Block-structured Bayesian source estimation model for magnetoencephalography signals

Kai Miyazaki[1], Okito Yamashita[2,3], and Yoichi Miyawaki[1,4]

[1] Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo, Japan
[2] Advanced Telecommunications Research Institutes International, Kyoto, Japan
[3] RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo, Japan
[4] Center for Neuroscience and Biomedical Engineering (CNBE), The University of Electro-Communications, Tokyo, Japan

k.miyazaki@uec.ac.jp
yoichi.miyawaki@uec.ac.jp

- **Abstract.** The human brain consists of functionally specialized areas that process specific types of information and interact with one another. Magnetoencephalography (MEG) is a neuroimaging technique that captures brain activity at high temporal resolution. However, its spatial resolution is insufficient to accurately localize neural activation, and thus cannot capture the interactive nature of human brain activity. To resolve this issue, various MEG source estimation models have been proposed. In particular, models incorporating prior information from functional magnetic resonance imaging (fMRI), which offers superior spatial resolution, have improved source estimation accuracy. However, these models typically ignore the similarity of activity across different brain areas, and still fall short in tracking dynamic brain activity pattern at a sufficient spatial scale. In this study, we developed a block-structured model that integrates information about functional areas and the inter-areal relationships of the human brain into MEG source estimation based on a hierarchical Bayesian framework. We evaluated our model performance using simulation data with sequential activation across multiple brain areas. Results showed that our model outperformed conventional approaches in source estimation accuracy, suggesting that incorporating the functional areal information and inter-areal relationships may enhance MEG source estimation, enabling human neuroimaging at high spatiotemporal resolution.

Keywords: Magnetoencephalography, Inverse problem, Hierarchical Bayesian model, Block-structured covariance

1 Introduction

The human brain processes external information at fine spatial and temporal scales through dynamic and harmonious interaction among multiple brain areas. Neuroimaging techniques require a sufficient spatiotemporal resolution for tracking such neural processes. Functional magnetic resonance imaging (fMRI) and

magnetoencephalography (MEG) are widely used neuroimaging techniques for measuring brain activity. fMRI offers high spatial resolution, capable of even distinguishing cortical layer structures at the submillimeter level. However, its temporal resolution is limited due to its reliance on blood oxygenation changes. In contrast, MEG provides high temporal resolution by detecting magnetic fields generated by electrical neural activity, but its spatial resolution is constrained because the sensors are placed above the scalp and detect signals from multiple brain areas. Consequently, there are no neuroimaging techniques that can sufficiently track neural activity with both high spatial and temporal resolution.

To resolve this issue, previous studies have attempted to estimate the locations of neural sources in the brain that generate the measured MEG signals [1–3]. This approach is known as MEG source estimation, and various source estimation models have been developed and evaluated. The most widely used method moderates the ill-posedness of the inverse problem by introducing regularization techniques, as the number of MEG sensors is much smaller than the number of candidate source localizations. Minimum norm estimation (MNE) [1] employs squared regularization, resulting in stable but spatially diffuse solutions. In contrast, automatic relevance determination (ARD) [4, 5] imposes a sparse prior within Bayesian framework to estimate more focal sources. In addition, high spatial resolution information from fMRI can be incorporated to improve MEG source estimation. For example, fMRI-derived blood oxygenation contrasts can serve as prior information. Variational Bayesian multimodal encephalography (VBMEG) [3] incorporates such information as a hierarchical prior within a Bayesian framework. Although these models have achieved a certain level of accuracy in source localization, their estimation processes are limited in that they do not consider the functional organization of the human brain [6–8]. Specifically, they treat each cortical location independently, without incorporating the spatial similarity of neural activity within the same functional areas or interaction among multiple brain areas [9–12].

To address this limitation, we hypothesized that incorporating such structural and functional properties into the prior distribution for MEG source estimation would lead to improve its accuracy. Specifically, we incorporated information of functional area and inter-areal relationships into the model as a hierarchical prior within a variational Bayesian framework. The model is expected to produce source estimates that more faithfully reflect the underlying functional organization of the brain, thereby enabling accurate MEG source estimation.

2 Method

2.1 MEG source estimation

MEG signals were computed using a forward model based on the positional relationship between the signal source and the MEG sensors. The forward model is described as

$$\mathbf{b} = \mathbf{G}\mathbf{j} + \mathbf{e} \tag{1}$$

where $\mathbf{b}$ ($\mathbb{R}^{N\times T}$ matrix; N, number of MEG sensors; T,number of time points) denotes the measured MEG signal, $\mathbf{G}$ ($\mathbb{R}^{N\times M}$ matrix; M, number of vertices on the cortical surface) is the leadfield matrix that describes how each cortical source current contributes to the MEG signal at each sensor location, $\mathbf{j}$ ($\mathbb{R}^{M\times T}$ matrix) represents the cortical source current, and $\mathbf{e}$ ($\mathbb{R}^{N\times T}$ vector) is the sensor noise modeled as a zero-mean multivariate Gaussian random variable with covariance $\beta^{-1}\boldsymbol{\Sigma}_{\epsilon}$. $\mathbf{G}$ was computed using VBMEG software.

MEG source estimation is to estimate $\mathbf{j}$ given $\mathbf{b}$ and $\mathbf{G}$ in Eq. (1). However, because the number of vertices typically exceeds the number of MEG sensors, it is an ill-posed problem and a unique solution cannot be obtained. To resolve this issue, various models have been proposed. In this paper, we developed a Bayesian estimation model that incorporates group information based on functional areas in the brain and inter-areal relationships as a hierarchical prior distribution. To evaluate whether these properties improve source estimation accuracy, we compared four models: (1) a hierarchical Bayesian model assuming that cortical surface vertices are treated independently of each other ("hierarchical automatic relevance determination" model), (2) a hierarchical Bayesian model incorporating only functional areal grouping (i.e., grouping vertices according to cortical parcellation; "hierarchical grouped automatic relevance determination" model), (3) a hierarchical Bayesian model incorporating both functional areal grouping and inter-areal relationships (i.e., grouping vertices according to cortical parcellation and defining interactions between groups; "block-structured hierarchical Bayesian estimation" model), and (4) a model with L2-norm regularization (minimum norm estimation, MNE) as a widely used standard method.

Hierarchical automatic relevance determination. The Bayesian estimation model introduces both the likelihood and the prior distribution for $\mathbf{j}$ as Gaussian distributions:

$$P(\mathbf{b}\,|\mathbf{j},\beta) \sim N(\mathbf{Gj},\beta^{-1}\boldsymbol{\Sigma}_{\epsilon}) \tag{2}$$

$$P(\hat{\mathbf{j}}\,|\boldsymbol{\alpha}) \sim N(0,\mathbf{A}^{-1}) \tag{3}$$

where $\beta^{-1}\boldsymbol{\Sigma}_{\epsilon}$ represents the sensor noise variance, and $\mathbf{A}$ is the $\mathbb{R}^{M\times M}$ source covariance matrix. If we put a uniform prior on the diagonal elements $\boldsymbol{\alpha}$ of the source covariance $\mathbf{A}$, we will obtain automatic relevance determination (ARD) [4, 5]. To improve the source estimation accuracy, a hierarchical prior can be placed on the elements of $\boldsymbol{\alpha}$. We call this model hierarchical automatic relevance determination (hARD). hARD places a conjugate prior distribution on the elements $\boldsymbol{\alpha}$ in the form of the following Gamma distribution:

$$P(\boldsymbol{\alpha}|\,\boldsymbol{\alpha_0},\boldsymbol{\gamma_0}) = \prod_{i=1}^{M} P(\alpha_i|\alpha_{i0},\gamma_{i0}) \sim \mathrm{Gamma}(\alpha_i;\alpha_{i0},\gamma_{i0}) \tag{4}$$

where α_{i0} represents the relevance parameter of each vertex, and the γ_{i0} represents the degree of freedom. Based on these probability distributions, the posterior probability distribution is then defined as:

$$P(\hat{\mathbf{j}}|\mathbf{b},\mathbf{G},\mathbf{A},\beta,\boldsymbol{\alpha_0},\boldsymbol{\gamma_0}) = N(\hat{\mathbf{j}}\,|\boldsymbol{m},\boldsymbol{\Sigma}_{\mathrm{t}}). \tag{5}$$

To estimate approximate posterior distribution, we used the variational Bayesian estimation and mean field approximation [13, 14]. The mean $\boldsymbol{m}$ and covariance $\boldsymbol{\Sigma_t}$ are given by

$$\boldsymbol{m} = \boldsymbol{\Sigma}_t \beta \mathbf{G}^{\mathrm{T}} \boldsymbol{\Sigma}_\epsilon \mathbf{b} \tag{6}$$

$$\boldsymbol{\Sigma}_t = \beta^{-1}(\mathbf{G}^{\mathrm{T}} \boldsymbol{\Sigma}_\epsilon \mathbf{G} + \beta^{-1}\mathbf{A})^{-1}, \tag{7}$$

respectively. The estimated current $\hat{\mathbf{j}}$ is obtained by calculating the expectation of this posterior distribution. hARD optimizes the hyperparameter as the following update rules:

$$\gamma_\beta = \frac{NT}{2} \tag{8}$$

$$(\beta^{\mathrm{new}})^{-1} = \gamma_\beta^{-1}\left[\frac{1}{2}\sum_{t=1}^{T}\{(\mathbf{b}_t - \mathbf{G}\hat{\mathbf{j}}_t)^{\mathrm{T}}\boldsymbol{\Sigma}_\epsilon(\mathbf{b}_t - \mathbf{G}\hat{\mathbf{j}}_t)\} + \frac{1}{2}\{\mathrm{Tr}(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_\epsilon^{-1}\mathbf{G})\beta^{-1}(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_\epsilon^{-1}\mathbf{G} + \beta^{-1}\mathbf{A})^{-1}\}\right] \tag{9}$$

$$\gamma_{i0}^{\mathrm{new}} = \gamma_{i0} + \frac{T}{2} \tag{10}$$

$$(\alpha_i^{\mathrm{new}})^{-1} = w_i \alpha_{i0}^{-1} + (1 - w_i)\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\mathbf{j}}_{ti}^2 + \sigma_{tii}^2\right) \tag{11}$$

$$w_i = \frac{\gamma_{i0}}{\gamma_{i0} + \frac{T}{2}} \tag{12}$$

where σ_{tii}^2 is diagonal element of $\boldsymbol{\Sigma}_t$ in Eq. (7), and w_i is the prior weight, which controls the balance between the first and second terms of Eq. (11) based on the confidence of hyperparameter of hierarchical prior (in the case of ARD, w_i is 0). By iteratively applying the update rules shown in Eqs. (6–12) until convergence, the model parameters are optimized and an approximate posterior distribution of the cortical source current is obtained. Figure 1 (a) shows the graphical illustration of hARD.

Hierarchical grouped automatic relevance determination. hARD treats each vertex independently and ignores the fact that vertices within the same functional area of the brain may exhibit similar activation due to shared processing roles. To incorporate this principle of functional organization into the model, we developed a Bayesian estimation model that introduces grouping of vertices based on cortical parcellation. We call this model hierarchical grouped automatic relevance determination (hgARD). The likelihood and prior distribution for $\mathbf{j}$ is the same as Eqs. (1-2), and the diagonal elements $\boldsymbol{\alpha}$ of the source covariance matrix are shared within the group as

$$\boldsymbol{\alpha} = (a_1, \dots, a_1, a_2, \dots, a_2, a_k, \dots a_k) \tag{13}$$

where k represents the number of groups. The number of repetitions of each a_k in $\boldsymbol{\alpha}$ corresponds to the number of vertices assigned to the k-th group. The conjugate prior distribution for this grouped element $\boldsymbol{\alpha}$ is given by the following Gamma distribution:

$$P(\boldsymbol{\alpha}|\, \boldsymbol{\alpha_0}, \boldsymbol{\gamma_0}) = \prod_{g=1}^{k} P(\alpha_g|\alpha_{g0}, \gamma_{g0}) \sim \mathrm{Gamma}(\alpha_g; \alpha_{g0}, \gamma_{g0}). \tag{14}$$

Based on these probability distributions, the posterior probability distribution and the mean $\boldsymbol{m}$ and covariance $\boldsymbol{\Sigma_t}$ are same as Eqs. (5-7). hgARD optimizes the hyperparameters using an update rule as in hARD. The update rules are similar to Eq. (10-12) but are modified according to Eq. (13). In update rules of Eq. (10-12), γ_{i0}, α_{i0} and T are replaced by γ_{g0} α_{g0} and $N_g T$, respectively, where N_g is the number of vertices belonging to the g-th group. Moreover, the term $\left(\hat{\mathbf{j}}_{ti}^2 + \sigma_{tii}^2\right)$ is replaced by the group-wise sum $\sum_i\left(\hat{\mathbf{j}}_{ti}^2 + \sigma_{tii}^2\right)$. By iteratively applying the update rules until convergence, the model parameters are optimized and an approximate posterior distribution of the cortical source current is obtained. Figure 1 (b) shows the graphical illustration of hgARD.

Block-Structured Hierarchical Bayesian estimation model. hARD and hgARD treat only the diagonal elements of the source covariance matrix that represent source current variance at each cortical location, and ignore the off-diagonal elements that represent relationships between functional areas. To consider a full source covariance matrix, the conjugate prior distribution should be Wishart distribution [5]. Although a full covariance matrix can potentially account for the covariance structure of the source currents, it is impractical to estimate due to the excessive number of parameters. To mitigate this difficulty, we developed a Bayesian estimation model that incorporates grouping not only for the diagonal elements but also for the off-diagonal elements of the source covariance matrix. We call this model Block-Structured Hierarchical Bayesian estimation (BHBE) model, which corresponds to the model (3). The likelihood and prior distribution for $\mathbf{j}$ is the same as Eqs. (1-2), and the source covariance matrix $\mathbf{A}$ is defined as $k \times k$ block matrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \dots & \mathbf{A}_{1,k} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \dots & \mathbf{A}_{2,k} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{A}_{k,1} & \mathbf{A}_{k,2} & \dots & \mathbf{A}_{k,k} \end{bmatrix} \tag{15}$$

where the diagonal blocks $\mathbf{A}_{l,l}$ and the off-diagonal blocks $\mathbf{A}_{l,l'}$ are defined as

$$\mathbf{A}_{l,l} = \begin{bmatrix} a_g^2 & a_{gg} & \dots & a_{gg} \\ a_{gg} & a_g^2 & \dots & a_{gg} \\ \vdots & \dots & \ddots & \vdots \\ a_{gg} & a_{gg} & \dots & a_g^2 \end{bmatrix} \tag{16}$$

$$\mathbf{A}_{l,l'} = \begin{bmatrix} a_{gg'} & a_{gg'} & \dots & a_{gg'} \\ a_{gg'} & a_{gg'} & \dots & a_{gg'} \\ \vdots & \dots & \ddots & \vdots \\ a_{gg'} & a_{gg'} & \dots & a_{gg'} \end{bmatrix}, \tag{17}$$

respectively ($g \neq g'$ and $g, g' = 1,2, \dots, k$). According to the method proposed by Archakov and Hansen [15], the block-structured covariance matrix $\mathbf{A}$ can be decomposed as

$$\mathbf{A} = \mathbf{QDQ}^{\mathrm{T}} \tag{18}$$

where $\mathbf{Q}$ is an orthonormal matrix and $\mathbf{D}$ is a $\mathbb{R}^{M\times M}$ block-diagonal matrix. $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \begin{bmatrix} \boldsymbol{v}_{n_1} & \mathbf{0} & \dots & \mathbf{0} & \boldsymbol{v}_{n_1\perp} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{v}_{n_2} & \dots & \mathbf{0} & \mathbf{0} & \boldsymbol{v}_{n_2\perp} & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots & \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{v}_{n_\mathrm{k}} & \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{v}_{n_k\perp} \end{bmatrix} \tag{19}$$

where $\boldsymbol{v}_{n_g}$ is $\mathbb{R}^{n_g\times 1}$ column vector with entries $\frac{1}{\sqrt{n_g}}$ and $\boldsymbol{v}_{n_g\perp}$ is $\mathbb{R}^{n_g\times n_g-1}$ matrix that is orthogonal to $\boldsymbol{v}_{n_g}$ such that $\boldsymbol{v}_{n_g\perp}^{\mathrm{T}}\boldsymbol{v}_{n_g} = 0$ and $\boldsymbol{v}_{n_g\perp}^{\mathrm{T}}\boldsymbol{v}_{n_g\perp} = \mathbf{I}_{n_g-1}$. $\mathbf{D}$ is defined as

$$\mathbf{D} = \begin{bmatrix} \mathbf{C} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \lambda_1\mathbf{I}_{n_1-1} & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \lambda_k\mathbf{I}_{n_k-1} \end{bmatrix} \tag{20}$$

where $\mathbf{C}$ is a $\mathbb{R}^{k\times k}$ positive definite matrix representing the relationship between the groups, and λ_g is a positive scalar. The elements of the original covariance matrix $\mathbf{A}$ can then be expressed as

$$c_{gg'} = \begin{cases} a_g^2 + (n_g - 1)a_{gg'}, & \text{for } g = g' \\ a_{gg'}\sqrt{n_g n_{g'}}, & \text{for } g \neq g' \end{cases}, \tag{21}$$

$$\lambda_g = a_g^2 - a_{gg}\ . \tag{22}$$

Here the diagonal blocks of $\mathbf{A}$ are determined by the diagonal elements of $\mathbf{C}$ and λ_g and the off-diagonal blocks of $\mathbf{A}$ are controlled by the off-diagonal elements of $\mathbf{C}$. From Eq. (22), λ_g represents the difference between the diagonal and off-diagonal elements within group g, reflecting how much the variances a_g^2 of the diagonal elements deviate from covariance a_{gg} of the off-diagonal elements within the group. This indicates that as λ_g increases, correlations among vertices within the group decreased, resulting in greater variability in the estimated currents among those vertices. In the current definition of $\mathbf{A}$ in Eq. (15), the elements are arranged in group-wise order (e.g., $\mathrm{diag}(\mathbf{A}) = (a_{11}, \dots, a_{11}, a_{22}, \dots, a_{22} \dots a_{kk}, \dots, a_{kk})$). However, in actual data, the vertex grouping doesn't necessarily follow such an ordered structure, and the elements of $\mathbf{A}$ may appear in irregular arrangement (e.g., $\mathrm{diag}(\mathbf{A}) = (a_{11}, a_{22}, a_{kk}, a_{11}, a_{kk}, \dots a_{11})$). To align the group ordering to the vertex ordering, we introduced the permutation matrix $\mathbf{P}$, and defined the prior distribution as

$$P(\hat{\mathbf{j}}\,|\mathbf{C}, \lambda) \sim N(0, \mathbf{A}^{-1}) \sim N(0, (\mathbf{PQDQ^T P^T})^{-1}). \tag{23}$$

The conjugate prior distributions for this hyperparameter $\mathbf{C}$ and λ are given by the following Wishart and Gamma distributions:

$$P(\mathbf{C}|\mathbf{C_0}, \nu_0) \sim \mathcal{W}(\mathbf{C};\ \mathbf{C_0}, \nu_{c0}) \tag{24}$$

$$P(\lambda|\lambda_0,\gamma_0) = \prod_{g=1}^{k} P(\lambda_g|\lambda_{g0},\gamma_{g0}) \sim \mathrm{Gamma}(\lambda_g;\lambda_{g0},\gamma_{g0}) \tag{25}$$

where λ_{g0} is a hyperparameter that represents the expected difference between the variance and covariance within the group, the γ_{g0} is the degree of freedom and the $\mathbf{C}_0 \in \mathbb{R}^{k\times k}$ defines the covariance between groups. Based on these probability distributions, the posterior mean $\boldsymbol{m}$ and covariance $\boldsymbol{\Sigma_t}$ are given by Eqs. (8,9), respectively. Whereas the previous study [16] used Markov-chain Monte Carlo algorithm to explore the posterior distribution, we employed the variational Bayesian estimation and the mean field approximation. The hyperparameters λ and $\mathbf{C}$ are then optimized by the following update rules

$$\nu_{c0}^{\mathrm{new}} = \nu_{c0} + T \tag{26}$$

$$\mathbf{C} = w_{cg}\mathbf{C_0}^{-1} + (1-w_{cg})\frac{1}{T}\sum_{t=1}^{T}(\mathbf{Q}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}(\hat{\mathbf{j}}_t\hat{\mathbf{j}}_t^{\mathrm{T}} + \boldsymbol{\Sigma}_t)\mathbf{QP})_{1:k,1:k} \tag{27}$$

$$w_{cg} = \frac{T}{\nu_{c0}+T} \tag{28}$$

$$\gamma_{g0}^{\mathrm{new}} = \gamma_{g0} + \frac{(n_g-1)T}{2} \tag{29}$$

$$\left(\lambda_g^{\mathrm{new}}\right)^{-1} = w_{\lambda g}\lambda_{g0}^{-1} + (1-w_{\lambda g})\frac{1}{(n_g-1)T}\sum_{t=1}^{T}\sum_i(\mathbf{P}^{\mathrm{T}}(\hat{\mathbf{j}}_t\hat{\mathbf{j}}_t^{\mathrm{T}} + \boldsymbol{\Sigma}_t)\mathbf{QP})_{k+1+\sum_{f=1}^{g-1}n_f:k+1+\sum_{f=1}^{g}n_f,k+1+\sum_{f=1}^{g-1}n_f:k+1+\sum_{f=1}^{g}n_f} \tag{30}$$

$$w_{\lambda g} = \frac{\gamma_{g0}}{\gamma_{g0}+\frac{(n_g-1)T}{2}} \tag{31}$$

where $(\cdot)_{1:n,1:n}$ denotes the submatrix formed by selecting rows and columns from 1 to n. In this formulation, the number of blocks determines the number of the groups (cortical parcellation), and the hyperparameter $\mathbf{C}_0$ and λ_{g0} determine the relationship between and within the groups, respectively. Figure 1 (c) shows the graphical illustration of the proposed model.

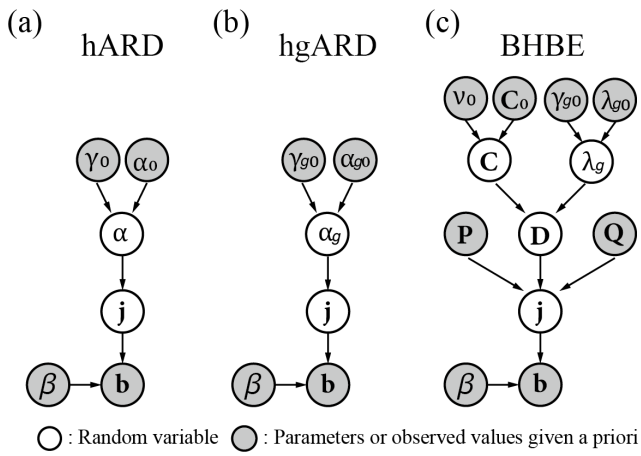

Fig. 1. Graphical representations of the Bayesian models.

2.2 Experiments

We performed MRI and MEG experiments to obtain the cortical surface model (from MRI data) and the head position in MEG scanner (from MEG data) to be used in the simulation analysis according to the following procedures.

Participant. One individual took part in both MRI and MEG experiments. The participant gave written informed consent before the experiment. The procedure was approved by the institutional review board of the University of Electro-Communications, Advanced Telecommunications Research Institute International (ATR) Brain Activity Imaging Center, and Ethical Review Board of Osaka University Hospital. The experiment was conducted under corrected vision with MRI-compatible glasses.

MRI experiment. Anatomical images were obtained using an MRI system (3-Tesla MAGNETOM Prismafit, Siemens) located at the ATR Brain Activity Imaging Center. The magnetization-prepared rapid-acquisition gradient-echo (MP-RAGE) sequence (208 sagittal slices; TR, 2,250 ms; TE, 3.06 ms; FA, 9 deg; FOV, 256 × 256 mm; voxel size, 1.0 × 1.0 × 1.0 mm) was used.

MEG experiment. The participant's head and the positions of the five marker coils (three on the forehead and two on each earlobe) were first recorded using a 3D scanner and a position sensor system (FastSCAN Cobra, Polhemus Inc., USA). These markers were used to calibrate the head position relative to the MEG sensors inside the MEG scanner (PQ1160C, Yokogawa Electric Co., Japan). Then, the participant was asked to lie on their back in the MEG scanner and remain as still as possible. The current was then applied to the five marker coils to generate a magnetic field, and the magnetic field was measured with 160-channel superconducting quantum interference device (SQUID) sensors in the MEG scanner. The spatial relationship between the MEG sensors and the marker coils was computed by solving an inverse problem based on the measured magnetic field. Using the position information of the marker coils, the participant head location relative to the MEG sensor locations was estimated. Finally, the high-resolution anatomical MRI images of the participant were aligned to the participant head locations in the MEG scanner.

Cortical parcellation for grouping. We employed a standard brain atlas provided by the Human Connectome Project (HCP) dataset [7], which divides the cortical surface into 180 areas per hemisphere, resulting in 360 areas across both hemispheres, each of which serves as a group information for MEG source estimation in this study. To assign these areas for the participant cortical surface, we used the Freesurfer software (http://surfer.nmr.mgh.harvard.edu/) to normalize the participant's MRI structure data to the standard brain space and the cortical surface was extracted from the normalized participant brain using the polygon model with 15,002 vertices at the gray/white matter outer boundary. Then each vertex was assigned with a label for the corresponding area.

There were exceptional vertices that did not correspond to any of the HCP-defined areas, which were located near the corpus callosum. These vertices were labeled with the 361st area (1,110 vertices were identified in total).

2.3 Simulation data generation and the hyperparameter

Cortical source current model. To evaluate source estimation accuracy, we simulated temporally localized and spatially distinct cortical activations across early visual areas. In this simulation, three regions of interest (ROIs) were defined: V1, V2, and V3, and they were sequentially activated as shown in the left panel of Fig. 2. Each ROI's cortical source current followed a sinusoidal waveform with a half-cycle duration of 25 ms, peaking at 13 ms (V1), 25 ms (V2), and 37 ms (V3), respectively. This waveform was assigned to each ROI in left or right hemisphere and its polarity was reversed in the opposite hemisphere (the positive activation was assigned to the left hemisphere in condition 1 and the polarity was reversed in condition 2).

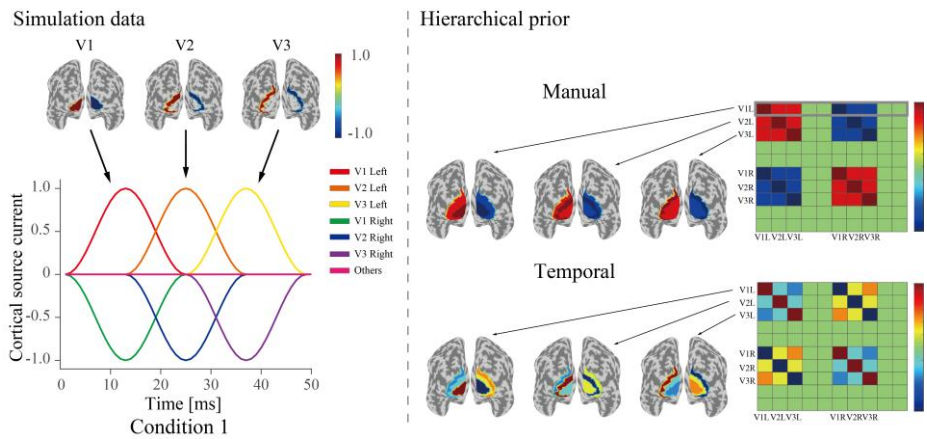

Fig. 2. Schematic view of cortical source current model and hierarchical prior.

Sensor noise model. The noise was independently added to each MEG sensor, each time point, and each trial. It was sampled from a Gaussian distribution with a mean of 0 and standard deviation σ. The value of σ was adjusted such that the resulting signal-to-noise ratio (SNR) varied from -5 dB to 5 dB with 5 dB steps. The SNR was calculated using the following equation:

$$GFP_t = \sum_{n=1}^{N} b_{n,t}^2 \tag{32}$$

$$\mathrm{SNR} = \frac{1}{T_r}\left(\frac{\sum_t^{I_s} GFP_t/|I_s|}{\sum_t^{I_n} GFP_t/|I_n|} - 1\right) \tag{33}$$

where GFP_t represents the global field power at time t, $b_{n,t}$ denotes the MEG signal recorded from the nth sensor at time t, T_r denotes the number of trials, and I_s and I_n

correspond to the task and noise time windows, defined as periods during which the brain was assumed to be active and inactive, respectively. The resulting SNR was then converted to decibels (dB) using $10\log_{10}(\mathrm{SNR})$. To compute the SNR, the task time window was defined as 1-25 ms, while the noise time window was defined as 26–50 ms using only the added noise components. One simulation dataset in this scenario consisted of five runs, each containing ten trials per condition, resulting in a total of 100 trials (50 trials per condition).

Hyperparameters for hierarchical priors. To define the hyperparameters for hierarchical priors, we used two types of covariance matrices (Fig. 2, right): (1) a manually designed covariance matrix, and (2) a temporal covariance matrix. For the manually designed covariance matrix, the diagonal elements corresponding to the active ROIs (V1, V2, and V3) were set to 1, and all others to 0. The off-diagonal elements were set as follows: 0.8 for ROI pairs within the same hemisphere (e.g., left V1 and left V2), –1 for ROI pairs in opposite hemispheres corresponding to the same functional area (e.g., left V1 and right V1), and –0.8 for ROI pairs in opposite hemispheres corresponding to different functional areas (e.g., left V1 and right V2). The temporal covariance matrix was computed by calculating the covariance across time from the simulated cortical source current. These two covariance matrices were used as hyperparameters $\mathbf{C}_0$ in the hierarchical Bayesian models. For hARD and hgARD that omit to model inter-areal relationships, only the diagonal elements of these matrices were retained. In this simulation, the ideal value of λ_g is zero because the source current is uniform within each ROI. However, directly setting it to zero would cause a division by zero during the estimation process. Therefore, we set the value of λ_g to 1.0×10^{-3} in this study. To examine the effect of the hierarchical prior in estimation, the prior weight was varied from 0 to 0.05 with 0.01 steps.

Quantitative evaluation of models. To evaluate the estimation accuracy, we computed the area under the precision-recall curve (APR) [17] at the peak activation time point of each ROI, using the absolute values of the estimated currents normalized by the maximum across all vertices at the corresponding time point. In addition, we computed the spatiotemporal correlation between the estimated current and the cortical source current across all vertices and time points to assess the overall agreement. Before computing each metric, the estimated source currents were averaged across all 50 trials for each condition (i.e., condition 1 and condition 2).

3 Results

3.1 Simulation data analysis

Figure 3 shows the spatial distribution of the estimated current. The color represents estimated current amplitudes at the time points indicated at the top. The color scale was symmetrically adjusted based on the absolute maximum of the estimated current

amplitude. Values between 0 ±10% of the absolute maximum of the estimated current amplitude were not displayed on the map, to emphasize the most salient regions of activation. Representative results at -5 dB SNR under condition 1 using a prior weight of 0.03 for the two types of hierarchical priors are shown in this figure. The results demonstrate that BHBE accurately estimated the current distribution, closely matching the true source in both spatial location and polarity. In contrast, the other models produced scattered estimates with frequent polarity reversals in neighboring vertices.

Figure 4 shows the waveforms of the estimated current. The results show that BHBE accurately reconstructed the three distinct waveforms corresponding to V1, V2 and V3. In contrast, the other models produced noisy estimates, with a slight increase in amplitudes around the activation period of V1 (1–25 ms), but failed to reconstruct waveforms corresponding to the activation periods of V2 and V3.

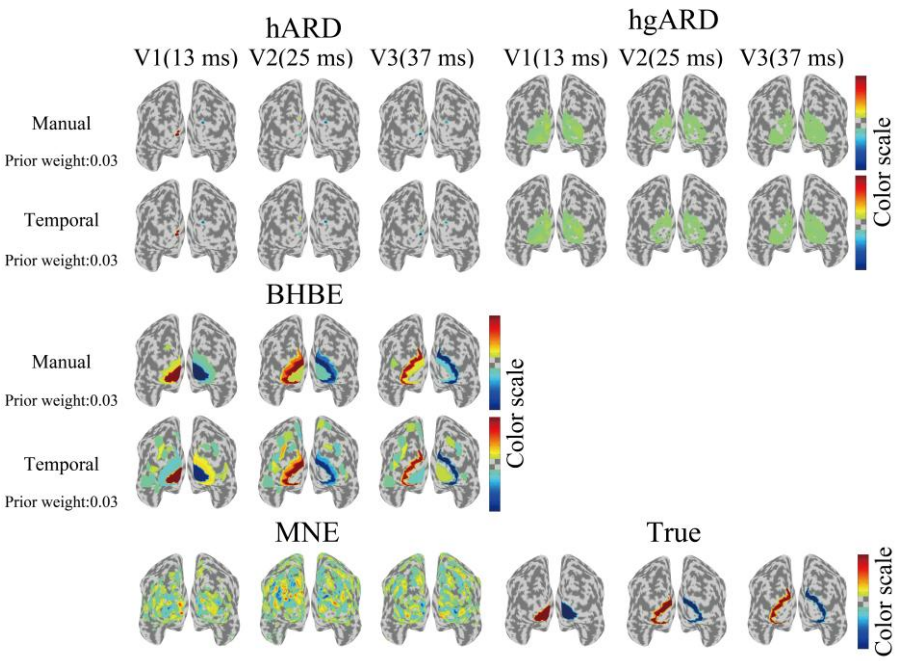

Fig. 3. Estimated cortical current maps

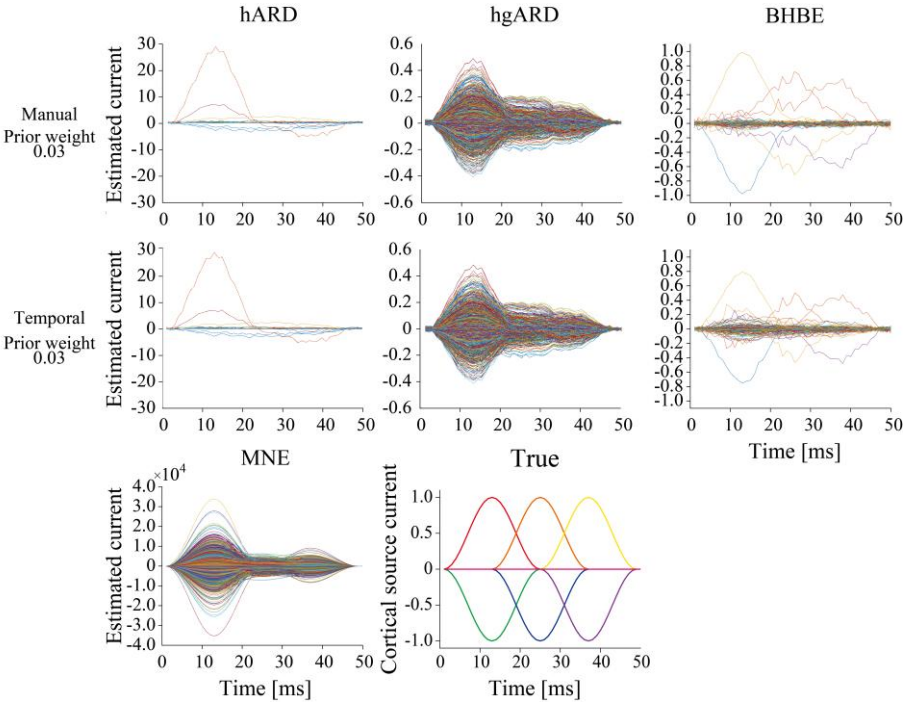

Fig. 4. Waveforms of estimated cortical current

Figure 5 shows the APR plotted against the prior weight (results from condition 1 are shown as representative examples). Since MNE is not a hierarchical Bayesian model, a single APR value was obtained and plotted across all prior weight for comparison. The results show that BHBE achieved higher APR values for all tested ROIs, indicating that estimated sources closely match the true source distribution, than all other models.

Figure 6 shows spatiotemporal correlation between the true and estimated cortical currents, plotted against the prior weight. The results demonstrate that BHBE outperformed the other models irrespective of the prior weight, indicating that our proposed model can accurately estimate the source cortical current with high fidelity in both space and time.

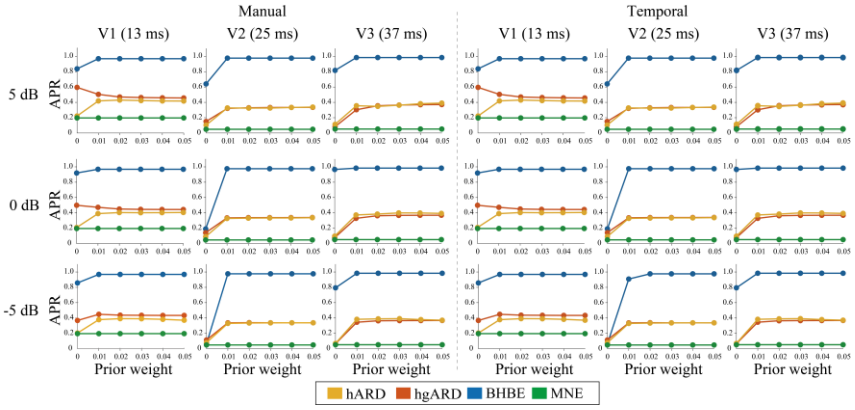

Fig. 5. APR of the estimated cortical current at the peak time point of each ROI

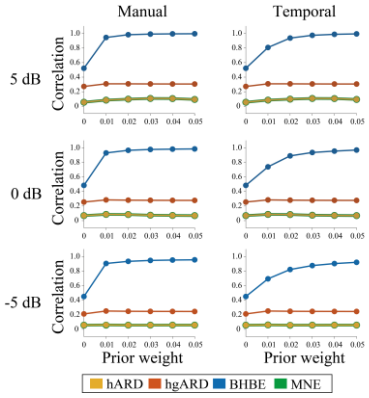

Fig. 6. Spatiotemporal correlation between the true and the estimated source cortical current

4 Conclusion and Discussion

In this study, we proposed a block-structured hierarchical Bayesian estimation (BHBE) model to improve the source estimation accuracy by incorporating both functional areal grouping and inter-areal relationships. We compared it with models that incorporated only functional areal grouping and neither of them. According to the simulation results, BHBE outperformed the other models in source estimation accuracy. These findings suggest that incorporating functional areal grouping and inter-areal relationships into MEG source estimation models can significantly enhance the accuracy of cortical source current estimation. Although the formulation is different from this study, a block-structured model has also been explored in a previous study [16]. A key difference lies in the type of information incorporated: whereas the previous study used physical relationships by grouping spatially adjacent vertices into blocks, our model incorporates functional relationships by grouping vertices according to a widely accepted cortical parcellation derived from fMRI data [7]. As a result, our approach may

be able to capture inter-areal dependencies that are not necessarily spatially adjacent but are functionally connected.

However, there might be a possibility that the present simulation setups were suitable for BHBE, potentially biasing the results of source estimation. In this study, the cortical source current was defined such that all vertices within each ROI exhibited uniform activation with the same amplitude, which may favor our model's framework of assigning the shared hyperparameters to vertices within the same group. To further test whether our model can generalize to various situations, it would be necessary to evaluate our model with more complex patterns of source cortical current. Although such complex activation patterns were not examined in this study, the flexible prior structure of our model may help address such situations. Theoretically, increasing the value of λ_g reduces the correlation of the estimated current within the same group, thereby increasing the variance of the estimated currents within the group. Thus, by adjusting λ_g our model is expected to flexibly accommodate a range of activation patterns, from homogeneous to more complex heterogeneous ones.

Although we used either a manually designed covariance matrix or a temporal covariance matrix as the hierarchical prior in this study, other sources of information may also be valuable for real data analysis. For example, fMRI experiments could be performed with the same experiment design as MEG, allowing the spatial and temporal covariance structures of the fMRI signals to be used as hierarchical prior distributions. By incorporating such neuroscientifically grounded priors to define inter-areal relationships, we might be able to analyze when and where each brain area is active and interacts during the processing of sensory information. Thus, our model framework may offer new opportunities for capturing dynamics of human neural activity with high spatiotemporal resolution.

Acknowledgments. This study was partially funded by Research Fellow of Japan Society for the Promotion of Science (JSPS) Grant Number 24KJ1127, JSPS KAKENHI Grant Number JP18KK0311, JP20H00600, JP25H01138, a research granted from Murata Science and Education Foundation, the Naito Foundation, and Moonshot Program 9 under Grant JPMJMS2291.

Disclosure of Interests. We have no competing interests.

References

1. Hämäläinen, M S, Ilmoniemi, R J: Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42 (1994). https://doi.org/10.1007/BF02512476.
2. Iwaki, S., Ueno, S.: Weighted minimum-norm source estimation of magnetoencephalography utilizing the temporal information of the measured data. J. Appl. Phys. 83, 6441–6443 (1998). https://doi.org/10.1063/1.367732.
3. Sato, M., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M.: Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage. 23, 806–826 (2004). https://doi.org/10.1016/j.neuroimage.2004.06.037.

4. MacKay, D.J.C.: Bayesian Interpolation. Neural Comput. 4, 415–447 (1992). https://doi.org/10.1162/neco.1992.4.3.415.
5. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer New York, NY (2006).
6. Destrieux, C., Fischl, B., Dale, A., Halgren, E.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage. 53, 1–15 (2010). https://doi.org/10.1016/j.neuroimage.2010.06.010.
7. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Essen, D.C.V.: A multi-modal parcellation of human cerebral cortex. Nature. 536, 171–178 (2016). https://doi.org/10.1038/nature18933.
8. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106, 1125–1165 (2011). https://doi.org/10.1152/jn.00338.2011.
9. Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 34, 537–541 (1995). https://doi.org/10.1002/mrm.1910340409.
10. Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., Salmelin, R.: Dynamic imaging of coherent sources: Studying neural interactions in the human brain. Proc. Natl. Acad. Sci. 98, 694–699 (2001). https://doi.org/10.1073/pnas.98.2.694.
11. Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V.: Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. 100, 253–258 (2003). https://doi.org/10.1073/pnas.0135058100.
12. Fox, M.D., Snyder, A.Z., Zacks, J.M., Raichle, M.E.: Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. Nat. Neuro-sci. 9, 23–25 (2006). https://doi.org/10.1038/nn1616.
13. Parisi, G.: Statistical Field Theory. addison (1988).
14. Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H.C., Sa-dato, N., Kamitani, Y.: Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders. Neuron. 60, 915–929 (2008). https://doi.org/10.1016/j.neuron.2008.11.004.
15. Archakov, I., Hansen, P.R.: A Canonical Representation of Block Matrices with Applications to Covariance and Correlation Matrices. Rev. Econ. Stat. 106, 1099–1113 (2024). https://doi.org/10.1162/rest_a_01258.
16. Creal, D., Kim, J.: Bayesian estimation of block covariance matrices. Available SSRN 4098134. (2022).
17. Davis, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd international conference on Machine learning - ICML '06. pp. 233–240. ACM Press, Pittsburgh, Pennsylvania (2006). https://doi.org/10.1145/1143844.1143874.