

Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework

Jean Daunizeau,^{a,b,c,d,k,*} Christophe Grova,^e Guillaume Marrelec,^{b,c,f,k} Jérémie Mattout,^{a,g,h,k}
Saad Jbabdi,^{b,i,k} Mélanie Pélégrini-Issac,^{b,c,k} Jean-Marc Lina,^{b,d,f,j,k} and Habib Benali^{b,c,f,k}

^aWellcome Department of Imaging Neuroscience, London, UK

^bINSERM U678, Paris F-75013, France

^cUniversité Pierre et Marie Curie, Faculté de Médecine Pitié-Salpêtrière, Paris F-75013, France

^dCentre de Recherches Mathématiques, Montréal, Québec, Canada

^eMontreal Neurological Institute, Montréal, Québec, Canada

^fUniversité de Montréal, MIC/UNF, Montreal, Canada H3W 1W5

^gCEA/SHFJ, Orsay, France

^hINSERM U821, Dynamique Cérébrale et Cognition, 69000, Lyon, France

ⁱFMRIB Lab., Oxford, UK

^jEcole de Technologie Supérieure, Montréal, Québec, Canada

^kIFR49, Paris, France

Received 21 April 2006; revised 12 December 2006; accepted 3 January 2007
Available online 15 February 2007

In this work, we propose a symmetrical multimodal EEG/fMRI information fusion approach dedicated to the identification of event-related bioelectric and hemodynamic responses. Unlike existing, asymmetrical EEG/fMRI data fusion algorithms, we build a joint EEG/fMRI generative model that explicitly accounts for local coupling/uncoupling of bioelectric and hemodynamic activities, which are supposed to share a common substrate. Under a dedicated assumption of spatio-temporal separability, the spatial profile of the common EEG/fMRI sources is introduced as an unknown hierarchical prior on both markers of cerebral activity. Thereby, a devoted Variational Bayesian (VB) learning scheme is derived to infer common EEG/fMRI sources from a joint EEG/fMRI dataset. This yields an estimate of the common spatial profile, which is built as a trade-off between information extracted from EEG and fMRI datasets. Furthermore, the spatial structure of the EEG/fMRI coupling/uncoupling is learned exclusively from the data. The proposed data generative model and devoted VBEM learning scheme thus provide an un-supervised well-balanced approach for the fusion of EEG/fMRI information. We first demonstrate our approach on synthetic data. Results show that, in contrast to classical EEG/fMRI fusion approach, the method proved efficient and robust regardless of the EEG/fMRI discordance level. We apply the method on EEG/fMRI recordings from a patient with epilepsy, in order to identify brain areas involved during the generation of epileptic spikes. The results are validated using intracranial EEG measurements.

© 2007 Elsevier Inc. All rights reserved.

Introduction

Because of the complementary temporal and spatial resolutions of electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), combining measurements originating from both modalities may reveal fine spatio-temporal structures of neuronal activity that would otherwise remain undetected if the analyses were conducted using data from only one modality. This fusion of information is essential to understand the physiological processes mediating the treatment of a cognitive task or spontaneous brain activity.

The main cause of EEG measurements is likely to be the post-synaptic cortical currents associated to the large pyramidal neurons, which are oriented perpendicular to the cortical surface (Nunez, 1981). Even though fMRI is believed to reveal some complementary features of neuronal activity (Nunez and Silberstein, 2000; Mukamel et al., 2005), it is only an indirect measure thereof, through metabolism, oxygenation and blood flow. Despite the increasing amount of literature in the field of neuro-vascular coupling (see, for a recent example (Riera et al., 2006)), none of the existing biophysical models specifies precisely what is meant by the “neural activity” that drives the hemodynamic response. Therefore, these models cannot tell us what aspect of neural information processing is reflected by the BOLD signal. As a matter of fact, neural information processing within a given cortical unit can be described along many different dimensions, and its relationship with existing neurophysiological processes can be characterized on different scales, for example, local field potentials versus spiking activity, excitatory versus inhibitory postsynaptic potentials or different types of receptor at synapses (Stefan et al., 2004).

* Corresponding author. The Wellcome Department of Imaging Neuroscience, Institute of Neurology, UCL, 12 Queen Square, London WC1N 3BG, UK. Fax: +44 207 813 1445.

E-mail address: j.daunizeau@fil.ion.ucl.ac.uk (J. Daunizeau).

Available online on ScienceDirect (www.sciencedirect.com).

Sophisticated animal studies that combine multielectrode recordings with fMRI (Puce et al., 1997; Logothetis et al., 2001; Logothetis and Wandell, 2004) or with optical imaging techniques (Mathiesen et al., 1998; Martindale et al., 2003) have started to address these issues. However, the quantitative contribution of each of the neurophysiological processes that occur within the “active” areas to both EEG and fMRI measurements is nowadays mostly unknown. As a consequence, the state-of-the-art (invasive) experimental evidences are probably not understood so as to ground a robust neurophysiologically informed model of the neuro-vascular coupling (Gonzales et al., 2001; Stefan et al., 2004; Daunizeau et al., 2005a).

Nevertheless, we might define “neuronal activity” *operationally*: it is the set of dynamic processes that characterizes the nodes belonging to the brain network *specifically* involved in the treatment of a given event (e.g. a cognitive/sensory/motor task or spontaneous brain activity) (Friston, 2005b). This allows us to consider *event-related* (ER) EEG and fMRI datasets as different *measures* of this “neuronal activity”, since the ER response is indeed defined as the reproducible EEG or fMRI signature that is systematically consecutive to the appearance of a single event (Friston, 2005a). In that perspective, the bioelectric and metabolic activities (understood as ER responses), as detected by EEG and fMRI, do not necessarily match.

The “neuronal activity” ζ can then be decomposed into two non-orthogonal sub-spaces ζ_{EEG} and ζ_{fMRI} that correspond to the part of ζ that contributes to EEG and fMRI datasets, respectively (Pieger and Greenblatt, 2001). The intersection ζ_1 of ζ_{EEG} and ζ_{fMRI} (see Fig. 1) defines the “common substrate” of neuronal activity. Conversely, ζ_2 (respectively ζ_3) denotes the subspace of neuronal activity detected by EEG (respectively fMRI) that does not contribute to fMRI (respectively EEG) measurements. This decomposition formalizes the apparent coupling/uncoupling process occurring between bioelectric and hemodynamic ER responses.

Then, the following question arises: what should we expect to learn about neuronal activity by combining EEG and fMRI information? Since no information about ζ_2 (resp. ζ_3) is available from the fMRI (resp. EEG) data set, no multimodal procedure will provide a finer characterization of this activity subspace than a unimodal EEG (resp. fMRI) data analysis. By contrast, it seems reasonable to expect that such a multimodal approach should benefit from the complementarity of EEG and fMRI information about the subspace ζ_1 to refine our knowledge about the commonalities of the bioelectric and hemodynamic ER responses.

Most current EEG/fMRI fusion strategies mainly rely on the introduction of constraints derived from a preliminary analysis of

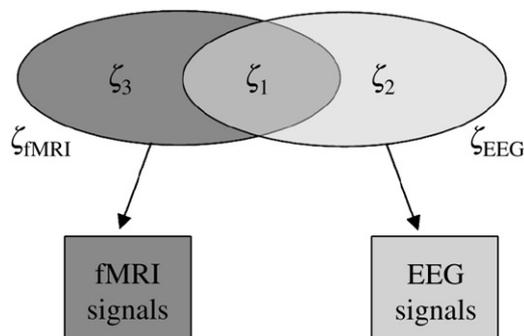


Fig. 1. Formalization of the coupling/uncoupling between brain activities observed using multimodal ER EEG/fMRI experiments (adapted from Pieger and Greenblatt, 2001).

fMRI into the EEG source reconstruction problem (Liu et al., 1998; Babiloni et al., 2003; Ahlfors and Simpson, 2004). Stated in Pieger and Greenblatt’s formulation (Pieger and Greenblatt, 2001), these approaches try to explain EEG data using an estimate $\hat{\zeta}_{\text{fMRI}}$ of ζ_{fMRI} as a prior information on ζ_{EEG} . They are *asymmetrical*, since they do not consider EEG and fMRI data sets as equivalent and do not analyze them jointly (Trujillo et al., 2001). Importantly, the estimation bias arising from a potential discrepancy between ζ_{EEG} and ζ_{fMRI} will mainly depend on the exibility in the way the fMRI priors are introduced (Mattout et al., 2006).

On the opposite, the objectives of a *symmetrical* well-balanced approach for the fusion of multimodal EEG/fMRI information are twofold. First, the approach should be able to identify the part of EEG and fMRI signals that conveys complementary information about the common substrate of the underlying brain activity. Second, it should exploit such information, as extracted from joint data sets, in order to decrease the uncertainty relative to our knowledge of this common subspace of ER responses.

As a consequence, a symmetrical fusion approach would require the explicit definition of the common neuronal substrate that elicits both EEG and fMRI measurements. This entails building a model that encompasses our knowledge about the link between bioelectric and hemodynamic activities. Here, we propose to define this neuro-vascular coupling through the common properties exhibited by the “active” areas specifically contributing to both event-related EEG and fMRI measurements. Due to our lack of precise knowledge regarding the coupling between the temporal properties of bioelectric and hemodynamic ER responses, we restricted these common properties to the spatial profile (i.e. the position and the extent) of the common EEG and fMRI sources. In other words, the only common substrate (or subspace) of bioelectric and hemodynamic “activities” (or ER responses) is defined as the spatial support of the EEG/fMRI common signal generators.

Noticeably, the dual estimation of the bioelectric and hemodynamic ER responses does not circumvent the potential difficulties of the inverse problems related to each modality. On the one hand, reconstructing cortical sources current density from EEG measurements is a well known intrinsically spatial ill-posed inverse problem (Baillet et al., 2001). This has a temporal drawback: we do not know what the temporal dynamics of the true bioelectric sources are. On the other hand, estimating the hemodynamic response using deconvolution approaches is a difficult temporal inverse problem due to potential overlapping of fMRI responses to consecutive events (Marrelec et al., 2003a). Similarly, this implies uncertainty in the localization of the hemodynamic activity.

The hierarchical EEG/fMRI generative model that we propose here is a generalization of the extended sources mixing model (Daunizeau et al., 2005b), whose key aspect lies in the separation between spatial and temporal characteristics of brain activity. The coupling model is restricted to the spatial profile of ER bioelectric and hemodynamic responses. The model does not impose any constraint regarding an eventual coupling between temporal dynamics of the physiological processes underlying EEG and fMRI measurements. Moreover, the hierarchical model explicitly accounts for potential discrepancies between the bioelectric and hemodynamic activation sites. Indeed, rather than formally constraining the analysis, the objective of the approach is to find the common spatial substrate that may explain some features of the joint data sets; the remaining characteristics (e.g., in the temporal domain) of the activity that are required to explain some specific features of the respective datasets are estimated without any EEG/fMRI coupling constraint.

Despite its somewhat heuristic aspect, this joint EEG/fMRI information fusion approach is both robust to the lack of information about the neuro-vascular coupling and flexible enough to allow for later incorporation of further knowledge about the generation of bioelectric and hemodynamic ER responses.

This paper comprises various sections. In Model of spatially concordant ER responses, we show how to rely on the spatio-temporal decomposition of cerebral activity to formalize the coupling/decoupling of bioelectric and hemodynamic event-related responses (Separation of space and time). We also expose the statistical assumptions required to specify the associated graphical generative model (Specification of the hierarchical model). In Learning the model: variational Bayesian learning scheme, we describe the Variational Bayesian (VB) algorithm that we developed in order to make statistical inference on the model. Evaluation presents elements of evaluation of the whole approach: (1) assessment of the method’s performance using simulated data and (2) illustration of the approach in the context of interictal spike localization from EEG/fMRI data simultaneously acquired on a patient with focal epilepsy. In Discussion, we discuss the main aspects of the model and associated VB inference scheme and replace the whole approach in the context of the general problem of information fusion.

Notations

In the following, \mathbf{X}^T , \mathbf{X}_i , \mathbf{X}_{ij} and $\text{tr}(\mathbf{X})$ indicate the transpose of matrix \mathbf{X} , the i th vector column of \mathbf{X} , the scalar element of the i th column and j th row of \mathbf{X} and the trace of \mathbf{X} , respectively. $(x_i)_{1 \leq i \leq n}$ denotes the $n \times 1$ vector whose entries are x_i . \mathbf{I}_n and $\mathbf{0}_n$ stand for the $n \times n$ identity matrix and the $n \times 1$ null vector, respectively. For any $n \times 1$ vector \mathbf{x} , $\text{Diag}(\mathbf{x})$ denotes the $n \times n$ diagonal matrix whose diagonal is \mathbf{x} . By contrast, $\text{diag}(\mathbf{X})$ denotes the $n \times 1$ vector containing the diagonal entries of the $n \times n$ matrix \mathbf{X} . \otimes denotes the Kronecker product, Δ denotes the Laplacian operator and “ \propto ” relates two expressions that are proportional. The cardinal of any set \mathcal{V} is written $\text{card}[\mathcal{V}]$. For two variables x and y , $x|y$ stands for “ x given y ”, $p(x)$ for the probability of x , $\langle x \rangle$ for its expectation and \hat{x} for its estimate. $\mathcal{N}(\mathbf{m}, \mathbf{V})$ is the Gaussian probability density function (pdf) with mean \mathbf{m} and covariance matrix \mathbf{V} and $\mathcal{G}(a, b)$ is the Gamma pdf with a degrees of freedom (d.o.f.) and shape parameter b . Given an $a \times b$ matrix \mathbf{X} , let us also define an $ab \times 1$ vector-valued function denoted by $\text{vec}(\mathbf{X})$ such that

$$\text{vec}(\mathbf{X}) = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_b \end{pmatrix}.$$

Model of spatially concordant ER responses

In the following, we assume that a single type of event is involved in the experiment. We then provide a generative model that can account for both fMRI and EEG event-related data in a hierarchical fashion.

The EEG and fMRI forward problems

On the one hand, solving the EEG inverse problem within the so-called distributed framework amounts to finding a unique solution to the following linear system (Dale and Sereno, 1993):

$$\mathbf{M} = \mathbf{G}\mathbf{J} + \mathbf{E}, \quad (1)$$

where \mathbf{M} stands for the $p \times t_1$ matrix of scalp (ER potential) EEG data set ($p \sim 10^2$: number of sensors, $t_1 \sim 10^2$: number of time samples), \mathbf{E} is an additive measurement noise, \mathbf{J} is the $n \times t_1$ matrix of the unknown time courses of the dipoles ($n \sim 10^4$: number of dipoles distributed on the cortical surface) and \mathbf{G} is the $p \times n$ gain matrix (forward operator) associated with the position and orientation of the dipoles.

\mathbf{J} is the voxelwise description of the bioelectric ER response. \mathbf{G} is obtained by solving the so-called forward problem (de Munck, 1988) for a given set of dipoles with fixed position and orientation (distributed perpendicular to the cortical surface). Each column \mathbf{G}_j of \mathbf{G} indicates the putative contribution of dipole j to the scalp data (its so-called forward field).

On the other hand, the general linear model (GLM) links the stimulation paradigm to the fMRI measurements through the hemodynamic response function (HRF) (Marrelec et al., 2001)¹:

$$\mathbf{Y} = \mathbf{B}\mathbf{h} + \mathbf{F}, \quad (2)$$

where \mathbf{Y} is the $t_2 \times n$ matrix of voxelwise fMRI measurements ($t_2 \sim 10^2$: number of time samples, $n \sim 10^4$: number of voxels), \mathbf{F} is an additive measurement noise, \mathbf{h} is the $k \times n$ matrix of the unknown HRF in each voxel (k : order of the convolution model), and \mathbf{B} is the $t_2 \times k$ design matrix, consisting of the lagged stimulus onset covariates, i.e.:

$$\mathbf{B} = \begin{pmatrix} x_k & \dots & x_1 \\ \vdots & & \\ x_{t_2+k-1} & \dots & x_{t_2} \end{pmatrix}, \quad (3)$$

given that $(x_i)_{1 \leq i < t_2+k}$ is the event time course.

Separation of space and time

We further assume that cerebral activity is structured by a set of active areas (brain regions) that are characterized by their temporal coherence. Therefore, let us consider a given parcelling of the cortical surface into q anatomically and functionally homogeneous clusters (see Appendix B). The so-called extended sources mixing model (see previous work (Daunizeau et al., 2005a)) then associates each parcel $P_i (i=1, \dots, q)$ with its single temporal dynamics using the following hierarchical model of bioelectric sources \mathbf{J} :

$$\mathbf{J} = \text{Diag}(\mathbf{w}^{\text{EEG}})\mathbf{C}\mathbf{X} + \mathbf{R}, \quad (4)$$

where \mathbf{X} is an unknown $q \times t_1$ matrix made of the q time courses of the q parcels, \mathbf{C} is the known $n \times q$ matrix describing the cortex parcelling ($C_{ji}=1$ if $j \in P_i$, and $C_{ji}=0$ otherwise.), \mathbf{w}^{EEG} is a $n \times 1$ unknown vector, and \mathbf{R} is a residual bioelectric activity that cannot be explained using the extended sources mixing model. In this formulation, the temporal dynamics of the i th dipole is defined as the time course of the parcel to which it belongs, weighted by a scalar w_i^{EEG} . The vector \mathbf{w}^{EEG} expresses the relative within-region distribution of current intensity. It describes the (time-invariant) spatial profile of each active extended cortical source. On the opposite, \mathbf{X} embodies the temporal features of the bioelectric sources

¹ This formulation does not refer to the standard GLM, as proposed in (Friston et al., 1995), where the design matrix is defined by regressors that model the different types of stimuli and are constructed by convolution of the stimulus with a canonical hemodynamic response function, while regression coefficients represent effects sizes. Rather, this forward model is a discretization of the unknown ER hemodynamic impulse response function.

\mathbf{J} . Eq. (4) thus enforces a spatio-temporal separability of the bioelectric activity by means of the parcelling \mathbf{C} .

In a similar fashion, let us write the extended sources mixing model for the hemodynamic ER response:

$$\mathbf{h} = \mathbf{Z}\mathbf{C}^T \text{Diag}(\mathbf{w}^{\text{fMRI}}) + \mathbf{L}, \quad (5)$$

where \mathbf{Z} is an unknown $k \times q$ matrix containing the HRF temporal shape of the q parcels, \mathbf{w}^{fMRI} is an also unknown $n \times 1$ vector associated with the spatial profile of the hemodynamic activity sources, and \mathbf{L} is the residual discrepancy from the extended sources mixing model.

Furthermore, we assume that the common feature of bioelectric and hemodynamic ER responses is the spatial profile of the sources seen by both neuroimaging modalities. This assumption may be formalized using the following equation:

$$\mathbf{w}^{\text{EEG}} = \mathbf{w}^{\text{fMRI}} = \mathbf{w}. \quad (6)$$

Hence, \mathbf{X} (resp. \mathbf{Z}) represents the bioelectric (resp. hemodynamic) temporal dynamics of activity sources common to EEG and fMRI, whereas residuals \mathbf{R} and \mathbf{L} are associated to the activity sources specific of EEG and fMRI, respectively. Since the physical units of bioelectric and hemodynamic activities are carried by the time course variables \mathbf{X} and \mathbf{Z} , the common substrate of cerebral activity \mathbf{w} is a dimensionless quantity. Fig. 2 illustrates this “spatially concordant” ER responses model.

Specification of the hierarchical model

In a Bayesian perspective, any uncertainty associated with an unknown quantity is to be modeled through a probability density function (pdf). In the previous section, we described the hierarchical

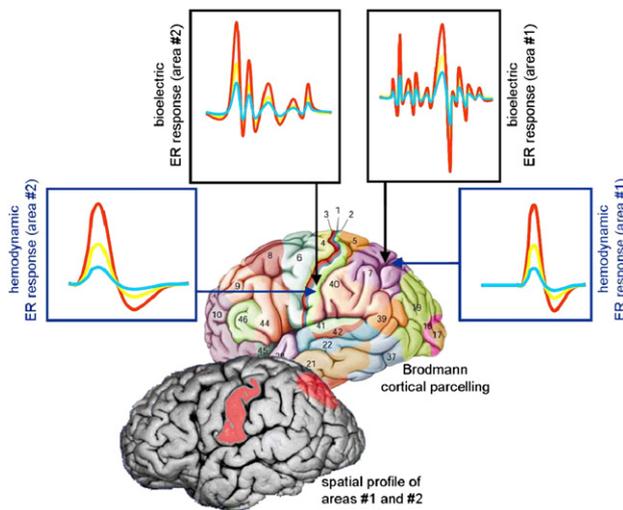


Fig. 2. Schematic illustration of spatio-temporal decomposition entailed by the spatially concordant ER responses model. Two active areas are depicted, with their (time invariant) spatial profile. These exhibit coherent bioelectric and hemodynamic ER responses (the colored time courses are associated with different voxels belonging to the same parcel), which are supposed to contribute to EEG and fMRI measurements, respectively.

observation model. Now, we focus on additional assumptions about the conditional dependencies between the different model parameters.

From now on, we will refer to \mathcal{M} as the graphical model depicted in Fig. 3. This model encompasses all assumptions associated with the hierarchical generative model. The Markov properties of the graphical model \mathcal{M} enables us to write the joint posterior pdf of the model parameters as the following product of conditional pdf:

$$\begin{aligned} p(\mathbf{J}, \mathbf{h}, \mathbf{w}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta} | \mathbf{M}, \mathbf{Y}, \mathcal{M}) &\propto p(\mathbf{M} | \mathbf{J}, \alpha_1, \mathcal{M}) p(\mathbf{Y} | \mathbf{h}, \alpha_2, \mathcal{M}) \\ &\quad p(\mathbf{J} | \mathbf{w}, \mathbf{X}, \epsilon_1, \mathcal{M}) p(\mathbf{h} | \mathbf{w}, \mathbf{Z}, \epsilon_2, \mathcal{M}) \\ &\quad p(\mathbf{w} | \gamma, \mathcal{M}) p(\mathbf{X} | \beta_1, \mathcal{M}) \\ &\quad p(\mathbf{Z} | \beta_2, \mathcal{M}) p(\boldsymbol{\theta} | \mathcal{M}), \end{aligned}$$

where $\boldsymbol{\theta}$ is the set of precision hyperparameters of the graphical model \mathcal{M} :

$$\boldsymbol{\theta} = (\alpha_1, \alpha_2, \epsilon_1, \epsilon_2, \beta_1, \beta_2, \gamma).$$

Each conditional pdf corresponds to prior (and independent) assumptions regarding the probabilistic link between the model parameters (see (Marrelec et al., 2003b, 2004) for details about graphical models in fMRI). The following subsections are devoted to the derivation of this conditional pdf.

Likelihoods

The data likelihoods result from statistical assumptions about the measurement noise. We consider \mathbf{E} and \mathbf{F} (cf Eq. (1) and (2)) as realizations of independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and (unknown) precisions α_1 and α_2 , respectively. Then, EEG and fMRI data likelihoods yield:

$$\begin{cases} p(\mathbf{M} | \mathbf{J}, \alpha_1, \mathcal{M}) = \prod_{i=1}^{I_1} \mathcal{N}\left(\mathbf{G}\mathbf{J}_i, \frac{1}{\alpha_1} \mathbf{I}_p\right) \\ p(\mathbf{Y} | \mathbf{h}, \alpha_2, \mathcal{M}) = \prod_{j=1}^n \mathcal{N}\left(\mathbf{B}\mathbf{h}_j, \frac{1}{\alpha_2} \mathbf{I}_{I_2}\right). \end{cases} \quad (7)$$

Prior densities on bioelectric and hemodynamic activities

Specifying such prior densities amounts to defining the pdf of \mathbf{J} and \mathbf{h} conditional on the parameters associated with their respective spatio-temporal decomposition.

We assume that the residuals \mathbf{R} and \mathbf{L} are the realizations of i.i.d. Gaussian random variables with zero mean and (unknown) precisions ϵ_1 and ϵ_2 . This assumption is strictly equivalent to the definition of the following conditional pdf:

$$\begin{cases} p(\mathbf{J} | \mathbf{w}, \mathbf{X}, \epsilon_1, \mathcal{M}) = \prod_{i=1}^{I_1} \mathcal{N}\left(\text{Diag}(\mathbf{w})\mathbf{C}\mathbf{X}_i, \frac{1}{\epsilon_1} \mathbf{I}_n\right) \\ p(\mathbf{h}^* | \mathbf{w}, \mathbf{Z}^*, \epsilon_2, \mathcal{M}) = \prod_{j=1}^k \mathcal{N}\left(\text{Diag}(\mathbf{w})\mathbf{C}\mathbf{Z}_j^*, \frac{1}{\epsilon_2} \mathbf{I}_n\right), \end{cases} \quad (8)$$

where $\mathbf{h}^* \equiv \mathbf{h}^T$ and $\mathbf{Z}^* \equiv \mathbf{Z}^T$.

Prior densities of the coupling model parameters

To set the prior pdf on variables \mathbf{X} , \mathbf{Z}^* and \mathbf{w} , we make specific assumptions regarding the expected behavior of brain activity:

- a neuronal macrocolumn located next to an active macrocolumn is expected to be also partially active (Nunez, 1981). As a consequence, the probability of activation of each macrocolumn depends on its neighborhood. We *a priori* expect that the Laplacian of \mathbf{w} is zero:

$$\langle \Delta w \rangle = 0. \quad (9)$$

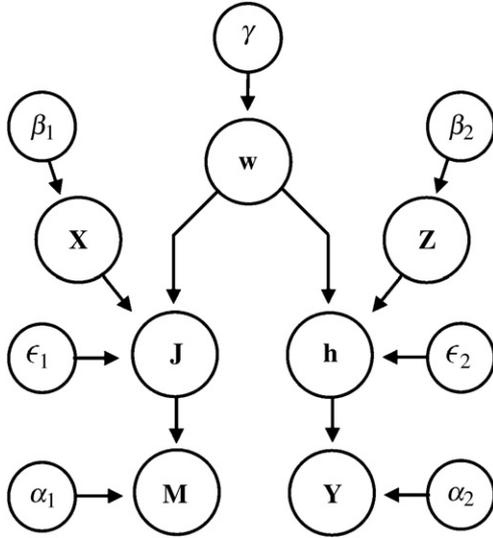


Fig. 3. Graph representing the hierarchical relations between the EEG/fMRI data generative model parameters. \mathbf{w} is the common spatial profile of the sources, \mathbf{X} (respectively \mathbf{Z}) is the temporal dynamics of the bioelectric (respectively hemodynamic) activity. \mathbf{J} is the time course of the bioelectric activity. \mathbf{h} is the hemodynamic activity. \mathbf{M} (resp. \mathbf{Y}) contains the EEG (resp. fMRI) measurements. The other nodes are the precision parameters that are associated with prior assumptions about the expected structure of brain activity, and noise measurements.

Under Gaussian assumption, this is strictly equivalent to the following prior pdf for the spatial field \mathbf{w} (Gössl et al., 2001):

$$p(\mathbf{w}|\gamma, \mathcal{M}) = \mathcal{N}\left(\mathbf{0}_n, \frac{1}{\gamma}(\mathbf{S}^T\mathbf{S})^{-1}\right), \quad (10)$$

where γ is the (unknown) precision of the Laplacian field of \mathbf{w} , and \mathbf{S} is the $n \times n$ discrete Laplacian operator defined as:

$$\begin{cases} S_{jj} = 1 \\ S_{jj'} = -\frac{1}{\text{card}[\mathcal{V}_j]} \text{ if } j' \in \mathcal{V}_j \\ S_{jj'} = 0 \text{ otherwise,} \end{cases} \quad (11)$$

where \mathcal{V}_j is the neighborhood of voxel j .

• EEG and fMRI are known to provide oversampled measures of bioelectric and hemodynamic activities, respectively (Nunez, 1981; Heeger, 2002). This implies some temporal smoothness in the time courses of bioelectric and hemodynamic activities of the cortical parcels. Here, we *a priori* expect that the second temporal derivatives of \mathbf{X} and \mathbf{Z} are zero (Marrelec et al., 2001, 2003b, 2004):

$$\left\langle \frac{\partial^2 \mathbf{X}}{\partial t^2} \right\rangle = \left\langle \frac{\partial^2 \mathbf{Z}}{\partial t^2} \right\rangle = 0. \quad (12)$$

Under Gaussian assumption, this is equivalent to the following prior pdf for \mathbf{X} and \mathbf{Z} (Marrelec et al., 2003a):

$$\begin{cases} p(\text{vec}(\mathbf{X})|\beta_1, \mathcal{M}) = \mathcal{N}\left(\mathbf{0}_{qt_1}, \frac{1}{\beta_1}(\mathbf{T}_1^T\mathbf{T}_1)^{-1}\right) \\ p(\text{vec}(\mathbf{Z}^*)|\beta_2, \mathcal{M}) = \mathcal{N}\left(\mathbf{0}_{kq_1}, \frac{1}{\beta_2}(\mathbf{T}_2^T\mathbf{T}_2)^{-1}\right), \end{cases} \quad (13)$$

where β_1 and β_2 are the (unknown) precisions of the second temporal derivatives of \mathbf{X} and \mathbf{Z} , and \mathbf{T}_1 (resp. \mathbf{T}_2) is a $qt_1 \times qt_1$ (resp. $kq \times kq$) matrix such that:

$$\begin{cases} T_{ii} = -2 \\ T_{ij} = 1 \text{ if } j = i \pm (q + 1) \\ T_{ij} = 0 \text{ otherwise.} \end{cases} \quad (14)$$

Prior densities on scaling hyperparameters

All precision hyperparameters are unknown quantities. Therefore, in a full Bayesian approach, we need to specify prior densities on each of them. These were chosen such that our graphical model \mathcal{M} belongs to conjugate exponential models, which are easier to manipulate (Gelman et al., 1998).

As for the variances of the EEG/fMRI measurement noise, let us assume that we have at our disposal two data windows (\mathbf{M}_0 , of size $p \times t_3$ in EEG, and \mathbf{Y}_0 of size $n \times t_4$ in fMRI) containing only i.i.d. noise. In practice, one can use scalp recordings preceding bioelectric responses in EEG, and fMRI data corresponding to regions that do not belong to grey matter. A direct consequence of the assumption of Gaussian noise is that α_1 given \mathbf{M}_0 (resp. α_2 given \mathbf{Y}_0) behaves as a Gamma variate. The definition of their respective prior pdf thus pertains to the derivation of the conditional pdf $p(\alpha_1|\mathbf{M}_0)$ and $p(\alpha_2|\mathbf{Y}_0)$:

$$\begin{aligned} p(\alpha_1|\mathcal{M}) &= p(\alpha_1|\mathbf{M}_0) = \mathcal{G}(a_1, b_1) \\ p(\alpha_2|\mathcal{M}) &= p(\alpha_2|\mathbf{Y}_0) = \mathcal{G}(a_2, b_2), \end{aligned} \quad (15)$$

where parameters (a_1, b_1) and (a_2, b_2) are such that:

$$\begin{cases} a_1 = \frac{pt_3}{2}, & b_1 = \frac{\text{tr}(\mathbf{M}_0^T\mathbf{M}_0)}{2} \\ a_2 = \frac{nt_4}{2}, & b_2 = \frac{\text{tr}(\mathbf{Y}_0^T\mathbf{Y}_0)}{2}. \end{cases} \quad (16)$$

Finally, since we have no prior information about the remaining precision hyperparameters, we consider noninformative Jeffreys' priors (uniform pdf over the log-hyperparameter) (Kass and Wassermann, 1996):

$$\begin{cases} p(\epsilon_1|\mathcal{M}) \propto (\epsilon_1)^{-1} \\ p(\epsilon_2|\mathcal{M}) \propto (\epsilon_2)^{-1} \\ p(\beta_1|\mathcal{M}) \propto (\beta_1)^{-1} \\ p(\beta_2|\mathcal{M}) \propto (\beta_2)^{-1} \\ p(\gamma|\mathcal{M}) \propto (\gamma)^{-1}. \end{cases} \quad (17)$$

All assumptions listed in Specification of the hierarchical model and associated with the spatially concordant event-related response model form the graphical (hierarchical) model that is summarized by the graph represented in Fig. 3.

Learning the model: variational Bayesian learning scheme

There are two main goals in Bayesian learning. The first one is to provide the posterior distribution over the model parameters; the second one, to provide a quantitative feedback on the relevance of

Table 1

Functional forms of the variational pdf of the localization parameters of the graphical model \mathcal{M} , which are Gaussian densities

Functional form	Mean and covariance matrix
$q_J(\text{vec}(\mathbf{J})) = \mathcal{N}(\text{vec}(\boldsymbol{\mu}_J), \mathbf{I}_l \otimes \boldsymbol{\Sigma}_J)$	$\boldsymbol{\mu}_J = \boldsymbol{\Sigma}_J \left(\frac{a_E}{b_E} \mathbf{G}^T \mathbf{M} + \frac{a_R}{b_R} \text{Diag}(\boldsymbol{\mu}_W) \mathbf{C} \boldsymbol{\mu}_X \right)$ $\boldsymbol{\Sigma}_J = \left(\frac{a_E}{b_E} \mathbf{G}^T \mathbf{G} + \frac{a_R}{b_R} \mathbf{I}_n \right)^{-1}$
$q_h(\text{vec}(\mathbf{h})) = \mathcal{N}(\text{vec}(\boldsymbol{\mu}_h), \mathbf{I}_n \otimes \boldsymbol{\Sigma}_h)$	$\boldsymbol{\mu}_h = \boldsymbol{\Sigma}_h \left(\frac{a_E}{b_E} \mathbf{B}^T \mathbf{Y} + \frac{a_L}{b_L} \boldsymbol{\mu}_Z \mathbf{C}^T \text{Diag}(\boldsymbol{\mu}_W) \right)$ $\boldsymbol{\Sigma}_h = \left(\frac{a_E}{b_E} \mathbf{B}^T \mathbf{B} + \frac{a_L}{b_L} \mathbf{I}_2 \right)^{-1}$
$q_X(\text{vec}(\mathbf{X})) = \mathcal{N}(\text{vec}(\boldsymbol{\mu}_X), \boldsymbol{\Sigma}_X)$	$\text{vec}(\boldsymbol{\mu}_X) = \frac{a_R}{b_R} \boldsymbol{\Sigma}_X (\mathbf{I}_l \otimes \text{Diag}(\boldsymbol{\mu}_W) \mathbf{C}) \text{vec}(\boldsymbol{\mu}_J)$ $\boldsymbol{\Sigma}_X = \left(\frac{a_R}{b_R} \mathbf{I}_l \otimes \mathbf{Q} + \frac{a_X}{b_X} \mathbf{T}_1^T \mathbf{T}_1 \right)^{-1}$
$q_Z(\text{vec}(\mathbf{Z}^*)) = \mathcal{N}(\text{vec}(\boldsymbol{\mu}_{Z^*}), \boldsymbol{\Sigma}_Z)$	$\text{vec}(\boldsymbol{\mu}_{Z^*}) = \frac{a_L}{b_L} \boldsymbol{\Sigma}_Z (\mathbf{I}_k \otimes \text{Diag}(\boldsymbol{\mu}_W) \mathbf{C}) \text{vec}(\boldsymbol{\mu}_{h^*})$ $\boldsymbol{\Sigma}_Z = \left(\frac{a_L}{b_L} \mathbf{I}_k \otimes \mathbf{Q} + \frac{a_Z}{b_Z} \mathbf{T}_2^T \mathbf{T}_2 \right)^{-1}$
$q_W(\mathbf{w}) = \mathcal{N}(\boldsymbol{\mu}_W, \boldsymbol{\Sigma}_W)$	$\boldsymbol{\mu}_W = \boldsymbol{\Sigma}_W \text{diag} \left(\frac{a_R}{b_R} \boldsymbol{\mu}_J^T \mathbf{C} \boldsymbol{\mu}_X + \frac{a_L}{b_L} \boldsymbol{\mu}_h \mathbf{C} \boldsymbol{\mu}_{Z^*} \right)$ $\boldsymbol{\Sigma}_W = \left(\frac{a_R}{b_R} \mathbf{P}_1 + \frac{a_L}{b_L} \mathbf{P}_2 + \frac{a_W}{b_W} \mathbf{S}^T \mathbf{S} \right)^{-1}$

any data generative model \mathcal{M} regarding the observed data. This second step is achieved by calculating the model evidence $p(\mathbf{Y}, \mathbf{M} | \mathcal{M})$ in order to perform model comparison.

As for most useful generative models, the posterior dependencies between the parameters of our model \mathcal{M} renders $p(\mathbf{Y}, \mathbf{M} | \mathcal{M})$ a quantity difficult to calculate. The main principle of the variational Bayesian (VB) approach is to approximate the joint posterior pdf of all unknown parameters with a simpler distribution. This approximation usually posits further independence assumptions than those implied by the original generative model. In this work, we propose to use the so-called mean-field approximation (Ghahramani, 1995), which states that the joint posterior pdf can be factorized as the product of the marginal (approximate) posterior pdf of all nodes of the graphical model \mathcal{M} :

$$q(\boldsymbol{\theta}) \approx q_J(\mathbf{J}) q_h(\mathbf{h}) q_X(\mathbf{X}) q_Z(\mathbf{Z}) q_W(\mathbf{w}) q_\theta(\boldsymbol{\theta}), \quad (18)$$

where $\boldsymbol{\theta}$ is the vector of precision parameters ($\boldsymbol{\theta} = \{\alpha_1, \alpha_2, \epsilon_1, \epsilon_2, \beta_1, \beta_2, \gamma\}$), $\boldsymbol{\Theta}$ denotes the full set of nodes ($\boldsymbol{\Theta} = \{\mathbf{J}, \mathbf{h}, \mathbf{X}, \mathbf{Z}, \mathbf{w}, \boldsymbol{\theta}\}$) and $q(\cdot)$ denotes the variational approximation of any marginal posterior pdf $p(\cdot | \mathbf{Y}, \mathbf{M}, \mathcal{M})$. The structure of the graphical model \mathcal{M} (see Fig. 3) implies the separability of the joint posterior variational pdf $q_\theta(\boldsymbol{\theta})$:

$$q_\theta(\boldsymbol{\theta}) = q_{\alpha_1}(\alpha_1) q_{\alpha_2}(\alpha_2) q_{\epsilon_1}(\epsilon_1) q_{\epsilon_2}(\epsilon_2) q_{\beta_1}(\beta_1) q_{\beta_2}(\beta_2) q_\gamma(\gamma). \quad (19)$$

The mean-field approximation enables us to decompose the model evidence $p(\mathbf{Y}, \mathbf{M} | \mathcal{M})$ the following way:

$$\ln p(\mathbf{Y}, \mathbf{M} | \mathcal{M}) = \underbrace{\langle \ln p(\boldsymbol{\Theta}, \mathbf{Y}, \mathbf{M} | \mathcal{M}) \rangle_{\Pi q(\cdot)}}_{\mathcal{F}(q)} + \sum S(q(\cdot)) + \mathcal{D}_{KL}(q(\boldsymbol{\Theta}), p(\boldsymbol{\Theta} | \mathbf{Y}, \mathbf{M}, \mathcal{M})), \quad (20)$$

where $S(\cdot)$ is the Shannon entropy of any pdf, $\mathcal{D}_{KL}(\cdot)$ is the Kullback–Leibler divergence, “ \cdot ” denotes each node of the graphical model, and is therefore associated to some marginal pdf.

The quantity $\mathcal{F}(q)$ is called the negative free energy. Since the model evidence is fixed for a given generative model \mathcal{M} , maximizing the negative free energy is equivalent to minimizing

the Kullback–Leibler divergence between the mean-field approximation of the joint posterior pdf and the true posterior joint pdf. In other words, the higher the negative free energy, the closer to the true marginal posterior pdf the variational $q(\cdot)$.

Then, the functional form of the $q(\cdot)$ falls off from the maximization of the free energy:

$$\frac{\partial \mathcal{F}(q)}{\partial q(\cdot)} = 0 \Rightarrow \ln q(\cdot) \propto \langle \ln p(\boldsymbol{\Theta}, \mathbf{y} | \mathcal{M}) \rangle_{\Pi q(\cdot)}, \quad (21)$$

where the expectation is taken under the product of the marginals of the Markov blanket of each node.

The VB learning scheme is simply the iterative optimization of the negative free energy through updates of the sufficient statistics of the variational marginal pdf of each node of the graphical model \mathcal{M} . At convergence, the negative free energy is then used as an approximation (a lower bound) of the model evidence, and may be used for model comparison purposes. In other words, maximizing the negative free energy actually allows us to both find the marginal posterior of the model parameters and compute the model evidence.

Under the mean-field approximation, the variational posterior pdf of location and scale (precision) parameters of the model \mathcal{M} are Gaussian and Gamma densities, respectively. The sufficient statistics of each variational posterior marginal pdf $q(\cdot)$ (mean and covariance matrix for Gaussian pdf, degrees of freedom and shape parameter for Gamma pdf) are given in Tables 1 and 2.

Matrices \mathbf{P}_1 and \mathbf{P}_2 (size $n \times n$), and \mathbf{Q} (size $q \times q$) are defined the following way:

$$\begin{cases} \mathbf{P}_1 = \text{Diag} \left(\text{diag}(\mathbf{C}(\boldsymbol{\mu}_X \boldsymbol{\mu}_X^T) \mathbf{C}^T) + \sum_{j=1}^l \mathbf{C} \text{diag}(\boldsymbol{\Sigma}_X^{[j]}) \right) \\ \mathbf{P}_2 = \text{Diag} \left(\text{diag}(\mathbf{C}(\boldsymbol{\mu}_Z^T \boldsymbol{\mu}_Z) \mathbf{C}^T) + \sum_{j=1}^k \mathbf{C} \text{diag}(\boldsymbol{\Sigma}_Z^{[j]}) \right) \\ \mathbf{Q} = \mathbf{C}^T \left(\text{Diag}(\text{diag}(\boldsymbol{\Sigma}_W)) + \text{Diag}(\boldsymbol{\mu}_W)^2 \right) \mathbf{C}, \end{cases} \quad (22)$$

where $\boldsymbol{\Sigma}^{[i]}$ is the i th diagonal block (size $q \times q$) of matrix $\boldsymbol{\Sigma}$.

Table 2

Functional forms of the variational pdf of the scaling parameters of the graphical model \mathcal{M} , which are Gamma densities

Functional form	Parameters (d.o.f. and shape)
$q_{z_1}(\alpha_1) = \mathcal{G}(a_E, b_E)$	$a_E = \frac{pt_1}{2} + a_1$ $b_E = \frac{1}{2} \text{tr} \left(t_1 \Sigma_J \mathbf{G}^T \mathbf{G} + (\mathbf{M} - \mathbf{G} \mu_J)^T (\mathbf{M} - \mathbf{G} \mu_J) \right) + b_1$
$q_{z_2}(\alpha_2) = \mathcal{G}(a_F, b_F)$	$a_F = \frac{nt_2}{2} + a_2$ $b_F = \frac{1}{2} \text{tr} \left(n \Sigma_h \mathbf{B}^T \mathbf{B} + (\mathbf{Y} - \mathbf{B} \mu_h)^T (\mathbf{Y} - \mathbf{B} \mu_h) \right) + b_2$
$q_{\epsilon_1}(\epsilon_1) = \mathcal{G}(a_R, b_R)$	$a_R = \frac{nt_1}{2}$ $b_R = \frac{1}{2} \text{tr} \left[t_1 \Sigma_J + (\mu_J - \text{Diag}(\mu_W) \mathbf{C} \mu_X)^T (\mu_J - \text{Diag}(\mu_W) \mathbf{C} \mu_X) \right]$ $\quad + \frac{1}{2} \text{tr} \left[\Sigma_X \left(\mathbf{I}_{t_1} \otimes (\mathbf{C}^T \text{Diag}(\mu_W)^2 \mathbf{C} + \mathbf{Q}) \right) + \mu_X^T \mathbf{Q} \mu_X \right]$
$q_{\epsilon_2}(\epsilon_2) = \mathcal{G}(a_L, b_L)$	$a_L = \frac{nk}{2}$ $b_L = \frac{1}{2} \text{tr} \left[n \Sigma_h + (\mu_h - \mu_Z \mathbf{C}^T \text{Diag}(\mu_W))^T (\mu_h - \mu_Z \mathbf{C}^T \text{Diag}(\mu_W)) \right]$ $\quad + \frac{1}{2} \text{tr} \left[\Sigma_Z \left(\mathbf{I}_k \otimes (\mathbf{C}^T \text{Diag}(\mu_W)^2 \mathbf{C} + \mathbf{Q}) \right) + \mu_Z \mathbf{Q} \mu_Z^T \right]$
$q_{\beta_1}(\beta_1) = \mathcal{G}(a_X, b_X)$	$a_X = \frac{qt_1}{2}$ $b_X = \frac{1}{2} \text{tr} \left[\Sigma_X \mathbf{T}_1^T \mathbf{T}_1 \right] + \frac{1}{2} \text{vec}(\mu_X)^T \mathbf{T}_1^T \mathbf{T}_1 \text{vec}(\mu_X)$
$q_{\beta_2}(\beta_2) = \mathcal{G}(a_Z, b_Z)$	$a_Z = \frac{qk}{2}$ $b_Z = \frac{1}{2} \text{tr} \left[\Sigma_Z \mathbf{T}_1^T \mathbf{T}_1 \right] + \frac{1}{2} \text{vec}(\mu_Z^*)^T \mathbf{T}_1^T \mathbf{T}_1 \text{vec}(\mu_Z^*)$
$q_\gamma(\gamma) = \mathcal{G}(a_W, b_W)$	$a_W = \frac{n}{2}$ $b_W = \frac{1}{2} \text{tr} \left[\Sigma_W \mathbf{S}^T \mathbf{S} \right] + \frac{1}{2} \mu_W^T \mathbf{S}^T \mathbf{S} \mu_W$

Algorithmically, the Variational Bayesian learning scheme simply updates the sufficient statistics of the variational marginal pdf $q(\cdot)$ by successively applying the relations listed in Tables 1 and 2.

At convergence of the VB algorithm, we may define the parameter estimates as their expectation under their variational posterior pdf. Since we also quantify the uncertainty affecting these parameters (e.g., the variance of the variational posterior Gaussian pdf), we can build test statistics (e.g., Student's or Fisher's scores).

Evaluation

Numerical Monte Carlo simulations were performed in order to evaluate the proposed approach, which will be denoted by *BASTERF*, for Bayesian Spatio-Temporal Event-Related Fusion.

Simulations

In a sense, the BASTERF approach is a synthesis of analyzes of identification of ER bioelectric and hemodynamic responses already existing in EEG and in fMRI, respectively. Indeed, one of the objectives of the BASTERF approach is to characterize these responses *via* the estimation of time courses of current density (sources \mathbf{J} of the distributed model in EEG) and voxelwise hemodynamic responses (HRF \mathbf{h} of the GLM in fMRI).

Noticeably, the VB learning scheme provides us with the posterior marginal pdf of these quantities (see Table 1). The latter are used as estimates of modality-specific ER responses.

In the simulation series, the estimations of \mathbf{J} and \mathbf{h} were compared to usual estimators that were not dedicated to multimodal EEG/fMRI information fusion, i.e.,

- for EEG: minimum norm estimator (MNE) (Hauk, 2004), LORETA (Pascual-Marqui et al., 1994), ReML² (Mattout et al., 2006), and BASTA (Daunizeau et al., 2005a), which is the “pure EEG” analogue of BASTERF;
- for fMRI: a simple maximum likelihood approach (denoted as ML), and a fully Bayesian HRF estimation (BHE) method, which introduces a temporal smoothness prior assumption similar to that proposed in Prior densities of the coupling model parameters (Marrelec et al., 2003a).

In addition, we compared BASTERF to a classical asymmetrical multimodal EEG/fMRI information fusion approach, which consists of an fMRI-weighted minimum norm approach (WMN) (Liu et al., 1998). Note that the comparison of BASTERF and WMN with the “mono-modality” methods was only intended to

² The ReML approach to the EEG inverse problem included both i.i.d. and spatial Markov processes prior assumptions on the cortical bioelectric ER response using a mixture of prior covariance matrices. Their respective contribution is then assessed by estimating their optimal weight using the restricted maximum-likelihood scheme. The method is currently available as part of the SPM 5 software (SPM, 2005).

give an idea of the added-value (the gain or loss of information) of the “multi-modality” approaches.

BASTERF can be distinguished from established EEG or fMRI generative models by the number and diversity of its model parameters. Therefore, we chose to define a very simple simulation environment, in order to bring to light the behavior of the BASTERF approach with respect to the main issue of multimodal EEG/fMRI information fusion: the possible mismatch between bioelectric and hemodynamic activities.

Simulation environment

A cortical surface was extracted from a structural MRI of a normal subject and down-sampled to about 500 vertices ($n=458$). The $p=128$ EEG sensors were co-registered with the anatomical frame of reference of the subject, and the gain matrix was calculated using a three-sphere analytical model (Moshier and Leahy, 1998) using the SPM software (SPM, 2005).

The parameters of the simulation environment are given in Table 3.

We ran four series of simulations, mimicking four levels of concordance/discordance between bioelectric and hemodynamic sources:

- (S1): perfect concordance. For each simulation, two spatially extended bioelectrically and hemodynamically active sources were randomly chosen on the cortex;
- (S2): mixed concordance. For each simulation, two bioelectrically active sources were randomly selected on the cortex, one of them being also a hemodynamically active source. Stated another way, only one of the two sources that elicited an electrical potential measured in EEG also yielded a hemodynamic response;
- (S3): mixed concordance. For each simulation, two hemodynamically active sources were randomly chosen on the cortex, one of them was also bioelectrically active. In other words, one of the two fMRI existing sources also generated an electrical potential measured in EEG;
- (S4): weak concordance. For each simulation, one bioelectrically active source and one hemodynamically active source (non spatially concordant) were randomly chosen on the cortex. This was an extreme situation: the spatial support common to bioelectric and hemodynamic activity sources was the empty set. However, a given portion of the inactive sources support was common to both EEG and fMRI modalities.

For each series, 50 source configurations were simulated. Each source was randomly drawn on the cortical surface, and had a spatial extent of approximately 2 cm^2 . Time courses of

Table 3
Parameters of the simulation environment

EEG	fMRI
Sampling frequency: 1 kHz	Sampling frequency: 1 Hz ($TR=1 \text{ s}$)
Number of sensors: $p=128$	Size of simulated HRF: $k=30$
Number of time samples: $t_1=41$	Number of time samples: $t_2=870$
Gain matrix \mathbf{G} : 3 spheres (analytic)	Paradigm: random (~ 100 occurrences)
Noise \mathbf{E} : Gaussian i.i.d.	Noise \mathbf{F} : Gaussian i.i.d.
Signal-to-noise ratio: $SNR_{\text{EEG}}=1$	Signal-to-noise ratio: $SNR_{\text{fMRI}}=0.01$
	Number of dipoles/voxels: $n=458$
	Number of cortical parcels: $q=21$

bioelectrically active sources were half-sine curves (15 ms period). When two bioelectric sources were present (series (S1) and (S2)), one of the two sine time courses was shifted by 5 ms with respect to the other one. Each hemodynamic response was modeled by the temporal dynamics of a damped oscillator analogous to the canonical HRF as implemented in the SPM software (SPM, 2005) (one major positive peak and one minor negative peak 5 s and 12 s post-stimulus, respectively). When two hemodynamically active sources were present (series (S1) and (S3)), one of the two HRF was shifted by 7 s with respect to the other one.

The cortical parcelling required by both BASTA and BASTERF approaches was derived using the procedure described in Appendix B in order to obtain $q=21$ anatomically connected clusters.

Evaluation metrics

The quality of the estimation of bioelectric and hemodynamic ER responses was evaluated using the following criteria:

- the sum of squared errors (SSE).
- The localization error (LE), defined as the mean geodesic distance between the reconstructed and the simulated sources.
- The probability of correctly labelling a source as active/inactive, quantified using the area under the ROC curve.

These estimation adequacy scores were calculated for all estimated bioelectric and hemodynamic ER responses. We refer the interested reader to the Appendix C of this paper for the details concerning their implementation.

The behavior of estimators that are not dedicated to multimodal EEG/fMRI information fusion, i.e., LORETA, MNE, ReML and BASTA (for EEG), and ML and BHE (for fMRI), is not expected to be affected by the different simulations configurations (series S1, S2, S3 and S4). Hence, these approaches provide reference estimators to assess the behavior of multimodal information fusion procedures, i.e., WMN for estimating bioelectric activity and BASTERF for estimating both bioelectric and hemodynamic activity.

Results

Fig. 4 shows an example of source configuration from series (S2). Spatial supports of both brain activity markers (interpolated on the non-reduced cortical mesh), as well as the corresponding time courses are presented. In this simulation, both areas were bioelectrically active. The only hemodynamically non-active area was that located on the left frontal lobe. The simulated time courses of the two active areas respectively peaked at 21 and 26 ms, the latter being the one also exhibiting a hemodynamic ER response.

On this example, the fMRI-weighted MNE EEG inverse approach failed to recover the left frontal active source (not seen by fMRI), in contradistinction to the other methods (see bioelectrical ER response cortical maps in Fig. 4). However, only BASTA and BASTERF approaches prove able to recover the simulated time courses with accuracy (see highly noisy estimated time course on Fig. 4, more specifically for the fMRI-weighted MNE). Moreover, BASTERF benefitted from fMRI partial concordant information when estimating the left parieto-temporal source dynamics. As a consequence, its spatial and temporal characterization is better using BASTERF than using BASTA.

Fig. 5 (resp. 6 and 7) show the empirical cumulative distribution function of the SSE (resp. LE and ROC) score corresponding to the estimations of bioelectric ER response \mathbf{J} for

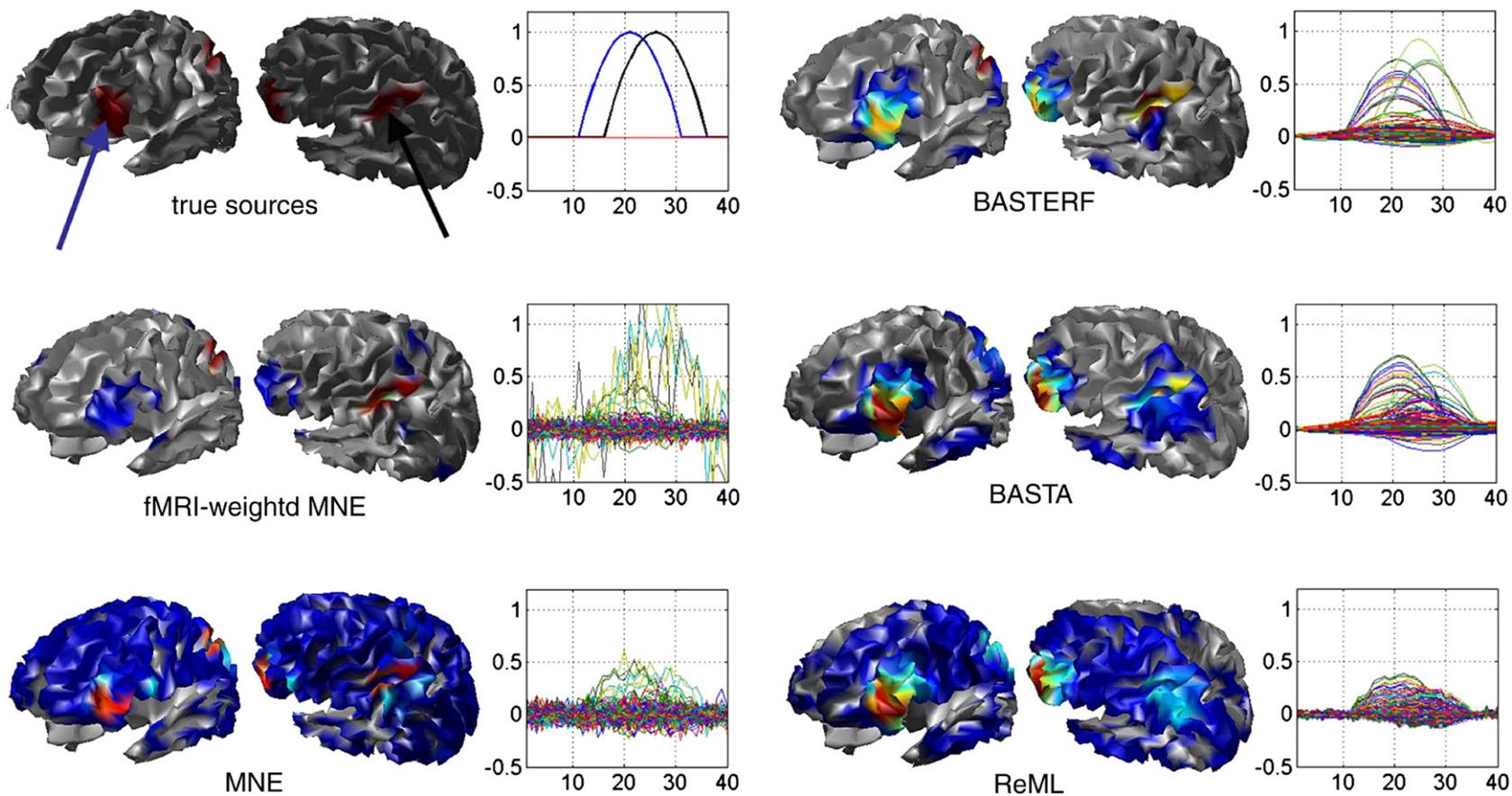


Fig. 4. Example of source reconstruction for the series of simulations (S2): bio-electric ER responses. From left to right and top to bottom: simulated EEG sources, BASTERF, fMRI-weighted MNE, BASTA, MNE and ReML reconstructions. We did not include LORETA for this example because of the poor quality of its estimation. The cortical maps show the summed power of the voxelwise time courses of estimated bioelectric ER responses, using normalized scales. The maps have been thresholded at the 0.5% quantile of the spatial distribution. The butter fly plots show the estimated bioelectric ER responses time courses.

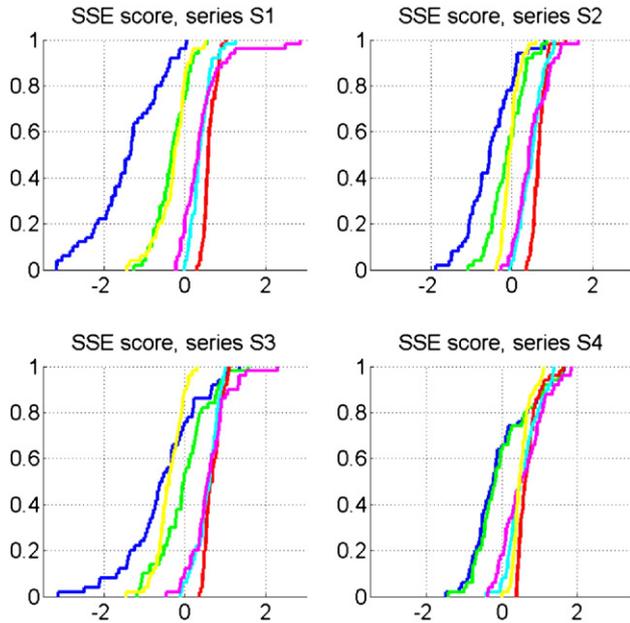


Fig. 5. Evaluation of the estimation of bioelectric ER response \mathbf{J} : SSE score. Diagrams show the empirical cumulative distribution function (cdf) of the SSE score over the four series of simulations. x - and y -axis represent the (log) SSE value and corresponding cdf, respectively. From left to right and top to bottom: S1, S2, S3 and S4. On each diagram, BASTERF is in blue, BASTA is in green, LORETA is in red, MNE is in turquoise, ReML is in violet and WMN is in yellow.

each of the four series of simulations (S1, S2, S3, S4). Means of these adequacy scores are also presented in Tables 4–6.

Means of SSE and LE adequacy scores associated to the estimation of the hemodynamic ER response \mathbf{h} for each of the four series of simulations (S1, S2, S3, S4) are also presented in Tables 7 and 8. We did not report the ROC score values in any table, since all simulations exhibited a ROC score of 1 for all methods. Neither did we plot the empirical cdf of the adequacy indices because they did not reveal any significant difference between the three methods.

First of all, we note that the estimation capabilities of the mono-modality approaches (BASTA, MNE, ReML for EEG and ML for fMRI) are stable whatever the level of EEG-fMRI discordance (i.e. the simulation series). Second, whenever there is a significant difference between the different approaches, BASTERF exhibits the best results for all adequacy scores (i.e. SSE, LE and ROC) (Figs. 6–8).

As for EEG, the general ranking of the mono-modality approaches is the following: BASTA > ReML > MNE > LORETA.

Table 4

Means of SSE score calculated on the estimation of the bioelectric ER response \mathbf{J} for the four series of simulations: S1, S2, S3 and S4 (and total average over all simulation series)

	BASTERF	BASTA	LORETA	MNE	ReML	WMN
(S1)	0.33	0.79	1.87	1.54	2.00	0.80
(S2)	0.69	0.98	1.99	1.68	1.77	0.99
(S3)	0.83	1.20	2.03	1.81	2.08	0.72
(S4)	1.30	1.33	2.10	1.93	2.10	1.73
Average	0.79	1.07	1.99	1.74	1.99	1.06

Table 5

Means of LE score (geodesic distance in mm) calculated on the estimation of the bioelectric ER response \mathbf{J} for the four series of simulations: S1, S2, S3 and S4 (and total average over all simulation series)

	BASTERF	BASTA	LORETA	MNE	ReML	WMN
(S1)	56.5	74.8	107.2	78.8	78.6	60.4
(S2)	55.8	68.0	113.8	71.3	74.5	84.8
(S3)	18.7	31.3	101.1	35.5	36.7	21.4
(S4)	27.7	27.5	84.0	39.5	30.8	102.2
Average	39.7	50.4	101.5	56.3	55.1	67.2

Besides, we note that LORETA exhibited the worst behavior in all simulations, whatever the adequacy score (except for the simulation series (S4), for which the fMRI-weighted MN algorithm behaved even worst). Moreover, the three adequacy scores showed a significant difference between BASTERF and BASTA, except for the simulation series (S4), where the two methods behaved similarly. These observations indicate that the improved localization properties of BASTERF are mainly due to the contribution of the fMRI information, whenever the latter is sufficiently concordant with the bioelectric ER response.

As for the adequacy of the estimation of the hemodynamic ER response, the simulations did not show any significant effect of the estimation scheme. Surprisingly, the estimation of the hemodynamic ER responses \mathbf{h} seems to be slightly more perturbed by the configuration (S3) than for (S4) (cf Table 7). This point will be made clearer later on.

Detailed analysis of the empirical cdf of adequacy scores SSE, LE and ROC allowed us to understand the dissimilarity of behavior between asymmetrical (WMN) and symmetrical approaches for the fusion of multimodal EEG/fMRI information.

On one hand, it is worth noticing that the quality of the bioelectric ER response estimation of the algorithm WMN decreased as the EEG/fMRI discordance increased:

- perfect concordance (S1): the fMRI-based constraint introduced in the WMN approach was relevant. The estimation of the bioelectric ER response $\hat{\mathbf{J}}_{\text{WMN}}$ was better than that of all mono-modality approaches, and equivalent to that of BASTERF;
- mixed concordance (series (S2) and (S3)): the fMRI-derived constraint was not fully relevant. The quality of the bioelectric ER response estimate $\hat{\mathbf{J}}_{\text{WMN}}$ got closer to that of the MNE approach. More precisely, for this kind of EEG/fMRI fusion approach, it seemed more compromising to omit part of the bioelectric active sources (series (S2)) than to define an fMRI-derived prior containing spurious sources (series (S3)). This particular behavior has already been reported in previous evaluations of WMN-like EEG/fMRI fusion approaches (Liu

Table 6

Means of ROC score calculated on the estimation of the bioelectric ER response \mathbf{J} for the four series of simulations: S1, S2, S3 and S4 (and total average over all simulation series)

	BASTERF	BASTA	LORETA	MNE	ReML	WMN
(S1)	1.00	0.97	0.67	0.89	0.94	1.00
(S2)	0.97	0.95	0.62	0.85	0.90	0.93
(S3)	0.99	0.95	0.70	0.90	0.92	1.00
(S4)	0.95	0.94	0.71	0.86	0.93	0.85
Average	0.98	0.95	0.68	0.88	0.92	0.95

Table 7

Means of SSE score calculated on the estimation of the hemodynamic ER response \mathbf{h} for the four series of simulations: S1, S2, S3 and S4 (and total average over all simulation series)

	BASTERF	ML	BHE
(S1)	0.07	0.74	0.24
(S2)	0.06	0.74	0.22
(S3)	0.12	0.73	0.24
(S4)	0.06	0.74	0.22
Average	0.08	0.74	0.23

et al., 1998; Ahlfors and Simpson, 2004; Daunizeau et al., 2005a). We will discuss this point further in Discussion;

- weak concordance (series (S4)): the constraint from fMRI in the WMN approach was irrelevant. The estimation of the bioelectric activity $\hat{\mathbf{J}}_{\text{WMN}}$ was strongly biased. As a consequence, WMN performed worst than any mono-modality algorithm (except LORETA).

On the other hand, the bioelectric ER response estimation of the approach BASTERF was relatively stable, whatever the EEG/fMRI concordance/discordance level (and w.r.t. all adequacy scores).

Of course, the BASTERF estimation of \mathbf{J} was perturbed by the EEG/fMRI concordance level (cf Fig. 5), and this effect can be seen clearly on the empirical cdf associated to the ROC score (cf Fig. 7). That is, estimation of bioelectric and hemodynamic ER responses is most efficient with a common spatial support that represents all cerebral activity sources. However, in opposition to WMN-like asymmetrical EEG/fMRI fusion approaches, BASTERF proved able to downweight the influence of fMRI (resp. EEG) data in the estimation of the bioelectric (resp. hemodynamic) activity when the EEG/fMRI concordance became contentious. As a consequence, the behavior of BASTERF (w.r.t. its bioelectric ER response estimation abilities) is similar to that of BASTA in situations of weak concordance between EEG and fMRI (simulation series S4). This might also explain the relative comparison between the hemodynamic ER response estimation capabilities of BASTERF for (S3) and (S4) series.

Application to interictal spike localization for one patient with focal epilepsy

Interictal spikes are transient events, characteristic of epilepsy, that occur between seizures. They are generated by the brain without any clinical signs, thus making multimodal imaging studies feasible.

In order to illustrate BASTERF on clinical data, we selected EEG/MRI data from a 32-year-old woman with drug-resistant focal epilepsy admitted at the Montreal Neurological Institute for presurgical investigation. The patient underwent simultaneous EEG-

Table 8

Means of LE score (geodesic distance in mm) calculated on the estimation of the hemodynamic activity sources \mathbf{h} for the four series of simulations: S1, S2, S3 and S4 (and total average over all simulation series)

	BASTERF	ML	BHE
(S1)	14.9	17.7	17.5
(S2)	16.9	15.7	19.2
(S3)	17.1	19.1	17.9
(S4)	17.5	18.8	17.7
Average	16.6	17.8	18.1

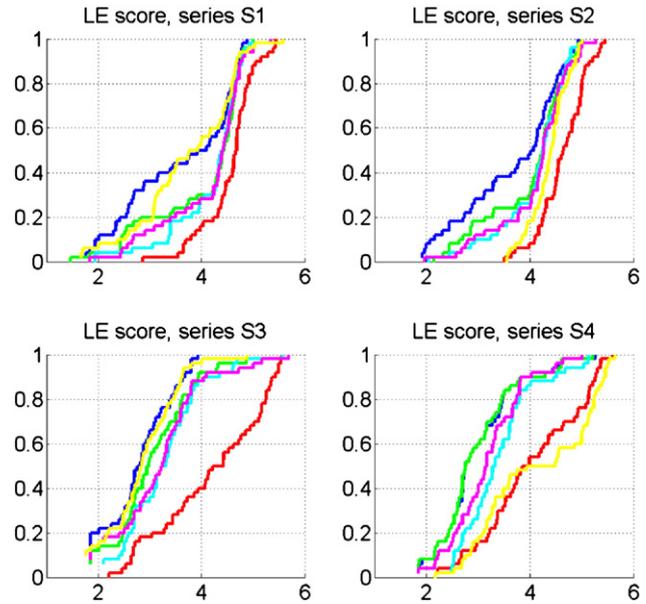


Fig. 6. Evaluation of the estimation of bioelectric ER response \mathbf{J} : LE score. Diagrams show the empirical cumulative distribution function (cdf) of the LE score over the four series of simulations. X and y-axis represent the (log) LE value and corresponding cdf, respectively. From left to right and top to bottom: S1, S2, S3 and S4. On each diagram, BASTERF is in blue, BASTA is in green, LORETA is in red, MNE is in turquoise, ReML is in violet and WMN is in yellow.

fMRI examination as described in (Gotman et al., 2004). An additional high resolution EEG recording session was performed outside the scanner immediately afterwards. This second EEG recording will be referred to as the “prolonged EEG” to distinguish it from the EEG recorded during fMRI scanning. Written informed consent was obtained in accordance with the regulations of the Research Ethics Board of the Montreal Neurological Institute and Hospital.

During the complete presurgical investigation, intracranial electrodes were also implanted. In this study, we will use intracranial analysis as a partial validation tool (Alarcon et al., 1994) to assess BASTERF results.

Data acquisition

The data acquisition protocol is described in details in (Gotman et al., 2004). The EEG-fMRI sessions were carried out in a 1.5 T Siemens Sonata scanner (Siemens, Erlangen, Germany) using 21 electrodes and an EMR32 amplifier recording at a sampling rate of 1 kHz (Schwarzer, Munich, Germany). A standard EPI fMRI sequence was used (voxel dimensions $5 \times 5 \times 5$ mm, 25 slices, 64×64 matrix, TE=50 ms, TR=3 s, ip angle 90°) and an anatomical scan was also acquired prior to fMRI recording (170 sagittal slices, 1 mm slice thickness, 256×256 matrix, TE=9.2 ms, TR=22 ms, ip angle 30°). The fMRI data were acquired in runs of 120 images lasting approximately 6 min each, followed by a short pause. The scanning session lasted for approximately 2 h in total, with 12 runs of fMRI data acquired. Following the EEG-fMRI scanning session, the patient was taken directly from the scanner to the clinical EEG department. Extra electrodes were added according to the 10–10 standard leading to a total of 44 electrodes. This prolonged EEG session lasted for approximately 45 min (sampling rate: 200 Hz).

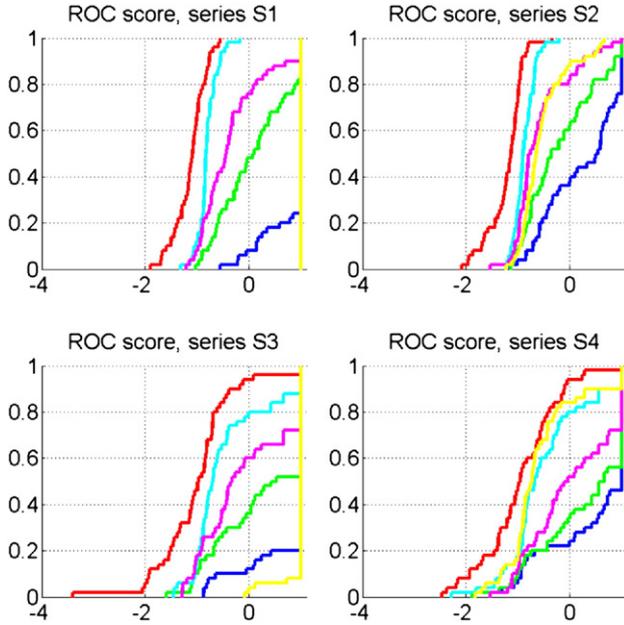


Fig. 7. Evaluation of the estimation of bioelectric ER response \mathbf{J} : ROC score. Diagrams show the empirical cumulative distribution function (cdf of the ROC score over the four series of simulations. x and y -axis represent the $\log(p/1-p)$ value and corresponding cdf, respectively (where p is the area under the ROC curve). From left to right and top to bottom: S1, S2, S3 and S4. On each diagram, BASTERF is in blue, BASTA is in green, LORETA is in red, MNE is in turquoise, ReML is in violet and WMN is in yellow.

Data preprocessing

The artefact induced on the EEG recording by the gradient switching during fMRI scanning was removed using the FEMR software (Schwarzer; Hoffmann et al. 2000).

An expert epileptologist detected manually the interictal spikes on both EEG recordings, i.e. simultaneous and prolonged EEG sessions. Spikes marked in the EEG acquired in the scanner were used to define the ER paradigm required for the identification of the hemodynamic response from fMRI data (i.e., the \mathbf{B} matrix, cf Eq. (3)). The spikes marked in the prolonged EEG recording were averaged. The resulting scalp ER potential was used for bioelectric ER response estimation in our approach (i.e. the measurements matrix \mathbf{M}) since it had more electrodes and was much less noisy than the EEG recorded inside the scanner.

The white matter/grey matter interface was segmented from the anatomical MRI scan using the BrainVISA software (BrainVISA, 2005). The cortical surface was then down-sampled in order to obtain a 4000 vertices mesh³. The EEG gain matrix was estimated using a three-sphere analytical model of the BrainStorm software (BrainStorm, 2005).

fMRI data were first motion-corrected, realigned and spatially smoothed (6 mm FWHM) using the BIC software package⁴. Then, fMRI data were prewhitened with an autoregressive filter of order

³ This was done using the `reducepatch.m` MATLAB (MATLAB, 2005) function.

⁴ BIC: Brain Imaging Center of the MNI: <http://www.bic.mcgill.ca/software/>.

1, and low frequency drifts were removed from the signal by fitting a third-order polynomial (Worsley et al., 2002).

Moreover, in order to define a common spatial support for both bioelectric and hemodynamic activities, volumic preprocessed fMRI data were interpolated on the cortical mesh extracted from the anatomical MRI (Grova et al., 2006) (see Appendix A). Surface fMRI data were then temporally linearly interpolated to 1 Hz pseudo-acquisition frequency.

Results

Fig. 8 shows the spatial distribution of a Student score $\mathbf{T}^{\text{BASTERF}}$ derived from the VB estimation of the spatial profile of the bioelectrically and hemodynamically active sources \mathbf{w} :

$$\mathbf{T}^{\text{BASTERF}} = \left(\frac{|\mu_{w_i}|}{\sqrt{(\Sigma_{w,i})_{1 \leq i \leq n}}} \right), \quad (23)$$

where μ_{w_i} and Σ_{w_i} are the first two moments (mean and covariance matrix) of the variational posterior pdf $q_{w_i}(\mathbf{w})$ of the temporal invariant \mathbf{w} . At convergence of the algorithm, these moments are given by the corresponding expressions in Table 1.

Four anatomical regions were found to be simultaneously active according to EEG and fMRI data (Fig. 8). Sources found in

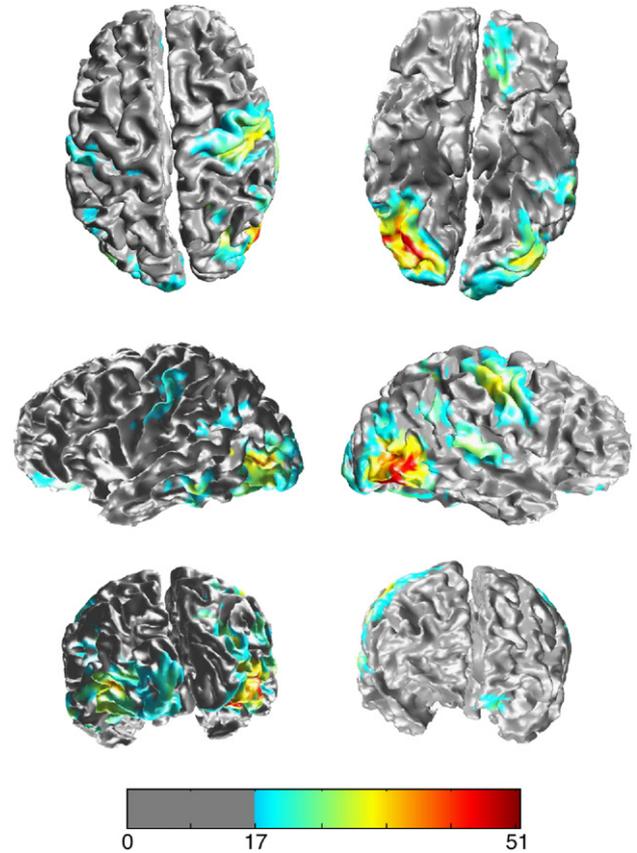


Fig. 8. Spatial distribution of a Student score $\mathbf{T}^{\text{BASTERF}}$ derived from the VB estimation of the spatial profile of the bioelectrically and hemodynamically active sources \mathbf{w} . From left to right and top to bottom: top, bottom, left, right, back and front views of the cortical surface. The three most significant sources are localized on left and right occipital lobes and on the right post-central gyrus. Moreover, one supplementary activated region appears onto the left post-central gyrus (see intracranial EEG results in Fig. 10).

the right and left occipital lobes and in the right post-central gyrus were the most significant ones, whereas an additional source located in the left post-central gyrus was also observed. Time courses corresponding to estimated bioelectric and hemodynamic activities within these four regions are shown in Fig. 9. Notice that all estimated hemodynamic time courses showed mainly a negative BOLD response (peaking between 4 and 6 s after the spike), that could be denoted as an fMRI deactivation in agreement with EEG sources.

Fortunately, implanted intracranial EEG electrodes were close to the four regions identified by BASTERF and these intracranial EEG recordings were then used to confirm our results (Fig. 10). First of all, these intracranial recordings were reviewed by an experienced electroencephalographer in order to infer which intracranial pattern of interictal spiking was the most likely to correspond to the activity visible on the scalp EEG, i.e., the type of spike we localized using BASTERF. Typical interictal spikes were manually detected on one particular epidural electrode located on the cortex (in yellow in Fig. 10(a)) and used to average the signal over all contacts of all implanted electrodes. Electrode contacts found close to the four regions identified by BASTERF are represented in Fig. 10(a) using the same color code as that used in Fig. 9. Average intracranial EEG signals corresponding to these four regions as well as all the other contacts are represented in Fig. 10(b). Intracranial EEG recordings confirmed the involvement of both occipital areas and the right post-central gyrus during the generation and the propagation of the interictal spikes analyzed in this study. Noticeably, intracranial EEG recordings showed also some activity elsewhere in regions not detected by BASTERF.

The cortical currents and the measured intracranial EEG electric potentials correspond to distinct physical quantities. Indeed, the former is likely to be the physical cause of the latter, i.e. intracranial EEG observations should be predicted by the electric potential field generated by the spatial distribution of cortical currents.

For instance, let us consider the polarity inversion seen on the 10 active plots located in the left occipital lobe (blue time courses

on Fig. 10): all sensors showed a highly synchronized activity for the main peak in addition to a gradient of amplitude from positive to negative values. This may be interpreted as the expression of a dipolar electric potential field generated approximatively at the barycenter of the active plots, i.e. in the middle of the left occipital BASTERF active source. Noticeably, this active area appeared to be spatially spread on two cortical parcels with highly correlated positive (at the main peak) time courses (see Fig. 9).

Intracranial EEG measurements also exhibited significant activity (related to the typical spikes analyzed in this study) close to the right occipital lobe, as well as behind the right post-central gyrus (cf green and red electrode contacts in Fig. 10(a)). However, we did not observe any polarity inversion on these intracranial signals, suggesting probably more distant generators, that could still be located close to the two most significant regions of the $T^{BASTERF}$ map shown on Fig. 8. It was not possible to further confirm such a hypothesis, because of the sparsity of the intracranial measurements available in these regions (especially in the right post-central gyrus).

Moreover, according to the intracranial electrodes located close to the least significant BASTERF region, i.e. left post-central gyrus (cf cyan electrode contacts in Fig. 10(a)), no significant epileptiform activity was observed at the time of the spike likely to be detected on the surface.

Furthermore, both intracranial EEG (cf Fig. 10(b)) and estimated cortical currents (Fig. 9(b)) seemed to exhibit a similar temporal scenario: the epileptiform activity started in the right occipital region, and later on spread in left occipital and right post-central gyrus regions, the last ones being mostly temporally coherent. Hence, the time dynamics of the estimated cortical currents, despite their enforced smoothness, are likely to be adequately rendered. These results seem to express a strong agreement between the BASTERF results and the intracranial EEG findings.

Last but not least, let us notice that the hemodynamic ER responses did not obey the same chronology as that of the bioelectric ER responses. From the point of view of fMRI, the right occipital hemodynamic ER response peak appeared later than that

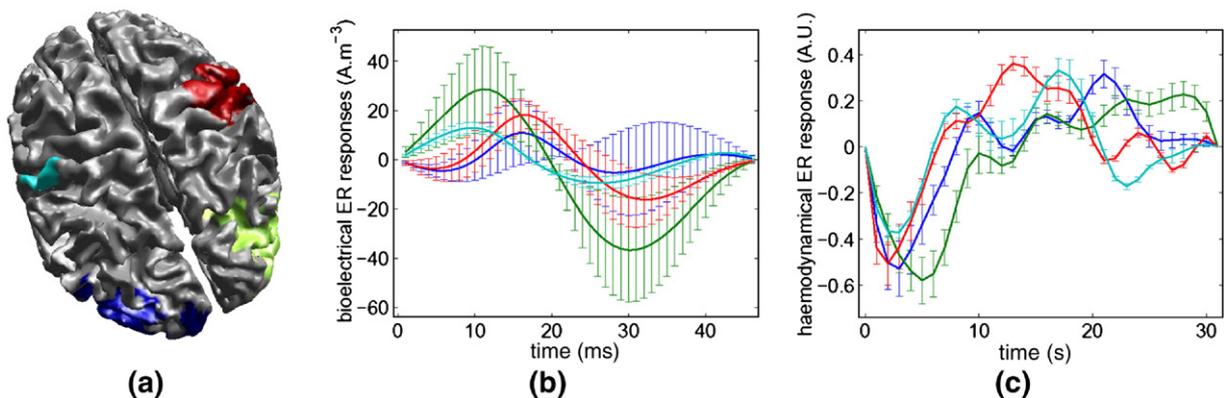


Fig. 9. Time courses of the bioelectric and hemodynamic activities of the common EEG/fMRI sources. (a) Regions found commonly active using BASTERF. (b) Mean estimated time courses of the bioelectric activity of these four regions (x- and y-axis: current density in $A \cdot m^{-3}$ and time in ms, respectively). (c) Mean estimated time courses of the hemodynamic activity of these four regions (x- and y-axis: amplitude of BOLD signal change in arbitrary units and time in s, respectively). On all subgyrus, green=right occipital region, blue=left occipital region, red=right post-central gyrus and turquoise=left post-central gyrus. Errorbars show half the standard deviation over the set of voxels belonging to the four activated regions.

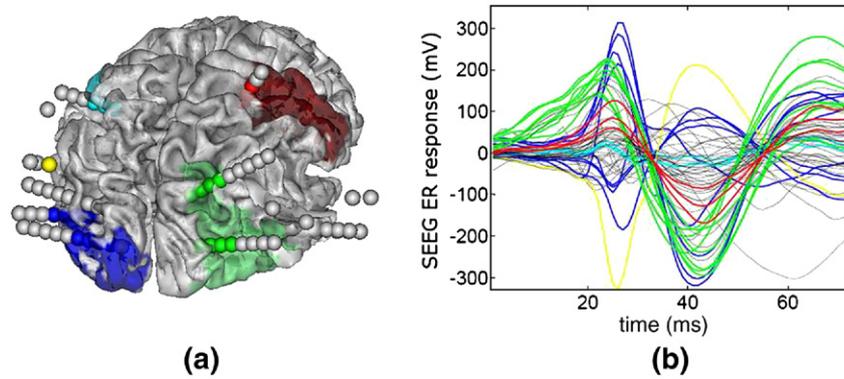


Fig. 10. Intracranial EEG results. (a) Superimposition of intracranial electrodes position on the cortical surface extracted from the anatomical MRI. The contact used to average the intracranial EEG signals is represented in yellow. The contacts located close to the four regions identified by BASTERF are represented using the same colors as in Fig. 9 (green=right occipital region, blue=left occipital region, red=right post-central gyrus and turquoise=left post-central gyrus). The remaining contrasts are represented in grey. Regions found commonly active using BASTERF are also represented using the same color code on the cortical surface. (b) Temporal dynamics of the bioelectric ER potential, as measured by implanted electrodes (yellow=electrode used to select spikes for averaging, green=right occipital focus, blue=left occipital focus, red=right post-central gyrus and turquoise=left post-central gyrus and black=remaining electrodes). Intracranial EEG results confirmed the presence of an epileptic discharge close to the three regions identified by BASTERF, i.e. localized into the right and left occipital lobes, as well as behind the right post-central gyrus (cf Fig. 8).

of the two other activated regions (according to fMRI temporal resolution and sampling rate). This point will be discussed further in Discussion.

Discussion

The spatially concordant ER responses model

In this work, we proposed a symmetrical approach to EEG/fMRI information fusion, which aims at estimating both bioelectric and hemodynamic ER responses. We therefore relied upon the spatially concordant ER responses model, which can be summarized as follows:

- a set of general and reliable prior assumptions about the expected features of brain activity are introduced *via* the data generative model \mathcal{M} . For instance, the prior spatio-temporal separability of \mathbf{J} and \mathbf{h} enables us to enforce a common and robust representation for both bioelectric and hemodynamic ER responses;
- coupling or uncoupling between EEG/fMRI is enabled through the notion of spatial support common to both bioelectric and hemodynamic ER responses. Importantly, given this profile, the structure of the generative model implies the independence of these two different types of brain activity. Moreover, at the limit, $\mathbf{w} \rightarrow \mathbf{0}_n$, equations related to EEG and fMRI data observation (Eqs. (1) and (2)) are decoupled. In other words, the classical EEG and fMRI data generative models pertain to the “uncoupling” limiting case for the spatially concordant ER responses model. Let us note that this situation is *a priori* assumed since the prior pdf of \mathbf{w} has zero mean. As a consequence, the structure of the EEG/fMRI coupling is learned exclusively from the data;
- the specification of the data generative model \mathcal{M} relies on a graphical (hierarchical) formalization of the problem, whose structure guides the variational approximations required for full Bayesian inference.

As a matter of fact, within our generative model, the spatio-temporal factorization is a practical route to define the common subspace of bioelectric and hemodynamic ER responses, i.e. the spatial support of the EEG/fMRI common sources. Moreover, through the parcelling, it offers “for free” a dimension reduction of the problem. This dimension reduction, in some way, is mandatory for a truly balanced EEG/fMRI information fusion approach. Indeed, the efficiency of any parameter estimation depends on the balance between the data likelihood and the prior. That is, the influence of the data in the estimation depends on the “quantity of information” available for each parameter. Now, fMRI provides a substantial amount of information regarding the spatial properties of the ER response. Therefore, in order to build a balanced symmetric information fusion, we have to reduce the “ill-posedness” of the EEG inverse problem. That is the “last but not least” consequence of the spatio-temporal factorization. Thus, this model for spatially concordant ER responses substantiates a true unsupervised well-balanced symmetrical EEG/fMRI information fusion approach.

Lastly, note that the introduction of the residuals \mathbf{R} and \mathbf{L} in the generative model has been motivated by the potential existence of modality-specific “neuronal activities” (cf subspaces ζ_2 and ζ_3 on Fig. 1). However, these residuals also potentially model some discrepancy between the actual common sources dynamics and their assumed spatio-temporal decomposition. *a priori*, it seems difficult to expect a particular spatial structure for these potential residuals. For that reason, we assumed \mathbf{R} and \mathbf{L} were realizations of i.i.d. Gaussian random variables, which is the least compromising hypothesis we could make. As a consequence, the VB learning scheme is exactly returning a minimum norm estimate of these quantities. As such, it is likely that the residuals are underestimated, which might be a potential limit of the current approach.

The variational Bayesian learning scheme

The VB learning scheme relies on the specification of the prior conditional pdf of the generative model parameters. In our

generative model, we have made use of non-informative prior pdf (for instance, Jeffreys priors on precision hyperparameters). Indeed, the building of non-informative distributions is the heart of a controversial debate: how should we parameterize our prior ignorance? Jeffreys priors are derived from formal rules (invariance principles) for choosing non-informative priors associated to families of parameters (location, scale, ...). Unfortunately, these priors are *improper*, i.e. unnormalized. The practical alternative to Jeffreys-like priors is the use of diffuse but proper prior pdf. In our case, this pertains to specify vague Gamma prior pdf for all precision parameters, which would embody the expected scale of magnitude of all the unknown hidden states (\mathbf{J} , \mathbf{h} , \mathbf{w} , \mathbf{X} and \mathbf{Z}). This would theoretically prevent the posterior from its potential impropriety. However, when improper priors lead to badly behaved posteriors, it is a warning that the problem itself may be hard; in this situation diffuse proper priors are likely to lead to similar difficulties (Kass and Wassermann, 1996).

A second argument in favor of our use of non-informative priors for precision hyperparameters may come from both the intuitive understanding of the hierarchical generative model and the simulation series. The observation Eqs. (1) and (2), combined with the informative (Gamma) prior on the measurement noises (cf Eq. (15)) are very likely to enforce a certain scale of magnitude for the estimated voxelwise dynamics \mathbf{J} and \mathbf{h} . Then, the conditional pdf given by Eqs. (8), (10) and (13) can be rewritten as a zero mean prior pdf for \mathbf{J} and \mathbf{h} , with a given covariance structure. This covariance structure is decomposed into spatial and temporal components (through the spatio-temporal factorization), which have to match the (data) imposed scale of \mathbf{J} and \mathbf{h} . In other words, the spatial and temporal variability of \mathbf{J} and \mathbf{h} is parameterized by a mixture of covariance matrices, which has to be “adjusted” through the estimation of the precision hyperparameters. Two limiting situations may then occur: either there is enough information in the data to fit the hyperparameters, or there is not. In the latter case, one may claim that the posterior Gamma hyperparameter pdf derived from the VB learning scheme may reflect any potential lack of data information, by showing, for instance, a high posterior variance of the precision hyperparameters. However, the simulation series did not show any striking difference between the posterior variances of the measurement noise precisions (whose marginal posterior are assured to be proper) and the other precision hyperparameters (whose posterior might, *a priori*, be improper).

Another comment is to be made about the VB inversion of the generative model. The hierarchical prior on \mathbf{w} is derived by making use of the discrete Laplacian operator \mathbf{S} (cf Prior densities of the coupling model parameters), which is rank deficient (i.e. of rank $n-1$). As a consequence, the prior Gaussian pdf of \mathbf{w} is formally degenerate, i.e. one of the principal axis of its prior covariance matrix has an infinite variance. Nevertheless, this does not lead to an improper posterior pdf, because its posterior covariance matrix is full-rank (see Table 1). An equivalent remark can be made for \mathbf{X} and \mathbf{Z} , whose hierarchical prior made use of the discrete second temporal derivative operators \mathbf{T}_1 and \mathbf{T}_2 . Despite the fact that this rank-deficiency does not invalidate the use of application of the VB framework, it may be a problem for the actual calculation of the variational free energy, for which one may have to resort to some full-rank approximation of the above-mentioned operators.

Furthermore, we should shed some light on the VB learning of the time invariant \mathbf{w} , which is the EEG/fMRI coupling key quantity. The common spatial profile of brain activity \mathbf{w} is estimated

according to a trade-off in fitting both the EEG and fMRI data. More precisely, the functional form of the variational marginal posterior pdf $q_{\mathbf{w}}$ is such that:

- this trade-off between the two terms of attach to the bioelectric and hemodynamic activities, respectively, is related to a measure of uncertainty associated to these cerebral activity markers (via the precision hyperparameters ϵ_1 and ϵ_2 , cf Table 1). In other words, the influence of the EEG or fMRI datasets in the estimation of their common spatial profile \mathbf{w} is an increasing function of the plausibility of the information they contain;
- the variational posterior covariance matrix of \mathbf{w} is a decreasing function of the variational posterior covariance matrices of \mathbf{X} and \mathbf{Z} (cf Table 2, Eq. (22)). This phenomenon illustrates the conservation of the total uncertainty associated with the subsystem (\mathbf{w} , \mathbf{X} , \mathbf{Z}), which is a commonly observed characteristic of dual variables. This enforces a weighting that penalizes the voxels/dipoles whose temporal (bioelectric and/or hemodynamic) characteristics are highly uncertain. In other words, the iterative estimation of \mathbf{w} may be considered as a selecting process of brain areas whose characteristics are the least uncertain given the observed EEG/fMRI datasets.

By construction, the estimated common spatial profile \mathbf{w} should reveal the bioelectrically and hemodynamically active areas that are characterized with no uncertainty by the joint EEG/fMRI datasets. \mathbf{w} then yields a quantitative image of the true fusion of multimodal information.

Indeed, the simulation series highlighted the expected behaviour of the invariant spatial profile \mathbf{w} . Let us consider the difference between WMN-like and BASTERF fusion approaches, which is most prominent for the simulation series (S4). These series have been described as “weak concordance” situations because despite the fact that no active source was contributing to both EEG and fMRI datasets, both datasets partially agree about the inactive sources. This puts into light another difference between the proposed approach and the more classical fMRI-weighted minimum norm approach to the EEG inverse problem: the VB inference scheme is explicitly estimating the uncertainty of the model parameters associated to *both* EEG and fMRI data generative models. As a consequence, the areas of the brain that do not contribute to any of the EEG and fMRI datasets are considered as *certainly inactive*. However, the areas that are expressed in one of the two datasets will be associated to a higher local variance. Intuitively, the algorithm has then to choose between these *less certainly inactive areas* to explain the respective datasets. This is not the case for fMRI-weighted minimum norm approaches, because these do not consider the local uncertainty of the fMRI prior.

The variational Bayesian learning scheme further enables us to test for the significance of the estimated common spatial profile \mathbf{w} . Such a test might reveal the “common spatial support” of bioelectric and hemodynamic activities. By *spatial support of activity*, we mean the infinite set of *significantly* activated voxels. Hence, the *common* spatial support is the intersection of bioelectrically and hemodynamically active voxels. This classification would allow us to localize the regions that exhibit a strong bioelectric/hemodynamic coupling for a given subject and a given experiment.

However, the mean-field approximation implies an underestimation of the degree of uncertainty associated to the estimations of the graphical model parameters. For instance, this may explain the large T-values obtained in the epileptic patient study (cf Fig. 8). As a

consequence, it might be difficult to propose an absolute way of thresholding the map \mathbf{w} . This particular point should be tackled with great care, and might benefit from an exhaustive comparison with classical Bayesian sampling approaches (such as Gibbs sampling).

The cortical parcelling

Up to now, we did not discuss the specific influence of the parcelling (\mathbf{C}), which constitutes, in this work, a fixed prior assumption about the structure of cortical activity.

As a matter of fact, the adequacy of the local spatio-temporal factorization assumption (within the parcels dynamics) is definitely dependent on the cortical parcelling itself. Noticeably, this sensitivity has been evaluated when first proposing the spatio-temporal model for the EEG inverse problem in (Daunizeau et al., 2005b). The conclusions were the following: as long as each parcel dynamics can be described using one approximated time course, the model is actually fine. Moreover, even when at least one prior cortical parcel contains more than one active source, the subsequent perturbation on the estimated spatial profile \mathbf{w} is “acceptable”, in the sense that the corresponding cortical source estimation ($\mathbf{J} = \text{Diag}(\mathbf{w}) \mathbf{C}\mathbf{X}$) is most of the time better than the standard LORETA (Pascual-Marqui et al., 1994), MN and WMN methods.

Nevertheless, the sensitivity of the whole approach to the potential inadequacy of the prior cortical parcelling still remains to be properly evaluated. The simulation series S1, S2, S3 and S4 were conducted such that no cortical parcel contained more than one active simulated source. Therefore, we did not evaluate the behavior of the whole approach when the prior hypothesis of temporal coherence of the parcels was mainly irrelevant. Although important, the question of estimating such a cortical parcel falls beyond the scope of this article. Several approaches have already been proposed to derive a cortical parcelling based on functional neuroimaging data (see (Flandin et al., 2002; Bellec et al., 2004) for fMRI and (Daunizeau et al., 2004; Lapalme et al., 2006) for EEG). However, no multimodal approach has yet been proposed. We are currently investigating the putative usefulness of such an approach, which, interestingly, might be consecutive to the inability of the model to describe properly the cortical voxelwise ER response. Indeed, in that case, the data itself will “require” a more adequate parcelling to be properly modelled.

Noticeably, the data generative model \mathcal{M} introduces \mathbf{J} and \mathbf{h} as perturbations of their respective spatio-temporal decomposition (cf Eqs. (4) and (5)). Hence, the residuals \mathbf{R} and \mathbf{L} enable us to relax the estimation of \mathbf{J} and \mathbf{h} and prevent it from inappropriate constraint of temporal coherence. As a consequence, it is possible to explicitly look for the parcelling that is associated with a maximal intra-parcel temporal coherence, according to both EEG and fMRI data. This optimization scheme may be introduced during the iterations of the VB algorithm, by specifying a prior pdf⁵ for the new unknown variable \mathbf{C} . Another practical alternative would be to compare different models associated to a reduced set of given parcellings using the variational free energy (approximation of the model evidence).

Comments on the observed neurovascular coupling

The validation part of this work has involved the use of multimodal epilepsy data, and relied on the available intracranial

EEG measurements to draw some conclusions regarding the behavior of the whole approach. First of all, one must emphasize the limitations of the intracranial EEG neuroimaging modality: it is based on local measurements. As a consequence, intracranial EEG does not provide us with any information about any potential activity remote to the implanted electrodes. Moreover, a quantitative comparison of intracranial EEG signals and reconstructed cortical currents would involve a dedicated forward model, exactly equivalent to that underlying the calculation of the EEG gain matrix (Chang et al., 2005). Particularly, this model should inform us about how to distort the amplitude of the intracranial EEG signals by rendering the spatial shape of the electrical potential field that could have been generated by the cortical currents. Though highly relevant, such a quantitative analysis of the intracranial EEG measurements is beyond the scope of this article.

Nevertheless, despite these concerns, the intracranial EEG measurements allowed us to (partially) validate the results of the BASTERF analysis of the epilepsy data, and to compare the bioelectric and the hemodynamic ER responses in the (clinical) context of interictal spikes localization. The latter lead us to make two striking observations:

1. The local hemodynamic ER responses do not obey the same chronology as the bioelectric ER responses;
2. A significant bioelectric ER response is locally associated to a negative hemodynamic ER response, i.e. a *local deactivation*.

The concern related to the temporal precedence of the ER responses has already been reported in (Bandettini et al., 1997). This has outstanding implications for any causality analysis within the active network: one may draw complete opposite conclusions about the causal relations of the nodes by relying on either bioelectric or hemodynamic ER responses. Since the hemodynamic response is partly driven by biophysical processes that are independent of the underlying neuronal activity (e.g., dynamics of the muscular cells which regulate the vasodilation), we are inclined to favor EEG-related analysis in any inference regarding the causal relations within the active network. In other words, one should always complement any fMRI analysis with EEG recordings in order to be able to infer both the nodes and the links of the active network.

The second point brings to light the difficulty in linking the temporal dynamics of bioelectric and hemodynamic activities. As a matter of fact, very few works have specifically focused on experimental evidence for candidate processes to deactivation. Deactivation is commonly interpreted as a transient metabolism suspension related to the event (Raichle et al., 2000). Relying on the “sharp-slow wavy” pattern of the scalp ERP of this patient, the fMRI deactivation may be interpreted by an increase in the inhibitory activity, which is believed to reduce the net presynaptic activity (Arthurs and Boniface, 2003). One explanation is that strong local inhibition is playing a role in keeping these epileptic discharges from spreading, and therefore could represent the most prominent phenomenon from the metabolic point of view (Stefanovic et al., 2005). We have no evidence, however, for whether this putative inhibitory activity is the cause or the consequence of the observed epileptiform activity.

Noticeably, these results would have unlikely been obtained by using a method enforcing a prior link between the time courses of the bioelectric and hemodynamic ER responses (e.g.: through an extension of the Balloon model (Riera et al., 2006)). This is mainly because the existing neurovascular coupling models do not

⁵ For example, this pdf could be a multinomial pdf (living in the labels space), whose mode might be defined as the Brodmann cortical parcelling.

entail specific mechanisms of deactivation. Hence, it may be of interest to primarily develop models able to generate coupling/uncoupling for both activation and deactivation processes, in order to design efficient and robust EEG/fMRI information fusion approaches.

Acknowledgments

The authors are grateful to Dr. Will Penny for fruitful discussions, to Dr. Jean Gotman for providing the data and to Dr. Eliane Kobayashi for marking spikes and helping with the intracranial EEG data analysis.

Appendix A. Anatomically informed interpolation of fMRI data on the cortical surface

The purpose of such an interpolation is to associate an fMRI time series to each vertex of the cortical surface defining the EEG distributed source model. Designing an optimal scheme to interpolate fMRI raw data on the cortical surface relies on a trade-off between choosing large enough interpolation kernels, because of the distributed nature of the hemodynamic response, and avoiding mixing data issued from different anatomical structures. The definition of these interpolation kernels should be robust to slight mis-alignments or distortions between anatomical and functional data.

Given these requirements, the method proposed in (Grova et al., 2006) automatically adjusts the level of such a trade-off, by defining interpolation kernels around each vertex of the cortical surface using a geodesic Voronoï diagram. Geodesic 3D distance to take into account cortical morphology. Interpolation kernels are then generated around each node of the surface using a geodesic Voronoï diagram, which ensures by construction the constraint of proximity around each node and the robustness to mis-registration. To integrate functional data over a sufficiently large area, the spatial support used to generate the Voronoï diagram consisted of the mask of the gray matter tissue segmented from the MRI, after a morphological dilatation of 3 mm, in order to take into account the spatial resolution of fMRI data.

To summarize, the different steps used to perform such an interpolation were the following: (1) segmentation of the cortical surface from the anatomical MRI, (2) definition of the anatomical and functional masks used as spatial supports for the interpolation, (3) construction of interpolation kernels using geodesic Voronoï diagrams starting from each node of the cortical surface and (4) integration of the fMRI signal at each node of the cortical surface.

Appendix B. Obtaining the cortical parcelling

In this work, we implicitly assumed that the cortical parcelling (C) was common to both data sets. This constraint enabled us to make a straight comparison of the temporal dynamics of bioelectric and hemodynamic activities of the parcels. Cortical parcelling methods from EEG (Daunizeau et al., 2004; Lapalme et al., 2006) or fMRI (Bellec et al., 2004; Simon et al., 2004) data sets already exist. However, no process has been yet developed to obtain such a parcelling from the structure of both bioelectric and hemodynamic activities. Therefore, we propose to use a very simple anatomical parcelling analogous (in the sense that it is not derived from any functional neuroimaging data) to Brodmann parcelling.

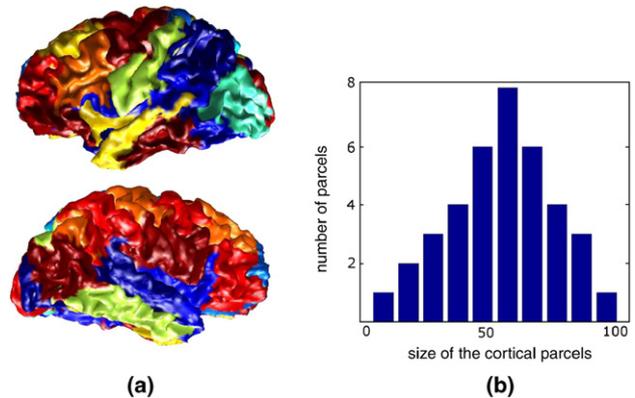


Fig. 11. Anatomical parcelling for BASTERF application. (a) Cortical field of handles associated with the parcels (all vertices belonging to the same parcel are represented using the same color). (b) Histogram of the distribution of the cortical parcel sizes.

In the clinical application presented in this article, this anatomical parcelling was obtained using a two-step procedure dedicated to cut the cortical mesh into connected regions. The first step consisted in down-sampling the cortical mesh up to 39 vertices⁶. Then, a region-growing algorithm handled neighboring vertices of the 4000-vertices mesh to the 39 seeds. Fig. 11 shows the cortical parcelling and the observed distribution of the sizes of the cortical parcels (i.e. their number of vertices).

Appendix C. Adequacy metrics

The sum of squared errors (SSE) was defined as:

$$\text{SSE}(\hat{\mathbf{J}}) = \frac{\sum_{i=1}^{t_1} \|\mathbf{J}_i - \hat{\mathbf{J}}_i\|^2}{\sqrt{\sum_{i=1}^{t_1} \|\mathbf{J}_i\|^2 \sum_{i=1}^{t_1} \|\hat{\mathbf{J}}_i\|^2}} \quad (24)$$

where \mathbf{J} and $\hat{\mathbf{J}}$ are the bioelectric ER responses simulated and estimated, respectively. The application to the hemodynamic ER response is straightforward.

The localization error (LE) was an estimation of the cortical (geodesic, i.e. non-Euclidean) distance between the center of the simulated source and the global maximum of the squared cortical bioelectric/hemodynamic ER response at the peak of the simulated source. Whenever two sources were simulated (simulations series (S1) and (S2) for EEG, (S1) and (S3) for fMRI), we reported the mean localization error.

The ROC analysis was conducted such as to compare the cortical binary mask corresponding to the simulated areas and the power map \mathbf{P}_J defined as:

$$\mathbf{P}_J = \left(\sum_{i=1}^{t_1} J_{i,t}^2 \right)_{i=1, \dots, n} \quad (25)$$

for the bioelectric ER response (and equivalently for the hemodynamic ER response). The area under the ROC curve is

⁶ This was done using the reducepatch.m MATLAB (MATLAB, 2005) function.

then the probability to correctly disambiguate between an activated and an inactivated elementary dipole, on the basis of the power map \mathbf{P}_j . Note that this index provides an accurate estimation of the global spatial behavior of an estimator (in terms of both its localization and extent estimation capabilities).

References

- Ahlfors, S.P., Simpson, G.U., 2004. Geometrical interpretation of fMRI-guided MEG/EEG. *NeuroImage* 22, 323–332.
- Alarcon, G., Guy, C., Walker, S., Elwes, R., Polkey, C., 1994. Intracerebral propagation of interictal activity in partial epilepsy: implications for source localization. *J. Neurol. Neurosurg. Psychiatry* 57, 435–449.
- Arthurs, O.J., Boniface, S.J., 2003. What aspect of the fMRI BOLD signal best reflects the underlying electrophysiology in human somatosensory cortex? *Clin. Neurophysiol.* 114, 1203–1209.
- Babiloni, F., Babiloni, C., Carducci, F., Romani, G.L., Rossini, P.M., Angelone, L., Cincotti, F., 2003. Multimodal integration of high-resolution EEG and functional magnetic resonance imaging data: a simulation study. *NeuroImage* 19, 1–15.
- Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. *IEEE Signal Process. Mag.* 18, 14–30.
- Bandettini, P.A., Kwong, K.K., Davis, T.L., Tootell, R.B.H., Wong, E.C., Fox, P.T., Belliveau, J.W., Weissko, R.M., Rosen, B.R., 1997. Characterization of cerebral blood oxygenation and flow changes during prolonged brain activation. *Hum. Brain Mapp.* 5, 93–109.
- Bellec, P., Marrelec, G., Posé, C., Jolivet, O., Pélégrini-Issac, M., Doyon, J., Benali, H., 2004. An assisted data-driven procedure to select functionally connected brain regions in fMRI. *Proceedings of the 10th International Conference on Functional Mapping of the Human Brain*. Available on NeuroImage CD-Rom, volume 22, Budapest, Hungary.
- BrainStorm, 2005. BrainStorm software. Available at <http://neuroimage.usc.edu/brainstorm/>.
- BrainVISA, 2005. BrainVISA software. Available at <http://brainvisa.info/>.
- Chang, N., Gulrajani, R., Gotman, J., 2005. Dipole localization using simulated intracerebral EEG. *Clin. Neurophysiol.* 116, 2707–2716.
- Dale, A.M., Sereno, M., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5, 162–176.
- Daunizeau, J., Mattout, J., Goulard, B., Lina, J.M., Benali, H., 2004. Data-driven cortex parcelling: a regularization tool for the EEG/MEG inverse problem. *Second 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'04)*, pp. 1343–1346. Arlington, VA.
- Daunizeau, J., Grova, C., Mattout, J., Marrelec, G., Clonda, D., Goulard, B., Lina, J.M., Benali, H., 2005a. Assessing the relevance of fMRI-based prior in the EEG inverse problem: a Bayesian Model Comparison approach. *IEEE Trans. Sign. Proc.* 53, 3461–3472.
- Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., Lina, J.M., 2005b. Bayesian spatio-temporal approach for EEG source reconstruction: conciliating ECD and distributed models. *IEEE Trans. Biomed. Eng.* 53, 503–516.
- de Munck, J.C., 1988. The potential distribution in a layered spheroidal volume conductor. *J. Appl. Phys.* 64, 464–470.
- Flandin, G., Kherif, F., Pennec, X., Rivière, D., Ayache, N., Poline, J.B., 2002. Parcellation of brain images with anatomical and functional constraints for fMRI data analysis. *Proc. IEEE ISBI*, pp. 907–910. Washington, DC, USA, Jul.
- Friston, K.J., 2005a. Models of brain function in neuroimaging. *Annu. Rev. Psychol.* 56, 57–87.
- Friston, K.J., 2005b. A theory of cortical responses. *Philos. Trans. R. Soc.* 360, 815–836.
- Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., Turner, R., 1995. Analysis of fMRI time-series revisited. *NeuroImage* 2, 45–53.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1998. *Bayesian Data Analysis*. Chapman and Hall, Londres.
- Ghahramani, Z., 1995. Factorial learning and the EM algorithm. *Adv. Neural Inf. Process. Syst.* 7, 617–624.
- Gonzales, A.S.L., Blanke, O., Lantz Thut, G., Grave de Peralta Merendez, R., 2001. The use of functional constraints for the neuromagnetic inverse problem: alternatives and caveats. *Int. J. Bioelectromagn.* 3, 103–114.
- Gössl, C., Auer, D., Fahrmeir, L., 2001. Bayesian spatiotemporal inference in functional magnetic resonance imaging. *Biometrics* 57, 554–562.
- Gotman, J., Benar, C.G., Dubeau, F., 2004. Combining EEG and fMRI in epilepsy: methodological challenges and clinical results. *J. Clin. Neurophysiol.* 21, 179–192.
- Grova, C., Makni, S., Flandin, G., Ciuciu, P., Gotman, J., Poline, J.B., 2006. Anatomically informed interpolation of fMRI data on the cortical surface. *NeuroImage* 31, 1475.
- Hauk, O., 2004. Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. *NeuroImage* 21, 1612–1621.
- Heeger, D.J., 2002. What fMRI tells us about neuronal activity? *Nature* 3, 142–151.
- Kass, R.E., Wassermann, L., 1996. The selection of prior distributions by formal rules. *J. Am. Stat. Assoc.* 91, 1343–1370.
- Lapalme, E., Lina, J.M., Mattout, J., 2006. Data-driven parcelling and entropic inference in MEG. *NeuroImage* 30, 160–171.
- Liu, A.K., Belliveau, J.W., Dale, A.M., 1998. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalographic data: Monte Carlo simulations. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8945–8950.
- Logothetis, N.K., Wandell, B.A., 2004. Interpreting the BOLD signal. *Annu. Rev. Physiol.* 66, 735–769.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neuro-physiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.
- Marrelec, G., Benali, H., Ciuciu, P., Poline, J.B., 2001. Bayesian estimation of the hemodynamic response function in functional MRI. In: Fry, R. (Ed.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Workshop Proceedings. AIP, Melville, pp. 229–247.
- Marrelec, G., Benali, H., Ciuciu, P., Pélégrini-Issac, M., Poline, J.B., 2003a. Robust Bayesian estimation of the hemodynamic response function in event-related BOLD fMRI using basic physiological information. *Hum. Brain Mapp.* 19, 1–17.
- Marrelec, G., Ciuciu, P., Pélégrini-Issac, M., Benali, H., 2003b. Estimation of the hemodynamic response function in event-related functional MRI: directed acyclic graphs for a general Bayesian inference framework. In: Taylor, C.J., Noble, J.A. (Eds.), *Information Processing in Medical Imaging*. Springer-Verlag, Berlin, pp. 635–646.
- Marrelec, G., Ciuciu, P., Pélégrini-Issac, M., Benali, H., 2004. Estimation of the hemodynamic response function in event-related functional MRI: Bayesian networks as a framework for efficient Bayesian modelling and inference. *IEEE Trans. Med. Imag.* 23, 959–967.
- Martindale, J., Mayhew, J., Berwick, J., Jones, M., Martin, C., Johnston, D., Redgrave, P., Zheng, Y., 2003. The hemodynamic impulse response to a single neural event. *J. Cereb. Blood Flow Metab.* 23, 546–555.
- Mathiesen, C., Caesar, K., Akgoren, N., Lauritzen, M., 1998. Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *J. Physiol.* 512, 555–566.
- MATLAB. MATLAB software, 2005. Available at <http://www.mathworks.com/>.
- Mattout, J., Philipps, C., Rugg, M.D., Friston, K.J., 2006. MEG source localization under multiple constraints: an extended Bayesian framework. *NeuroImage* 30, 753–767.
- Mosher, J.C., Leahy, R.M., 1998. Recursive MUSIC: a framework for EEG and MEG source localization. *IEEE Trans. Biomed. Eng.* 45, 1342–1354.
- Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., Malach, R., 2005. Coupling between neuronal ring, field potentials, and fMRI in human auditory cortex. *Science* 5, 951–954.

- Nunez, P.L., 1981. *Electric Fields of the Brain*. New York Press, Oxford.
- Nunez, P.L., Silberstein, R.B., 2000. On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? *Brain Topogr.* 13, 79–96.
- Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low Resolution Electromagnetic Tomography: a new method for localizing electrical activity in the brain. *Int. J. Psychophysiol.* 18, 49–65.
- Pieger, P.J., Greenblatt, R.E., 2001. Nonlinear analysis of multimodal dynamic brain imaging data. *Int. J. Bioelectromagn.* 3 (Available at <http://www.ijbem.org/volume3/number1/toc.htm>).
- Puce, A., Allison, T., Spencer, S.S., Spencer, D.D., McCarthy, G., 1997. Comparison of cortical activation evoked by faces measured by intracranial field potentials and functional MRI: two cases studies. *Hum. Brain Mapp.* 5, 298–305.
- Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Schuman, G.L., 2000. A default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98, 67–682.
- Riera, J., Wan, X., Jimenez, J.C., Kawashima, R., 2006. Nonlinear local electrovascular coupling: I. A theoretical model. *Hum. Brain Mapp.* 25, 1025–1041.
- Simon, O., Kherif, F., Flandin, G., Poline, J.B., Rivière, D., Mangin, J.F., Le Bihan, D., Dehaene, S., 2004. Automatized clustering and functional geometry of human parietofrontal networks for language, space, and number. *NeuroImage* 23, 1192–1202.
- SPM, 2005. SPM 5 software. Available at <http://www.fil.ion.ucl.ac.uk/spm/>.
- Stefan, K., Harrison, L.E., Penny, W.D., Friston, K., 2004. Biophysical models of fMRI responses. *Curr. Opin. Neurobiol.* 14, 629–635.
- Stefanovic, B., Warnking, J.M., Kobayashi, E., 2005. Hemodynamic and metabolic responses to activation, deactivation and epileptic discharges. *NeuroImage* 28, 205–215.
- Trujillo, N.J., Martinez, E., Melie, L., Valdès-Sosa, P.A., 2001. Bayesian model for fMRI and EEG/MEG neuroimage fusion. *Int. J. Bioelectromagn.* 3 (Available at <http://www.ijbem.org/volume3/number1/toc.htm>).
- Worsley, K.J., Liao, C., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. *NeuroImage* 15, 1–15.