A BAYESIAN APPROACH TO INTRODUCING ANATOMO-FUNCTIONAL PRIORS IN THE EEG / MEG INVERSE PROBLEM

Le présent chapitre a été publié et correspond à la référence suivante

Résumé — Dans ce chapitre, nous présenterons une nouvelle approche concernant l’estimation des amplitudes de sources dipolaires dans un modèle distribué en imagerie MEG at EEG. Cette méthode consiste en l’introduction d’informations a priori de nature spatiale et temporelle comme un remède possible à ce problème inverse mal posé. Un schéma de régularisation quadratique permet la préservation de discontinuités entre les amplitudes de certaines sources considérées comme a priori non corrélées ; notamment entre des dipôles situés sur des parois opposées d’un sillon. De plus, nous introduisons des contraintes de douceur temporelle dans l’évolution des amplitudes des sources à des échelles de temps bien plus petites que celles des processus cognitifs sous-jacents. Ces a priori sont aisément développés au sein d’un formalisme Bayésien, produisant un estimateur du Maximum a Posteriori (MAP) de l’activité électrique cérébrale. Nous présentons des résultats issus de simulations numériques de l’EEG et comparés à ceux issus de méthodes de régularisation quadratique ainsi que d’une technique de norme minimale généralisée appelée LORETA (Low Resolution Electromagnetic Tomography).

Abstract — In this paper, we present a new approach to the recovering of dipole magnitudes in a distributed source model for MEG and EEG imaging. This method consists in introducing spatial and temporal a priori information as a cure to this ill-posed inverse problem. A non quadratic spatial regularization scheme allows the preservation of dipole moment discontinuities between some a priori non-correlated sources, for instance when considering dipoles located on both sides of a sulcus. Moreover, we introduce temporal smoothness constraints on dipole magnitude evolution at time scales smaller than those of cognitive processes. These priors are easily integrated into a Bayesian formalism, yielding a maximum a posteriori (MAP) estimator of brain electrical activity. Results from EEG simulations of our method are presented and compared with those of classical quadratic regularization and a now popular generalized minimum-norm technique called LORETA (Low Resolution Electromagnetic Tomography).
Introduction

As a tool for human brain mapping of cognitive functions, functional brain imaging long term objectives are twofold: localizing groups of neurons involved in some simple cognitive tasks such as those dealing with a perceptual process, and then identifying dynamic links between these groups in order to hopefully better understand “how does the brain work?” [1].

Once the problem is put in such a way, it is obvious that functional imaging modalities should ideally offer excellent space and time resolutions. Defining space and time resolution is not an easy task and may still be considered as an open issue. Nevertheless, spatial and temporal accuracy should be at least better than 5 mm and 5 ms respectively.

Current functional brain imaging techniques such as Position Emission Tomography (PET), Single Photon Computed Tomography (SPECT), and Functional Magnetic Resonance Imaging (f-MRI) possess excellent spatial resolutions but fail to offer at the same time temporal resolution better than, say, one minute in PET/SPECT and one second in f-MRI ([2],[3]). This time limitation is mainly due to metabolic considerations, like the evolution of radioactive isotope concentration in blood with PET and SPECT. f-MRI time resolution is still very promising, but technical problems remain.

Electromagnetic imaging of the brain is the only functional imaging modality that is liable to offer excellent time resolution. In principle, one could be able to obtain functional images of the brain electrical activity up to every millisecond (which roughly corresponds to the data sampling rate). Unfortunately, spatial resolution is limited and moreover, very different source arrangements can lead to the same electroencephalographic (EEG) and magnetoencephalographic (MEG) measurements. Owing to this non-uniqueness and also to numerical instability, the inverse problem consisting in finding out a source map from data is said to be ill-posed. Many researchers have proposed methods to regularize this issue. Very often, these latter consist in finding one or two best-fitting current dipoles according to data or, while using a distributed source model, give minimum-norm solutions characterized by spatial over-smoothing. These methods do not explicitly introduce anatomical or physiological constraints in the source reconstruction process, and so may yield unrealistic solutions. In this paper, we propose an inversion method that introduces some a priori information and works out at some spatio-temporal regularization. For instance, when consistent with physiology we show how it may be possible to introduce discontinuities between magnitudes of sources
located on both sides of a sulcus. We also introduce temporal smoothness constraints on source intensities with respect to the data sampling rate. We particularly insist on the fact that this spatio-temporal information yields better and less time consuming results than spatial regularization only.

This paper is divided as follows: in section II, we review the formalism of the inverse problem with a distributed source model, and briefly describe the different approaches currently in use. In section III, we expose the a priori information we wish to introduce in the reconstruction process. These spatio-temporal priors result from multimodal registration of MRI and PET images. In section IV, we show how it is possible to introduce such an information in a Bayesian formalism with a maximum a posteriori (MAP) estimator of dipole activities. Then we describe a non-optimal and deterministic algorithm used to achieve a minimum of the energy function associated to the Gibbs formulation of the conditional probability distribution of dipole magnitudes given the data. Section 5 describes concisely the quadratic regularization (QR) and the generalized minimum-norm (LORETA) methods that will be used in section 6, where results from a simulated auditory experiment are shown and compared with both those from the QR and LORETA methods.

EEG / MEG source imaging as an inverse problem

The idea of using EEG and MEG as genuine brain imaging techniques is quite simple ([4], [5]). There are two complementary ways to measure the electromagnetic signature of the neural electric activity outside the skull: cutaneous electrodes measuring electric potentials - EEG, and supraconductive coils measuring some components of the magnetic field on the surface of the head - MEG. Thus, one can investigate whether it is possible to reconstruct neural ionic currents as sources of these data. Fundamentally, the main asset of EEG/MEG imaging is that one could obtain a mapping of neuronal electric activity with the same time resolution as the data sampling rate (typically 1 ms to 4 ms).

In a distributed source model using current dipoles with given positions and orientations, it is possible to gather all the sources in a row vector so as to obtain the following linear system:

$$M_n = G J_n + b_n,$$ (1)
where $\mathbf{M}_n$ is a $N_M$ row vector containing (EEG or MEG) measurements at time $n$, and $\mathbf{G}$ is a $N_M \times N$ matrix containing global information about the head model, the source set pattern and measurement positions on the surface of the skull. $\mathbf{J}_n$ is a $N$ row vector containing the $n^{th}$ time sample of dipole magnitudes. Finally, data are corrupted by an additive measurement noise $\mathbf{b}_n$. This is a standard writing for many linear forward problems in image reconstruction.

Thus, the $\mathbf{G}$ matrix and the $N$ source positions make up a framework for the following inverse problem: find the $N$ sources that best explain the EEG and/or MEG measurements. It is well known that this is an ill-posed problem:

**the solution is not unique**: radically different source configurations may offer the same excellent fit to data. Moreover, dipoles radially oriented with regards to the surface of the skull are “silent” in MEG measurements;

**the solution is not stable**: as matrix $\mathbf{G}$ is nearly singular, or in other words the linear system associated to (1) - i.e. $\mathbf{M}_n = \mathbf{G} \mathbf{J}_n$ - is ill-conditioned [38], there is little control over the propagation of measurement errors from the data to the solution. Regularization techniques are applied so as to counter the bad-conditioning of the linear operator by introducing complementary information on parameters to be estimated [7-9].

Owing to non-uniqueness, the problem should be confined to one for which the solution space dimension is reduced by the introduction of restricting criteria. The first approach consists in searching for a single or few dipoles whose forward solution to (1) is a best fit (generally in a least squares sense) to measurements ([10], [11]). Another technique can be found in [14], where Mosher et al. propose a method derived from the widely used antenna processing Multiple Signal Classification algorithm (MUSIC, [15]), which consists in estimating the active source locations by computing the projection of a fixed number of current dipoles onto the noise subspace of the measurement space. Greenblatt in [16] proposes an extension of this method to the case where there are more expected sources than measurement data, so there is no possible subspace decomposition of measurement space. His method consists in building a metric based on the covariance matrix estimated from data time slices. Then, the source space is projected through this metric onto the measurement space. A weighted forward matrix is computed via a singular value decomposition (SVD) of the matrix whose columns are the projections, through this metric, of the original source space basis onto the measurement space. This yields estimated source locations that mostly contribute to the
signal component in data. The final result is a weighted pseudo-inverse of the forward matrix \( G \) scanning through data time series to estimate sources.

In [12], a minimum-norm solution to (1) using a distributed source model (i.e. \( N \gg 1 \)) is proposed. Recently Pascual-Marqui et al. have presented a generalized minimum-norm solution to the inverse problem that appears to overcome some of the limitations of other minimum-norm methods that tend to yield solutions with most of the active dipoles at the surface of the source space as a priority (because of the decrease of lead field intensities with depth, see section V) [13].

As a primary conclusion, it appears that most of the time, single dipole models fail to give acceptable results when space and time overlapping in the source pattern is very likely. That is especially the case with the late components of evoked potentials where unique dipole models may yield inconsistent solutions in regard to physiology [17]. On the other hand, distributed sources minimum-norm solutions have over-smoothed magnitude patterns and do not respect specific brain anatomical constraints like sulcus borders. In [18], one can find a first attempt to constrain solutions on anatomical surfaces, deduced from a preliminary MRI study of the subject, which greatly improves the quality of the solution. Such an approach has been extensively used in the recent CURRY source reconstruction software in which one has the possibility to constrain dipoles perpendicularly to the cortex surface [19].

What we propose here is to introduce some additional \textit{a priori} information into the regularization process so as to obtain the most probable and satisfying solution in regard to some physiological and anatomical criteria in connection with a given cognitive experiment.
What kind of relevant spatio-temporal a priori information could be introduced?

In this section, we make an overview of the a priori assumptions we make while searching for a physiologically consistent solution to the EEG/MEG inverse problem.

Anatomical constraints

What we propose here is to reduce the dimension of the source space by introducing anatomical constraints on dipole locations and orientations. Such an approach has already been adopted by researchers and is getting widely used [18-20]. The main motivation is to prevent solutions from being found in unrealistic locations (i.e. in the white matter for instance). In our method, we propose an extension of Dale and Sereno's approach which consists in distributing dipoles on the cortex surface owing to high resolution MRI anatomical images. Moreover, we adopt the same restricting hypothesis as in [18] by demanding that dipoles have fixed orientations but changing magnitudes during a given experiment. More precisely, we use the fact that the only neuronal electric activity that can be measured outside the skull by EEG and MEG comes mainly from groups of neurons organized in macrocolumns perpendicularly to the cortex surface [21].

Once these assumptions are made, we build a distributed source model as a patchwork where dipoles are distributed on little surfaces that locally approximate some parts of the cortex surface. These patches are located using previous PET or f-MRI experiments with the same subject, resulting in “average” images of active brain areas during a given task. This means we make the strong assumption that metabolically active areas are closely related to active neuron groups [22]. As an example, Fig. 1 shows active areas isolated from PET data in an auditory experiment [23].

Figure 1 — Cortex patches representing a priori brain active areas isolated owing to anatomical MRI and PET imaging in an auditory experiment. These patches are located in temporal auditory and motor cortex areas. (NI) is the Nasion-Inion line. Black dots on every patch are for dipole locations.
A priori active areas may also be deduced from cognitive considerations, allowing the introduction of fugitive sources in areas that will not appear on any metabolic imaging modality. Here, the source space is designed with “patches” that are tangent to the cortex surface. It is then possible to distribute $N$ dipoles on these planes so as to create the source space. The number of dipoles per plane may vary, and could be proportional to the patch surface for instance. Such a modelization needs an accurate multimodal registration procedure - which we won't describe here - in order to obtain effective data fusion between those different imaging techniques. In addition to those anatomical constraints, we make some further spatio-temporal assumptions on dipole magnitudes distribution and evolution according to physiology.

Physiological constraints

We introduce two supplementary constraints as follows : firstly, one can safely consider that relevant frequencies of neural electrical activity are inferior to $100Hz$, and since measurements are over-sampled, it is assumed that dipole magnitude evolution is smooth in time. In the next section, we will see how this prior can be introduced as some kind of a temporal low-pass filter on $J_n$.

Secondly, although we assume that dipoles on the same surface patch have strongly related activities, it could also be possible that dipoles located on two different surface patches have completely separated activities. For instance, one can suppose that activities of dipoles located on two different walls of a sulcus, in some cases, could be really functionally non-interdependent depending on the nature of physiological constraints.

As a conclusion, these priors lead to the idea that the source space is made of locally homogeneous and smoothly time evolving areas (cortex patches) that can be considered partially or totally independent from one to another depending on physiological aspects. After having created some spatio-temporal neighborhood for each dipole - and when considering dipole magnitudes as pixel intensities of a reconstructed image - assuming non-correlated activities between some dipoles means that discontinuities may exist in the source image despite the fact that most of the inversion techniques produce spatially over-smoothed solutions. This over-smoothing is due to the usual practice to operate some kind of spatial low-pass filtering onto the reconstructed image, in order to make the balance right between algorithmic stability (noise averaging by some spatio-temporal low-pass filter) and
information preservation. As a lot of informative contents usually lie in higher frequency ranges - like plosive consonants in many languages and sharp edges in pictures - one can imagine that a compromise is to be found in some kind of adaptive filtering.

In order to break free from this and to proceed with a *selective image filtering*, we show in the next section that all the previous spatio-temporal *a priori* information can easily be introduced in a Bayesian framework while using a MAP estimator of dipole magnitudes. More particularly, we will see that it is possible to regularize this inverse problem while, at the time, respecting all the anatomo-functional priors, and in particularly preserving selected “edges”.

**A Bayesian framework for spatio-temporal regularization of the EEG / MEG inverse problem**

**MAP estimation**

In this section, we build a MAP estimator of dipole magnitudes which takes all the previous *a priori* information into account. Such a Bayesian technique consists in finding, for a given time sample \( n \), an estimator of current dipole magnitudes \( \hat{J}_n \) that maximizes the posterior probability distribution of \( J_n \) given the measurements \( M_n \). This estimator can be written as:

\[
\hat{J}_n = \max_{J_n} p(J_n|M_n).
\]  

(2)

The MAP estimator of dipolar activity is the most probable one with regards to measurements and *a priori* considerations. Indeed, according to Bayes law we have:

\[
p(J_n|M_n) \propto p(M_n|J_n) \cdot p(J_n),
\]  

(3)

where \( p(M_n|J_n) \) is the likelihood of \( J_n \) given measurements \( M_n \) - in other words, \( p(M_n|J_n) \) contains the forward model between the source and the measurement space and especially noise characteristics - and \( p(J_n) \) is the *a priori* probability distribution of \( J_n \) which gathers the above priors.

The MAP estimate is a popular Bayesian formulation of regularization techniques according to the general scheme of balancing between trust to data and fidelity to priors. In
this scheme, the posterior probability distribution can be written owing to an energy function in terms of $\mathbf{J}_n$ that can be associated to the probability distributions in (3):

$$p(\mathbf{J}_n|\mathbf{M}_n) = \frac{1}{Z} \exp(-U(\mathbf{J}_n)),$$

(4)

$Z$ is a normalization constant called partition function, and $U(\mathbf{J}_n) = U_1(\mathbf{J}_n) + \lambda U_2(\mathbf{J}_n)$ where $U_1(\mathbf{J}_n)$ and $U_2(\mathbf{J}_n)$ are energy functions associated with $p(\mathbf{M}_n|\mathbf{J}_n)$ and $p(\mathbf{J}_n)$ respectively; $\lambda$ is the tuning parameter between the likelihood and the prior term [9]. Then the MAP estimator of dipole magnitudes becomes:

$$\hat{\mathbf{J}}_n = \min_{\mathbf{J}_n} U(\mathbf{J}_n).$$

(5)

If measurement noise is assumed to be white, Gaussian and zero-mean, one can write $U_1(\mathbf{J}_n)$ as:

$$U_1(\mathbf{J}_n) = \lVert \mathbf{M}_n - \mathbf{G} \mathbf{J}_n \rVert^2.$$

(6)

Note that correlation between sensor noise components could easily be introduced. For instance, let $\mathbf{R}$ stands for the noise correlation matrix, then $U_1(\mathbf{J}_n)$ can be written as follows:

$$U_1(\mathbf{J}_n) = (\mathbf{M}_n - \mathbf{G} \mathbf{J}_n)^T \mathbf{R}^{-1} (\mathbf{M}_n - \mathbf{G} \mathbf{J}_n).$$

(1)

As a prior term, $U_2(\mathbf{J}_n)$ is written: $U_2(\mathbf{J}_n) = U_s(\mathbf{J}_n) + U_t(\mathbf{J}_n)$ where $U_s(\mathbf{J}_n)$ introduces spatial priors and $U_t(\mathbf{J}_n)$ temporal ones. We shall first focus on spatial constraints and come back to temporal ones later on.
Spatial constraints

A widely used formalism for the prior spatial energy function $U_s(J_n)$ consists in writing it as a sum of potential functions in terms of dipole magnitudes (and more generally pixel values). Thus, many approaches to spatial regularization may be adopted, depending on the form of the potential $\Phi$-function.

A classical inverse problem regularization using spatial priors derives from [24]. It consists in using the $L_2$-norm of the dipole intensity gradients as a prior term in the $U_s(J_n)$ energy function. This means that the image to be reconstructed presents a priori smooth pixel intensity variations (and consequently sharp edges are not likely). The corresponding $\Phi$-function is quadratic ($\Phi(u) \propto u^2$, dashed curve in Fig. 2), which means that emergence of discontinuities will be severely punished. That is why such a choice for spatial priors yields over-smoothed and blurred reconstruction of the real image, and would not respect our assumptions.

As seen in section 3, we would like the dipole image to be composed of smooth patches possibly separated by discontinuities. Such an image reconstruction problem is strongly linked to many previous efforts in image restoration or reconstruction using Markovian Random Fields (MRF). MRF processes are really convenient to build piecewise constant image models while preserving discontinuities at the same time. In [25], an image is modeled by two Markov processes: the first one is linked to pixel intensities, while the second one is an underpinning binary non-interacting line process. This latter is used to point at the probable locations of intensity jumps and to suspend smoothness constraints (that is to say regularization) in their vicinity. Then, the non-linear energy function associated to the MAP estimator depends on both pixel intensities and the binary line process. So, the spatial term of the prior energy function $U_s(J_n)$ can be written as:

![Figure 2 — Prior potential functions for different regularization methods. The point-dash curve represents the quadratic regularization method. The thick curve is the broken parabola for a threshold at $u = 8$, where $u$ stands for the intensity gradient norm. The other curves are the Lorentzian potential functions used in our method. Note the different shapes for $K = 4$ to $K=10$, where $K$ is a scaling factor depending on whether discontinuities are likely or not.](c:\users\silvin\phi.eps)
\[ U_s(J_n) = \sum_{k=1}^{N} \Phi(\nabla J_{n|k}, l_k), \]  

where \( \nabla J_{n|k} \) denotes the \( k \)th element of spatial gradient vector \( \nabla J_n \) and \( l_k \) stands for the binary line process at point \( k \).

Asymptotic global optimization both on \( J_n \) and \( l \) can be carried out by simulated annealing (SA). Nevertheless, such a method suffers from being very time consuming and practically, convergence is not guaranteed specially for continuous processes with large neighborhood like dipole moments. Indeed, the estimated value of one dipole magnitude depends on many other dipole values because \( G \) has wide support [26]. So, one must seek a deterministic, although non-optimal way to find a MAP estimate.

In [27], Blake and Zisserman propose a broken parabola form for the \( \Phi \)-function which gives rise to limited cost for higher gradient values (see Fig. 2)), and so may allow the restoration of intensity edges in the dipole image with noise sensitivity lowering at the same time. This is the so-called weak membrane model of a pixel field. This was the first attempt to produce an edge-preserving regularization solely on pixel intensities without using an explicit line process \( l \). This method has been used in the EEG/MEG inverse problem in [28] where optimization is carried out by gradient descent. However, one should pay extreme attention to algorithmic stability considerations as the potential \( \Phi \)-function is non differentiable and as \( G \) has wide support as mentioned before [26].

The previous idea has been extended in [29] using non linear functions of the dipole intensity gradients, \( \Phi(\nabla J_n) \). For instance, \( \Phi(u) \) could be of the following Lorentzian form:

\[ \Phi(u) = \frac{u^2}{1 + \left( \frac{u}{K} \right)^2}, \]  

where \( K \) is a scaling factor.

Anatomical priors can now be introduced by locally choosing different scale factors \( K \) in (8), depending on how much a discontinuity is likely between two different cortex patches as seen in section II. That is to say that \( K \)-indexed potential functions \( \Phi_K(u) \) are chosen locally depending on whether they are applied to gradients between a priori “non correlated” dipoles or not [31]. Fig. 2 shows \( \Phi_K \) functions for \( K \) values varying between 4 and 10. One can see
that for every $\Phi_K$, the cost of high intensity gradients in the reconstructed image is limited which allows preservation of discontinuities when necessary.

More precisely, gradient values inferior to the local scaling factor $K$ have a quadratic energetic cost as $\Phi_K(u)$ can be approximated by $u^2$, while strong gradients in regard to $K$ have a finite cost $K^2$. When comparing the regularization processes for two distinct $K$ values, $K_1$ and $K_2$ with $K_2 << K_1$, it is clear that $\Phi_{K_2}(u)$ represents a lower cost than $\Phi_{K_1}(u)$ for every gradient value $u$. So, by locally assigning different values to $K$ ($K=K_1$ or $K=K_2$), it is possible to set the local behavior of the algorithm: spatial variations of dipole moments will be more smoothed in areas where $K_1$ is the scaling factor, than in regions scaled with $K_2$.

In the context of dipole reconstruction, by assigning different values to $K$, among $K_1$ and $K_2$, from one dipole location to another, one can translate the physiological constraints seen above (section III). Then, $U_s(J_n)$ is written as:

$$U_s(J_n) = \sum_{k=1}^{N} \left( \Phi^1_x(\nabla_x J_{n|k}) + \Phi^1_y(\nabla_y J_{n|k}) \right)$$

where $x$ (res. $y$) indexed values are for horizontal (res. vertical) gradients between cortical dipoles along the surface patch.

Temporal constraints

Now, temporal constraints can be introduced by means of another energy function, $U_t(J_n)$. As seen in section II, we make the assumption that dipole magnitudes are slowly evolving with regard to the sampling frequency. The use of non-linear temporal constraints with MRF has already been widely spread by researchers in automatic movement detection [32]. Some previous work in biomedical engineering has shown that the introduction of temporal constraints brought great improvement in electrocardiography source imaging [33].

We propose here a quadratic cost function with a very simple geometric origin. Therefore, we use $\hat{J}_{n-1}$ as unique temporal neighbor of $\hat{J}_n$. Assuming that $\hat{J}_{n-1}$ and $\hat{J}_n$ may be really close to each other means that the orthogonal projection of $\hat{J}_n$ on the hyperplane $E_{\hat{J}_{n-1}}$ perpendicular to $\hat{J}_{n-1}$ is “small”. Thus, $U_t(J_n)$ may be written as:
\[ U_t(\mathbf{J}_n) = \beta \| \mathbf{P}^\perp_{n-1} \mathbf{J}_n \|^2. \]  

(10)

Where \( \mathbf{P}^\perp_{n-1} \) is the projector onto \( E^\perp_{J_{n-1}} \):

\[ \mathbf{P}^\perp_{n-1} = \mathbf{I} - \hat{\mathbf{J}}_{n-1} \hat{\mathbf{J}}_{n-1}^T / \| \hat{\mathbf{J}}_{n-1} \|^2, \]

(11)

with \( \mathbf{I} \) as the \( \mathbb{R}^N \) identity matrix.

As a primary conclusion, we may write that the MAP estimator of dipole moments at time \( n \) is given by minimizing \( U(\mathbf{J}_n) \), with:

\[ U(\mathbf{J}_n) = \| \mathbf{M}_n - \mathbf{G} \mathbf{J}_n \|^2 + \lambda \left( \sum_{k=1}^{N} \left( \Phi^i_x(\nabla_x \mathbf{J}_{n|k}) + \Phi^i_y(\nabla_y \mathbf{J}_{n|k}) \right) + \beta \| \mathbf{P}^\perp_{n-1} \mathbf{J}_n \|^2 \right). \]

(12)

\( \lambda \) is the parameter allowing the weighting of the prior part of the global energy with regard to the data attachment term, while \( \beta \) tunes the contribution of the temporal priors to the global \textit{a priori} term. The values of such parameters depend mostly on the supposed amount of noise in measurements. The lower the signal to noise ratio (SNR) the less the algorithm is allowed to trust data, and at the same time, the more it should pay attention to priors as guiding lines for an acceptable solution.

MAP estimation procedure

Zeroing \( U(\mathbf{J}_n) \) derivatives with respect to every \( \mathbf{J}_n \) components lead, after appropriate arrangement of terms, to the following so-called normal equations:

\[ \left( \mathbf{G}' \mathbf{G} + \lambda \left( \Delta^n + \beta \mathbf{P}^\perp_{n-1} \right) \right) \mathbf{J}_n = \mathbf{G}' \mathbf{M}_n, \]

(13)

where

\[ \Delta^n = -\nabla^i_x \mathbf{B}^n_x \nabla_x - \nabla^i_y \mathbf{B}^n_y \nabla_y, \]

(2)

is a weighted Laplacian, and

\[ \mathbf{B}^n_x = \text{diag} \left[ b^n_{x|k} \right]_{k=1\ldots N} \]

(14)

with

\[ b^n_{x|k} = \frac{\phi^i_x(\nabla_x \mathbf{J}_{n|k})}{2 \nabla_x \mathbf{J}_{n|k}}. \]

(3)
Thus, (13) is a non-linear equation in $J_n$ (because of the $\Delta^n$ dependence on $J_n$). We use a deterministic sub-optimal algorithm inspired by an iterative resolution procedure called ARTUR [34]. The iterative process goes as follows:

1) For the $n^{th}$ time sample, compute the projection matrix $P_{n-1}$. 
2) Estimate scaling parameters $K_1$ and $K_2$. 
3) Then the iterative minimization process goes:

An initial estimate of $J_n$ is computed, we call it $\hat{J}_n^0$ (step $m = 0$). We will come back later to its estimation. At step $m+1$, compute the weighted Laplacian from the $b^n_{x,m}$ and $b^n_{y,m}$ coefficients derived from the preceding value $\hat{J}_n^m$. By positing a new matrix $S_{n+1} = G'G + \lambda_1(\Delta + \beta P_{n-1}P_{n-1}^\perp)$, compute $\hat{J}_n^{m+1}$ such that:

\[
\hat{J}_n^{m+1} = (S_{n+1}^{-1})G'M_n.
\]

This iterative process is repeated until a suitable termination criterion is met. Typically, one can check the relative variation between two following $J_n$ estimates, which we write:

\[
\frac{\|\hat{J}_n^{m+1} - \hat{J}_n^m\|}{\|\hat{J}_n^{m+1}\|} < F_{\text{step}},
\]

where $F_{\text{step}}$ is a constant we set to $10^{-7}$. 

Coming back on choosing the starting point $\hat{J}_n^0$, it is very clear that the closer $\hat{J}_n^0$ is to $J_n$, the better is the $\hat{J}_n$ estimate and the quicker the algorithm will converge. Thus, we recommend as initial conditions:

\[
\begin{align*}
n = 1: & \quad \hat{J}_1^0 = (G'G + \lambda_0 \nabla \nabla)^{-1}G'M_1, \\
n > 1: & \quad \hat{J}_n^0 = \hat{J}_{n-1}.
\end{align*}
\]

Note that $\hat{J}_1^0$ is a MAP estimate of dipole magnitudes with minimum gradient $L_2$-norm priors and for $n > 1$, $\hat{J}_n^0$ follows the assumption of temporal smoothness. One could object that such initial solutions may propagate possible estimation errors due to non optimal minimization from one time sample to another. This is effectively true but from a practical point of view this initialization procedure has given satisfactory results through many simulations using different source configurations and noise levels.

As far as $K$ factors estimation is concerned, note that if $K_l$ is too small in regard to gradients in $J_n$, this will create spurious spikes on cortex patches and so local smoothness
assumptions would not be verified. Similarly, if \( K_2 \) is too large, discontinuities would not be preserved and regularization would produce an over-smoothed image. On the contrary, discontinuities would be systematically created between cortex patches if \( K_2 \) was too small, and this would be detrimental to algorithm stability as it is very likely that the majority of the areas are not active at the same time. That is why \( K_1 \) and \( K_2 \) are computed from the previous estimate \( \hat{J}_{n-1} \) and refreshed for every time sample as:

\[
K_1 = 3 \times \max \left( \left| u^{n-1}_x \right|, \left| u^{n-1}_y \right| ; i = 1...N \right)
\]

\[
K_2 = \frac{1}{8} \times \max \left( \left| u^{n-1}_x \right|, \left| u^{n-1}_y \right| ; i = 1...N \right), \text{where } u^{n-1}_x = \nabla_x \hat{J}_{n-1} \text{ and } u^{n-1}_y = \nabla_y \hat{J}_{n-1} \quad . \quad (16)
\]

One of the main drawbacks of regularization techniques is that they need some well chosen tuning parameters in order to be effective. Such parameters are sometimes called hyperparameters. Some of them, like \( K_1 \) and \( K_2 \) in (16), may effectively be set owing to intuitive considerations because of their direct link with “objective” size orders (like dipole magnitudes). This is no more the case when one has to choose hyperparameters like \( \lambda_0 \) in (15) or \( \lambda \) and \( \beta \) in (13) which are used as weighting factors between two terms. Broadly speaking, as far as global energy functions for MAP estimation are concerned, these coefficients are used to set a balance between a likelihood term - quantifying the faithfulness to raw data - and an a priori term with regards to estimator behavior as we have seen before. Some great efforts are made in order to estimate hyperparameters from data ([29], [35]). A maximum likelihood estimation from data of these parameters may sometimes offer effective results but may also lead to non tractable algorithms. Here, parameters \( \lambda_0, \lambda \) and \( \beta \) are chosen on an empirical basis. That is to say that in the future, efforts should be made to estimate these latter from data.

As a conclusion, we have seen in this section how convenient it is to use a Bayesian framework to build an estimator of dipole activity: spatio-temporal a priori information has been introduced by locally defining potential functions that may vary from one point to another depending on anatomically and physiologically driven discontinuity priors in the dipole magnitudes image. Temporal smoothness was introduced in a very simple way by means of geometrical similarity between two successive dipolar activities. MAP estimation is achieved by a deterministic algorithm. This non optimal procedure may fall into local minima of the highly non convex global energy function but is really easy to implement, and moreover has given good results.

We now present briefly the quadratic regularization and LORETA methods.
The Quadratic Regularization and LORETA methods

Quadratic regularization

Such an approach can be formulated in a Bayesian framework too. Thus, if we write the posterior distribution energy of $J_n$ with regards to measurements $M_n$ we have:

$$U_{QR}(J_n) = \|M - G J_n\|^2 + \lambda \|\nabla J_n\|^2.$$

(6)

Here, quadratic priors mean that the norm of the MAP estimate deviation from data should be as small as possible, and the second term enforces smoothness. Looking for a stationary point of $U_{QR}(J_n)$ means that $J_n^{QR}$ checks the following normal equation:

$$G^t G J_n^{QR} + \nabla^t \nabla = \lambda,$$

(7)

where one can note that $\nabla^t \nabla$ approximates a 1D Laplacian and finally,

$$J_n^{QR} = \left(G^t G + \lambda \nabla^t \nabla\right)^{-1} G^t M_n.$$

(8)

LORETA

While using generic minimum-norm methods, it has been shown that deeper sources could not be recovered because dipoles located at the surface of the source space with smaller magnitudes would be privileged for the same EEG/MEG data set. That is the reason why one could think of some kind of lead field normalization in order to give all the sources, close to the surface and deeper ones, the same opportunity of being nicely reconstructed by a minimum-norm technique. LORETA is a recent method that proceeds in this way.

LORETA is “true 3-dimensional Tomography (...), which has a relatively low spatial resolution” [13]. The inverse solution given by LORETA “corresponds to the smoothest current density capable of explaining the measured data”. This method has given nice results in noiseless simulations, but also in the analysis of some real data sets.

Basically, LORETA gives a minimum-norm restricted solution of

$$M_n = G J_n.$$

(9)

Nevertheless and strictly speaking, one should care about the presence of additive noise in $M_n$ in realistic situations, and should preferably seek a restricted least square solution of the linear
The LORETA method goes like this: first build up the non negative definite matrix $\tilde{N} = WB'BW$; where $B$ corresponds to a discrete spatial Laplacian, and $W$ is a lead field normalization matrix. While LORETA is a full 3D reconstruction method, it is also possible to restrict the source space dimension. Here, we will use the same cortex patchwork restriction as before. This is a restricted 3D reconstruction as dipoles are constrained on the cortex surface. But LORETA allows such a procedure as it is explicitly described in [13]. So, in this pseudo 2D case, a discrete Laplacian is computed over the 4 closest neighbors of a given dipole. The $W$ matrix is a diagonal $N \times N$ matrix:

$$W = \text{diag}(w_i; i = 1, \ldots, N),$$

with

$$w_i = \|G(i)\|,$$

where $G(i)$ is the $i^{th}$ column of the $G$ matrix.

The LORETA procedure consists in finding the minimum $\tilde{N}$ norm solution $\hat{J}_n^L$ to (12), which writes:

$$\hat{J}_n^L = \min_{J_n} \|J_n^t\tilde{N}J_n\| \text{ under constraint } M_n = GJ_n.$$

The $w_i$ coefficients make up the lead field normalization in order to overcome the systematic deeper dipoles zeroing of minimum-norm solutions as seen above.

By choosing “a dense grid” of sources (e.g. $N > N_M$ but Pascual et al. recommend $N >> N_M$), a simplified solution to (14) exists and is unique. It is given by:

$$\hat{J}_n^L = \tilde{N}^{-1}G^t(GN^{-1}G^t)^+M_n,$$

where $A^+$ denotes the Moore-Penrose (MP) pseudoinverse of matrix $A$ [38]. This method does not explicitly refer to regularization but nevertheless, one can consider that the introduction of the MP pseudoinverse, yielding a SVD of the matrix to be inverted, is very similar to the approach found in [15] and [16]. Anyway, it appears that this MP pseudoinverse is of great practical importance because of the bad conditioning of the matrix to be inverted.

We will now present some simulation results from MAP estimations and compare them to those given by quadratic regularization and by LORETA.
Results

In this section, results from Bayesian estimation of dipole activities in an EEG experiment simulation are shown. We will see how the introduction of temporal constraints greatly improves the quality of the results with regards to spatial edge preserving regularization only. We will compare these outcomes with the QR and the LORETA methods.

Simulation framework

We have chosen to isolate 16 *a priori* active areas in a spherical 3 shell head model from the previous auditory experiment data. Table 1 gathers information about the different media of this head modeling [37]. The corresponding cortex patchwork is shown on Fig. 1. 128 dipoles (8 dipoles per patch) are spread perpendicularly over these surfaces. For simplicity, we have chosen to put the same number of dipoles on every patch, and the EEG recording is calculated on a 65 electrodes set.

<table>
<thead>
<tr>
<th></th>
<th>Radius (cm)</th>
<th>Conductivity ($\Omega^{-1}.m^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brain</strong></td>
<td>10.7</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Skull</strong></td>
<td>11.3</td>
<td>0.0042</td>
</tr>
<tr>
<td><strong>Skin</strong></td>
<td>12.2</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*Table 1 — The 3 shell spherical head model characteristics used in the simulation process.*

The amount of additive white Gaussian noise is controlled via SNR defined as follows:

\[
\text{SNR} = 20\log\left(\frac{\sigma_{\mathbf{M}_n}}{\sigma_{\mathbf{b}_n}}\right) \text{ (dB)},
\]

where $\mathbf{M}_n = \mathbf{G}\mathbf{J}_n$, and $\sigma_{\mathbf{M}_n}$ (res. $\sigma_{\mathbf{b}_n}$) is $\mathbf{M}_n$ (res. $\mathbf{b}_n$) standard deviation. SNR is set to 20dB for every experiment. Then, a very simple activation sequence for dipole magnitudes is designed by means of a damped sine curve (Fig. 3c) in order to obtain evoked potentials like EEG measurements. This sequence will be distributed through active dipoles with different scale ratios and delays. So, we have designed a simulation sequence using 4 active homogeneous patches with delayed source activities. Dipoles on the other 12 planes are set to zero. Fig. 3a is a convenient 2D way of looking at these data. The 128 dipoles are arranged in
a row vector and the x-axis (scaled from 1 to 16) represents the patches on which they are distributed. Time samples are on the y-axis and dipole magnitudes are displayed owing to a gray-scaled colormap. The whole simulated activation sequence goes like this:

- both planes number 13 and 15 start at time zero, with an arbitrary normalized scale (maximum magnitude is 1),
- plane number 14 is then activated with a 20 samples time delay,
- then finally, plan 10 starts at time 40, with a scaling ratio of 1 to 5 (maximum amplitude is 0.2).

While there are 200 time samples, Fig. 3b shows a 3D meshing of this dipolar spatio-temporal mapping from time sample 20 to 80.

Such a simulation may appear too unrealistic, and it surely is. But since the electromagnetic brain imaging inverse problem is dramatically ill-posed, this simulation does show how effective it might be to introduce well-chosen priors to counter EEG blurring tendencies.
QR results

Results from the EEG experiment simulation are shown in Fig. 4. One can see that the reconstructed dipole activity is over-smoothed over the whole set of dipoles that have contributions to potentials close to active ones in the simulated sequence. On Fig. 4d - showing dipole contours at time sample 90 - we can see for instance how dipoles that are silent in the original activation sequence now have parasitic magnitudes because of the smoothing behavior of this quadratic regularization.

Figure 4 — Results of the QR method.
(a) a grayscale representation of this treatment.
(b) a 3D meshing of the result between time samples 20 and 80.
(c) mean dipole time course on plan 10.
(d) dipole profile at time sample 90.

An example of active dipole time course on patch 10 is on Fig. 4c. Finally, this blurred source reconstruction from blurred and noisy data appears very clearly on Fig. 4a and Fig. 4b where one can check that original edges are heavily smoothed.
LORETA results

Results are shown in Fig. 5. The global activity shape is still smooth, but to a lesser extent than for the previous method. This appears more clearly on Fig. 5d: some edges are better reconstructed than before and null dipoles are almost well recovered. But at the same time spurious spikes appear. These are due to the high noise sensitivity on data of the solution descended from the associated algebraic manipulation, that makes the assumption that measurements should be noiseless (see (28)). Moreover, this noise sensitivity is enhanced because of the pseudoinverse method, which is a harsh way to regularize the inverse problem.

MAP estimation results

Here in our MAP estimate source patches are considered as non correlated with each other. So we suppose that there may be some sharp intensity edges between two neighboring
source planes; that is to say that border gradients are scaled with a $K_2$ ratio, while the ones between two dipoles on the same plane are scaled with a $K_1$ ratio (remember that $K_1 > K_2$).

Spatial regularization (S-MAP)

Without considering any temporal constraints for reconstruction, results are shown on Fig. 6. Most edges are nicely recovered and there is very little spurious activity on silent dipoles. Fig. 6c shows the temporal profile of active dipoles. It is very clear that while the global temporal shape has been preserved, sharp magnitudes variations sometimes occur (see for instance between time samples 50 and 100). The entire spatial dipole profile at time 90 is shown on Fig. 6d: the main edges are recovered but the algorithm failed to display edges between planes 13 to 15.

---

Figure 6 — Results of the S-MAP method.
(a) a grayscale representation of this treatment.
(b) a 3D meshing of the result between time samples 20 and 80.
(c) mean dipole time course on plan 10.
(d) dipole profile at time sample 90.

One important point is the time consumption of such an iterative algorithm. While the mean iteration number per time sample is about 6, the algorithm might run up to about 34
steps to find a “stationary” solution. Very often, this occurs when our algorithm fails to find a solution close to the one found for the previous time sample, and falls into another local minimum, “qualitatively” very far from the previous one. The whole 200 time samples treatment takes about 4500 seconds on a Pentium 90 computer.

Spatio-temporal regularization (ST-MAP)

Results obtained when adding temporal constraints to spatial ones are shown in Fig.7. They may appear qualitatively very similar to those given by the S-MAP. Looking closer to active dipole time profile (Fig. 7c), it appears that the temporal evolution of dipoles is smoother. Moreover, all edges are almost perfectly recovered (see for instance Fig. 7b), and the mean number of iteration falls to 4 per time sample with a much smaller standard deviation from one time sample to another.

Figure 7 — Results of the ST-MAP method.
(a) a grayscale representation of this treatment.
(b) a 3D meshing of the result between time samples 20 and 80.
(c) mean dipole time course on plan 10.
(d) dipole profile at time sample 90.
Thus, the total computation time is reduced to 3500 seconds, which is about 22% less than with spatial regularization only. This shows that well chosen priors, such as temporal constraints in addition to spatial ones, greatly improve the quality of the solution. And at the same time, they tend to help stabilizing the algorithm so that it seeks a solution in the vicinity of the ones found for previous time samples, and finally computing time is decreased.

Primary conclusions

Temporal constraints play a major role in both qualitative and quantitative enhancement of reconstruction results with regards to spatial priors only. Despite heavy blurring on potentials, it is now possible to recover sharp edges in the source space. Quadratic regularization of course fails to recover these discontinuities and introduces some spurious activity on the whole source space. LORETA brings some improvement to this, and may offer satisfactory results as an initialization procedure for the MAP algorithm.

Finally, one should pay very much attention to the way a dipole activation sequence might be considered as acceptable or not. Indeed, computing a fit to data may lead to very serious mistakes and it should not be considered as the only way of trusting the reconstructed sources. As displayed in Table 2, poor solutions may offer excellent fit to data. This is especially the case when choosing a very small regularization parameter in the QR method: source reconstruction is made using solely noisy data, so it yields unacceptable solutions but with a perfect fit to data.

One should remember that the EEG/MEG inverse problem is dramatically ill-posed and that trusting data too much will unquestionably lead to nonsense solutions.

<table>
<thead>
<tr>
<th></th>
<th>Data fit</th>
<th>Reconstruction error</th>
</tr>
</thead>
<tbody>
<tr>
<td>QR - ( \lambda_0 = 0.5 )</td>
<td>almost 100 %</td>
<td>&gt; &gt; 100 %</td>
</tr>
<tr>
<td>QR - ( \lambda_0 = 2000 )</td>
<td>97.6 %</td>
<td>27.5 %</td>
</tr>
<tr>
<td>LORETA</td>
<td>97.8 %</td>
<td>36.7 %</td>
</tr>
<tr>
<td>S-MAP - ( \lambda = 3 \times 10^5 ), ( \beta = 0 )</td>
<td>97.3 %</td>
<td>11 %</td>
</tr>
<tr>
<td>ST-MAP - ( \lambda = 3 \times 10^5 ), ( \beta = 0.5 )</td>
<td>97.8 %</td>
<td>6.3 %</td>
</tr>
</tbody>
</table>

Table 2 — Data fit compared to error on dipole magnitudes for different regularization methods. Data fit (res. reconstruction error) stands for the mean of the following relative error \( 100 \times \frac{\|M_n - GJ_n\|}{\|M_n\|} \) over all the time samples.
Conclusions and Final Remarks

As minimum-norm solutions to electromagnetic brain imaging give rise to spatially over-smoothed and sometimes unrealistic source configurations, one might address a new and more sophisticated way to solve this inverse problem. It could be considered as very intuitive to introduce some *a priori* information in the inversion procedure to help finding anatomically and physiologically consistent source sets.

Such an approach has been developed in this paper in a Bayesian framework. It consists in building a MAP estimate of dipolar activity that takes into account spatial and temporal priors. The use of MRF models for dipolar spatio-temporal landscapes is very convenient and allows the recovery of intensity discontinuities despite heavy data blurring and noise. Thus, final dipole images are made of small cortex patches that are considered as independent from each other whenever necessary. This method brings significant improvements to the spatial resolution of EEG/MEG compared to minimum-norm solutions with a distributed source model. Then, note that a minimum-norm solution such as LORETA's would certainly be of particular help to initiate the iterative minimization process of MAP estimation.

But as this article was aimed at the presentation of a reconstruction method, it is now important to stress on some remaining practical problems. For the treatment of real data sets, realistic head models should be used in EEG, together with an effective fusion procedure between MRI images - so as to draw a reliable and fine cortex patchwork - and another functional imaging modality to isolate very probable active brain areas. These problems give rise to very great efforts through the whole EEG/MEG imaging community and improvements are constantly emerging.

Finally, since high temporal resolution is the main advantage of EEG and MEG, it is likely that a better use of the whole data set as global information about source evolution and dynamics would bring crucial help to distributed dipoles reconstruction methods.
ACKNOWLEDGEMENT

We gratefully acknowledge Pierre Chavel, Director of the Physique des Images department in the Institut d’Optique Théorique et Appliquée in Orsay, for his constant support and his precious help with the manuscript; Bernard Renault, Director of the LENA lab in Paris, for having initiated EEG and MEG source reconstruction in LENA; Francisco Varela, Jacques Martinerie and Jean-Philippe Lachaux at LENA for many fruitful discussions on the material presented here. We would also like to thank Cécile Mico and Gildas Marin at IOTA for their contribution to the treatment of cortical images. Very special thanks to Robert Fripp.
REFERENCES


