Multimodal Functional Imaging Using fMRI-Informed Regional EEG/MEG Source Estimation

Wanmei Ou[1], Aapo Nummenmaa[2], Matti Hämäläinen[2], and Polina Golland[1]

[1] Computer Science and Artificial Intelligence Laboratory, MIT, USA
[2] Athinoula A. Martinos Center for Biomedical Imaging, MGH, USA
wanmei@csail.mit.edu

Abstract. We propose a novel method, fMRI-Informed Regional Estimation (FIRE), which utilizes information from fMRI in E/MEG source reconstruction. FIRE takes advantage of the spatial alignment between the neural and the vascular activities, while allowing for substantial differences in their dynamics. Furthermore, with the regional approach, FIRE can be efficiently applied to a dense grid of sources. Inspection of our optimization procedure reveals that FIRE is related to the re-weighted minimum-norm algorithms, the difference being that the weights in the proposed approach are computed from both the current estimates and fMRI data. Analysis of both simulated and human fMRI-MEG data shows that FIRE reduces the ambiguities in source localization present in the minimum-norm estimates. Comparisons with several joint fMRI-E/MEG algorithms demonstrate robustness of FIRE in the presence of sources silent to either fMRI or E/MEG measurements.

Keywords: EEG, MEG, fMRI, Inverse Problem, Expectation-Maximization.

1 Introduction

The principal difficulty in interpreting Electroencephalography (EEG) and magnetoencephalography (MEG) data stems from the ill-posed electromagnetic inverse problem: infinitely many spatial current patterns give rise to identical measurements [10]. Additional assumptions on the spatial current patterns must be incorporated into the reconstruction process to obtain a unique estimate [3].

In addition to the general assumptions about the spatial current patterns such as minimum energy (or ℓ_2-norm), specific prior knowledge about activation locations can be obtained from other imaging modalities. Among them, functional Magnetic Resonance Imaging (fMRI) provides the most relevant information for the reconstruction due to its good spatial resolution. fMRI measures the hemodynamic activity, which indirectly reflects the neural activity measured by E/MEG. Extensive studies of neurovascular coupling have demonstrated similarity in spatial patterns of these two types of activations [13]. However, the timecourses of

the neural and the vascular activities differ substantially, and their exact relationship is yet to be characterized in full. In addition to the differences in their physiological origins, E/MEG and fMRI have different sensitivity characteristics. For example, a brief transient neural activity may be difficult to detect in fMRI while sustained weak neural activity may lead to relatively strong fMRI signals but might have a poor signal-to-noise ratio in E/MEG.

The most straightforward way to incorporate fMRI information into E/MEG inverse estimation is the fMRI-weighted Minimum-Norm Estimation (fMNE), [1,12]. This method uses a thresholded Statistical Parametric Map (SPM) from fMRI analysis to construct weights for the standard Minimum-Norm Estimation (MNE), leading to significant improvements when the SPM is accurate. However, the weights depend on arbitrary choices of the threshold and weighting parameters. Moreover, these weights are assumed to be time independent causing excessive bias in the estimated source timecourses. Sato *et al.* [15] combined the Automatic Relevance Determination (ARD) framework and fMNE to achieve more focal estimates. In this method, which we will refer to as fARD, the parameters of a hyperprior are set based on the thresholded SPM. In addition to the arbitrary choice of the threshold similar to that in fMNE, the estimates computed via fARD are often unstable, especially in the regions where the vascular activity is weak.

Here, we propose a novel method, the fMRI-Informed Regional Estimation (FIRE), to improve the accuracy of the E/MEG source estimates. Since the relationship between the dynamics of the evoked neural and the evoked vascular signals is largely unknown, we only model the similarity of spatial patterns in the two processes, as opposed to the Kalman-filter approach in [6]. Furthermore, we expect that the shape of the activation timecourses varies across brain regions, especially for the neural activation timecourses. To account for this fact, FIRE treats the temporal dynamics in different brain regions independently. In other words, there is no constraint imposing similarity of the activation timecourses across regions. We assume the shape of the activation timecourses to be constant within a brain region, modulated by a set of location-specific latent variables. The regions are chosen based on subject-specific cortical parcellation [7]. Handling the temporal dynamics of the two types of activities separately while exploiting their common spatial pattern helps to preserve the temporal resolution of E/MEG and to achieve accurate source localization.

The prior on the latent variables encourages spatially smooth current estimates within a brain region. The prior also encourages the number of activated regions to be small, similar to the ARD approach [17], except that our prior is region-based rather than location-based. Both the activation timecourse model and the choice of brain regions in FIRE are similar to those employed in recent work by Daunizeau *et al.* [4]. However, Daunizeau *et al.* aim to symmetrically infer brain activities visible in either EEG or fMRI data, resulting in an extra random variable to model the vascular activity. The confidence of the estimated brain activities reduces when there are discrepancies between the EEG and the fMRI measurements. Furthermore, due to the complexity of this model, the estimation is limited to a coarse source space. Instead of aiming at a

symmetrical inference, we focus on the estimation of current sources. We incorporate the fMRI information to reduce ambiguities in source localization usually present in E/MEG source estimation.

To fit the model to the data, we use the coordinate descent method, alternating between the estimation of current sources and of other model parameters. This iterative update scheme is similar to the re-weighted MNE methods such as FOCal Underdetermined System Solver (FOCUSS) [8]. In contrast to the re-weighted MNE, in our method the weights are jointly determined using both the estimated neural activity and the vascular activity measured by fMRI. Moreover, the estimates at different time points influence each other. The computation of the weights is related to problems arising in continuous Gaussian mixture modeling, which can be efficiently optimized using the Expectation-Maximization (EM) algorithm [5].

In the following, we first discuss the model underlying FIRE, the inference procedure, and the implementation details. We then present the experimental comparisons between FIRE and prior methods for joint E/MEG-fMRI analysis using both simulated and human data, followed by a discussion and conclusions.

2 Methods

2.1 Neurovascular Coupling and Data Models

We assume that the source space comprises N discrete locations on the cortex parcelled into K brain regions. We denote the set indexing the discrete locations in region k by P_k and the cardinality of P_k by N_k.

Fig. 1 illustrates our model. The shape of the source timecourses is identical within a region but varies across regions. Specifically, we let $\mathbf{u}_k$ and $\mathbf{v}_k$ be the
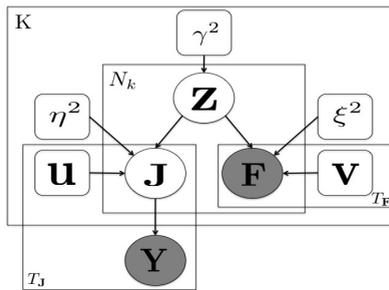
Fig. 1. Graphical interpretation of FIRE. The hidden activity $\mathbf{z}$ models the neurovascular coupling relationship. The hidden current source distribution $\mathbf{J}$ is measured by E/MEG, producing observation $\mathbf{Y}$. $\mathbf{F}$ denotes fMRI measurements. Vectors $\mathbf{u}$ and $\mathbf{v}$ are the unknown neural and vascular waveforms in a certain brain region, respectively. The inner plate represents N_k vertices in region k; the outer plate represents K regions. The bottom left and right plates represent $T_{\mathbf{J}}$ and $T_{\mathbf{F}}$ time points in the neural and the vascular measurements, respectively.

unknown waveforms in region k, associated with neural and vascular activities, respectively. We model the neural and the vascular activity strength through a hidden vector variable $\mathbf{z} = [z_1, z_2, \cdots, z_N]^{\mathrm{T}}$. The continuous scalar z_n indicates the activation strength at location n on the cortical surface. Thus, the probabilistic model for the neural activation timecourse $\mathbf{j}_n$ and the vascular activation timecourse $\mathbf{f}_n$ at location n in region k can be expressed as

$$p\left(\mathbf{j}_n, \mathbf{f}_n | z_n; \mathbf{u}_k, \mathbf{v}_k, \eta_k^2, \xi_k^2\right) = \mathcal{N}\left(\mathbf{j}_n; z_n \mathbf{u}_k, \eta_k^2 \mathbf{I}\right) \mathcal{N}\left(\mathbf{f}_n; z_n \mathbf{v}_k, \xi_k^2 \mathbf{I}\right), \qquad (1)$$

where η_k^2 and ξ_k^2 are noise variances. We construct all matrices such that each row represents a location or a sensor and each column represents a particular time point. Thus, we let $N \times T_{\mathbf{J}}$ matrix $\mathbf{J} = [\mathbf{j}_1, \mathbf{j}_2, \cdots, \mathbf{j}_N]^{\mathrm{T}}$ be the neural current on the cortex for all $T_{\mathbf{J}}$ time points. We assume that the vascular signal $\mathbf{f}_n$ at location n is directly observable through fMRI. We let $N \times T_{\mathbf{F}}$ matrix $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_N]^{\mathrm{T}}$ be the fMRI measurements on the cortex over $T_{\mathbf{F}}$ time points. Note that our neurovascular coupling model captures only the spatial alignment between the two types of activities; it does not impose temporal similarity between the signals.

The neural currents $\mathbf{j}_n$ are detected with E/MEG described by the standard observation model. We let $M \times T_{\mathbf{J}}$ matrix $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \cdots, \mathbf{y}(T_{\mathbf{J}})]$ be the E/MEG measurements at all $T_{\mathbf{J}}$ time points. Column t of matrix $\mathbf{J}$, $\mathbf{j}(t)$, denotes the neural current distribution at time t. The quasi-static Maxwell's equations imply that E/MEG signals at time t are instantaneous linear combinations of the currents at different locations:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{j}(t) + \mathbf{e}(t) \qquad \forall\, t = 1, 2, \cdots, T_{\mathbf{J}}, \qquad (2)$$

where $\mathbf{e}(t)$ is the measurement noise. The $M \times N$ forward matrix $\mathbf{A}$ is determined by the electromagnetic properties of the head, the geometry of the sensors, and the locations of the sources. With spatial whitening in the sensor space, $\mathbf{e}(t) \sim \mathcal{N}(0, \mathbf{I})$. The number of sources N ($\sim 10^3 - 10^4$) is much larger than the number of measurements M ($\sim 10^2$), leading to an infinite number of solutions satisfying Eq. (2) even for $\mathbf{e}(t) = \mathbf{0}$. In general, $\mathbf{j}_n$ should be modeled as three timecourses corresponding to the three Cartesian components of the current. However, due to the columnar organization of the cortex, we can further constrain the current orientation to be perpendicular to the cortical surface and consider a scalar timecourse at each location.

2.2 Priors and Parameter Settings

To encourage the activation patterns to be smooth within a region, we impose a prior on the modulating variables. Specifically, we define $\mathbf{z}_k = \{z_n\}_{n \in P_k}$ and assume $\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \gamma_k^2 \mathbf{\Phi}_k)$, where the variance γ_k^2 indicates the activation strength in region k, and $\mathbf{\Phi}_k$ is a fixed matrix that acts as a regularizer by penalizing the sum of squared differences between neighboring locations. This spatial prior is particularly important for the brain regions where vascular activity is too weak to measure, but the neural activity can be detected by E/MEG.

Our $\mathbf{\Phi}_k$ is similar to the regularizer used in the Low Resolution Brain Electromagnetic Tomography (LORETA) [14], except that we apply $\mathbf{\Phi}_k$ to individual brain regions while LORETA's spatial regularizer is applied to the whole brain. We assume separate variance γ_k^2 for different brain regions since the strength of current is expected to vary significantly between regions with and without active sources. This choice is similar to the recent work in the application of ARD to E/MEG reconstruction [15,17], except that their work assumes independent γ^2 for each location in the brain.

Since the forward model $\mathbf{A}$ is underdetermined, the current distribution $\mathbf{J}$, produced by our neurovascular coupling model, can fully explain the E/MEG data. In other words, without the noise term η_k^2 (i.e., $\mathbf{j}_n = z_n \mathbf{u}_k$), the fMRI data can exert too much influence on the reconstruction results. Although we can estimate the noise variance of the current source timecourses η_k^2 by extending the inference procedure, we find the corresponding estimate unstable without a prior. Based on preliminary empirical testing, we fix $\eta_k^2 = 1$. With proper temporal whitening of the fMRI data, we can also assume that $\xi_k^2 = \eta_k^2$. Fixing $\eta_k^2 = \xi_k^2$ helps to significantly reduce the computational burden of the estimation.

To summarize, our model can be mathematically expressed as

$$p(\mathbf{Y}, \mathbf{J}, \mathbf{F}, \mathbf{z}; \mathbf{\Theta}) = p(\mathbf{Y}|\mathbf{J})p(\mathbf{J}, \mathbf{F}|\mathbf{z}; \mathbf{\Theta})p(\mathbf{z}; \mathbf{\Theta}), \tag{3}$$

where $\mathbf{\Theta} = [\theta_1, \theta_2, \cdots, \theta_K]$ is the combined set of parameters, and $\theta_k = \{\mathbf{u}_k, \mathbf{v}_k, \gamma_k^2\}$. $p(\mathbf{Y}|\mathbf{J})$ is the E/MEG data model in Eq. (2). $p(\mathbf{J}, \mathbf{F}|\mathbf{z}; \mathbf{\Theta})$ is our neurovascular coupling model in Eq. (1), and $p(\mathbf{z}; \mathbf{\Theta})$ is the prior on $\mathbf{z}$. Therefore,

$\log p(\mathbf{Y}, \mathbf{J}, \mathbf{F}, \mathbf{z}; \mathbf{\Theta}) = \sum_{t=1}^{T_\mathbf{J}} \log \mathcal{N} (\mathbf{y}(t); \mathbf{A}\mathbf{j}(t), \mathbf{I}) +$
$\sum_{k=1}^{K} \sum_{n=1}^{N_k} \log \left[\mathcal{N} \left(\mathbf{j}_n; z_n \mathbf{u}_k, \eta_k^2 \mathbf{I} \right) \mathcal{N} \left(\mathbf{f}_n; z_n \mathbf{v}_k, \xi_k^2 \mathbf{I} \right) \right] + \sum_{k=1}^{K} \log \mathcal{N} \left(\mathbf{z}_k; \mathbf{0}, \gamma_k^2 \mathbf{\Phi}_k \right).$

2.3   Inference

Our goal is to estimate the current source $\mathbf{J}$ and the timecourses $\mathbf{u}$ and $\mathbf{v}$. We treat the activation strength $\mathbf{z}$ as an auxiliary variable, and marginalize it out in the analysis. We formulate the inference as

$$\{\mathbf{J}^*, \ \mathbf{\Theta}^*\} = \underset{\mathbf{J}, \ \mathbf{\Theta}}{\arg\max} \log p(\mathbf{Y}, \mathbf{J}, \mathbf{F}; \mathbf{\Theta})$$

$$= \underset{\mathbf{J}, \ \mathbf{\Theta}}{\arg\max} \log \int_{\mathbf{z}} p(\mathbf{Y}, \mathbf{J}, \mathbf{F}, \mathbf{z}; \mathbf{\Theta}) d\mathbf{z} = \underset{\mathbf{J}, \ \mathbf{\Theta}}{\arg\max} \log \left(p(\mathbf{Y}|\mathbf{J})p(\mathbf{J}, \mathbf{F}; \mathbf{\Theta}) \right). \tag{4}$$

With marginalization of $\mathbf{z}$, $p(\mathbf{J}, \mathbf{F}; \mathbf{\Theta})$ acts as the prior for $\mathbf{J}$. Since both $\mathbf{J}$ and $\mathbf{F}$ are linear functions of $\mathbf{z}$, $p(\mathbf{J}, \mathbf{F}; \mathbf{\Theta})$ is a continuous Gaussian mixture model.

The difficulty in inference with the proposed model is caused by the intertwining between space and time, reflected by the intersection of the temporal plates and the spatial plates in Fig. 1. That is because the output of a given E/MEG sensor is a mixture of signals from the entire source space. Hence, the inference

must be performed for all time points and all locations simultaneously. FIRE is thus substantially more computationally demanding compared to standard temporally independent E/MEG estimation and voxel-wise fMRI analysis.

Due to the special structure in our model when one set of variables is fixed, we can derive an efficient gradient descent method with two alternating steps. In the first step, we fix $\boldsymbol{\Theta}$ and derive a closed-form solution for $\mathbf{J}$. In the second step, we fix $\mathbf{J}$ and show that $\boldsymbol{\Theta}$ can be efficiently estimated through the EM algorithm. Section 5 discusses an alternative approach to this inference problem.

For a fixed $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Theta}}$, $p(\mathbf{Y}, \mathbf{J}, \mathbf{F}; \widehat{\boldsymbol{\Theta}})$ is a jointly-Gaussian distribution. Thus, the estimate of $\mathbf{J}$ is the conditional mean:

$$\widehat{\mathbf{J}} = \arg\max_{\mathbf{J}} \log p(\mathbf{Y}, \mathbf{J}, \mathbf{F}; \widehat{\boldsymbol{\Theta}}) = E\left[\mathbf{J} | \mathbf{Y}, \mathbf{F}; \widehat{\boldsymbol{\Theta}}\right] = \boldsymbol{\Gamma}_{\mathbf{w},\mathbf{J}}^{\mathrm{T}} \boldsymbol{\Gamma}_{\mathbf{w}}^{-1} \mathbf{w}, \tag{5}$$

where $\mathbf{w}^{\mathrm{T}} = \left[(\mathrm{vec}(\mathbf{Y}))^{\mathrm{T}} (\mathrm{vec}(\mathbf{F}))^{\mathrm{T}}\right]$ includes both E/MEG and fMRI measurements. Operator $\mathrm{vec}(\cdot)$ concatenates adjacent columns of a matrix. $\boldsymbol{\Gamma}_{\mathbf{w}}$ is the covariance matrix of $\mathbf{w}$, and $\boldsymbol{\Gamma}_{\mathbf{w},\mathbf{J}}$ is the cross-covariance matrix between $\mathbf{w}$ and $\mathrm{vec}(\mathbf{J})$. Thus, E/MEG and fMRI measurements jointly determine the estimate of the neural activity. Eq. (5) is similar to the standard MNE solution [10], but it also includes the correlation between $\mathbf{Y}$ and $\mathbf{F}$ and the correlation among different time points of $\mathbf{J}$.

For a fixed $\mathbf{J} = \hat{\mathbf{J}}$, we optimize the parameters $\boldsymbol{\Theta}$:

$$\widehat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \log p(\widehat{\mathbf{J}}, \mathbf{F}; \boldsymbol{\Theta}). \tag{6}$$

As shown in Fig. 1, when the current distribution $\mathbf{J}$ is fixed, the E/MEG measurement $\mathbf{Y}$ does not provide additional information for the parameter estimation. Since the parameters for different regions are independent for a fixed $\hat{\mathbf{J}}$, the estimates for different regions can be obtained independently. Furthermore, parameter $\boldsymbol{\Theta}$ can be efficiently estimated using the EM algorithm by re-introducing the latent variable $\mathbf{z}$, which is the auxiliary variable describing activation strength. This method can be thought of as an extension of the EM algorithm for probabilistic PCA [16] to two sets of data [2]. For region k, the parameter estimates $\widehat{\theta}_k$ can be obtained by optimizing the lower bound of the log-probability:

$$\log p\left(\{\widehat{\mathbf{j}}_n, \mathbf{f}_n\}_{n \in P_k}; \theta_k\right) \geq \int_{\mathbf{z}_k} q(\mathbf{z}_k) \log p\left(\{\widehat{\mathbf{j}}_n, \mathbf{f}_n\}_{n \in P_k}, \mathbf{z}_k; \theta_k\right) d\mathbf{z}_k, \tag{7}$$

where $q(\mathbf{z}_k) = p\left(\mathbf{z}_k | \{\widehat{\mathbf{j}}_n, \mathbf{f}_n\}_{n \in P_k}; \widehat{\theta}_k\right)$ is the posterior probability computed in the E-step. Since $\{\widehat{\mathbf{j}}_n, \mathbf{f}_n\}_{n \in P_k}$ and $\mathbf{z}_k$ are jointly-Gaussian distributed for a fixed $\widehat{\theta}_k$, $q(\mathbf{z}_k)$ is also a Gaussian distribution. We use $\langle \cdot \rangle_q$ to denote the expectation with respect to the posterior distribution $q(\mathbf{z}_k)$, i.e., $\langle \cdot \rangle_q \triangleq \left[\cdot | \{\widehat{\mathbf{j}}_n, \mathbf{f}_n\}_{n \in P_k}; \widehat{\theta}_k\right]$. Since the M-step depends only on quantities related to the first- and the second-

order statistics of $\mathbf{z}_k$, we only need to update those quantities in the E-step:

$$\langle\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}\rangle_q \leftarrow \left[\frac{1}{\gamma_k^2}\boldsymbol{\Phi}_k^{-1} + \left(\frac{\mathbf{u}_k^{\mathrm{T}}\mathbf{u}_k + \mathbf{v}_k^{\mathrm{T}}\mathbf{v}_k}{\eta_k^2}\right)\mathbf{I}\right]^{-1}$$

$$\langle\mathbf{z}_k\rangle_q \leftarrow \frac{\langle\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}\rangle_q}{\eta_k^2}\left[\left(\mathbf{u}_k^{\mathrm{T}}\widehat{\mathbf{j}}_1 + \mathbf{v}_k^{\mathrm{T}}\mathbf{f}_1\right), \cdots, \left(\mathbf{u}_k^{\mathrm{T}}\widehat{\mathbf{j}}_{N_k} + \mathbf{v}_k^{\mathrm{T}}\mathbf{f}_{N_k}\right)\right]^{\mathrm{T}}$$

$$\langle\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\Phi}_k^{-1}\mathbf{z}_k\rangle_q \leftarrow \langle\mathbf{z}_k\rangle_q^{\mathrm{T}}\boldsymbol{\Phi}_k^{-1}\langle\mathbf{z}_k\rangle_q + \mathrm{tr}\left(\boldsymbol{\Phi}_k^{-1}\langle\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}\rangle_q\right).$$

In the M-step, we fix $q(\mathbf{z}_k)$ and optimize Eq. (7). With some algebra, we arrive at the update equations for the model parameters:

$$\widehat{\mathbf{u}}_k \leftarrow \frac{\sum_{n\in P_k}\langle z_n\rangle_q\widehat{\mathbf{j}}_n}{\mathrm{tr}(\langle\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}\rangle_q)} \;,\;\; \widehat{\mathbf{v}}_k \leftarrow \frac{\sum_{n\in P_k}\langle z_n\rangle_q\mathbf{f}_n}{\mathrm{tr}(\langle\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}\rangle_q)} \;,\;\; \text{and} \;\; \widehat{\gamma_k^2} \leftarrow \frac{\langle\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\Phi}_k^{-1}\mathbf{z}_k\rangle_q}{N_k}. \tag{8}$$

We iterate the EM algorithm until convergence which usually takes less than ten iterations. We then re-estimate $\mathbf{J}$ according to Eq. (5).

To summarize, the algorithm proceeds as follows:

(i) Initialize $\hat{\mathbf{J}}$ as the MNE estimate.
(ii) Until convergence:
    1. Compute $\widehat{\boldsymbol{\Theta}}$ using the EM algorithm: E-step for the hidden variable $\mathbf{z}$ followed by M-step for the model parameters $\boldsymbol{\Theta}$.
    2. Update $\hat{\mathbf{J}}$ according to Eq. (5) for $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Theta}}$.

3   Implementation

For the computation of the forward matrix $\mathbf{A}$, we need to specify the E/MEG forward model and the source space. We employ the single-compartment boundary-element model for the MEG forward computations [9]. The source space is confined to a mesh on the cortical surface with approximately 5-mm resolution, corresponding to about 5000 vertices per hemisphere.

The functional regions are defined by parceling the cortical folding pattern using the FreeSurfer software, resulting in 35 parcels per hemisphere [7]. The boundaries of adjacent parcels are defined along sulci. We merge adjacent parcels that contain fewer than 30 vertices. Our neurovascular coupling model requires an orientation reference for each brain region. Here, we set the orientation reference to be the largest left singular vector of the matrix formed by the outward cortical normals within a region.

We apply the standard preprocessing to fMRI data, then estimate the hemo-dynamic response function (HRF) at each voxel with a 15-bin finite impulse response regressor covering a 20-s time window using the FS-FAST software (MGH, Boston, MA). The estimated HRF is used as $\mathbf{f}_n$ in our model. For a source space of $N \sim 10^4$ vertices and timecourses of $T_{\mathbf{J}} \sim 10^2$ and $T_{\mathbf{F}} \sim 10^1$ samples, FIRE takes less than 20 iterations until the energy function reduces less than 0.1% from the energy of the previous iteration. In each iteration of

the coordinate descent algorithm, the estimate of $\boldsymbol{\Theta}$ takes 30 seconds, while the estimate of $\mathbf{J}$ takes 4 minutes on a standard PC (2.8 GHz CPU and 8 GB RAM), leading to the total run time of approximately 1.5 hours. Estimating $\mathbf{J}$ involves an inversion of an $(MT_{\mathbf{J}} + NT_{\mathbf{F}}) \times (MT_{\mathbf{J}} + NT_{\mathbf{F}})$ dense symmetric matrix $\boldsymbol{\Gamma}_{\mathbf{w}}$, which is too large to store in memory. Instead, we employ the conjugate gradient descent method to solve the corresponding system of linear equations. It usually takes 100 iterations until convergence.

4 Results

We first compare FIRE to MNE, fMNE, and fARD using simulated data. We then extend the comparison to human MEG and fMRI data from a somatosensory study.

4.1 Simulation Studies

To simulate MEG measurements, we created two patches on the cortical sheet, with current source orientation along the outward normal to the cortical surface. Shown in the lateral-occipital view of the right hemisphere (Fig. 2), Patch A contains 20 vertices and is located in the inferior parietal region. Patch B contains 32 vertices and is located in the superior parietal region. We simulated neural and vascular timecourses in these two patches for three different scenarios: no silent activity, silent vascular activity, and silent neural activity. In the two cases with silent activities, we kept the activity of patch B unchanged while silencing neural or vascular activity in patch A. The simulated neural signals are shown as solid black lines in the rightmost column of Fig. 2. The activation maps corresponding to the peaks of the two simulated neural signals are shown in the first column.

For the forward calculations, we employed the sensor configuration of the 306-channel Neuromag VectorView MEG system used in our human studies and added Gaussian noise to the signals. The resulting signals have a SNR of 3 dB, within the typical SNR range of real MEG data. Since the two patches are close in the highly folded cortex and they exhibit neural activity during overlapping time intervals, it is particularly difficult to obtain accurate current source estimates.

Columns two to five in Fig. 2 depict the current estimates using different methods. Following [12], the fMNE weighting parameters are set to 1 and 0.1 for active and inactive fMRI locations, respectively. Since the estimates from different methods are not directly comparable in amplitude, the threshold for each method is chosen to be $1/6$ of the maximum absolute value of the corresponding current estimates $\mathbf{J}^*$. The rightmost column in Fig. 2 presents the estimated timecourses (dashed) of the most active vertex, in terms of energy, in both patches.

No Silent Activity. As shown in Fig. 2(a), the MNEs extend across adjacent gyri. fMNE, fARD, and FIRE correctly localize the two patches at the peak activation, but FIRE provides a better estimate of the spatial extent of the activations. The fARD estimate is unstable, as reflected by the large fluctuations in the estimated timecourses in patch B (green).

Silent Vascular Activity. When the vascular activity in patch A is silent, fMNE shows excessive bias towards patch B. Without a large weight, the amplitude of the estimated timecourses (blue) in patch A is significantly lower than the corresponding estimates in patch B. It would be therefore easy to miss neural activation in patch A when interpreting the results (column three in Fig. 2(b)). In contrast, by combining neural and vascular information in the re-weighted scheme, FIRE avoids such a bias. Its estimate in patch A (column five) is similar to that obtained from MNE (column one). As the weights for patch B increase and the weights for patch A decrease in the fARD update, the estimate in patch B explains the activation in patch A. As shown in the timecourse panel, the estimated timecourse in patch B (green) is similar in shape to the simulated timecourse in patch A (black solid). The change of sign is due to the fact that the outward normals for patch A and patch B are in approximately opposite directions.

Silent Neural Activity. As shown in Fig. 2(c), all methods can correctly localize the neural activity in patch B, except for the small false positive in patch A for fARD. By assigning identical weights to patches A and B, fMNE estimates a timecourse for patch A (blue) that is noisier than the corresponding one produced by FIRE (red). FIRE suppresses the weights for patch A since the current estimates in that patch are close to zero; its results are closer to the simulations.

4.2   Median-Nerve Experiments

We also tested the method using human experimental data. The median nerve at the right wrist was stimulated according to an event-related protocol, with a random inter-stimulus-interval ranging from 3 to 14 s. This stimulus activates a complex cortical network [11], including the contralateral primary somatosensory cortex (cSI) and bilateral secondary somatosensory cortices (cSII and iSII).

MEG and fMRI data were acquired in separate sessions. The MEG measurements were acquired using a 306-channel Neuromag VectorView MEG system. A 200-ms baseline before the stimulus was used to estimate the noise covariance matrix of the MEG sensors. An average signal, computed from approximately 100 trials, was used as the input to each method. The fMRI images were acquired using a Siemens 3T machine (TR=1.5 s, $64{\times}64{\times}24$, $3{\times}3{\times}6\,\mathrm{mm}^3$, single channel head coil). Anatomical images, from a 3T scanner, were used to construct the source space and the forward model.

In the leftmost column in Fig. 3, approximate locations for cSI (solid), cSII (dashed), and iSII (dashed) are highlighted on the fMRI activation maps ($p \leq$ 0.005). Given the expected activations, we partitioned the post-central region into two regions, separately covering cSI and cSII. Note that in the noisy SPM, the sites of fMRI activations do not exactly agree with the locations of the expected current sources.

Columns two to five in Fig. 3 present the estimates at 75 ms after stimulus onset, during which cSI, cSII, and iSII should be activated. The threshold was set separately for each hemisphere since the activation in iSII is much weaker
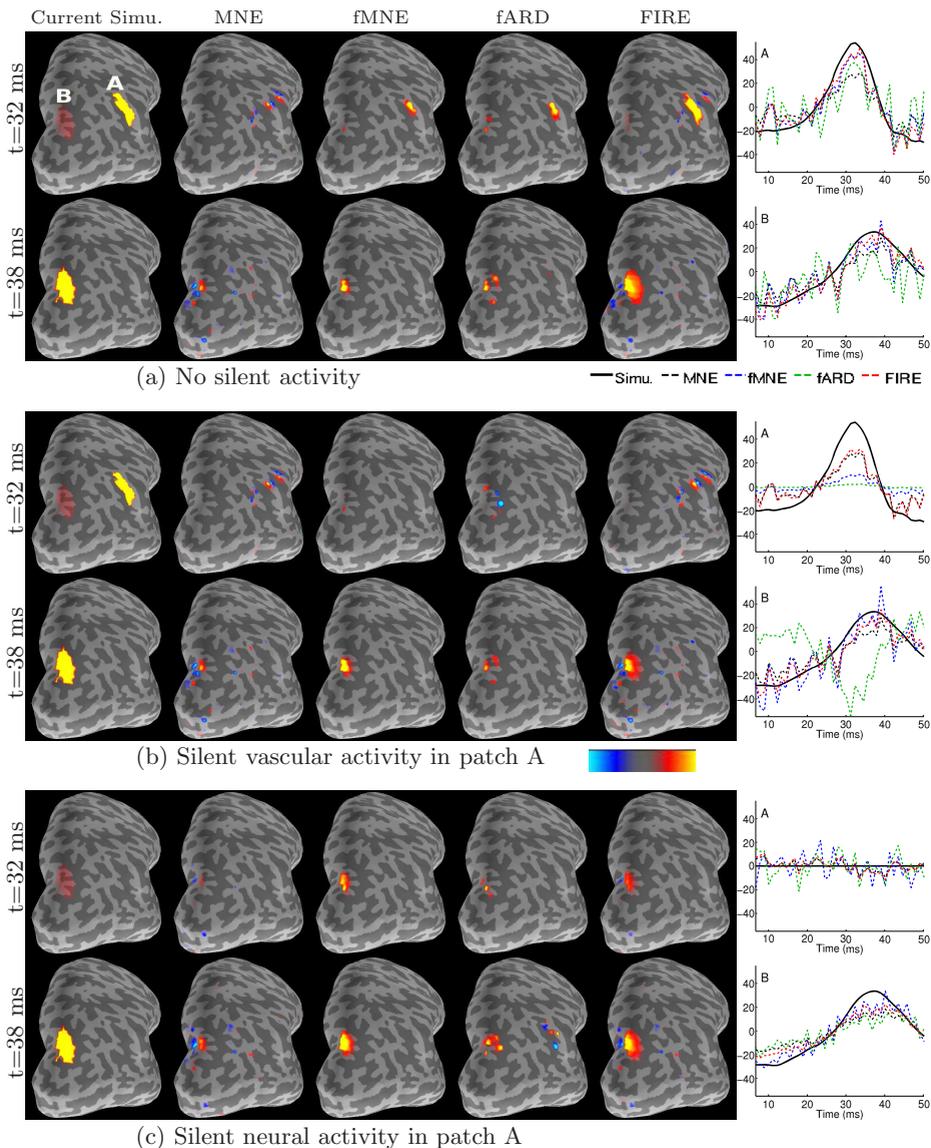
Fig. 2. Current source estimates in three scenarios. Lateral-occipital view of the right hemisphere is shown. Patch A and patch B are highlighted in the top left panel; the rest of the figures follow the same convention. (a) Neither neural nor vascular activity is silent. (b) Vascular activity in patch A is silent. (c) Neural activity in patch A is silent. The first column illustrates the simulated current distributions with a selected threshold at the peak activations. The next four columns show the estimates from MNE, fMNE, fARD, and FIRE. Hot/cold colors correspond to outward/inward current flow. The rightmost column shows the simulated (black solid) and the estimated (dashed) timecourses from the most active vertices in patch A (top) and B (bottom) for the corresponding methods.
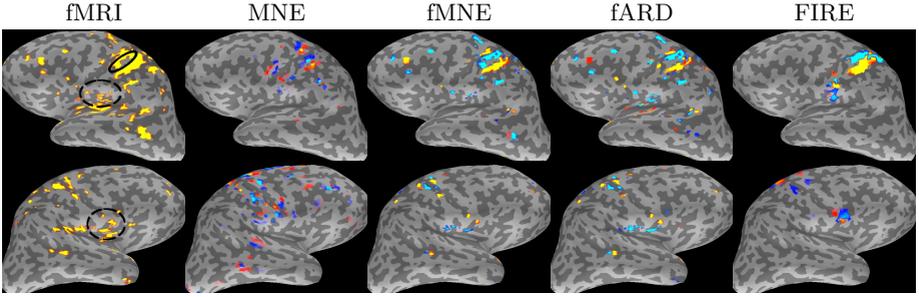
Fig. 3. Human median-nerve experiments. In the first column, approximate locations for cSI (solid), cSII (dashed), and iSII (dashed) are highlighted on the fMRI activation maps. Columns two to five show the current estimates obtained via MNE, fMNE, fARD, and FIRE at 75 ms after stimulus onset. Hot/cold colors indicate outward/inward current flow.

than that in cSI and cSII. For each method, the threshold is set to be 1/6 of the maximum absolute value of the corresponding current estimates. MNE produces a more diffuse estimate, including physiologically unlikely activations at the gyrus anterior to the cSI area. In contrast, FIRE pinpoints cSI on the post-central gyrus. With the prior knowledge from fMRI, the detected cSII and iSII activations using fMNE, fARD, and FIRE are within the expected areas. The fMNE and fARD show stronger weighting towards the fMRI, reflected by the activations in the temporal lobes. Due to the highly folded cortex and uncertainties in MRI-fMRI registration, fMRI cannot distinguish between the walls of the central sulcus and the post-central sulcus, causing both walls to show strong vascular activity after mapping of the fMRI volume onto the cortex. Hence, fMNE, fARD, and FIRE estimates extend to both sulcal walls mentioned above.

5   Discussion

The coupling of spatial and temporal domains in the joint fMRI-E/MEG analysis constrained many previous models to operate on a coarse source space. The use of region-based neurovascular coupling model proposed in this paper reduces the computational burden, leading to a tractable reconstruction in a densely sampled source space similar to that typically used in MNE. Since the objective function is not convex, FIRE depends on the initialization. We believe MNE estimate is a reasonable choice for initialization since it is unbiased.

As an alternative to the coordinate descent inference procedure proposed in this work, one could treat both $\mathbf{J}$ and $\mathbf{z}$ as latent variables in the EM framework. Since $\mathbf{J}$, $\mathbf{z}$, and the measurements are jointly Gaussian, the posterior distribution of the latent variables is also Gaussian, leading to a closed-form update. Similar to the derivations in Eq. (8), the M-step updates depend on the second order statistics of the latent variables. Since $\mathbf{J}$ is not fixed in this EM procedure, the estimate at each location depends on the estimate at all other locations in the

source space, as opposed to the region-independent estimation when $\mathbf{J}$ is fixed. Therefore, the computation of the second-order statistics is infeasible except for an extremely coarse discretization of the source space.

The estimation of $\mathbf{u}_k$ and $\mathbf{v}_k$ is closely related to the canonical correlation analysis (CCA). CCA seeks vectors to project two high dimensional data sets ($\{\hat{\mathbf{j}}_n\}_{n \in P_k}$ and $\{\mathbf{f}_n\}_{n \in P_k}$ in our case) to a low dimensional space so as to maximize the correlation coefficient. The probabilistic interpretation of CCA has been established in [2].

Our neurovascular coupling model is designed for fixed-orientation current estimates, since the latent-variable model assumes that the spatial concordance of neural and vascular activities is characterized by a scalar. For free-orientation current estimates, the neurovascular coupling model would have to be adjusted to handle the correspondence between the current flow in three directions and a single vascular activation timecourse at a certain location. Moreover, FIRE assumes a single activation waveform pair, $\mathbf{u}$ and $\mathbf{v}$, in a region. The validity of this assumption depends on the size of the region and the distance between two activation sources. We cannot directly extend FIRE to multiple activation waveform pairs per region, since such an extension does not capture the fact that the shape of the vascular activation timecourses from two distinct sources is often highly similar but the neural processes are different. In the situation where there are two distinct current sources in one region, our preliminary results demonstrate that FIRE can localize the two current sources, but the estimated timecourses are combinations of the true timecourses. We defer the extension for free-orientation estimate and the extension for multiple activation sources per region to future work.

6   Conclusions

In contrast to most joint fMRI-E/MEG models, FIRE explicitly takes into account the inherent differences in the data measured by E/MEG and fMRI. The corresponding estimates can be efficiently computed with an iterative procedure which bears similarity with re-weighted MNE methods, except that the weights are based on both the current estimates in the last iteration and the fMRI data via the proposed neurovascular coupling model. This construction of the weights reduces the excessive sensitivity to fMRI present in many joint fMRI-E/MEG analysis methods, leading to more accurate current estimates as demonstrated by analysis of both simulated and human data.

References

1. Ahlfors, S., Simpson, G.: Geometrical interpretation of fMRI-guided MEG/EEG inverse estimates. NeuroImage 22, 323–332 (2004)
2. Bach, F., Jordan, M.: A probabilistic interpretation of canonical correlation analysis. Technical Report 688, UC Berkeley (2005)
3. Baillet, S., et al.: Electromagnetic brain mapping. IEEE Sig. Proc. Mag. (2001)
4. Daunizeau, J., et al.: Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. NeuroImage 36, 69–87 (2007)
5. Dempster, A., et al.: Maximum likelihood from incomplete data via the EM algorithm. J. of Roy. Stat. Soc. B 39, 1–38 (1977)
6. Deneux, T., Faugeras, O.: EEG-fMRI fusion of non-triggered data using Kalman filtering. In: ISBI, pp. 1068–1071 (2006)
7. Fischl, B., et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355 (2002)
8. Gorodnitsky, I., Rao, B.: Sparse signal reconstruction from limited data using FO-CUSS: a re-weighted MNE algorithm. IEEE Trans. Sig. Proc. 45, 600–616 (1997)
9. Hämäläinen, M., Sarvas, J.: Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Biomed. Eng. 36, 165–171 (1989)
10. Hämäläinen, M., et al.: Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65, 413–497 (1993)
11. Hari, R., Forss, N.: Magnetoencephalography in the study of human somatosensory cortical processing. Philos. Trans. R. Soc. Lond. B 354, 1145–1154 (1999)
12. Liu, A., et al.: Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. PNAS 95, 8945–8950 (1998)
13. Logothetis, N., Wandell, B.: Interpreting the BOLD signal. Annu. Rev. Physiol. 66, 735–769 (2004)
14. Pascual-Marqui, R., et al.: Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65 (1994)
15. Sato, M., et al.: Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage 23, 806–826 (2004)
16. Tipping, M., Bishop, C.: Probabilistic principal component analysis. J. Royl. Stat. Soc. B 61, 611–622 (1999)
17. Wipf, D., Nagarajan, S.: A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44, 947–966 (2009)