SYMMETRICAL EEG/FMRI FUSION WITH SPATIALLY ADAPTIVE PRIORS USING VARIATIONAL DISTRIBUTION APPROXIMATION

Martin Luessi[1], *S. Derin Babacan*[1], *Rafael Molina*[2], *James R. Booth*[3], *Aggelos K. Katsaggelos*[1]

[1]Dept. of EECS
[3]Dept. of Com. Sciences and Disorders
Northwestern University, Evanston, IL 60208, USA

[2]Departamento de Ciencias
de la Computación e I.A.
Universidad de Granada, 18071 Granada, Spain

ABSTRACT

In this paper, we propose a symmetrical EEG/fMRI fusion algorithm which combines EEG and fMRI by means of a common generative model. The use of a total variation (TV) prior as well as spatially adaptive temporal priors enables adaptation to the local characteristics of the estimated responses. We utilize an approximate variational Bayesian framework and obtain a fully automatic fusion algorithm. Simulation results demonstrate that the proposed algorithm outperforms existing EEG/fMRI fusion methods.

Index Terms— EEG, fMRI, total variation (TV), variational Bayesian methods.

1. INTRODUCTION

Functional neuroimaging allows for spatial and temporal localization of neurological activity. It is therefore an invaluable tool for gaining a better understanding of the brain. In recent years, functional magnetic resonance imaging (fMRI) has become a prominent neuroimaging method as it enables accurate localization of neurological activity, typically within millimeters. However, the temporal resolution of fMRI is in the order of seconds, which limits its use for analyzing the temporal aspects of neurological activity. Electroencephalography (EEG) is another functional neuroimaging method, which measures the electrical activity in the brain via electrodes placed on the surface of the scalp. While EEG can achieve submillisecond resolution, the spatial resolution is considered poor as it is generally not possible to determine the source of the activity using EEG recordings alone. This limitation is due to the fact that EEG source localization using the distributed source framework [1] involves solving the following ill-posed inverse problem

$$\mathbf{M} = \mathbf{LS} + \eta_1, \tag{1}$$

where $\mathbf{M}$ is a $m \times t_1$ matrix containing the EEG recordings obtained from m electrodes placed on the scalp, $\mathbf{S}$ is a $n \times t_1$ matrix representing the responses of n normal-oriented current dipoles distributed on the cortical surface, $\mathbf{L}$ is a known $m \times n$ forward operator, which can be calculated using anatomical information, and η_1 is a $m \times t_1$ matrix representing noise. Clearly, estimating $\mathbf{S}$ is an ill-posed problem, since the number of dipoles n is typically several orders of magnitudes larger than the number of electrodes m (a typical scenario is $m = 64$, $n = 5000$). In order to find a unique solution it is necessary to make additional assumptions, which can be incorporated in the form of deterministic regularization terms, or in the form of priors when a Bayesian formulation is used. In our work we follow the approach where spatial information from fMRI is utilized in order to find a more meaningful solution to the inverse problem. The idea is that the sources of hemodynamic activity measured by fMRI often coincide with sources of electrical activity measured by EEG. This concept was used in the fMRI weighted minimum norm algorithm [2] and more recently in [3], where a Bayesian framework is used to "fuse" fMRI and EEG in a symmetrical fashion. We follow the symmetrical approach and augment it in several ways leading to significant improvements in estimation quality. More specifically, we use a total variation (TV) prior for the spatial profile of the cortical activity, which leads to spatially adaptive spatial smoothness and consequently to the preservation of sharp boundaries at the edges of active brain regions. We also utilize spatially adaptive temporal priors, allowing for adaptation of the degree of temporal smoothness according to the estimated activity in different brain regions. We utilize approximate variational Bayesian inference methods and obtain a fully automatic EEG/fMRI fusion algorithm. Using simulations we demonstrate that the proposed algorithm outperforms existing EEG/fMRI fusion methods.

We use the following notation throughout this paper: $\mathbf{A}_{i\cdot}$ denotes a row vector containing the elements of the i-th row of $\mathbf{A}$, while $\mathbf{A}_{\cdot i}$ is a column vector containing the elements of the i-th column of $\mathbf{A}$. The operator $\mathrm{diag}\,(\mathbf{A})$ extracts the main diagonal of $\mathbf{A}$ as a column vector, whereas $\mathrm{Diag}\,(\mathbf{a})$ is a diagonal matrix with $\mathbf{a}$ as its diagonal. The operator $\mathrm{vec}\,(\mathbf{A})$ vectorizes $\mathbf{A}$ by stacking its columns and $\otimes$ denotes the Kronecker product.

The paper is organized as follows. In Section 2, the hierarchical generative model is introduced. The Bayesian inference procedure is shown in Section 3. Experimental results are presented in Section 4 and conclusions are drawn in Section 5.

2. HIERARCHICAL GENERATIVE MODEL

In this section we define a joint hierarchical generative model which explains both, the EEG and fMRI observations. Since there is no generally accepted model for the coupling between hemodynamic and current responses, we resort to the symmetrical EEG/fMRI fusion model by Daunizeau et al. [3]. The model assumes that the electric activity can be described by q temporally coherent parcels, each having a distinct response waveform. This assumption is formalized by

$$\mathbf{S} = \mathrm{Diag}\left(\mathbf{w}^{\mathrm{EEG}}\right)\mathbf{CX} + \mathbf{R}_1, \tag{2}$$

where $\mathbf{C}$ is a $n \times q$ matrix defining the parceling ($\mathbf{C}_{ij} = 1$ if $i \in P_j$, the j-th parcel, $\mathbf{C}_{ij} = 0$ otherwise), $\mathbf{w}^{\mathrm{EEG}}$ is a $n \times 1$ vector representing the unknown spatial profile of electric activity, $\mathbf{X}$ is a $q \times t_1$ matrix with the unknown response waveforms, and $\mathbf{R}_1$ is a $n \times t_1$ matrix representing residual activity which cannot be explained by the model.

In order to obtain a similar hierarchical description for fMRI,

we first introduce the fMRI observation model [4], which models the measurements by

$$\mathbf{Y} = \mathbf{B}\mathbf{H} + \eta_2, \tag{3}$$

where $\mathbf{Y}$ is the $t_2 \times n$ matrix containing the fMRI measurements at n locations on the cortical surface (we assume here that the locations coincide with the locations of the EEG current dipoles), $\mathbf{H}$ is a $k \times n$ matrix representing the hemodynamic response function (HRF) of length k at each location, η_2 is the $t_2 \times n$ matrix with additive noise, and $\mathbf{B}$ is a known $t_2 \times k$ convolution matrix describing the experimental protocol, as defined in [4]. By making the same coherency assumptions as for EEG, a hierarchical description of the HRFs for all locations is obtained

$$\mathbf{H} = \mathbf{Z}\mathbf{C}^T \mathrm{Diag}\left(\mathbf{w}^{\mathrm{fMRI}}\right) + \mathbf{R}_2, \tag{4}$$

where $\mathbf{Z}$ is a $k \times q$ matrix containing the unknown HRFs of the parcels, $\mathbf{w}^{\mathrm{fMRI}}$ is a $n \times 1$ vector describing the spatial profile, and $\mathbf{R}_2$ is a $k \times n$ matrix representing the modeling residual. In order to establish a connection between the imaging modalities, a common spatial profile is assumed, i.e.,

$$\mathbf{w}^{\mathrm{EEG}} = \mathbf{w}^{\mathrm{fMRI}} = \mathbf{w}. \tag{5}$$

2.1. Bayesian Modeling

We proceed by defining a graphical model. As depicted in Fig. 1 and explained next. By defining a probability density function (pdf) for each node of the model we can obtain the joint pdf and consequently derive a Bayesian inference procedure, as described in the next section. The acquisition noise for EEG and fMRI is assumed to be
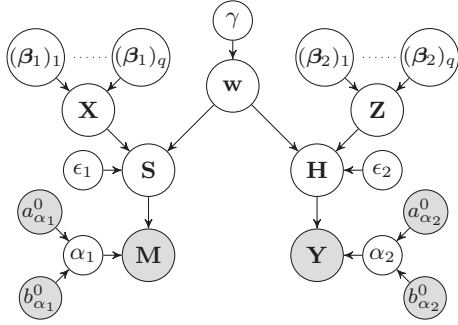

Fig. 1. Graphical Model (gray: known, white: unknown)

zero mean, i.i.d., Gaussian, resulting in the following observation probabilities

$$\mathrm{p}\left(\mathbf{M}|\mathbf{S}, \alpha_1\right) = \prod_{i=1}^{t_1} \mathcal{N}\left(\mathbf{M}_{\cdot i}|\mathbf{L}\mathbf{S}_{\cdot i}, \alpha_1^{-1}\mathbf{I}_m\right), \tag{6}$$

$$\mathrm{p}\left(\mathbf{Y}|\mathbf{H}, \alpha_2\right) = \prod_{i=1}^{n} \mathcal{N}\left(\mathbf{Y}_{\cdot i}|\mathbf{B}\mathbf{H}_{\cdot i}, \alpha_2^{-1}\mathbf{I}_{t_2}\right), \tag{7}$$

where α_1 and α_2 are the noise precisions for EEG and fMRI, respectively. We make the same assumptions about the modeling residuals and obtain the following hierarchical priors for the current responses and HRFs

$$\mathrm{p}\left(\mathbf{S}|\mathbf{X}, \mathbf{w}, \epsilon_1\right) = \prod_{i=1}^{t_1} \mathcal{N}\left(\mathbf{S}_{\cdot i}|\mathrm{Diag}\left(\mathbf{w}\right)\mathbf{C}\mathbf{X}_{\cdot i}, \epsilon_1^{-1}\mathbf{I}_n\right), \tag{8}$$

$$\mathrm{p}\left(\mathbf{H}|\mathbf{Z}, \mathbf{w}, \epsilon_2\right) = \prod_{i=1}^{k} \mathcal{N}\left(\mathbf{H}_{i\cdot}^T|\mathrm{Diag}\left(\mathbf{w}\right)\mathbf{C}\mathbf{Z}_{i\cdot}^T, \epsilon_2^{-1}\mathbf{I}_n\right). \tag{9}$$

The Bayesian formulation allows for an elegant way to model prior knowledge about the nature of the problem by means of prior distributions. For the problem at hand, we know that event related processing in the brain occurs in a number of specialized brain regions. Consequently it can be assumed that the spatial profile $\mathbf{w}$ contains sharp boundaries between active and inactive regions. We incorporate this *a priori* assumption by utilizing a total variation (TV) prior given by the Gibbs distribution

$$\mathrm{p}\left(\mathbf{w}|\gamma\right) = \frac{1}{\mathrm{Z}(\gamma)} \exp\left(-\gamma \mathrm{TV}\left(\mathbf{w}\right)\right), \tag{10}$$

where $\mathrm{Z}(\gamma)$ is the partition function and $\mathrm{TV}(\cdot)$ is the energy function of the form $\mathrm{TV}(\mathbf{w}) = \sum_{c\in\mathcal{C}} \mathrm{V}_c(\mathbf{w})$, where $\mathcal{C}$ denotes a set of cliques and V_c is a potential function defined on a clique. We define a discrete version of the TV integral for $\mathbf{w}$ as follows

$$\mathrm{TV}(\mathbf{w}) = \sum_{i=1}^{n} \sqrt{\sum_{j\in c_i, i\neq j} (\mathbf{w}_i - \mathbf{w}_j)^2}, \tag{11}$$

where c_i denotes a clique of size 3 for the i-th dipole. The cliques are obtained by a simple heuristic algorithm which selects two neighbors for each vertex from the graph describing a triangular tessellation of the cortical surface. Due to space limitations, the details of the algorithm are omitted.

In order to be able to estimate the parameter γ, we need to approximate the partition function $\mathrm{Z}(\gamma)$. Note that we can calculate the partition function for a single location on the cortex as

$$\int\int \exp\left(-\gamma\sqrt{u^2 + v^2}\right) du dv = \frac{2\pi}{\gamma^2}. \tag{12}$$

Using this we approximate $\mathrm{p}\left(\mathbf{w}|\gamma\right)$ in (10) as

$$\mathrm{p}\left(\mathbf{w}|\gamma\right) = c\gamma^{\varphi n} \exp\left(-\gamma \mathrm{TV}\left(\mathbf{w}\right)\right), \tag{13}$$

where c is a constant and φ is a parameter with a value of $\varphi = 2.0$ if all locations are assumed to be independent. Due to the dependence between the locations, we empirically find that using $\varphi = 1.0$ improves the performance of the algorithm.

We also make the *a priori* assumption that the hemodynamic response functions and the current responses are smooth in the temporal dimension. This assumption can be expressed by a Gaussian prior which penalizes the second order temporal derivative; a prior of this form was also used in [3, 4]. In contrast to previous work, we assume that the temporal characteristics of the responses vary across different brain regions. We model this by utilizing a separate Gaussian prior for every parcel, i.e.,

$$\mathrm{p}\left(\mathbf{X}|\boldsymbol{\beta}_1\right) = \prod_{i=1}^{q} \mathcal{N}\left(\mathbf{X}_{i\cdot}^T|\mathbf{0}, (\boldsymbol{\beta}_1)_i^{-1}\left(\mathbf{T}_{(t_1,1)}^T \mathbf{T}_{(t_1,1)}\right)^{-1}\right), \tag{14}$$

$$\mathrm{p}\left(\mathbf{Z}|\boldsymbol{\beta}_2\right) = \prod_{i=1}^{q} \mathcal{N}\left(\mathbf{Z}_{\cdot i}|\mathbf{0}, (\boldsymbol{\beta}_2)_i^{-1}\left(\mathbf{T}_{(k,1)}^T \mathbf{T}_{(k,1)}\right)^{-1}\right), \tag{15}$$

where $\mathbf{T}_{(a,b)}$ is a $ab \times ab$ matrix given by

$$(\mathbf{T}_{(a,b)})_{ij} = \begin{cases} -2 & \text{if} \quad i = j, \\ 1 & \text{if} \quad j = i \pm b, , \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are $q \times 1$ vectors with per-parcel smoothness precision hyperparameters. The use of separate hyperparameters allows for spatially adaptive temporal smoothness, i.e., the model can reduce the degree of temporal smoothness in active regions while enforcing more smoothness in inactive regions.

Following the Bayesian approach we proceed by defining priors for all hyperparameters of the model. We use gamma densities for

the noise precisions, i.e.,

$$p(\alpha_1) = \Gamma\left(\alpha_1|a_{\alpha_1}^0, b_{\alpha_1}^0\right), \quad p(\alpha_2) = \Gamma\left(\alpha_2|a_{\alpha_2}^0, b_{\alpha_2}^0\right), \quad (17)$$

where the shape and scale parameters are obtained from pre-stimulus data (see [3]). We make no assumptions about the remaining hyperparameters and consequently use noninformative Jeffreys priors given by

$$p(\theta_i) \propto (\theta_i)^{-1} \quad \forall \quad \theta_i \in \theta, \quad (18)$$

where $\theta = \{(\boldsymbol{\beta}_1)_1 \ldots (\boldsymbol{\beta}_1)_q, (\boldsymbol{\beta}_2)_1 \ldots (\boldsymbol{\beta}_2)_q, \epsilon_1, \epsilon_2, \gamma\}$.

3. BAYESIAN INFERENCE

Inference is based on the posterior distribution

$$\begin{aligned}
p(\Theta|\mathbf{M}, \mathbf{Y}) \propto & p(\mathbf{M}|\mathbf{S}, \alpha_1) p(\mathbf{S}|\mathbf{X}, \mathbf{w}, \epsilon_1) p(\mathbf{X}|\boldsymbol{\beta}_1) \\
& \times p(\mathbf{Y}|\mathbf{H}, \alpha_2) p(\mathbf{H}|\mathbf{Z}, \mathbf{w}, \epsilon_2) p(\mathbf{Z}|\boldsymbol{\beta}_2) \quad (19) \\
& \times p(\mathbf{w}|\gamma) p(\theta),
\end{aligned}$$

where $\Theta = \{\mathbf{S}, \mathbf{H}, \mathbf{w}, \mathbf{X}, \mathbf{Z}, \alpha_1, \alpha_2\} \cup \theta$. Since calculating $p(\Theta|\mathbf{M}, \mathbf{Y})$ is intractable, we employ the mean field approximation, i.e., we find a distribution which factorizes over the nodes of the graphical model

$$\begin{aligned}
q(\Theta) = & q(\mathbf{S}) q(\mathbf{H}) q(\mathbf{w}) q(\mathbf{X}) q(\mathbf{Z}) q(\alpha_1) q(\alpha_2) \\
& \times q(\epsilon_1) q(\epsilon_2) q(\gamma) \prod_{i=1}^{q} q((\boldsymbol{\beta}_1)_i) \prod_{i=1}^{q} q((\boldsymbol{\beta}_2)_i). \quad (20)
\end{aligned}$$

The approximate posterior is found by performing a variational minimization of the Kullback-Leibler (KL) divergence between the true posterior and $q(\Theta)$, a technique known as the Variational Bayesian (VB) method. As it turns out, the form of the TV prior prevents us from finding $q(\Theta)$ analytically; we resort to the method proposed in [5] to find a lower bound to $p(\mathbf{w}|\gamma)$ for which the variational minimization can be performed. First, consider the inequality

$$\sqrt{ab} \leq \frac{a+b}{2} \Rightarrow \sqrt{a} \leq \frac{a+b}{2\sqrt{b}}, \quad (21)$$

for $a \geq 0$ and $b \geq 0$. We proceed by defining for $\mathbf{w}$, γ, and an $n \times 1$ vector $\mathbf{u} \in (\mathbb{R}^+)^n$, the following functional:

$$\begin{aligned}
& \mathrm{M}(\mathbf{w}, \mathbf{u}, \gamma) = \\
& c\gamma^{\varphi n} \exp\left(-\frac{\gamma}{2} \sum_{i=1}^{n} \frac{\left[\sum_{j \in c_i, i \neq j} (w_i - w_j)^2\right] + u_i}{\sqrt{u_i}}\right). \quad (22)
\end{aligned}$$

By comparing (13) with (22) and inequality (21) with $a = \sum_{j \in c_i, i \neq j} (w_i - w_j)^2$ and $b = u_i$, we obtain

$$p(\mathbf{w}|\gamma) \geq \mathrm{M}(\mathbf{w}, \mathbf{u}, \gamma). \quad (23)$$

Using (23) in (19), we obtain a lower bound to the posterior, i.e.,

$$p(\Theta|\mathbf{M}, \mathbf{Y}) \geq \mathrm{M}(\Theta, \mathbf{u}, \mathbf{M}, \mathbf{Y}). \quad (24)$$

As in the regular VB framework, we proceed by performing a variational minimization of the KL divergence. However, instead of the true posterior (19) we utilize its lower bound and obtain

$$q(\Theta_i) \propto \exp\left(\mathbb{E}_{\Theta \setminus \Theta_i}[\ln \mathrm{M}(\Theta, \mathbf{u}, \mathbf{M}, \mathbf{Y})]\right), \quad (25)$$

where the expectation is calculated over all variables except the variables of interest. It can be shown that sequentially updating the sufficient statistics of all approximate marginal distributions results in a monotonically decreasing upper bound to the KL divergence between $q(\Theta)$ and the true posterior; for details refer to [5]. Due to limited space, we only show the approximate marginal pdfs which

are different than those in [3]. Using (25) we find

$$q(\mathbf{X}) = \mathcal{N}(\mathrm{vec}(\mathbf{X})|\mathrm{vec}(\langle\mathbf{X}\rangle), \Sigma_{\mathbf{X}}), \quad (26)$$

$$\mathrm{vec}(\langle\mathbf{X}\rangle) = \langle\epsilon_1\rangle\Sigma_{\mathbf{X}}\left(\mathbf{I}_{t_1} \otimes \mathbf{C}^T \mathrm{Diag}(\langle\mathbf{w}\rangle)\right)\mathrm{vec}(\langle\mathbf{S}\rangle),$$

$$\Sigma_{\mathbf{X}} = \left(\langle\epsilon_1\rangle(\mathbf{I}_{t_1} \otimes \mathbf{Q}) + \mathbf{T}_{(t_1,q)}^T(\mathbf{I}_{t_1} \otimes \mathrm{Diag}(\langle\boldsymbol{\beta}_1\rangle))\mathbf{T}_{(t_1,q)}\right)^{-1},$$

and

$$q(\mathbf{Z}) = \mathcal{N}\left(\mathrm{vec}\left(\mathbf{Z}^T\right)|\mathrm{vec}\left(\langle\mathbf{Z}\rangle^T\right), \Sigma_{\mathbf{Z}}\right) \quad (27)$$

$$\mathrm{vec}\left(\langle\mathbf{Z}\rangle^T\right) = \langle\epsilon_2\rangle\Sigma_{\mathbf{Z}}\left(\mathbf{I}_k \otimes \mathbf{C}^T \mathrm{Diag}(\langle\mathbf{w}\rangle)\right)\mathrm{vec}\left(\langle\mathbf{H}\rangle^T\right)$$

$$\Sigma_{\mathbf{Z}} = \left(\langle\epsilon_2\rangle(\mathbf{I}_k \otimes \mathbf{Q}) + \mathbf{T}_{(k,q)}^T(\mathbf{I}_k \otimes \mathrm{Diag}(\langle\boldsymbol{\beta}_2\rangle))\mathbf{T}_{(k,q)}\right)^{-1}$$

where

$$\mathbf{Q} = \mathbf{C}^T\left(\mathrm{Diag}(\mathrm{diag}(\Sigma_{\mathbf{w}})) + \mathrm{Diag}(\langle\mathbf{w}\rangle)^2\right)\mathbf{C}. \quad (28)$$

For the spatial profile $\mathbf{w}$ we obtain

$$q(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\langle\mathbf{w}\rangle, \Sigma_{\mathbf{w}}), \quad (29)$$

$$\langle\mathbf{w}\rangle = \Sigma_{\mathbf{w}}\mathrm{diag}\left(\langle\epsilon_1\rangle\langle\mathbf{S}\rangle\langle\mathbf{X}\rangle^T\mathbf{C}^T + \langle\epsilon_2\rangle\langle\mathbf{H}\rangle^T\langle\mathbf{Z}\rangle\mathbf{C}^T\right),$$

$$\begin{aligned}
\Sigma_{\mathbf{w}}^{-1} = & \langle\epsilon_1\rangle\mathbb{E}\left[\mathrm{Diag}\left(\mathrm{diag}(\mathbf{C}\mathbf{X}\mathbf{X}^T\mathbf{C}^T)\right)\right] + \langle\gamma\rangle\left(\mathbf{N}_1^T W(\mathbf{u})\mathbf{N}_1\right) \\
& + \langle\epsilon_2\rangle\mathbb{E}\left[\mathrm{Diag}\left(\mathrm{diag}(\mathbf{C}\mathbf{Z}^T\mathbf{Z}\mathbf{C}^T)\right)\right] + \langle\gamma\rangle\left(\mathbf{N}_2^T W(\mathbf{u})\mathbf{N}_2\right),
\end{aligned}$$

where $\mathbf{N}_1$ is a $n \times n$ matrix defined as

$$(\mathbf{N}_1)_{ij} = \begin{cases} 1 & \text{if } i = j, \\ -1 & \text{if vertex } j \text{ is the 1st neighbor of vertex } i \\ 0 & \text{otherwise,} \end{cases} \quad (30)$$

$\mathbf{N}_2$ is defined the same way but for the second neighbor, and $W(\mathbf{u})$ is an $n \times n$ diagonal matrix where the i-th element on the main diagonal is given by $1/\sqrt{u_i}$. Note that spatial adaptivity is introduced through the matrix $W(\mathbf{u})$, as it controls the degree of spatial smoothness for each vertex. The pdfs of the temporal smoothness precision hyperparameters are found to be gamma distributions; they are given by

$$q((\boldsymbol{\beta}_1)_i) = \Gamma((\boldsymbol{\beta}_1)_i|(a_{\boldsymbol{\beta}_1})_i, (b_{\boldsymbol{\beta}_1})_i), \quad (a_{\boldsymbol{\beta}_1})_i = \frac{t_1}{2}, \quad (31)$$

$$(b_{\boldsymbol{\beta}_1})_i = \frac{1}{2}\mathbb{E}\left[\mathbf{X}_{i\cdot}\mathbf{T}_{(t_1,1)}^T\mathbf{T}_{(t_1,1)}\mathbf{X}_{i\cdot}^T\right],$$

and

$$q((\boldsymbol{\beta}_2)_i) = \Gamma((\boldsymbol{\beta}_2)_i|(a_{\boldsymbol{\beta}_2})_i, (b_{\boldsymbol{\beta}_2})_i), \quad (a_{\boldsymbol{\beta}_2})_i = \frac{k}{2}, \quad (32)$$

$$(b_{\boldsymbol{\beta}_2})_i = \frac{1}{2}\mathbb{E}\left[\mathbf{Z}_{\cdot i}^T\mathbf{T}_{(k,1)}^T\mathbf{T}_{(k,1)}\mathbf{Z}_{\cdot i}\right].$$

Also $q(\gamma)$ is found to be gamma distributed, with the pdf given by

$$q(\gamma) = \Gamma(\gamma|a_\gamma, b_\gamma), \quad a_\gamma = \varphi n, \quad b_\gamma = \sum_{i=1}^{n}\sqrt{u_i}. \quad (33)$$

The inference algorithm iteratively updates the sufficient statistics of all marginal pdfs and performs a minimization with respect to $\mathbf{u}$ during every iteration. The minimizer can be found in closed form since (22) is convex with respect to $\mathbf{u}$; it is given by

$$u_i = \mathbb{E}\left[\sum_{j \in c_i, i \neq j}(\mathbf{w}_i - \mathbf{w}_j)^2\right]. \quad (34)$$

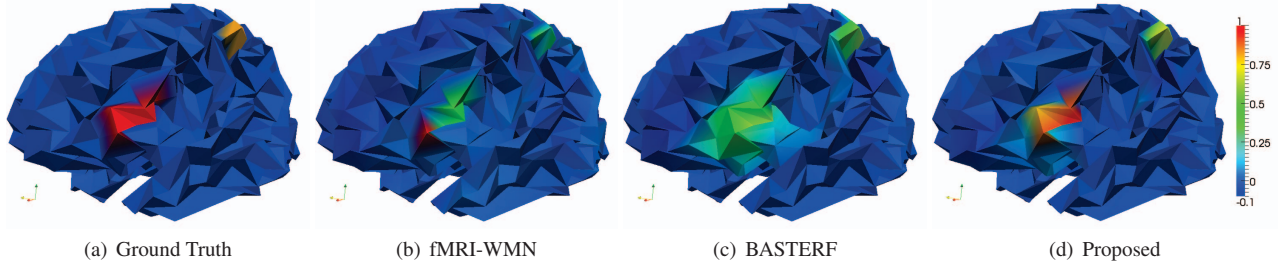The algorithm iterates until a convergence criterion is satisfied.

(a) Ground Truth   (b) fMRI-WMN   (c) BASTERF   (d) Proposed

Fig. 2. Current amplitudes estimated for one simulation of scenario S1 projected onto the cortex at the peak activity of the first EEG source ($t = 20$). It can be seen that the proposed algorithm allows for sharper boundaries and the estimated currents are closer to the ground truth.

4. EXPERIMENTAL RESULTS

We perform a number of simulations using synthetic data in order to evaluate the performance of the proposed method. The simulation methodology is similar to the one used in [3] and consists of generating synthetic data by placing sources at random locations on the cortical surface. For each simulation, two sources are generated; each source can be electrically active, hemodynamically active, or both, depending on the simulation scenario. The different simulation scenarios allow us to assess the robustness of our algorithm in situations where the neurological activity of a source is only detectable via one imaging modality. The first scenario (S1) represents the ideal case where both sources are hemodynamically and electrically active. Scenarios S2 and S3 are mixed cases where one source is either hemodynamically or electrically inactive, respectively. In the last scenario (S4) one source is electrically active and the other source is hemodynamically active; the scenario can therefore be considered the worst case scenario since the EEG source localization does not benefit from the information provided by fMRI. Mean squared error (MSE) scores averaged over 25 simulations per scenario are shown in Table 1. As can be seen the proposed method outperforms existing state-of-the-art algorithms, i.e., the fMRI weighted minimum norm algorithm (fMRI-WMN) [2] as well as the "BASTERF" algorithm [3]. The performance of the proposed method is also illustrated in Fig. 2, where the estimated currents are projected onto the cortex. The spatial adaptivity of the proposed algorithm increases the estimation accuracy and allows for sharper boundaries between active and inactive regions. It is clear that the proposed method provides the most accurate source localization result among the evaluated methods.

	fMRI-WMN	BASTERF	Proposed
S1	1.666×10^{-3}	1.578×10^{-3}	0.723×10^{-3}
S2	2.383×10^{-3}	2.196×10^{-3}	1.808×10^{-3}
S3	0.881×10^{-3}	0.813×10^{-3}	0.311×10^{-3}
S4	1.520×10^{-3}	1.209×10^{-3}	1.198×10^{-3}

Table 1. Table 1: Average MSE **S**

The parameters used for the simulations are as follows: $m = 64$, $n = 1028$, $t_1 = 50$, $t_2 = 1000$, $k = 30$, $q = 32$, $\varphi = 1.0$, Peak-SNR EEG $= 500$, Peak-SNR fMRI $= 1.0$. The parcellation was obtained by aligning the cortical mesh with an anatomical atlas [6]. The sources consist of compact clusters of 8 dipoles. For EEG half sine curves with a length of 20 time points are used as source waveforms, while a canonical HRF shape is used for fMRI.

5. CONCLUSIONS

We propose a symmetrical EEG/fMRI fusion algorithm which utilizes spatially adaptive priors. The algorithm combines the modalities by means of a common generative model, representing a formal way of including fMRI information into the ill-posed EEG source localization problem. Spatial adaptivity is achieved by means of a TV prior, which is used for the spatial profile, as well as spatially adaptive temporal priors, which enable the algorithm to adapt the degree of temporal smoothness according to the responses in different brain regions. Since calculating the posterior is intractable and the form of the TV prior prevents calculating the mean field approximation to the posterior analytically, we resort to an approximate variational Bayesian method [5] to draw inference. Experimental results show that our algorithm outperforms existing EEG/fMRI fusion methods, in terms of MSE, as well as, preservation of sharp edges in the solution.

6. REFERENCES

[1] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Mod. Phys.*, vol. 65, no. 2, pp. 413–497, Apr 1993.

[2] A. K. Liu, J. W. Belliveau, and A. M. Dale, "Spatiotemporal imaging of human brain activity using functional mri constrained magnetoencephalography data: Monte carlo simulations," *Proc. Natl. Acad. Sci. U.S.A*, vol. 95, no. 15, pp. 8945–8950, July 1998.

[3] J. Daunizeau, C. Grova, G. Marrelec, J. Mattout, S. Jbabdi, M. Pélégrini-Issac, J. M. Lina, and H. Benali, "Symmetrical event-related eeg/fmri information fusion in a variational bayesian framework," *NeuroImage*, vol. 36, no. 1, pp. 69–87, May 2007.

[4] G. Marrelec, H. Benali, P. Ciuciu, and J. B. Poline, "Bayesian estimation of the hemodynamic response function in functional mri," *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, vol. 617, no. 1, pp. 229–247, 2002.

[5] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Parameter estimation in tv image restoration using variational distribution approximation," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 326–339, 2008.

[6] S. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, H. Johansenberg, P. Bannister, M. DeLuca, I. Drobnjak, and D. Flitney, "Advances in functional and structural mr image analysis and implementation as fsl," *NeuroImage*, vol. 23, pp. S208–S219, 2004.