Doctoral Dissertation

Towards Low-Cost and High-Reproducible MEG Current Source Reconstruction using Meta-Analysis fMRI Data

Keita Suzuki

Program of Information Science and Engineering
Graduate School of Science and Technology
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Engineering

Keita Suzuki

Thesis Committee:
Supervisor   Kazushi Ikeda
(Professor, Division of Information Science)
Yoshinobu Sato
(Professor, Division of Information Science)
Motoaki Kawanabe
(Professor, Division of Information Science)
Jun Morimoto
(Professor, Graduate School of Informatics, Kyoto University)
Makoto Fukushima
(Assistant Professor, Division of Information Science)
Okito Yamashita
(Department Head, CBI, ATR Neural Information Analysis Laboratories)

Towards Low-Cost and High-Reproducible MEG Current Source Reconstruction using Meta-Analysis fMRI Data*

Keita Suzuki

Abstract

Magnetoencephalography (MEG) offers a unique way to noninvasively investigate millisecond-order cortical activities by mapping sensor signals (magnetic fields outside the head) to cortical current sources using current source reconstruction methods. Current source reconstruction is defined as an ill-posed inverse problem, since the number of sensors is less than the number of current sources. One powerful approach to solving this problem is to use functional MRI (fMRI) data as a spatial constraint, although it boosts the cost of measurement and the burden on subjects. Here, we show how to use the meta-analysis fMRI data in two different ways instead of the individually recorded fMRI data. In the first approach, single meta-analysis fMRI data was imported as prior information of the hierarchical Bayesian estimation to mitigate the differences between the meta-analysis and individual data. In the second approach, a combination of multiple meta-analysis fMRI data was automatically selected based on MEG data and used as prior information of currents. Using simulations, we found the performances of both approaches were better than conventional methods. Notably, the first approach showed better performance than the estimation using low-quality individual fMRI data. By applying experimental data of a face recognition task, we qualitatively confirmed that group analysis results using the first approach showed a tendency similar to the results using the individual fMRI

---

data. Using the second approach, in addition to reconstructing a plausible current map, it also estimated related terms derived from meta-analysis data. Our results indicate that the use of meta-analysis fMRI data improves current source reconstruction without additional measurement costs and leads to reproducible results because of omitting the contaminated prior information.

fMRIのメタ分析結果を利用した低コストかつ高い再現性を持つ MEG 電流源推定問題に関する研究*

鈴木 啓大

内容梗概

　脳磁図 (MEG) は非侵襲性，ミリ秒オーダーの時間分解能，そして比較的高い空間分解能を同時に達成できる唯一の方法である．ここで，MEG の計測データは頭外部のセンサーによって計測されるため，空間分解能の高さは電流源の推定手法に依存する．観測センサ数よりも推定電流源数が多いことから，この推定は劣決定逆問題として定義される．この問題を解くための手法の一つとして，機能的 MRI(fMRI) 計測により得られる高い空間分解能を持つデータを電流源推定の空間制約として使用することが提案されている．しかしながら，MEG 計測に加えて fMRI 計測も要するため，そのコストと被験者への負担は大きく増加する．したがって，我々は計測によって得られる fMRI データの代わりに，多くの研究結果から得られるメタ分析結果を空間制約として使用することを提案する．ここで，メタ分析結果の組み込み方には二通りの方法を用意した．一つ目のアプローチは，階層変分ベイズ推定の枠組みにメタ分析結果を組み込む方法である．階層変分ベイズ推定の効果により，メタ分析結果と個人のデータとの不一致が緩和されることが期待される．二つ目のアプローチは，複数のメタ分析結果の組み合わせを観測データから自動的に決定し，これを制約情報として組み込む方法である．シミュレーション実験の結果により，両アプローチは一般的に用いられている電流源推定手法よりも良い推定解が得られることが確認できた．特に，階層変分ベイズ推定を用いた枠組みでは，低いクオリティの fMRI データを使用する場合よりもかえって推定が向上することが確認できた．さらに，実データを用いた検証では，実際の fMRI データを制約条件として推定する場合と非常に類似した結果が得られた．我々の結果は，メタ分析結果を用いることにより電流源推定を低コ

ストかつ，場合によっては実際の fMRI データを使用するよりも高い再現性を持って行うことができることを示唆している．

キーワード

脳磁図, 逆問題, 電流源推定, メタ分析, 階層ベイズ

Contents

List of Figures

List of Tables

1. Introduction

Non-invasive neuroimaging measurements can record human brain activities while avoiding the risks of brain surgery. It is a tremendous feature in revealing our brain functions (Baillet, 2017) and also the treatment of brain dysfunctions caused by diseases such as autism (Kikuchi et al., 2016), schizophrenia (Rojas, 2019), epilepsy (Stefan and Trinka, 2017), and so on (Uhlhaas et al., 2017). Today, major neuroimaging modalities are broadly divided into two types: those that measure hemodynamic response and electromagnetic signals.

The modalities that measure hemodynamic response include functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS). In particular for fMRI, due to its high spatial resolution and the established way of reporting based on the standard brain (MNI-ICBM152), a massive amount of research is reported, and consequently, meta-analysis studies are popular (Laird et al., 2011; Smith et al., 2009; Dockès et al., 2020; Yarkoni et al., 2011). However, the temporal resolution of fMRI is fundamentally limited because of the slow hemodynamic response to neural activities.

On the other hand, magnetoencephalography (MEG) and electroencephalogram (EEG) detect fluctuations of magnetic/electric fields caused by synchronized activities of cortical pyramidal neurons. Therefore, these modalities record neural activity induced signals directly with millisecond-order temporal resolution (M. hamalainen, Hari, Ilmoniemi, Knuutila, and Lounasmaa, 1993). Here, the process of measuring current sources using sensors is formulated as the forward problem. Conversely, we can discuss the measured activity on a cortex by mapping signals from sensors to sources; known as the inverse problem or current source reconstruction (Baillet, Mosher, and Leahy, 2001; Gross et al., 2013).

Therefore, by solving the inverse problem of MEG/EEG, we can obtain neurophysiological data with not only high temporal but also adequate spatial resolution. Furthermore, fMRI and its meta-analysis data can be combined in the inverse problem solution to enhance the spatial resolution.

In this article, we focused on MEG rather than EEG because of its higher spatial accuracy. However, we note that our proposals and discussions are also applicable to EEG.

1.1 Forward problem

The purpose of the forward model is to formulate the relationships between the activities of current sources and sensor signals. It consists of three subcomponents: the source model, the head model, and the measurement model.

For the first, what the source model describes is the electrophysiology of neurons and the assumption on currents. When excitatory postsynaptic potentials (EPSPs) are generated at apical dendrites, intracellular currents flow from apical dendrites to the soma of a pyramidal cell. This currents flow is known as primary currents. MEG mainly records the primary currents while the primary currents also generated extracellular currents called secondary currents. Although the primary currents of a single neuron are too small to generate observable signals, dendrites of cortical pyramidal cells are arranged perpendicularly to the cortical surface. Consequently, synchronously fluctuating dendric currents enable to observe electromagnetic signals, and synchronized activity of tens of thousands of neurons is regarded as a current dipole. The dipole fitting model is one of the source models and it assumes that only a few numbers of dipoles are activated simultaneously. A key problem with this model is that the number of sources must be decided a priori and the result approximates currents as few dipoles even if it is distributed. On the other hand, distributed source models assume that each vertex of the head model has a dipole and model the distribution of them. Our study model the sources as a distributed source model.

The shape of the subject's head is modeled as a 3D polygonal model. In general, current sources are assumed on vertices of the cortex piece of the polygon. The simplest one is a spherical model specified by the origin of the sphere and its radius. However, it does not consider the shapes of the cortex. Therefore, structural information derived from MRI data is used to construct a realistic model. Then, the number of vertices controls the difficulty of the inverse problem in a trade-off between the ill-posedness and the structural constraint. The more vertices induce not only better structural constraint but also worse ill-posed condition. Although it depends on the purpose of the study, thousands of current sources are commonly assumed (Henson et al., 2009).

The measurement model reflects the type, positions, and orientations of the MEG sensors. The type of MEG sensors is categorized as magnetometer or

gradiometer. The magnetometer uses a coil and detects magnetic fields, including many signals outside the head. The gradiometer uses two coils in the opposite direction and detects differential signals of them. Hence, magnetic fields far from coils cancel each other, and only magnetic fields generated close to the coils are detected by the gradiometer. Most MEG sensors are equipped with both magnetometer and gradiometer. Their sensor information is incorporated in the volume conductor model.

Given the source model, the head model, and the sensor model, the volume conductor model formulates the equations that describe the relationships between the current sources and the sensors by solving Maxwell's equations under quasi-static conditions. They are numerically calculated using boundary-element methods or finite-element methods. Consequently, the forward model is summarized into a matrix called the lead field describing the relationships between sources and sensors as linear equations. When discussing the inverse problem, we regard a lead field matrix is known.

1.2 Inverse problem

The current source reconstruction methods using the distributed source model rely on solutions of the inverse problem, which is an ill-posed problem requiring constraints in addition to data, since the number of observed signals is only a few hundred, whereas the number of considered current sources is in the thousands. This problem is similarly defined for electroencephalogram (EEG), which observes changes in the electric field caused by neural activity. Many studies have proposed various types of additional constraints and ways to incorporate them in the mathematical model. The most straightforward but powerful one is minimum norm estimation (MNE) (Hämäläinen et al., 1993; Hämäläinen and Ilmoniemi, 1994; Wang et al., 1992). The MNE solution is obtained by minimizing a cost function consisting of a data-fitting error and the L2 norm of currents. Moreover, a number of extensions of MNE have been proposed: incorporating a Laplacian filter resulting in a spatial smoothness solution (LORETA and its family; Pascual-Marqui et al., 1994, Pascual-Marqui, 2007, 2002), weighting the lead field matrix to consider depth structure (wMNE; Lin et al., 2006), considering data-adaptive prior distribution on each source to obtain a spatially sparse

3

solution (Champagne; Cai et al., 2021; Owen et al., 2012; Wipf et al., 2010), and in the data-adaptive scheme, composing a current covariance matrix as a linear combination of candidate components (multiple sparse priors; Friston et al., 2008). For more details and recent reviews of MEG/EEG source reconstruction, refer to the article or book (He et al., 2018; Sekihara and Nagarajan, 2015).

On the other hand, when the functional constraint is considered, fMRI data are ideal due to the characteristics of high spatial resolution. Dynamic statistical parametric mapping (dSPM; Dale et al., 2000) uses an fMRI statistical map as a hard constraint of the source current variance by assuming that the prior current variances are fixed according to the fMRI statistical map. The hierarchical variational Bayesian (hVB) estimation approach (Sato et al., 2004) similarly uses an fMRI statistical map but incorporates it as a soft constraint of current variance by assuming that the prior current variances are random variables with probabilistic distributions having parameters according to the fMRI statistical map (i.e., they incorporate the fMRI map into the hierarchical prior of the current variance). Therefore, even if the intrinsic difference in the measurement principles of fMRI and MEG causes a discrepancy between the spatial patterns of the fMRI activation map and the source current map (Kaneoke, 2006), the hVB approach is able to resolve such discrepancies.

Although the fMRI prior information leads to significant improvement in terms of spatial accuracy, there are practical problems. First, it requires additional cost to measure the fMRI data. It increases not only the measurement cost but also the burden on the subject. This burden also causes a decrease in measurement quality and leads to a second issue. It is also inapplicable to a published MEG dataset without fMRI records. Furthermore, when targeting modalities with lower measurement costs, such as an EEG or an optically pumped magnetometer (OPM) (Boto et al., 2018), the measurement cost of fMRI will become relatively more serious. As a second problem, the reliability of individual-level statistical maps is not always sufficient, especially for cognitive or emotional tasks that induce smaller signal changes compared with visual and motor tasks (Elliott et al., 2020). The quality of fMRI data is strongly influenced by the subject's condition and noise caused by a system and physiology (Geissler et al., 2007; Welvaert and Rosseel, 2013).
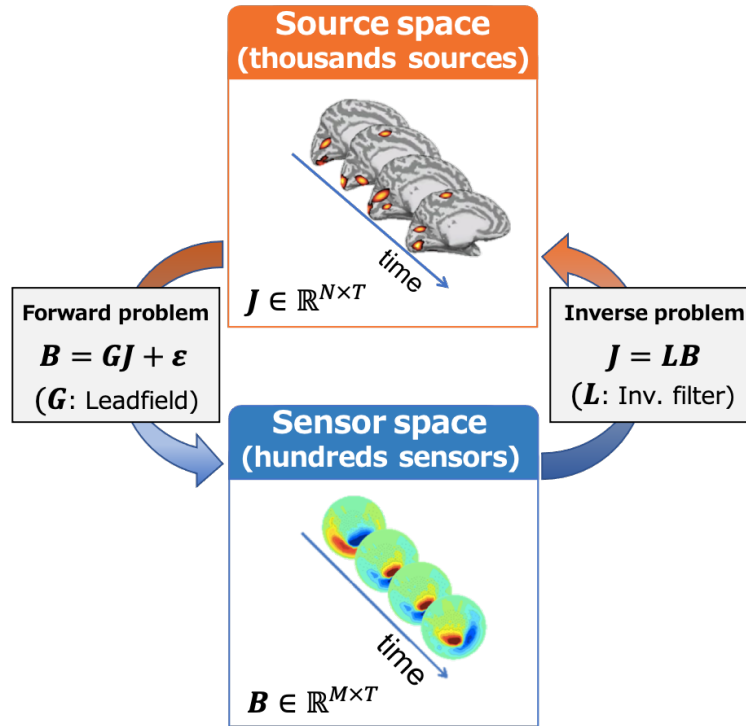
Figure 1.1. Illustration of the forward and inverse problem. The forward problem is modeling the lead field matrix $\boldsymbol{G}$. It is calculated using the sensor positions, types, and a head model. Under the modeled $\boldsymbol{G}$, the observation process of magnetic fields' time series $\boldsymbol{B}$ (M sensors by T time points) are expressed by the multiplication of $\boldsymbol{G}$ and currents' time series $\boldsymbol{J}$ (N sources by T time points) with adding measurement noise $\boldsymbol{\varepsilon}$. On the other hand, the goal of the inverse problem is summarized in deriving the inverse filter $\boldsymbol{L}$. Then, the currents' time series are reconstructed by the multiplication of $\boldsymbol{L}$ and $\boldsymbol{B}$. Note that $\boldsymbol{J}$ and $\boldsymbol{B}$ are visualized as a topography here using a head model and sensor positions, although actual data are time series signals.

1.2.1 Our contribution

In this study, to solve the issues mentioned above, we propose two different approaches using meta-analysis fMRI data as a prior information instead of an individual prior. The meta-analysis fMRI data form a statistical map synthesized

from thousands of published fMRI studies. We used the meta-analysis results available from the Neurosynth open-source project (Yarkoni et al., 2011).

A key idea of the first approach is that we took single meta-analysis data as the hierarchical prior distribution of the current variance in hVB estimation (Sato et al., 2004) rather than the prior current variance as in the dSPM approach, since the former provides an adaptive way to incorporate fMRI information based on MEG measurements. The feasibility of meta-analysis fMRI data as prior information was assessed by comparing the results of source reconstruction with those estimated using an individual fMRI prior. The results were quantitatively and qualitatively evaluated using simulated and experimental data, respectively. For the simulation, we conducted multi-modal data generation. Both fMRI and MEG data were generated based on the same ground truth to mimic a situation in which these are obtained from an individual. In order to bring the statistical properties of simulated data closer to those of real data, this simulation adopted real data to the extent possible. Using simulated data, we studied the effectiveness of a meta-analysis fMRI prior compared with a contaminated individual fMRI prior. We also investigated the appropriate range of an important hyper-parameter we call prior weight. Using experimental data, we conducted group analysis using the meta-analysis fMRI prior. Furthermore, we verified whether it is possible to select optimal meta-analysis fMRI prior based on a statistical criterion for both simulated and experimental data.

As mentioned in the discussion chapter, the selection of the meta-analysis data is one of the most critical points for the first approach. Moreover, the results of meta-analysis data selection suggest the possibility of the combined meta-analysis data as prior information. Therefore, we propose another approach that uses meta-analysis data as prior information but the selection of them is automatically determined using observation data. For this approach, named relevant spatial prior (RSP) estimation, we conducted evaluations using another simulation setting and the same experimental data.

1.3 Dissertation outline

The outline of this dissertation is illustrated in Fig. 1.2. The main topic is our two approaches for the current source reconstruction, hVB and RSP. After the

Introduction chapter, we mathematically described the inverse problem in *Mathematical formulation of the source reconstruction* chapter. Here, we tried to explain both approaches in the general framework. Next, we introduced *Materials and Methods* common to both approaches. After that, the thesis is branched into two flows. Both flows include details of simulated and experimental data and the results of the analysis. Finally, flows are merged into the *Discussion* chapter. Supplemental information is summarized into several appendix chapters.
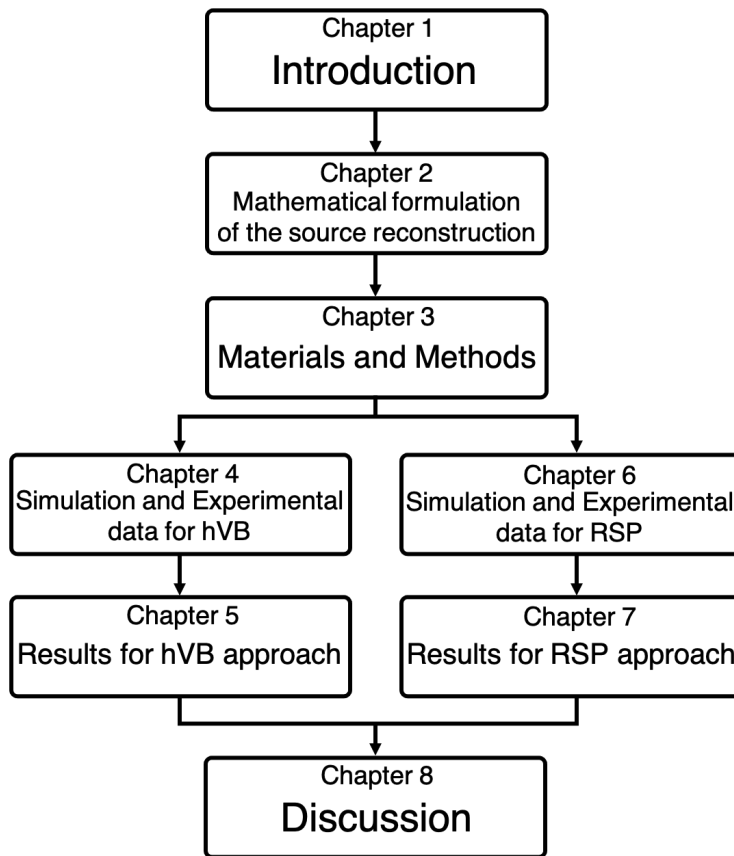
Figure 1.2. Outline of the dissertation.

2. Mathematical formulation of the source reconstruction

In this chapter, we formulate the inverse problem of source reconstruction as mathematical expressions. In particular, we introduce that the problem is generally explained as the differences of the prior of the current covariance in Bayesian perspective. Here we show the notations for convenience.

Table 2.1. List of mathematical notations

Symbol	Definition
$\mathbf{1}$	Vector consists of ones.
$\boldsymbol{I}$	Identity matrix.
$P(x)$	Probability distribution of random variable x.
$P(x\|y)$	Conditional probability distribution of x given y.
$\mathcal{N}(x\|\mu, \sigma^2)$	Univariate Gaussian distribution over x with mean μ and variance σ^2.
$\mathcal{N}(\boldsymbol{x}\|\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Multivariate Gaussian distribution over $\boldsymbol{x}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.
$\Gamma(\gamma)$	Gamma function defined as $\int_0^\infty t^{\gamma-1}e^{-t}dt$.
$\Gamma(\alpha\|\alpha_0, \gamma_0)$	Gamma distribution over α with mean α_0 and degree of freedom γ_0. Defined as $\frac{1}{\alpha\Gamma(\gamma_0)}(\alpha\gamma_0/\alpha_0)^{\gamma_0}e^{-\alpha\gamma_0/\alpha_0}$.

2.1 Bayesian formulation of current source reconstruction

From the Bayesian perspective, the inverse problem is commonly formulated as

$$P(\boldsymbol{j}|\boldsymbol{b}) = \frac{P(\boldsymbol{b}|\boldsymbol{j})P_0(\boldsymbol{j})}{P(\boldsymbol{b})}, \tag{2.1}$$

where $\boldsymbol{j} = \{j_n | n = 1 : N\}$ is neural activities on N dipoles, and $\boldsymbol{b} = \{b_m | m = 1 : M\}$ is observed MEG recordings by M sensors. Since the purpose of current source reconstruction is to estimate current source $\boldsymbol{j}$ from the observed magnetic

field data $\boldsymbol{b}$, a posterior distribution $P(\boldsymbol{j}|\boldsymbol{b})$ is what we want. The right-hand side is composed of a likelihood $P(\boldsymbol{b}|\boldsymbol{j})$, a prior of currents $P_0(\boldsymbol{j})$, and a marginal likelihood $P(\boldsymbol{b})$. As mentioned in the section 1.1, we assume that the currents $\boldsymbol{j}$ and MEG recordings $\boldsymbol{b}$ are linearly related through the lead field matrix $\boldsymbol{G} = \{G_{mn}|m = 1 : M, n = 1 : N\}$ (Hämäläinen et al., 1993) as

$$\boldsymbol{b} = \boldsymbol{G}\boldsymbol{j} + \varepsilon, \tag{2.2}$$

where ε is the observation noise. When we assume the observation noise to be white noise, ε follows a Gaussian distribution with the diagonal covariance $\sigma^2\boldsymbol{I}$. Then, the likelihood function is written as

$$P(\boldsymbol{b}|\boldsymbol{j}) = \mathcal{N}(\boldsymbol{b}|\boldsymbol{G}\boldsymbol{j}, \sigma^2\boldsymbol{I}). \tag{2.3}$$

The above likelihood is commonly assumed for most source reconstruction methods. Hence, the differences in methods are mainly explained as the differences in the assumptions on the prior distribution $P_0(\boldsymbol{j})$. When we assume that the prior distribution is Gaussian with mean 0 and some covariance matrix $\boldsymbol{A}^{-1}$, then

$$P_0(\boldsymbol{j}) = \mathcal{N}(\boldsymbol{j}|\boldsymbol{0}, \boldsymbol{A}^{-1}). \tag{2.4}$$

Under this prior distribution, the right-hand side of Eq.(2.1) follows a Gaussian. Therefore, the left-hand side of Eq.(2.1) is a Gaussian distribution as well. Here we express the posterior distribution using mean $\bar{\boldsymbol{j}}$ and covariance $\boldsymbol{\Lambda}^{-1}$ as

$$P(\boldsymbol{j}|\boldsymbol{b}) = \mathcal{N}(\boldsymbol{j}|\bar{\boldsymbol{j}}, \boldsymbol{\Lambda}^{-1}). \tag{2.5}$$

To calculate $\bar{\boldsymbol{j}}$ and $\boldsymbol{\Lambda}^{-1}$, we simplify the exponential part of Eq.(2.5) as

$$-\frac{1}{2}(\boldsymbol{j} - \bar{\boldsymbol{j}})^\top\boldsymbol{\Lambda}(\boldsymbol{j} - \bar{\boldsymbol{j}}) = -\frac{1}{2}\boldsymbol{j}^\top\boldsymbol{\Lambda}\boldsymbol{j} + \boldsymbol{j}^\top\boldsymbol{\Lambda}\bar{\boldsymbol{j}} + C. \tag{2.6}$$

C denotes terms unrelated to $\boldsymbol{j}$. Next, when we focus on $\boldsymbol{j}$, denominator of the right-hand side of Eq.(2.1) is ignored. Hence, the exponential part of the right-hand side of Eq.(2.1) is reformulated as a quadratic expression for $\boldsymbol{j}$ using Eq.(2.3) and (2.4) as

$$-\frac{1}{2}\{\sigma^{-2}(\boldsymbol{b} - \boldsymbol{G}\boldsymbol{j})^\top(\boldsymbol{b} - \boldsymbol{G}\boldsymbol{j}) + \boldsymbol{j}^\top\boldsymbol{A}\boldsymbol{j}\} = -\frac{1}{2}\boldsymbol{j}^T(\boldsymbol{A} + \sigma^{-2}\boldsymbol{G}^\top\boldsymbol{G})\boldsymbol{j} + \sigma^{-2}\boldsymbol{j}^\top\boldsymbol{G}^\top\boldsymbol{b} + C. \tag{2.7}$$

By comparing Eq.(2.6) and (2.7), mean and precision (inverse of covariance) of the posterior Eq.(2.5) are calculated as

$$\boldsymbol{\Lambda} = \boldsymbol{A} + \frac{1}{\sigma^2}\boldsymbol{G}^\top\boldsymbol{G}, \tag{2.8}$$

$$\bar{\boldsymbol{j}} = \frac{1}{\sigma^2}\boldsymbol{\Lambda}^{-1}\boldsymbol{G}^\top\boldsymbol{b}. \tag{2.9}$$

$\bar{\boldsymbol{j}}$ is a minimum mean squared error (MMSE) estimation and also it corresponds to a maximum a posteriori (MAP) estimation because mean and mode of a Gaussian distribution are same. Practically, the size of $\boldsymbol{\Lambda}^{-1}$ in Eq.(2.9) is N by N (N is thousands) and this inversion is computationally heavy, therefore it is reformulated using the matrix inversion $(\boldsymbol{X}^{-1}+\boldsymbol{Y}^\top\boldsymbol{Z}^{-1}\boldsymbol{Y})^{-1}\boldsymbol{Y}^\top\boldsymbol{Z}^{-1} = \boldsymbol{X}\boldsymbol{Y}^\top(\boldsymbol{Y}\boldsymbol{X}\boldsymbol{Y}^\top+\boldsymbol{Z})^{-1}$,

$$\begin{aligned}\bar{\boldsymbol{j}} &= \boldsymbol{A}^{-1}\boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{A}^{-1}\boldsymbol{G}^\top + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{b} \\ &= \boldsymbol{A}^{-1}\boldsymbol{G}^\top\boldsymbol{S}^{-1}\boldsymbol{b}, \end{aligned} \tag{2.10}$$

while $\boldsymbol{S} = (\boldsymbol{G}\boldsymbol{A}^{-1}\boldsymbol{G}^\top + \sigma^2\boldsymbol{I})$ is called model data covariance and is a covariance of marginal likelihood $P(\boldsymbol{b})$.

Here, $\frac{1}{\sigma^2}\boldsymbol{\Lambda}^{-1}\boldsymbol{G}^\top = \boldsymbol{A}^{-1}\boldsymbol{G}^\top\boldsymbol{S}^{-1}$ is an inverse filter $\boldsymbol{L}$ because it maps observations $\boldsymbol{b}$ to estimated currents $\bar{\boldsymbol{j}}$. Furthermore, characteristics of an inverse filter are summarized in $\boldsymbol{A}$, the precision matrix of the prior distribution. It means that differences in methods reflect differences in knowledge of current distribution.

2.2 Temporal expansion

In the previous section 2.1, we assumed that we have data at only a single time point. However, data has many time points in the real situation. Therefore, we expand the formulations to the temporal direction.

When we estimate currents of T time points $\boldsymbol{J} = (\boldsymbol{j}_1, ..., \boldsymbol{j}_T) \in \mathbb{R}^{N \times T}$ using temporally independent and identically distributed (i.i.d) observations $\boldsymbol{B} = (\boldsymbol{b}_1, ..., \boldsymbol{b}_T) \in \mathbb{R}^{M \times T}$, then a Bayesian theorem Eq.(2.1), a likelihood Eq.(2.3), and a prior distribution Eq.(2.4) are rewritten as

$$P(\boldsymbol{J}|\boldsymbol{B}) = \frac{P(\boldsymbol{B}|\boldsymbol{J})P_0(\boldsymbol{J})}{P(\boldsymbol{B})}, \tag{2.11}$$

$$P(\boldsymbol{B}|\boldsymbol{J}) = \prod_{t=1}^{T} P(\boldsymbol{b}_t|\boldsymbol{j}_t)$$

$$= \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{b}_t|\boldsymbol{G}\boldsymbol{j}_t, \sigma^2 \boldsymbol{I}), \tag{2.12}$$

$$P_0(\boldsymbol{J}) = \prod_{t=1}^{T} P_0(\boldsymbol{j}_t)$$

$$= \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{j}_t|\boldsymbol{0}, \boldsymbol{A^{-1}}). \tag{2.13}$$

By following similar calculations in section 2.1 using above two equations, the same inverse filter is derived. Then, estimation of T time points $\bar{\boldsymbol{J}} = (\bar{\boldsymbol{j}}_1...\bar{\boldsymbol{j}}_T) \in \mathbb{R}^{N \times T}$ is calculated as

$$\bar{\boldsymbol{J}} = \frac{1}{\sigma^2} \boldsymbol{A}^{-1} \boldsymbol{G}^\top \boldsymbol{B} \tag{2.14}$$

$$= \boldsymbol{A}^{-1} \boldsymbol{G}^\top \boldsymbol{S}^{-1} \boldsymbol{B}. \tag{2.15}$$

2.3 Minimum norm estimation

The minimum norm estimation (MNE) is one of the most classical and commonly used method (Hämäläinen et al., 1993; Hämäläinen and Ilmoniemi, 1994; Wang et al., 1992). While it is originally derived in the context of constrained optimization problems, the solution of MNE can be characterized by the below prior

$$P_0(\boldsymbol{j}) = P(\boldsymbol{j}|\alpha)$$

$$= \mathcal{N}(\boldsymbol{j}|\boldsymbol{0}, \alpha^{-1}\boldsymbol{I}), \tag{2.16}$$

where $\alpha\boldsymbol{I}$ is a diagonal precision matrix. The MAP estimation is calculated using Eq.(2.9) as

$$\bar{\boldsymbol{j}} = \frac{1}{\sigma^2} \left(\frac{1}{\alpha}\boldsymbol{I} + \frac{1}{\sigma^2}\boldsymbol{G}^\top\boldsymbol{G} \right)^{-1} \boldsymbol{G}^\top\boldsymbol{b}$$

$$= \left(\frac{\sigma^2}{\alpha}\boldsymbol{I} + \boldsymbol{G}^\top\boldsymbol{G} \right)^{-1} \boldsymbol{G}^\top\boldsymbol{b}. \tag{2.17}$$

11

As the prior Eq.(2.16) indicated, MNE assumed that current dipoles are independent of each other and there is no knowledge on amplitudes of them.

If we have specific knowledge of each dipole, it is simply incorporated in diagonal elements of the precision matrix of the prior distribution. Then, the prior distribution is

$$
\begin{aligned}
P_0(\boldsymbol{j}) &= P(\boldsymbol{j}|\boldsymbol{\alpha}) \\
&= \mathcal{N}(\boldsymbol{j}|\boldsymbol{0}, \boldsymbol{A}^{-1}),
\end{aligned}
\tag{2.18}
$$

where $\boldsymbol{A} = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_N)$ is the diagonal precision matrix of the currents and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$ is a vector of diagonal elements. Dynamic statistical parametric mapping (dSPM; Dale et al., 2000) uses an fMRI statistical map to determine this $\boldsymbol{A}$. Therefore, the estimated current map of dSPM reflects the fMRI data as a hard constraint. However, if there is a discrepancy between the spatial patterns of the fMRI activation map and the source current map due to the difference in the measurement principles of them (Kaneoke, 2006), the result also reflects such errors. In section 2.4, we introduce another way to use a prior data as a soft constraint.

2.4 Hierarchical variational Bayesian estimation

For the hierarchical variational Bayesian estimation (hVB; Sato et al., 2004), we assume the observation noise $\boldsymbol{\varepsilon}$ follows an identically independent Gaussian distribution $\mathcal{N}(\boldsymbol{0}, (\beta\boldsymbol{\Phi})^{-1})$, where β is a scaling parameter and $\boldsymbol{\Phi}$ is a noise precision matrix scaled by β ($\boldsymbol{\Phi}$ is normalized to satisfy $\mathrm{tr}(\boldsymbol{\Phi}) = M$ and is typically determined from pre-stimulus rest period measurements),.

To incorporate spatial data as a soft constraint, the hVB model assumes the prior distribution in a hierarchical way:

$$
\begin{aligned}
P_0(\boldsymbol{J}|\boldsymbol{\Lambda}) &= P_0(\boldsymbol{J}|\boldsymbol{A}, \beta) \\
&= \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{j}_t|\boldsymbol{0}, (\beta\boldsymbol{A})^{-1}),
\end{aligned}
\tag{2.19}
$$

$$P_0(\boldsymbol{A}) = \prod_{n=1}^{N} P_0(\alpha_n)$$

$$= \prod_{n=1}^{N} \Gamma(\alpha_n | \alpha_{n0}, \gamma_0), \tag{2.20}$$

$$\Gamma(\alpha | \alpha_0, \gamma_0) \equiv \frac{1}{\alpha \Gamma(\gamma_0)} (\alpha \gamma_0 / \alpha_0)^{\gamma_0} e^{-\alpha \gamma_0 / \alpha_0}, \tag{2.21}$$

$$P_0(\beta) = 1/\beta. \tag{2.22}$$

$\Gamma(\alpha | \alpha_0, \gamma_0)$ represents a Gamma distribution with mean α_0 and degree of freedom γ_0, where $\Gamma(\gamma_0) \equiv \int_0^\infty t^{\gamma_0 - 1} e^{-t} dt$ is the Gamma function. Here, α_{n0} is determined based on fMRI information as $\alpha_{n0}^{-1} \propto \mathrm{fMRI}_n^2$, where fMRI_n is an imported statistical value of the nth dipole. Equation (2.22) is the prior distribution of the scaling parameter β, defined as a non-informative prior distribution.

Using the likelihood function Eq. (2.12) and the prior distributions Eq. (2.19), (2.20), (2.22), we compute the joint posterior distribution of the current $\boldsymbol{J}$, prior precision $\boldsymbol{A}$, and scaling parameter β given the magnetic field $\boldsymbol{B}$ as follows:

$$P(\boldsymbol{J}, \boldsymbol{A}, \beta | \boldsymbol{B}) = \frac{P(\boldsymbol{B} | \boldsymbol{J}, \beta) P_0(\boldsymbol{J} | \boldsymbol{A}, \beta) P_0(\boldsymbol{A}) P_0(\beta)}{P(\boldsymbol{B})}. \tag{2.23}$$

However, it cannot be solved analytically because the marginal likelihood $P(\boldsymbol{B})$ involves integration over $\boldsymbol{A}$. Therefore, the approximated posterior distribution is calculated using the variational Bayesian method. The resulting algorithm consists of iterative updating of the posterior distribution $Q(\boldsymbol{J}, \beta)$ and $Q(\boldsymbol{A})$ (Appendix A).

We used the VBMEG toolbox (Takeda et al., 2019) to conduct the hVB estimation. In the VBMEG implementation, the hyperparameter called the prior weight controls the relative confidence between the fMRI information and the observed MEG. The prior weight w is defined as

$$w = \frac{\gamma_0}{\gamma_0 + T/2}, \tag{2.24}$$

where γ_0 is the degree of freedom of the hierarchical prior distribution. The prior weight ranges from 0 to less than 1. This confidence parameter appears in updating the current variance α_n^{-1}:

$$\alpha_{np}^{-1} = w \alpha_{n0}^{-1} + (1 - w) \left\langle \frac{\beta}{T} \sum_{t=1}^{T} j_{n,t}^2 \right\rangle_{Q(\boldsymbol{J}, \beta)}, \tag{2.25}$$

13

where $<>$ indicates taking the expectation with respect to the posterior distributions of $\boldsymbol{J}$ and β. The left-hand side denotes the updated nth current's variance α_{np}^{-1}. The first term of the right-hand side means fMRI information, and the second term is current sources calculated from observed MEG data with T time points. When we set the prior weight close to 1, the resulting prior distribution tends to have peaks at the places where the fMRI information has a high value. This indicates strong confidence in the fMRI information in the estimation. On the other hand, if it is set to a small value, the result tends to emerge according to the MEG data. To represent the strongest confidence in the fMRI information, we set $w = 0.99999999$. Note that if the prior weight is 0, then we use only the fMRI information as the initial value of α_{np}^{-1}. Even if the prior information includes false positive activities, hVB estimation is able to suppress them by setting the proper value on the prior weight (Yoshioka et al., 2008).

Our first approach incorporates a meta-analysis fMRI data instead of an individual fMRI data using hVB estimation. The incorporated meta-analysis fMRI data is selected by experimenter beforehand. As discussed in Discussion chapter, it might bring a selection bias on results.

2.5 Relevant spatial prior estimation

To overcome the selection bias and express prior data as the combination of multiple meta-analysis data, we introduce relevant spatial prior (RSP) estimation. It is inspired by the formulation of multiple sparse priors (MSP; Friston et al., 2008) and dependent relevance determination (DRD; Wu et al., 2019). MSP is one of the source reconstruction methods that incorporate multiple prior information such as a depth structure, a smoothness between dipoles, and also the minimum

norm constraint by constructing the covariance of prior distribution as

$$\boldsymbol{A}^{-1} = w_1\boldsymbol{C}_1 + \ldots + w_p\boldsymbol{C}_P$$

$$= \sum_{p=1}^{P} w_p\boldsymbol{C}_p, \tag{2.26}$$

$$w_p = \exp(u_p), \tag{2.27}$$

$$u_p \sim \mathcal{N}(m, k^{-1}), \tag{2.28}$$

$$\boldsymbol{u} \sim \mathcal{N}(m\boldsymbol{1}, k\boldsymbol{I}), \tag{2.29}$$

where P is the total number of prior data, $\boldsymbol{C}_p \in \mathbb{R}^{N \times N}$ is the pth covariance matrix, w_p is the weight of the pth covariance, and u_p is the log of w_p. Each covariance $\boldsymbol{C}_p$ is determined using each prior knowledge (e.g., for MNE, $\boldsymbol{C}_p = \alpha^{-1}\boldsymbol{I}$) and the weighted sum of them is used as the covariance $\boldsymbol{A}^{-1}$. As above indicated, each u_p is independently follows Gaussian with parameters m and k and exponential of it is used as weight to restrict the weight to a positive value. MSP estimates posterior mean of currents and weights iteratively using restricted maximum likelihood estimation (Patterson and Thompson, 1971). As a result, weights of non-contributed priors are automatically pruned by the effect of ARD (Neal, 1996).

RSP estimation is an expansion of MSP. It can incorporate multiple spatial prior data using a structure of them. We defined that the covariance of RSP as

$$\boldsymbol{A}^{-1} = \text{diag}\left(\sum_{p=1}^{P} w_p\boldsymbol{d}_p\right)$$

$$= \text{diag}\left(\boldsymbol{D}\boldsymbol{w}\right), \tag{2.30}$$

$$w_p = \exp(u_p), \tag{2.31}$$

$$\boldsymbol{u} \sim \mathcal{N}(m\boldsymbol{1}, \boldsymbol{K}), \tag{2.32}$$

where $\boldsymbol{d}_p = (d_{p1}, \ldots, d_{pN})^T \in \mathbb{R}^N$ is a spatial data vector and $\boldsymbol{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_P) \in \mathbb{R}^{N \times P}$ is a whole dataset. Each data vector $\boldsymbol{d}_p$ is determined using spatial data such as fMRI data. Hence, the covariance matrix $\boldsymbol{A}^{-1}$ is composed of a weighted combination of spatial prior data. Furthermore, unlike MSP, RSP can incorporate the relevance structure between data such as similarities of spatial patterns into

15

$\boldsymbol{K}$. Because $\boldsymbol{K}$ is a covariance matrix of the Gaussian process prior Eq.(2.32), it can express the relationships between spatial data. Here, $\boldsymbol{K}$ is defined as

$$\boldsymbol{K} = \begin{pmatrix} k_{1,1} & \dots & k_{P,P} \\ \vdots & \ddots & \vdots \\ k_{P,1} & \dots & k_{P,P} \end{pmatrix}, \tag{2.33}$$

where

$$k_{i,j} = \rho \exp\left(-\frac{\text{dist}(\boldsymbol{d}_i, \boldsymbol{d}_j)}{2l^2}\right). \tag{2.34}$$

$k_{i,j}$ includes hyper parameter ρ and l, and also dist function. ρ and l are common to all $k_{i,j}$, and control amplitude of $k_{i,j}$ and dist, respectively. Function dist measures some distance between two data. It can be designed using any feature of the data. If we designed dist function as

$$\text{dist}(\boldsymbol{d}_i, \boldsymbol{d}_j) = \begin{cases} 0 & (i = j) \\ \infty & (i \neq j), \end{cases} \tag{2.35}$$

then $\boldsymbol{K}$ becomes a scalar matrix and RSP is equivalent to MSP. By designing the dist function properly, estimated weights of similar spatial data are kept together by the effect of DRD (Wu et al., 2019).

Our second approach incorporates a set of meta-analysis data and estimates the weights of them using RSP.

3. Materials and Methods

In this chapter, we introduce the key materials and methods used to implement our approaches (hVB approach and RSP approach).

3.1 Forward modeling using VBMEG

We constructed a brain model and a lead field matrix using the MATLAB opensource toolbox VBMEG (`https://vbmeg.atr.jp`). The individual brain model is constructed based on a subject ' s T1-MRI image. First, we construct a polygon model of the cortical surface using FreeSurfer (`http://surfer.nmr.mgh.harvard.edu/`). From the polygon model, we select 10,004 vertices as the current sources based on the predefined coordinate in the standard brain (MNI-ICBM152). As a result, source locations of different subjects are aligned to those of the standard brain while the sulcus-gyrus structure of each subject was maintained. This allows an easy comparison of the estimated source currents across subjects for each source. Therefore, we can proceed to group analyses on the source currents without any transformation. Next, we construct a 1-shell (cerebrospinal fluid) head conductivity model. Based on the model, we make a lead field matrix by solving the Maxwell equations with a boundary element method. For more information on preparing the brain model and lead field matrix using VBMEG, see Takeda et al., 2019.

3.2 Importing fMRI information

The statistical results of fMRI data were imported into a subject ' s brain model using VBMEG. Original statistical maps (t-value for individual fMRI data and z-value for meta-analysis fMRI) are defined on voxels, and thus they were transformed to the cortical surface using an inverse-distance weighted interpolation method. Imported fMRI information were used to calculate parameters in the probability distribution of the prior current variances as described in chapter 2.

3.3 Meta-analysis fMRI prior

A statistical map synthesized by Neurosynth (Yarkoni et al., 2011) (`https://neurosynth.org`) was imported to calculate the hierarchical prior using VBMEG. Neurosynth infers the regions associated with a term by mining thousands of papers and conducting a type of multilevel kernel density analysis (Wager et al., 2009). Therefore, the term must be selected by the experimenter when obtaining the meta-analysis results. In order to generate the effective meta-analysis data as prior information, the term must be related to the task of a MEG experiment. However, there are many choices in the term for a task. Therefore, for hVB approach, we selected the term from the highest layer of the ontology of each task domain (Thompson and Fransson, 2017). Namely, the term "visual" was chosen for the visual domain task and "motor" was chosen for the motor domain task (see discussion on our choice of terms). For RSP approach, we tried to select the related meta-analysis data from a set of data.

After generating the meta-analysis statistical maps, they were imported to the subject's brain model in the same way as the individual fMRI prior. The synthesized meta-analysis maps (Fig. 3.1) were employed as the hierarchical fMRI prior. Note, the results of hVB approach uses meta-analysis data synthesized using Neurosynth v0.4 and RSP approach uses Neurosynth v0.7.
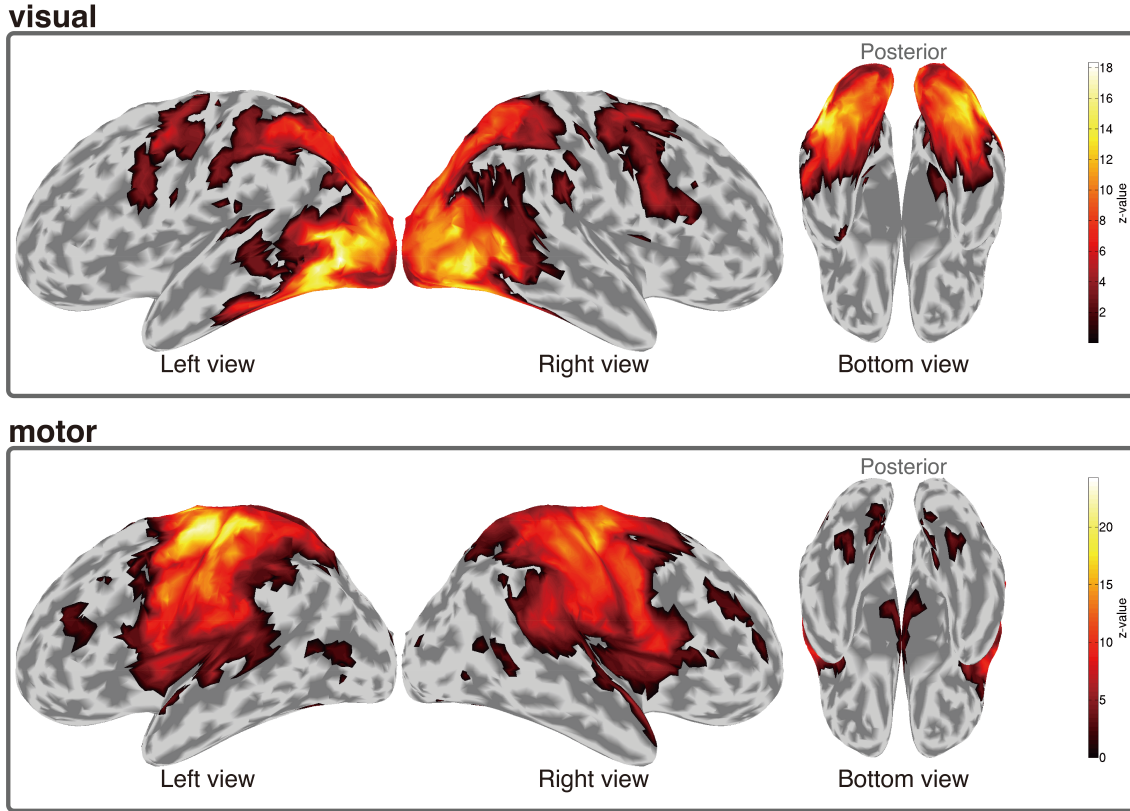
Figure 3.1. Meta-analysis priors synthesized by the terms "visual" and "motor." These are synthesized using Neurosynth (v0.4) and imported into VBMEG ' s coordinate, and they are used for the prior information of current source reconstruction. In hVB approach, the visual prior is used for the visual domain task, and the motor prior is used for the motor domain task.

3.4 Performance evaluation of simulated data

Inspired by Owen et al. (2012), we evaluated the performance of source reconstruction using the aggregate score of source-level spatial and temporal correlations, defined as

$$\text{AggregateScore} = \frac{\text{SpatialCorrelation} + \text{TemporalCorrelation}}{2}. \tag{3.1}$$

Since it is the average of Pearson ' s correlation coefficients, it takes a value from -1 to 1. Here, the correlation coefficient of spatial maps was calculated between

19

the simulated current map and the reconstructed one for each time point, and they were averaged along the time axis:

$$\text{SpatialCorrelation} = \frac{1}{T} \sum_t^T \text{corrcoef}(|\boldsymbol{j}_t^{\text{true}}|, |\boldsymbol{j}_t^{\text{recon}}|), \qquad (3.2)$$

where T is the number of time points (250 for hVB simulations and 50 for RSP simulations) and $|\boldsymbol{j}_t^{\text{true}}|$ and $|\boldsymbol{j}_t^{\text{recon}}|$ are the simulated current map and the reconstructed current map at time t, respectively.

On the other hand, to calculate temporal correlation, we matched each true source to the reconstructed source that has maximal Pearson's correlation within 10 mm. Here, we define that $\boldsymbol{j}_n^{\text{true}}$ and $\boldsymbol{j}_n^{\text{recon}}$ are true (simulated) and reconstructed current time series at source n. Since most of the true sources have no activation, we cannot simply calculate the averaged correlation coefficient among sources as spatial correlation Eq. (3.2). Therefore, we computed the temporal correlation using the following sequential source matching procedure:

1. Pearson's correlation coefficients between $\boldsymbol{j}_n^{\text{true}}$ and $\boldsymbol{j}_n^{\text{recon}}$ for all 10,004 sources were calculated. As a result, a correlation coefficient matrix ($10,004 \times 10,004$) was derived.

2. The Euclidean distance matrix between all 10,004 sources was calculated. Then, this matrix was thresholded at 10 mm. By applying this mask to the correlation coefficient matrix computed in procedure 1, the source pairs with over 10-mm distance were excluded.

3. The highest one of the remaining values was selected. Then, the selected true source and the estimated one were considered a pair. The rows and columns of the selected sources were removed from the masked correlation matrix.

4. Procedure 3 was iterated until all true sources that have a non-zero activity were paired.

After the matching procedure was completed, we took the average of correlations of selected pairs and defined it as the temporal correlation in Eq. (3.1). This procedure may result in true sources that do not have a matching target, and

these were excluded from the analysis. We note that even in the worst case, such sources accounted for only 0.004% of all sources and had almost no effect on the results.

4. Simulation and Experimental data for hVB

We explain the details of the simulation and experimental data to confirm the hVB approach.

4.1 Data simulation

To evaluate the feasibility of meta-analysis fMRI data as the prior information of current source reconstruction, we generated individual fMRI and MEG data with two task conditions and several difficulty settings. Both fMRI and MEG data were generated based on the same ground truth, which we call a common truth. This procedure aggregated various datasets of several modalities (structural MRI, fMRI, resting-state fMRI, diffusion MRI, and MEG sensor locations).

We generated two sets of data focusing on visual and motor experiments. The simulated dataset of the visual domain was made using multi-modal datasets of a face recognition task, which is publicly available from `https://openneuro.org/datasets/ds000117/versions/1.0.3/` (Wakeman and Henson, 2015). The simulated dataset of the motor domain was made using the left-foot movement task from the Human Connectome Project (Van Essen et al., 2012).

For all simulated data, we used the same brain model constructed from an individual's data (subject No. 2 of Wakeman & Henson, 2015) using VBMEG.

4.1.1 Common truth

First, we randomly selected five subjects from both datasets (subject No. 2, 6, 9, 15, 16 for the visual domain, and No. 115724, 144933, 481951, 896879, 902242 for the motor domain). FMRI preprocessing, including realignment, slice timing correction, normalization, and smoothing, were conducted for each subject using the same parameters as those in the original study (we referred to Wakeman and Henson (2015) for visual data and Van Essen et al. (2012) for motor data). Then, task conditions against a baseline were contrasted. For the visual domain dataset, we employed the face image recognition task (famous condition). For the motor domain dataset, we employed the left-foot movement task (LFoot condition). Then, we acquired one contrasted statistical map for each subject of both tasks.

Next, group analysis was conducted using contrasted statistical maps of five subjects for each task. Finally, the resulted statistical maps of group analysis were thresholded with three levels of significance: p < 0.001, 0.01, or 0.05. These thresholded maps were used as the common truth for both fMRI and MEG data generation. Thresholding controlled the density of the common truth map and modulated the cancellation index of the MEG data (Ahlfors et al., 2010). These procedures are presented in the middle part of Fig. 4.1. To prevent the simulated data from being highly dependent on the results of an individual, the common truth was built based on the results of group analysis. All of the preprocesses and statistical tests were performed using spm8 (`https://www.fil.ion.ucl.ac.uk/spm/software/spm8/`).

4.1.2 FMRI data generation

We generated fMRI data with various signal qualities based on the common truth by adding the resting-state brain activity as background noise (upper part of Fig. 4.1). We designed a virtual fMRI task design. It was an event-related paradigm. A stimulus was designed to be presented for 2 sec with 20 sec intervals. We set the total number of fMRI scans to 240, with 2 sec of repetition time. Then the designed time series were convolved with the canonical hemodynamic response function (HRF) to obtain the time series of task-related brain activities. The amplitude of the time series on all voxels was determined by weighting with the common truth map, resulting in simulated fMRI data during the designed task. The fMRI time series was further contaminated by adding the experimental resting-state fMRI activity (see Ogawa, Aihara, Shimokawa, & Yamashita, 2018, for data acquisition) as noise to take temporal modulations of each brain region into account (Valente et al., 2009). We varied the amount of noise so that the contrast-to-noise ratio (CNR) ranged from 0.5 to 10. CNR is defined as the ratio between the task-related signal amplitude A and a standard deviation of the task-irrelevant noise signal σ (Baumgartner et al., 2000; Welvaert and Rosseel, 2013):

$$\text{CNR} = \frac{A}{\sigma}. \tag{4.1}$$

Finally, general linear model (GLM) analysis was conducted on the generated task fMRI time series data using SPM8 (Friston et al., 1994), resulting in statistical

maps of various CNR levels, which served as individual fMRI priors for current estimation. Examples of the common truth and simulated individual fMRI priors for CNR=0.5, 1, 3, 5, and 10 are shown in Fig. 4.2. The statistical maps in the right panel gradually approach the common truth map as CNR increases. Using these CNR-controlled priors, we investigated the effects of the quality of the individual fMRI data on the current reconstruction that were comparable to the meta-analysis fMRI data.

In Fig. C.1 and Fig. C.2, we report the distribution of CNR for all simulation settings. These were calculated using coefficients and residuals of GLM, as CNR = coeff.(condition) / mean square of residuals × scaling factor. The scaling factor is the maximum value of the task design time series (after convolution with HRF). As these figures indicate, the maximum CNRs of each setting were configured to the desired value. To be compatible with more research, we also show the percent signal change (PSC), which is a popularly used quality measure (Pernet, 2014; Welvaert and Rosseel, 2013). It was also calculated using parameters of GLM as PSC = coeff.(condition) / coeff.(constant) × scaling factor × 100.

4.1.3 MEG data generation

The locations and amplitudes of dipoles were defined based on the common truth spatial map imported into VBMEG's coordinate. Because it is difficult to design the temporal evolution among distributed dipoles by hand, we designed it using the ROI-level whole-brain current simulation. The simulation was conducted using the connectome obtained from diffusion MRI data (it was averaged over 13 subjects, see Endo, Hiroe, & Yamashita, 2020, for data acquisition), and the Larter-Breakspear neural mass model (Breakspear et al., 2003; Larter et al., 1999) with the AAL parcellation (details of the ROI-level whole-brain current simulation are following Endo, Hiroe, & Yamashita, 2020). To obtain vertex-level whole-brain current time series, we assumed the same waveforms for vertices belonging to a single ROI. Then by weighting the whole-brain waveforms with the common truth of the task (visual or motor), we generated the current source time series of the desired task. Observed MEG time series were calculated by multiplying the lead field matrix and adding Gaussian noises. The length of a single trial was set to 250 msec. These procedures are presented in the lower part

of Fig. 4.1.

The quality of MEG data was defined as the signal-to-noise ratio (SNR) of trial-averaged data, and it was controlled by varying the number of trials. Here, SNR of MEG data is defined as

$$\text{SNR} = 10\log_{10}\frac{\sigma_s^2}{\sigma_n^2}, \tag{4.2}$$

where σ_s^2 is a variance of signals during the task and σ_n^2 is a variance of noises. We configured the SNR of single-trial MEG data to be 0 [dB]. Consequently, the SNR of trial-averaged data became 0, 5, and 10 [dB] when the number of trials was 1, 3, and 10, respectively.

It is well known that current sources at nearby vertices with opposite dipole orientations cancel each other, resulting in nearly zero MEG signals. Consequently, it is assumed that the more vertices that are activated coherently, the more difficult source reconstruction is. To characterize such a difficulty of the source reconstruction, we computed the weighted cancellation index (wCI), which is an extended version of a previously proposed cancellation index (Ahlfors et al., 2010) :

$$\text{wCI} = 1 - \frac{\left[\sum_m^M (\sum_n^N g_{mn} j_n)^2\right]^{1/2}}{\sum_n^N \left[\sum_m^M (g_{mn} j_n)^2\right]^{1/2}}, \tag{4.3}$$

where g_{mn} is an element of a lead field matrix $\boldsymbol{G}$ and j_n is a current at source n. For the simulated current time series, we calculated the wCI at each time point and averaged them over time. The wCI of the simulated data obtained from the thresholding of common truth with p < 0.001, 0.01, and 0.05 was 0.77, 0.88, and 0.90 for the visual task and 0.40, 0.71, and 0.83 for the motor task, respectively.

To summarize the above, we generated MEG data by varying three SNR levels and three wCI levels; thus, in total, there were nine simulation settings for each task.

4.1.4 Current source reconstruction methods for comparison

In the simulation studies, we compared the source currents estimated by hVB with different priors and also other source imaging methods. Estimation using hVB was conducted with the meta-analysis priors (synthesized with the terms

"visual" and "motor"), the simulated individual fMRI priors (with CNR=0.5, 1, 3, 5, 10), the common truth prior, and without a prior.

When we used the meta-analysis prior, we basically selected the relevant prior for the task ("visual" prior for visual domain task and "motor" prior for motor domain task). However, we also used irrelevant priors for the tasks, namely "visual" prior for motor domain task and "motor" prior for visual domain task.

The individual fMRI priors obtained from different CNRs were considered. We also considered the common truth as the prior information of the ideal case. When there is no prior information, we can only use the uniform spatial pattern. This means we set the same values on all initial values of the hierarchical prior α_{n0}^{-1} in Eq.(2.20). In this thesis, we call this uniform spatial pattern as "uniform prior" of hVB estimation. If the uniform prior was used with a small prior weight, the estimated current variances tend to be sparse due to the effect of ARD (Neal, 1996).

To evaluate the significance of fMRI information, we also applied weighted MNE (wMNE) and Champagne as benchmark methods. Both wMNE and Champagne were run using the functions from Brainstorm software (Tadel et al., 2011; `http://neuroimage.usc.edu/brainstorm`) and NUTMEG software (Dalal et al., 2004; `http://nitrc.org/projects/nutmeg/`), respectively.

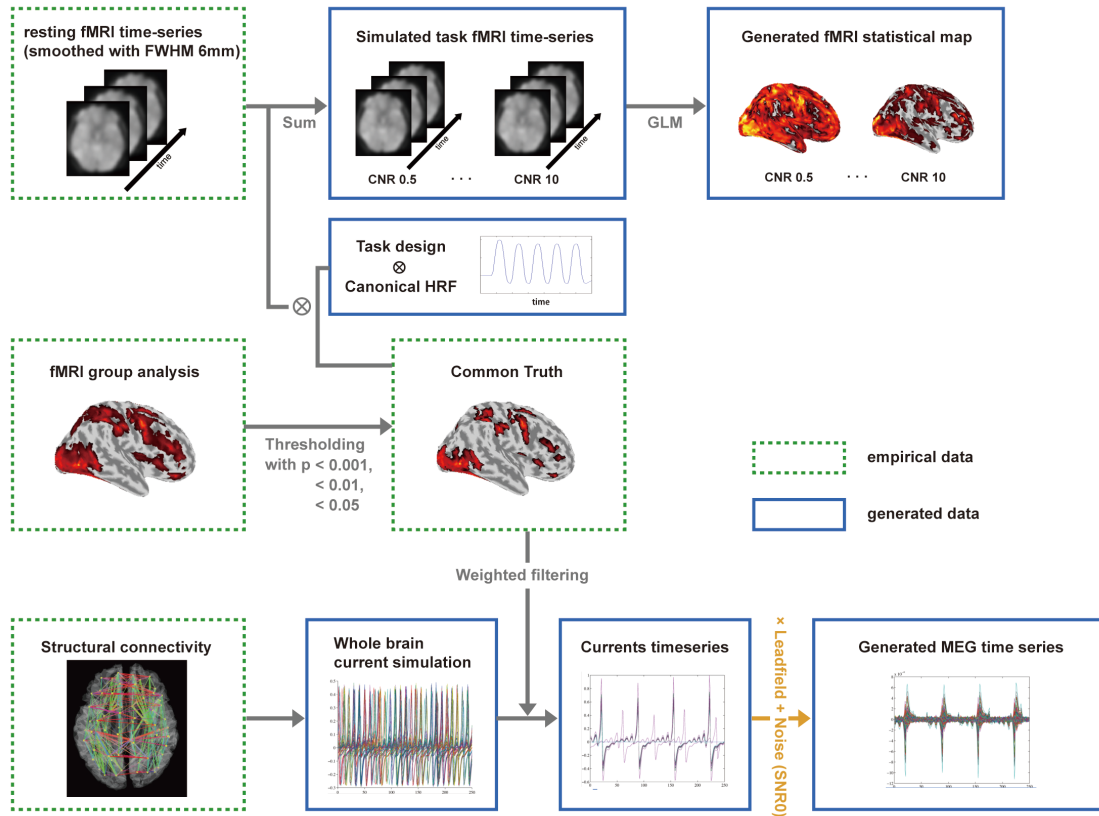We applied spatial smoothness filtering for all the above methods (see Appendix B).

Figure 4.1. FMRI (upper part) and MEG (lower part) data-generation process from the common truth (middle part). Simulated individual fMRI priors are taken as the hierarchical prior of hVB estimation for simulated MEG data. Both modalities are generated based on the common truth map. Density of the common truth is controlled by thresholding the p-value ($p < 0.001$, 0.01, or 0.05). It affects the density of the generated fMRI statistical map as well as the cancellation of MEG data. The SNR of single-trial MEG data is fixed at 0 [dB], and total SNR is controlled by the number of trials (only a single trial is shown in the figure).
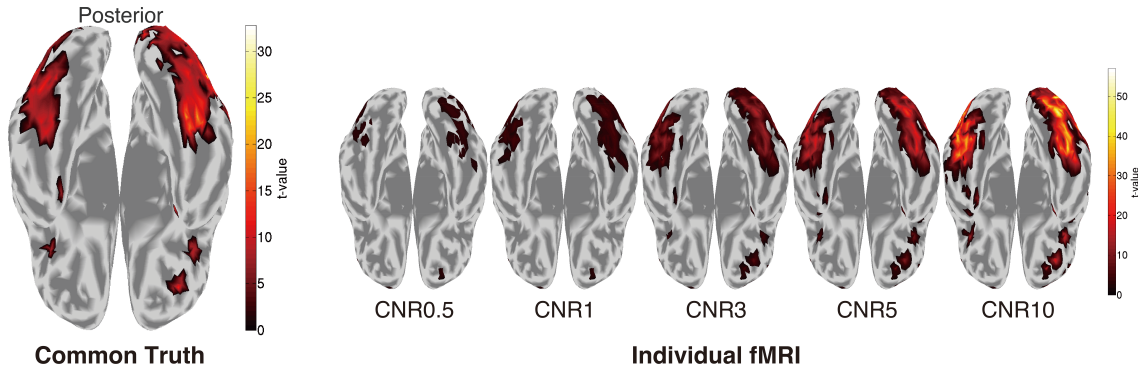
Figure 4.2. Examples of the common truth map and simulated individual fMRI priors. Here, the common truth is computed for the face image recognition task with thresholding p < 0.01. Individual fMRI data are generated from the common truth with controlled contamination. As the individual fMRI images show, statistical maps gradually clarify the common truth map with increasing CNR. They are imaged from the bottom view to observe the change in statistical maps in the fusiform face area.

4.2 Experimental data

We analyzed the multi-subject, multi-modal neuroimaging dataset for face processing recorded by Wakeman and Henson (2015). It was the same dataset that was used to generate simulated data of the visual domain. This dataset contains the evoked responses of 16 subjects to three types of face stimuli: famous, unfamiliar, and scrambled. MEG, EEG, electro-oculograms, and electro-cardiograms were simultaneously recorded at 1,100 Hz with an Elekta Neuromag Vectorview 306 system (Helsinki). T1 images and fMRIs were also collected with a Siemens 3T TIM TRIO (Siemens, Erlangen, Germany). These data are stored in the Brain Imaging Data Structure (BIDS) format (http://bids.neuroimaging.io/). We preprocessed MEG data in the same way as done previously (Takeda et al., 2019). Schematic descriptions are visualized in the left box of Fig. 4.3. As a result, the face (famous and unfamiliar) and the scrambled conditions were compiled.

We constructed a t-map by contrasting all stimulus conditions (face, scrambled) against the baseline using SPM8. The t-map was computed for each subject

and imported as individual fMRI priors (Fig. 4.4). On the other hand, the meta-analysis prior was synthesized using Neurosynth with the term "visual," and it was common for all subjects (Fig. 3.1). We carried out source reconstruction using two types of priors, an individual fMRI prior and the meta-analysis prior, for each subject. The prior weight parameter was set to 0.3 for both priors.

Using all the subjects' source currents, we conducted a group analysis and examined the statistical differences of the current amplitudes between the face and scrambled conditions. The method of group analysis followed our previous work (Takeda et al., 2019). Briefly, we calculated the stimulus-triggered average of the estimated source currents for each subject and condition. Then for each source and time, we compared the 16 subjects' current amplitudes between the face and scrambled conditions using a paired t-test. Finally, the multiple comparison problem was solved using Storey & Tibshirani (2003)'s method (false discovery rates were controlled at 0.05). All of the procedures were computed separately for results with an individual fMRI prior or the meta-analysis prior. The graphical description of whole group analysis procedure is visualized in Fig. 4.3.
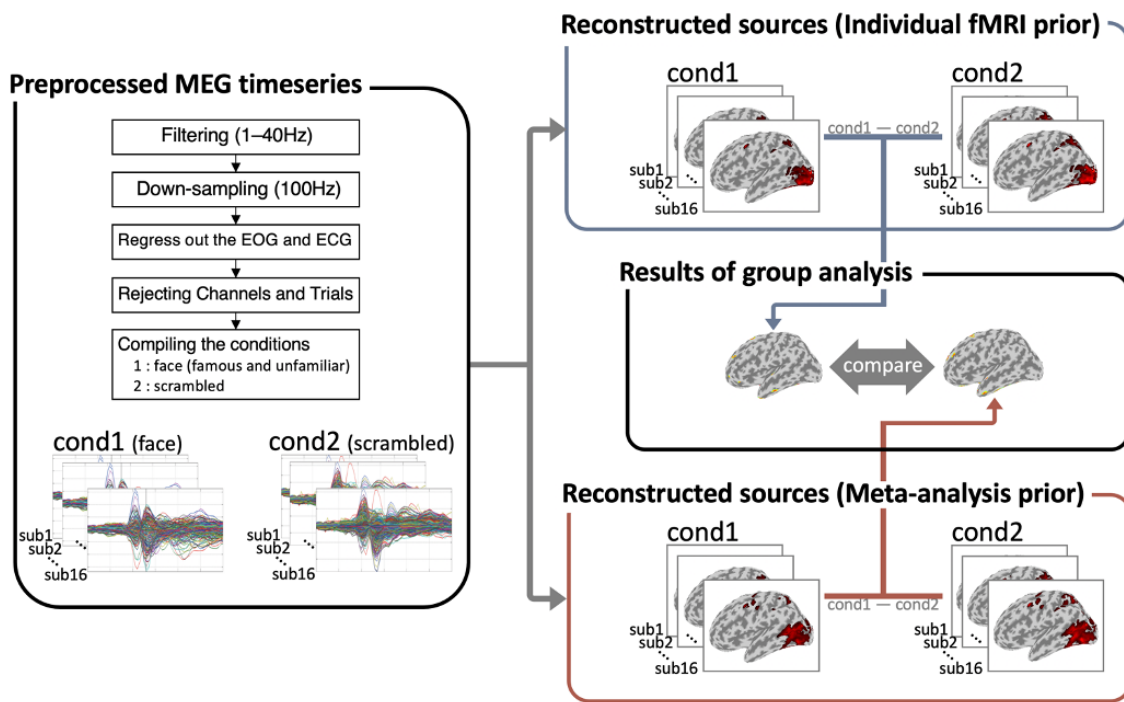
Figure 4.3. Graphical explanation of the group analysis process. The preprocess of MEG data (in the left box) and the flow of the group analysis are displayed.
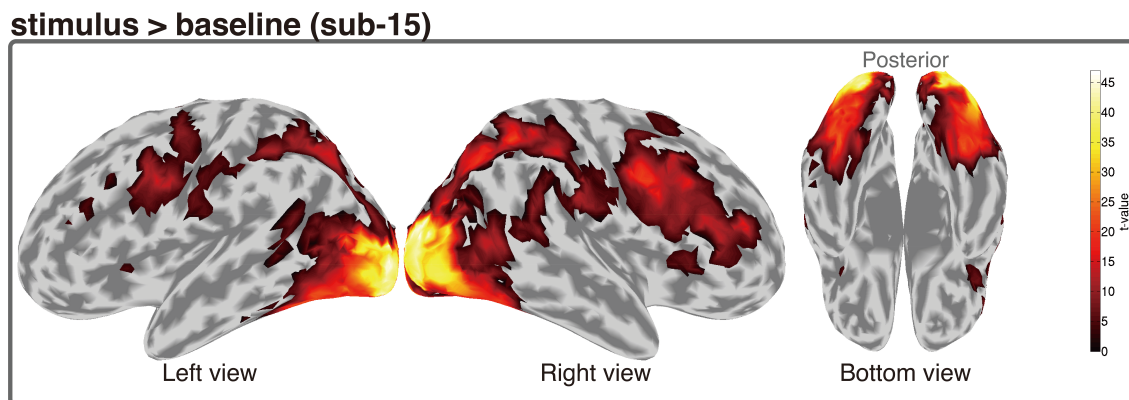
stimulus > baseline (sub-15)

Figure 4.4. Imported individual fMRI prior of sub-15. This is computed by contrasting all stimulus conditions (face (famous and unfamiliar), scrambled) against a baseline.

5. Results for hVB approach

5.1 Simulated data

We first studied the effects of the prior weight parameter for both task domains (visual and motor). We then evaluated how effective the meta-analysis prior was compared to the individual fMRI priors using the appropriate prior weight. Finally, data-driven model selection was attempted using a statistical criterion.

5.1.1 Effects of prior weight parameter

We first studied the effects of the prior weight parameter on reconstructed sources. We repeatedly conducted current source reconstruction using the hVB method by varying the prior weight from 0 to 0.99999999 (denoted as 1). The full results with all simulation settings are reported in Fig. 5.1 for the visual task and Fig. 5.2 for the motor task. For comparison, we also plot the results with the worst (CNR=0.5) and the best (CNR=10) individual fMRI priors. In addition, the results using the common truth prior are reported.

For both tasks, the scores of the meta-analysis fMRI prior were very consistently located between the simulated individual fMRI priors when the prior weight and the term ("visual" or "motor") were properly selected. For example, for the results of $p < 0.05$ in the visual task (top 3 panels of Fig. ??), when the meta-analysis prior "visual" was used with a prior weight 0.3, the aggregate scores were consistently higher than the results of CNR0.5 and lower or comparable to CNR10. We can find similar results for all conditions when using the relevant prior ("visual" for the visual and "motor" for the motor task) and the optimal prior weight settings. The appropriate range of the prior weight tends to be high when the cancellation is high ($p < 0.05$, with wCI of 0.90 for the visual task and 0.83 for the motor task). Conversely, when the cancellation is low ($p < 0.001$, wCI of 0.77 for the visual task and 0.40 for the motor task), the appropriate range of the prior weight is low as well. These results indicate that estimations require strong prior information when the cancellation is severe. This is because strong cancellation induces very few observation signals.
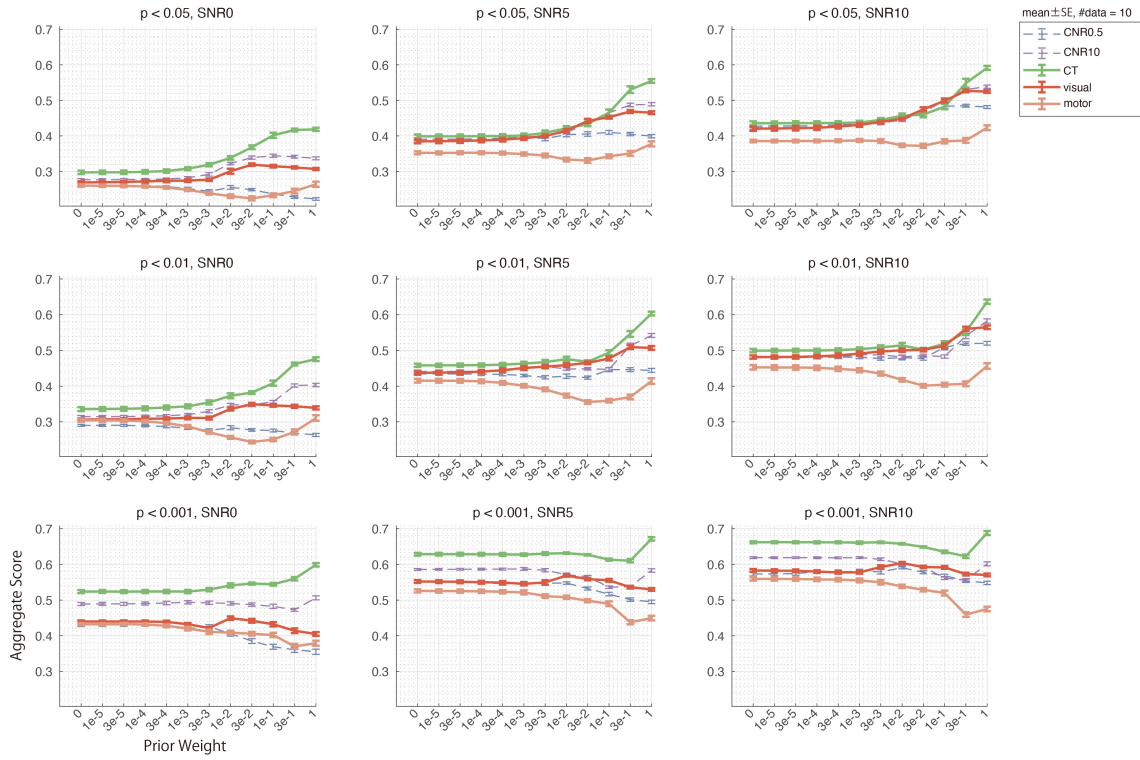
Figure 5.1. Scores for the visual task with various prior weight parameters. Results of the hVB estimation with the worst and the best individual fMRI priors (CNR=0.5 and 10), the common truth prior, and the meta-analysis priors (synthesized using terms "visual" and "motor") are shown. Vertical axes indicate the aggregate score (upper is better), and horizontal axes indicate the value of the prior weight parameter. The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of all simulation settings (3 SNR by 3 wCI (denoted as thresholds of the common truth)).
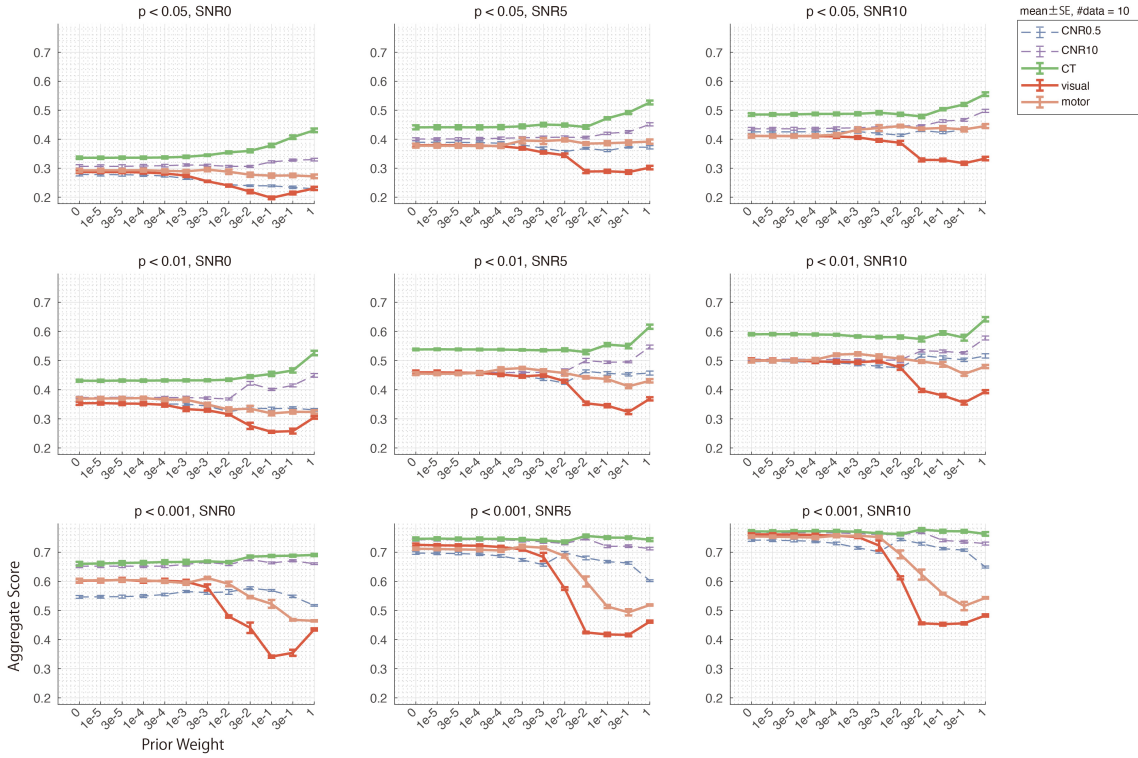
32

Figure 5.2. Scores for the motor task with various prior weight parameters. Results of the hVB estimation with the worst and the best individual fMRI priors (CNR=0.5 and 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are shown. Vertical axes indicate the aggregate score (upper is better) and horizontal axes indicate the value of the prior weight parameter. The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of all simulation settings (3 SNR by 3 wCI (denoted as thresholds of the common truth)).

5.1.2 Comparison between meta-analysis prior and other priors

We conducted a close comparison between the meta-analysis prior and the individual fMRI prior with the properly fixed prior weight parameters. We focused on one simulation setting (characterized by SNR and wCI), which had a level of difficulty similar to the experimental data (face recognition task data (Wakeman

and Henson, 2015)). Since we do not know the actual values of the SNR (Eq.(4.2)) and the wCI (Eq.(4.3)) for the experimental data, we approximated those values as follows. A variance of signals during the task σ_s^2 and a variance of noises σ_n^2 were calculated as the variance of trial-averaged signals and the variance of signals before task onset, respectively. The wCI was calculated using the t-map of task fMRI data of face−baseline contrast. Hence, j_n was substituted with the t-value. The calculated SNR was 4.88 [dB], and wCI was 0.95. Therefore, we selected the simulation setting with SNR = 5 [dB] and the threshold of the common truth as p < 0.05, corresponding to wCI = 0.90 (face image recognition task) and 0.83 (left-foot movement task).

The results are presented in Fig. 5.3. As a control analysis, we also compared the results of the meta-analysis prior with those of the uniform prior and the benchmark methods, wMNE and Champagne. The MNE and Champagne methods used their default hyperparameters, although the hVB methods used the best prior weight parameter found through investigation using a grid search. In this search, hVB was conducted using all prior weight values as Figs. 5.1 and 5.2, and then assessed using the aggregate score to select the best prior weight. Reconstructed current maps of the visual domain task are also shown in Fig. 5.1.2. We can see that the sparser the estimated source, the higher the amplitude is recovered.

For the visual domain task, the result of the relevant meta-analysis prior ("visual") shows a better score than that for the individual fMRI prior with CNR=1. For the motor domain task, the result of the relevant meta-analysis prior ("motor") shows a better score than that for the individual fMRI prior with CNR=0.5. On the other hand, such an individual fMRI prior with low-CNR was better than the benchmark methods. However, if we use irrelevant meta-analysis data, the performance is degraded. Therefore, meta-analysis fMRI data synthesized using the relevant term has comparable efficiency with the individual fMRI prior with CNR 0.5–1.

We also reported results assessed using spatial correlation (Fig. E.1) and temporal correlation (Fig. E.2) for the same simulation setting as that in Fig. 5.3. Although the results assessed using temporal correlation show little difference between priors, spatial correlation shows clearer differences. Hence, the differences

in prior data mainly affect the estimated spatial maps rather than the time series.

To show the results under various conditions, the aggregate scores for all nine simulation settings for the visual and motor tasks are shown in Fig. E.3 and E.4, respectively. For both visual and motor tasks, the hVB results using the appropriate meta-analysis fMRI prior were comparable or better than the scores of low-CNR fMRI priors. Counterintuitively, when $p < 0.001$ for the motor task, the results of the meta-analysis fMRI data of "visual" are comparable to those of "motor." This is due to the fact that the results of the fMRI group analysis used to generate the simulated data had activation on vertices that can be covered by both "visual" and "motor" meta-analysis data when $p < 0.001$.

Additionally, the results of the Champagne method show a comparable score to hVB using individual fMRI data when the simulated current map has a sparse setting ($p < 0.001$). In particular, this method shows consistent scores regardless of SNR. These characteristics of the Champagne method seem to be caused by the effect of sparse estimation (Wipf et al., 2010). These results indicate that the fMRI data are particularly helpful when sources are densely activated.

Moreover, we also conducted current source reconstruction with the meta-analysis fMRI prior synthesized using more detailed terms such as "face" and "object recognition" for the visual domain task and "hand" and "foot" for the motor domain task. Results are shown in Fig. E.5 and E.6. For comparison, the scores of "visual" and "motor" were plotted (the same as in Fig. 5.3). In these cases, "face" and "foot" seem to be suitable for the visual (face image recognition) and the motor (left-foot movement) tasks, respectively. Actually, only the results with the "foot" term improved the performance, whereas the result of the "face" term was comparable to "visual."
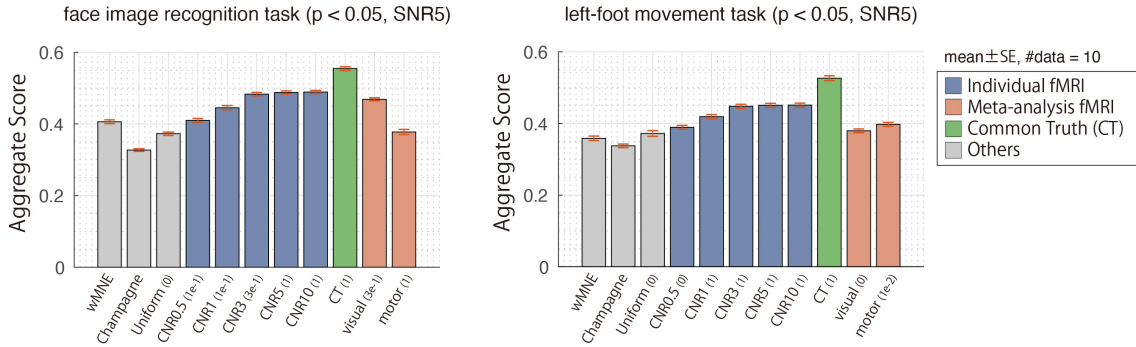
Figure 5.3. Comparison of aggregate scores with various methods for the visual task (left) and the motor task (right). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are reported. These priors are used with the best prior weight parameter denoted in parentheses. Vertical axes indicate the aggregate score (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of a realistic simulation setting (p<0.05, SNR5) determined using experimental data. The results of wMNE and Champagne are also shown as a control.
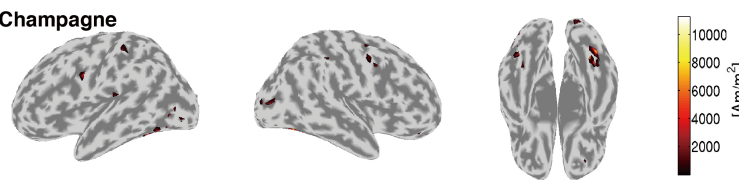
Ground truth current map

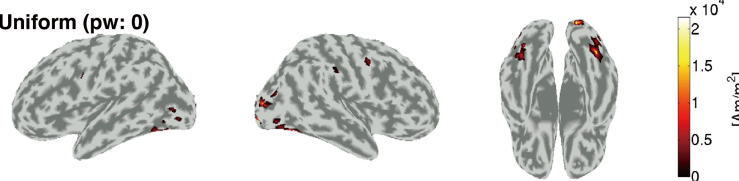Left view

Right view

Posterior
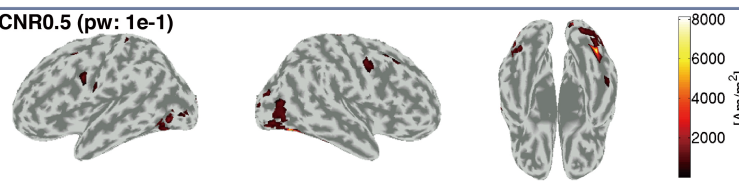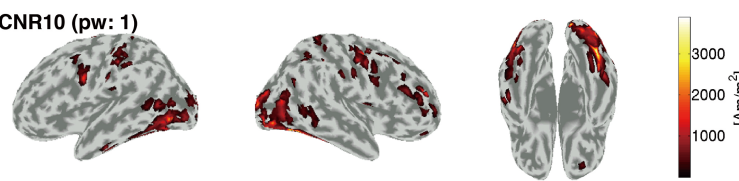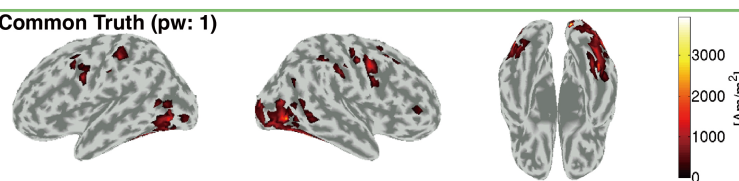L   R

Bottom view
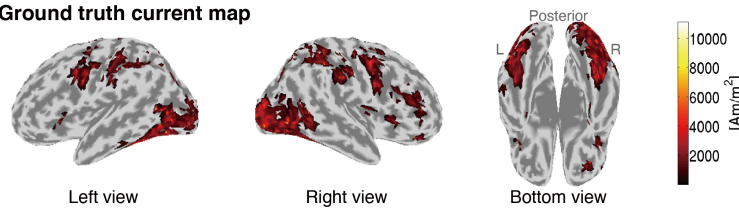
wMNE
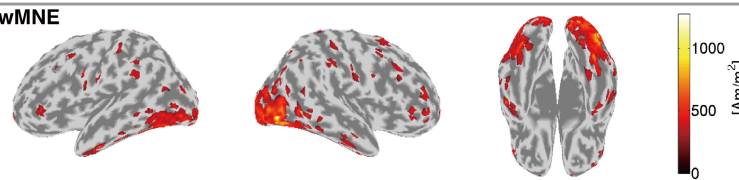
Champagne

Uniform (pw: 0)

CNR0.5 (pw: 1e-1)

CNR10 (pw: 1)

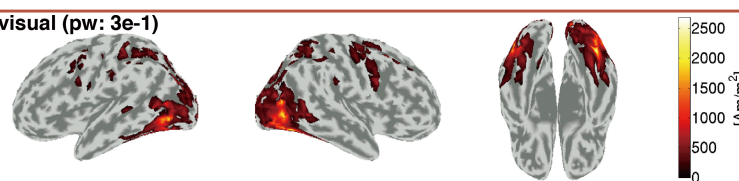Common Truth (pw: 1)

visual (pw: 3e-1)

Figure 5.4. Reconstructed current maps of the visual domain task with various methods. The ground truth current map (top) and reconstructed maps are shown. To visualize the maps, amplitudes are averaged along the time axis. The prior weights are set to the same values as those in Fig. 5.3. For all figures, activities over 10 % of their maximum value are displayed.

5.1.3 Model selection using free energy

In the neuroscience literature, free energy is one of the popularly used statistical criteria for model selection (Friston et al., 2003; Fukushima et al., 2015; Wipf and Nagarajan, 2009). In particular, free energy is used to select prior information of source reconstruction in the multiple sparse priors method (Friston et al., 2008; Henson et al., 2010). Inspired by these studies, we attempted to select hyperparameters (prior weight and meta-analysis fMRI data) for hVB estimation using free energy (see Appendix A for the details of free energy in hVB approximation). First, we studied whether the prior weight parameter could be selected using free energy. The effects of the prior weight on free energy and the aggregate score are reported in Fig. 5.5 As the results indicate, free energies were negatively correlated with prior weights, but they were not related to aggregated scores. This tendency was shown for other prior data as well. Thus, the prior weight parameter could not be selected based on free energy.
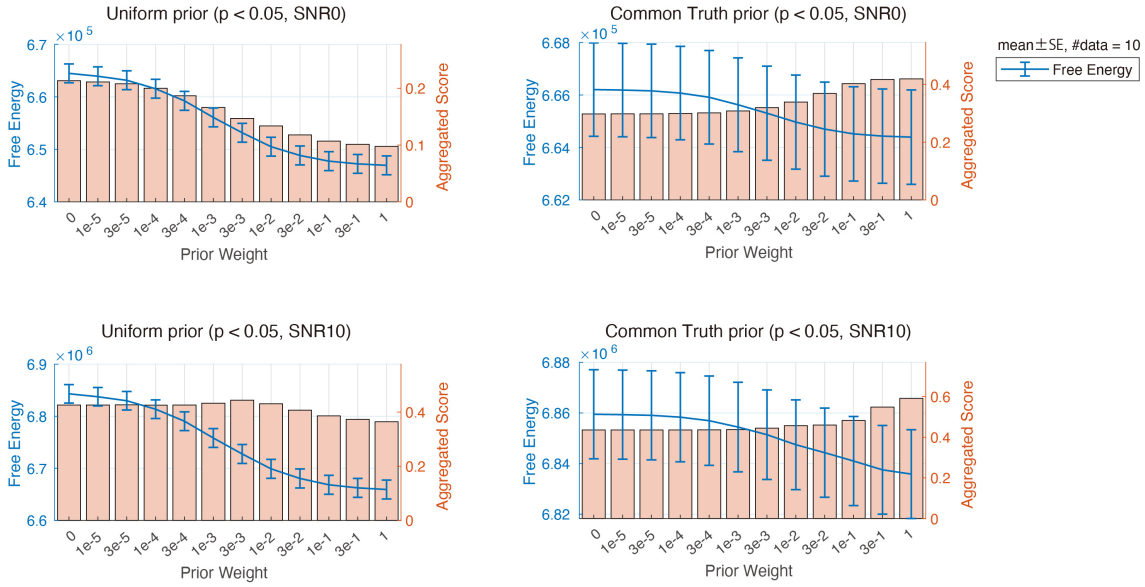
Figure 5.5. Free energy comparison of various prior weight parameters for simulated data. We display results of the uniform (left) and common truth (right) prior for the simulated visual task, with $p < 0.05$, and the best and worst SNR settings. Free energies are plotted as blue lines and aggregate scores are shown as bar plots. Left and right vertical axes indicate the free energy and the aggregate score, respectively. Horizontal axes indicate the value of the prior weight parameter. The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. Note that a similar tendency is observed for the simulated motor task data.

Next, we considered the meta-analysis fMRI data selection. Here, the prior weight was pre-fixed to 0.3 for all prior data. For both visual and motor tasks, we showed the results with SNR0 and 10, respectively (Fig. 5.6 and 5.7). These figures indicate strong relationships between free energies and aggregated scores regardless of SNR. These results support the potential of selecting the prior data using free energy.

39

Figure 5.6. Free energy comparison of various priors for simulated data (face image recognition task). Free energies are plotted as blue lines and aggregate scores are shown as bar plots. Left and right vertical axes indicate the free energy and the aggregate score, respectively. The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. All results are derived using prior weight 0.3. We display results of the face image recognition task thresholded with p < 0.05. Left panel is for SNR = 0 and right panel is for SNR=10.
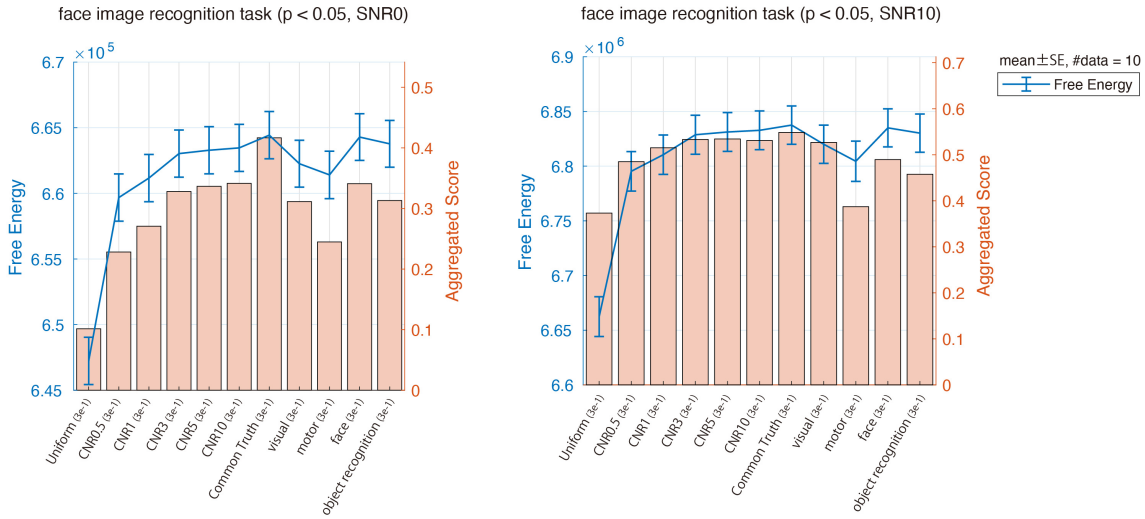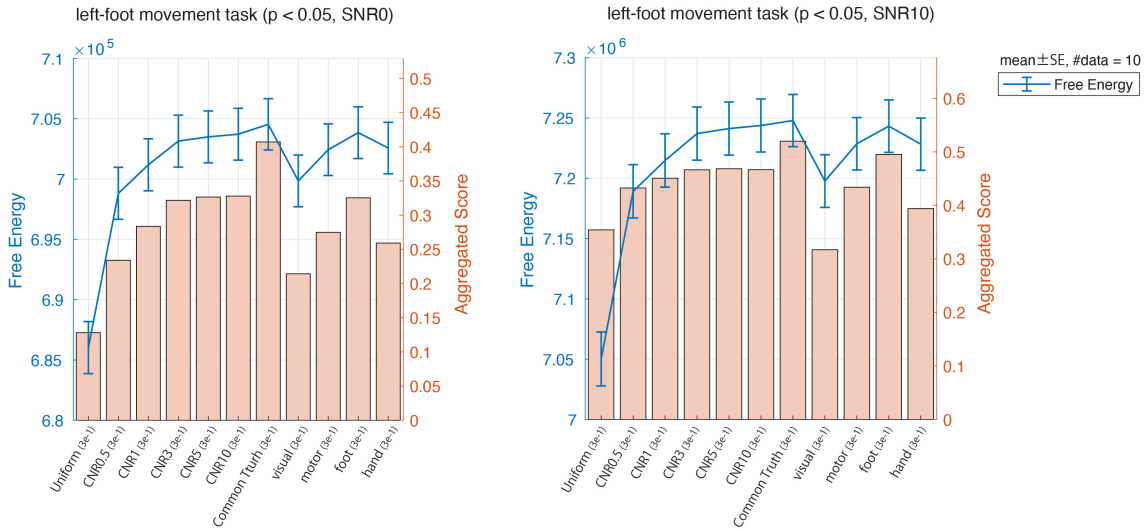
Figure 5.7. Free energy comparison of various priors for simulated data (left-foot movement task). Free energies are plotted as blue lines and aggregate scores are shown as bar plots. Left and right vertical axes indicate the free energy and the aggregate score, respectively. The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. All results are derived using prior weight 0.3. We display results of the left-foot movement task thresholded with $p < 0.05$. Left panel is for $SNR = 0$ and right panel is for $SNR=10$.

5.2 Experimental data

Since we do not know the ground-truth of current sources for real experimental data, we evaluated the reproducibility of the statistical maps obtained from group analyses. We compared the individual fMRI prior with the meta-analysis fMRI prior ("visual"). The prior weight parameters for the individual fMRI prior and the meta-analysis fMRI prior were each set to 0.3, following our previous work (Takeda et al., 2019) and the above simulation results (according to Fig. 5.3, which is the simulated face image recognition task with the realistic setting, 0.3 was the best prior weight parameter). Finally, we attempted data-driven model selection as well as simulated data.

5.2.1 Group analysis

First, we checked the currents reconstructed using the individual fMRI prior and the meta-analysis prior. The reconstructed currents time series and maps of one example subject are shown in Fig. 5.8 and Fig. 5.9. These reconstructions were run for all subjects in the same way. Next, we conducted the group analysis of reconstructed current sources by contrasting the face condition to the scrambled condition. Fig. 5.10 shows the number of sources exhibiting significant differences between the face and scrambled conditions along the time axis. Although the largest difference was observed at 0.17 sec for both priors, the detected source of the meta-analysis prior was slightly smaller than the individual fMRI prior. The time series of detected sources were quite similar, although we found a slight difference at the second peak.

In Fig. 5.11, we compared the significant t-value maps of both priors at 0.17 sec, and they looked very similar to each other. They show significance at the right Fusiform Face Area (FFA) and right insular cortex. The significant differences at the right FFA are consistent with previous studies that reported that this area exhibits face-selective responses (Grill-Spector et al., 2017, 2004; Jas et al., 2018; Rossion et al., 2018; Wakeman and Henson, 2015). The statistical map obtained using the meta-analysis prior shows a significant difference on the left fusiform area and the left insula, which are not observed in that obtained from the individual fMRI prior.

In Appendix F, we considered the effect of signal leakage using the resolution kernel analysis. This showed that reconstructed currents in the insular cortex can be leaked from the fusiform gyrus.
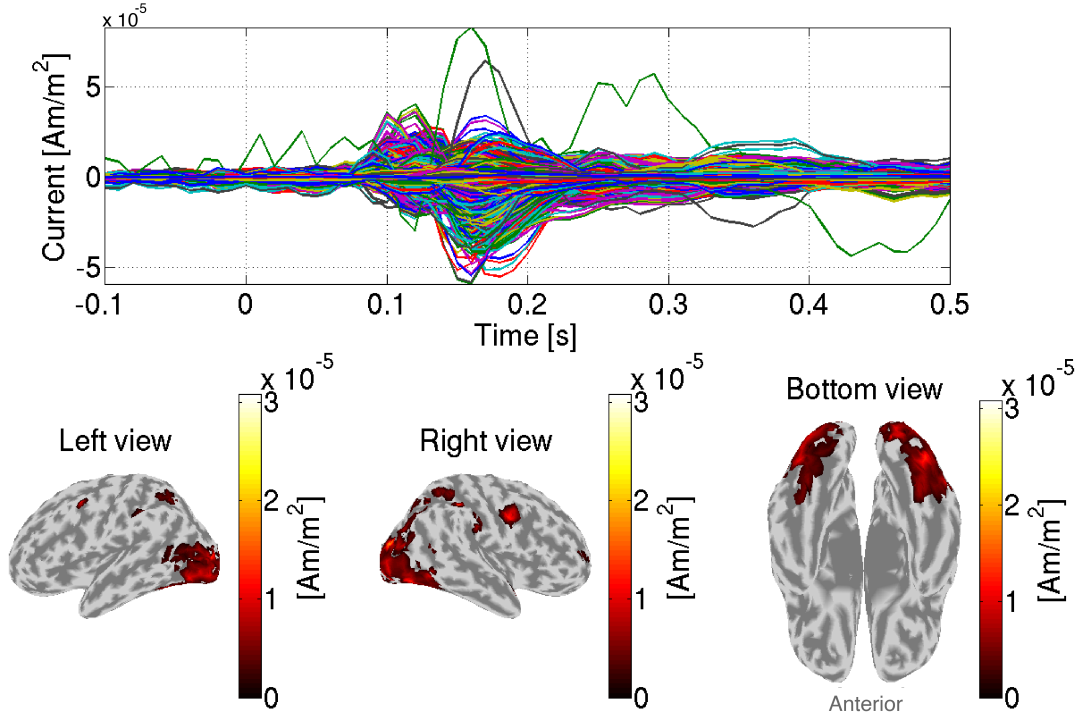
Figure 5.8. Stimulus-triggered average of source currents estimated in face condition (sub-15) using individual fMRI prior. The prior weight is set to 0.3. Its time series (top) and amplitudes averaged within 0–0.3 sec (bottom) are shown. In the bottom figures, activities over 30% of their maximum value are displayed.

Stimulus–triggered average of source currents estimated using meta-analysis fMRI prior (sub-15)
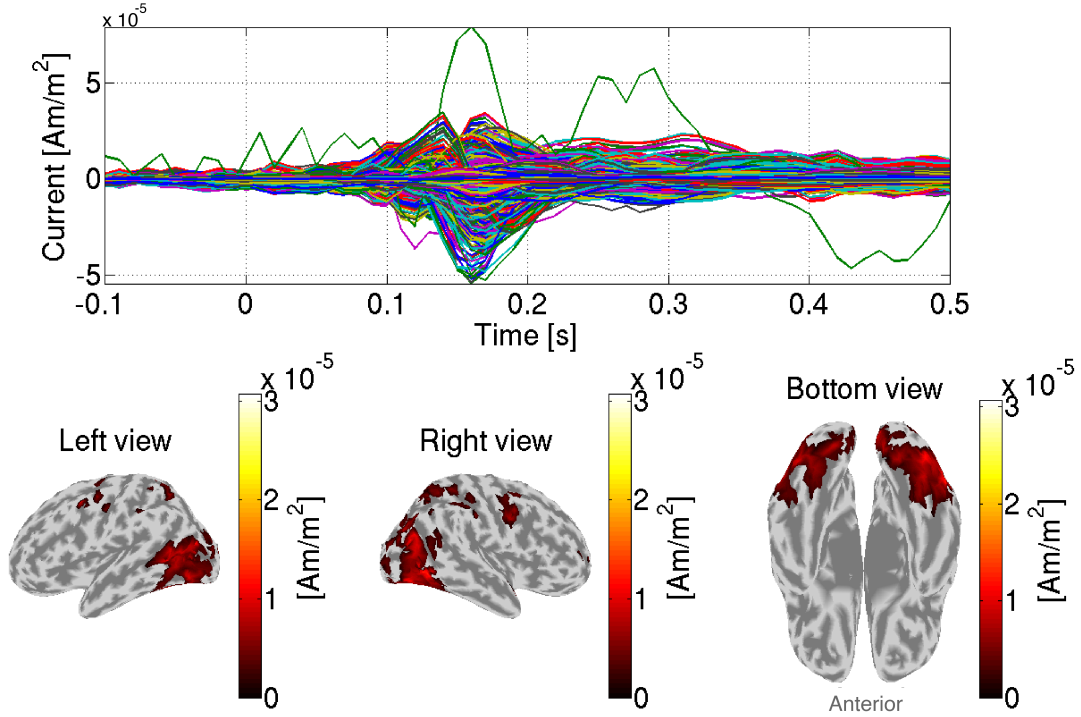
Figure 5.9. Stimulus-triggered average of source currents estimated in face condition (sub-15) using meta-analysis fMRI prior. The prior weight is set to 0.3. Its time series (top) and amplitudes averaged within 0–0.3 sec (bottom) are shown. In the bottom figures, activities over 30% of their maximum value are displayed.
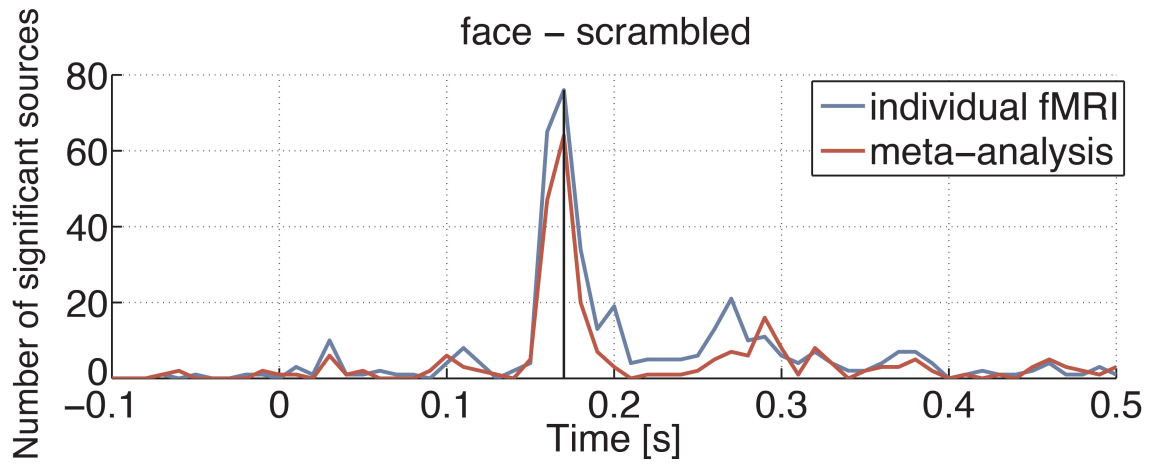
Figure 5.10. Differences in current amplitudes between face and scrambled conditions along the time axis. The number of sources exhibiting significant differences (q < 0.05) are shown for both individual fMRI and meta-analysis fMRI priors.

Significance map with individual fMRI prior
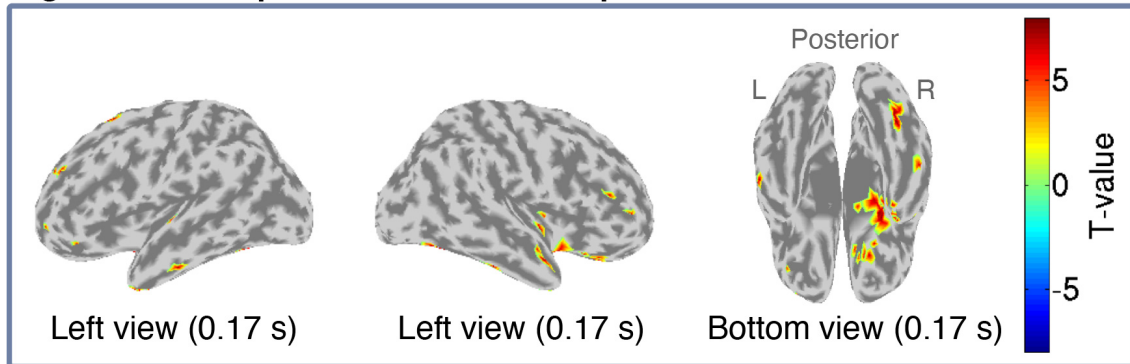
Significance map with meta-analysis fMRI prior
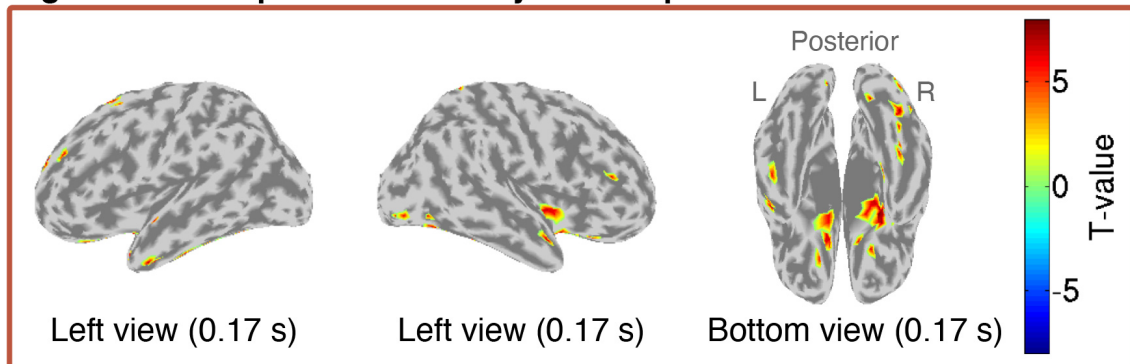
Figure 5.11. Significant t-value maps of the individual fMRI prior (top) and the meta-analysis prior (bottom) at 0.17 sec.

5.2.2 Model selection using free energy

We attempted free energy based model selection as well as simulated data. We selected candidate meta-analysis dataset including relevant ("visual," "face," and "object recognition") and irrelevant ("motor") terms. All hVB estimations were conducted with prior weight = 0.3. The results are shown in Fig. 5.12. Because baselines of subjects are different, between-prior differences are slight in the left panel. Therefore, we corrected each subject's baseline to zero to clarify the between-prior differences (right panel). This figure shows that the free energy of the meta-analysis prior of the "motor" term is higher than the subject's own fMRI data and the meta-analysis data of "visual," although this is a face

recognition task. However, when "motor" was used as a prior of the source reconstruction, little activity was observed in the visual processing area. Besides, the results of the group analysis showed almost no significant difference between conditions. Therefore, the results of the "motor" prior supported by free energy were inappropriate. Compared to the results of simulated data, this does not support the prior data selection using free energy.

Next, we attempted the same selection using trial-averaged data. The results are shown in Fig. 5.13. Note that we applied the hVB algorithm to all single-trial data rather than trial-averaged data because our previous study showed higher intra-subject reproducibility in the former setting. This figure visualized that ranks of free energy were arranged in the order of "visual," "Individual fMRI," "face," "object recognition," and "motor." Moreover, individual fMRI data and all relevant terms were superior to the irrelevant term "motor." These supported the plausibility of the free energy based prior data selection.

FE comparison for experimental data (single-trial setting)
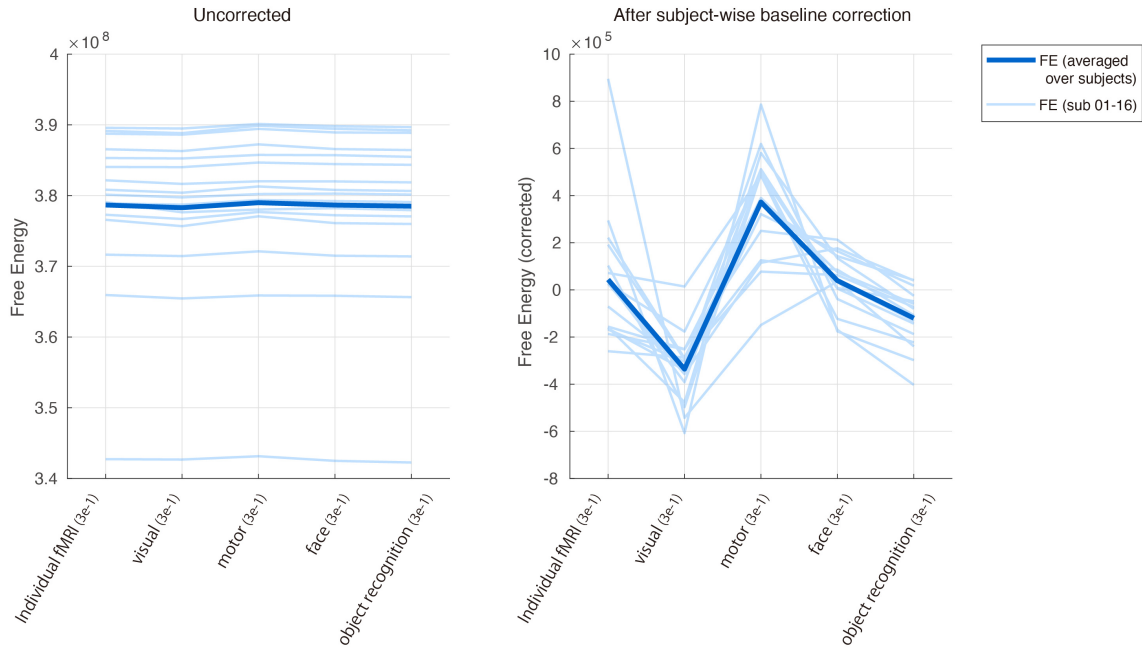
Figure 5.12. Free energy comparison of various priors for experimental data using the single-trial setting. Uncorrected (left) and subject-wise baseline corrected free energies are plotted. Both panels show each subject's free energy (plotted as light blue lines) and averaged one (plotted as a dark blue line). All results are derived using prior weight 0.3. Note higher free energy guarantees higher lower bounds of the evidence (see Appendix A).

FE comparison for experimental data (trial-averaging setting)
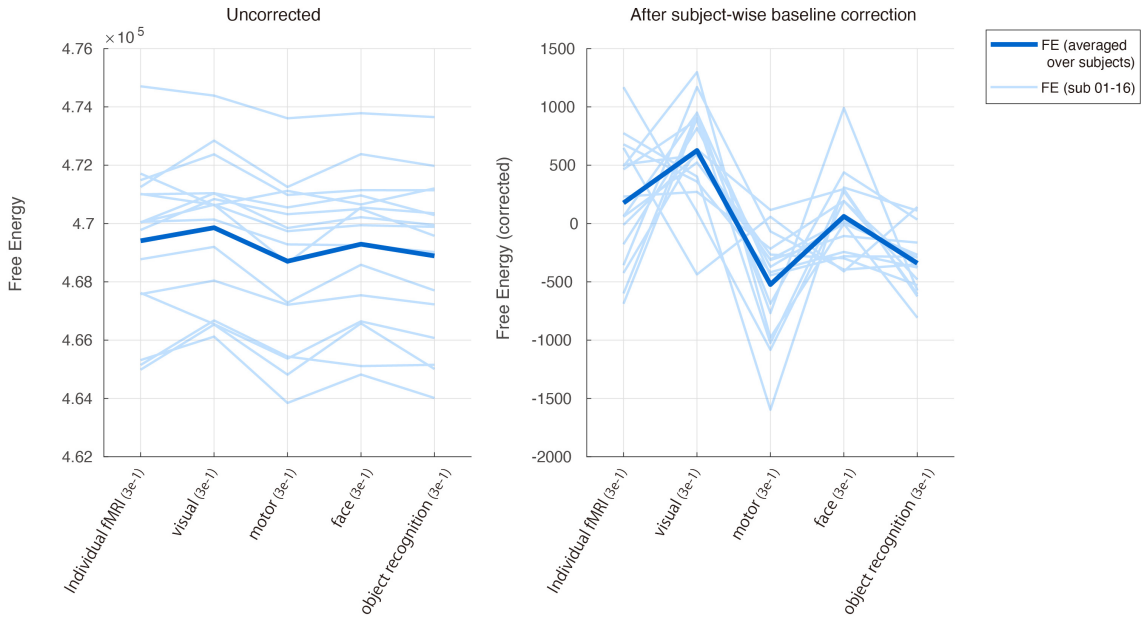
Figure 5.13. Free energy comparison of various priors for experimental data using the trial-averaging setting. Uncorrected (left) and subject-wise baseline corrected free energies are plotted. Both panels show each subject's free energy (plotted as light blue lines) and averaged one (plotted as a dark blue line). All results are derived using prior weight 0.3. Note higher free energy guarantees higher lower bounds of the evidence (see Appendix A).

6. Simulation and Experimental data for RSP

We explain the details of the simulation and experimental data to confirm the RSP approach.

6.1 Data simulation

6.1.1 Constructing meta-analysis dataset and covariance matrix

For the first, we prepared a meta-analysis dataset using Neurosynth to generate the simulated data for evaluation of RSP approach. Since there are thousands of terms in Neurosynth, we referred to a study that pruned and clustered all terms in Neurosynth into 7 domains with less than 150 terms using hierarchical clustering (W.H.Thompson & P.Fransson 2017). Using this result, we constructed three datasets of 15, 25, and 129 terms, respectively. All selected terms are lined up in Table. D.1.

Next, hyper parameters and dist function of Eq.(2.35) are defined. We set $\rho = 9$, $l = 2$, and dist function as cosine similarity (Note that cosine similarity does not meet the axiom of distance but useful to evaluate similarity of data). Hence, we derived the covariance matrix K for each dataset. Fig. 6.1 shows K for the dataset of 15 terms. We can see there are clusters in matrices because terms are arranged by domains.

To compare RSP with MSP, we also prepared a scalar covariance matrix for each datasets.
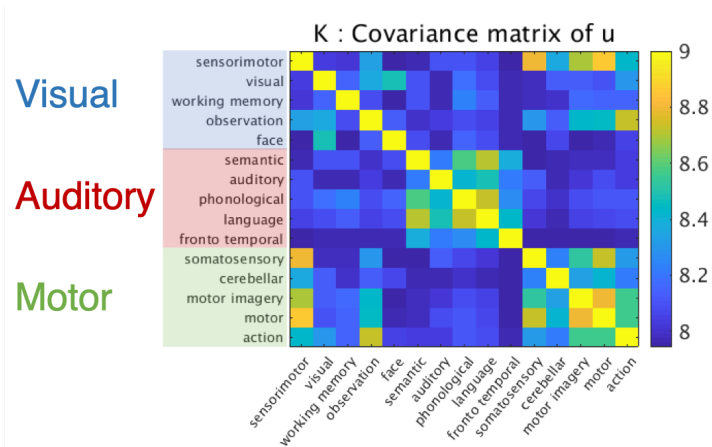
Figure 6.1. Covariance matrix K (dataset of 15 terms). The covariance matrix K of u is imaged. It reflects a relevance structure of terms. Terms are arranged and colored by domains (blue : Visual, red : Auditory, green : Motor).

6.1.2 Data generation

We generated simulated data by following the observation process as shown in Fig. 6.2 . Here, hyper parameter $m = -5$, $\sigma^2 = -50^2$ (no observation noise), and the time length $T = 50$.

To evaluate our RSP approach, we estimated $\boldsymbol{u}$, $\boldsymbol{w}$, $\boldsymbol{J}$, and hyper parameters using a observation $\boldsymbol{B}$.
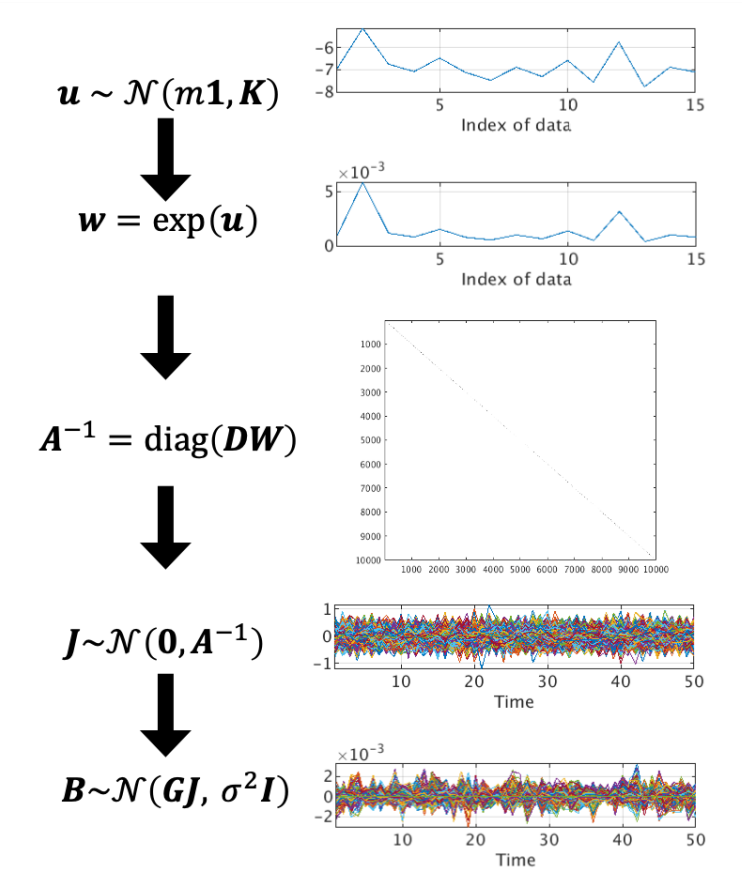
Figure 6.2. Data-generation process for RSP approach.

6.2 Experimental data

We analyzed the same face processing dataset as hVB approach with same pre-processing (see section 4.2). Here, we only focused on a subject (sub-8), the face condition. Moreover, we set a time of interest (ToI) because RSP assumes the stationarity. The ToI is set to the temporal peak (0.16 sec after onset) $\pm$ 25 msec. ToI of all trials is cut out and assembled to construct long observation data. Consequently, we prepared a MEG observation time series with the number of timepoints $T = 3438$. A graphical explanation of this process is shown in Fig. 6.3.

We conducted RSP estimation using a meta-analysis dataset with 15 terms

and covariance K of them (as same as Fig. 6.1). RSP estimation is also conducted using scalar covariance matrix (equivalent to MSP). Then, results of RSP and MSP were compared.
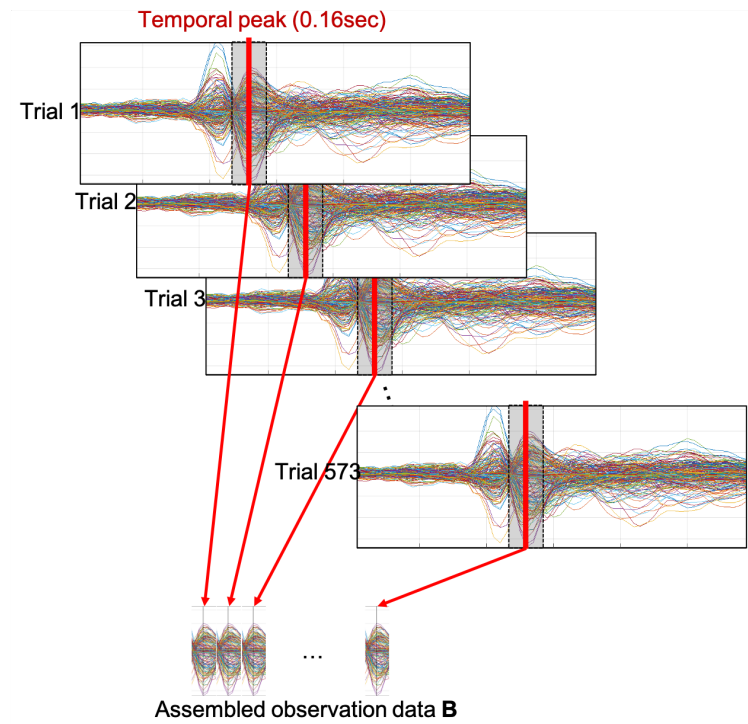
Figure 6.3. Preprocessing of the experimental data.

7. Results for RSP approach

7.1 Simulated data

We studied the reconstruction performance of RSP and MSP by changing the size of the dataset from 15 to 129 (Fig. 7.1)). For the benchmark, the results of MNE and hVB were also evaluated. HVB was computed only for 15 dataset due to the computational cost. The term of meta-analysis data for hVB was selected based on free energy, and prior weight parameter was set to 0.03. Here, we evaluated three reconstructed values; $\bar{\boldsymbol{B}}$, $\bar{\boldsymbol{J}}$, and $\bar{\boldsymbol{w}}$, where $\bar{\boldsymbol{B}}$ is calculated as $\boldsymbol{G}\bar{\boldsymbol{J}}$. Since MNE and hVB do not have the parameter $\bar{\boldsymbol{w}}$, $\bar{\boldsymbol{B}}$ and $\bar{\boldsymbol{J}}$ were drawn. As the figure indicated, every method reconstruct $\bar{\boldsymbol{B}}$ perfectly except for hVB. The reason for the perfect reconstruction might come from the inverse problem. Namely, there is an infinite number of solutions that meet the forward model. However, for the hVB, $\bar{\boldsymbol{B}}$ and also $\bar{\boldsymbol{J}}$ were significantly lower than others. This is because of the simulation setting. The data were simulated using multiple meta-analysis fMRI data; therefore, it is a disadvantageous condition for hVB, which can use only one piece of information. This result suggested that hVB is inadequate for tasks such as stimulating complex brain functions (e.g., movie watching (ES Finn and PA Bandettini, 2021)).

Next, when we look at $\bar{\boldsymbol{J}}$, we cannot see differences between RSP and MSP, while MNE shows lower performance than them. It indicates that RSP and MSP outperform the no-functionally constrained method, although the data include complex brain activation.

Finally, we can find differences between RSP and MSP in $\bar{\boldsymbol{w}}$. When the number of terms is large, RSP is slightly better than MSP. Furthermore, when the number of terms is small, some estimations of MSP are wrong. It indicates that the possibility of RSP in terms of stability of estimation.
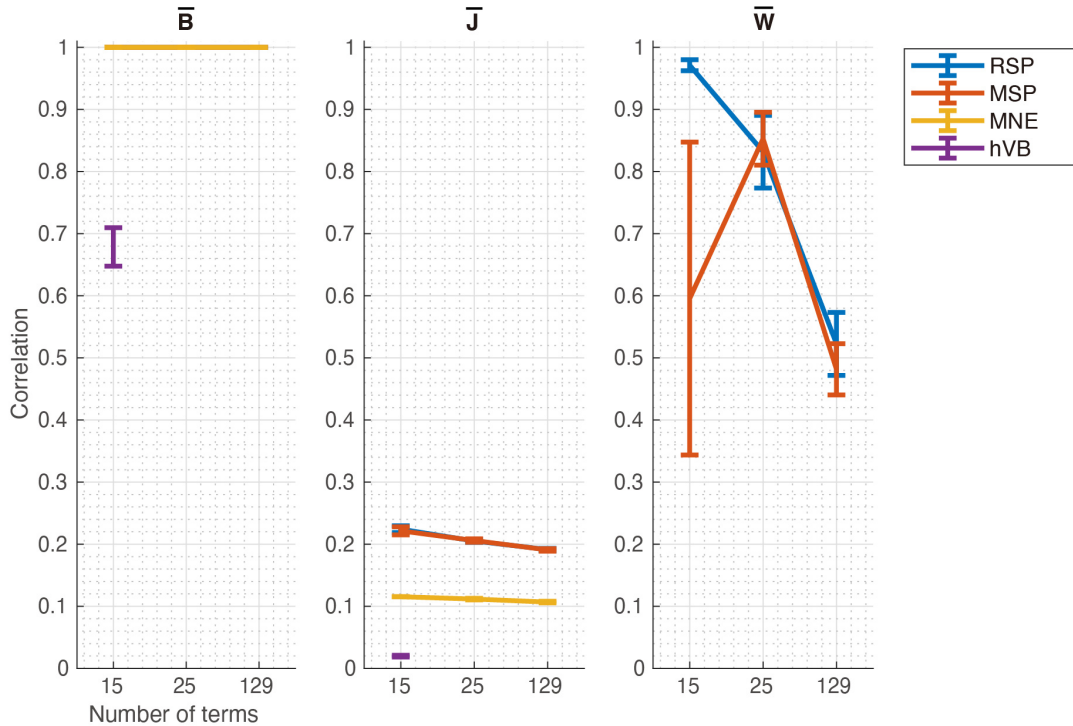
Figure 7.1. Comparing the results of RSP and MSP. For the benchmark, MNE and hVB are reported as well in $\bar{B}$ and $\bar{J}$. The y-axis of $\bar{w}$ indicates Pearson's correlation coefficients. The y-axis of $\bar{B}$ and $\bar{J}$ is temporally averaged correlation coefficients defined in section 3.4.

7.2 Experimental data

We compared currents maps and weights $\bar{w}$ estimated using RSP and MSP for face processing data. Reconstructed currents maps (Fig. 7.2) show similar activations on FFA. However, there are several differences between them. First, MSP reconstructed more currents on the dorsal stream than RSP. Second, RSP indicated activations on the insular cortex similar to the result of group analysis at almost the same time point using hVB (Fig. 5.11). As the same discussion in Appendix F, it might be caused by the signal leakage from FFA.

These differences in reconstructed maps are interpretable by analyzing esti-

mated weights (Fig. 7.3). As the figure indicated, there is a significant difference between $\bar{w}$ of RSP and MSP. While MSP estimated only on "visual," RSP estimated not only "visual" but also "face." This difference caused the difference in reconstructed sources because meta-analysis fMRI data "visual" includes both the ventral and dorsal streams.
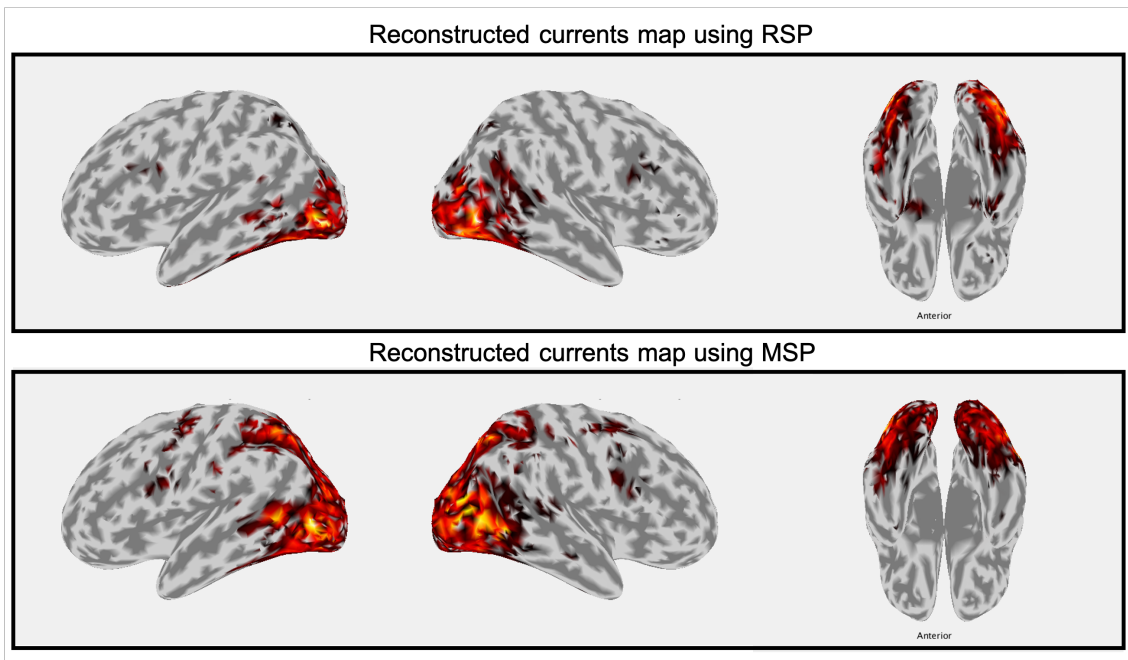
Figure 7.2. Reconstructed currents maps using RSP and MSP. To visualize the maps, amplitudes are averaged along the time axis. For both figures, activities over 10 % of their maximum value are displayed.
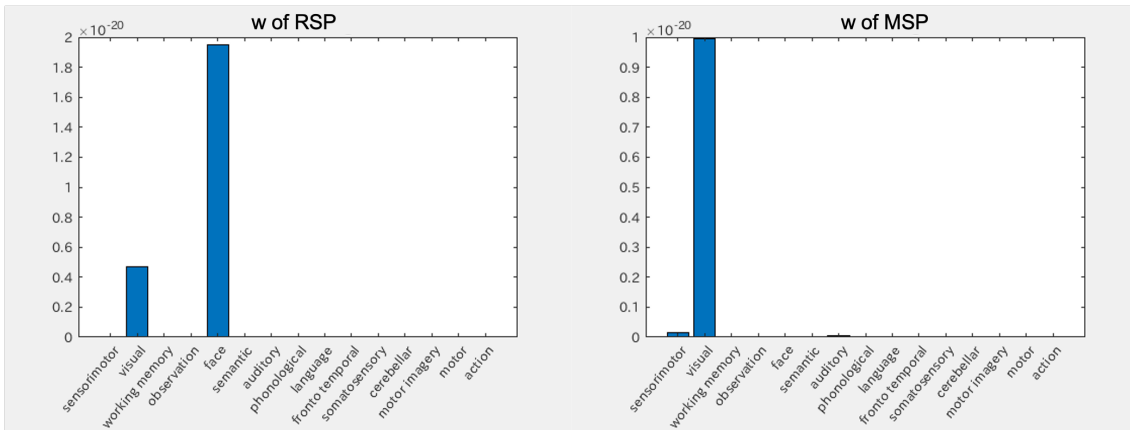
Figure 7.3. Comparing estimated w of RSP and MSP.

8. Discussion

In the present study, we proposed the use of meta-analysis fMRI data for MEG current source reconstruction. The meta-analysis fMRI data is a statistical map synthesized from thousands of published fMRI studies. Here, we introduced two different approaches.

8.1 Discussions on hVB approach

For the first, we used the meta-analysis results available from the Neurosynth open-source project (Yarkoni et al., 2011) as the hierarchical prior distribution of the current variance in the hierarchical variational Bayesian estimation method (Sato et al., 2004) because this approach offers an adaptive way to incorporate fMRI information based on MEG measurements. This hVB approach was quantitatively evaluated using simulations of a visual task and a motor task. As a result, we discovered that source reconstruction performance using the meta-analysis prior was comparable to that using an individual fMRI prior with CNR = 1 and CNR = 0.5 for the visual and motor tasks, respectively. We also confirmed the fMRI prior outperformed the MNE, Champagne, and hVB estimations with a uniform prior (no fMRI information). In particular, when the simulated currents are spatially dense (p < 0.05), or the SNR of MEG data is high, hVB using the low-CNR fMRI prior was better than the sparse estimations of the Champagne method. Using the experimental data of a real face recognition task (Wakeman and Henson, 2015), we qualitatively confirmed that group analysis results obtained from the meta-analysis fMRI prior were similar to those obtained from individual fMRI priors.

Our simulations were based on multi-modal data generation, where both whole-brain fMRI and MEG time series data were generated from the common truth activation map. We defined the common truth activation map for a specific task, which was the group analysis of results obtained from real fMRI experimental data of five subjects during the task, representing the spatial pattern of task-specific neuronal activation. From the common truth map, whole-brain fMRI time series data were generated using the virtually designed event-related paradigm and resting-state fMRI time series as noise. We modified the amount of

resting-state fMRI noise to obtain the individual fMRI prior maps of various signal quality, defined as CNR. MEG time series was generated by applying the lead field matrix to simulated current time series data with added noise. The current time series data for a specific task were generated by weighting the whole-brain current time series (simulated using the connectome dynamics model) with the spatial map of the task (common truth map). This procedure allows us to remove any biases of the cancellation effect among dipoles caused by the hand design. However, the simulated current time series seems to be more difficult than the empirical one when we do not adjust the wCI. This is because the connectome dynamics model was simulated with coarse parcellation (AAL atlas), and the time series were common within each ROI. Moreover, the coherence of simulated resting-state data tends to be higher than the stimulus evoked response. To make the difficulty realistic under such parcellation, we modulated wCI of the generated data by thresholding the common truth with $p < 0.001, 0.01$, or 0.05. We also modulated observation noises of MEG data with SNR $= 0, 5$, or 10 [dB] by varying the number of trials. Then, simulations were conducted with nine combinations of settings (3 SNR by 3 wCI).

Inspired by Owen et al. (2012), we assessed the results of source reconstruction using the aggregate score (Eq. (3.1)). This metric averaged the sum of spatial and temporal correlation coefficients between simulated and reconstructed currents. By comparing the evaluation using the aggregate score (Fig. 5.3), spatial correlation (Fig. E.1), and temporal correlation (Fig. E.2), we confirmed that the differences in prior data mainly affect the estimated spatial maps rather than the temporal time series. This is natural because we incorporated the spatial prior data in the covariance of the current map.

From the simulation studies of the visual and motor tasks, we confirmed that the current source reconstruction performance of hVB using the meta-analysis fMRI prior was better than that using the low-CNR individual fMRI prior. For the visual domain task (left of Fig. 5.3), the result of the relevant meta-analysis prior shows a better score than the individual fMRI prior with CNR=1. For the motor domain task (right of Fig. 5.3), the result of the relevant meta-analysis prior shows a better score than for the individual fMRI prior with CNR=0.5. Referring to Figs. C.1 and C.2, CNR1 and CNR0.5 of the $p < 0.05$ condition corresponds

to PSC0.7 and PSC0.6 at the maximum activated voxel, respectively. Note that these low PSC values could be observed as the worst-case in a real-task fMRI study (e.g., Drobyshevsky et al., 2006, reported similar PSCs for the cognitive and emotional tasks). We also confirmed that the hVB method using the meta-analysis fMRI prior was superior to the control methods without fMRI priors, such as the hVB method using the uniform prior, wMNE, and Champagne. Especially, by comparing with the Champagne method, we confirmed that the meta-analysis prior for the hVB method was helpful when the source spatial density was high ($p < 0.05$). Moreover, the meta-analysis fMRI data synthesized using an irrelevant term, such as "motor" for the visual task or "visual" for the motor task, resulted in degraded performance. Therefore, we concluded that it is worth using the meta-analysis fMRI prior synthesized using relevant terms when the individual fMRI data are missing or of low quality. This also means that the results equivalent to the use of a low-CNR fMRI prior can be obtained without fMRI measurement, and in that case, the reduced cost could be used to enhance MEG measurement. It should be noted that we used the default hyperparameters for the control methods (MNE and Champagne) while the optimized parameters for the hVB methods. Thus, the performance comparison may not be fair. Although their performance may be improved using parameter fine-tuning (Bertrand et al., 2019; Cai et al., 2021), it is beyond the scope of this study.

The prior weight parameter is the most important parameter for hVB estimation, which affects the quality of current source reconstruction. Therefore, we studied the proper range of the prior weight parameter using simulated data by the grid-search strategy. For both visual and motor tasks (Fig. 5.1 and Fig. 5.2), the optimal prior weight tends to be high when the cancellation is high ($p < 0.05$) regardless of the type of fMRI priors, while the optimal prior weight is low when the cancellation is low ($p < 0.001$). These results indicate that the current source with a higher cancellation setting requires stronger prior information because strong cancellation induces fewer MEG observations.

Using the experimental data of a real face-recognition task, we qualitatively confirmed that group analysis results of the meta-analysis fMRI prior have a similar tendency with the results of an individual fMRI prior. We set the prior weight parameter for the meta-analysis fMRI prior to 0.3 based on the simulation

study of the visual task and set the parameter for the individual fMRI prior to 0.3 according to the previous work (Takeda et al., 2019). For the significance time series of both priors, the largest difference was observed at the same time point, although there were small differences (Fig. 5.10). In Fig. 5.11, we compare the significant t-value maps of both priors at that time. Both of them show significance at the right FFA, although they were derived using different types of priors. The significance at the right FFA is consistent with previous studies reporting that this area exhibits face-selective responses (Grill-Spector et al., 2017, 2004; Jas et al., 2018; Rossion et al., 2018; Wakeman and Henson, 2015). In these figures, we can observe activities on the insula as well. We hypothesized that they occurred due to signal leakage (Brookes et al., 2012; Colclough et al., 2015; Palva et al., 2018; Sato et al., 2018). To verify this, we calculated the resolution kernel for both priors (Sekihara et al., 2005). As a result of the analysis, the activities on the insula could be observed due to the signal leakage from the ventral occipitotemporal cortex (Appendix F). These results corresponded to our previous study (Takeda et al., 2019). Such similarities in the results between individual fMRI and meta-analysis fMRI suggest the possibility of substituting a meta-analysis prior for the individual prior in a neuroscience study.

It is difficult to subjectively select the term used to synthesize the meta-analysis fMRI data. This came from the results with detailed terms such as "face" and "object recognition" for the visual domain task and "hand" and "foot" for the motor domain task (Figs. E.5 and E.6). The meta-analysis prior synthesized using the term "foot" significantly improved the score (comparable to the high-CNR fMRI priors) for the left-foot movement task. However, against our intuition, the result with the term "face" was not improved, and "object recognition" was lower than the result with the term "visual." These results might be due to the following reasons. In most visual tasks, the stimulus-induced activation occurs from the primary visual cortex, and it is then transferred to the higher visual cortex. Hence, the entire visual cortex is activated. This situation is favorable for the meta-analysis prior synthesized using the term from the highest layer of the ontology. Therefore, the term "face" was not a better selection than the term "visual." On the other hand, the left-foot movement task with the $p > 0.05$ setting mainly activated left-foot-related areas. Therefore, the term from

the lower layer of the ontology was suitable for this situation. In the results, only "foot" was a good prior, whereas both "hand" and "foot" were selected from the lower layer. This seems natural because the meta-analysis with the term "hand" is not directly related to the left-foot-related areas. Although these obvious results indicate that the meta-analysis data synthesized using detailed terms might improve the reconstruction, they also imply the difficulty of making the appropriate data selection for the meta-analysis.

In the neuroscience literature, free energy is one of the popularly used statistical criteria for model selection (Friston et al., 2003; Fukushima et al., 2015; Wipf and Nagarajan, 2009). In particular, free energy is used to select prior information of source reconstruction in the multiple sparse priors method (Friston et al., 2008; Henson et al., 2010). Inspired by these studies, we attempted to select hyperparameters (prior weight and meta-analysis fMRI data) for hVB estimation using free energy. First, we studied whether the prior weight parameter could be selected using free energy. The effects of the prior weight on free energy and the aggregate score are reported in Fig. 5.5. As the results indicate, free energies were negatively correlated with prior weights, but they were not related to aggregated scores. This tendency was shown for other prior data as well. We speculated that this is because when the prior weight is low, the estimated currents can be fitted to the observation data independent of the prior data and thus free energy also has a high value. Thus, we concluded that prior weight cannot be selected based on free energy. Next, we considered the prior data selection when the prior weight was fixed to 0.3 for all priors. For both visual and motor tasks, we showed the results with SNR0 and 10, respectively (Figs. 5.6 and 5.7). These figures indicate strong relationships between free energies and aggregated scores regardless of SNR. These results support the potential of selecting the prior data using free energy. Therefore, we also plotted free energies with the prior weight of 0.3 for the experimental data (Fig. 5.12). The figure shows that the free energy of the meta-analysis prior of the "motor" term is higher than the subject's own fMRI data and the meta-analysis data of "visual," although this is a face recognition task. However, when "motor" was used as a prior in the source reconstruction, little activity was observed in the visual processing area. In addition, the results of the group analysis showed almost no significant difference between conditions.

Therefore, the results of the "motor" prior supported by free energy were inappropriate. Compared to the results of simulated data, this does not support the prior data selection using free energy. Note that we applied the hVB algorithm to all single-trial data rather than trial-averaged data because our previous study showed higher intra-subject reproducibility in the former setting. However, when we applied the trial-averaged data and checked the applicability of free energy for the prior data selection, we obtained plausible results similar to the results of simulated data (Fig. 5.13). This is probably because the SNR of the single-trial data of real data was much worse than the SNR0 of the simulation. To confirm the applicability of prior data selection based on free energy, further investigation into the impact of a single-trial setting on free energy computation is required. Furthermore, since we need to repeatedly run the hVB algorithm for all candidate prior data to compute free energy, the computational cost will be significantly increased in proportion to the number of candidates.

Care must be taken when using the meta-analysis fMRI prior for scientific findings. Use of prior information biases the results of current source reconstruction. Since there are numerous combinations of the meta-analysis prior and the prior weight value, the choice depends on the data analyst, and results may be chosen after trials. To mitigate such results selection, we need a guideline on how to use the method and report the results. For now, we recommend that meta-analysis fMRI data be synthesized using relevant and general (conservative) terms such as "visual" for visual tasks or "motor" for motor tasks. This comes from the above discussion on term selection. However, there is no solution to decide the prior weight value systematically, although it has been shown to have a significant impact on the results as well (Figs. 5.1 and 5.2). Therefore, trial-and-error efforts are indispensable to tune the prior weight. However, this recommendation might be updated by the progress made in free-energy-based prior data selection. For example, it may be possible to select the term based on the free energy using pre-fixed prior weight value. On the other hand, this method is ready for applications such as brain-machine interface (BMI), where the direct goal is to improve decoding accuracy. In this case, we believe there is no problem in selecting a term that will improve the decoding accuracy as long as the generalization performance is evaluated.

8.2 Discussions on RSP approach

As above discussions reveal, one of the most critical issues in hVB approach is the selection bias. To overcome it, we introduced second approach, RSP estimation. While hVB utilizes ARD to suppress error activation on dipoles of prior data, RSP selects the combination of prior data itself. This mitigates the selection bias because we do not need to select prior data.

As the results of the simulation studies indicated, RSP improved the reconstruction significantly than MNE and hVB (Fig. 7.1). Furthermore, it was suggested that RSP estimate latent variable w more stably than MSP although the performance of current reconstruction itself is equivalent to MSP. Both RSP and MSP were robust to increasing the number of prior data in terms of reconstructed currents J.

The results of the experimental data indicated interesting differences between RSP and MSP (Fig. 7.2). First, MSP reconstructed more currents on the dorsal stream than RSP. Second, RSP indicated activations on the insular cortex similar to the result of group analysis at almost the same time point using hVB (Fig. 5.11). As discussed above, it might caused by the signal leakage from FFA. These differences in reconstructed maps are interpretable by analyzing estimated weights (Fig. 7.3). As the figure indicated, there is a significant difference between $\bar{w}$ of RSP and MSP. While MSP estimated only on "visual", RSP estimated not only "visual" but also "face". This difference caused the difference in reconstructed sources because meta-analysis fMRI data "visual" includes both of the ventral and dorsal streams.

The RSP estimation using meta-analysis fMRI dataset have another possibility related to interpretable estimation. Because all of meta-analysis fMRI data are labelled by terms, we might discuss the estimation based on selected data. If this is established, it could lead to the ability to reverse inference at MEG time resolution.

8.3 General discussions

In this study, we used the results of meta-analysis from Neurosynth. However, many useful and practical platforms for fMRI meta-analysis have also been pro-

posed and are promising as the prior information of source reconstruction. The BrainMap group supplies the ICA maps of large-scale meta-analysis fMRI data (`https://www.brainmap.org/icns/`). These are well-suited to our present recommendation of selecting a conservative prior for hVB approach because ICA maps are already classified to a few intrinsic connectivities (Laird et al., 2011; Smith et al., 2009). This feature is also suitable for RSP approach because it might help interpretations of selected prior data. Another option is NeuroQuery (`https://neuroquery.org`). Because of its ability to predict brain activity from a short sentence (Dockès et al., 2020), it seems to be suitable for applications such as decoding newly designed tasks. While this contradicts our recommendation of hVB approach to generate a conservative prior, it is still attractive for applications such as BMI.

In terms of reproducibility, meta-analysis fMRI is promising. In recent years, the study of task fMRI has been a problem because of its poor reproducibility (Elliott et al., 2020). When individual fMRI data are used as the prior for source reconstruction, the reconstructed maps are accumulated with uncertainty included in the fMRI data. On the other hand, although the meta-analysis prior does not take into account individual variability, it is not affected by the reproducibility issue of individual data. Therefore, it is not possible to conclude which is better, but in terms of reproducibility, there is potential in using meta-analysis data as prior information. This is supported by the fact that the results of the two group analyses were very similar (Fig. 5.11).

Our approach is also easily applicable to the source reconstruction problems of EEG. As demonstrated previously (Takeda et al., 2019), hVB estimation for EEG data is supported in VBMEG software. And RSP can use EEG data as same as MEG because formulations are same for EEG. We expect the low measurement cost of EEG to make it suitable for combination with meta-analysis fMRI data.

The use of the meta-analysis fMRI prior would be a significant step in the development of MEG studies. This is because the advent of OPM has shown remarkable results in reducing the cost of MEG measurements and expanding the measurement targets (Boto et al., 2018; Hill et al., 2019; Lin et al., 2019; Tierney et al., 2020). However, the number of sensors in the current OPM devices is limited to dozens due to physical constraints (Hill et al., 2020). Therefore, it

would be important to gain the ability to add prior information that is comparable to a low-CNR individual fMRI prior without the cost of measurement.

Appendix

A. Hierarchical variational approximation

The hVB approximates the joint posterior distribution Eq. (2.23) using the distribution $Q(\boldsymbol{J}, \beta, \boldsymbol{A})$ (Sato et al., 2004). It is obtained by maximizing the following free energy $F(Q)$ under the independent decomposition assumption that $Q(\boldsymbol{J}, \beta, \boldsymbol{A}) = Q(\boldsymbol{J}, \beta)Q(\boldsymbol{A})$:

$$F(Q) = \int Q(\boldsymbol{J}, \beta, \boldsymbol{A}) \log \left[\frac{P(\boldsymbol{J}, \beta, \boldsymbol{A}, \boldsymbol{B})}{Q(\boldsymbol{J}, \beta, \boldsymbol{A})} \right] d\boldsymbol{J} d\beta d\boldsymbol{A}$$

$$= \log P(\boldsymbol{B}) - \mathrm{KL}[Q(\boldsymbol{J}, \beta, \boldsymbol{A}) || P(\boldsymbol{J}, \beta, \boldsymbol{A} | \boldsymbol{B})]. \tag{A.1}$$

Here, maximizing the free energy guarantees maximizing the lower bound of the evidence $P(\boldsymbol{B})$ (Bishop, 2006). As a result, the maximization of free energy is realized as a sequential optimization for below $Q(\boldsymbol{J}, \beta)$ and $Q(\boldsymbol{A})$ (Sato et al., 2004):

$$Q(\boldsymbol{J}, \beta) = \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{j}_t | \bar{\boldsymbol{j}}_t, (\beta \boldsymbol{S})^{-1}) \Gamma(\beta | \beta_p, \gamma_{\beta p}), \tag{A.2}$$

$$\boldsymbol{S} = \boldsymbol{G}^{\top} \boldsymbol{\Phi} \boldsymbol{G} + \bar{\boldsymbol{A}}, \tag{A.3}$$

$$\bar{\boldsymbol{j}}_t = \boldsymbol{S}^{-1} \boldsymbol{G}^{\top} \boldsymbol{\Phi} \boldsymbol{B}_t, \tag{A.4}$$

$$\beta_p^{-1} = \frac{1}{TN} \mathrm{Tr} \left[\left(\boldsymbol{G} \bar{\boldsymbol{A}}^{-1} \boldsymbol{G}^{\top} + \boldsymbol{\Phi}^{-1} \right)^{-1} \sum_{t=1}^{T} \boldsymbol{B}_t \boldsymbol{B}_t^{\top} \right], \tag{A.5}$$

$$\gamma_{\beta p} = \frac{TN}{2}, \tag{A.6}$$

$$Q(\boldsymbol{A}) = \prod_{n=1}^{N} \Gamma(\alpha_n | \alpha_{np}, \gamma_p), \tag{A.7}$$

$$\gamma_p = \gamma_0 + \frac{T}{2}, \tag{A.8}$$

$$\alpha_{np}^{-1} = \left[\frac{\gamma_0}{\alpha_{n0}} + \frac{1}{2} \sum_{t=1}^{T} \left(\beta_p \bar{j}_{n,t}^2 + S_{n,n}^{-1} \right) \right] \gamma_p^{-1}, \tag{A.9}$$

where $\bar{\boldsymbol{j}}_t$ is the mean of posterior current distribution at time point t and $\bar{\boldsymbol{A}} = \mathrm{diag}(\alpha_{1p}, \alpha_{2p}, \ldots, \alpha_{Np})$ is the posterior precision matrix of the currents. $\bar{j}_{l,m}$ and

$S_{l,m}$ are the l, m elements of $\bar{\boldsymbol{j}}_t$ and $\boldsymbol{S}$, respectively. When we introduce the prior weight Eq. (2.24) to Eq. (A.9), we can derive Eq. (2.25). The hVB algorithm evaluates $Q(\boldsymbol{J}, \beta)$ and $Q(\boldsymbol{A})$ iteratively until the free energy converges.

B. Spatial smoothness constraint

When conducting current source reconstruction using hVB, a smoothness constraint is incorporated (Yoshioka et al., 2008). To do this, we modeled current sources as

$$\boldsymbol{j} = \boldsymbol{W}\boldsymbol{z}, \tag{B.1}$$

where $\boldsymbol{W}$ is a spatial filter having a Gaussian profile with a full width at half maximum (FWHM) of 8 mm, and $\boldsymbol{z}$ is an auxiliary variable. Then, the forward model Eq.(2.2) can be replaced by

$$\begin{aligned}
\boldsymbol{b} &= \boldsymbol{G}\boldsymbol{W}\boldsymbol{z} + \boldsymbol{\varepsilon} \\
&= \boldsymbol{G}_w \boldsymbol{z} + \boldsymbol{\varepsilon},
\end{aligned} \tag{B.2}$$

where $\boldsymbol{G}_w = \boldsymbol{G}\boldsymbol{W}$ is a smoothed lead field matrix. Therefore, the source reconstruction becomes the problem of estimating $\boldsymbol{z}$ using $\boldsymbol{G}_w$. For a fair comparison, we filtered the estimations of wMNE and Champagne using the same $\boldsymbol{W}$. Note we confirmed that this process did not bring a disadvantage for any method in terms of either temporal or spatial assessment.
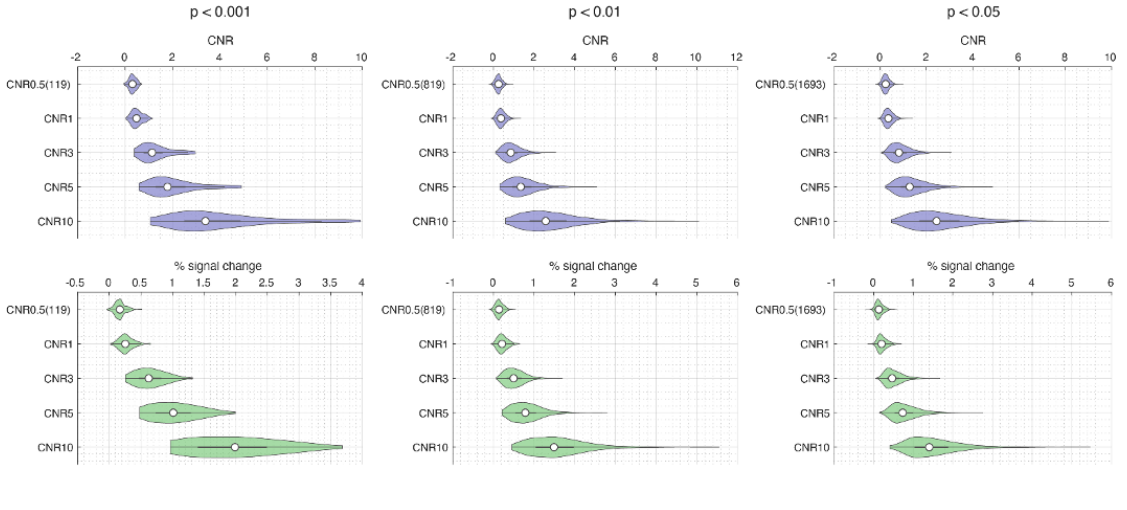
C. Distribution of CNR and PSC

Figure C.1. Distribution of calculated CNR and PSC for the visual simulation setting. The upper figures show CNR, and the lower figures show PSC. Each column corresponds to a different significance level, p<0.001, 0.01, or 0.05 from left to right. We report results only on truly activated voxels. The number in parentheses of CNR0.5 denotes the number of voxels, and it is the same for other CNRs because it is modulated by significance level. The boxplot inside the violins represent the interquartile range, and the white dot shows the median value.
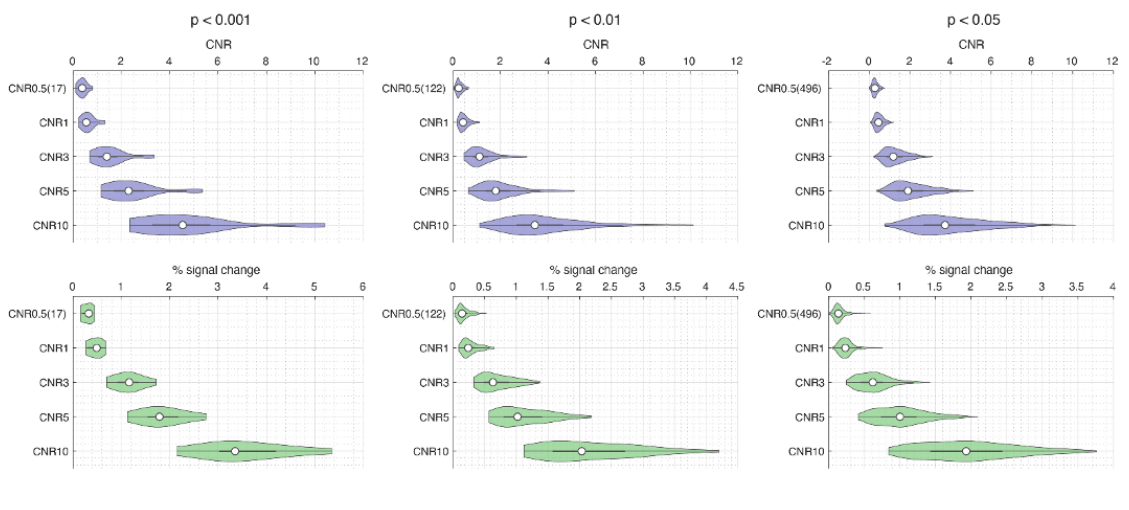
Figure C.2. Distribution of calculated CNR and PSC for the motor simulation setting. The upper figures show CNR, and the lower figures show PSC. Each column corresponds to a different significance level, p<0.001, 0.01, or 0.05 from left to right. We report results only on truly activated voxels. The number in parentheses of CNR0.5 denotes the number of voxels, and it is the same for other CNRs because it is modulated by significance level. The boxplot inside the violins represent the interquartile range, and the white dot shows the median value.

D. List of meta-analysis dataset

Table D.1: List of meta-analysis dataset of 129 terms

Domain	Term
Default Mode	affect
Default Mode	arousal
Default Mode	assessment
Default Mode	autobiographical memory
Default Mode	awareness
Default Mode	basal ganglia
Default Mode	belief
Default Mode	confidence
Default Mode	context
Default Mode	default mode
Default Mode	detection

Default Mode	distress
Default Mode	dorsal attention
Default Mode	emotion
Default Mode	emotion regulation
Default Mode	empathy
Default Mode	episodic memory
Default Mode	evaluation
Default Mode	examination
Default Mode	fear
Default Mode	frontostriatal
Default Mode	hippocampal
Default Mode	identification
Default Mode	inference
Default Mode	inhibitory
Default Mode	interacting
Default Mode	knowledge
Default Mode	limbic
Default Mode	memory
Default Mode	memory encoding
Default Mode	memory retrieval
Default Mode	mentalizing
Default Mode	mind
Default Mode	morphological
Default Mode	navigation
Default Mode	negative affect
Default Mode	olfactory
Default Mode	perspective
Default Mode	practice
Default Mode	prefrontal temporal
Default Mode	recognition memory
Default Mode	recollection
Default Mode	reward
Default Mode	salience
Default Mode	self
Default Mode	self referential
Default Mode	semantic memory
Default Mode	sexual
Default Mode	social
Default Mode	social cognition
Default Mode	stress
Default Mode	subcortical
Default Mode	sustained attention
Default Mode	task positive
Default Mode	temporal parietal
Visual	attentional control

Visual	body
Visual	competing
Visual	face
Visual	meaningful
Visual	motion
Visual	object recognition
Visual	observation
Visual	occipitotemporal
Visual	orienting
Visual	parietal frontal
Visual	planning
Visual	prefrontal parietal
Visual	preparation
Visual	reasoning
Visual	recognition
Visual	representation
Visual	retrieval
Visual	sensorimotor
Visual	spatial attention
Visual	switching
Visual	visual
Visual	visual attention
Visual	visuomotor
Visual	visuospatial
Visual	working memory
Fronto Parietal	attention
Fronto Parietal	choice
Fronto Parietal	cognitive control
Fronto Parietal	conflict
Fronto Parietal	decision making
Fronto Parietal	domain general
Fronto Parietal	executive control
Fronto Parietal	expected
Fronto Parietal	feedback
Fronto Parietal	fronto parietal
Fronto Parietal	goal
Fronto Parietal	learning
Fronto Parietal	maintenance
Fronto Parietal	monitoring
Fronto Parietal	motivational
Fronto Parietal	prediction
Fronto Parietal	response inhibition
Fronto Parietal	risk
Auditory	access
Auditory	auditory

Auditory	categorization
Auditory	coding
Auditory	fronto temporal
Auditory	integration
Auditory	language
Auditory	language comprehension
Auditory	lexical
Auditory	naming
Auditory	phonological
Auditory	reading
Auditory	semantic
Auditory	speech
Auditory	speech production
Auditory	task control
Auditory	temporal frontal
Motor	action
Motor	action observation
Motor	cerebellar
Motor	coordination
Motor	discrimination
Motor	endogenous
Motor	eye movement
Motor	imagery
Motor	imitation
Motor	intelligence
Motor	mirror neuron
Motor	motor
Motor	motor control
Motor	motor imagery
Motor	orientation
Motor	rhythm
Motor	somatosensory

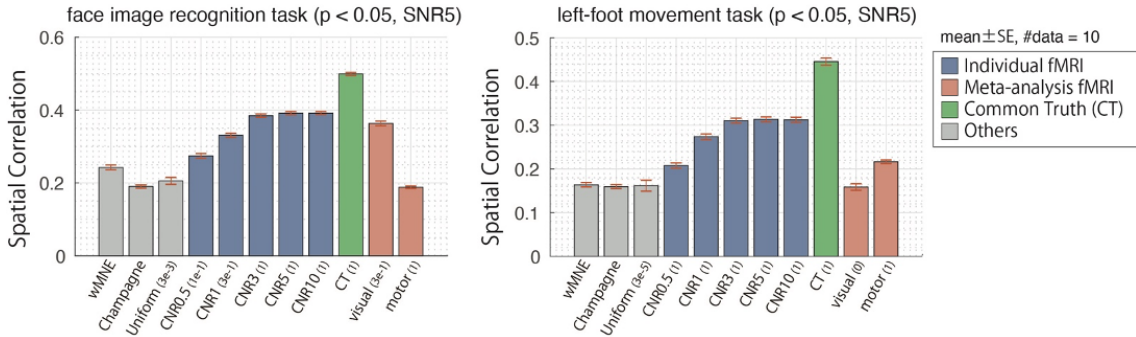# E. Supplementary figures of hVB results

Figure E.1. Comparison of spatial correlation with various methods for the visual task (left) and the motor task (right). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are reported. These priors are used with the best prior weight parameter denoted in parentheses. Vertical axes indicate the correlation coefficient of the spatial map (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of a realistic simulation setting (p<0.05, SNR5) determined using experimental data. The results of wMNE and Champagne are also shown as a control.
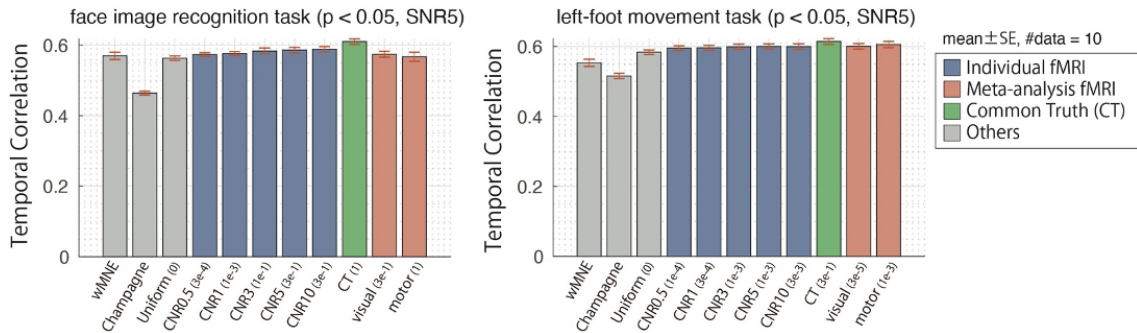
Figure E.2. Comparison of temporal correlation with various methods for the visual task (left) and the motor task (right). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are reported. These priors are used with the best prior weight parameter denoted in parentheses. Vertical axes indicate the correlation coefficient of the time series (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of a realistic simulation setting (p<0.05, SNR5) determined using experimental data. The results of wMNE and Champagne are also shown as a control.
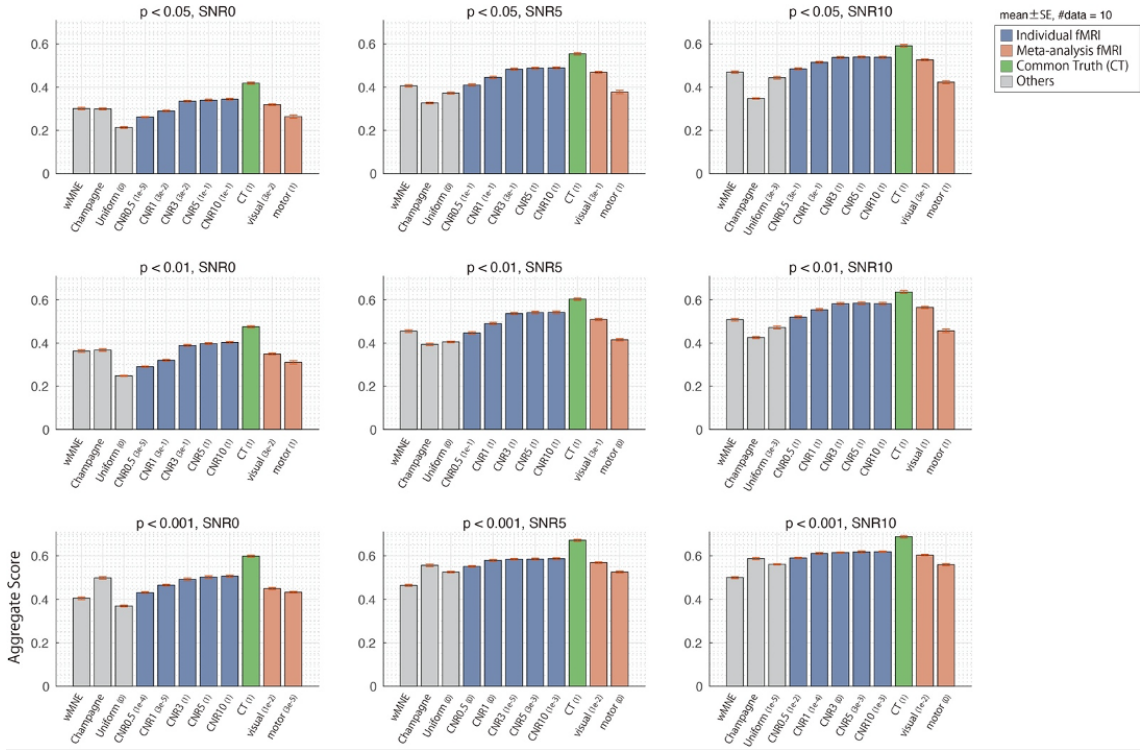
Figure E.3. Comparison of aggregate scores with various methods for all simulation settings (visual task). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are reported. These priors are used with the best prior weight parameter denoted in parentheses. Vertical axes indicate the aggregate score (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of all simulation settings (3 SNR by 3 wCI (denoted as thresholds of the common truth)).
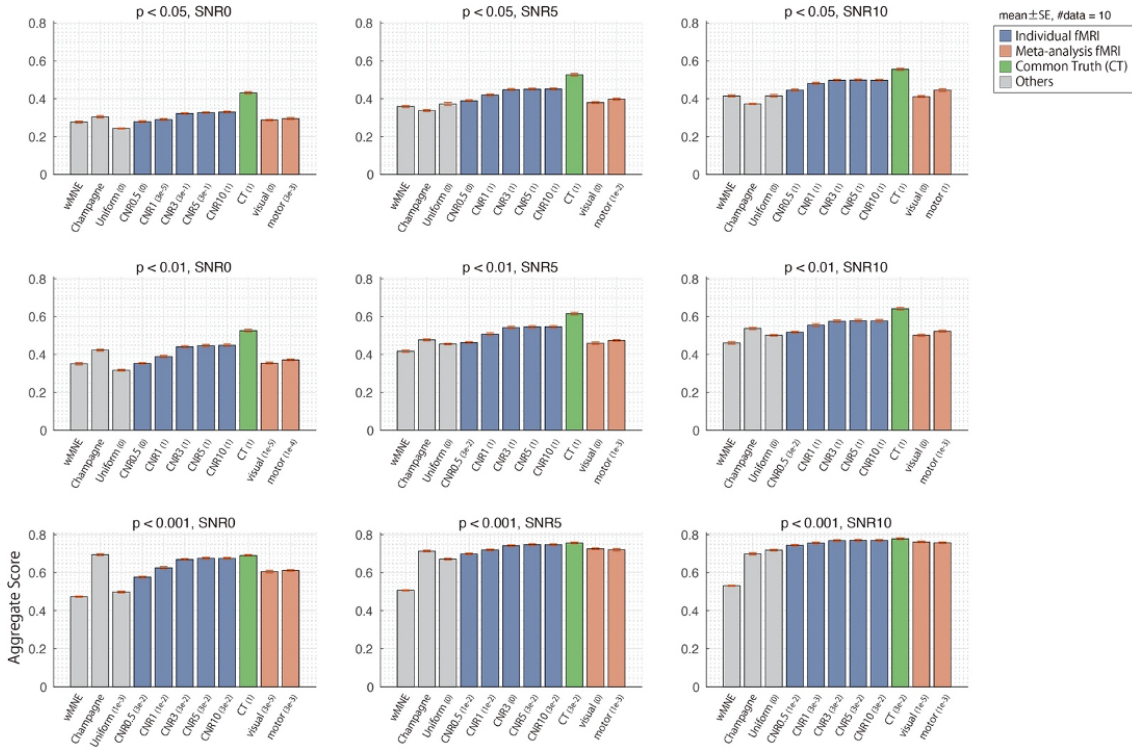
76

Figure E.4. Comparison of aggregate scores with various methods for all simulation settings (motor task). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "visual" and "motor") are reported. These priors are used with the best prior weight parameter denoted in parentheses. Vertical axes indicate the aggregate score (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of all simulation settings (3 SNR by 3 wCI (denoted as thresholds of the common truth)).
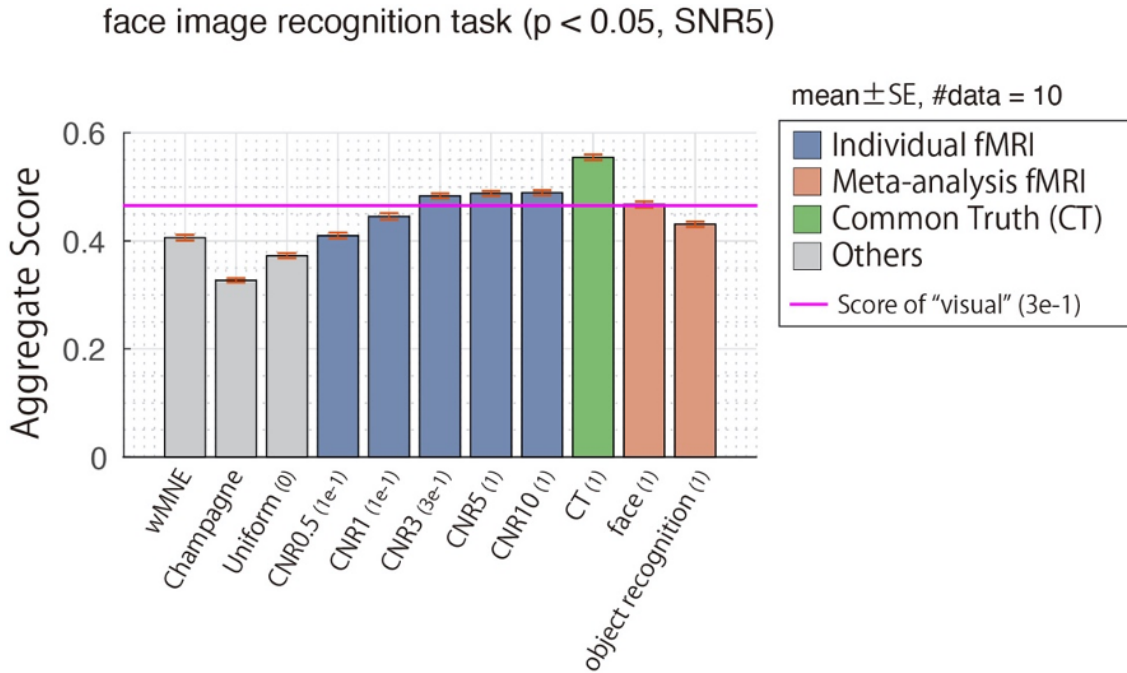
Figure E.5. Score comparison of meta-analysis prior synthesized using detailed terms (visual task). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "face" and "object recognition") are reported. These priors are used with the best prior weight parameter denoted in parentheses. The horizontal line colored magenta indicates the aggregate score of the term "visual" with prior weight 3e-1. Vertical axes indicate the aggregate score (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of a realistic simulation setting (p<0.05, SNR5) determined using experimental data. The results of wMNE and Champagne are also shown as a control.
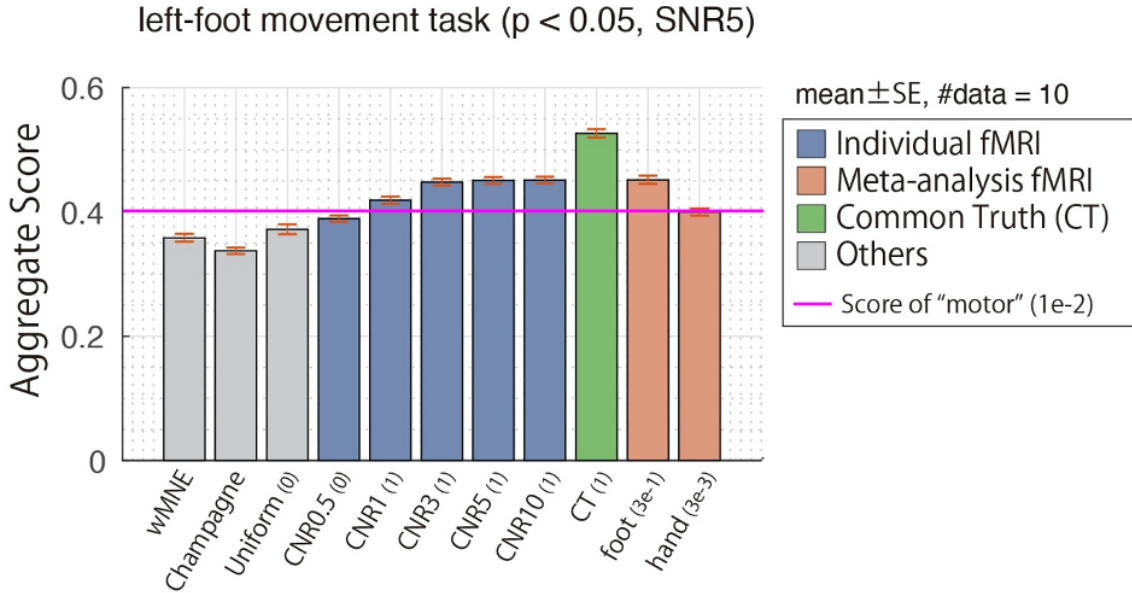
Figure E.6. Score comparison of meta-analysis prior synthesized using detailed terms (motor task). The results of the hVB estimation with the uniform prior, the individual fMRI priors (CNR=0.5, 1, 3, 5, 10), the common truth prior, and the meta-analysis priors (synthesized using the terms "foot" and "hand") are reported. These priors are used with the best prior weight parameter denoted in parentheses. The horizontal line colored magenta indicates the aggregate score of the term "motor" with prior weight 1e-2. Vertical axes indicate the aggregate score (upper is better). The error bar represents standard error of the mean over ten different Monte Carlo repetitions for the data generation. We display results of a realistic simulation setting (p<0.05, SNR5) determined using experimental data. The results of wMNE and Champagne are also shown as a control.

F. Signal leakage analysis using resolution kernel

Because of the ill-posedness of the source reconstruction problem, there are linear relations between reconstructed sources. Therefore, signals are spread from real activity to other regions and cause spurious activations. This phenomenon is called signal leakage or source leakage (Brookes et al., 2012; Colclough et al., 2015; Palva et al., 2018; Sato et al., 2018). To consider the impact of signal

leakage, we employed the resolution kernel (Sekihara et al., 2005).

In the absence of noise, the forward model is $\boldsymbol{b} = \boldsymbol{Gj}$ using observations $\boldsymbol{b}$, lead field matrix $\boldsymbol{G}$, and real currents $\boldsymbol{j}$. Then, reconstructed currents $\bar{\boldsymbol{j}}$ is calculated as

$$\begin{aligned} \bar{\boldsymbol{j}} &= \boldsymbol{Lb} \\ &= \boldsymbol{LGj}, \end{aligned} \tag{F.1}$$

where $\boldsymbol{L}$ is an inverse filter. Here, we define $\boldsymbol{R} = \boldsymbol{LG}$ as the resolution kernel. As the above equation indicated, $\boldsymbol{R}$ tells us how the real currents spread in the reconstructed one. For example, the i-th column of this matrix indicates the signal spread of the i-th current of $\boldsymbol{j}$. Although ideal $\boldsymbol{R}$ is a diagonal matrix, it is impossible due to the ill-posedness.

We computed a resolution kernel derived with the meta-analysis fMRI prior ("visual") and plotted a column of the right ventral occipitotemporal cortex (VOTC). This region is a part of the fusiform gyrus and is known for the face-selective response (Jonas et al., 2016). As the figure displayed, many signals could be leaked from the right VOTC to the right insular cortex. It suggested that the reconstructed currents on the insula were leakage from the fusiform gyrus (Fig. 5.11). Note that similar results were obtained with the inverse filter of an individual fMRI prior.
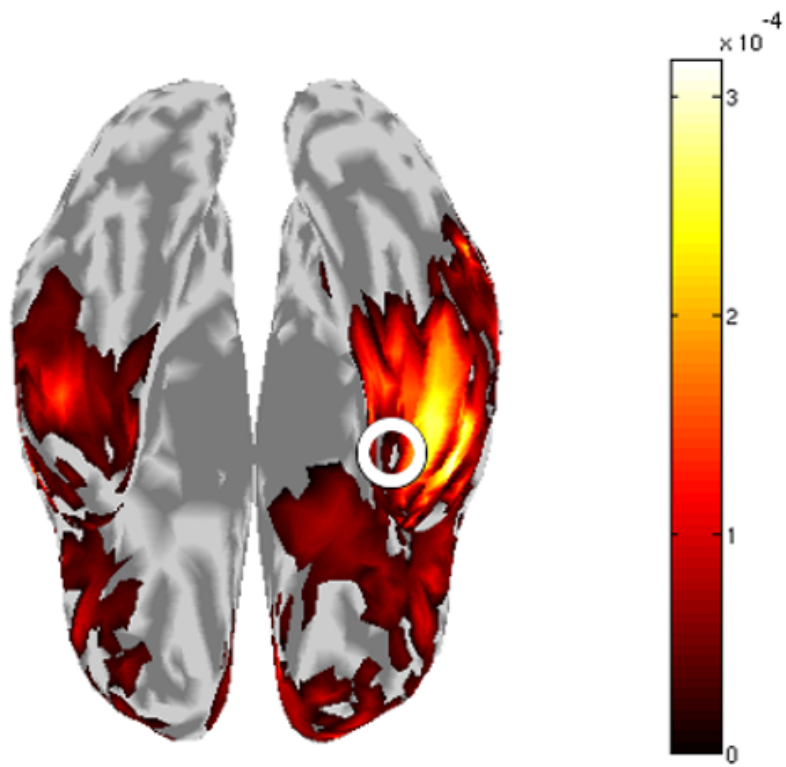
Figure F.1. Resolution kernel at the ventral occipitotemporal cortex is displayed on the cortex. Intensity indicates the degree of signal leakage from the right ventral occipitotemporal cortex (marked as a white circle). It is calculated using the resolution kernel derived with the meta-analysis fMRI prior ("visual"). This figure is of sub-15 of experimental data.

Acknowledgements

References

Ahlfors, S.P., Han, J., Lin, F.H., Witzel, T., Belliveau, J.W., Hämäläinen, M.S., Halgren, E., 2010. Cancellation of EEG and MEG signals generated by extended and distributed sources. Hum. Brain Mapp. 31, 140–149. `https://doi.org/10.1002/hbm.20851`

Baillet, S., 2017. Magnetoencephalography for brain electrophysiology and imaging. Nat. Neurosci. 20, 327–339. `https://doi.org/10.1038/nn.4504`

Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Process. Mag. 18, 14–30. `https://doi.org/10.1109/79.962275`

Baumgartner, R., Ryner, L., Richter, W., Summers, R., Jarmasz, M., Somorjai, R., 2000. Comparison of two exploratory data analysis methods for fMRI: Fuzzy clustering vs. principal Component analysis. Magn. Reson. Imaging 18, 89–94. `https://doi.org/10.1016/S0730-725X(99)00102-2`

Bertrand, Q., Massias, M., Gramfort, A., Salmon, J., 2019. Handling correlated and repeated measurements with the smoothed multivariate square-root Lasso, in: Advances in Neural Information Processing Systems. Curran Associates, Inc., pp. 3959–3970.

Bishop, C.M., 2006. Pattern recognition and machine learning, Information science and statistics. Springer, New York, NY.

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., Barnes, G.R., Bowtell, R., Brookes, M.J., 2018. Moving magnetoencephalography towards real-world applications with a wearable system. Nature 555, 657–661. `https://doi.org/10.1038/nature26147`

Breakspear, M., Terry, J.R., Friston, K.J., 2003. Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a

nonlinear model of neuronal dynamics. Neurocomputing 52–54, 151–158. `https://doi.org/10.1016/s0925-2312(02)00740-3`

Brookes, M.J., Woolrich, M.W., Barnes, G.R., 2012. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. Neuroimage 63, 910–920. `https://doi.org/10.1016/j.neuroimage.2012.03.048`

Cai, C., Hashemi, A., Diwakar, M., Haufe, S., Sekihara, K., Nagarajan, S.S., 2021. Robust estimation of noise for electromagnetic brain imaging with the champagne algorithm. Neuroimage 225, 117411. `https://doi.org/10.1016/j.neuroimage.2020.117411`

Colclough, G.L., Brookes, M.J., Smith, S.M., Woolrich, M.W., 2015. A symmetric multivariate leakage correction for MEG connectomes. Neuroimage 117, 439–448. `https://doi.org/10.1016/j.neuroimage.2015.03.071`

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

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., Louis, S., 2000. Dynamic Statistical Parametric Mapping : Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity. Neuron 26, 55–67. `https://doi.org/10.1016/S0896-6273(00)81138-1`

Dockès, J., Poldrack, R.A., Primet, R., Gözükan, H., Yarkoni, T., Suchanek, F., Thirion, B., Varoquaux, G., 2020. Neuroquery, comprehensive meta-analysis of human brain mapping. Elife 9, 1–34. `https://doi.org/10.7554/eLife.53385`

Drobyshevsky, A., Baumann, S.B., Schneider, W., 2006. A rapid fMRI task battery for mapping of visual, motor, cognitive, and emotional function. Neuroimage 31, 732–744. `https://doi.org/10.1016/j.neuroimage.2005.12.016`

Elliott, M., Knodt, A., Ireland, D., Morris, M., Poulton, R., Ramrakha, S., Sison, M., Moffitt, T., Caspi, A., Hariri, A., 2020. What is the Test-Retest Reliability of Common Task-fMRI Measures? New Empirical Evidence and a Meta-Analysis. Biol. Psychiatry 87, S132–S133. `https://doi.org/10.1016/j.biopsych.2020.02.356`

Endo, H., Hiroe, N., Yamashita, O., 2020. Evaluation of Resting Spatio-Temporal Dynamics of a Neural Mass Model Using Resting fMRI Connectivity and EEG Microstates. Front. Comput. Neurosci. 13, 1–11. `https://doi.org/10.3389/fncom.2019.00091`

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. Neuroimage 39, 1104–1120. `https://doi.org/10.1016/j.neuroimage.2007.09.048`

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302. `https://doi.org/10.1016/S1053-8119(03)00202-7`

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J. -P, Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp. 2, 189–210. `https://doi.org/10.1002/hbm.460020402`

Fukushima, M., Yamashita, O., Knösche, T.R., Sato, M. aki, 2015. MEG source reconstruction based on identification of directed source interactions on whole-brain anatomical networks. Neuroimage 105, 408–427. `https://doi.org/10.1016/j.neuroimage.2014.09.066`

Geissler, A., Gartus, A., Foki, T., Tahamtan, A.R., Beisteiner, R., Barth, M., 2007. Contrast-to-noise ratio (CNR) as a quality parameter in fMRI. J. Magn. Reson. Imaging 25, 1263–1270. `https://doi.org/10.1002/jmri.20935`

Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. 7, 555–562. `https://doi.org/10.1038/nn1224`

Grill-Spector, K., Weiner, K.S., Kay, K., Gomez, J., 2017. The Functional Neuroanatomy of Human Face Perception. Annu. Rev. Vis. Sci. 3, 167–196. https://doi.org/10.1146/annurev-vision-102016-061214

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M., Schoffelen, J.M., 2013. Good practice for conducting and reporting MEG research. Neuroimage 65, 349–363. https://doi.org/10.1016/j.neuroimage.2012.10.001

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. https://doi.org/10.1103/RevModPhys.65.413

Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42. https://doi.org/10.1007/BF02512476

He, B., Sohrabpour, A., Brown, E., Liu, Z., 2018. Electrophysiological Source Imaging: A Noninvasive Window to Brain Dynamics. Annu. Rev. Biomed. Eng. 20, 171–196. https://doi.org/10.1146/annurev-bioeng-062117-120853

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. https://doi.org/10.1002/hbm.20956

Henson, R.N., Mattout, J., Phillips, C., Friston, K.J., 2009. Selecting forward models for MEG source-reconstruction using model-evidence. Neuroimage 46, 168–176. https://doi.org/10.1016/j.neuroimage.2009.01.062

Hill, R.M., Boto, E., Holmes, N., Hartley, C., Seedat, Z.A., Leggett, J., Roberts, G., Shah, V., Tierney, T.M., Woolrich, M.W., Stagg, C.J., Barnes, G.R., Bowtell, R.R., Slater, R., Brookes, M.J., 2019. A tool for functional brain imaging with lifespan compliance. Nat. Commun. 10, 1–11. https://doi.org/10.1038/s41467-019-12486-x

Hill, R.M., Boto, E., Rea, M., Holmes, N., Leggett, J., Coles, L.A., Papastavrou, M., Everton, S.K., Hunt, B.A.E., Sims, D., Osborne, J., Shah, V., Bowtell, R., Brookes, M.J., 2020. Multi-channel whole-head OPM-MEG: Helmet design and a comparison with a conventional system. Neuroimage 219, 116995. `https://doi.org/10.1016/j.neuroimage.2020.116995`

Jas, M., Larson, E., Engemann, D.A., Leppäkangas, J., Taulu, S., Hämäläinen, M., Gramfort, A., 2018. A reproducible MEG/EEG group study with the MNE software: Recommendations, quality assessments, and good practices. Front. Neurosci. 12, 1–18. `https://doi.org/10.3389/fnins.2018.00530`

Kaneoke, Y., 2006. Magnetoencephalography: In search of neural processes for visual motion information. Prog. Neurobiol. 80, 219–240. `https://doi.org/10.1016/j.pneurobio.2006.10.001`

Kikuchi, M., Yoshimura, Y., Mutou, K., Minabe, Y., 2016. Magnetoencephalography in the study of children with autism spectrum disorder. Psychiatry Clin. Neurosci. 70, 74–88. `https://doi.org/10.1111/pcn.12338`

Laird, A.R., Fox, P.M., Eickhoff, S.B., Turner, J.A., Ray, K.L., McKay, D.R., Glahn, D.C., Beckmann, C.F., Smith, S.M., Fox, P.T., 2011. Behavioral interpretations of intrinsic connectivity networks. J. Cogn. Neurosci. 23, 4022–4037.

Larter, R., Speelman, B., Worth, R.M., 1999. A coupled ordinary differential equation lattice model for the simulation of epileptic seizures. Chaos 9, 795–804. `https://doi.org/10.1063/1.166453`

Lin, C.H., Tierney, T.M., Holmes, N., Boto, E., Leggett, J., Bestmann, S., Bowtell, R., Brookes, M.J., Barnes, G.R., Miall, R.C., 2019. Using optically pumped magnetometers to measure magnetoencephalographic signals in the human cerebellum. J. Physiol. 597, 4309–4324. `https://doi.org/10.1113/JP277899`

Lin, F.H., Belliveau, J.W., Dale, A.M., Hämäläinen, M.S., 2006. Distributed current estimates using cortical orientation constraints. Hum. Brain Mapp.

27, 1–13. `https://doi.org/10.1002/hbm.20155`

Neal, R.M., 1996. Bayesian Learning for Neural Networks. Springer-Verlag, Berlin, Heidelberg.

Ogawa, T., Aihara, T., Shimokawa, T., Yamashita, O., 2018. Large-scale brain network associated with creative insight: Combined voxel-based morphometry and resting-state functional connectivity analyses. Sci. Rep. 8, 1–11. `https://doi.org/10.1038/s41598-018-24981-0`

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. `https://doi.org/10.1016/j.neuroimage.2011.12.027`

Palva, J.M., Wang, S.H., Palva, S., Zhigalov, A., Monto, S., Brookes, M.J., Schoffelen, J.M., Jerbi, K., 2018. Ghost interactions in MEG/EEG source space: A note of caution on inter-areal coupling measures. Neuroimage 173, 632–643. `https://doi.org/10.1016/j.neuroimage.2018.02.032`

Pascual-Marqui, R.D., 2007. Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1 exact, zero error localization 1–16.

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA) technical details. Methods Find. Exp. Clin. Pharmacol.

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. `https://doi.org/https://doi.org/10.1016/0167-8760(84)90014-X`

Patterson, H.D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545–554. `https://doi.org/10.1093/biomet/58.3.545`

Pernet, C.R., 2014. Misconceptions in the use of the General Linear Model applied to functional MRI: A tutorial for junior neuro-imagers. Front. Neurosci. 8, 1–12. `https://doi.org/10.3389/fnins.2014.00001`

Rojas, D.C., 2019. Review of schizophrenia research using MEG. Magnetoencephalography From Signals to Dyn. Cortical Networks Second Ed. 1121–1146. `https://doi.org/10.1007/978-3-030-00087-5_41`

Rossion, B., Jacques, C., Jonas, J., 2018. Mapping face categorization in the human ventral occipitotemporal cortex with direct neural intracranial recordings. Ann. N. Y. Acad. Sci. 1426, 5–24. `https://doi.org/10.1111/nyas.13596`

Sato, M.A., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical Bayesian estimation for MEG inverse problem. Neuroimage 23, 806–826. `https://doi.org/10.1016/j.neuroimage.2004.06.037`

Sato, Masashi, Yamashita, O., Sato, Masa aki, Miyawaki, Y., 2018. Information spreading by a combination of MEG source estimation and multivariate pattern classification. PLoS One 13, 1–28. `https://doi.org/10.1371/journal.pone.0198806`

Sekihara, K., Nagarajan, S.S., 2015. Electromagnetic brain imaging: a bayesian perspective. Springer.

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. `https://doi.org/10.1016/j.neuroimage.2004.11.051`

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. 106, 13040–13045. `https://doi.org/10.1073/pnas.0905267106`

Stefan, H., Trinka, E., 2017. Magnetoencephalography (MEG): Past, current and future perspectives for improved differentiation and treatment of epilepsies. Seizure 44, 121–124. `https://doi.org/10.1016/j.seizure.2016.10.028`

Storey, J.D., Tibshirani, R., 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 100, 9440–9445. `https://doi.org/10.1073/pnas.1530509100`

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: A user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011. `https://doi.org/10.1155/2011/879716`

Takeda, Y., Suzuki, K., Kawato, M., Yamashita, O., 2019. MEG Source Imaging and Group Analysis Using VBMEG. Front. Neurosci. 13, 1–12. `https://doi.org/10.3389/fnins.2019.00241`

Thompson, W.H., Fransson, P., 2017. Spatial confluence of psychological and anatomical network constructs in the human brain revealed by a mass meta-analysis of fMRI activation. Sci. Rep. 7, 1–11. `https://doi.org/10.1038/srep44259`

Tierney, T., Levy, A., Barry, D., Meyer, S., Shigihara, Y., Everatt, M., Mellor, S., Lopez, J.D., Bestmann, S., Holmes, N., Roberts, G., Hill, R., Boto, E., Leggett, J., Shah, V., Brookes, M., Bowtell, R., Maguire, E., Barnes, G., 2020. Mouth magnetoencephalography: A unique perspective on the human hippocampus 17–19. `https://doi.org/10.1101/2020.03.19.998641`

Uhlhaas, P.J., Liddle, P., Linden, D.E.J., Nobre, A.C., Singh, K.D., Gross, J., 2017. Magnetoencephalography as a Tool in Psychiatric Research: Current Status and Perspective. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 2, 235–244. `https://doi.org/10.1016/j.bpsc.2017.01.005`

Valente, G., De Martino, F., Filosa, G., Balsi, M., Formisano, E., 2009. Optimizing ICA in fMRI using information on spatial regularities of the sources. Magn. Reson. Imaging 27, 1110–1119. `https://doi.org/10.1016/j.mri.2009.05.036`

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., 2012. The Human Connectome Project: A data acquisition perspective. Neuroimage 62, 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

Wager, T.D., Lindquist, M.A., Nichols, T.E., Kober, H., Van Snellenberg, J.X., 2009. Evaluating the consistency and specificity of neuroimaging data using meta-analysis. Neuroimage 45, S210–S221. https://doi.org/10.1016/j.neuroimage.2008.10.061

Wakeman, D.G., Henson, R.N., 2015. A multi-subject, multi-modal human neuroimaging dataset. Sci. data 2, 150001. https://doi.org/10.1038/sdata.2015.1

Wang, J.-., Williamson, S.J., Kaufman, L., 1992. Magnetic source images determined by a lead-field analysis: the unique minimum-norm least-squares estimation. IEEE Trans. Biomed. Eng. 39, 665–675. https://doi.org/10.1109/10.142641

Welvaert, M., Rosseel, Y., 2013. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. PLoS One 8. https://doi.org/10.1371/journal.pone.0077089

Wipf, D., Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. Neuroimage 44, 947–966. https://doi.org/10.1016/j.neuroimage.2008.02.059

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. https://doi.org/10.1016/j.neuroimage.2009.06.083

Wu, A., Koyejo, O., Pillow, J., 2019. Dependent relevance determination for smooth and structured sparse regression. J. Mach. Learn. Res. 20, 1–43.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods 8, 665–670. `https://doi.org/10.1038/nmeth.1635`

Yoshioka, T., Toyama, K., Kawato, M., Yamashita, O., Nishina, S., Yamagishi, N., Sato, M.A., 2008. Evaluation of hierarchical Bayesian method through retinotopic brain activities reconstruction from fMRI and MEG signals. Neuroimage 42, 1397–1413. `https://doi.org/10.1016/j.neuroimage.2008.06.013`

Publication list

Journal papers

Takeda, Y., Suzuki, K., Kawato, M., Yamashita, O., 2019. MEG Source Imaging and Group Analysis Using VBMEG. Front. Neurosci. 13, 1–12. `https://doi.org/10.3389/fnins.2019.00241`

Suzuki, K., Yamashita, O., 2021. MEG current source reconstruction using a meta-analysis fMRI prior. Neuroimage 236, 118034. `https://doi.org/10.1016/j.neuroimage.2021.118034`

Conference (Refereed)

Suzuki, K., Yamashita, O., 2018. Meta-analysis fMRI data helps MEG current estimation. Organization for Human Brain Mapping (OHBM 2018). June 2018.

Other publications and conferences (Not Refereed)

Suzuki, K., Yamashita, O., 2017. 脳内の構造的結合情報を利用した MEG 電流源推定手法の提案. 第一回ヒト脳イメージング研究会. September 2017. (In Japanese)

Suzuki, K., Yamashita, O., 2018. 構造的ネットワークを制約に加えた安静時電流源推定法の提案. IEEE CISJ (Neurocomputing). February 2018. (In Japanese)

Suzuki, K., Yamashita, O., 2018. メタ分析 fMRI データを制約に用いる MEG 電流源推定方法の提案. Japan Human Brain Mapping Society. March 2018. (In Japanese)

Suzuki, K., Yamashita, O., 2019. メタ解析 fMRI データを空間制約情報に用いた MEG 信号源推定方法. IEEE CISJ (Neurocomputing). June 2019. (In Japanese)

Suzuki, K., Yamashita, O., 2019. MEG group analysis using common fMRI activity among subjects. APSIPA BioSiPS Workshop. August 2019.

Award

IEEE Computational Intelligence Society Japan Chapter Young Researcher Award (Neurocomputing). March 2020.