J$ | $H_\beta$ | $H_A$ | $H_\eta$ | $H_q$ | $F$ |
|---|---|---|---|---|---|---|---|
| *Simulation 2* (Sparse) | | | | | | | |
| hVB | 10.19 | −0.33 | −0.00 | – | – | −0.04 | 9.83 |
| dynamic hVB | 10.23 | −1.39 | −0.00 | −0.62 | −0.00 | −0.02 | 8.19 |
| *Simulation 2* (Dense) | | | | | | | |
| hVB | 10.15 | −0.35 | −0.00 | – | – | −0.04 | 9.77 |
| dynamic hVB | 10.23 | −1.38 | −0.00 | −0.64 | −0.00 | −0.02 | 8.19 |
| Real data (averaged across eleven subjects) | | | | | | | (mean ± s.d.) |
| hVB | 9.72 | −0.29 | −0.00 | – | – | −0.03 | 9.40 ± 0.04 |
| dynamic hVB | 10.05 | −1.22 | −0.00 | −0.67 | −0.00 | −0.03 | 8.13 ± 0.22 |

The free energy in the hVB method was obtained by specifying $\mathbf{A} = \mathbf{0}$ in $H_J$ and setting $H_A$ and $H_\eta$ to zero. The free energy values were not compared in *Simulation 1*, because a fair comparison is impossible when a (improper) non-informative prior is imposed on one method but not the other. In *Simulation 1*, the non-informative priors of $\boldsymbol{\eta}_{1:N}$ were used only in the dynamic hVB method.

configurations, and 2) although all of the active sources are not reconstructed in spatially dense source settings, our method identifies at least one single source interaction for every pair of ROIs with true effective connectivity. Moreover, while the nRMSE score of the sources was rather high in the dense setting (0.86), a reasonable score (0.39) was obtained by computing the nRMSE from the region-wise source time courses. These results indicate that, even when the sources are densely distributed over the cortex, ROI-wise activity and effective connectivity can be appropriately estimated. We emphasize that these estimation performances were obtained under realistic simulation settings, which introduce significant discrepancies between our model formulation and the data generation process. In particular, the assumed dynamic model was based on linear autoregressive modeling, whereas the data were generated from a network of non-linear neural mass models. The limitation of our source dynamics modeling manifests as false positive effective connectivity in the dense setting. This indicates a deficiency in the assumed model, since it also appeared when the MAR model was fitted to the true source time courses. It should be mentioned that the above discussion of *Simulation 2* was based on estimates informed by the fMRI prior. Our analysis with and without the fMRI prior confirmed that the prior certainly improves the estimation accuracy in both sparse and dense source configurations.

Applying our method to real data collected during a face perception task, we obtained physiologically plausible estimates that were appropriately consistent among subjects. The estimated source positions were mainly located along the occipitotemporal cortex, which contains the ventral part of the face-selective regions IOG and FG (Haxby et al., 2000). The estimated effective connectivity largely overlapped with the ventral visual pathways mediating transmission between the lower and higher visual areas during object recognition (Goodale and Milner, 1992). Since the face-selective regions have been extensively studied in the literature (Haxby et al., 1994, 2000; Kanwisher et al., 1997), they are useful for properly evaluating the source distributions estimated from the methods. In contrast, how these face-selective regions dynamically interact is not known in detail, although several studies have investigated anatomical connectivity (Gschwind et al., 2012; Pyles et al., 2013), functional connectivity (Davies-Thompson and Andrews, 2012), and effective connectivity (Fairhall and Ishai, 2007) among these areas. Therefore, although our estimated effective connectivity roughly corresponds to the ventral visual pathways, rigorous evaluation of the existence and directions of the source interactions is beyond the scope of this paper. For improved future evaluation, the consistency of the estimates in the present study could be examined using other neuroimaging modalities with higher spatial resolution than MEG, such as the electrocorticogram (ECoG).

**Table 4**
Model comparison between dMRI-based and random connectivity constraints in *Simulation 2*.

| | Sparse | | Dense | |
|---|---|---|---|---|
| | dMRI | Random | dMRI | Random |
| Free energy ($\times 10^5$) | 8.19 | 8.05 | 8.19 | 8.03 |
| Current sources | | | | |
| AUC | 1.00 | 1.00 | 1.00 | 0.99 |
| nRMSE | 0.15 | 0.15 | 0.86 | 0.87 |
| MAR matrix | | | | |
| AUC | 1.00 | 0.31 | 0.91 | 0.41 |

By simultaneously estimating the source amplitudes and interactions, the proposed dynamic method (the dynamic hVB method) achieved superior overall results to those of the non-dynamic methods (MNE, LCMV, and the hVB method), which sequentially estimate the amplitudes and interactions. One crucial difference exists between the two classes of methods; in the dynamic method, the spatiotemporal patterns of the source activity are constrained by the MAR model informed by knowledge of the anatomical brain networks, while the non-dynamic methods impose no temporal constraint on the sources. The richness of the prior on the sources in the dynamic method realizes a more precise source reconstruction, eventually leading to more accurate source interactions. Prior knowledge on the existence of anatomical connections further improves the performance of the dynamic method. Specifically, this prior may reciprocally compensate for the estimated magnitudes of the source amplitudes and interactions, since the estimation steps of the current sources and their associated MAR coefficients are mutually dependent under the variational Bayesian algorithm. In contrast, the non-dynamic methods do not update the reconstructed current sources after estimating the MAR coefficients. Therefore, the estimated connectivity is more sensitive to source reconstruction errors, because if two reconstructed sources are once incorrectly assigned to pairs of anatomically non-connected locations, the non-dynamic methods no longer estimate their effective connectivity in successive iterations. The modeling and algorithmic differences between the dynamic and non-dynamic methods are highlighted by the higher physiological plausibility and inter-subject consistency in the dynamic method, when these methods were applied to the face perception dataset.

We found that the estimation accuracies of the hVB and dynamic hVB methods were not necessarily reflected by the free energy criterion. The slightly improved log likelihood in the dynamic hVB method was masked by the heavily penalized complexity terms of the current sources and MAR matrix in Eq. (A.18). Nevertheless, the dynamic hVB method provided better estimates than the hVB method (see e.g. Table 2), possibly because the free energy becomes a much looser lower bound of the log model evidence in the dynamic hVB method than in the hVB method. In particular, the temporal independence assumption of $Q(\mathbf{J}_{1:T})$ is exact in the hVB method but is approximated in the dynamic hVB method. Conversely, we confirmed that the free energy was useful for optimal selection of models within the dynamic hVB method, when comparing the dMRI-based connectivity constraint with a random connectivity constraint and when determining the local delay constant $\tau$ in *Simulation 1*. However, the best local delay constant $\tau$ was not determined by the free energy criterion in *Simulation 2*, in which the local dynamics of the data-generating model (i.e., the neural mass model) were non-linear and more complex than those of our estimation model (the MAR model with a single delay component). This problem might be resolved by refining the dynamic source model. In addition to the best $\tau$ in *Simulation 2*, appropriate threshold of anatomical connectivity cannot be selected by the free energy criterion. Showing conditions that the free energy cannot be used for model comparison requires a more extensive study in the future.

Using the proposed method, we can estimate the effective connectivity without requiring the selection of a limited number of ROIs as network nodes. This was achieved by extending the source-dimensional autoregressive models adopted in previous dynamic

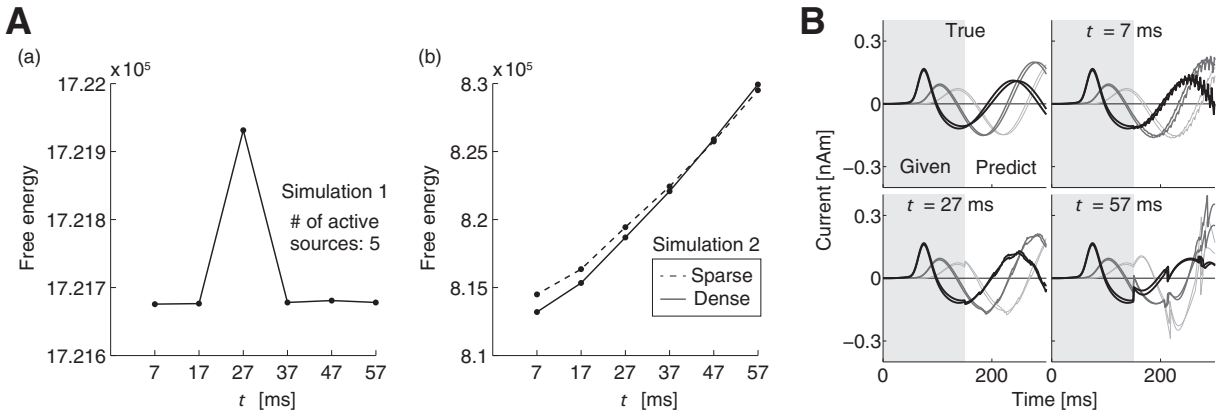


**Fig. 8.** (A) Free energy as a function of local delay constant $\tau$ in a typical example of *Simulation 1* (a) and in *Simulation 2* (b). (B) Current sources predicted by the estimated dynamic model at different local delay constant $\tau$. Colors of the time courses are explained in Fig. 2. The true sources (top-left) were generated from the neural mass models in the sparse setting of *Simulation 2* with an input signal different from the original (the first differential of the original input). From the given former half of the time courses, the latter half were predicted by forward simulation of the estimated dynamic model, varying the local delay constants ($\tau = 7, 27, 57$ ms).

methods of MEG/EEG source reconstruction (Galka et al., 2004; Yamashita et al., 2004; Lamus et al., 2012; Fukushima et al., 2012), in contrast to the approach of Olier et al. (2013) where they assumed a full MAR model in a low-dimensional latent space. The non-diagonal MAR coefficients newly introduced in the proposed method are derived from the existence and lengths of anatomical connections. These coefficients allow long-range source interactions to be estimated within whole-brain anatomical networks. Our method adopts an exploratory approach to estimating the effective connectivity, which complements the confirmatory approach of DCM (Friston et al., 2003; David et al., 2006), in which the effective connectivity between brain regions is inferred using a procedure of Bayesian model comparison.

When incorporating the anatomical connectivity information of dMRI into the MAR model, we require a threshold value of fiber counts to binarize the anatomical connectivity. We must also determine the conduction velocity and the local delay constant that specify the time lags between two sources. We selected a small threshold value of fiber counts ($10^{-4}$) to minimize false negative effective connectivity. Our model assumes that effective connectivity exists only between anatomically connected pairs of sources. Therefore, a higher threshold would increase false negative effective connectivity by increasing the number of false negative anatomical connections. By choosing a small threshold, we reduced the possibility of false negative anatomical connections, while increasing the number of MAR coefficients to be estimated. Although more MAR coefficients should increase the false positive effective connectivity, this effect is partially negated by the sparse prior applied to the MAR coefficients. The conduction velocity and local delay constant were set to 6 m/s and 27 ms, respectively. In *Simulation 2*, we investigated the impact of these parameters on the estimation accuracy, and found that the estimation performance was relatively insensitive to changes in both parameters. This indicates that appropriate estimates could be obtained if the time lags in the MAR model were only rough approximations.

The hyperparameters $m_0$ and $\gamma_0$, controlling the magnitude and the reliability of the fMRI prior, respectively, must be carefully determined. Such care is required because non-negligible differences might exist between the spatial patterns of the electrocortical activity and those observed in fMRI. Increasing the dependence on the fMRI prior may not improve the results. Indeed, imposing a large weighting on the spatially dense fMRI prior (by setting $\gamma_0$ high) reduced the estimation accuracy in the sparse source setting of *Simulation 2*, where a moderate value of $\gamma_0$ provided higher estimation accuracy. Therefore, in applying the methods to experimental data, these hyperparameters were selected to yield reasonable evaluation scores in both the sparse and the dense simulation settings.

Various forms of spatial prior information or constraints can be unified within the framework of covariance component estimation (Friston et al., 2008; Wipf and Nagarajan, 2009). Viewed from this perspective, the proposed method uses source covariance components with focal spatial support and estimates their hyperparameters with help of the fMRI prior (Sato et al., 2004). The proposed method is informed by these spatial priors incorporated into the covariance component estimation, and also by the source activities predicted by the dynamic source model. By combining multiple forms of covariance components as in Friston et al. (2008), the proposed method could be further informed by other types of spatial priors based on, for example, minimum norm, smoothness, and depth constraints.

Representing the source dynamics by linear autoregressive modeling would introduce errors to the estimated (actually non-linear) source interactions. When the interactions were governed by realistic non-linear network dynamics, the limitation of our dynamic model emerged as false positive estimates of the effective connectivity. This observation confirms the necessity of a non-linear extension of the linear MAR source model to embrace a wide range of dynamic phenomena. In future work, we will develop a means of incorporating non-linear neurobiological dynamics into the source model, while preserving the tractability of estimation in the high-dimensional source space.

## Acknowledgments

The authors are grateful to Dr. Mitsuo Kawato (ATR CNS) for his insightful suggestions and Mr. Ryosuke Hayashi (ATR NIA) for his help with data pre-processing. We also would like to thank Dr. Daniel Wakeman (MGH) and Dr. Richard Henson (MRC CBU) for providing the real experimental data. This research was supported by a contract with the National Institute of Information and Communications Technology entitled, 'Development of network dynamics modeling methods for human brain data simulation systems' (173) and by the Japan Society for the Promotion of Science through Grant-in-Aid for JSPS Fellows (23-3907).

## Appendix A. Update rules of the estimation algorithm

In this appendix, the update rules for the current sources and the model parameters are derived in detail. The update equations are presented in order of the updating steps in the estimation algorithm (**A**-step, $\boldsymbol{\eta}$-step, **q**-step, **J**-step, and $\beta$-step). For notational simplicity, $\mathbf{J}_t$

and $\overline{\mathbf{J}}_t$ for $t \le 0$ are included in the update equations, although they are set to zero in the model, and the subscripts in expectation terms are omitted.

The **A**-step updates the approximate posterior distribution of the MAR matrix $Q(\mathbf{A})$. Substituting **A** for $x$ in Eq. (16) and calculating its expectation term, we obtain a sum of linear and quadratic terms of $\mathbf{a}_n$, where $\mathbf{a}_n^T$ denotes the non-zero entries of the $n$-th row of **A**. Therefore, $Q(\mathbf{A})$ is a product of Gaussian distributions:

$$Q(\mathbf{A}) = \prod_{n=1}^{N} \mathcal{N}(\mathbf{a}_n | \overline{\mathbf{a}}_n, \overline{\mathbf{U}}_n). \tag{A.1}$$

The covariance $\overline{\mathbf{U}}_n$ and the mean $\overline{\mathbf{a}}_n$ are derived as

$$\begin{cases} \overline{\mathbf{U}}_n = \{\mathrm{diag}(\overline{\boldsymbol{\eta}}_n) \\ \qquad + \overline{\beta}\, \overline{q}_n \sum_{t=1}^{T} \left( \sum_{l=1}^{L} \mathbf{E}_{n,\Delta_l} .* \left\langle \mathbf{J}_{t-\Delta_l} \mathbf{J}_{t-\Delta_l}^{\mathrm{T}} \right\rangle \right)_{C_n, C_n} \Big\}^{-1} \\ \overline{\mathbf{a}}_n = \overline{\mathbf{U}}_n\, \overline{\beta}\, \overline{q}_n \sum_{t=1}^{T} \left( \left( \sum_{l=1}^{L} \mathbf{e}_{n,\Delta_l} .* \overline{\mathbf{J}}_{t-\Delta_l} \right)_{C_n} \overline{\mathbf{J}}_{n,t} \right), \end{cases} \tag{A.2}$$

where $\overline{\boldsymbol{\eta}}_n$, $\overline{\beta}$, $\overline{q}_n$, $\overline{\mathbf{J}}_{t-\Delta_l}$, and $\overline{\mathbf{J}}_{n,t}$ are the means of the approximate posterior distributions, which are updated in the previous iteration, and $C_n$ is a set of indices corresponding to the cortical locations anatomically connected to the $n$-th source position. A dot product .* denotes an element-wise multiplication of a pair of vectors or matrices. The $N$-dimensional column vector $\mathbf{e}_{n,\Delta_l}$ is an indicator of binary values. If the locations corresponding to the index entries of $\mathbf{e}_{n,\Delta_l}$ are anatomically connected to the $n$-th source position with time lag $\Delta_l$, the relevant indices in $\mathbf{e}_{n,\Delta_l}$ are set to one. Another indicator, the $N$-dimensional matrix $\mathbf{E}_{n,\Delta_l}$ is computed as $\mathbf{e}_{n,\Delta_l}\, \mathbf{e}_{n,\Delta_l}^{\mathrm{T}}$. Here, the terms subscripted $C_n$, $C_n$ and $C_n$ in Eq. (A.2) can be considered as the covariance and mean, respectively, of a vector of current sources that directly affects $J_{n,t}$.

The $\boldsymbol{\eta}$-step updates the approximate posterior distribution of the inverse variances of the MAR coefficients $Q(\boldsymbol{\eta}_{1:N})$. Substituting $\boldsymbol{\eta}_{1:N}$ for $x$ in Eq. (16) and calculating its expectation term, we obtain a sum of linear and logarithmic terms of $\eta_{n,k}$. Therefore, $Q(\boldsymbol{\eta}_{1:N})$ is a product of gamma distributions:

$$Q(\boldsymbol{\eta}_{1:N}) = \prod_{n=1}^{N} \prod_{k=1}^{K_n} \Gamma\left(\eta_{n,k} | \overline{\eta}_{n,k}, g_{n,k}\right) \tag{A.3}$$

whose shape parameter $g_{n,k}$ and mean $\overline{\boldsymbol{\eta}}_{n,k}$ are given by

$$\begin{cases} g_{n,k} = g_0 + \dfrac{1}{2} \\ \overline{\eta}_{n,k} = g_{n,k} \left\{ g_0\, \eta_0^{-1} + \dfrac{1}{2} \left( \left\langle \mathbf{a}_n\, \mathbf{a}_n^{\mathrm{T}} \right\rangle \right)_{k,k} \right\}^{-1}. \end{cases} \tag{A.4}$$

The **q**-step updates the approximate posterior distribution of the inverse variance of the scaled current noise $Q(\mathbf{q})$. Substituting **q** for $x$ in Eq. (16) and calculating its expectation term, we obtain a sum of linear and logarithmic terms of $q_n$. Therefore, $Q(\mathbf{q})$ is a product of gamma distributions:

$$Q(\mathbf{q}) = \prod_{n=1}^{N} \Gamma(q_n | \overline{q}_n, \gamma_n) \tag{A.5}$$

whose shape parameter $\gamma_n$ and mean $\overline{q}_n$ are given by

$$\begin{cases} \gamma_n = \gamma_0 + \dfrac{1}{2}\, T \\ \overline{q}_n = \gamma_n \Big\{ \gamma_0\, \overline{\nu}_n + \dfrac{1}{2}\, \overline{\beta} \sum_{t=1}^{T} \Big( \left\langle \mathbf{J}_t\, \mathbf{J}_t^{\mathrm{T}} \right\rangle \\ \qquad - \Big( \sum_{l=1}^{L} \left( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \right) \Big) \overline{\mathbf{J}}_t^{\mathrm{T}} - \overline{\mathbf{J}}_t \Big( \sum_{l=1}^{L} \left( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \right) \Big)^{\mathrm{T}} \\ \qquad + \left\langle \Big( \sum_{l=1}^{L} \left( \mathbf{A}_l\, \mathbf{J}_{t-\Delta_l} \right) \Big) \Big( \sum_{l=1}^{L} \left( \mathbf{A}_l\, \mathbf{J}_{t-\Delta_l} \right) \Big)^{\mathrm{T}} \right\rangle \Big)_{n,n} \Big\}^{-1}. \end{cases} \tag{A.6}$$

$\overline{\mathbf{A}}_l$ and $\overline{\mathbf{J}}_{t-\Delta_l}$ are the means of the approximate posterior distributions, where $\overline{\mathbf{A}}_l$ is obtained from $\overline{\mathbf{a}}_n$ for $n \in \{1{:}N\}$ (see Eq. (5) for the definition of $\mathbf{A}_l$).

The **J**-step updates the approximate posterior distributions of the current sources $Q(\mathbf{J}_t)$ for $t \in \{1{:}T\}$. Substituting $\mathbf{J}_t$ for $x$ in Eq. (16) and calculating its expectation term, we obtain linear and quadratic terms of $\mathbf{J}_t$. Therefore, $Q(\mathbf{J}_t)$ is a Gaussian distribution:

$$Q(\mathbf{J}_t) = \mathcal{N}\left(\mathbf{J}_t | \overline{\mathbf{J}}_t, \overline{\beta}^{-1} \overline{\mathbf{V}}_t\right) \tag{A.7}$$

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

$$\begin{cases} \overline{\mathbf{V}}_t = \left( \mathbf{G}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{G} + \hat{\mathbf{V}}_t^{-1} \right)^{-1} \\ \overline{\mathbf{J}}_t = \overline{\mathbf{V}}_t \left( \mathbf{G}^{\mathrm{T}} \mathbf{S}^{-1} \mathbf{B}_t + \hat{\mathbf{V}}_t^{-1} \hat{\mathbf{J}}_t \right) \end{cases} \tag{A.8}$$

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

$$\hat{\mathbf{V}}_t = \begin{cases} \left( \mathrm{diag}(\overline{\mathbf{q}}) + \sum_{l'=1}^{L'} \left\langle \mathbf{A}_{l'}^{\mathrm{T}} \mathrm{diag}(\overline{\mathbf{q}})\, \mathbf{A}_{l'} \right\rangle \right)^{-1}, & \text{for } 1 \le t \le T - \Delta_1 \\ \left( \mathrm{diag}(\overline{\mathbf{q}}) \right)^{-1}, & \text{for } T - \Delta_1 < t \le T \end{cases}$$

$$\hat{\mathbf{J}}_t = \begin{cases} \hat{\mathbf{V}}_t \Big( \mathrm{diag}(\overline{\mathbf{q}}) \sum_{l=1}^{L} \left( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \right) + \sum_{l'=1}^{L'} \left( \overline{\mathbf{A}}_{l'}^{\mathrm{T}}\, \mathrm{diag}(\overline{\mathbf{q}})\, \overline{\mathbf{J}}_{t+\Delta_{l'}} \right) \\ \qquad - \sum_{l'=1}^{L'} \left\langle \mathbf{A}_{l'}^{\mathrm{T}}\, \mathrm{diag}(\overline{\mathbf{q}}) \Big( \sum_{l=1(l \ne l')}^{L} \mathbf{A}_l\, \overline{\mathbf{J}}_{t-\Delta_l+\Delta_{l'}} \Big) \right\rangle \Big), \\ \qquad\qquad \text{for } 1 \le t \le T - \Delta_1 \\ \sum_{l=1}^{L} \left( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \right), \qquad \text{for } T - \Delta_1 < t \le T. \end{cases} \tag{A.9}$$

The variables denoted with an upper bar are obtained from the means of the approximate posterior distributions. $L$ is the total number of time lags. $L'$ equals a number of integers within the range of $t + \Delta_{L'} \le T < t + \Delta_{L'+1}$ for $T - \Delta_L < t \le T$, and $L' = L$ for $1 < t \le T - \Delta_L$. $\overline{\mathbf{V}}_t$ and $\overline{\mathbf{J}}_t$ in Eq. (A.8) is calculated from the following computationally efficient form:

$$\begin{cases} \overline{\mathbf{V}}_t = (\mathbf{I} - \mathbf{K}\mathbf{G}) \hat{\mathbf{V}}_t \\ \overline{\mathbf{J}}_t = \hat{\mathbf{J}}_t + \mathbf{K} \left( \mathbf{B}_t - \mathbf{G} \hat{\mathbf{J}}_t \right) \end{cases} \tag{A.10}$$

where $\mathbf{K} = \hat{\mathbf{V}}_t \mathbf{G}^{\mathrm{T}} \left( \mathbf{G} \hat{\mathbf{V}}_t \mathbf{G}^{\mathrm{T}} + \mathbf{S} \right)^{-1}$ is a gain matrix for the sensor measurements and **I** is the identity matrix.

The $\beta$-step updates the approximate posterior distribution of the scaling parameter $Q(\beta)$. Substituting $\beta$ for $x$ in Eq. (16) and calculating its expectation term, we obtain linear and logarithmic terms of $\beta$. Therefore, $Q(\beta)$ is a gamma distribution:

$$Q(\beta) = \Gamma\left(\beta | \overline{\beta}, \gamma_\beta\right) \tag{A.11}$$

whose shape parameter $\gamma_\beta$ and mean $\overline{\beta}$ are given by

$$\begin{cases} \gamma_\beta = \frac{1}{2}MT \\ \overline{\beta} = \gamma_\beta \Big\{ \frac{1}{2}\sum_{t=1}^{T} \mathrm{tr}\Big( \mathbf{S}^{-1} \Big( \mathbf{B}_t \mathbf{B}_t^{\mathrm{T}} - (\mathbf{G}\overline{\mathbf{J}}_t)\mathbf{B}_t^{\mathrm{T}} - \mathbf{B}_t (\mathbf{G}\overline{\mathbf{J}}_t)^{\mathrm{T}} \\ \qquad + (\mathbf{G}\overline{\mathbf{J}}_t)(\mathbf{G}\overline{\mathbf{J}}_t)^{\mathrm{T}} \Big) \Big) + \frac{1}{2}\sum_{t=1}^{T} \mathrm{tr}\Big( \mathrm{diag}(\overline{\mathbf{q}}) \Big( \overline{\mathbf{J}}_t\, \overline{\mathbf{J}}_t^{\mathrm{T}} \\ \qquad - \Big( \sum_{l=1}^{L} \Big( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \Big) \Big) \overline{\mathbf{J}}_t^{\mathrm{T}} - \overline{\mathbf{J}}_t \Big( \sum_{l=1}^{L} \Big( \overline{\mathbf{A}}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \Big) \Big)^{\mathrm{T}} \\ \qquad + \Big\langle \Big( \sum_{l=1}^{L} \Big( \mathbf{A}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \Big) \Big) \Big( \sum_{l=1}^{L} \Big( \mathbf{A}_l\, \overline{\mathbf{J}}_{t-\Delta_l} \Big) \Big)^{\mathrm{T}} \Big\rangle \Big) \Big) \Big\}^{-1}. \end{cases} \tag{A.12}$$

The expectation terms in the update equations are evaluated as follows. These terms, which are used in the **A**-step, $\boldsymbol{\eta}$-step, and **q**-step, are derived as

$$\Big\langle \mathbf{J}_{t-\Delta_l}\, \mathbf{J}_{t-\Delta_l}^{\mathrm{T}} \Big\rangle = \overline{\mathbf{J}}_{t-\Delta_l}\, \overline{\mathbf{J}}_{t-\Delta_l}^{\mathrm{T}} + \overline{\mathbf{V}}_{t-\Delta_l}, \tag{A.13}$$

$$\Big\langle \mathbf{a}_n\, \mathbf{a}_n^{\mathrm{T}} \Big\rangle = \overline{\mathbf{a}}_n\, \overline{\mathbf{a}}_n^{\mathrm{T}} + \overline{\mathbf{U}}_n, \tag{A.14}$$

$$\Big\langle \mathbf{J}_t\, \mathbf{J}_t^{\mathrm{T}} \Big\rangle = \overline{\mathbf{J}}_t\, \overline{\mathbf{J}}_t^{\mathrm{T}} + \overline{\mathbf{V}}_t. \tag{A.15}$$

After some rearrangement, the final term of the **q**-step in Eq. (A.6) is evaluated with the above expectation terms as:

$$\begin{aligned} &\Big( \Big\langle \Big( \sum_{l=1}^{L} \Big( \mathbf{A}_l\, \mathbf{J}_{t-\Delta_l} \Big) \Big) \Big( \sum_{l=1}^{L} \Big( \mathbf{A}_l\, \mathbf{J}_{t-\Delta_l} \Big) \Big)^{\mathrm{T}} \Big\rangle \Big)_{n,n} \\ &\quad = \mathrm{tr}\Big( \Big\langle \mathbf{a}_n\, \mathbf{a}_n^{\mathrm{T}} \Big\rangle \Big( \sum_{l=1}^{L} \mathbf{E}_{n,\Delta_l} .* \Big\langle \mathbf{J}_{t-\Delta_l}\, \mathbf{J}_{t-\Delta_l}^{\mathrm{T}} \Big\rangle \Big)_{C_n, C_n} \Big). \end{aligned} \tag{A.16}$$

The trace of the final term of the $\beta$-step in Eq. (A.12) is derived from Eq. (A.16) by replacing the expectation of the quadratic term of the current sources with its mean. The remaining expectation terms reside in Eq. (A.9) for the **J**-step. These are obtained by computing the following expectation term for all pairs of $l'$ and $l$:

$$\Big\langle \mathbf{A}_{l'}^{\mathrm{T}}\, \mathrm{diag}(\overline{\mathbf{q}})\, \mathbf{A}_l \Big\rangle = \sum_{n=1}^{N} \overline{q}_n \Big\langle \widetilde{\mathbf{a}}_{l'n} \widetilde{\mathbf{a}}_{ln}^{\mathrm{T}} \Big\rangle, \tag{A.17}$$

where $\widetilde{\mathbf{a}}_{ln}^{\mathrm{T}}$ denotes the $n$-th row of $\mathbf{A}_l$. The expectation in the right-hand side of Eq. (A.17) is obtained by first expanding Eq. (A.14) into an $N$-dimensional matrix (based on the original source index of $\mathbf{a}_n$). Then, the matrix entries not corresponding to row and column source indices that anatomically connect the $n$-th source index with time lags $\Delta_{l'}$ and $\Delta_l$, respectively, are padded with zeros.

The free energy $F$, maximized during the iterative updates, can be compactly written as

$$\begin{aligned} F &= \Big\langle \log \frac{P(\mathbf{B}_{1:T}, \mathbf{J}_{1:T}, \beta, \mathbf{A}, \boldsymbol{\eta}_{1:N}, \mathbf{q})}{Q(\mathbf{J}_{1:T}, \beta, \mathbf{A}, \boldsymbol{\eta}_{1:N}, \mathbf{q})} \Big\rangle \\ &= \Big\langle \sum_{t=1}^{T} \log P(\mathbf{B}_t | \mathbf{J}_t, \beta) \Big\rangle + \Big\langle \sum_{t=1}^{T} \log \frac{P\Big( \mathbf{J}_t | \mathbf{J}_{t-\Delta_1}, \ldots, \mathbf{J}_{t-\Delta_L}, \beta, \mathbf{A}, \mathbf{q} \Big)}{Q(\mathbf{J}_t)} \Big\rangle \\ &\quad + \Big\langle \log \frac{P(\beta)}{Q(\beta)} \Big\rangle + \Big\langle \log \frac{P(\mathbf{A}|\boldsymbol{\eta}_{1:N})}{Q(\mathbf{A})} \Big\rangle + \Big\langle \log \frac{P(\boldsymbol{\eta}_{1:N})}{Q(\boldsymbol{\eta}_{1:N})} \Big\rangle + \Big\langle \log \frac{P(\mathbf{q})}{Q(\mathbf{q})} \Big\rangle \\ &= L + H_J + H_\beta + H_A + H_\eta + H_q. \end{aligned} \tag{A.18}$$

When the data are well-fitted by the model parameters, a high expected log likelihood function $L$ is obtained. The other terms are the negative values of the KL divergences between the approximate posteriors and the priors, which penalize the total model complexity (Attias, 1999).

## Appendix B. Bias correction of AUC

To interpret AUC as an unbiased detection accuracy score, one should provide the same number of active and inactive sources (or connections) in the ROC analysis (Grova et al., 2006). In our simulations, the number of actives was actually far less than the number of inactives, as usually assumed in distributed source methods. Thus, to compute the AUC, we randomly and sequentially extracted the inactives until the number of actives and inactives matched. Note that randomly extracting the inactives from the whole brain significantly underestimates the false positive rate. Therefore, as conducted by Grova et al. (2006), we computed two types of AUC; one by randomly extracting the inactives from the close neighborhood of actives ($\mathrm{AUC}_{\mathrm{close}}$), the other by extracting the far local maxima of false positives ($\mathrm{AUC}_{\mathrm{far}}$). Note that, if the spatial neighborhood contain a much larger proportion of inactives relative to actives (as found in *Simulation 2*), false positives estimated within the neighborhood are less penalized than those estimated beyond the neighborhood. The final score of AUC was defined as the mean of $\mathrm{AUC}_{\mathrm{close}}$ and $\mathrm{AUC}_{\mathrm{far}}$ both computed from the average scores of 50 different drawings from a set of inactives. $\mathrm{AUC}_{\mathrm{close}}$ and $\mathrm{AUC}_{\mathrm{far}}$ for sources were computed similarly to Grova et al. (2006). By an equivalent procedure, $\mathrm{AUC}_{\mathrm{close}}$ and $\mathrm{AUC}_{\mathrm{far}}$ for connections were obtained (with some modifications, as described below).

$\mathrm{AUC}_{\mathrm{close}}$ was used to assess the reconstruction accuracy of the spatial extent of sources or connections. The spatial neighborhood in computing $\mathrm{AUC}_{\mathrm{close}}$ was specified by a 5 cm-radius sphere centered on the actives. The area of this sphere is comparable to the spatial neighborhood adopted in Grova et al. (2006). When calculating this score for connections, the spatial neighborhood of a connection was defined as connections of which seed and target sources are included in the spatial neighborhoods of the seed and target sources of the centered connection, respectively.

$\mathrm{AUC}_{\mathrm{far}}$ was computed to detect false positives far from the actives. To compute this score for sources, a coarse cortical parcel was initially selected at random. Within this parcel subtracted by the area occupied by the spatial neighborhood of the actives, the false positive source with maximum magnitude was then extracted. The coarse cortical parcel was obtained by a nearest-neighbor manner parcellation of the whole cortex into 86 parcels, the number of parcels specified in Grova et al. (2006). When calculating this score for connections, we extracted the false positive connection with maximum absolute MAR coefficient from a randomly selected pair of parcels, whose connections were not included in the spatial neighborhood of the actives. When selecting parcels or pairs of parcels, we excluded in advance those candidates lacking a source or anatomical connection beyond the spatial neighborhood.

## Appendix C. Source reconstruction steps of LCMV

In the LCMV beamformer, the following spatial filter weights $F_n$ of the $n$-th source were used in source reconstruction:

$$F_n = \Big( G_n^{\mathrm{T}}\, \mathbf{C}^{-1} G_n \Big)^{-1} G_n^{\mathrm{T}}\, \mathbf{C}^{-1}. \tag{C.1}$$

$G_n$ is the $n$-th column of the lead field matrix **G**, with the source orientations fixed to the directions used when generating the simulation data (i.e., perpendicular to the cortical surface). We computed the sensor noise covariance matrix **C** from the activity duration with a regularization term:

$$\mathbf{C} = \hat{\mathbf{C}} + \lambda\, \mathbf{I}, \tag{C.2}$$

where $\hat{\mathbf{C}}$ is a regularization-free covariance matrix and λ is a regularization constant computed as one tenth of $\mathrm{tr}\left(\hat{\mathbf{C}}\right)$ divided by the number of sensors. Location biases of the source images were removed by normalizing the weights of the spatial filter $F_n$ (Sekihara et al., 2005):

$$\overline{F}_n = \frac{F_n}{\sqrt{F_n^T F_n}}. \tag{C.3}$$

The reconstructed sources over the whole brain $\overline{\mathbf{J}}_t$ were obtained as

$$\overline{\mathbf{J}}_t = \overline{\mathbf{F}}\,\mathbf{B}_t, \tag{C.4}$$

where $\mathbf{B}_t$ denotes the sensor measurements and $\overline{\mathbf{F}}$ is a matrix consisting of the spatial filters $\overline{F}_n$ ($n = 1, \ldots, N$). The reconstructed sources in Eq. (C.4) were computed using NUTMEG (Dalal et al., 2011).

## Appendix D. Details of the neural mass model simulations

This appendix details the functional network of neural mass models used for generating the stimulus-evoked responses in *Simulation 2*. The equations of this network model were integrated by a standard Runge–Kutta method with an integration time step of 0.1 ms (the generated waveforms were afterwards downsampled to 1 kHz). The parameter settings used in each component of the network model, namely, the neural mass model, the extrinsic coupling, and the input signal, are provided below.

The neural mass model was based on a model of two subpopulations (David and Friston, 2003). The convolution kernels for the excitatory and the inhibitory synapses, represented as $(H_e t / \tau_e)\exp(-t / \tau_e)$ and $(H_i t / \tau_i)\exp(-t / \tau_i)$ for $t \geq 0$, respectively, were assumed in both subpopulations. The parameters $\tau_e$ and $\tau_i$ for the two subpopulations were set to the values adopted by David and Friston (2003). The parameter specifying the relative proportion of the subpopulation with larger $\tau_e$ and $\tau_i$ was set to 0.7. Under this parameter setting, the activity peaks occurred at 100 ms and 170 ms following the input, mimicking the stimulus-evoked responses during face perception. For each subpopulation, the parameters $H_e$ and $H_i$ were derived by dividing the values of $H_e\tau_e$ and $H_i\tau_i$ in David et al. (2005) by the above-specified values of $\tau_e$ and $\tau_i$, respectively. The parameters governing the intrinsic couplings and the sigmoid functions of the neural mass model were set according to David et al. (2005), while the relative strength among the intrinsic couplings was specified by the setting popularly used in the literature (Jansen and Rit, 1995; David and Friston, 2003; David et al., 2006; Spiegler et al., 2010).

The extrinsic coupling parameters in the network model were set to reproduce typical waveforms of stimulus-evoked responses. Following David et al. (2005), who extensively investigated the relationships between extrinsic coupling parameters and response waveforms, we set these parameters to 50 and 10 for forward and backward connections, respectively. The time lags in the extrinsic couplings were determined solely from the fiber transmission delay. The local delay constant $\tau$ in Eq. (4) was omitted because it had been embodied in the convolution kernels of the neural mass model.

The input signal to the network model was generated by a gamma density function and a set of discrete cosine functions, as described in David et al. (2006). The shape and scale parameters of the gamma density function were set to 64 and 1024, respectively, yielding activity peaks at 100 ms and 170 ms. The parameters in the discrete cosine functions were set to those adopted in David et al. (2006).

## Appendix E. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.09.066.

## References


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

Baillet, S., Garnero, L., 1997. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans. Biomed. Eng. 44, 374–385.

Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Process. Mag. 18, 14–30.

Barber, D., Chiappa, S., 2007. Unified inference for variational Bayesian linear Gaussian state-space models. Adv. Neural Inf. Process. Syst. 19 (NIPS 2006), pp. 81–88.

Beal, M.J., 2003. Variational algorithm for approximate Bayesian inference(Ph.D. thesis) Univ. College London, London, UK.

Bolstad, A., Van Veen, B.D., Nowak, R., 2009. Space-time event sparse penalization for magneto-/electroencephalography. Neuroimage 46, 1066–1081.

Brookes, M.J., Hale, J.R., Zumer, J.M., Stevenson, C.M., Francis, S.T., Barnes, G.R., Owen, J.P., Morris, P.G., Nagarajan, S.S., 2011. Measuring functional connectivity using MEG: methodology and comparison with fcMRI. Neuroimage 56, 1082–1104.

Cheung, B.L.P., Riedner, B.A., Tononi, G., Van Veen, B.D., 2010. Estimation of cortical connectivity from EEG using state-space models. IEEE Trans. Biomed. Eng. 57, 2122–2134.

Dalal, S.S., Zumer, J.M., Guggisberg, A.G., Trumpis, M., Wong, D.D.E., Sekihara, K., Nagarajan, S.S., 2011. MEG/EEG source reconstruction, statistical evaluation, and visualization with NUTMEG. Comput. Intell. Neurosci. 2011.

Dale, A., Liu, A., Fischl, B., Buckner, R., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Dannhauer, M., Lämmel, E., Wolters, C.H., Knösche, T.R., 2013. Spatio-temporal regularization in linear distributed source reconstruction from EEG/MEG: a critical evaluation. Brain Topogr. 26, 229–246.

Daunizeau, J., Friston, K.J., 2007. A mesostate-space model for EEG and MEG. Neuroimage 38, 67–81.

Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., Lina, J.M., 2006. Bayesian spatio-temporal approach for EEG source reconstruction: conciliating ECD and distributed models. IEEE Trans. Biomed. Eng. 53, 503–516.

Daunizeau, J., Grova, C., Marrelec, G., Mattout, J., Jbabdi, S., Pélégrini-Issac, M., Lina, J.M., Benali, H., 2007. Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. Neuroimage 36, 69–87.

David, O., Friston, K., 2003. A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage 20, 1743–1755.

David, O., Harrison, L., Friston, K., 2005. Modelling event-related responses in the brain. Neuroimage 25, 756–770.

David, O., Kiebel, S., Harrison, L., Mattout, J., Kilner, J., Friston, K., 2006. Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage 30, 1255–1272.

Davies-Thompson, J., Andrews, T.J., 2012. Intra- and interhemispheric connectivity between face-selective regions in the human brain. J. Neurophysiol. 108, 3087–3095.

de Pasquale, F., Della Penna, S., Snyder, A.Z., Marzetti, L., Pizzella, V., Romani, G.L., Corbetta, M., 2012. A cortical core for dynamic integration of functional networks in the resting human brain. Neuron 74, 753–764.

Deco, G., Jirsa, V., McIntosh, A.R., Sporns, O., Kötter, R., 2009. Key role of coupling, delay, and noise in resting brain fluctuations. Proc. Natl. Acad. Sci. U. S. A. 106, 10302–10307.

Fairhall, S.L., Ishai, A., 2007. Effective connectivity within the distributed cortical network for face perception. Cereb. Cortex 17, 2400–2406.

Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2, 56–78.

Friston, K.J., Harrison, L.M., Penny, W.D., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K.J., Harrison, L.M., Daunizeau, J., Kiebel, S.J., Phillips, C., Trujillo-Barreto, N.J., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. Neuroimage 39, 1104–1120.

Fukushima, M., Yamashita, O., Kanemura, A., Ishii, S., Kawato, M., Sato, M., 2012. A state-space modeling approach for localization of focal current sources from MEG. IEEE Trans. Biomed. Eng. 59, 1561–1571.

Galka, A., Yamashita, O., Ozaki, T., Biscay, R., Valdés-Sosa, P., 2004. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. Neuroimage 23, 435–453.

Ghosh, A., Rho, Y., McIntosh, A., Kötter, R., Jirsa, V., 2008. Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput. Biol. 4, e1000196.

Goodale, M.A., Milner, A., 1992. Separate visual pathways for perception and action. Trends Neurosci. 15, 20–25.

Grova, C., Daunizeau, J., Lina, J.M., Bénar, C.G., Benali, H., Gotman, J., 2006. Evaluation of EEG localization methods using realistic simulations of interictal spikes. Neuroimage 29, 734–753.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., Vuilleumier, P., 2012. White-matter connectivity between face-responsive regions in the human brain. Cereb. Cortex 22, 1564–1576.

Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42.

Hämäläinen, M., Hari, R., Ilmoniemi, 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.

Haxby, J., Horwitz, B., Ungerleider, L., Maisog, J., Pietrini, P., Grady, C., 1994. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J. Neurosci. 14, 6336–6353.

Haxby, J., Hoffman, E., Gobbini, M., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.

Henson, R.N., Flandin, G., Friston, K.J., Mattout, J., 2010. A parametric empirical Bayesian framework for fMRI-constrained MEG/EEG source reconstruction. Hum. Brain Mapp. 31, 1512–1531.
Henson, R.N., Wakeman, D.G., Litvak, V., Friston, K.J., 2011. A parametric empirical Bayesian framework for the EEG/MEG inverse problem: generative models for multi-subject and multi-modal integration. Front. Hum. Neurosci. 5, 1–16.
Hillebrand, A., Barnes, G.R., 2002. A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. Neuroimage 16, 638–650.
Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M., Engel, A.K., 2012. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat. Neurosci. 15, 884–892.
Honey, C.J., Kötter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc. Natl. Acad. Sci. U. S. A. 104, 10240–10245.
Hui, H.B., Pantazis, D., Bressler, S.L., Leahy, R.M., 2010. Identifying true cortical interactions in MEG using the nulling beamformer. Neuroimage 49, 3161–3174.
Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., Chang, C., 2013. Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80, 360–378.
Jansen, B., Rit, V., 1995. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol. Cybern. 366, 357–366.
Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.
Lamus, C., Hämäläinen, M.S., Temereanca, S., Brown, E.N., Purdon, P.L., 2012. A spatiotemporal dynamic distributed solution to the MEG inverse problem. Neuroimage 63, 894–909.
Matsuura, K., Okabe, Y., 1995. Selective minimum-norm solution of the biomagnetic inverse problem. IEEE Trans. Biomed. Eng. 42, 608–615.
Mosher, J.C., Lewis, P.S., Leahy, R.M., 1992. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans. Biomed. Eng. 39, 541–557.
Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: forward solutions for inverse methods. IEEE Trans. Biomed. Eng. 46, 245–259.
Neal, R.M., 1996. Bayesian learning for neural networks. Springer, New York.
Nummenmaa, A., Auranen, T., Hämäläinen, M.S., Jääskeläinen, I.P., Sams, M., Vehtari, A., Lampinen, J., 2007. Automatic relevance determination based hierarchical Bayesian MEG inversion in practice. Neuroimage 37, 876–889.
Nunez, P.L., Srinivasan, R., 2006. Electric fields of the brain: the neurophysics of EEG. Oxford University Press, New York.
Olier, I., Trujillo-Barreto, N.J., El-Deredy, W., 2013. A switching multi-scale dynamical network model of EEG/MEG. Neuroimage 81, 262–287.
Ou, W., Hämäläinen, M.S., Golland, P., 2009. A distributed spatio-temporal EEG/MEG inverse solver. Neuroimage 44, 932–946.
Ou, W., Nummenmaa, A., Ahveninen, J., Belliveau, J.W., Hämäläinen, M.S., Golland, P., 2010. Multimodal functional imaging using fMRI-informed regional EEG/MEG source estimation. Neuroimage 52, 97–108.
Owen, J.P., Wipf, D.P., Attias, H.T., Sekihara, K., Nagarajan, S.S., 2009. Accurate reconstruction of brain activity and functional connectivity from noisy MEG data. Conf. Proc. 31st IEEE Eng. Med. Biol. Soc, pp. 65–68.
Owen, J.P., Wipf, D.P., Attias, H.T., Sekihara, K., Nagarajan, S.S., 2012. Performance evaluation of the Champagne source reconstruction algorithm on simulated and real M/EEG data. Neuroimage 60, 305–323.
Palva, S., Palva, J.M., 2012. Discovering oscillatory interaction networks with M/EEG: challenges and breakthroughs. Trends Cogn. Sci. 16, 219–230.
Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24, 5–12.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
Portin, K., Vanni, S., Virsu, V., Hari, R., 1999. Stronger occipital cortical activation to lower than upper visual field stimuli. Neuromagnetic recordings. Exp. Brain Res. 124, 287–294.
Pyles, J.A., Verstynen, T.D., Schneider, W., Tarr, M.J., 2013. Explicating the face perception network with white matter connectivity. PLoS One 8, e61611.
Sato, M., 2001. Online model selection based on the variational Bayes. Neural Comput. 13, 1649–1681.
Sato, M., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical Bayesian estimation for MEG inverse problem. Neuroimage 23, 806–826.
Scherg, M., Von Cramon, D., 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalogr. Clin. Neurophysiol. 62, 32–44.
Schmitt, U., Louis, A.K., Darvas, F., Buchner, H., Fuchs, M., 2001. Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data. IEEE Trans. Med. Imaging 20, 314–324.
Schoffelen, J.M., Gross, J., 2009. Source connectivity analysis with MEG and EEG. Hum. Brain Mapp. 30, 1857–1865.
Sekihara, K., Sahani, M., Nagarajan, S.S., 2005. Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. Neuroimage 25, 1056–1067.
Spiegler, A., Kiebel, S.J., Atay, F.M., Knösche, T.R., 2010. Bifurcation analysis of neural mass models: impact of extrinsic inputs and dendritic time constants. Neuroimage 52, 1041–1058.
Stephan, K.E., Tittgemeyer, M., Knösche, T.R., Moran, R.J., Friston, K.J., 2009. Tractography-based priors for dynamic causal models. Neuroimage 47, 1628–1638.
Taulu, S., Simola, J., 2006. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys. Med. Biol. 51, 1759–1768.
Tononi, G., Sporns, O., Edelman, G.M., 1994. A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc. Natl. Acad. Sci. U. S. A. 91, 5033–5037.
Tournier, J., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage 35, 1459–1472.
Trujillo-Barreto, N.J., Aubert-Vázquez, E., Penny, W.D., 2008. Bayesian M/EEG source reconstruction with spatio-temporal priors. Neuroimage 39, 318–335.
Uutela, K., Hämäläinen, M., Somersalo, E., 1999. Visualization of magnetoencephalographic data using minimum current estimates. Neuroimage 10, 173–180.
Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880.
Wipf, D., Nagarajan, S.S., 2009. A unified Bayesian framework for MEG/EEG source imaging. Neuroimage 44, 947–966.
Wipf, D.P., Owen, J.P., Attias, H.T., Sekihara, K., Nagarajan, S.S., 2010. Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. Neuroimage 49, 641–655.
Woolrich, M.W., Stephan, K.E., 2013. Biophysical network models and the human connectome. Neuroimage 80, 330–338.
Yamashita, O., Galka, A., Ozaki, T., Biscay, R., Valdés-Sosa, P., 2004. Recursive penalized least squares solution for dynamical inverse problems of EEG generation. Hum. Brain Mapp. 21, 221–235.
Yoshioka, T., Toyama, K., Kawato, M., Yamashita, O., Nishina, S., Yamagishi, N., Sato, M., 2008. Evaluation of hierarchical Bayesian method through retinotopic brain activities reconstruction from fMRI and MEG signals. Neuroimage 42, 1397–1413.