



# SEEG dipole source localization based on an empirical Bayesian approach taking into account forward model uncertainties

Steven Le Cam, Radu Ranta, Vairis Caune, Gundars Korats, Laurent Koessler, Louis Maillard, Valérie Louis-Dorr

## ► To cite this version:

Steven Le Cam, Radu Ranta, Vairis Caune, Gundars Korats, Laurent Koessler, et al.. SEEG dipole source localization based on an empirical Bayesian approach taking into account forward model uncertainties. *NeuroImage*, 2017, 153, pp.1-15. 10.1016/j.neuroimage.2017.03.030 . hal-01505700

**HAL Id: hal-01505700**

**<https://hal.science/hal-01505700>**

Submitted on 11 Apr 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SEEG dipole source localization based on an empirical Bayesian approach taking into account forward model uncertainties

S. Le Cam<sup>a,b,\*</sup>, R. Ranta<sup>a,b</sup>, V. Caune<sup>a,b</sup>, G. Korats<sup>a,b,c</sup>, L. Koessler<sup>a,b</sup>, L. Maillard<sup>a,b,d</sup>,  
V. Louis-Dorr<sup>a,b</sup>

<sup>a</sup>Université de Lorraine, CRAN, UMR 7039, 54500 Vandoeuvre-lès-Nancy, France

<sup>b</sup>CNRS, CRAN, UMR 7039, France

<sup>c</sup>Ventspils University College, 101 Inženieru iela, LV-3601, Ventspils, Latvia

<sup>d</sup>CHU Nancy, Neurology Department, 54000 Nancy, France

---

## Abstract

Electromagnetic brain source localization consists in the inversion of a forward model based on a limited number of potential measurements. A wide range of methods has been developed to regularize this severely ill-posed problem and to reduce the solution space, imposing spatial smoothness, anatomical constraint or sparsity of the activated source map. This last criteria, based on physiological assumptions stating that in some particular events (*e.g.*, epileptic spikes, evoked potential) few focal area of the brain are simultaneously actives, has gained more and more interest. Bayesian approaches have the ability to provide sparse solutions under adequate parametrization, and bring a convenient framework for the introduction of priors in the form of probabilistic density functions. However the quality of the forward model is rarely questioned while this parameter has undoubtedly a great influence on the solution. Its construction suffers from numerous approximation and uncertainties, even when using realistic numerical models. In addition, it often encodes a coarse sampling of the continuous solution space due to the computational burden its inversion implies. In this work we propose an empirical Bayesian approach to take into account the uncertainties of the forward model by allowing constrained variations around a prior physical model, in the particular context of SEEG measurements. We demonstrate on simulations that the method enhance the accuracy of the source time-course estimation as well as the sparsity of the resulting source map. Results on real signals prove the applicability of the method in real contexts.

**Keywords:** Electrical source imaging; Stereo-electroencephalography (SEEG); Inverse problem; Variational Bayes; Intracerebral electrical Stimulation (ICS); Interictal spikes

---

---

\*Corresponding author

Email address: [steven.le-cam@univ-lorraine.fr](mailto:steven.le-cam@univ-lorraine.fr) (S. Le Cam)

"NOTICE: this is the author's version of a work that was accepted for publication in NeuroImage. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in NeuroImage, DOI: 10.1016/j.neuroimage.2017.03.030"

## 1. Introduction

Electromagnetic brain source localization consists in reconstructing the underlying cerebral activity from the electroencephalographic measurements. The recorded electromagnetic field is generated by synchronized activities of local and compact sets of neuronal cells, each set being considered as an equivalent dipolar source at the macroscopic level. The number of such equivalent sources is far higher than the number of recordings, making the reconstruction problem severely ill-posed. While the source localization problem has been extensively studied from non-invasive measurements, we have been recently working on solving this problem from intracranial measurements, based on Stereo-EEG (SEEG) setup (Caune et al. (2014, 2013); Le Cam et al. (2014)). SEEG consists of several electrode shafts (containing each up to eighteen aligned recording contacts) inserted in the brain volume within the structure of interests. These recordings are often considered as direct measurements of local field potential in the direct neighborhood of the contacts. However the recorded potentials are still produced by the mixing of several sources, including the projection of far sources of strong amplitudes. Thus a localization procedure is still of interest, while very few works have been developed on the subject (Chang et al. (2005); Yvert et al. (2005)). In comparison to surface EEG, the SEEG modality offers high Signal to Noise Ratio (SNR), because the sensors are closer to the generators and the signal do not have to propagate through the skull barrier. More precise identification of deep activated structures and reconstruction of their time-courses is expected, facilitating the subsequent functional or effective connectivity analysis of the underlying brain network, usually carried out from non-invasive EEG/MEG setup (David et al. (2006); Coito et al. (2015)). In a recent study (Caune et al. (2014)), we have thoroughly questioned the feasibility of a source localization from these intracranial measurements. In particular, we emphasized some important guidelines for confident electrical source localization in this particular context of aligned contacts on electrode shafts implanted in clusters.

While the temporal resolution of the (S)EEG modality is high, its spatial resolution is poor, and the set of activated sources cannot be determined based on the measurements alone (Nunez and Srinivasan (2006); Baillet et al. (2001)). A wide range of (physiological) constraints/priors can be imposed to regularize the problem, resulting in an abundant literature on this topic. Electrical Source Imaging (ESI) approaches have met high interests (Baillet et al. (2001); Michel et al. (2004)), offering visually interpretable results and direct comparison with imaging modalities (fMRI, MRI, PET, etc...). They are based on distributed source models, in which a high number of dipoles are placed on a regular grid in the brain volume (possibly restricted to particular area, usually to the cortical mantle). Under the commonly accepted hypothesis

of purely resistive propagation media, the inversion becomes a linear but severely ill-posed problem. Most of the methods mainly differs in the spatial/temporal constraints that are imposed on the underlying distribution of sources, imposing spatial smoothness of the source map (Loreta or sLoreta (Pascual-Marqui (2007))), or enforcing focal solutions (FOCUSS) (Gorodnitsky and Rao (1997)). In particular, the Bayesian framework has been extensively studied in this context (Baillet et al. (2001); Phillips et al. (2005); Trujillo-Barreto et al. (2008); Wipf and Nagarajan (2009); Lucka et al. (2012)), where hyper-parameters can be introduced to control the optimization of the model parameters.

In each of the mentioned studies, the electromagnetic projections gains mapping the sources to the electrode contacts (*i.e.*, the lead-field matrix) are supposed to be known. While the efforts have been focused on the localization of the sources and on the estimation of their time-courses, the accuracy of the propagation model should be also questioned, as it highly impacts the precision of the resulting source reconstruction. The computation of sophisticated propagation models based on heavy numerical approaches (Finite Element Models (FEM)) has been widely explored and promoted in the last few years. In this purpose, a precise segmentation of the brain areas (scalp, skull, white and gray matters) is required, and for each of these areas the conductivities are to be estimated (Acar et al. (2016)). We can mention also the consideration of anisotropy, mostly within the white matter where the directions of the fibers is considered to influence the electric propagation (Wolters et al. (2006); Bangerter et al. (2010)). These models rely on a large amount of parameters and pre-processing steps, each of them bearing their own uncertainties. The utility of such complex model is currently a challenging topic as some discordant studies are appearing, stating that fitted sphere models yield similar source localization performance than more elaborated but still uncertain numerical models (Biro et al. (2014); Caune et al. (2014)). In addition, the solution space is continuous while the lead-field used in the distributed model encodes the projections for a finite number of dipole positions, so it cannot capture all the possible projections for every single location in the brain volume, leading to a basis mismatch problem. Chi et al. (2011) demonstrate that the quality of the decomposition of a sparse physical field is considerably degraded in the presence of basis mismatch (a high number of basis elements are recruited to reconstruct the data), even when the assumed basis corresponds to a fine-grained sampling of the parameter space. In our particular application, it means that keeping the lead-field fixed can results in scattered solutions spread over a high number of dipolar elements.

Very few notable works have explored the opportunity to estimate simultaneously the source localization and the forward model. Hansen et al. (2016) exploit a corpus of forward models with varying conductivity parameters built from the structural scans of 16 subjects. Low-resolution anatomic information are extracted from this corpus through a Principal Component Analysis, and are used as *priors* within a Bayesian approach to produce estimate of the sources from the EEG measurements of a new participant (for which no anatomical information are available). Acar et al. (2016) directly optimize the conductivities of a numerical forward model computed on the MRI of the patient. The set of sources, assumed to be focal, are identified as independent dipolar-like scalp maps based on an Independent Component Analysis, and the estimation of the sources and of the conductivity parameters are iterated until convergence.

Unlike these two last studies, we do not tackle the brain tissue conductivities estimation, we rather propose an optimization of the lead-field coefficients themselves within a Variational Bayes (VB) framework. The proposed method is to be directly compared with the SOFOMORE approach (Stahlhut et al. (2011)), where the source and the lead-field coefficients are simultaneously estimated within a hierarchical Bayesian framework. Independent univariate Gaussian priors of equal variance are attributed to each lead-field coefficient within a column, and conjugate prior distributions (gamma distributions) are assigned to these priors. The designing of the prior densities is, by their own admission, not properly carried out in their paper. In particular, no physiological constraint taking account of the source to contacts distance and of the sensors placement are embedded, and the plausibility of the posterior estimates are not explicitly addressed.

The key difference with this previous work is the introduction of multivariate Gaussian likelihoods over each column of the lead-field (*i.e.*, the forward field of one dipole on the lead-field grid), with non-diagonal precision matrix specifically designed for each column. These precision matrices explicitly encode the dependencies between the projection gains of each single forward field. This is of particular interest in the context of the SEEG recording setup, where aligned contacts are implanted in clusters, yielding high dependencies between the projection gains of a given source on the sensors. To be more specific, a gain variation of a source on a given sensor, either due to a variation of the source position/orientation, or to a forward model parameter error, will induce a correlated gain variation on its neighboring sensors. We then constrain the posterior means to remain close to a confident physical model prior, preventing from non-physiological posterior estimates. In contrast to Stahlhut et al. (2011), we give a particular attention to the designing of the lead-field column precision priors. We provide a simple and fast method for their computation, considering both the uncertainties due to the sampling of the source space and to the forward model parameter errors. Finally, a post-optimization strategy is proposed to map the obtained lead-field posterior estimates within the source space.

The Gaussian modeling of the model errors, as imperfect as it is, reflects the numerous uncertainties related to the resolution of the forward problem. Besides, such priors introduce additional degree of freedom for the Bayesian algorithms to express their sparse inclinations. Indeed, under adequate parametrization, Bayesian optimization strategies are known to produce sparse solutions (e.g. Sparse Bayesian Learning (SBL), restricted Maximum Likelihood (reML), Automatic Relevance Determination (ARD) (Tipping (2001))). This aspect is observed in Stahlhut et al. (2011), and is confirmed by our experiments. While sparse priors might not be desirable for all application, this is relevant when trying to localize a limited number of focal sources, such as inter-ictal epileptic events or in the case of controlled experiments implying localized evoked potentials (*e.g.*, auditory evoked potential (Kiebel et al. (2008))), which can be summarized to the activation of few brain structures.

In the section 2.1, we present the data model and the probabilistic Bayesian framework. In section 2.2, the Variational Bayesian (VB) scheme is used for maximizing the log-likelihood of the model, simultaneously providing estimates of the sources and of the columns of the lead-field. In this section we also explain how the precision matrices over the lead-field columns can be computed, and we expose the Equivalent Current

Dipole (ECD) fitting to the posterior lead-field columns. In the evaluation section 3, we provide simulation results for different scenarios of underlying dipole configurations and noise. In particular, a FEM is used to produce the simulated data, while the inversion is based on a coarse One Sphere Model (OSM), simulating the uncertainties of the lead-field matrix both in term of source space sampling and of forward model imprecision. Finally, we demonstrate the reliability of the localization results in the context of real SEEG recordings of intracranial stimulations and interictal spikes.

## 2. Methods

In this section we present the general Bayesian formulation of the source localization problem, we specify the parameters of the model and we briefly remind the classical approach for maximizing the data likelihood of the model. We then introduce a probabilistic modeling of the lead-field columns, and we explicit their optimizations through the VB scheme. In the paper are given the main theoretical aspects of the method as well as the analytic expressions for the model parameter updates, more detailed developments of the equations being given in the appendix.

### 2.1. Bayesian modeling

It is widely assumed that a patch of synchronized neuronal cells can be approximated by a single equivalent current dipole positioned within this region. The potential recorded on a contact is then modeled as a sum of linear and instantaneous projection of a high number  $N_s$  of dipolar activities covering the brain volume (commonly restricted to the cortical columns). Let consider  $\mathbf{X} \in \mathbb{R}^{N_c \times T}$  the  $N_c$  measurements of length  $T$  due to  $N_s$  sources  $\mathbf{S} \in \mathbb{R}^{N_s \times T}$ :

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \boldsymbol{\varepsilon} \quad (1)$$

where  $\mathbf{S}$  encodes the amplitudes of the sources, the columns of  $\mathbf{A} \in \mathbb{R}^{N_c \times N_s}$  represent the projection coefficients from the  $N_s$  fixed dipoles to the  $N_c$  contacts (thus encoding the dipole positions and orientations), and  $\boldsymbol{\varepsilon}$  is an additive noise. Within a Bayesian framework, both  $\mathbf{S}$  and  $\boldsymbol{\varepsilon}$  are usually considered as random variables, generally assumed Gaussian. More specifically, general Gaussian scale mixtures are proposed for  $\mathbf{S}$  with arbitrary covariance components corresponding to given *priors* on the dependencies between sources. In this paper we will make the assumptions that the sources are independent, without loss of generality (Wipf and Nagarajan (2009)). Let introduce the diagonal precision matrix (*i.e.*, the inverse of the covariance matrix)  $\Gamma_{\mathbf{S}} = \text{diag}(\gamma_{\mathbf{S}})$ , with  $\gamma_{\mathbf{S}} = [\gamma_{s_1}, \dots, \gamma_{s_{N_s}}]$  and such that  $p(\mathbf{S}|\Gamma_{\mathbf{S}}) = \mathcal{N}(\mathbf{0}, \Gamma_{\mathbf{S}}^{-1})$ . Let assume  $\boldsymbol{\varepsilon}$  as a spatially and temporally homogeneous white Gaussian noise. Its diagonal precision matrix is written  $\Gamma_{\boldsymbol{\varepsilon}} = \gamma_{\boldsymbol{\varepsilon}} I_{N_c}$  ( $I_N$  being the  $N \times N$  identity matrix), and the data likelihood then writes:

$$p(\mathbf{X}|\mathbf{S}, \boldsymbol{\theta}) \propto \exp -\frac{\gamma_{\boldsymbol{\varepsilon}}}{2} (\mathbf{X} - \mathbf{A}\mathbf{S})^\top (\mathbf{X} - \mathbf{A}\mathbf{S}) \quad (2)$$

with  $\top$  the matrix transposition. The joint distribution of  $\mathbf{X}$  and  $\mathbf{S}$  with respect to the set of parameters  $\boldsymbol{\theta} = \{\Gamma_{\mathbf{S}}, \gamma_{\boldsymbol{\varepsilon}}\}$  writes:

$$p(\mathbf{X}, \mathbf{S} | \theta) = p(\mathbf{X} | \mathbf{S}, \theta) p(\mathbf{S} | \Gamma_{\mathbf{S}}) \quad (3)$$

Under this parametrization and considering that  $\Gamma_{\mathbf{S}}$  and  $\gamma_{\epsilon}$  are fixed and known, the posterior probability  $p(\mathbf{S} | \mathbf{X}, \theta)$  is also Gaussian with precision and mean given by:

$$\hat{\Gamma}_{\mathbf{S}} = \Gamma_{\mathbf{S}} + \gamma_{\epsilon} \mathbf{A}^{\top} \mathbf{A} \quad (4)$$

$$\hat{\mathbf{S}} = \gamma_{\epsilon} \hat{\Gamma}_{\mathbf{S}}^{-1} \mathbf{A}^{\top} \mathbf{X} \quad (5)$$

and in this case the Maximum a Posteriori (MAP) solution is given by the source posterior mean  $\hat{\mathbf{S}}$ . When  $\Gamma_{\mathbf{S}}$  is diagonal, this solution is equivalent to the  $\ell_2$ -Weighted minimum norm solution (Dale and Sereno (1993)).

If  $\Gamma_{\mathbf{S}}$  and  $\Gamma_{\epsilon}$  are unknown, these parameters can be estimated through an Expectation-Maximization (EM) procedure maximizing the marginal likelihood of the model (*i.e.*, maximizing eq.(3) integrated over  $\mathbf{S}$ , the *hidden data*) (Bishop et al. (2006)). This algorithm is sometimes called Sparse Bayesian Learning (SBL) (Sato et al. (2004); Ramírez et al. (2006); Wipf and Rao (2007)), because it naturally provides sparse solutions by shrinking to zero the variance of irrelevant sources (Tipping (2001); Friston et al. (2008)). The initial set of source candidates can be allowed to be very large, and we can rely on the optimization procedure to prune the superfluous elements and select the most appropriate set of sources.

## 2.2. Simultaneous optimization of $\mathbf{A}$ and $\mathbf{S}$

On one hand, the simultaneous reconstruction of both the sources and the projection coefficients makes the EEG source reconstruction even more ill-posed, the solution space being broaden. On the other hand, the construction of the forward model involves lots of hypothesis, model reductions and uncertainties. Moreover the lead-field matrix  $\mathbf{A}$  represents a sampling of the source space, and there is few chances that the true source positions (in term of equivalent dipole) will exactly match those of the  $\mathbf{A}$  grid. By introducing precision parameters over the forward matrix  $\mathbf{A}$ , we allow variations of its columns, and thus let the model be data-driven, *i.e.*, adjust  $\mathbf{A}$  to the true underlying propagation parameters of the medium. In the following, the columns of  $\mathbf{A}$  will then be considered as latent variables and we will propose a Variational Bayesian (VB) approach to optimize the evidence of the model. Benefiting from the sparse property of the Bayesian approaches, we look for solutions that are not spread over a high number of fixed dipoles, but rather focused on a small number of dipoles with optimized lead-field projections.

### 2.2.1. Forward model likelihood

In the distributed approach, a lead-field is generated based on a given forward model, which can vary from simple analytical models (Infinite Homogeneous Model (IHM) or spherical models) to complex numerical model based on Boundary Element Modeling (BEM) or Finite Element modeling (FEM). Each column of the lead-field encodes the projection of a given elementary dipole within the brain volume. The positions of these dipoles are often assumed to be restricted to the cortical mantle, with orientations orthogonal to the cortical mantle surface. However in this work, we



do not impose *prior* anatomical constraint for the source localization.  $N_p$  positions (*i.e.*,  $N_s = 3 * N_p$  sources) are defined on a regular grid with a spatial resolution of  $\delta_p$  over the three spatial dimensions within the brain volume. The step of the grid  $\delta_p$  (usually of the order of few millimeters) then defines the number of considered dipole positions  $N_p$  in the lead-field, usually up to several thousands. This model-based lead-field matrix, denoted  $\mathbf{A}_0$  in the following, will be now considered as a *prior* and we will let the algorithm optimize the values of  $\mathbf{A}$ . We assume that the columns of  $\mathbf{A}$  follow independent Gaussian distributions around the prior model  $\mathbf{A}_0$ :

$$p(\mathbf{A}|\mathbf{A}_0, \Gamma_{\mathbf{A}}) = \prod_{j=1}^{N_s} \mathcal{N}(\mathbf{a}_j | \mathbf{a}_{0j}, \Gamma_{\mathbf{a}_j}^{-1}) \quad (6)$$

$\Gamma_{\mathbf{A}}$  contains  $N_s$  precision matrices  $\Gamma_{\mathbf{a}_j}$  of size  $N_c \times N_c$ , constraining the projection coefficients in each of the  $N_s$  columns  $\mathbf{a}_j$  to remain close to the prior model  $\mathbf{a}_{0j}$  with a given degree of freedom encoded in  $\Gamma_{\mathbf{a}_j}$ . These matrices can both describe the uncertainty on the position of the dipole as well as the uncertainty of the used forward model (tissue segmentation, conductivities...).

We have tested several empirical methods to shape these precision parameters, based on the computation of second order statistics of forward model clusters representing the variability of the projections for each dipole (*e.g.*, by sampling the surrounding source space projection of each dipole using an analytical model (IHM or OSM), by computing the projection of each dipolar elements using several forward models and/or by sampling the conductivity parameters in given ranges (Hansen et al. (2016))). Such exploratory methods turned out to be computationally heavy, does not guaranty that the source space is covered and neither prevent from high overlapping between neighboring dipolar elements.

We propose here a fast empirical method inspired from the subspace correlation metric used in the MUSIC-type algorithms. To take account of the model uncertainties, we start by computing  $M$  forward models  $\{\mathbf{A}_0^m, m \in \{1, \dots, M\}\}$  (*e.g.*, by sampling the model parameters in physiological ranges (Hansen et al. (2016))). The columns of the prior mean forward model  $\mathbf{A}_0$  are set as the mean of these  $M$  model columns:

$$\mathbf{a}_{i0} = 1/M * \sum_{m=1}^M \mathbf{a}_{i0}^m \quad (7)$$

From each forward model  $\mathbf{A}_0^m$ , and for each of the three orientations of a given position, we start by computing the closest (in term of correlation) dipole projections that can be coded within each of the surrounding three-dipole subspaces ( $K = 26$  if we consider all the grid points on the cube surrounding the current position). This can be done by a normalized back-projection of a regression of the current dipole projection  $\mathbf{a}_{i0}^m$  on each of the  $K$  3-D surrounding subspaces  $\mathbf{A}_{k0}^m \in \mathbb{R}^{N_c \times 3}$ :

$$\mathbf{a}_{ik}^m = \frac{\mathbf{A}_{k0}^m (\mathbf{A}_{k0}^{m+} \mathbf{a}_{i0}^m)}{\|\mathbf{A}_{k0}^{m+} \mathbf{a}_{i0}^m\|} \quad (8)$$

with  $\mathbf{A}_{k0}^{m+}$  the Moore-Penrose pseudo-inverse of the matrix  $\mathbf{A}_{k0}^m$ . The inverse covariance matrix of the resulting  $K \times M$  dipole projections gives the precision matrix  $\Gamma_{\mathbf{a}_i}$  for each column  $\mathbf{a}_i$  of  $\mathbf{A}$ :



$$\Gamma_{\mathbf{a}_i}^{-1} = \frac{1}{M * K - 1} \sum_{m=1}^M \sum_{k=1}^K (\mathbf{a}_{ik}^m - \mathbf{a}_{i0})(\mathbf{a}_{ik}^m - \mathbf{a}_{i0})^T \quad (9)$$

We insist on the importance of considering non-diagonal precision matrix, the non-null covariance imposing constraint on the re-estimation of the projection values, imposing them to remain close from the prior physiologically plausible projection pattern encoded in  $\mathbf{A}_0$ . This will be further illustrated in the evaluation section 3. Provided that the source space has been adequately sampled, this empirical procedure is expected to bring a blanket of the source space, while restricting the influence of each dipolar dictionary element to their respective vicinity.

### 2.2.2. Variational Bayesian optimization

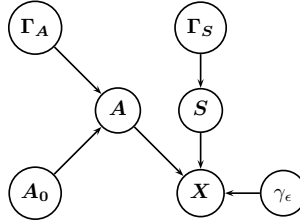


Figure 1: Hierarchical Model

The set of latent variables of the problem is now the couple  $\{\mathbf{S}, \mathbf{A}\}$ . The joint distribution of  $\mathbf{X}$ ,  $\mathbf{S}$  and  $\mathbf{A}$  writes:

$$p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) = p(\mathbf{X} | \mathbf{A}, \mathbf{S}, \theta) p(\mathbf{S} | \Gamma_S) p(\mathbf{A} | \mathbf{A}_0, \Gamma_A) \quad (10)$$

with  $\theta = \{\Gamma_S, \gamma_\epsilon, \mathbf{A}_0, \Gamma_A\}$  the set of model parameters, as illustrated fig 1. The optimization of this set through the EM procedure becomes intractable because of couplings between  $\mathbf{A}$  and  $\mathbf{S}$ . We then propose a VB optimization scheme by introducing a set of candidate distributions  $q(\cdot) \in \Omega$  such that  $q(\mathbf{S}, \mathbf{A}) = q(\mathbf{S})q(\mathbf{A})$ , with  $q(\mathbf{S}) = \prod_{i=1}^{N_s} q(\mathbf{s}_i)$  and  $q(\mathbf{A}) = \prod_{i=1}^{N_s} q(\mathbf{a}_i)$ . The VB optimization procedure fits  $q(\mathbf{A}, \mathbf{S})$  to the true posterior distribution  $p(\mathbf{A}, \mathbf{S} | \mathbf{X}, \theta)$  so that the Kullback-Leibler distance between these two distributions vanishes. This is equivalent to finding the distributions  $\{q(\mathbf{S}), q(\mathbf{A})\}$  maximizing the quantity  $\mathcal{F}(q) = \int_{\mathbf{A}, \mathbf{S}} q(\mathbf{A}, \mathbf{S}) \ln \frac{p(\mathbf{X}, \mathbf{A}, \mathbf{S})}{q(\mathbf{A}, \mathbf{S})}$ , called the free energy, which is a lower bound of the model evidence (Friston et al. (2008); Wipf and Nagarajan (2009)). This results in a version of the SBL algorithm for simultaneous estimation of  $\mathbf{A}$  and  $\mathbf{S}$ , which we will denote as VBLF (Variational Bayes for Lead-Field re-estimation).

Within the VB scheme, the candidate distributions are iteratively updated by evaluating the marginal expected values of the model joint probability (Bishop et al. (2006)):

$$\ln q(\mathbf{s}_i) \propto \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{A}) \prod_{j \neq i} q(\mathbf{s}_j)} \quad (11)$$

$$\ln q(\mathbf{a}_i) \propto \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}) \prod_{j \neq i} q(\mathbf{a}_j)} \quad (12)$$

Under the Gaussian prior parametrization adopted in this paper both for  $\mathbf{S}$  and  $\mathbf{A}$ , the approximated posteriors  $q(\mathbf{a}_j)$  of each column are Gaussians with precisions parameters  $\hat{\Gamma}_{\mathbf{a}_j}$  and means  $\hat{\mathbf{a}}_j$ , given as (see Appendix A):

$$\hat{\Gamma}_{\mathbf{a}_j} = \Gamma_{\mathbf{a}_j} + \gamma_\epsilon \hat{E}_{\mathbf{s}_j} \mathbf{I}_{N_c} \quad (13)$$

$$\hat{\mathbf{a}}_j = \hat{\Gamma}_{\mathbf{a}_j}^{-1} (\Gamma_{\mathbf{a}_j} \mathbf{a}_{0j} + \gamma_\epsilon \Lambda_j) \quad (14)$$

where we define:

$$\hat{E}_{\mathbf{s}_j} = (\hat{\mathbf{s}}_j \hat{\mathbf{s}}_j^\top + T \hat{\gamma}_{\mathbf{s}_j}^{-1}) \quad (15)$$

$$\Lambda_j = \mathbf{X} \hat{\mathbf{s}}_j^\top - \hat{\mathbf{A}}^{\setminus j} (\hat{\Gamma}_{\mathbf{s}_j}^{\setminus j} + \hat{\mathbf{S}}^{\setminus j} \hat{\mathbf{s}}_j^\top) \quad (16)$$

where, for any matrix  $\mathbf{M}$ , we write  $\mathbf{M}_{\cdot j}$  the  $j$ th column of  $\mathbf{M}$ ,  $\mathbf{M}^{\setminus j}$  the matrix  $\mathbf{M}$  minus its  $j$ th column and  $\mathbf{M}^{\setminus j}$  the matrix  $\mathbf{M}$  minus its  $j$ th line.  $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_{N_s}]$  is the VB updated version of the lead-field matrix  $\mathbf{A}$ .

The posterior estimates of  $\mathbf{S}$  must now includes the posterior covariances on the column of  $\mathbf{A}$ . We introduce the diagonal matrix  $\mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}} \in \mathbb{R}^{N_s \times N_s}$  containing the traces of the lead-field column posterior covariance matrices on its diagonal:  $\mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}} = \text{trace}(\hat{\Gamma}_{\mathbf{a}_i}^{-1})$ . The posterior mean and covariance of  $\mathbf{S}$  now write:

$$\hat{\Gamma}_{\mathbf{S}} = \Gamma_{\mathbf{S}} + \gamma_\epsilon (\hat{\mathbf{A}}^\top \hat{\mathbf{A}} + \mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}}) \quad (17)$$

$$\hat{\mathbf{S}} = \gamma_\epsilon \hat{\Gamma}_{\mathbf{S}}^{-1} \hat{\mathbf{A}}^\top \mathbf{X} \quad (18)$$

Updates of the source precisions  $\gamma_{s_i}$  and noise precision  $\gamma_\epsilon$  are obtained by a derivation of the expected marginal likelihood (with respect to  $q(\mathbf{A})$  and  $q(\mathbf{S})$ ) as:

$$\gamma_{s_i}^{-1} = \frac{\hat{E}_{\mathbf{s}_i}}{T} \quad (19)$$

$$\gamma_\epsilon^{-1} = \frac{1}{TN_c} \|\mathbf{X} - \hat{\mathbf{A}} \hat{\mathbf{S}}\|_F^2 + \frac{1}{N_c} \text{trace}(\hat{\mathbf{A}} \hat{\Gamma}_{\mathbf{S}}^{-1} \hat{\mathbf{A}}^\top) + \frac{1}{TN_c} \sum_{i=1}^{N_s} \hat{E}_{\mathbf{s}_i} \mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}}^{-1} \quad (20)$$

The algorithm consists in iteratively computing the equations (13) to (18) (the E-step) and (19) to (20) (the M-step) until convergence.

An important aspect of the variational approach is that it brings a quantitative way to carry out model selection through the free energy quantity  $\mathcal{F}$ . Considering  $\mathbf{A}$  as a latent variable, the free energy formulation becomes (see Appendix B):

$$\mathcal{F}(q) = \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}), q(\mathbf{A})} - \langle \ln q(\mathbf{S}) \rangle_{q(\mathbf{S})} - \langle \ln q(\mathbf{A}) \rangle_{q(\mathbf{A})} \quad (21)$$

with

$$\begin{aligned} \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}), q(\mathbf{A})} &= \langle \ln p(\mathbf{X} | \mathbf{S}, \mathbf{A}, \gamma_\epsilon) \rangle_{q(\mathbf{S}), q(\mathbf{A})} + \langle \ln p(\mathbf{S} | \Gamma_{\mathbf{S}}) \rangle_{q(\mathbf{S})} \\ &\quad + \langle \ln p(\mathbf{A} | \mathbf{A}_0, \Gamma_{\mathbf{A}}) \rangle_{q(\mathbf{A})} \end{aligned} \quad (22)$$

$\langle \cdot \rangle_q$  denoting the expectation with respect to the distribution  $q$ .

### 2.2.3. Posterior forward field mapping

The Bayesian algorithm described above provides us with forward fields  $\hat{\mathbf{a}}_i$  of dipoles  $\hat{\mathbf{d}}_i$  with time-courses  $\hat{\mathbf{s}}_i$ . However the positions and orientations of these Equivalent Current Dipoles (ECD)  $\hat{\mathbf{d}}_i$  remain unknown. Indeed our method does not restrict these sources to the prior lead-field grid, and a post-optimization procedure is needed to fulfill our objective of mapping them in the brain volume. In addition, the dipole  $\hat{\mathbf{d}}_i$  is not necessarily unitary since  $\mathbf{A}$  and  $\mathbf{S}$  are simultaneously optimized and no such constrain is imposed in the model. Hence it is important to note that  $\hat{\mathbf{A}}$  cannot be considered as a proper lead-field matrix in its most general acceptation<sup>1</sup>; the same applies for the time-course  $\hat{\mathbf{s}}_i$  which indeed represent the actual current density of the dipolar source  $\hat{\mathbf{d}}_i$  up to a scalar factor. The identification of these parameters (position, orientation and scalar factor) can be solved by an Equivalent Current Dipole procedure by fitting  $\hat{\mathbf{a}}_i$  to a confident propagation model (Scherg et al. (1999); Caune et al. (2014)). The localization accuracy will still be subject to the inherent uncertainties of the used model, however this ECD fitting step maps the sources in the continuous solution space and the solutions are not constrained to the lead-field grid. The constant factor scaling the chosen post-fitting propagation model to the posterior lead-field  $\hat{\mathbf{a}}_i$  solves the scalar ambiguity between the true source amplitude and the corresponding unitary-dipole projection column.

In practice for each posterior forward field  $\hat{\mathbf{a}}_i$ , the optimization algorithm (Sequential Quadratic Programming) is initialized with the parameters of the prior ECD associated to  $\mathbf{a}_{0i}$ , the new ECD being constrained to remain close to its prior version through the model-driven prior precision matrix  $\Gamma_{\mathbf{a}_i}$ . The outcomes of this procedure are the parameters of the fitted dipole  $\hat{\mathbf{d}}_i$  (position, orientation and scaling factor). Theoretically, this ECD assignment step has to be carried out for each of the  $N_s$  columns of  $\mathbf{A}$ . However such procedure would be neither computationally tractable neither relevant for silent or low amplitude sources (for which the posterior forward field might be unreliable due to the relative low SNR). Thus it will be applied only to the posterior columns  $\{\hat{\mathbf{a}}_i\}$  associated to the set of active sources, *i.e.*, those associated to the most powerful time-courses  $\{\hat{\mathbf{s}}_i\}$  (as defined in the evaluation section 3), resulting in a limited set of identified dipoles  $\hat{\mathbf{d}}$ . An illustration of the global localization procedure including this post-optimization step is given fig.2.

## 3. Evaluation

We first validate the method on simulations and we compare the performance of our algorithm with several algorithms of the literature (SBL, SOFOMORE, Champagne and RAP-MUSIC). We then evaluate the methods when facing real data, first in the context of intracranial stimulation (ICS) and finally for the localization of interictal spikes event.

---

<sup>1</sup>*i.e.*, representing the projection of unitary dipoles on the electrodes

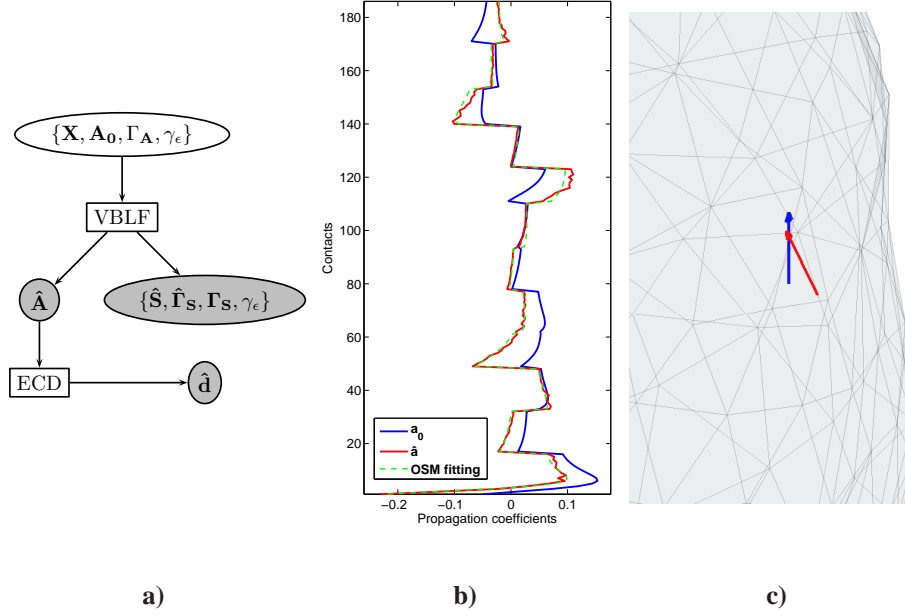


Figure 2: a) Work-flow of the method including the ECD identification step, with  $\hat{\mathbf{d}}$  the set of identified dipoles. b) Example of a posterior lead-field column mean  $\hat{\mathbf{a}}$  (red) superimposed with its OSM fitting (dashed green) and its prior version  $\mathbf{a}_0$  (blue). c) Prior dipole (blue) and fitted dipole (red), 3.1mm distant one from the other in this case (with a prior lead-field resolution of 1cm).

### 3.1. Simulations

We use the adult MNI-ICBM152 averaged model (Fonov et al. (2011)) which is freely available for download on the web<sup>2</sup>. The forward model is computed using a five compartment Finite Element Model (FEM) for 509 source positions regularly sampling the gray matter volume. For more details on how the FEM has been computed, please refer to Caune et al. (2014).

To simulate the uncertainties of the forward model, we use One Sphere Models (OSM (Yao (2000))) to solve the inverse problem and build the lead-field prior  $\mathbf{A}_0$  (with conductivities of 0.33 S/m).  $M = 17$  different OSM are computed using locally fitted spheres at 17 points<sup>3</sup> of the upper skull mesh, for which closed-form are available (Zhang (1995)). For each of these 17 models, the lead-field is computed for a regular grid of freely oriented dipoles with 1cm resolution, situated within a 2cm envelope of the gray matter mesh and with no overlap with the 509 previously sampled positions used from the forward FEM computation (in all,  $N_p = 2214$  possible dipole positions). The prior matrix  $\mathbf{A}_0$  and the associated precision matrices  $\{\Gamma_{\mathbf{a}_i}\}$  are then

<sup>2</sup><http://www.ucl.ac.uk/medphys/research/borl/intro/headmodels/adult>

<sup>3</sup>One on top of the head, 8 around the horizontal plane passing through the nasion, and one at mid-distance of each surface line linking these 8 previous points to the top one

computed using the procedure described section 2.2.1.

We have simulated a realistic SEEG implantation with the help of the neuro-physiologists of the Nancy CHU. The number of electrode shafts equals 12, on which 186 contacts are spread (about 14 to 18 per electrodes). Our prior OSM lead-field  $\mathbf{A}_0$  encodes the projections in the three directions for each source point, and is then of size  $N_c \times 3 \times N_p = 186 \times 6642$ .

Either three or five equivalent dipoles with random positions and orientations are selected among the set of 509 dipoles. We affect real time-courses to these seeded dipoles, taken from the SEEG recordings of an epileptic patient on bipolar montage (difference of two consecutive contacts of a same electrode shaft, assumed to catch the activity of a local source). A set of 200 windows with time-length  $256ms$  (128 samples) have been selected, corresponding to background activities, inter-epileptic events, as well as various representative periods of the ictal events (from spike bursting to paroxysmic rhythmic activities). Additive spatially independent white noise is simulated on the contacts, with standard deviations taken as:

$$\sigma_e = \frac{\|\mathbf{X}_s\|_F}{T * N_c * n_s * 10^{L_{SNR}/20}} \quad (23)$$

with  $\mathbf{X}_s$  the potential due to the  $n_s$  simulated sources, and  $L_{SNR}$  the desired SNR in dB which varies in the set  $\{10, 5, 0, -5\}$  in our simulations. The division by  $n_s$  ensure that the SNR for each dipole respectively is comparable on average from a simulation to another across the two source configurations  $n_s = \{3, 5\}$ .

We evaluate the benefit of re-estimating the lead-field  $\mathbf{A}$  by comparing the localization results based on the proposed VBLF approach (section 2.2.2) with a SBL approach (section 2.1). We also compare the performances with SOFOMORE<sup>4</sup>. While gaussian priors of equal variance are attributed to each projection gain of a given lead-field column in its original version (Stahlhut et al. (2011)), we set the variance values as estimated by our method (*i.e.*, the diagonal values of the covariance matrices as computed in equation (9)). SOFOMORE is then a particular case of our approach with diagonal lead-field column precisions. This comparison is carried out to assess the importance of explicitly encoding physical model-driven constraints within the lead-field priors. We also compare the results using RAP-MUSIC (Mosher and Leahy (1998)) informed with the true number of sources, and another empirical Bayesian approach called Champagne (Wipf et al. (2010)) for which the true standard-deviation of the noise has been informed. For SBL, RAP-MUSIC, and Champagne, a local OSM lead-field is built: for each position of the lead-field grid, the three cartesian projections are given by the OSM model (among the  $M = 17$  previously computed) whose fitting point on the skull mesh is the closest to the considered grid point.

### 3.1.1. Evaluation criteria

The performances of the different algorithms are evaluated based on two complementary metrics. The first one is the Distance of Localization Error (DLE) (Yao and Dewald (2005); Becker et al. (2014)), which measure the localization accuracy of the

---

<sup>4</sup>used in its non-informative form (gamma hyper-parameters set to 0)

true sources while penalizing the presence of spurious sources (false positives). The DLE is given by:

$$\text{DLE} = \frac{1}{2n_s} \sum_{i \in \mathcal{S}} \min_{j \in \hat{\mathcal{S}}} \|x_i - \hat{x}_j\| + \frac{1}{2\hat{n}_s} \sum_{j \in \hat{\mathcal{S}}} \min_{i \in \mathcal{S}} \|x_i - \hat{x}_j\| \quad (24)$$

with  $\mathcal{S}$  and  $\hat{\mathcal{S}}$  the set of simulated sources and of estimated sources respectively,  $x_i$  the position of the  $i^{\text{th}}$  simulated source and  $\hat{x}_j$  the position of the  $j^{\text{th}}$  estimated source, and  $n_s$  is the number of simulated sources while  $\hat{n}_s$  is the number of estimated sources. In this paper, a dipole is included in the solution  $\hat{\mathcal{S}}$  if the standard-deviation of its estimated time-course is above a threshold given as the tenth of the largest standard-deviation in the current solution.

The second metric is the  $\mathbf{A}'$  metric (Snodgrass and Corwin (1988); Wipf et al. (2010)) which take into account both the ratio of successful source localization (hits) and the ratio of false positives. A simulated source is considered as recovered (true positive) if at least one dipole in the solution  $\hat{\mathcal{S}}$  is less than  $1\text{cm}$  far from this dipole. The hits rate is the number of true positives divided by the total number of simulated sources  $n_s$ . Any dipoles in the solution  $\hat{\mathcal{S}}$  farther than  $1\text{cm}$  from all the true dipoles are considered as false positive. To define a false positive rate, a maximum number of possible false positives has to be defined. This maximum number has been arbitrarily set to 100 in Wipf et al. (2010). In this paper we decide to set it to 48, which is the highest number of false positives obtained in our simulations (Champagne algorithm with 5 dipoles in the  $-5\text{dB}$  case). Given these hit rates ( $HR$ ) and false positives rates ( $FPr$ ), the  $\mathbf{A}'$  metric is computed as (Snodgrass and Corwin (1988)):

$$\mathbf{A}' = \begin{cases} 0.5 + (HR - FPr)(1 + HR - FPr)/(4HR(1 - FPr)), & \text{if } HR \geq FPr \\ 0.5 + (FPr - HR)(1 + FPr - HR)/(4FPr(1 - HR)), & \text{if } FPr > HR \end{cases} \quad (25)$$

In addition, we assess the ability of the methods in estimating the dipoles time-courses and projections on the contacts using the correlation coefficients between the estimations and the ground-truth. Finally, we also evaluate the sparsity of the methods by providing the number of dipoles  $\hat{n}_s$  selected in the final solution  $\hat{\mathcal{S}}$ .

### 3.1.2. Results

We iterate 400 simulations for each configuration (number of sources and noise level). For all the algorithm, the variance of the noise is initialized using a minimum description length (MDL) approach (except for RAP-MUSIC where the exact number of sources is given and Champagne to which the true noise variance is given). For all empirical Bayes algorithm, the prior source precision  $\Gamma_{\mathcal{S}}$  is set as the identity matrix.

The performances in term of time-course and lead-field reconstructions (fig 3, rows 1 & 2) demonstrate the good performance of our approach. VBLF is indeed consistently the most accurate in median over all configurations, with significantly higher time-course correlation rates ( $\rho_t$ ) compared to the other methods. The method is particularly accurate in estimating the forward field coefficients, reaching correlation rates

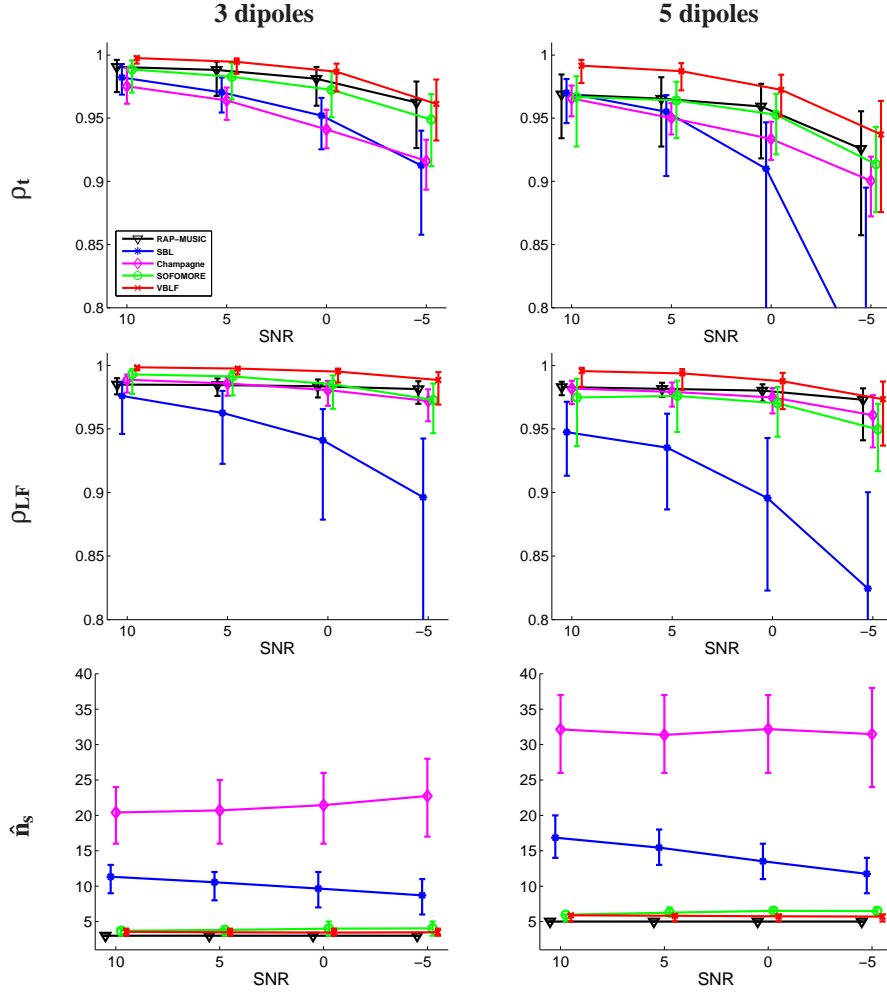


Figure 3: Performance evaluation for 3 (first column) and 5 (second column) simultaneous dipoles. Medians of the correlation coefficients of the estimated time-course and lead-field projections with the ground-truth are given on resp. first ( $\rho_t$ ) and second ( $\rho_{LF}$ ) rows. The median number of significant dipoles in the solution ( $\hat{n}_s$ ) for each method are given on the third row. Error bars corresponding to first and third quartiles are also given.

$\rho_{LF}$  over 0.99 for 3 and 5 dipoles from 10 to 0dB SNR. If SOFOMORE manages to compete with both RAP-MUSIC and Champagne in the 3 dipoles case, it is less accurate than these two fixed lead-field methods in reconstructing the lead-field columns in the 5 dipoles case. This proves that it might be inappropriate to give degrees of freedom to the lead-field coefficients if the optimization procedure is not carefully constrained. On fig.4 are given an example of lead-field coefficients re-estimates (*i.e.*, including the post-ECD scaling procedure) for both VBLF and SOFOMORE, for the 5 dipoles/10dB simulation case. Starting with an OSM prior, VBLF proves to be able to produce rel-



evant forward field posterior mean close to the true (simulated) one, here given by a FEM. The unconstrained version (*i.e.*, SOFOMORE, diagonal precision matrix over the columns of  $\mathbf{A}$ ) brings less accurate estimates (see fig.4, 2<sup>nd</sup> row 3<sup>rd</sup> column in particular, errors are visible between the blue and dashed green lines). Such bad estimates interfere with the estimates of the other simultaneous dipoles, impacting significantly the localization results.

One can notice the good performance of the Champagne algorithm in this reconstruction task, this method being originally dedicated to robust estimation of the source time-courses (Wipf et al. (2010)). RAP-MUSIC also manage to produce reliable estimates of the lead-field and time-courses, the method is successful in identifying the 3D subspaces best fitting the projections of each simulated dipole.

The performances of both SOFOMORE and VBLF in estimating the true number of sources are consistently very near the RAP-MUSIC reference (true number of sources informed), with a small but significant advantage for our VBLF approach considering the tight quartile values (see the  $\hat{n}_s$  metric, third row of fig 3). These methods are able to retrieve the true number of sources even in case of strong interferences (up to 5 simultaneous dipoles) and noise level (-5dB on average for each dipole), while Champagne and SBL provides large over-estimations of the number of underlying sources. This can be explained by the basis mismatch issue as discussed in (Chi et al. (2011)), the true forward field of the sources being not perfectly aligned with any columns of the lead-field matrix, a high number of lead-field elements are necessary for the algorithm to reconstruct the observations. The re-estimation of these lead-field columns in SOFOMORE or VBLF brings a data-driven fitting of the lead-field to the true source projection, avoiding such scattering effects.

The performances in term of localization precision (DLE) and  $A'$  metric are given fig 5. The informed version of RAP-MUSIC is expected to provide accurate results in this context of additive white Gaussian noise. The Champagne approach also provides very accurate localization results, with DLE below or very near to 1cm within each  $n_s$  and SNR configurations. However as shown by the rather high number of significantly powerful sources in the solution (see fig 5), the number of false positives can be high as shown by the  $A'$  metric, especially in the  $n_s = 5$  dipoles case.

After applying the ECD fitting step (as described section 2.2.3), VBLF brings the best DLE performance, in concordance with the high accuracy of the source lead-field estimates. The DLE is decreased of about 2mm in median with respect to SOFOMORE over all configurations, which can be considered as a significant gain in accuracy considering that the DLE are consistently below 1cm (representing roughly a gain in performance of about 20%). The true number of sources being accurately estimated by VBLF, very few false positives are produced, explaining the high  $A'$  values. This is not observed for SOFOMORE, for which the rather low number of false positives are compensate by higher localization errors due to inaccuracies in the source lead-field estimates. The superiority of our method with respect to the SOFOMORE approach demonstrates the importance of constraining the dipole projections through the introduction of non-null off-diagonal values in each  $\Gamma_{a_i}$ .

A last metric is given by the final values of the free energy function (equation (21)) which describe the relevance of the model to the data (Friston et al. (2008)). This lower bound of the model evidence can be evaluated for SBL, SOFOMORE and VBLF. While

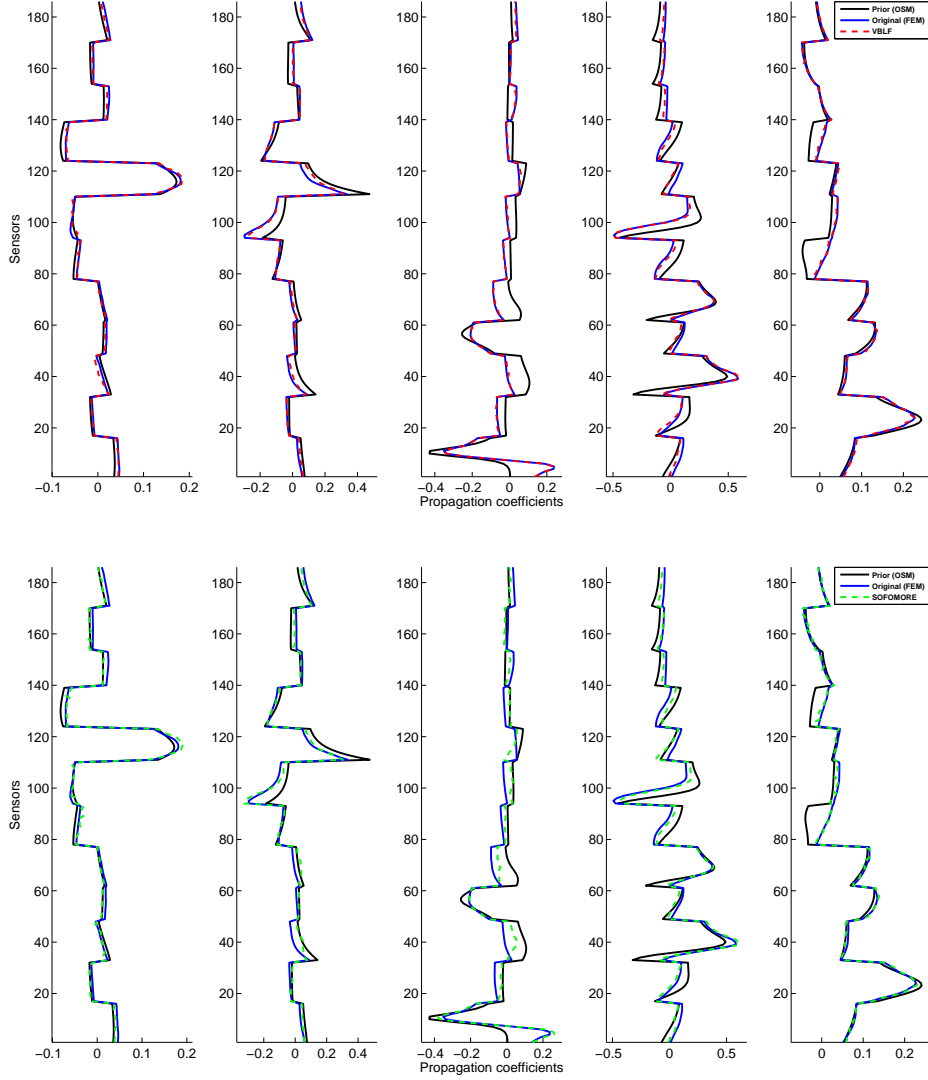


Figure 4: An illustration of the lead-field coefficients estimates when using VBLF (top/dashed red) or SOFOMORE (bottom/dashed green), in the 5 dipoles case - 10dB (one figure per dipole). The estimates are super-imposed with the true FEM lead-field (blue) and the prior column (black).

the actual values of this function has no meaning in itself, its comparison from a model to another can be used to determine which model is the most appropriate to explain the data (fixed lead-field (SBL), diagonal lead-field column covariance matrices (SOFOMORE), and non-diagonal lead-field column covariance matrices (VBLF)). Using the values of the SBL free energy estimates as a reference (setting it to 0), VBLF provides

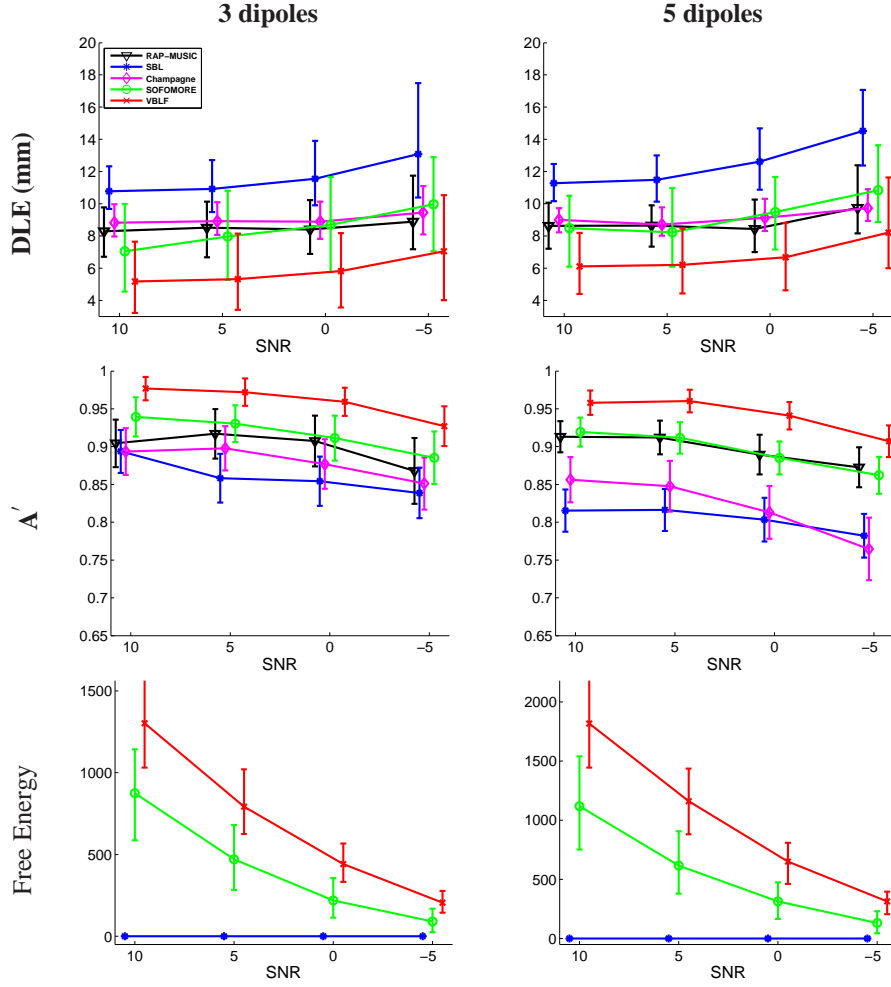


Figure 5: Performance evaluation for 3 (first column) and 5 (second column) simultaneous dipoles. Median of the DLE and  $A'$  metrics are given resp. on the first and second rows. On the third row is given the relative median values of the Free Energy for VBLF and SOFOMORE with respect to the SBL reference values. Error bars of first and third quartiles are superimposed.

the highest free energy values over all the configuration and confirms its relevance for this simulated source localization task (bottom figures of fig 3).

### 3.2. Intracranial Stimulations

The localization method is tested on Intracranial Stimulation (ICS) data, where two consecutive electrode contacts are used to inject a current in the brain volume. We will consider a data set from a 40-year-old man with presumed bitemporal lobe epilepsy. He was implanted with ten depth electrodes in the right temporal lobe and insular cortex

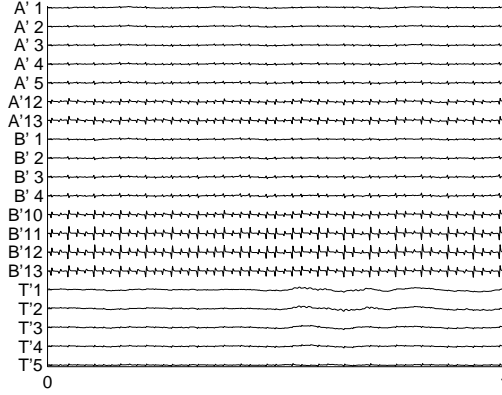


Figure 6: One second of SEEG signals recorded during an ICS session on TB'8-9. These 20 channels are used for the localization of the stimulation.

and four in the left temporal lobe. The reference was FPz surface electrode from the classical 10-20 system. The depth contacts coordinates were automatically determined using the procedure described in Hofmanis et al. (2011). We consider stimulations injected in the left (less implanted) hemisphere, using the contacts of the electrodes TB'.

Three ICS data-set are used, and correspond to a profound (TB'2-3), intermediate (TB'4-5) and superficial (TB'8-9) stimulation site. The stimulation is applied for about 4s for each site, and consists in a train of 55 biphasic pulses/s. We have extracted 3s of signal centered within the stimulation window, and we cut it in 12 sub-windows of 250ms each (thus 13 to 14 pulses per window). The localization is applied separately on each sub-windows. While it is true that the ICS source is stronger than physiological sources, it may not be the dominant activity recorded by far electrodes on which local physiological activities are super-imposed. On figure 6, the time-courses of the potentials generated by the ICS dipoles for the stimulation site TB'8-9 are given. On the contact T'1, distant of 36mm from the stimulation site, the potentials due to the ICS source are visible but is clearly weaker than the surrounding physiological activity.

The lead-field matrix is built based on 17 locally fitted OSM using a 1cm resolution grid (like in the simulations), the spheres being fitted on 17 well distributed points on the upper part of the inner skull. The contacts of the electrode TB' used for the stimulation are not considered. We aim at localizing the stimulation based only on a limited set of distant contacts. We restrict the used measurements to those of the hemisphere of the stimulation, thus 20 channels spread on 3 electrode shafts (A', B', T', see fig 6). Using electrode shafts can pose a conditioning problem, as tackled in previous works (Caune et al. (2014)). We arrived to the conclusion that at least a dozen contacts from 3 different electrode shafts are needed to successfully carry out a brain source localization, such conditions being met in the particular case of this data set.

The position errors  $\epsilon_p$  are given in table 1, averaged over the 12 sub-windows localization results. After convergence, each algorithm generates solutions whose powers

	TB'2-TB'3			TB'4-TB'5			TB'8-TB'9		
	$\varepsilon_p$	$\rho_t$	$\rho_{LF}$	$\varepsilon_p$	$\rho_t$	$\rho_{LF}$	$\varepsilon_p$	$\rho_t$	$\rho_{LF}$
SBL	$3.0 \pm 0.2$	$0.98 \pm 0.01$	$0.93 \pm 0.01$	$5.0 \pm 0.6$	$0.96 \pm 0.01$	$0.97 \pm 0.01$	$4.8 \pm 0.3$	$0.95 \pm 0.02$	$0.93 \pm 0.02$
VBLF	$1.8 \pm 0.9$	$0.97 \pm 0.01$	$0.97 \pm 0.01$	$3.0 \pm 0.5$	$0.95 \pm 0.04$	$0.99 \pm 0.01$	$3.1 \pm 0.7$	$0.96 \pm 0.03$	$0.99 \pm 0.02$
SOFOMORE	$1.3 \pm 0.7$	$0.95 \pm 0.02$	$0.97 \pm 0.01$	$3.1 \pm 0.7$	$0.97 \pm 0.03$	$0.99 \pm 0.01$	$2.3 \pm 0.8$	$0.95 \pm 0.02$	$0.99 \pm 0.02$
Champagne	$4.8 \pm 0.08$	$0.93 \pm 0.04$	$0.92 \pm 0.02$	$4.6 \pm 1.2$	$0.97 \pm 0.01$	$0.95 \pm 0.03$	$5.4 \pm 0.9$	$0.95 \pm 0.03$	$0.90 \pm 0.04$
RAP-MUSIC	$7.6 \pm 1.9$	$0.97 \pm 0.01$	$0.97 \pm 0.00$	$5.9 \pm 0.0$	$0.99 \pm 0.00$	$0.97 \pm 0.00$	$8.8 \pm 0.27$	$0.98 \pm 0.01$	$0.98 \pm 0.00$

Table 1: Performance values for the localization of three ICS dipoles (averaged over 12 windows of 250ms each) for five different algorithms: SBL, VBLF, SOFOMORE, Champagne and RAP-MUSIC. Position errors  $\varepsilon_p$  are given in mm. Correlation coefficients  $\rho_t$  between the mean of the closest (<1cm) significant source time-courses and the potential recorded on the electrode of stimulation next to the stimulated contacts. Correlation coefficients  $\rho_{LF}$  between the mean lead-field projections of the closest significant sources (re-estimated in the case of VBLF and SOFOMORE) and an estimation of the true projection of the dipole of stimulation based on the One Sphere Model.

are concentrated around the position of the stimulation dipole. We retain the time-course and lead-field projection of the equivalent dipole of maximum power as this representing the stimulation dipole. In the table 1 are then given two additional performance values (averaged over the 12 sub-windows):  $\rho_t$  is the correlation coefficient between the estimated equivalent dipole time course and the potential recorded on a non-saturated (manually chosen) contact belonging to the electrode of stimulation.  $\rho_{LF}$  is the correlation coefficient between the estimated equivalent dipole lead-field projection with an OSM estimated projection of the dipole of stimulation using its true position and orientation.

The performances of all five algorithms (SBL, VBLF, SOFOMORE, Champagne as well as RAP-MUSIC) are evaluated (considering the post-optimization of section 2.2.3 for both SOFOMORE and VBLF). The number of sources in RAP-MUSIC has been informed through an MDL approach (1 to 3 sources are estimated depending on the analyzed window). All algorithms are successful in localizing the stimulation dipole (under 1cm on average with small standard-deviation), with consistent advantages for SBL and SOFOMORE for which the regressed dipole is about 2 to 3 mm from the stimulation site (taken as the middle of the two stimulated contacts). These results are competing with the localization using an ECD approach applied on the averaged stimulation peaks within a window of 2.5s (see Caune et al. (2014)). This can be taken as proofs that (i) the lead-field projection is indeed well recovered (further confirmed by the  $\rho_{LF}$  correlation values) and (ii) the OSM indeed provides a reliable approximation of the head propagation medium when a precision of the order of few millimeters is needed. Relying on recent studies on the topic (Birot et al. (2014); Caune et al. (2014)), OSM is reported to provide a reliable approximation of the real propagation medium and is accurate enough to reach precise localization errors of the order of few millimeters, especially in this context of intracranial SEEG recordings where we are far less impacted by errors in the skull:brain ratio estimate than in the case of scalp recordings (Acar and Makeig (2013); Acar et al. (2016)).

### 3.3. Epileptic interictal spikes

We now evaluate the methods on a data set of interictal spikes recorded from a 28 year-old woman with drug-resistant insulo-opercular epilepsy (already included in a

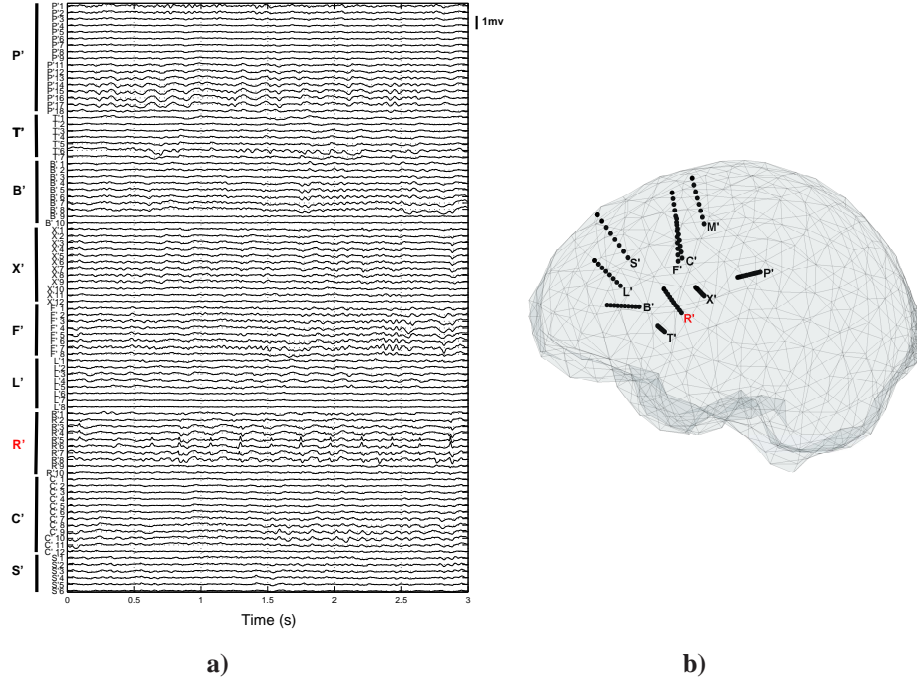


Figure 7: a) SEEG signals of interictal spikes in common reference montage (FPz scalp electrode). Epileptic spikes are maximally visible on the contacts 5 and 6 of the depth electrode R' (left central operculum). b) Mesh of the head and SEEG configurations. The labels are near the most profound contact of their corresponding electrode shaft, *i.e.*, the contact 1.

previous paper (Caune et al. (2014))). She gave her informed consent prior to participation. The intracranial EEG (SEEG) recording setup consists in ten electrode shafts implanted in the insulo-opercular regions (internal and lateral contacts, see fig.7(b)): P', cingulum/parietal operculum ; T', infero-anterior insula/ superior temporal gyrus ; B', anterior insula/pre-motor cortex; X', posterior insula/post-central gyrus; F', and L', anterior and posterior part of the inferior frontal sulcus/middle frontal gyrus; R', middle insula/central operculum; C', cingulum/middle frontal gyrus; S', superior frontal sulcus; M', supplementary motor area/superior frontal gyrus. The inspection of the SEEG data on bipolar montage reveals the presence of interictal epileptic spikes near the R' depth electrode, especially on the contacts R'5 and R'6, which also impact neighboring electrodes (especially T' and X'), though with rather low signal to noise ratios. Within the time windows of spikes occurrence (about 100ms length), few other activities of lower amplitudes are impacting the electrodes. We then expect to localize the main generator(s) within a focal area surrounding the electrode shaft R'. Three seconds of interictal recordings on the 100 available contacts are given fig. 7(a).

Several OSM are computed to build the lead-field matrix used for the inversion. 17 local spheres are fitted on the inner skull of the patient following the procedure described above, and we use a regular 3D grid with 7mm resolution restricted to the

implanted hemisphere for the forward model computation, yielding 2322 source positions (*i.e.*, 6966 columns in the forward matrix  $\mathbf{A}$ ). The coordinates of the depth contacts were obtained as for the ICS localization described above. All the contacts, at the exception of those of the electrode R', are used to perform the localization, *i.e.*, 90 recording channels.

The bipolar analysis of the R' shaft reveals a polarity inversion between R'6 and R'7, indicating a dipolar source situated in a plane between the two contacts and having a strong component oriented orthogonal to this plane. From the potential time course recorded on the contact R'6, 20 interictal spikes have been identified through a simple thresholding process and validated by the expert. The localization is carried out after averaging the 20 time windows of 64 samples (125ms), thus enhancing the signal to noise ratio of the inter-ictal activity of interest.

The same five algorithms were tested on this data set. Again, MDL is used for the initialization of RAP-MUSIC (resulting in 7 estimated sources). Figure 8 illustrates the localization map for each tested algorithm. The localization of the most intense equivalent dipole is concordant for all the methods, distant from 6.2mm to 9.8mm from the R'6 contact, and with a maximum distance of 6mm between the two most distant solutions (SBL with RAP-MUSIC). It is also consistent with the localization results obtained using an ECD approach in a recent paper (Caune et al. (2014)), and validated by the expert.

In particular, the localization map is very sparse for the VBLF approach, five co-activated dipoles (*i.e.*, with a standard deviation over the tenth of the most intense dipole) are found within the cloud of the electrodes, explaining co-occurring activities of low amplitudes. These dipoles are localized within the boundaries of the brain (gray matter) mesh. Champagne also provide a similarly sparse localization map, although yielding nine co-occurring equivalent dipoles. SOFOMORE as well as SBL are providing less sparse maps. Both approaches are estimating dipoles of relative high intensity and distant from the cloud of electrodes (out of the brain mesh in some case), while such results seem to be irrelevant for this given data for which the main activities are clinically presumed to take place near the electrode R'. Finally, VBLF and Champagne are producing very similar time courses for the main equivalent dipole, having their maximum peak at the same instant, while the activities estimated by the three other algorithms are delayed of a few samples (fig. 9 left).

When comparing the estimated lead-field projection given by the VBLF and the SOFOMORE approach (fig. 9 right), one can see the somewhat unplaussible estimate of the lead-field projection produced by SOFOMORE, while the estimate of the VBLF approach is more relevant. This is confirmed when applying the fitting of both estimated columns using the OSM. The resulting mean squared error of the OSM fitting is of 0.01 for the VBLF estimate while it reaches 0.08 in the case of SOFOMORE.

#### 4. Discussion

In this study we propose a distributed approach for simultaneous estimation of the brain source activities under controlled re-adjustment of their lead-field projections  $\mathbf{A}$ , based on a Variational Bayesian (VB) scheme. The re-estimation of the columns are



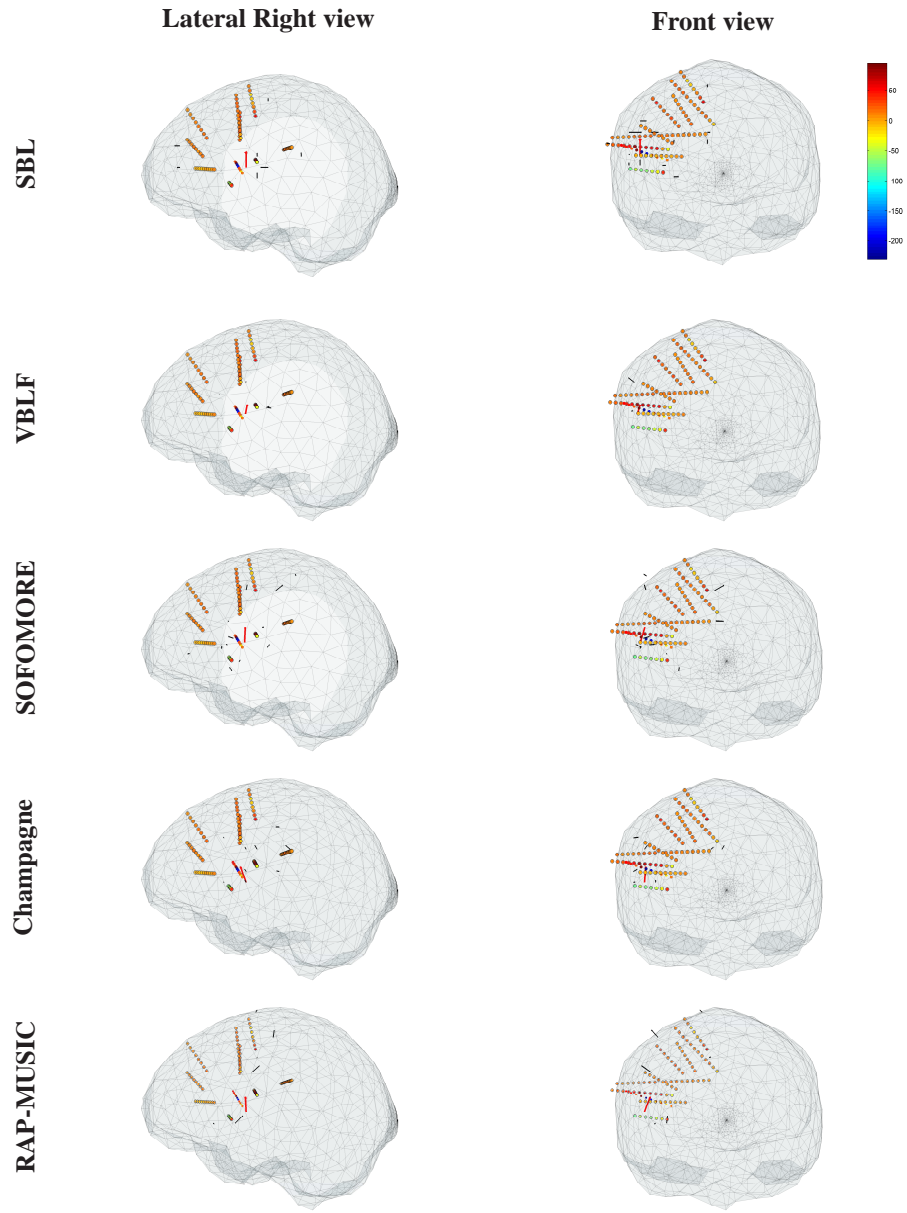


Figure 8: Localization maps of averaged inter-ictal spikes (20 events) over a window length of 125ms (64 samples), centered on the maximum values of each spike (as recorded on the electrode R'6). The potential values ( $\mu V$ ) on the electrodes are given as colors (dark blue for negative values to dark red for positive values, see colorbar), and the contacts used for the localization are black circled (all electrodes except R'6 on which the interictal spikes has been identified by the expert). Two views (lateral right and frontal) are given. The dipoles in red represent the equivalent dipoles identified as the estimated generators of the interictal spikes, while the dipoles in black represent the simultaneous activities estimated by each algorithm.

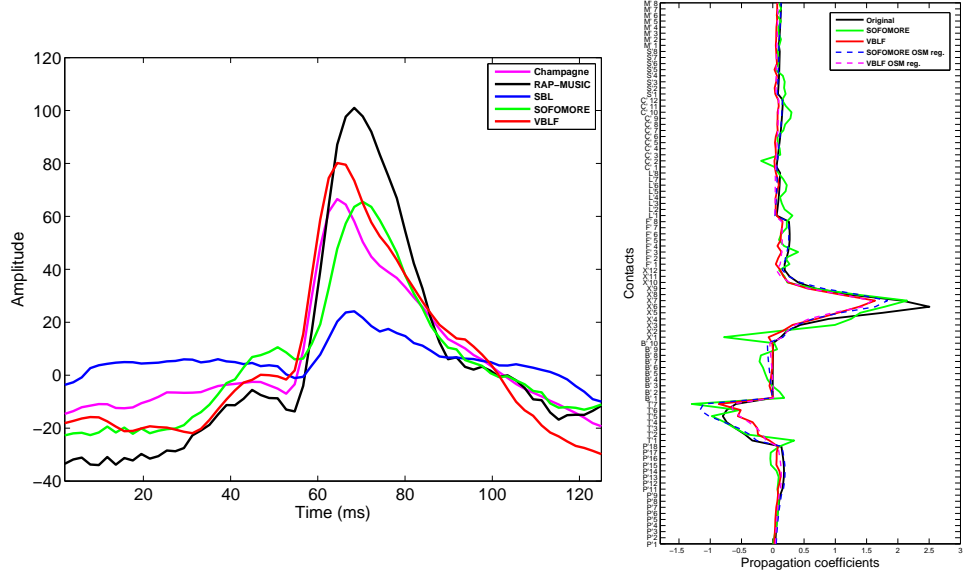


Figure 9: Left: estimated time courses for the equivalent dipoles of maximum power for each algorithm. Right: original projection coefficients on the 90 contacts for the elementary dipole chosen by both VBLF and SOFOMORE, along with the lead-field optimization given by each algorithm, on which are super-imposed their OSM regressions (dashed lines).

controlled through covariance matrices, constraining the projections to remain physiologically plausible, *i.e.*, preventing to deviate implausibly from prior projection patterns  $\mathbf{A}_0$  computed using a physical propagation model. We provide comparisons with several algorithms from the literature (SBL, SOFOMORE, Champagne, RAP-MUSIC) with several noise level configurations, giving quantitative facts on the benefits of taking account of the forward model uncertainties.

The simultaneous optimization of the source time-courses and of the lead-field projections implies that the estimated lead-field columns  $\hat{\mathbf{A}}$  do not correspond anymore with the projection of unitary dipolar elements as initially encoded in  $\mathbf{A}_0$ , in term of positions, orientations and (unitary) amplitudes. The ambiguity in the estimated source parameters can be disentangled by a fitting of the posterior columns based on an analytical model, such as the One Sphere model used in this study. In this perspective, the VBLF approach can be seen as a decomposition on a continuous dictionary where the dictionary atoms are iteratively learned from the observations (Yang et al. (2013)). The method separates the different sources contributing to the data, and optimizes their projection gains on the contacts through a constrained data-driven procedure. The modified projection gains are then mapped and scaled to their corresponding unitary dipolar components in the source space through a classical Scherg’s single-dipole inversion, providing the final source parameter estimates.

While this is a distributed approach in its formulation, our methods might be seen as an ECD approach where thousands of dipole candidates are simultaneously opti-

mized in their amplitudes, positions and orientations. This preclude the initialization issues (multi-start strategies, initial guess of the number of dipoles) encountered when dealing with the ECD approaches (Scherg et al. (1999); Kiebel et al. (2008); Caune et al. (2014)).

The gain in sparsity is significant when the lead-field is reconsidered (both for SOFOMORE and our method), and the size of the estimated source space is consistently very close to the true one in the simulated case. By offering such flexible framework, both the estimation of the equivalent dipoles projections and time-courses are enhanced, on condition that the correction of the projection coefficients is carefully controlled. Indeed the unconstrained re-evaluation scheme (*i.e.*, SOFOMORE) leads globally to less accurate estimation results than our constrained version (VBLF), especially for low SNR levels.

Our approach provides consistent results when applied on real data. First, its application to an Intracranial Stimulation (ICS) data set demonstrates its ability to provide accurate localization, time course estimation and lead-field projection estimates in a real context. Secondly, the method proves to produce plausible localization map in a case of inter-epileptic spikes localization. The result is concordant with the expertise of the neurologists. In this last case, the importance of using non-diagonal covariance matrix for modeling the lead-field column distribution is further illustrated, as the lead-field estimate produce by SOFOMORE is not physiologically relevant.

It is worth noticing that the localization accuracy could obviously be enhanced by introducing a finer grid of elementary dipole positions. However it would still remains a discrete sampling of the source space with uncertain model projections, thus prone to mismatch basis issues (Chi et al. (2011)). Also, it would result to a higher computational burden. To give a quantitative comparison, in the simulated case our VBLF scheme takes about 2 minutes for 1000 iterations with the considered lead-field grid of 1cm resolution (6642 columns), while the SBL approach takes about 1 minute for the same number of iterations. Besides, the variational Bayesian optimization scheme is reported to be slow. As a faster alternative, modified versions of the gradient descent algorithms inspired from MacKay (1992); Miele and Cantrell (1969); Zheng et al. (2015) could be adopted, through a direct gradient optimization of the free-energy function.

The analysis of the confidence we can place in a localization result must be deepened, in particular in this difficult context of SEEG recordings where the conditioning of the electrode shafts and their positions with regard to the structures of interest are sensitive questions. Such issues have been tackled in a previous study (Caune et al. (2014)) in the case where a single dominant dipole is to be identified, based on a dipole fitting (ECD) procedure. The case of multiple simultaneous sources can now be studied through the use of the proposed distributed source models, with the aim to provide a clinical added value on how the electrode shafts has to be implanted for reaching confident localization results covering the whole brain volume, or at least for the brain structures suspected to contain the sources of interest. In this perspective, extended validation campaign on real recordings of interictal spikes, or more likely on evoked potential activities (for which the source positions, time courses and instant of apparition are perfectly controlled), must be carried out in close collaboration with neurophysiologists.

## 5. Acknowledgments

The authors thank Prof. S. Colnat-Coulbois (Nancy) for the neurosurgical contributions, and the project ANR SEPICOT for their financial support.

## 6. References

- Acar, Z. A., Acar, C. E., Makeig, S., 2016. Simultaneous head tissue conductivity and EEG source location estimation. *NeuroImage* 124, 168–180.
- Acar, Z. A., Makeig, S., 2013. Effects of forward model errors on eeg source localization. *Brain topography* 26 (3), 378–396.
- Baillet, S., Mosher, J. C., Leahy, R. M., 2001. Electromagnetic brain mapping. *Signal Processing Magazine, IEEE* 18 (6), 14–30.
- Bangera, N. B., Schomer, D. L., Dehghani, N., Ulbert, I., Cash, S., Papavasiliou, S., Eisenberg, S. R., Dale, A. M., Halgren, E., 2010. Experimental validation of the influence of white matter anisotropy on the intracranial eeg forward solution. *Journal of computational neuroscience* 29 (3), 371–387.
- Becker, H., Albera, L., Comon, P., Gribonval, R., Wendling, F., Merlet, I., 2014. A performance study of various brain source imaging approaches. In: *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, pp. 5869–5873.
- Birot, G., Spinelli, L., Vulliémoz, S., Mégevand, P., Brunet, D., Seeck, M., Michel, C. M., 2014. Head model and electrical source imaging: a study of 38 epileptic patients. *NeuroImage: Clinical* 5, 77–83.
- Bishop, C. M., et al., 2006. *Pattern recognition and machine learning*. Vol. 1. springer New York.
- Caune, V., Le Cam, S., Ranta, R., Maillard, L., Louis-Dorr, V., 2013. Dipolar source localization from intracerebral SEEG recordings. In: *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, pp. 41–44.
- Caune, V., Ranta, R., Le Cam, S., Hofmanis, J., Maillard, L., Koessler, L., Louis-Dorr, V., 2014. Evaluating dipolar source localization feasibility from intracerebral SEEG recordings. *NeuroImage* 98, 118–133.
- Chang, N., Gulrajani, R., Gotman, J., 2005. Dipole localization using simulated intracerebral EEG. *Clinical neurophysiology* 116 (11), 2707–2716.
- Chi, Y., Scharf, L. L., Pezeshki, A., Calderbank, A. R., 2011. Sensitivity to basis mismatch in compressed sensing. *Signal Processing, IEEE Transactions on* 59 (5), 2182–2195.

- Coito, A., Plomp, G., Genetti, M., Abela, E., Wiest, R., Seeck, M., Michel, C. M., Vulliemoz, S., 2015. Dynamic directed interictal connectivity in left and right temporal lobe epilepsy. *Epilepsia* 56 (2), 207–217.
- Dale, A. M., Sereno, M. I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *Journal of cognitive neuroscience* 5 (2), 162–176.
- David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., Friston, K. J., 2006. Dynamic causal modeling of evoked responses in eeg and meg. *NeuroImage* 30 (4), 1255–1272.
- Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., Group, B. D. C., et al., 2011. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* 54 (1), 313–327.
- Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage* 39 (3), 1104–1120.
- Gorodnitsky, I. F., Rao, B. D., 1997. Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm. *Signal Processing, IEEE Transactions on* 45 (3), 600–616.
- Hansen, S. T., Hauberg, S., Hansen, L. K., 2016. Data-driven forward model inference for EEG brain imaging. *NeuroImage* 139, 249–258.
- Hofmanis, J., Caspary, O., Louis-Dorr, V., Maillard, L., 2011. Automatic depth electrode localization in intracranial space. In: 4th International Conference on Bio-inspired Systems and Signal Processing, Biosignals 2011. p. CDROM.
- Kiebel, S. J., Daunizeau, J., Phillips, C., Friston, K. J., 2008. Variational bayesian inversion of the equivalent current dipole model in EEG/MEG. *NeuroImage* 39 (2), 728–741.
- Le Cam, S., Caune, V., Ranta, R., Maillard, L., Koessler, L., Louis-Dorr, V., 2014. Influence of the stereo-EEG sensors setup and of the averaging on the dipole localization problem. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 1147–1150.
- Lucka, F., Pursiainen, S., Burger, M., Wolters, C. H., 2012. Hierarchical bayesian inference for the EEG inverse problem using realistic FE head models: depth localization and source separation for focal primary currents. *Neuroimage* 61 (4), 1364–1382.
- MacKay, D. J., 1992. Bayesian interpolation. *Neural computation* 4 (3), 415–447.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. *Clinical neurophysiology* 115 (10), 2195–2222.
- Miele, A., Cantrell, J., 1969. Study on a memory gradient method for the minimization of functions. *Journal of Optimization Theory and Applications* 3 (6), 459–470.

- Mosher, J. C., Leahy, R. M., 1998. Recursive music: a framework for EEG and MEG source localization. *Biomedical Engineering, IEEE Transactions on* 45 (11), 1342–1354.
- Nunez, P. L., Srinivasan, R., 2006. *Electric fields of the brain: the neurophysics of EEG*. Oxford university press.
- Pascual-Marqui, R. D., 2007. Discrete, 3D distributed, linear imaging methods of electric neuronal activity. part 1: exact, zero error localization. *arXiv preprint arXiv:0710.3341*.
- Phillips, C., Mattout, J., Rugg, M. D., Maquet, P., Friston, K. J., 2005. An empirical bayesian solution to the source reconstruction problem in EEG. *NeuroImage* 24 (4), 997–1011.
- Ramírez, R., Palmer, J., Makeig, S., Rao, B. D., Wipf, D. P., 2006. Analysis of empirical bayesian methods for neuroelectromagnetic source localization. In: *Advances in Neural Information Processing Systems*. pp. 1505–1512.
- Sato, M.-a., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical bayesian estimation for MEG inverse problem. *NeuroImage* 23 (3), 806–826.
- Scherg, M., Bast, T., Berg, P., 1999. Multiple source analysis of interictal spikes: goals, requirements, and clinical value. *Journal of Clinical Neurophysiology* 16 (3), 214–224.
- Snodgrass, J. G., Corwin, J., 1988. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General* 117 (1), 34.
- Stahlhut, C., Mørup, M., Winther, O., Hansen, L. K., 2011. Simultaneous EEG source and forward model reconstruction (SOFOMORE) using a hierarchical bayesian approach. *Journal of Signal Processing Systems* 65 (3), 431–444.
- Tipping, M. E., 2001. Sparse bayesian learning and the relevance vector machine. *The journal of machine learning research* 1, 211–244.
- Trujillo-Barreto, N. J., Aubert-Vázquez, E., Penny, W. D., 2008. Bayesian M/EEG source reconstruction with spatio-temporal priors. *Neuroimage* 39 (1), 318–335.
- Wipf, D., Nagarajan, S., 2009. A unified bayesian framework for MEG/EEG source imaging. *Neuroimage* 44 (3), 947–966.
- Wipf, D. P., Owen, J. P., Attias, H. T., Sekihara, K., Nagarajan, S. S., 2010. Robust bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *NeuroImage* 49 (1), 641–655.
- Wipf, D. P., Rao, B. D., 2007. An empirical bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Transactions on Signal Processing* 55 (7), 3704–3716.

- Wolters, C., Anwander, A., Tricoche, X., Weinstein, D., Koch, M., MacLeod, R., 2006. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *NeuroImage* 30 (3), 813–826.
- Yang, Z., Xie, L., Zhang, C., 2013. Off-grid direction of arrival estimation using sparse bayesian inference. *Signal Processing, IEEE Transactions on* 61 (1), 38–43.
- Yao, D., 2000. Electric potential produced by a dipole in a homogeneous conducting sphere. *IEEE transactions on bio-medical engineering* 47 (7), 964–966.
- Yao, J., Dewald, J. P., 2005. Evaluation of different cortical source localization methods using simulated and experimental EEG data. *Neuroimage* 25 (2), 369–382.
- Yvert, B., Fischer, C., Bertrand, O., Pernier, J., 2005. Localization of human supratemporal auditory areas from intracerebral auditory evoked potentials using distributed source models. *Neuroimage* 28 (1), 140–153.
- Zhang, Z., 1995. A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. *Physics in medicine and biology* 40 (3), 335.
- Zheng, Y., Fraysse, A., Rodet, T., 2015. Efficient variational bayesian approximation method based on subspace optimization. *Image Processing, IEEE Transactions on* 24 (2), 681–693.

## Appendix A. VB simultaneous optimization of $\mathbf{S}$ and $\mathbf{A}$

In this appendix we develop the expressions of the updates for the distributions  $q(\mathbf{S})$  and  $q(\mathbf{A})$ :

**Update of  $q(\mathbf{S})$ :**

$$\ln q(\mathbf{S}) \propto \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \Gamma_{\mathbf{S}}, \gamma_{\epsilon}, \mathbf{A}_0, \Gamma_{\mathbf{A}}) \rangle_{q(\mathbf{A})} \quad (\text{A.1})$$

$$\propto \langle \ln p(\mathbf{X} | \mathbf{S}, \mathbf{A}, \gamma_{\epsilon}) \rangle_{q(\mathbf{A})} + \ln p(\mathbf{S} | \Gamma_{\mathbf{S}}) \quad (\text{A.2})$$

where we have suppressed all the additive factors independent of  $\mathbf{S}$ . The expression for the updates of the Gaussian distribution  $q(\mathbf{S})$  then writes:

$$\hat{\Gamma}_{\mathbf{S}} = \Gamma_{\mathbf{S}} + \gamma_{\epsilon} (\hat{\mathbf{A}}^{\top} \hat{\mathbf{A}} + \mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}^{-1}}) \quad (\text{A.3})$$

$$\hat{\mathbf{S}} = \gamma_{\epsilon} \hat{\Gamma}_{\mathbf{S}}^{-1} \hat{\mathbf{A}}^{\top} \mathbf{X} \quad (\text{A.4})$$

with  $\mathbf{D}_{\hat{\Gamma}_{\mathbf{A}}^{-1}}$  is a  $N_s \times N_s$  diagonal matrix composed of the trace of the updated covariance matrix of the  $N_s$  columns of  $\mathbf{A}$ , and  $\hat{\mathbf{A}}$  is the updated version of  $\mathbf{A}$ .

**Update of  $q(\mathbf{A})$ :**



We write  $q(\mathbf{A})$  as the product of its marginals  $q(\mathbf{A}) = \prod_{j=1}^{N_s} q(\mathbf{a}_j)$ . The marginal likelihoods  $\{q(\mathbf{a}_j)\}_{j=1..N_s}$  are updated as:

$$\ln q(\mathbf{a}_j) \propto \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \Gamma_{\mathbf{S}}, \gamma_{\epsilon}, \Gamma_{\mathbf{A}}, \mathbf{A}_0) \rangle_{q(\mathbf{S}), q(\mathbf{A}^{\setminus j})} \quad (\text{A.5})$$

$$\propto \langle \ln p(\mathbf{X} | \mathbf{S}, \mathbf{A}, \gamma_{\epsilon}) \rangle_{q(\mathbf{S}), q(\mathbf{A}^{\setminus j})} + \ln p(\mathbf{a}_j | \mathbf{a}_{0j}, \Gamma_{\mathbf{a}_j}) \quad (\text{A.6})$$

with  $\mathbf{A}^{\setminus j}$  the matrix  $\mathbf{A}$  minus the column  $j$ .  $q(\mathbf{a}_j)$  is a gaussian distribution with the updated mean and precision parameters  $(\hat{\mathbf{a}}_j, \hat{\Gamma}_{\mathbf{a}_j})$ :

$$\hat{\Gamma}_{\mathbf{a}_j} = \Gamma_{\mathbf{a}_j} + \gamma_{\epsilon} \hat{E}_{\mathbf{s}_j} \mathbf{I}_{N_c} \quad (\text{A.7})$$

$$\hat{\mathbf{a}}_j = \hat{\Gamma}_{\mathbf{a}_j}^{-1} (\Gamma_{\mathbf{a}_j} \mathbf{a}_{0j} + \gamma_{\epsilon} \Lambda_j) \quad (\text{A.8})$$

where we define:

$$\hat{E}_{\mathbf{s}_j} = \hat{\mathbf{s}}_j \hat{\mathbf{s}}_j^{\top} + T \hat{\gamma}_{\mathbf{s}_j}^{-1} \quad (\text{A.9})$$

$$\Lambda_j = \mathbf{X} \hat{\mathbf{s}}_j^{\top} - \hat{\mathbf{A}}^{\setminus j} (\hat{\Gamma}_{\mathbf{S}_j}^{\setminus j} + \hat{\mathbf{S}}^{\setminus j} \hat{\mathbf{s}}_j^{\top}) \quad (\text{A.10})$$

With  $\hat{\Gamma}_{\mathbf{S}_j}^{\setminus j}$  the  $j^{\text{th}}$  column of  $\hat{\Gamma}_{\mathbf{S}}$  minus its  $j^{\text{th}}$  element,  $\hat{\mathbf{S}}^{\setminus j}$  the estimated source matrix  $\hat{\mathbf{S}}$  minus its  $j^{\text{th}}$  line.  $\hat{\mathbf{A}}$ , the updated version of  $\mathbf{A}$ , is composed of the  $N_s$  columns  $\{\hat{\mathbf{a}}_j\}$ .

#### Updates of $\Gamma_{\mathbf{S}}$ and $\gamma_{\epsilon}$ :

We derive the expected model likelihood  $\langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \Gamma_{\mathbf{S}}, \gamma_{\epsilon}, \mathbf{A}_0, \Sigma_{\mathbf{A}}) \rangle_{q(\mathbf{S}), q(\mathbf{A})}$  with respect to  $\Gamma_{\mathbf{S}}$  and  $\gamma_{\epsilon}$  to get the expression of their updates:

$$\gamma_{\mathbf{s}_i}^{-1} = \frac{\hat{E}_{\mathbf{s}_i}}{T} \quad (\text{A.11})$$

$$\gamma_{\epsilon}^{-1} = \frac{1}{TN_c} \|\mathbf{X} - \hat{\mathbf{A}} \hat{\mathbf{S}}\|_F^2 + \frac{1}{N_c} \text{trace}(\hat{\mathbf{A}} \hat{\Gamma}_{\mathbf{S}}^{-1} \hat{\mathbf{A}}^{\top}) + \frac{1}{TN_c} \sum_{i=1}^{N_s} \hat{E}_{\mathbf{s}_i} \text{trace}(\hat{\Gamma}_{\mathbf{a}_i}^{-1}) \quad (\text{A.12})$$

#### Appendix B. Free Energy

The Bayesian scheme is closely related to the optimization of the free-energy, whose expression is given as:

$$\mathcal{F}(q) = \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}, \mathbf{A})} - \langle \ln q(\mathbf{S}, \mathbf{A}) \rangle_{q(\mathbf{S}, \mathbf{A})} \quad (\text{B.1})$$

$$= \langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}, \mathbf{A})} - \langle \ln q(\mathbf{S}) \rangle_{q(\mathbf{S})} - \langle \ln q(\mathbf{A}) \rangle_{q(\mathbf{A})} \quad (\text{B.2})$$

with

$$\langle \ln p(\mathbf{X}, \mathbf{S}, \mathbf{A} | \theta) \rangle_{q(\mathbf{S}, \mathbf{A})} = \langle \ln p(\mathbf{X} | \mathbf{S}, \gamma_{\epsilon}) \rangle_{q(\mathbf{S})} + \langle \ln p(\mathbf{S} | \Gamma_{\mathbf{S}}) \rangle_{q(\mathbf{S})} + \langle \ln p(\mathbf{A} | \mathbf{A}_0, \Gamma_{\mathbf{A}}) \rangle_{q(\mathbf{A})} \quad (\text{B.3})$$

It brings a way to quantify how a model fits the data, and can then be used to compare model and select the best one in Bayesian sense. We develop the computation of each term as follows:

$$\langle \ln p(\mathbf{X}|\mathbf{S}, \gamma_\varepsilon) \rangle_{q(\mathbf{S})} = -\frac{TN_c}{2} \ln 2\pi \quad (\text{B.4})$$

$$\langle \ln p(\mathbf{S}|\Gamma_{\mathbf{S}}) \rangle_{q(\mathbf{S})} = -\frac{TN_s}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^T \mathbf{s}_t \Gamma_{\mathbf{S}} \mathbf{s}_t^\top + T \text{trace}(\Gamma_{\mathbf{S}} \hat{\Gamma}_{\mathbf{S}}^{-1}) \quad (\text{B.5})$$

$$\begin{aligned} \langle \ln p(\mathbf{A}|\mathbf{A}_0, \Gamma_{\mathbf{A}}) \rangle_{q(\mathbf{A})} &= -\frac{N_s N_c}{2} (\ln 2\pi) - \frac{1}{2} \sum_{j=1}^{N_s} [-\ln |\Gamma_{\mathbf{a}_j}| \\ &\quad + \text{trace}(\hat{\Gamma}_{\mathbf{a}_j}^{-1} \Gamma_{\mathbf{a}_j}) + (\hat{\mathbf{a}}_j - \mathbf{a}_{0j})^\top \Gamma_{\mathbf{a}_j} (\hat{\mathbf{a}}_j - \mathbf{a}_{0j})] \end{aligned} \quad (\text{B.6})$$

$$\langle \ln q(\mathbf{S}) \rangle_{q(\mathbf{S})} = -\frac{TN_s}{2} (\ln 2\pi + 1) + \frac{T}{2} \ln |\hat{\Gamma}_{\mathbf{S}}| \quad (\text{B.7})$$

$$\langle \ln q(\mathbf{A}) \rangle_{q(\mathbf{A})} = -\frac{N_s N_c}{2} (\ln 2\pi + 1) + \frac{1}{2} \sum_{j=1}^{N_s} \ln |\hat{\Gamma}_{\mathbf{a}_j}| \quad (\text{B.8})$$