



## Estimation of hyper-parameters for a hierarchical model of combined cortical and extra-brain current sources in the MEG inverse problem



Ken-ichi Morishige<sup>a,b,\*</sup>, Taku Yoshioka<sup>c</sup>, Dai Kawawaki<sup>d</sup>, Nobuo Hiroe<sup>d</sup>, Masa-aki Sato<sup>d</sup>, Mitsuo Kawato<sup>a,e</sup>

<sup>a</sup> Department of Intelligent Systems Design Engineering, Toyama Prefectural University, 5180 Kurokawa, Imizu-shi, Toyama 939-0398, Japan

<sup>b</sup> Cognitive Mechanisms Laboratories, Advanced Telecommunications Research Institute International, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan

<sup>c</sup> Chinou Jouhou Shisutemu Inc., 134 Chudoji-Minami-cho, Shimogyo-ku, Kyoto-shi, Kyoto 600-8813, Japan

<sup>d</sup> Neural Information Analysis Laboratories, Advanced Telecommunications Research Institute International, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan

<sup>e</sup> Brain Information Communication Research Laboratory Group, Advanced Telecommunications Research Institute International, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan

### ARTICLE INFO

#### Article history:

Accepted 6 July 2014

Available online 14 July 2014

#### Keywords:

Denoising

Hierarchical Bayesian method

MEG

Covert pursuit

### ABSTRACT

One of the major obstacles in estimating cortical currents from MEG signals is the disturbance caused by magnetic artifacts derived from extra-cortical current sources such as heartbeats and eye movements. To remove the effect of such extra-brain sources, we improved the hybrid hierarchical variational Bayesian method (hyVBED) proposed by Fujiwara et al. (NeuroImage, 2009). hyVBED simultaneously estimates cortical and extra-brain source currents by placing dipoles on cortical surfaces as well as extra-brain sources. This method requires EOG data for an EOG forward model that describes the relationship between eye dipoles and electric potentials. In contrast, our improved approach requires no EOG and less a priori knowledge about the current variance of extra-brain sources. We propose a new method, “extra-dipole,” that optimally selects hyper-parameter values regarding current variances of the cortical surface and extra-brain source dipoles. With the selected parameter values, the cortical and extra-brain dipole currents were accurately estimated from the simulated MEG data. The performance of this method was demonstrated to be better than conventional approaches, such as principal component analysis and independent component analysis, which use only statistical properties of MEG signals. Furthermore, we applied our proposed method to measured MEG data during covert pursuit of a smoothly moving target and confirmed its effectiveness.

© 2014 Elsevier Inc. All rights reserved.

### Introduction

One major obstacle in estimating cortical currents from MEG signals is the disturbance caused by magnetic artifacts derived from extra-cortical current sources such as heartbeats and eye movements. These artifacts can be orders of magnitude larger than signals from the brain. Thus, cortical current estimation is a difficult inverse problem with an extremely low signal-to-noise ratio. Principal component analysis (PCA) and independent component analysis (ICA) have been applied to remove the influence of extra-brain sources (Barbati et al., 2004; Bell and Sejnowski, 1995; Comon, 1994; Jung et al., 2000; Vigario, 1997; Vigario et al., 2000). Although these statistical methods are efficient in many circumstances, removing artifacts correlated with brain signals is, in principle, difficult. This is because these methods rely on either the orthogonality or statistical independence between cortical and extra-brain time-series signals, which may not be satisfied in many interesting situations. Additionally, these methods face difficulty in determining which principal and independent components are

“artifacts,” because the selection process of artifact components is often subjective.

Fujiwara et al. proposed the hybrid hierarchical variational Bayesian method (hyVBED, 2009), which simultaneously estimates not only cortical currents but also eye currents, thus giving it the ability to objectively isolate these currents based on a physical and physiological forward model rather than ad hoc statistical assumptions. Consequently, artifact components can be objectively selected and removed because they depend on physical dipole locations. However, this method requires just as many additional electrical or magnetic data as extra-brain sources for preparing the forward models. Measured MEG data are affected not only by eye movements but also by heartbeats and muscle activations. If all extra-brain components are to be removed, the hyVBED method requires so many electrical or magnetic sensors to be simultaneously placed on a subject's body that it becomes unrealistic and impractical as a denoising method for clinical applications (e.g., for a purpose of brain–machine interfaces).

In general hierarchical Bayesian frameworks, some hyper-parameters must be predetermined. They are usually selected under a constraint that maximizes the marginal likelihood or free energy. In the early stage of our research, we tried to find an optimal hyper-parameter set using this method. It was possible to select proper

\* Corresponding author at: 5180 Kurokawa Imizu-shi, Toyama 939-0398, Japan. Fax: +81 766 56 8030.

E-mail address: [kmorishi@pu-toyama.ac.jp](mailto:kmorishi@pu-toyama.ac.jp) (K. Morishige).

hyper-parameter sets for relatively large current sources such as eye and heartbeat artifacts, but selecting them for such small current sources as brain activities was difficult. Although the free-energy approach is theoretically elegant, it is neither robust nor practical for MEG data. For such reasons, we need to develop alternative methods to help us find the optimal hyper-parameters for the MEG inverse problem.

This paper proposes a less demanding, thus more practical, denoising method, “extra-dipole,” to overcome these difficulties associated with either statistical or physical/physiological methods. Our approach requires no additional sensor data other than MEG, and generally necessitates less a priori knowledge about the current variance of extra-brain sources. Instead of statistical and physical/physiological information, we used the fMRI spatial activation patterns as informative signals to find the optimal hyper-parameters. With these values, we evaluated our proposed method by estimating the cortical and extra-brain dipole currents from simulated and measured MEG data. By introducing appropriate constraints on prior information, our approach obtained a more accurate estimate of cortical and extra-brain currents from obfuscated MEG signals than previous methods. We especially demonstrated better performance of this method compared with conventional statistical methods, such as PCA and ICA, as well as the original variational Bayesian method of Sato et al. (2004), named VBMEG in this paper.

**Materials and methods**

*Cortical and extra-brain dipole current model*

Cortical electrical currents associated with neuronal activities induce weak transient magnetic fields. These current sources were modeled using  $I$  current dipoles with fixed positions and orientations, and these current moments were expressed as  $I$ -dimensional vector  $\mathbf{J}_{\text{brain}}$ . The

observed magnetic fields were generated by not only cortical currents but also various extra-brain source currents. Each extra-brain source was modeled using the three-dimensional resultant dipole current in the  $x$ - $y$ - $z$  direction. The number of extra-brain sources was  $K$ , and the extra-brain source was modeled using  $3K$  dipoles, and these current moments were expressed as  $3K$ -dimensional vector  $\mathbf{J}_{\text{artifact}}$ . When observation noise was ignored, the relationship between dipole currents and observed magnetic fields was expressed as follows:

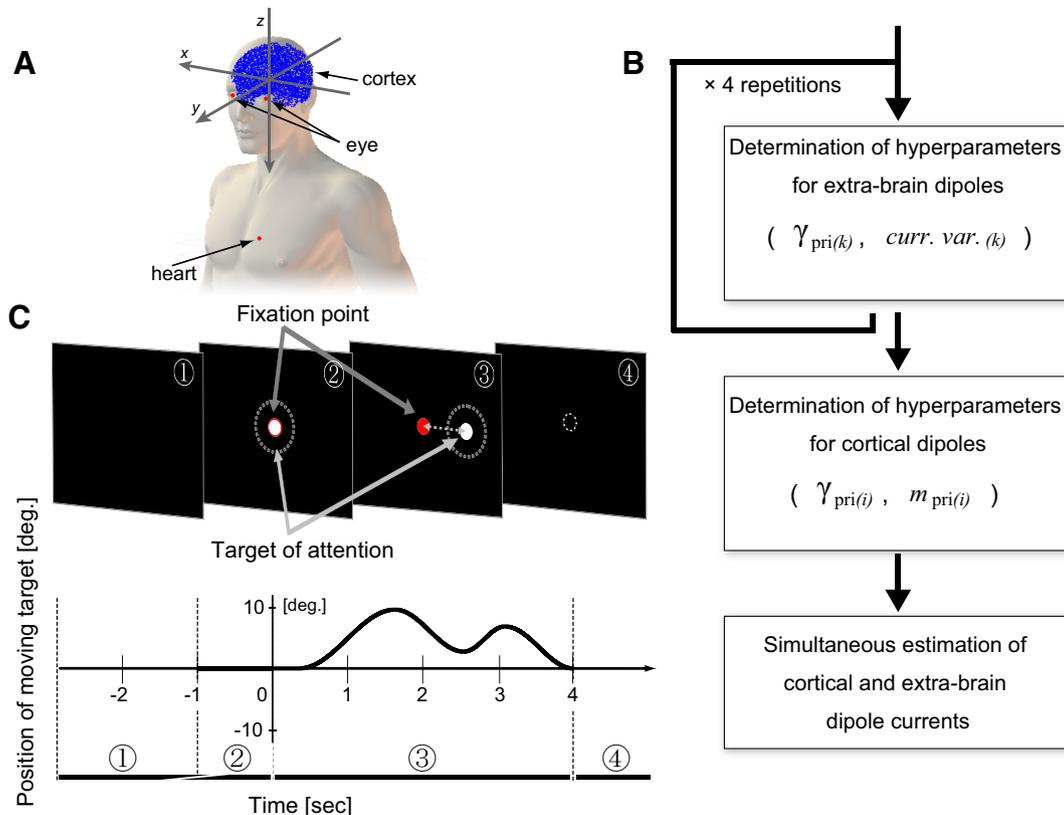
$$\mathbf{B}(t) = \mathbf{G}_{\text{brain}} \mathbf{J}_{\text{brain}}(t) + \mathbf{G}_{\text{artifact}} \mathbf{J}_{\text{artifact}}(t), \tag{1}$$

where  $\mathbf{B}(t)$  is an  $M$ -element vector for the magnetic field on  $M$  MEG channels.  $\mathbf{G}_{\text{brain}}$  denotes the lead field matrix for cortical current sources, which also represents in its row a magnetic field generated by a unit current moment. This lead field matrix was calculated by the linear Galerkin method using the boundary element method (Ferguson et al., 1994). The dipole current directions were assumed to be perpendicular to the cortical surface. A positive current was defined as one directed toward the inside of the cortex (Fig. 1A). These settings are based on knowledge that cortical currents induced by pyramidal cells play the major role of generating MEG signals that originate from brain activities.

We considered three extra-brain sources, the left and right eyeballs, and the heart, as the main noise source; nine dipoles (three extra-brain sources  $\times$  three directions) were thus located there. The lead field matrix for the eye and heart current sources  $\mathbf{G}_{\text{artifact}}$  was calculated by the Biot-Savart law.

*Hierarchical Bayesian estimation*

To estimate cortical and extra-brain source currents at the dipole locations, hierarchical Bayesian estimation (Variational Bayesian Multimodal Encephalography: VBMEG) has been used (Sato et al., 2004; Yoshioka et al., 2008). This method uses functional magnetic resonance



**Fig. 1.** (A) Cortical and extra-brain source dipole locations. (B) Experimental protocol. (C) Flow of selecting optimal hyper-parameters. Red and white circles show the fixation point and target of attention, respectively.

imaging (fMRI) information as prior knowledge for source estimation and yields more accurate source localization. Additionally, not only cortical but also extra-brain currents can be estimated simultaneously with prior information about the current intensities of extra-brain source currents.

We expressed current distributions of the  $L$  dipoles located on the cortical surface and the extra-brain sources as  $\mathbf{J} = [\mathbf{J}_{\text{brain}}, \mathbf{J}_{\text{artifact}}] = \{J_l | l = 1, \dots, L = I + 3K\}$ . In hierarchical Bayesian estimation, a Gaussian probability distribution  $P(\mathbf{B}|\mathbf{J})$ , where magnetic field  $\mathbf{B}$  is observed for a given current  $\mathbf{J}$ , is given by

$$P(\mathbf{B}|\mathbf{J}) \propto \exp\left[-\frac{\beta}{2} \|\mathbf{B} - \mathbf{G} \cdot \mathbf{J}\|^2\right], \quad (2)$$

where  $\beta$  and  $\mathbf{G}$  denote the inverse observation-noise variance and the lead field matrix  $\mathbf{G} = [\mathbf{G}_{\text{brain}}, \mathbf{G}_{\text{artifact}}]$ , respectively. The posterior probability distribution for source currents  $P_{\text{post}}(\mathbf{J}, \boldsymbol{\alpha}|\mathbf{B})$ , which represents the probability of source current  $\mathbf{J}$  and inverse current variance  $\boldsymbol{\alpha}$  under observed MEG data  $\mathbf{B}$  and prior information, is expressed as follows:

$$P_{\text{post}}(\mathbf{J}, \boldsymbol{\alpha}|\mathbf{B}) = \frac{P(\mathbf{B}|\mathbf{J})P_{\text{pri}}(\mathbf{J}|\boldsymbol{\alpha})P_{\text{pri}}(\boldsymbol{\alpha})}{P(\mathbf{B})}, \quad (3)$$

where  $P(\mathbf{B})$  denotes a marginal likelihood.  $P_{\text{pri}}(\mathbf{J}|\boldsymbol{\alpha})P_{\text{pri}}(\boldsymbol{\alpha})$  is the hierarchical prior distribution for the source current that is determined by the prior current variance.

The estimation results using the hierarchical Bayesian method depend largely on prior information that is given by a hierarchical prior distribution with two kinds of hyper-parameters. This implies that the selection of the two hyper-parameters must largely influence whether the estimation results are good or bad. Hierarchical prior distributions are introduced to the source current dipoles as follows:

$$P_{\text{pri}}(\mathbf{J}(t)|\boldsymbol{\alpha}_{\text{pri}}, \beta) \propto \exp\left[-\frac{\beta}{2} \mathbf{J}(t)^T \mathbf{A} \mathbf{J}(t)\right], \quad (4)$$

$$P_{\text{pri}}(\boldsymbol{\alpha}_{\text{pri}}) = \prod \text{Gamma}(\alpha_{\text{pri}(l)} | \bar{\alpha}_{\text{pri}(l)}, \gamma_{\text{pri}(l)}), \quad (5)$$

$$\text{Gamma}(\alpha | \bar{\alpha}, \gamma) \equiv \alpha^{-1} \left(\frac{\alpha \gamma}{\bar{\alpha}}\right)^\gamma \Gamma(\gamma)^{-1} e^{-\frac{\alpha \gamma}{\bar{\alpha}}}, \quad (6)$$

$$P_{\text{pri}}(\beta) = \frac{1}{\beta}, \quad (7)$$

where  $\mathbf{A}$  is a diagonal matrix with elements  $\boldsymbol{\alpha}_{\text{pri}} = \{\alpha_{\text{pri}(l)} | l = 1, \dots, L\}$ .  $\alpha_{\text{pri}(l)}$  is the inverse current variance corresponding to the  $l$ -th current dipole.  $\Gamma(\gamma)$  represents the Gamma function. A prior distribution of current variance  $\nu_{\text{pri}(l)} = \alpha_{\text{pri}(l)}^{-1}$  represents the prior distribution of  $J_l$  with mean current variance  $\bar{\nu}_{\text{pri}(l)} = \bar{\alpha}_{\text{pri}(l)}^{-1}$  and degree of freedom  $\gamma_{\text{pri}(l)}$ .

Hyper-parameter  $\bar{\nu}_{\text{pri}(i)}$  ( $i = 1, \dots, I$ ), which expresses the mean current variance of the cortical dipole, is defined using fMRI activation ( $t$ -value) and magnification parameter  $m_{\text{pri}}$  (Yoshioka et al., 2008).  $i$  denotes the index of the cortical dipole.  $\bar{\nu}_{\text{pri}(i)}$  is a monotonic increasing function of the fMRI  $t$ -value at each dipole. When the baseline variance is  $\nu_{\text{base}}$ ,  $\bar{\nu}_{\text{pri}(i)}$  is expressed as follows:

$$\bar{\nu}_{\text{pri}(i)} = \nu_{\text{base}} + (m_{\text{pri}} - 1) \nu_{\text{base}} f_{(i)}^2, \quad (8)$$

where  $f_{(i)}$  is a normalized fMRI-dipole  $t$ -value with a maximum value of 1. Consequently,  $\bar{\nu}_{\text{pri}(i)}$  takes values from  $\nu_{\text{base}}$  to  $m_{\text{pri}} \nu_{\text{base}}$ .

$\bar{\nu}_{\text{pri}(k)}$  ( $k = 1, \dots, 3K$ ) denotes the mean current variances of extra-brain dipoles located on the left and right eyeballs and the heart.  $k$  is the index of an extra-brain dipole.

In our previous proposed method, VBMEG and hyVBED/VBED, we empirically defined the hyper-parameters for cortical dipoles ( $m_{\text{pri}} = 100 \sim 500$ ,  $\gamma_{\text{pri}(i)} = 500, 100, 10$  or  $0$ ) (Aihara et al., 2012; Fujiwara et al., 2009; Sato et al., 2004; Shibata et al., 2007; Toda et al., 2011; Yoshimura et al., 2012). Unlike cortical currents, fMRI activation is not available for prior information of eye currents. Because it provides less knowledge about eye currents, hyVBED/VBED used a non-informative prior information for the eye dipoles. In contrast to these previous studies, a method must be developed not based on practical experience but on observed data to provide more precise estimates of dipole currents.

#### Hyper-parameters for extra-brain sources

The actual current intensities for extra-brain dipoles are generally unknown to MEG users. To determine the hyper-parameters for current variances of extra-brain source dipoles, we adopted the following two criteria. Criterion 1: Prior and posterior variances for each extra-brain source dipole should be the same, and Criterion 2: a high correlation coefficient should be obtained between the spatial patterns of fMRI activities and MEG estimation as the magnitude of the estimated mean current variance. A unique parameter can be determined so that it maximizes the correlation coefficient Criterion 2, among the parameter set that satisfies condition Criterion 1. The two criteria are based on the following assumptions:

**Criterion 1.** The prior and posterior variances for each extra-brain dipole should be the same.

Our proposed method assumed that the prior and posterior current variances ( $\nu_{\text{pri}(k)}$  and  $\nu_{\text{post}(k)}$ ) should have the same order of magnitude if prior information on the current intensities is reasonable. In other words, if the a priori information about the solution (prior variance) is appropriate, the obtained solution (posterior variance) should be the same order of magnitude as the a priori knowledge. By varying two hyper-parameters, prior current variance and  $\gamma_{\text{pri}(k)}$ , for each extra-brain source, we estimated the dipole currents and calculated the index as follows:

$$O_{\text{pri}(k)} = \log_{10}(\nu_{\text{pri}(k)}), \quad (9)$$

$$O_{\text{post}(k)} = \log_{10}(\nu_{\text{post}(k)}), \quad (10)$$

$$O_{\text{artifact}(k)} = O_{\text{post}(k)} - O_{\text{pri}(k)}, \quad (11)$$

where  $O_{\text{pri}(k)}$  and  $O_{\text{post}(k)}$  are the orders of magnitude for the prior and posterior current variances, respectively.  $O_{\text{artifact}(k)}$  is the difference between the prior and posterior current variances in the order of magnitude.  $\nu_{\text{pri}(k)}$  and  $\nu_{\text{post}(k)}$  are the parameters for the prior and posterior current variances. Note that  $\nu_{\text{post}}$  depends on prior information  $\nu_{\text{pri}}$  and  $\gamma_{\text{pri}}$  because  $\nu_{\text{post}}$  is estimated from MEG data using hierarchical Bayesian estimation. Several hyper-parameter sets of near-zero  $O_{\text{artifact}(k)}$  were selected and became candidates for the optimal hyper-parameter set.

**Criterion 2.** The spatial patterns of the fMRI activities and the estimated current power should be highly correlated.

fMRI activation areas are expected to have high correlations with neuronal activities. Here, we assume that the spatial pattern of fMRI is a reliable information source for a spatial pattern of estimated source currents for MEG data. The correlation coefficients between fMRI

activity patterns  $f_{(i)}$  and current intensities  $C_{(i)}$  were calculated as follows:

$$C_{(i)} = \sqrt{\frac{1}{T} \sum_{t=1}^T \hat{Z}_{(i)}^2(t)}, \quad (12)$$

$$\text{corr} = \frac{\sum_{i=1}^I (C_{(i)} - \bar{C})(f_{(i)} - \bar{f})}{\sqrt{\sum_{i=1}^I (C_{(i)} - \bar{C})^2} \sqrt{\sum_{i=1}^I (f_{(i)} - \bar{f})^2}}, \quad (13)$$

where  $\hat{Z}$  is a trial average of the estimated cortical currents (Supplementary information 1; The design of the “trial” is explained in more detail in the subsection “Task settings”),  $T$  is the sampling number, and  $\bar{f}$  and  $\bar{C}$  are the trial averages of fMRI activations  $\{f_{(i)} | i = 1, \dots, I\}$  and current intensities  $\{C_{(i)} | i = 1, \dots, I\}$ , respectively. Correlation coefficients  $\text{corr}$  were calculated using trial average values of the estimated cortical current and fMRI activations, and several hyper-parameter sets of  $\text{corr}$  that were as large as possible were selected and became candidates for the optimal hyper-parameter set.

In a nested and recursive parameter search algorithm shown in Fig. 1B, to find the optimal hyper-parameter set that satisfies Criterion 1 and maximizes Criterion 2 we set the obtained parameter set as initial values and repeated the parameter search process. This process required four repetitions until the obtained parameter set converged to the same values. Then, we regarded our obtained parameter set as the optimal hyper-parameter set for the extra-brain source dipoles. In general, current amplitudes of the eyes and heart are considered to be about 10 and 1000 nAm, respectively, with current variances of  $10^2$  and  $10^6$  [nAm] $^2$ , respectively (Katila et al., 1981; Nousiainen et al., 1986). These parameter search ranges for the eyes and heartbeat source dipoles were set sufficiently wide to contain the expected value. Their search ranges were  $10^{0.5}$ – $10^5$  [nAm] $^2$  and  $10^4$ – $10^{8.5}$  [nAm] $^2$ , respectively. A preliminary analysis of our data indicates that an overly large  $\gamma_{\text{pri}}$  (e.g.,  $>10^7$ ) causes identical prior and posterior current variances on a constant basis irrespective of prior current variance. To avoid this, the range of  $\gamma_{\text{pri}(k)}$  was changed to  $10^1$ – $10^6$ , and hierarchical Bayesian estimation was conducted using hyper-parameters as current variances and  $\gamma_{\text{pri}(k)}$  (prior current variance: 10 types,  $\gamma_{\text{pri}(k)}$ : 14 types, hyper-parameter total: 140 sets were the target of the global search).

#### Hyper-parameters for cortical current dipoles

For the cortical current dipoles, we also simultaneously selected hyper-parameter sets to satisfy Criterion 1 and 2. However, unlike extra-brain source dipoles, the cortical surface has many current dipoles. To reflect the difference, we revised Criterion 1:

**Criterion 1'.** The prior and posterior variances for each cortical dipole should be the same.

By varying hyper-parameters magnification parameter  $m_{\text{pri}}$  and reliability  $\gamma_{\text{pri}(i)}$  for the cortex, we estimated the dipole currents and calculated the index as follows:

$$O_{\text{pri}(i)} = \log_{10}(\nu_{\text{pri}(i)}), \quad (14)$$

$$O_{\text{post}(i)} = \log_{10}(\nu_{\text{post}(i)}), \quad (15)$$

$$O_{\text{brain}} = \frac{1}{I} \sum_{i=1}^I (O_{\text{post}(i)} - O_{\text{pri}(i)}), \quad (16)$$

where  $O_{\text{brain}}$  is the mean difference between the prior and posterior current variances in the order of magnitude. Several hyper-parameter sets of near-zero  $O_{\text{brain}}$  were selected and became candidates for the optimal hyper-parameter set.

The hyper-parameter set was selected to satisfy the Criterion 1' and maximizes Criterion 2 among that parameter set. The selected parameter set was regarded as optimal for the cortical dipoles. Hierarchical Bayesian estimation was conducted using hyper-parameters as  $m_{\text{pri}}$  and  $\gamma_{\text{pri}(i)}$ , and their search ranges were  $10^0$ – $10^6$  and  $10^1$ – $10^6$ , respectively ( $m_{\text{pri}}$ : 9 types,  $\gamma_{\text{pri}(i)}$ : 11 types, hyper-parameter total: 99 sets). The  $m_{\text{pri}}$  search range for the cortical dipoles was set sufficiently wide to contain values 10–500, which were commonly used in previous VBMEG applications (Aihara et al., 2012; Callan et al., 2010; Fujiwara et al., 2009; Sato et al., 2004; Shibata et al., 2007; Toda et al., 2011; Yoshimura et al., 2012; Yoshioka et al., 2008).

#### Flow of selecting hyper-parameters

Ideally, the optimal hyper-parameter set based on the two criteria should be prepared for each current dipole of the cortex and the extra-brain source. However, this process requires a huge amount of computational time because the total number of current dipoles exceeds 10,000. To reduce time, some approximation is required. Only one hyper-parameter set  $m_{\text{pri}}$  and  $\gamma_{\text{pri}}$  is prepared for many cortical dipoles. Furthermore, each optimal parameter set is determined in the order of extra-brain source and cortical dipoles (Fig. 1B), because the extra-brain sources have larger current intensities than the cortical ones (Hämäläinen et al., 1993). This sequential search method is expected to provide quick convergence to the same values compared with a simultaneous search for all hyper-parameter sets. Because the optimal hyper-parameters for the cortex were not yet determined on the selection step for the extra-brain source dipoles, tentative values were used ( $m_{\text{pri}} = 500$ ,  $\gamma_{\text{pri}(i)} = 500$ ; Yoshioka et al., 2008).

#### Evaluation of a selected magnification parameter using simulated MEG data

When evaluating the selected prior current variances for the extra-brain source dipoles from the simulated MEG data, it is only necessary to compare the actual and estimated current variances directly. However, because  $m_{\text{pri}}$  is not a prior current variance for the cortical dipoles but a magnification parameter of the fMRI activations ( $t$ -values), we must calculate the true  $m_{\text{pri}}$  value.

$J_{\text{var}}$  denotes the maximum value of the actual cortical current variances. We assume that the posterior current variance meets value consistency with  $J_{\text{var}}$ . The posterior current variance also satisfies the optimality criterion where prior and posterior current variances should be the same. The effects of the smoothing spatial filters are ignored. Under these conditions, the true value of  $m_{\text{pri}}$  is expressed as follows:

$$m_{\text{pri}} = J_{\text{var}} / (\nu_{\text{base}} \times \rho). \quad (17)$$

Because we calculated  $m_{\text{pri}}$  with Eq. (17), it became a true value. We then used it to evaluate the selected hyper-parameter set for the cortical dipoles.

#### Subjects

Three men (aged 27–30 years with normal or corrected-to-normal visual acuity) participated in both the MEG and fMRI experiments, which were approved by the Ethics and Safety Committee of Advanced Telecommunications Research Institute International (ATR). Written informed consent was obtained from each subject.

### Task settings

The subjects covertly pursued the target motion (white circle in Fig. 1C) with eyes gazed at the fixation point (red circle in Fig. 1C). We prepared three target motion patterns. The target motions were approximately sinusoidal, and its frequency was changed when the target velocity was zero (Supplementary information 2). The mean sinusoidal frequency was 0.5 Hz, and the peak velocity was  $10^\circ/\text{s}$ . Peak-to-peak amplitudes of the three target motions were 8.9, 8.9, and  $9.5^\circ$ , respectively. Subjects pressed a start button when they were ready (Fig. 1C (1)). The fixation point and target of attention were presented in the center of the monitor. The subjects were required to maintain fixation after the targets were presented for a variable length of time in the range of 0.5–1.5 s (Fig. 1C (2)). Then, the target of attention started to move in a horizontal direction. The subjects began to pursue the target of attention covertly (Fig. 1C (3)). The fixation point and target of attention disappeared after 4 s of moving, and subjects then took a short rest (Fig. 1C (4)). These processes were regarded as a “trial”. One session consisted of 30 repetitions of a trial for three target motions. Each subject performed 270 trials (30 trials  $\times$  3 moving target motions  $\times$  3 sessions = 270 trials). The three target motions were randomly presented within a session.

### Dipole locations of cortical and extra-brain current sources

We constructed a polygon cortical surface model for each subject using FreeSurfer software (version 5.0.0; <http://surfer.nmr.mgh.harvard.edu/>) (Dale et al., 1999) with a T1 structural image for the subjects. The number of cortical surface dipoles of subjects A, B, and C was 20,004. The cortical current sources were located on the vertex points of the cortical surface model, and the orientation of the current sources was perpendicular to the cortical surface. A positive current was defined as one directed toward the inside of the cortex (Yoshioka et al., 2008).

The eyeball has a steady electric polarity with a positive charge at the cornea and a negative charge at the retina (Miles, 1939). When the eyeball rotates, the rotation induces currents inside it. To express the currents that arise from the three-dimensional rotation, each eyeball current source can be described as  $x$ -,  $y$ -, and  $z$ -axial dipoles. We obtained the positions of each eyeball by visual inspection from T1 structural images and placed three dipoles at the center of each eyeball.

The main noise sources from the heartbeats were assumed to be the heart's arterial side. The precise location from the origin at the center of the brain is not known because it depends on the neck angle. Therefore, approximate positional values ( $(x, y, z) = (1.5, 0.0, 35.0)$  [cm]) were set as the main noise source of the heart's location.

### MEG

We used a whole-head, 208-channel system (MEG vision-PQ1400RM; Yokogawa Electric Co., Japan) for the MEG recordings at a 1000-Hz sampling frequency. Electro-oculograms (EOGs) were simultaneously recorded. The slow DC drift components of the observed MEG data were removed using reference sensors. For each subject, each trial was extracted from  $-0.4$  s before to 4.0 s after the onset of smooth target motion. We used the observed MEG data during only one of the three visual target stimuli for future analysis. From this analysis, we rejected trials whose MEG signals from all channels exceeded 95% of the recordable range. If a subject fell asleep during a session, we also rejected the data recorded during the entire session. The remaining trials of subjects A, B, and C were 88 (97.8%), 86 (95.6%), and 56 trials (62.2%), respectively. These remaining data were used for analysis.

### fMRI

We obtained structural and functional magnetic resonance images with a 1.5-T MR scanner (MAGNEX ECLIPSE 1.5 T Power Drive 250;

Shimadzu-Marconi, Japan) and used a block design for the fMRI experiment. One session consisted of four repetitions of test and rest blocks. In the test block, subjects covertly pursued the moving target for 4 s. After this period, they were allowed a 1-s rest. These periods were repeated alternately three times (15 s). In the rest block, the target was presented in the center of the screen for 15 s. Five sessions were conducted. A target motion was generated in the same way as in the MEG experiment. The fMRI images were preprocessed and analyzed using SPM8 (The Wellcome Department of Cognitive Neurology, UK). The fMRI activities during the covert pursuit eye movements (test block) were significantly increased in the cortical regions of the lateral occipital temporal cortex (LOTc), the intraparietal cortex (IPC), the precentral cortex (PreCC), and the medial superior frontal cortex (MSFC) than during fixation (rest block) ( $p < 0.05$ , FWE corrected, Supplementary information 3). These areas, which are related to the saccadic and smooth pursuit eye movements (Petit and Haxby, 1999), are also activated when subjects orient their attention to visually target motion and pursue it covertly within their visual fields (Corbetta et al., 1998; Kawawaki et al., 2006). These results indicate that we obtained reasonable fMRI responses. In this paper, these fMRI activity patterns were used as prior information of the cortical current variance for hierarchical Bayesian estimation.

### Simulated MEG data

We evaluated our proposed method with simulated MEG data that were generated by the cortical and extra-brain dipole model. Spatial localizations of the cortical sources were based on the experimental subject's fMRI activity pattern. Sinusoidal currents were set to cortical dipoles located on fMRI activation areas ( $p < 0.05$ , FWE corrected), and zero currents were set to other dipoles. Candidates of the cortical areas related to the covert pursuit task are LOTc, IPC, PreCC, and MSFC. Dipoles on each cortical area were set to sinusoids with different frequencies ( $\sqrt{3}$ ,  $\sqrt{5}$ ,  $\sqrt{7}$ , and  $\sqrt{11}$  Hz, respectively). Other dipoles on the fMRI activation area were set to  $\sqrt{2}$  Hz. In general, we considered the maximum value of the cortical current density to be in the range of 100–250 pAm/mm<sup>2</sup> (Hämäläinen et al., 1993; Jerbi et al., 2004). In this paper, we set the maximum current density on the fMRI activation areas in the range of using a maximum single dipole current amplitude of 200 pAm. We set the amplitudes of the cortical currents proportional to the  $t$ -values of the fMRI information. These dipole currents, which were assumed to have amplitude fluctuations, were added to the cortical currents. The standard deviation of the fluctuations was 10% of the maximum current amplitude (20 pAm) with white Gaussian noise (signal dependent noise). These currents were also added to the spontaneous activity with white Gaussian noise (signal independent noise), where the standard deviation is one-fiftieth of the maximum current amplitude (Aihara et al., 2012).

Main components of the eye dipole currents during covert pursuit tasks were assumed to be tremor and microsaccade. The former was expressed as white Gaussian noise, and the latter was expressed as a combined sinusoidal wave of  $\sqrt{2}$ ,  $\sqrt{3}$ ,  $\sqrt{5}$ ,  $\sqrt{7}$ , and  $\sqrt{11}$  Hz, which had significant correlations with cortical currents of all areas. The current values of the  $x$ - $y$ - $z$  directions located in the eye were independently distributed. The standard deviations of the current intensities were 7.07 nAm of the  $x$ - $y$ - $z$  directions of the left and right eyeballs.

The heart dipole currents were assumed to be white Gaussian noise, and the current values of the  $x$ - $y$ - $z$  directions located in the heart were independently distributed. Standard deviations of 3.15, 1.58, and 1.05  $\mu\text{Am}$  were set as the heart current intensities of the  $x$ - $y$ - $z$  directions (these heart current intensities were calculated based on results in Nousiainen et al., 1986; Supplementary information 4).

To investigate the level of the magnetic background noise, MEG data were recorded without a subject. The standard deviation calculated from the acquired data of all sensors was 70.29 fT. In this paper, we used this value for making background sensor (observation) noise

with a Gaussian distribution for the simulated data. Considering these conditions, the cortical and extra-brain dipole currents were generated. We calculated the magnetic field from these currents with a lead field matrix, added magnetic background noise to them, and generated magnetic field data as simulated MEG data (Fig. 2). The total number of trials was 90 (30 trials  $\times$  3 sessions = 90 trials), which was the same as the measured MEG data for a single moving target. The signal-to-noise ratio and the standard deviation of the generated MEG data were also the same as those for the measured MEG data (Table 1 and Eq. (S3) in Supplementary information 5).

*Evaluation of currents estimated from simulated MEG data*

To evaluate the cortical currents estimated from the simulated MEG data, we calculated the root mean square error between the actual and estimated currents using the following formula:

$$RMSE(\gamma_{pri}, m_{pri}) = \sqrt{\frac{1}{I} \frac{1}{T} \sum_{i=1}^I \sum_{t=1}^T (\bar{J}_{true(i)}(t) - \bar{J}_{(i)}(t))^2}, \quad (18)$$

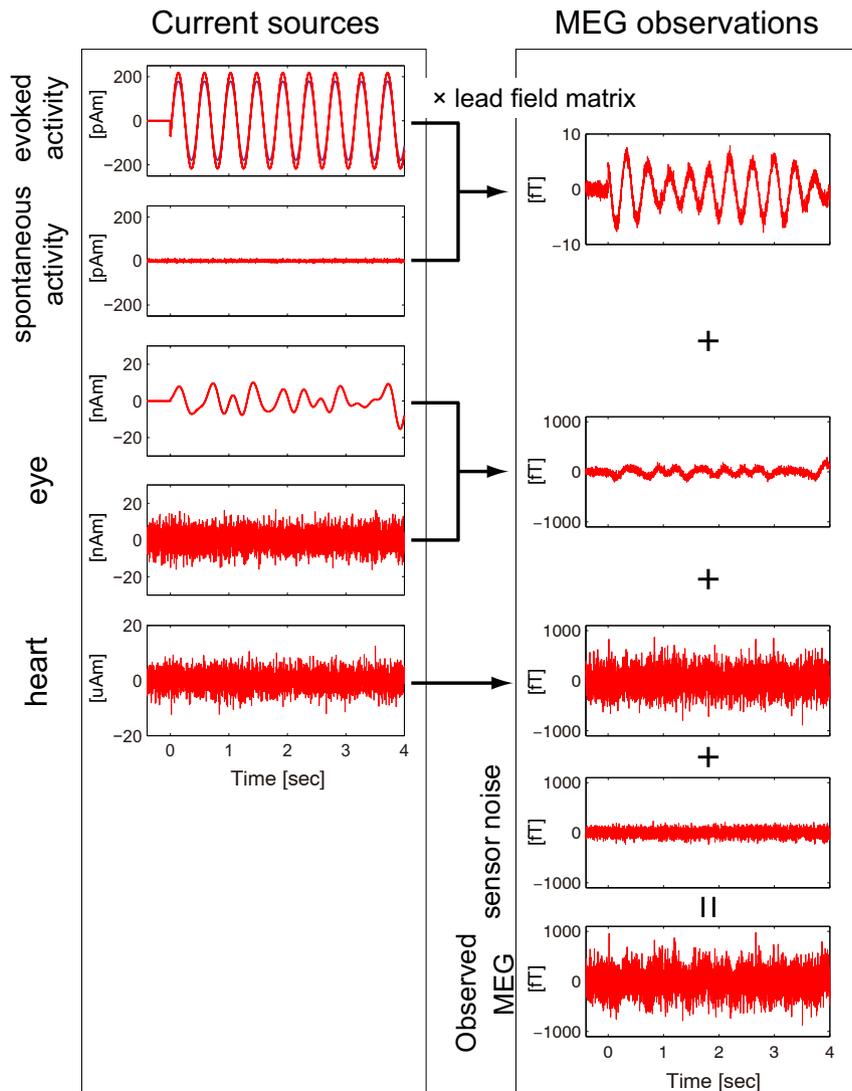
**Table 1**  
Standard deviation and signal-to-noise ratio for MEG data.

Subj. (trial num.)	S.D. [fT]		S/N ratio
	$\hat{B}$	$B$	
Simulated data (90)	33.38	184.90	0.033
Observed data A (88)	18.40	140.48	0.017
B (86)	24.73	151.54	0.027
C (56)	22.02	189.13	0.020

where  $I$  denotes the number of cortical dipoles,  $T$  denotes the sampling number,  $i$  is the index of the cortical dipole, and  $t$  denotes the sampling index.  $\bar{J}$  and  $\bar{J}_{true}$  denote the average of the estimated and actual cortical currents across trials.

*Artifact removal using PCA*

We applied PCA to the simulated MEG data. The first and second components had a characteristic where their spatial weights over sensor space were widely distributed. Because these components were



**Fig. 2.** Overview of simulated MEG data.

probably artifacts, we removed them (Supplementary information 6 and Fig. S3A). Then, we applied the standard VBMEG for only cortical dipoles to the subtracted MEG data.

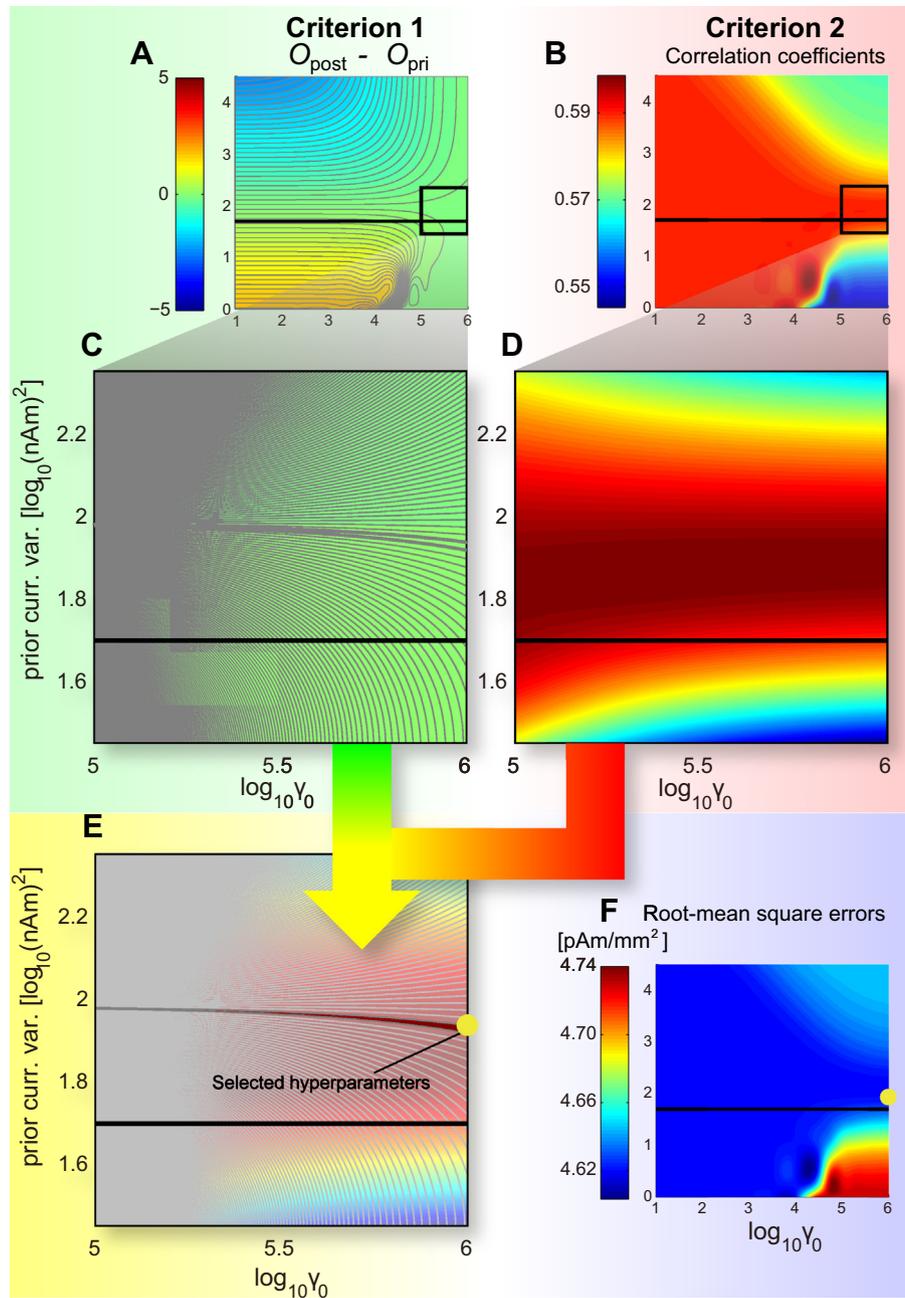
#### Artifact removal using ICA

We applied ICA to the simulated MEG data using the infomax ICA algorithm (Delorme and Makeig, 2004) and isolated 208 independent components. Two were identified as artifacts and removed because their spatial weights over sensor space were widely distributed (Supplementary information 6 and Fig. S3B). Then, we applied

the standard VBMEG for only cortical dipoles to the subtracted MEG data.

#### Calculation of mean magnetic fields arising from heartbeat

To calculate the mean profile of magnetic fields arising from heartbeats, the peak of the R-wave was identified for each heartbeat cycle in the record. In this peak detection process, artifactual ICA components of the 83rd MEG sensor data were used because this sensor was largely affected by heartbeats (Supplementary information 7). After detecting every peak of the R-wave, the observed MEG data was extracted from



**Fig. 3.** (A) Overview of the two-dimensional (2D) map of calculated order differences between prior and posterior current variances for simulated MEG data (e.g., y-direction of left eye). Thin gray lines show contours, and thick gray ones show edges of near-zero order of magnitude. Black line shows actual current variance. (B) Overview of the 2D map of correlation coefficients between cortical current intensities and fMRI activities. (C) and (D) are enlarged views of (A) and (B), respectively. (E) The selected hyper-parameter set based on two criteria. (F) Root mean square errors between estimated and actual currents. Yellow circle shows the selected hyper-parameter set.

– 150 ms before to 450 ms after the onset of the R-wave peak. The numbers of the extracted heartbeat cycles for subjects A, B, and C were 399, 298, and 260, respectively. Note that artifactual ICA components were only used for detecting the peak time location of the R-wave, and the observed MEG data were used for calculating the mean profile of the magnetic fields.

#### Goodness-of-fit evaluation for selecting the MEG data of the good/intermediate/bad fits

To evaluate the true and reconstructed MEG data for selecting the good/intermediate/bad fits, we calculated the goodness-of-fit using the following formula:

$$GOF_m = 1 - \frac{\sum_{t=1}^T (B_m^{true}(t) - B_m^{rec}(t))^2}{\sum_{t=1}^T (B_m^{true}(t) - \overline{B_m^{true}}(t))^2}, \quad (19)$$

where,  $B_m^{true}$  and  $B_m^{rec}$  indicate the true and reconstructed magnetic fields, respectively.  $m$  indicates the MEG channel index.

## Results

### Selecting hyper-parameters using simulated MEG data

To investigate whether our approach can properly select prior information, we applied it to the simulated MEG data. By systematically varying hyper-parameters for the extra-brain source dipoles, we calculated differences between prior and posterior current variances using Eq. (11). Fig. 3A shows an overview of the map of the calculated order differences, and Fig. 3C shows an enlarged view of it while taking an example of the result of the  $y$ -direction dipole of the left eye. Thin gray lines show the contours, and thick gray ones show the edges of the near-zero difference in order of magnitude. The black line shows the actual current variances used in the simulation.

The correlation coefficients between the cortical current intensities and the fMRI activities were also calculated using Eq. (13). Fig. 3B shows an overview of the map of the calculated correlation coefficients, and Fig. 3D shows an enlarged view of it. The yellow circle in Fig. 3E denotes the selected hyper-parameter set based on Criterion 1 and 2. The root mean square errors (RMSEs) between the estimated and actual cortical currents were calculated by Eq. (18). RMSEs were calculated for all hyper-parameter sets in the search range, and then the smallest value around the selected parameter set was taken, which indicates that our hyper-parameter estimation was effective. The selected prior current variance was fairly close to the actual current variance (compare the yellow circle and the black horizontal line in Fig. 3F; selected prior current variance:  $9.27^2$  nAm<sup>2</sup>, true current variance:  $7.06^2$  nAm<sup>2</sup>).

Our proposed method selected proper prior current variance at the  $y$ -direction dipole of the left eye as well as the other dipoles located on both eyes (Table 2). We also applied it to select the hyper-parameter set for the cortical dipoles. The selected magnification parameter was very close to the optimal value derived from Eq. (17), and the errors between the estimated and true currents were smaller than those calculated using other hyper-parameter sets (Table 2).

At the heart's  $x$ - and  $y$ -direction dipoles, the order of the selected current variances was the same level as the true values, but the selected current variance for the  $z$ -direction dipole tended to be larger than the true current variance (about three times larger than the true current intensity).

### Comparison of observed and reconstructed magnetic fields

To investigate the quality of the currents estimated from simulated MEG data, we reconstructed the MEG data from the estimated currents and compared them with the true MEG data. If the estimated currents

**Table 2**  
Selected hyper-parameters for simulated MEG data.

		True values	Selected values	
		Curr. var. [nAm] <sup>2</sup>	Curr. var. [nAm] <sup>2</sup>	Reliability $\gamma_0$
Left eye	(x)	7.07 <sup>2</sup>	11.27 <sup>2</sup>	10 <sup>6.00</sup>
	(y)	7.06 <sup>2</sup>	9.27 <sup>2</sup>	10 <sup>5.99</sup>
	(z)	7.07 <sup>2</sup>	7.59 <sup>2</sup>	10 <sup>5.32</sup>
Right eye	(x)	7.07 <sup>2</sup>	7.03 <sup>2</sup>	10 <sup>6.00</sup>
	(y)	7.07 <sup>2</sup>	8.75 <sup>2</sup>	10 <sup>6.00</sup>
	(z)	7.07 <sup>2</sup>	7.73 <sup>2</sup>	10 <sup>6.00</sup>
Heart	(x)	3154.96 <sup>2</sup>	2870.78 <sup>2</sup>	10 <sup>5.97</sup>
	(y)	1576.37 <sup>2</sup>	1399.59 <sup>2</sup>	10 <sup>5.00</sup>
	(z)	1051.83 <sup>2</sup>	3061.96 <sup>2</sup>	10 <sup>6.00</sup>
		True value	Selected values	
		$m_0$	$m_0$	$\gamma_0$
Cortex		12.95	13.06	10 <sup>6.00</sup>

were appropriate, we expect the reconstructed magnetic fields to resemble the true magnetic fields. Fig. 4A compares the reconstructed and true magnetic fields calculated from the mean cortical currents. The blue and red lines indicate the reconstructed and true magnetic fields, respectively. Fig. 4B shows the histogram for the number of the MEG sensor across all of goodness-of-fit scores. The most spatiotemporal patterns of the reconstructed magnetic fields resembled the true ones (Figs. 4A and B). Figs. 4C and D compare the eye and heart magnetic fields between the estimated and true ones. The magnetic fields calculated from eye currents had a spatiotemporal characteristic of large amplitudes around both eyes (Fig. 4C). The heart magnetic fields had a characteristic of large amplitudes throughout the MEG sensor space, especially at peripheral sensors (Fig. 4D). These results indicate that the reconstructed magnetic fields calculated from the estimated cortical, eye, and heart currents can reproduce the true ones well about the spatiotemporal characteristic.

### Spatiotemporal characteristics of estimated currents

To investigate the accuracy of the estimated current intensities and their distributions over cortical areas, we calculated the cortical current distribution using Eq. (12) and plotted them on the inflated cortical surface maps (Fig. 5). Figs. 5A and B shows spatial patterns of the true and estimated current distributions. The cortical areas with large estimated current intensities were very consistent with areas that have true cortical currents (correlation coefficient:  $r = 0.77$ ). The amplitudes of the estimated current intensities were also the same level as those of the true current intensities (the mean true and estimated current intensities on the fMRI activation area were 13.92 and 4.46 pAm, respectively). Our method estimated currents of the proper cortical areas fairly well.

To investigate the quality of the temporal patterns of the estimated cortical currents, we compared the temporal profiles of the mean cortical current densities with the true ones at each activation area. First, we calculated a trial average of the current densities for each cortical dipole. Second, we calculated the goodness-of-fit between the estimated and true currents. Third, we found the dipole index with the largest goodness-of-fit for each cortical area and used it for the analysis. Fig. 6 compares the estimated and true current densities for all activation areas, and the blue and red lines denote estimated and true values, respectively. The current dipoles of cortical areas LOTC, IPC, PreCC, and MSFC were respectively set as the sine waves of the following frequencies:  $\sqrt{3}$ ,  $\sqrt{5}$ ,  $\sqrt{7}$ ,  $\sqrt{11}$  Hz. Our proposed method expressed the characteristic of sine waves well with different frequencies on each activation area. Additionally, the true cortical currents have a characteristic where

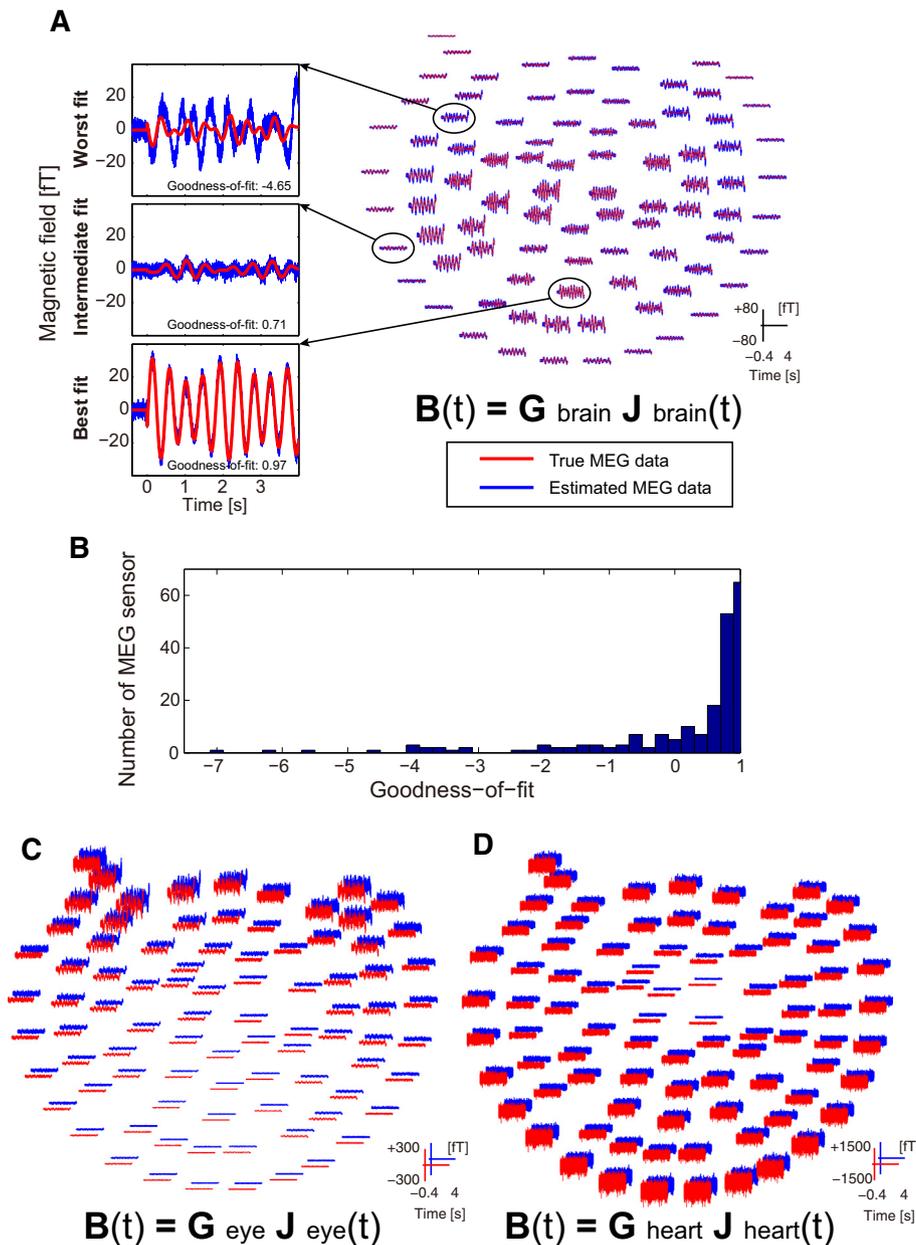
each dipole current has a different amplitude that is proportional to the fMRI activity information. The estimated cortical currents also expressed their differences in amplitudes.

Next, we scrutinized the correlations between the time series of estimated and true currents. We used scatter diagrams for each cortical area (MSFC, PreCC, IPC, and LOTC) to investigate the possible relationship between the estimated (x-axis) and true (y-axis) mean current densities. Each activation area had a high correlation between them ( $r = 0.57\text{--}0.99$ ), meaning that estimated currents reproduced true ones well (Supplementary information 8 and Fig. S5). MSFC was the most poorly correlated brain area with  $r = 0.57$  because it is close to the eyes and most significantly influenced by eye artifacts. The dipole currents of both eyes and the heart were also investigated using scatter diagrams. Highly positive correlations

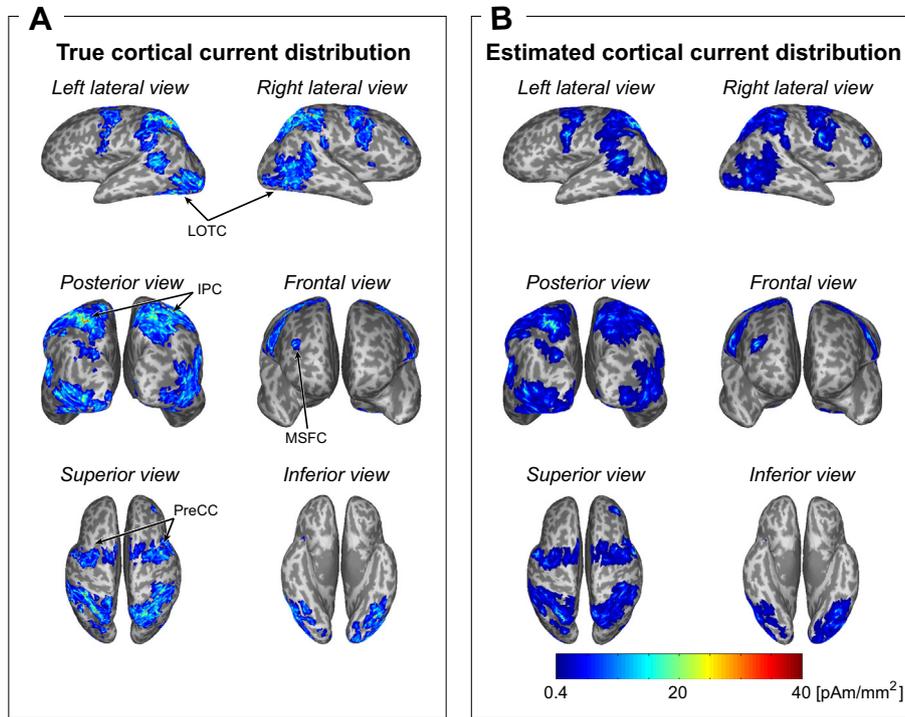
between the estimated and true currents were confirmed at the  $x\text{--}y$  directions of both eyes and the  $x\text{--}y$  directions of the heart (Supplementary information 8 and Fig. S6). The relationships between the estimated and the true current densities show a slope from 0.61 to 0.91 for the eye currents, and from 0.77 to 0.78 for the heart currents. Slopes smaller than 1 indicate that the estimated currents tended to have smaller values than the true ones, i.e., estimation gains were less than 1, which is reasonable (see Discussion).

#### Comparison of denoising performance among different methods

Our approach is expected to acquire less distorted cortical currents by simultaneously estimating not only cortical currents but also extra-brain source currents. To demonstrate this point quantitatively, we



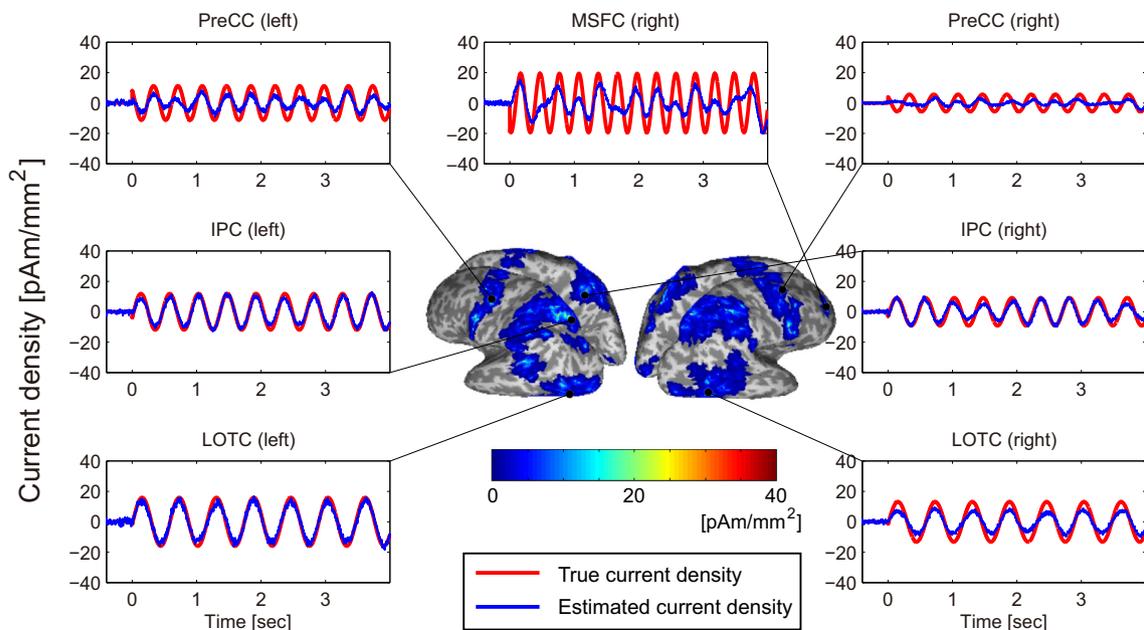
**Fig. 4.** Comparison of observed and reconstructed magnetic fields. (A) Observed and reconstructed magnetic fields were calculated from cortical currents and compared. The worst/intermediate/best fit MEG data were selected based on the Eq. (19). (B) Histogram for the number of the MEG sensor across all of goodness-of-fit scores. (C) Observed and reconstructed magnetic fields were calculated from both eye currents and compared. (D) Observed and reconstructed magnetic fields were calculated from heart currents and compared. 86 sensors from total 208 sensors were selected for the visualization purpose, which were distributed in almost equal distance.



**Fig. 5.** Cortical current distribution. (A) Current distribution calculated from true cortical currents on an inflated cortical map. (B) Current distribution calculated from estimated cortical currents on an inflated map.

compared the performance of our proposed method to conventional denoising methods. The first is the original VBMEG method (Sato et al., 2004), which is called here the “VBMEG” method. In this method, only the dipoles are located on the cortical surface and the dipole currents are estimated (Supplementary information 9 and Fig. S7A). The second is the “PCA” method, where the extra-brain components were removed from the MEG data by PCA in the preprocessing step, then

dipole currents were estimated by using the VBMEG (Supplementary information 9 and Fig. S7B). The third one is the “ICA” method, where the extra-brain components were removed from the MEG data by ICA in the preprocessing step, then dipole currents were estimated by using the VBMEG (Supplementary information 9 and Fig. S7C). The last one is our proposed method, “extra-dipole,” which places the dipoles on the cortical surface as well as on the extra-brain sources and



**Fig. 6.** Comparison of true and estimated cortical currents for a single dipole. A nine-period simple moving average was applied to the time series of estimated currents.

simultaneously estimates both types of dipole currents (Fig. 5B). The errors between the estimated and true cortical currents were calculated by Eq. (18). The errors of our proposed method were the smallest among the four methods compared (Fig. 7, “Correct prior information”). The original VBMEG method is not explicitly concerned about denoising unlike PCA, ICA, or extra-dipole methods, but is expected to possess some denoising capability because it puts spatial and directional constraints on cortical dipoles that should decrease the influences of extra-brain sources. When using the original VBMEG, however, some of the estimated cortical currents tend to represent the extra-brain components mistakenly (shown in red circles on Fig. S7A of Supplementary information 9). By estimating cortical currents and extra-brain dipole currents simultaneously, we could avoid this problem without any pre-processing steps unlike PCA and ICA.

For further confirmation of the advantage of our proposed method, we introduced false-positive fMRI prior information about positive activations on the cortical areas, even though such areas actually do not have them (Supplementary information 10). One characteristic of VB estimation is robustness against false-positive prior information (Sato et al., 2004). If the components of the cortical activities remained completely in the MEG data after the preprocessing steps of PCA and ICA, the error levels of the cortical currents were expected to be the same as the extra-dipole methods. However, when false prior information was used, the PCA and ICA errors exceeded those of the extra-dipole method (Fig. 7, “False positive prior information”). The errors of the four methods under the two different prior information conditions were quite similar with each other (compare Fig. 7 “Correct” and “False positive prior information”).

#### Applications to measured MEG data

To examine practical utility and estimation accuracy of our approach, we applied it to measured MEG data. We selected the hyper-parameter set based on Criterion 1 and 2 (Fig. 8 and Table 3). We used the selected parameters to estimate the cortical and extra-brain source currents. Katila et al. (1981) recorded ocular magnetic fields around a head when a subject was watching a light stimulus and estimated the eye dipole current moments from the observed magnetic fields. The estimated values were about 10 nAm. In our experiments, the mean current variances of both eyes estimated from the observed MEG data were  $6.25^2$ – $11.30^2$  nAm<sup>2</sup>, and these estimated values were similar to those of previous research with respect to order of magnitude (Fig. 9A, Table 3).

Fig. 9B (upper panel) shows the mean magnetic fields of a heart component (red line). The measured MEG data were extracted using the onset of the R-wave, and the mean profile of the magnetic fields

was calculated (see the Materials and method section for details). Fig. 9B (lower panel) shows the mean dipole currents located on the heart (orange, green, and cyan lines), which were also extracted using the same onset as the observed MEG. The orange, green, and cyan lines denote the components of the *x*-, *y*-, and *z*-directions, respectively. The heart components of the measured magnetic fields had a bipolar impulse profile in the range of R- to S-waves (Fig. 9B, upper panel). The estimated currents, especially the *y* and *z* components, also had the same characteristic in the same time window (Fig. 9B, lower panel). The measured magnetic fields changed slowly in the time window of Q- and T-waves, and the estimated heart currents also changed slowly there. Consequently, the estimated *y* and *z* components of heart currents expressed the characteristics of the measured magnetic fields.

We calculated the current intensities from the mean cortical currents during the covert pursuit task using Eq. (12) and plotted the values on the inflated cortical surface map (Fig. 10). LOTC, PreCC, IPC, and MSFC shared large current intensities among this experiment’s participants. These cortical areas were the same as the expected activation areas for the covert pursuit tasks from previous studies (Corbetta et al., 1998; Kawawaki et al., 2006). If the cortical currents were accurately estimated by our proposed method, their temporal patterns may include the time series correlated with kinematic target information such as positions and velocities. We investigated whether each cortical area possesses estimated currents representing such kinematic information of the target for attention. First, we calculated a trial average of each dipole current. Second, we calculated the correlation coefficients between the mean cortical currents and the target position and velocity. Third, the dipole index with the largest correlation coefficient for each cortical area was determined, and its time series of the mean cortical current was plotted in Fig. 11A (subject A). The target of attention began to move at the onset and moved smoothly and slowly for 4 s within the subject’s visual field. The visual evoked response was observed around 0.1 s from the onset as a negative deflection in several traces including the LOTC position trace of Fig. 11A. After that time point, the time series of the estimated currents was correlated with the target positions and/or velocities. We investigated all mean dipole currents to which two types of target information (positions and velocities) were the most similar, then we calculated the rate for each cortical area from all dipole currents of all participants and plotted a pie chart (Fig. 11B). In this analysis, we calculated these correlation coefficients using the data in a time range of 0.6–4.0 s to avoid the effects of visual evoked responses. If the correlation coefficient was below 0.2, we assumed that the dipole current had no resemblance to the target information and categorized it as “no correlation.” As a result, each activation area had two types of target information, and their rate depended on the activation areas. In the future, we need to investigate in more detail which dipoles on each activation area code the target information and whether they have predictive representation about the moving targets using decoding techniques.

#### Discussion

This paper proposes a practical MEG denoising method that requires no additional sensor data other than MEG, unlike the hyVBED method. We applied our method to simulated and measured MEG data that were contaminated by both eye movements and heartbeats. By introducing proper prior information, our approach more accurately estimated cortical and extra-brain source currents from the obfuscated MEG signals than the original VBMEG method. Additionally, we compared the performance of this method with such conventional statistical methods as PCA and ICA. It surpassed those that only used statistical properties of time-series signals (Fig. 7).

Fujiwara et al. (2009) recorded MEG and EOG signals while subjects performed horizontal, smooth pursuit eye movements (horizontal amplitude, 4°; frequency, 0.7 Hz). They calculated a forward model for both eyes from the EOG data and estimated the current moments of the eyes from the EOG and MEG data by the variational

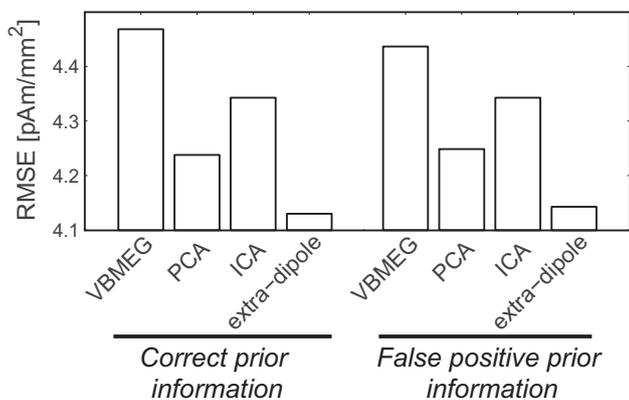
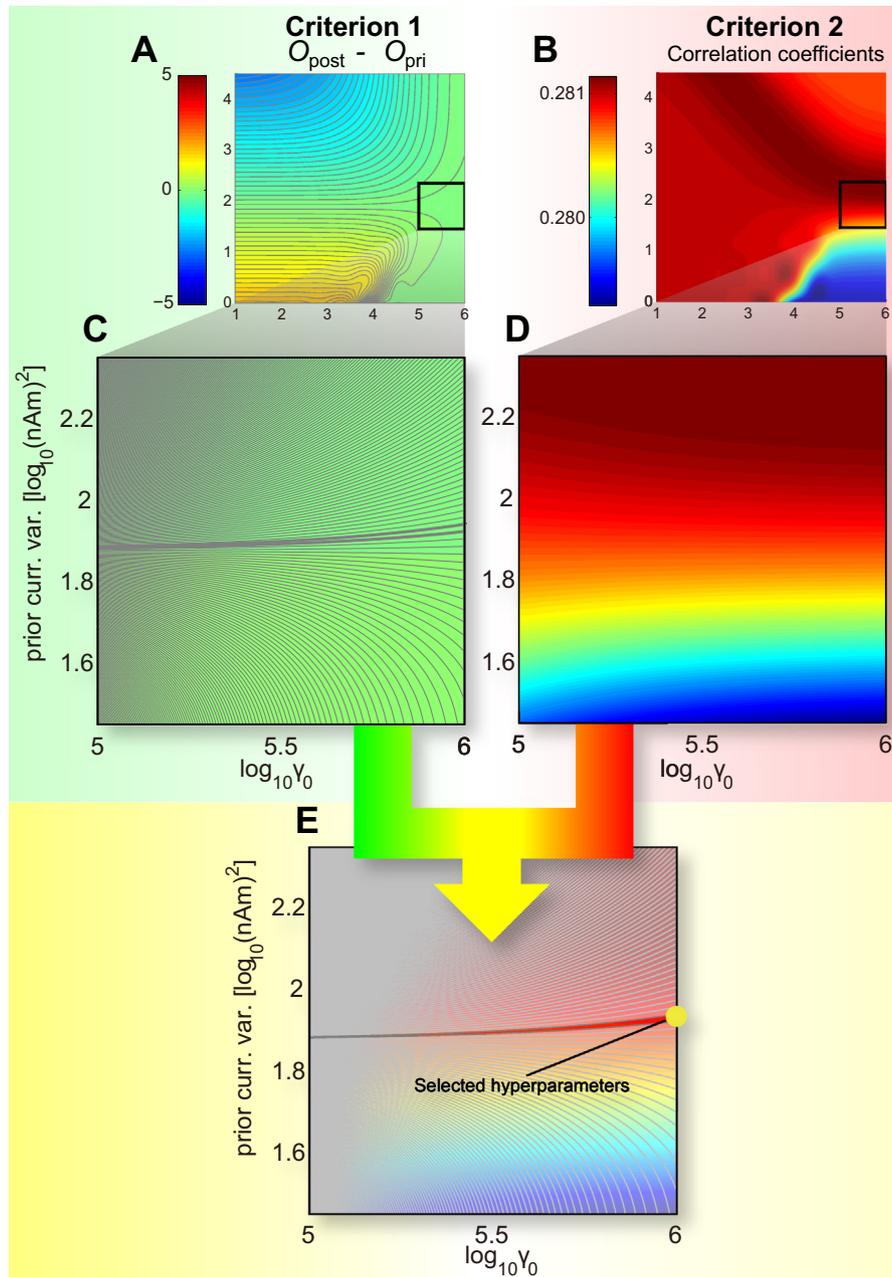


Fig. 7. Comparison of root mean square errors of estimated currents among four denoising methods. (A) shows the results under a condition of correct prior information. (B) shows the results under a condition of false-positive prior information.



**Fig. 8.** (A) Overview of the two-dimensional (2D) map of the calculated order differences between prior and posterior current variances for measured MEG data (e.g., y-direction of left eye). Thin gray lines show contours, and thick gray ones show edges of near-zero order of magnitude. (B) Overview of the 2D map of correlation coefficients between cortical current intensities and fMRI activities. (C) and (D) are enlarged views of (A) and (B), respectively. (E) The selected hyper-parameter set based on two criteria. Yellow circle shows the selected hyper-parameter set.

Bayesian method. The trial average of the estimated current amplitudes was in the range of 21.7–36.3 nAm. Unlike Fujiwara's experiment, this paper's task did not require subjects to move eyes overtly but only covertly. Thus, here involuntary eye movements such as tremors and microsaccades caused electrical currents at both eyes. The tremor and microsaccade amplitudes were about  $0.0042^\circ$  and  $0.25^\circ$ , respectively. Their vibrational frequencies were as high as 90 and 0.5 Hz, respectively (Martinez-Conde et al., 2004). The ratios of the charge movement currents between Fujiwara and our experiments were  $4^\circ \times 0.7 \text{ Hz} : 0.0042^\circ \times 90 \text{ Hz} = 1 : 0.135$  and  $4^\circ \times 0.7 \text{ Hz} : 0.25^\circ \times 0.5 \text{ Hz} = 1 : 0.04$ . When converting the estimated current intensities in their work with those

of our experiments, the eye current intensities arising from tremors are expected in the range of 2.93–4.90 nAm, and the eye current intensities arising from microsaccades are expected in the range of 0.97–1.62 nAm. The total intensity is about 3.90–6.52 nAm. The order in estimated magnitude was also the same as our results obtained with VBMEG augmented by extra-cortical dipoles.

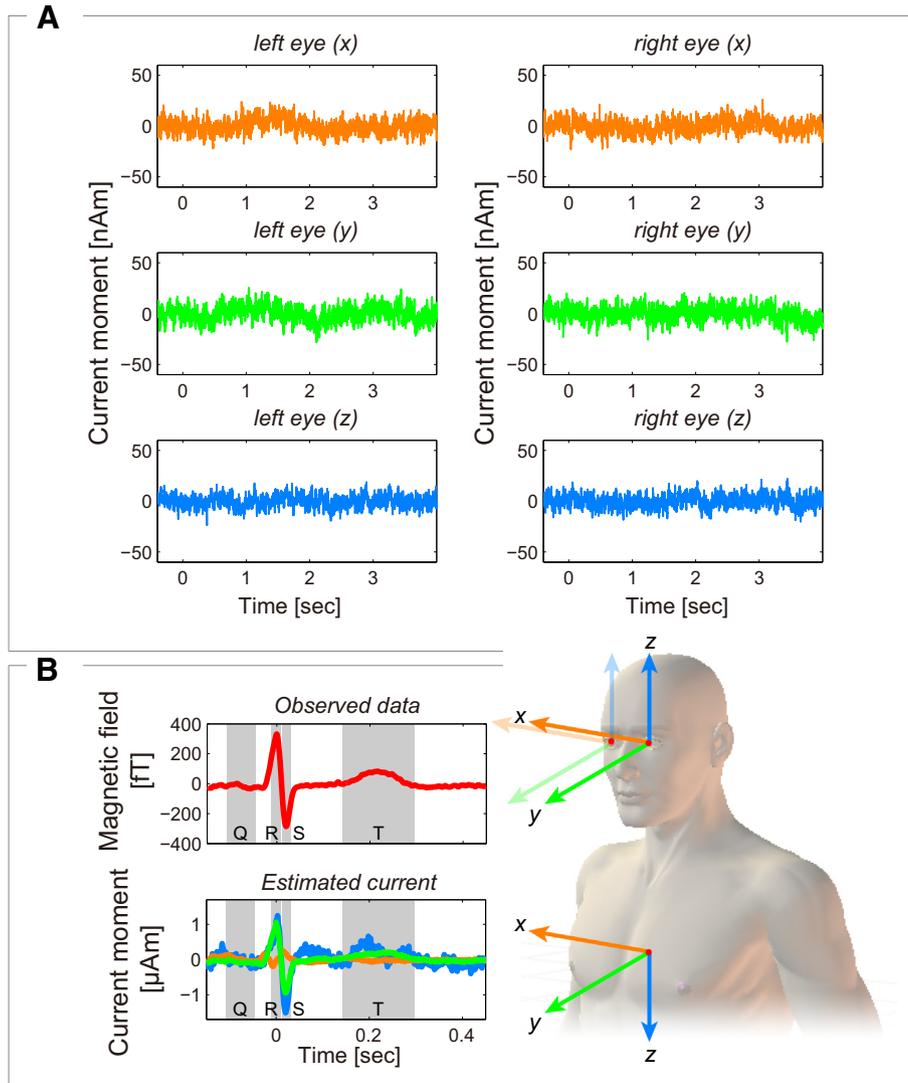
We calculated the mean values of the estimated current variance of the heart from simulated MEG data. The values were  $(x, y, z) = (2.87^2, 1.40^2, 3.06^2) \mu\text{Am}^2$ . The true current variances assumed in the simulation were  $(x, y, z) = (3.15^2, 1.58^2, 1.05^2) \mu\text{Am}^2$ . The estimated values of the x- and y-directions were the same level as the true ones, but the z-elements were larger than the true ones. The z-direction

**Table 3**  
Selected hyper-parameters for measured MEG data.

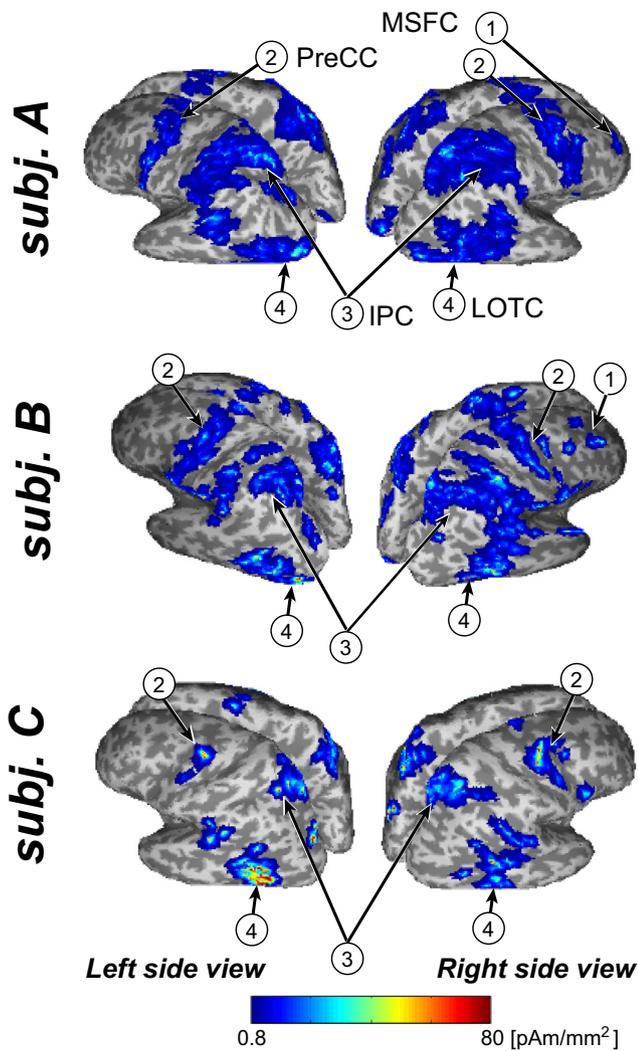
		Subj. A		Subj. B		Subj. C		Mean	
		Curr. var. [nAm] <sup>2</sup>	Reliability $\gamma_0$						
Left eye	(x)	9.20 <sup>2</sup>	10 <sup>6.00</sup>	6.93 <sup>2</sup>	10 <sup>6.00</sup>	9.64 <sup>2</sup>	10 <sup>6.00</sup>	8.67 <sup>2</sup>	10 <sup>6.00</sup>
	(y)	9.29 <sup>2</sup>	10 <sup>6.00</sup>	11.30 <sup>2</sup>	10 <sup>6.00</sup>	10.76 <sup>2</sup>	10 <sup>6.00</sup>	10.49 <sup>2</sup>	10 <sup>6.00</sup>
	(z)	7.55 <sup>2</sup>	10 <sup>6.00</sup>	6.25 <sup>2</sup>	10 <sup>6.00</sup>	6.41 <sup>2</sup>	10 <sup>5.98</sup>	6.76 <sup>2</sup>	10 <sup>5.99</sup>
Right eye	(x)	9.20 <sup>2</sup>	10 <sup>6.00</sup>	7.06 <sup>2</sup>	10 <sup>6.00</sup>	11.69 <sup>2</sup>	10 <sup>6.00</sup>	9.51 <sup>2</sup>	10 <sup>6.00</sup>
	(y)	8.79 <sup>2</sup>	10 <sup>6.00</sup>	8.04 <sup>2</sup>	10 <sup>6.00</sup>	13.03 <sup>2</sup>	10 <sup>5.99</sup>	10.19 <sup>2</sup>	10 <sup>6.00</sup>
	(z)	7.76 <sup>2</sup>	10 <sup>6.00</sup>	5.31 <sup>2</sup>	10 <sup>6.00</sup>	7.35 <sup>2</sup>	10 <sup>6.00</sup>	6.89 <sup>2</sup>	10 <sup>6.00</sup>
Heart	(x)	368.13 <sup>2</sup>	10 <sup>5.99</sup>	446.69 <sup>2</sup>	10 <sup>5.98</sup>	384.59 <sup>2</sup>	10 <sup>6.00</sup>	401.23 <sup>2</sup>	10 <sup>5.99</sup>
	(y)	445.66 <sup>2</sup>	10 <sup>6.00</sup>	449.78 <sup>2</sup>	10 <sup>6.00</sup>	444.63 <sup>2</sup>	10 <sup>5.99</sup>	446.69 <sup>2</sup>	10 <sup>6.00</sup>
	(z)	3243.40 <sup>2</sup>	10 <sup>6.00</sup>	3235.94 <sup>2</sup>	10 <sup>6.00</sup>	2805.43 <sup>2</sup>	10 <sup>5.99</sup>	3101.69 <sup>2</sup>	10 <sup>6.00</sup>
		Subj. A		Subj. B		Subj. C		Mean	
		$m_0$	$\gamma_0$	$m_0$	$\gamma_0$	$m_0$	$\gamma_0$	$m_0$	$\gamma_0$
Cortex		27.04	10 <sup>6.00</sup>	33.73	10 <sup>6.00</sup>	36.98	10 <sup>6.00</sup>	32.58	10 <sup>6.00</sup>

element of the lead field matrix was smaller than the others (values not shown in the paper). This result means that the MEG sensors have poor sensitivity to the z-direction magnetic fields. To express the observed magnetic field using only one dipole current, we selected the larger

current variance as prior information. This result might suggest that there is no need to locate the z-direction dipole of the heart in the first place. However, it might also suggest that we can locate extra dipoles because no negative effects were found at the cortical, eyes, and other heart



**Fig. 9.** Temporal patterns of extra-brain source currents estimated from measured MEG data. (A) Example of a single trial data of estimated eye currents (first trial data of subject A). (B) Example of trial average values of magnetic fields (red line) and estimated x–y–z heart currents (orange, green, and cyan lines) extracted at the onset of an R-wave.



**Fig. 10.** Spatial patterns of cortical currents estimated from measured MEG data (data of three subjects).

dipole currents. If a MEG system has not only axial- but also planar-gradiometer sensors, we can acquire rich sensitivity for the z-direction of the heart's magnetic fields (e.g., a whole-head 400-channel MEG system, Yokogawa Co., Japan). If the acquired MEG data contain the z-direction components, it is more meaningful to locate the z-direction dipole in the heart. It might be possible to obtain more precise dipole currents if axial- and planar-gradiometer sensor data were used for future data analysis. The current variance for each extra-brain source dipole was assumed to obey a single Gaussian distribution. If extra-brain sources have constant current intensities over entire trials, this assumption is appropriate (e.g., involuntary eye movements). However, heartbeats tend to have electrical impulse patterns in cycles, and the amplitude was larger than those of the others during R-wave duration. Thus, it is not always a proper assumption to express the properties of heart currents as a single Gaussian process. Under this inappropriate assumption, heart dipole currents were underestimated when larger values were expected, and the currents were overestimated when small values were expected. Therefore, the reconstructed magnetic fields from the estimated heart currents failed to express the amplitudes of the realistic magnetic fields adequately. To avoid this problem, we must improve the statistical model for extra-brain sources (e.g., using not a single Gaussian model but a Gaussian mixture model) to further improve the precision of the denoising method (Yoshioka et al., 2010).

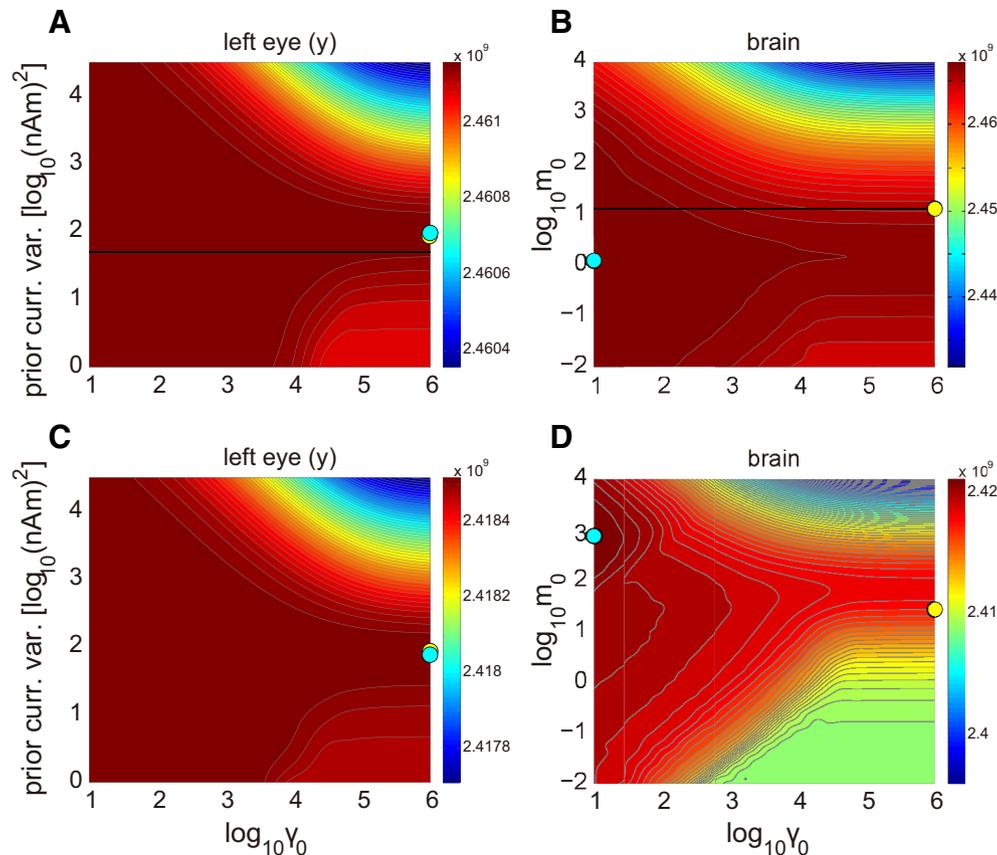
Conventional denoising methods such as PCA and ICA are heavily dependent on the subjective views of users who are required to select which principal and independent components are "artifacts." If too many components are considered artifacts, they might mistakenly remove the magnetic fields caused by brain activities. If too few are selected, they mistakenly attribute magnetic fields caused by artifacts to brain activities. In contrast, our proposed method is free of subjective views or decision criteria for artifacts because our extra-dipole method clearly represents the extra-brain sources using current dipoles, thus allowing it to reasonably isolate cortical and extra-brain activities without subjective decision criteria.

When false-positive fMRI prior information was introduced, the estimated currents with PCA and ICA have larger errors than those with the extra-dipole method. Cortical current estimation using VBMEG tends to have less sensitivity to false-positive prior information (Sato et al., 2004). Despite this property, the errors of PCA and ICA combined with VBMEG were larger. This means that PCA and ICA mistakenly or inappropriately removed some components with brain activities as "artifacts." Mistakenly selecting unnecessary principal and independent components causes improper denoising and current estimation. Our proposed method avoids these problems. It is difficult to remove artifacts correlated with brain signals because the PCA and ICA methods rely on orthogonality or statistical independence of time-series signals. In contrast, our method precisely isolates the cortical and extra-brain source components even if they have some correlation because this method calculates and uses the inverse filter from MEG signals on the cortical and extra-brain source currents. For example, it might be possible to apply this method to analyze MEG data during smooth pursuit tasks that require subjects to move both eyes overtly, not covertly. In such tasks, cortical activities have some correlation with eye movements, but our method is expected to acquire brain activities precisely. In the future, we will apply this method to various research issues that we could not investigate previously.

Why doesn't the conventional framework, which maximizes free energy, select proper hyper-parameters? Free energy reflects the goodness of the fit between the MEG sensor data and the Bayesian model. On the other hand, the index of Criterion 2 mainly reflects the similarity between prior information and the model. Each extra-brain source has a specific large physical property value, and therefore both free energy and Criterion 2 tend to rely on the informative prior, thus selecting large gamma values (Figs. 12A and C). The cortical current sources, which have about 20,000 dipoles located on the cortical surface, are controlled by only one fMRI magnification parameter:  $m_0$ . By greatly relying on  $m_0$ , the model resembles the fMRI activation pattern. Without relying on  $m_0$ , it fits the MEG sensor data. If hyper-parameters are selected by maximizing the free energy, the magnification parameter for the cortical currents tends to fit the MEG sensor data, and thus a small gamma is selected to satisfy the fMRI constraint with a lesser extent (Figs. 12B and D). However, it is empirically unstable for the Bayesian model to explain only the MEG data, and our simulation and experimental results support our theory and analysis. Our proposed method is robust because it selects the optimal hyper-parameters concerning both the MEG and fMRI data. Although our approach is not theoretically elegant but ad-hoc, it provides a robust and practical framework for the MEG inverse problem.

After applications of brain machine interfaces have been integrated into our daily future lives, EEG will be one dominant recording system for brain activities. However, weak EEG signals are susceptible to strong electrical noises in our everyday environments such as TVs, refrigerators, and muscle activities of body movements. However, a limit exists concerning the isolation of the effects of extra-brain sources using conventional methods with only statistical properties of time-series signals, such as PCA and ICA. In addition, hyVBED is not a realistic solution because additional sensor information, such as power-supply noises and EMG signals all over a user's body, is necessary for simultaneous recording. In contrast, our proposed method is a realistic approach because it





**Fig. 12.** Two-dimensional (2D) maps of free energy for simulated and measured MEG data. (A) 2D map for simulated MEG data (e.g., y-direction of left eye). (B) 2D map for simulated MEG data (brain). (C) 2D map for measured MEG data (e.g., subject A; y-direction of left eye). (D) 2D map for measured MEG data (subject A; brain). Black lines show actual current variance (simulated MEG data only). Cyan and yellow circles show selected hyper-parameter sets based on free energy and our proposed method, respectively.

supported by a contract with the Ministry of Internal Affairs and Communications entitled “Novel and innovative R&D making use of brain structures”, the National Institute of Information and Communications Technology (NICT), and “Development of BMI Technologies for Clinical Application” SRPBS by MEXT, and JSPS KAKENHI Grant Number 24700303.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.07.010>.

## References

- Aihara, T., Takeda, Y., Takeda, K., Yasuda, W., Sato, T., Otaka, Y., Hanakawa, T., Honda, M., Liu, M., Kawato, M., Sato, M., Osu, R., 2012. Cortical current source estimation from electroencephalography in combination with near-infrared spectroscopy as a hierarchical prior. *NeuroImage* 59 (4), 4006–4021.
- Barbati, G., Porcaro, C., Zappasodi, F., Rossini, P.M., Tecchio, F., 2004. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clin. Neurophysiol.* 115 (5), 1220–1232.
- Bell, A., Sejnowski, T.J., 1995. An information maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159.
- Callan, D., Callan, A., Sato, M., Kawato, M., 2010. Premotor cortex mediates perceptual performance. *NeuroImage* 51, 844–858.
- Comon, P., 1994. Independent component analysis a new concept? *Signal Process.* 36 (3), 287–314.
- Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., Shulman, G.L., 1998. A common network of functional areas for attention and eye movements. *Neuron* 21 (4), 761–773.
- Dale, A.M., Fischl, B.R., Sereno, M.I., 1999. Cortical surface-based analysis. I: segmentation and surface reconstruction. *NeuroImage* 9, 179–194.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of signal-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 34 (1), 9–21.
- Ferguson, A.S., Zhang, X., Stroink, G., 1994. A complete linear discretization for calculating the magnetic field using the boundary element method. *IEEE Trans. Biomed. Eng.* 46, 455–460.
- Fujiwara, Y., Yamashita, O., Kawawaki, D., Doya, K., Kawato, M., Toyama, K., Sato, M., 2009. A hierarchical Bayesian method to resolve an inverse problem of MEG contaminated with eye movement artifacts. *NeuroImage* 45, 393–409.
- Hämäläinen, M.S., 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.
- Jerbi, K., Baillet, S., Moshier, J.C., Nolte, G., Garnero, L., Leahy, R.M., 2004. Localization for realistic cortical activity in MEG using current multipoles. *NeuroImage* 22 (2), 779–793.
- Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37 (2), 163–178.
- Katila, T., Maniewski, R., Poutanen, T., Varpula, T., 1981. Magnetic fields produced by the human eye. *J. Appl. Phys.* 52 (2), 2565–2571.
- Kawawaki, D., Shibata, T., Goda, N., Doya, K., Kawato, M., 2006. Anterior and superior lateral occipito-temporal cortex responsible for target motion prediction during overt and covert visual pursuit. *Neurosci. Res.* 54 (2), 112–123.
- Martinez-Conde, S., Macknik, S.L., Hubel, D.H., 2004. The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* 5 (3), 229–240.
- Miles, W.R., 1939. The steady polarity potential of the human eye. *Proc. Natl. Acad. Sci. U. S. A.* 25 (1), 25–36.
- Morishige, K., Aihara, T., Kawato, M., Osu, R., Sato, M., 2011. Isolating cortical activities from artifacts in simulated EEG data during smooth pursuit eye movements. The 34th Annual Meeting of the Japan Neuroscience Society (Neuroscience 2011), Yokohama.
- Nousiainen, J.J., Lekkala, J.O., Malmivuo, J.A., 1986. Comparative study of the normal vector magnetocardiogram and vector electrocardiogram. *J. Electrocardiol.* 19 (3), 275–290.
- Petit, L., Haxby, J.V., 1999. Functional anatomy of pursuit eye movements in humans as revealed by fMRI. *J. Neurophysiol.* 82 (1), 463–471.
- Sato, M., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical Bayesian estimation for MEG inverse problem. *NeuroImage* 23, 806–826.
- Shibata, K., Yamagishi, N., Goda, N., Yoshioka, T., Yamashita, O., Sato, M., Kawato, M., 2007. The effects of feature attention on prestimulus cortical activity in the human visual system. *Cereb. Cortex* 18, 1664–1675.

- Toda, A., Imamizu, H., Kawato, M., Sato, M., 2011. Reconstruction of two-dimensional movement trajectories from selected magnetoencephalography cortical currents by combined sparse Bayesian methods. *NeuroImage* 54 (2), 892–905.
- Vigario, R.N., 1997. Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalogr. Clin. Neurophysiol.* 103, 395–404.
- Vigario, R.N., Särelä, J., Jousmäki, V., Hämäläinen, M.S., Oja, E., 2000. Independent component approach to analysis EEG and MEG recordings. *IEEE Trans. Biomed. Eng.* 47 (5), 589–593.
- Yoshimura, N., DaSalla, C.S., Hanakawa, T., Sato, M., Koike, Y., 2012. Reconstruction of flexor and extensor muscle activities from electroencephalography cortical currents. *NeuroImage* 59, 1324–1337.
- Yoshioka, T., Toyama, K., Kawato, M., Yamashita, O., Nishina, S., Yamagishi, N., Sato, M., 2008. Evaluation of hierarchical Bayesian method through retinotopic brain activities reconstruction from fMRI and MEG signals. *NeuroImage* 42, 1397–1413.
- Yoshioka, T., Morishige, K., Kawato, M., Sato, M., 2010. Gaussian mixture prior distribution on artifactual current for MEG inverse problem. The 33th Annual Meeting of the Japan Neuroscience Society (Neuroscience 2010), Kobe.