SYMMETRICAL EEG–FMRI IMAGING BY SPARSE REGULARIZATION

Thomas Oberlin *, Christian Barillot †, Rémi Gribonval †, Pierre Maurel †

* INP–ENSEEIHT and IRIT, University of Toulouse, Toulouse, France
† Inria – Irisa, Rennes, France

ABSTRACT

This work considers the problem of brain imaging using simultaneously recorded electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). To this end, we introduce a linear coupling model that links the electrical EEG signal to the hemodynamic response from the blood-oxygen level dependent (BOLD) signal. Both modalities are then symmetrically integrated, to achieve a high resolution in time and space while allowing some robustness against potential decoupling of the BOLD effect.

The novelty of the approach consists in expressing the joint imaging problem as a linear inverse problem, which is addressed using sparse regularization. We consider several sparsity-enforcing penalties, which naturally reflect the fact that only few areas of the brain are activated at a certain time, and allow for a fast optimization through proximal algorithms. The significance of the method and the effectiveness of the algorithms are demonstrated through numerical investigations on a spherical head model.

Index Terms— EEG-fMRI; multimodal imaging; structured sparsity; EEG inverse problem

1. INTRODUCTION

Over the past decades, numerous methods have been developed to non-invasively measure and analyze brain activity. Electroencephalography (EEG) and, to a lower extent, functional Magnetic Resonance Imaging techniques (fMRI), have been intensively used by neuroscientists, psychologists as well as physicians. Yet, the "brain activity" seen through each of these modalities is very different.

On one side, the EEG gives instantaneous measures of the brain electrical activity, produced by synchronized pyramidal neurons under the cortex. However reconstructing the electrical activity inside the brain is an ill-posed inverse problem, since these electrical currents are highly perturbed by diffusion across the skull and the scalp, and since one only measures them at a few locations. Even though many techniques manage to estimate this activity [1], their resolution

in space remains low. On the other side, functional imaging of the brain is achieved now intensively in neurosciences or in clinical studies through fMRI with an increased spatial precision. The most common fMRI procedure measures the Blood-oxygen-level dependent (BOLD) signal, which is related to the deoxygenation of the blood, and thus reveals with high spatial accuracy the regions of the brain involved in a cognitive task. However, the relatively long physiological processes that link the "neuronal activity" and the measured BOLD signal prevent the reconstruction of this activity with a high temporal resolution.

Numerous attempts have tried to combine both modalities, improving our knowledge about the neurophysiological coupling as well as the space-time resolution of the imaging techniques. Among them, asymmetrical methods consist in informing one modality with the other, by constraining the EEG inverse problem to a specific region determined by fMRI [2], or by using time-frequency patterns from the EEG in the fMRI data processing [3]. More recently, several authors proposed to integrate both modalities symmetrically, in order to prevent the results of one modality to be perturbed by the other one. Such approaches use neurovascular coupling models within a Bayesian framework [4] or stochastic filtering [5].

The approach described hereafter is in the same spirit: it builds a simplified linear model for the neurovascular coupling, and addresses the corresponding inverse problem through sparsity-regularized regression. The paper is organized as follows: Section 2 describes the coupling model and states the inverse problem, for which numerical analysis and algorithms are derived in Section 3. Section 4 is devoted to numerical experiments on simulated data.

2. A LINEAR SYMMETRICAL COUPLING OF EEG AND FMRI

2.1. Distributed source model in EEG

One usually represents the neuronal electrical activity as resulting from a certain number of electric dipoles, which create a current. Solving the quasi-static approximation of the Maxwell's equations allows to express the measurements as a known linear instantaneous mixing of these sources. The estimation of the dipole activity from the measurements is com-

monly referred as the (EEG) inverse problem. To solve this inverse problem, we will adopt the distributed source model [6], which consists in fixing a high number ($S \approx 10^4$) of dipoles (or sources) on the cortical surface, oriented perpendicularly to that surface, and in estimating the amplitude of these dipoles. In the discrete setting, the source activity will be denoted by the $S \times T$ matrix $X = X_{s,t}$, the dimensions corresponding to time (index t) and space (s), respectively. The measured activity, denoted by $E \in \mathbb{R}^{N \times T}$ ($N \ll S$ is the number of electrodes, T the number of time instants), writes:

$$E = GX + N_E \ , \tag{1}$$

where G is the mixing "leadfield" (or gain) matrix, of size $N \times S$, and N_E is a "noise" term, containing both measurement errors and physiological noise. The leadfield matrix can be computed accurately from a morphological image of the subject's brain and is thus assumed to be known. Note that in general the noise is not white, but can be considered as Gaussian with an estimated covariance matrix.

2.2. The neurovascular coupling

The main difficulty in symmetrical EEG–fMRI coupling lies in the complexity of the relation between the neuronal activity and the BOLD signal. Both quantities are related by a cascade of complex physiological processes that are not yet fully understood. A model of this relationship has been proposed in [7], through several differential equations involving physiological and anatomical parameters. A simplified model can be found in [8], which linearizes the "balloon model" of [7]. Here, the BOLD response is obtained by temporally convolving the neuronal activity with a function h similar to the canonical hemodynamic response function (HRF) used in standard BOLD fMRI.

Assuming this linear convolutive coupling, we can model the BOLD signal F as a function of the neuronal activity through operators Q and H in the following way:

$$F = QXH + N_F \ . \tag{2}$$

Typically, Q is a spatial sampling (or interpolation) operator, while $H \in \mathbb{R}^{T \times U}$ encodes the temporal convolution as well as a subsampling (in time). Its coefficients write $H_{ij} = h(t_i - u_j)$, with t_i and u_j the time instants corresponding to EEG and fMRI respectively, and h a canonical HRF. To simplify the problem, we will assume in the sequel that the BOLD signal is measured at the exact locations of the dipoles. We will thus neglect operator Q. To summarize, the electrical activity X of the sources is related to the BOLD signal F and the EEG measures E by the linear system:

$$\begin{cases} E &= GX + N_E \\ F &= XH + N_F \end{cases} \ . \tag{3}$$

2.3. Joint EEG–fMRI inverse problem

To reconstruct the activity X from measurements (E, F), assuming Gaussian noise one seeks to minimize the observation error in the least-square sense. Since the problem is *ill-posed* and admits infinitely many solutions, one needs to regularize it through some penalty function ϕ, which leads to the following optimization problem:

$$X^* = \arg\min_X \left(\frac{\alpha}{2} \|E - GX\|_F^2 + \frac{1-\alpha}{2} \|F - XH\|_F^2 + \lambda\phi(X) \right) \ , \tag{4}$$

where λ is a positive parameter tuning the tradeoff between the data-fidelity and the regularization terms, and $\alpha \in [0,1]$ balances the information used from each modality. Let us mention that the minimum is unique as soon as the penalty ϕ is convex, but may be achieved at several points, unless ϕ is strongly convex. We describe in the following section the algorithm used to solve problem (4), as well as the strategies to choose the penalty ϕ and the parameters α, λ.

3. SOLVING THE INVERSE PROBLEM

3.1. Proximal algorithms

The wide class of iterative-thresholding algorithms has been introduced as early as in the sixties, and has been intensively studied in the field of image/signal processing [9]. They aim to solve optimization problems where one term of the objective function is not differentiable. The minimization problem writes:

$$\min f(x) + \phi(x) \ , \tag{5}$$

where f is a differentiable function, with L-Lipschitz continuous gradient, and ϕ is proper and convex. When applied to problem (4), function f is quadratic with a Lipschitz constant of its gradient smaller than $L = \alpha \|G^*G\| + (1-\alpha) \|HH^*\|$, where $\|.\|$ denotes the spectral norm of a matrix. The Iterative Shrinkage-Thresholding Algorithm (ISTA) consists in updating the current solution x_k with:

$$x^{k+1} = \text{prox}_{\frac{1}{L}\phi}(x^k - \frac{1}{L}\nabla f(x^k)) \ , \tag{6}$$

where the proximal operator of the proper convex function ϕ is defined by $\text{prox}_\phi(x) = \arg\min_{y \in \mathbb{R}^n} \frac{1}{2} \|x - y\|_2^2 + \phi(y)$. One can prove that, whatever the initialization x^0, this algorithm (ISTA, also called FB for Forward-Backward proximal optimization) converges towards a minimizer, with asymptotic linear rate [9].

3.2. Choice of the penalty term

We will now mention some possible choices for the function ϕ, which must reflect the a priori information available on the true solution. The naive regularization uses a squared ℓ_2 norm, which tends to reduce the energy of the solution.

Unfortunately, this spreads the energy of the solution across all sources, being unrealistic with EEG data. Instead, it is reasonable to assume that sources are *sparse*, only a few of them being active at a certain time. The most obvious way of promoting sparsity is to set ϕ as the ℓ_0 pseudo-norm, which counts the number of non-zero coefficients. But the optimization problem becomes combinatorial, and in practice intractable. Instead, a wide literature suggests to use the ℓ_1 norm, defined by $\|X\|_1 = \sum_{s,t} |X_{s,t}|$. Using the ℓ_1 penalty makes problem (4) convex, which allows to use the ISTA algorithm, the proximal operator being the well-known soft-thresholding, defined for $\mu > 0$ as:

$$[\text{prox}_{\mu\|.\|_1}(x)]_j = \text{sgn}(x_j)(|x_j| - \mu)_+ = \left(1 - \frac{\mu}{|x_j|}\right)_+ x_j \ . \tag{7}$$

However, the ℓ_1 norm favors sparse but scattered sources and does not take into account the *temporal persistence* of the sources, namely their consistent behavior across time. Otherwise stated, the ℓ_1 norm considers the matrix X as a long one-dimensional vector. In order to enforce both temporal persistence and spatial sparsity of the sources, an alternative [6] is to use the ℓ_{12} mixed-norm, which writes

$$\|X\|_{12} = \sum_{s=1}^{S} \left(\sum_{t=1}^{T} X_{s,t}^2\right)^{\frac{1}{2}} \ , \tag{8}$$

whose proximal operator is a block soft-thresholding [10]:

$$[\text{prox}_{\mu\|.\|_{12}}(X)]_i = \left(1 - \frac{\mu}{\|X_i\|_2}\right)_+ X_i \ . \tag{9}$$

Other potential choices for the penalty ϕ include the 3-level mixed norm ℓ_{212} proposed by [11]. Note that in practice, these penalties must depend on some weights w_s and z_u, that compensate the fact that the columns of G (corresponding to dipoles) and the lines of H are not normalized. To lighten the notation, we do not consider the weights in the whole paper.

Many works in the past five years have underlined the sub-optimality of ℓ_1 penalty for sparse recovery [12]. In practice, one observes that the ℓ_1 norm does not promote the sparsest solutions, and that it introduces a bias in the estimation, i.e. it tends to overpenalize high coefficients. This issue is worsened in the case of the ℓ_{12} penalty, since in this case the ℓ_2 norm of a source $\|X_i\|_2$ is lowered, which promotes "flat" sources with constant amplitudes over time. To circumvent this drawback, we considered non-convex penalties and their associated proximal operators, such as the empirical Wiener shrinkage (EW) [13], also called nonnegative garrote in statistics. We obtained particularly good results with a variant of this operator, called structured EW (SEW), and defined by:

$$[S_\mu(X)]_i = \left(1 - \left(\frac{\mu}{\|X_i\|_2}\right)^2\right)_+ X_i \ . \tag{10}$$

3.3. Choice of the parameters

This section gives some heuristic to choose appropriate parameters λ and α, which are of course crucial for solving the inverse problem with high accuracy. We consider the case of ℓ_1 regularization, but the results hold for other choices of penalty function ϕ.

Parameter α balances the relative weight of both modalities: EEG and fMRI. To choose a suitable value for α, it seems natural to evaluate the confidence we have in each modality, thus to make α dependent on the noise levels of both measures. Assuming white Gaussian noise with variance σ_E^2 for EEG and σ_F^2 for fMRI, we propose to select α so as to equal the two data-fidelity terms in the functional:

$$\alpha^* = \frac{U\sigma_F^2}{T\sigma_E^2 + U\sigma_F^2} \ . \tag{11}$$

Choosing the regularization parameter λ is a common (yet hard) issue in linear inverse problems. We can however easily compute an upper bound on λ, from which the minimization (4) leads to the null solution. Suppose that we can estimate the penalty term of the "true" solution $\phi(X^*)$. If we have

$$\lambda < \lambda_M := \frac{\frac{\alpha}{2}\|E\|_F^2 + \frac{1-\alpha}{2}\|F\|_F^2}{\phi(X^*)} \ , \tag{12}$$
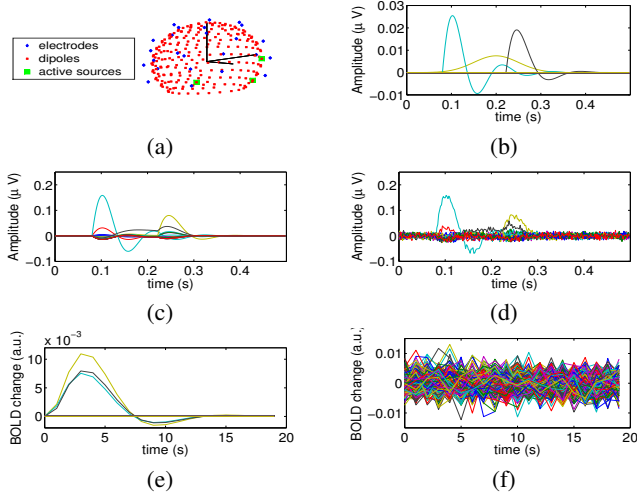
then the null solution will be preferred to the "true" one when minimizing (4). Conversely, a too low λ leads to unstable and non-sparse solutions. In practice, we obtain good results by setting $\lambda^* = 0.05\lambda_M$.

4. NUMERICAL EXPERIMENTS

4.1. Data and experiments

We use simulated data on a simple spherical 3-layer model. The sampling frequencies are set to 500 Hz for the EEG and 1 Hz for the fMRI. We consider 31 electrodes and 272 electric dipoles on the sphere representing the cortex. We generate three artificial electrical sources with damped waves, and generate the EEG and fMRI measures according to our linear model. A strong white Gaussian noise is then added to the measures, with input SNR of 2 dB for EEG and -18 dB for fMRI. The positions and time courses of the sources are depicted on Fig. 1, together with the ideal and noisy measures for both EEG and fMRI.
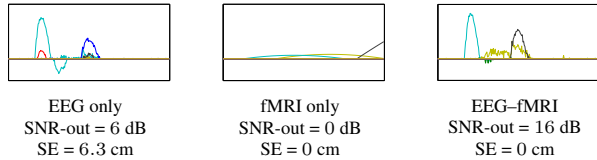
Let us first illustrate qualitatively the interest of using both EEG and fMRI on these simulated data. Fig. 2 shows the reconstructed sources, together with the corresponding reconstructed measurements, for EEG alone, fMRI alone, and the coupling. We always take a ℓ_1 penalty, and chose parameters α and λ so as to get the "best" results. For each test, we display the reconstruction error in terms of output SNR (Signal to noise ratio): if x is the true solution and $\hat{x}$ the estimated one, it writes SNR-out $= -20\log(\|x - \hat{x}\| / \|x\|)$. We also

Fig. 1. Location of the active sources (a), time course of the "true" sources (b), ideal (c) and noisy (d) EEG measurements, ideal (e) and noisy (f) fMRI measurements. The different colors correspond to different sources or sensors.

display the spatial error (SE), which consists of the sum of the Euclidean distances between the true sources and the estimated ones.

We observe that the EEG finds only one of the active sources, which leads to a quite poor spatial and temporal reconstruction. At the opposite, the fMRI–only solution finds the right source locations, but without any relevant temporal information. The combined solution we propose is satisfactory, since it manages to estimate the right source locations while giving a good temporal reconstruction.
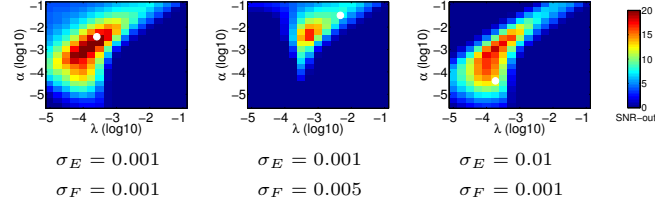

EEG only	fMRI only	EEG–fMRI
SNR-out = 6 dB	SNR-out = 0 dB	SNR-out = 16 dB
SE = 6.3 cm	SE = 0 cm	SE = 0 cm

Fig. 2. Time courses of the reconstructed sources, the different colors corresponding to different dipoles. Comparison between the EEG–only (left), fMRI–only (center) and combined (right) reconstructions.

4.2. Influence of the parameters

In order to investigate the sensitivity of the reconstruction to the parameters, as well as the relevance of values (α^*, λ^*), we display in Fig. 3 the accuracy of reconstruction as a function of the noise levels σ_E and σ_F. The SNR-out is color-coded, with highest values in red, and the value of (α^*, λ^*) (Sect. 3.3) is displayed as a white disc. We remark that our estimate (α^*, λ^*) gives a good order of magnitude, but does not exactly match the optimal parameters (i.e., the parameters achieving
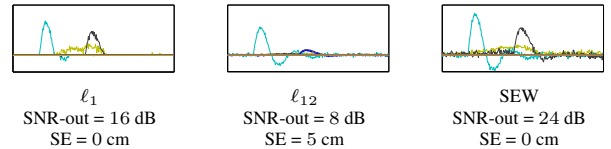
the highest SNR-out). We further observe that the area with high SNR-out are ellipsoidal, indicating a linear relationship between optimal λ and α which was enlighten in equation (12).

$\sigma_E = 0.001$	$\sigma_E = 0.001$	$\sigma_E = 0.01$
$\sigma_F = 0.001$	$\sigma_F = 0.005$	$\sigma_F = 0.001$

Fig. 3. Sensitivity to the parameters for different noise levels. The SNR-out is color-coded as a function of parameters (λ, α). The position of (λ^*, α^*) is displayed as a white disc.

4.3. Influence of the penalty term

Let us now investigate the influence of the penalty function ϕ. Figure 4 shows the reconstructed time-course of the dipoles for three different penalties: the standard ℓ_1 and ℓ_{12} norms, and the structured empirical Wiener shrinkage. Interestingly, the ℓ_{12} norm does not provide accurate results: although it estimates the right source locations, it tends to flatten their amplitude, which gives a poor SNR-out. The nonconvex SEW shrinkage seems to circumvent this problem, improving the reconstruction compared to the ℓ_1 penalty.
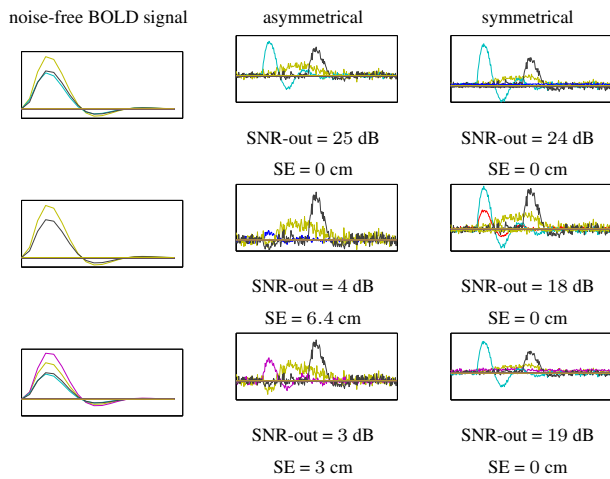

ℓ_1	ℓ_{12}	SEW
SNR-out = 16 dB	SNR-out = 8 dB	SNR-out = 24 dB
SE = 0 cm	SE = 5 cm	SE = 0 cm

Fig. 4. Influence of the penalty on the reconstruction. From left to right: ℓ_1, ℓ_{12} and SEW penalties.

4.4. Robustness to false positives or negatives

One of the major issues in EEG–fMRI is the errors in the coupling models, and more precisely the possible mismatches between EEG activity and BOLD. The major expected strength of direct symmetrical coupling is that such mismatches should not perturb too much the reconstruction. We investigate here this phenomenon, by artificially adding a false positive or a false negative activation in the simulated BOLD signal F. In the former, an artificial HRF-like response is added to F (equation (2)), that does not correspond to any EEG source, while in the latter the BOLD corresponding to one EEG source is artificially removed (i.e., one line of F is set to 0). Both noise-free BOLD signals are depicted on Fig. 5, together with the corresponding reconstructions. We compare the EEG–fMRI solution using the SEW, with an asymmetric

reconstruction, consisting in computing the fMRI–only solution, selecting the three main sources, and then inverting the EEG with only those sources. The results clearly show

Fig. 5. Influence of false positive or negative BOLD activations. The first row shows the reconstruction with full EEG–fMRI coupling, while second and third rows consider the case of BOLD false negative and positive activations, respectively. From left to right: noise-free simulated BOLD signal, asymmetrical, and symmetrical reconstructions.

that asymmetric reconstruction is very sensitive to false positives, since the corresponding source is automatically added in the reconstruction, the same effect being observed in case of a false negative. Instead, our symmetrical reconstruction appears to be quite stable.

5. CONCLUSION

This paper introduced a new symmetric coupling for EEG–fMRI imaging, based on recent developments in sparse regularization for inverse problems. The neuronal activity is reconstructed from both modalities symmetrically, through an efficient and convergent algorithm.

The main limitation of the approach is the assumption of a convolutive linear coupling between the electrical neuronal activity and the BOLD signal. Although this is quite restrictive, recent literature suggests that a linear model holds for a wide range of neuronal events. For potential application, this drawback can be circumvented by calibrating the temporal filter (our operator H) according to the subject and the task.

REFERENCES

[1] H. Becker, L. Albera, P. Comon, R. Gribonval, F. Wendling, I. Merlet, et al., "A performance study of various brain source imaging approaches," *Proc. ICASSP'2014*, pp. 5910–5914, 2014.

[2] A.K. Liu, J.W. Belliveau, and A.M. Dale, "Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte carlo simulations," *Proceedings of the National Academy of Sciences*, vol. 95, no. 15, pp. 8945–8950, 1998.

[3] L. Lemieux et al., "Event-related fMRI with simultaneous and continuous EEG: description of the method and initial case report," *Neuroimage*, vol. 14, no. 3, pp. 780–787, 2001.

[4] N.J. Trujillo-Barreto, E. Martinez-Montes, L. Melie-Garcia, and P. Valdes-Sosa, "A symmetrical Bayesian model for fMRI and EEG/MEG neuroimage fusion," *Int. J. Bioelectromagn*, vol. 3, no. 1, pp. 1998–2000, 2001.

[5] T. Deneux and O. Faugeras, "EEG-fMRI fusion of paradigm-free activity using Kalman filtering," *Neural computation*, vol. 22, no. 4, pp. 906–948, 2010.

[6] W. Ou, M. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, 2009.

[7] K.J. Friston, A. Mechelli, R. Turner, and C.J. Price, "Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics," *NeuroImage*, vol. 12, no. 4, pp. 466–477, 2000.

[8] P.A. et al. Robinson, "Bold responses to stimuli: dependence on frequency, stimulus form, amplitude, and repetition rate," *Neuroimage*, vol. 31, no. 2, pp. 585–599, 2006.

[9] P.L. Combettes and V.R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[10] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Foundations and Trends in Machine Learning*, vol. 4, no. 1, pp. 1–106, 2012.

[11] A. Gramfort, M. Kowalski, and M. Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in medicine and biology*, vol. 57, no. 7, pp. 1937, 2012.

[12] E.J. Candès, M.B. Wakin, and S.P. Boyd, "Enhancing sparsity by reweighted ℓ1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[13] K. Siedenburg and M. Dörfler, "Persistent time-frequency shrinkage for audio denoising," *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 29–38, 2013.