



Computational Neuroscience

DYNAMO: Concurrent dynamic multi-model source localization method for EEG and/or MEG

Javier M. Antelis*, Javier Minguez

Computer Science and Systems Engineering Department I3A, University of Zaragoza, Spain

HIGHLIGHTS

- ▶ Dynamic multi-model solution for the EEG/MEG source localization problem.
- ▶ A source model is a dynamic system, allowing the recursive estimation of the sources within the filter estimation framework.
- ▶ Simultaneous integration of different dynamic neural models with different parameters.
- ▶ The method does not require a prior definition neither of the number of sources or of the initial prior estimates.

ARTICLE INFO

Article history:

Received 6 May 2012

Received in revised form

17 September 2012

Accepted 17 September 2012

Keywords:

EEG

MEG

Source localization

Multi-model

IMM

ABSTRACT

This work presents a new dipolar method to estimate the neural sources from separate or combined EEG and MEG data. The novelty lies in the simultaneous estimation and integration of neural sources from different dynamic models with different parameters, leading to a dynamic multi-model solution for the EEG/MEG source localization problem. The first key aspect of this method is defining the source model as a dipolar dynamic system, which allows for the estimation of the probability distribution of the sources within the Bayesian filter estimation framework. A second important aspect is the consideration of several banks of filters that simultaneously estimate and integrate the neural sources of different models. A third relevant aspect is that the final probability estimate is a result of the probabilistic integration of the neural sources of numerous models. Such characteristics lead to a new approach that does not require a prior definition neither of the number of sources or of the underlying temporal dynamics, allowing for the specification of multiple initial prior estimates. The method was validated by three sensor modalities with simulated data designed to impose difficult estimation situations, and with real EEG data recorded in a feedback error-related potential paradigm. On the basis of these evaluations, the method was able to localize the sources with high accuracy.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Electrical and chemical activity of neurons in the brain results in extracellular currents, which generate electric potentials and magnetic fields that can be measured on the head surface through electroencephalography (EEG) and magnetoencephalography (MEG). EEG or MEG sensors measure the electrical or magnetic mixture of the temporal activity of groups of neurons in very different and separate brain regions. Therefore, clinical and functional interpretation of EEG and MEG signals entails the speculation of the active brain areas generating such signals. This speculation leads to the so-called EEG/MEG source localization problem: how to estimate the neural sources that generate the EEG/MEG signals (Baillet et al., 2001). The solution to the EEG/MEG

source localization problem is a challenge due to its inherent ill-definition: an infinite number of configurations of neural sources could explain the same EEG/MEG observations (Grech et al., 2007; Michel et al., 2004; Koles, 1998). Many studies have been carried out (see Baillet et al., 2001; Grech et al., 2007 for reviews) leading to two types of solutions: (i) dipolar methods, where neural sources are represented by low dimensional spaces; and (ii) imaging or distributed methods, where neural sources are represented by high dimensional spaces.

Most research has been directed towards distributed methods, since the problem becomes linear and can be addressed by non-parametric solutions (Hämäläinen and Ilmoniemi, 1994; Pascual-Marqui et al., 1994; Liu et al., 2002; Darvas et al., 2004) and by adaptive spatial filter solutions (Veen et al., 1997; Sekihara et al., 2005). These solutions can be readily interpreted as 3D images of brain activity. Few research has been directed towards dipolar methods due to the non-linear nature of the problem (addressed by parametric solutions with non-linear optimization algorithms;

* Corresponding author. Tel.: +34 976 762472; fax: +34 976 761914.

E-mail address: antelis@unizar.es (J.M. Antelis).

Mosher et al., 1992; Scherg, 1990) and dependence on the initial user specifications, such as the number of sources and initial conditions. However, dipolar methods provide a direct and easy-to-interpret mapping from the recorded EEG/MEG signals into a small set of source parameters, and are well suited for cases when the brain activity is focalized. The work presented herein is in line with dipolar methods.

Usual shortcomings of dipolar techniques are: (i) computation towards a solution that represents a unique hypothesis of neural sources, although many hypotheses could explain the same EEG/MEG signals; (ii) the dynamics of the neural sources are usually not addressed, when EEG/MEG signals are generated by neural sources with underlying time-variable dynamics; and (iii) no changes are allowed in the dimension of the solution space as a fixed number of sources is adjusted in advance, however, depending on the neural processes, the number of neural sources could change with time. Therefore, the estimation of EEG/MEG neural sources might be addressed as a multihypothesis, time-varying dimension and dynamic estimation problem. The present work takes existing studies a step further by considering a unified multi-model Bayesian framework to address the estimation of EEG/MEG neural sources.

Bayesian formalism has been used to address some of the aforementioned issues individually, for instance, to incorporate spatial and/or temporal prior information that constrains the solution, mainly in distributed methods (Baillet and Garnero, 1997; Schmidt et al., 1999; Trujillo-Barreto et al., 2008; Wipf and Nagarajan, 2009) and to a minor degree, in dipolar methods (Bertrand et al., 2001; Jun et al., 2006; Kiebel et al., 2008). Bayesian formalism has also been used to take into account specific dynamics for the neural sources, where Kalman filters have been used in distributed methods (Yamashita et al., 2004; Galka et al., 2004; Barton et al., 2009) and sequential Monte Carlo techniques in dipolar methods (Somersalo et al., 2003; Sorrentino et al., 2009). The idea of a multi-model approach has been applied in other works in terms of different anatomical brain compartments (Trujillo-Barreto et al., 2004) or in terms of different priors that constrain the solution (Mattout et al., 2006; Kiebel et al., 2008).

This paper proposes *Dynamo*, a framework for the estimation of neural sources based on the Integrating Multiple Model (Bar-Shalom et al., 2001) and on the Novel-Interacting Multiple Model (Quan et al., 2009) algorithms and that relies on the probabilistic integration of multiple dynamic dipolar models. Model estimation is recursively computed within a Bayesian filter framework, which is probabilistically integrated at each time to build the final estimation. The performance of the method was evaluated using simulated data (in three sensor modalities) as well as real EEG signals, and compared with other widely spread source localization techniques. The paper is organized as follows: Section 2 describes the dynamic EEG/MEG source localization problem, introduces a single dynamic source model and its Bayesian filter solution framework, defines the proposed dynamic multi-model source localization technique, and finally describes the evaluation methodology followed in both simulated and real scenarios. The performance of the technique and comparison with other source localization methods is presented in Section 3. Section 4 analyzes the results of the evaluations and the main properties of the *Dynamo* technique.

2. Methods and materials

2.1. Problem definition and models

2.1.1. Problem definition

Let the EEG and MEG measurements at time t be $\Phi_t = [\phi_t^1 \dots \phi_t^{N_\phi}]^T$ and $\mathbf{b}_t = [b_t^1 \dots b_t^{N_b}]^T$, where N_ϕ and N_b are the

number of electric and magnetic sensors on the head surface. Let the measurement vector be $\mathbf{m}_t = [\Phi_t, \mathbf{b}_t]^T$, with the assumption that there are N_t neural sources located in the brain. Each source is modeled as a current dipole, characterized by a position $\mathbf{r}_q \in \mathbb{R}^3$ and moment $\mathbf{q} \in \mathbb{R}^3$. The state of neural sources at time t is $\mathbf{x}_t = [\mathbf{r}_q^1, \mathbf{q}^1 \dots \mathbf{r}_q^{N_t}, \mathbf{q}^{N_t}]^T$, where $\mathbf{x}_t \in \mathbb{R}^{6 \cdot N_t}$. The aim is to estimate the neural sources \mathbf{x}_t at time t given: (i) the previous neural sources up to time $t-1$, $\mathbf{x}_{0:t-1}$, and (ii) the EEG and MEG measurements up to time t , $\mathbf{m}_{1:t}$. There are two physical models involved in this problem: (i) a transition model for EEG/MEG sources, specified in terms of a state space equation; and (ii) a measurement model, specified in terms of an observer equation.

2.1.2. Transition model: dynamics of EEG/MEG neural sources

This paper is in line with previous works in the context of dynamically-distributed source models (Yamashita et al., 2004; Galka et al., 2004; Barton et al., 2009) and dynamic dipolar source models (Somersalo et al., 2003; Sorrentino et al., 2007, 2008), which use a multivariate autoregressive source model for the source state:

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{t-T}) + \mathbf{w}_t \quad (1)$$

where $f(\cdot)$ is possibly a nonlinear function, T is the regression order and \mathbf{w}_t is the process noise. This model agrees with the common assumption that the brain activity follows a transition pattern towards neighboring areas. As in previous works (Somersalo et al., 2003; Sorrentino et al., 2007, 2008), this model is simplified by assuming a first order multivariate autoregressive linear model independent of time evolution. This leads to a transition model also referred to as linear random walk:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t \quad (2)$$

where \mathbf{w}_t is a Gaussian random vector with $p(\mathbf{w}) \sim \mathcal{N}(0, \mathbf{Q})$. This transition model is characterized by the noise covariance matrix \mathbf{Q} , which is a diagonal matrix whose elements represent the uncertainty in the position and moment of the sources due to the non-modeled real dynamics. Certainly, the real dynamics of neural sources are unknown and could be more complex than the random walk. The advantage of this model is the simplicity and generality when modeling specific dynamics.

2.1.3. Measurement model: the EEG/MEG forward problem

Given the neural sources, the EEG/MEG forward problem allows for the computation of electric potentials and magnetic fields on the head surface. This paper follows the classical forward solutions used to calculate EEG or MEG signals on the head surface (Mosher et al., 1999; Darvas et al., 2004):

$$\mathbf{m}_t = h(\mathbf{x}_t) + \mathbf{v}_t \quad (3)$$

$$\mathbf{m}_t = \begin{bmatrix} \Phi_t \\ \mathbf{b}_t \end{bmatrix} = \begin{bmatrix} h^{eg}(\mathbf{x}_t) \\ h^{meg}(\mathbf{x}_t) \end{bmatrix} + \begin{bmatrix} \mathbf{v}_t^{eg} \\ \mathbf{v}_t^{meg} \end{bmatrix} \quad (4)$$

where $h^{eg}(\cdot)$ and $h^{meg}(\cdot)$ are the functions that produce the electric Φ_t and magnetic \mathbf{b}_t measurements given the state \mathbf{x}_t of the neural sources, and \mathbf{v}_t^{eg} and \mathbf{v}_t^{meg} are the noises in the electric and magnetic observations, respectively. It must be noted that $h^{eg}(\cdot)$ and $h^{meg}(\cdot)$ are both nonlinear functions with respect to the position of the neural sources \mathbf{r}_q , but linear with respect to the moment \mathbf{q} . This formulation holds for both spherical and realistic head models. This paper adopts the three-shell spherical volume conductor model for EEG and the single-shell volume conductor model for MEG, with analytical equations given by Mosher et al. (1999):

$$\phi_t = h^{eg}(\mathbf{r}_\phi, \mathbf{r}_q, \mathbf{q}) = k_3^{eg}(\mathbf{r}_\phi, \mathbf{r}_q) \cdot \mathbf{q} \quad (5)$$

$$b_t = h^{meg}(\mathbf{r}_b, \mathbf{r}_q, \mathbf{q}) = k_1^{meg}(\mathbf{r}_b, \mathbf{r}_q) \cdot \mathbf{q} \quad (6)$$

where $k_3^{eg}(\mathbf{r}_\phi, \mathbf{r}_q)$ is the field kernel for a three-shell volume conductor model for EEG, $k_1^{meg}(\mathbf{r}_b, \mathbf{r}_q)$ is the field kernel for a one-shell volume conductor model for MEG, and \mathbf{r}_ϕ and \mathbf{r}_b represent the positions of the electric and magnetic sensors, respectively.

Assuming that the noises in the electric and magnetic observations are Gaussian and uncorrelated, thus $p(\mathbf{v}^{eg}) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{eg})$ and $p(\mathbf{v}^{meg}) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{meg})$ yield $p(\mathbf{v}) \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, with

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}^{eg} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{meg} \end{bmatrix} \quad (7)$$

The noise covariance matrix \mathbf{R} quantifies the uncertainty degree of the recorded EEG/MEG signals, i.e., the noise level measurement. As there is no a priori noise information for each sensor, the noise covariance matrices are estimated as $\mathbf{R}^{eg} = \Phi_{1:t} \Phi_{1:t}^T$ and $\mathbf{R}^{meg} = \mathbf{b}_{1:t} \mathbf{b}_{1:t}^T$, where each diagonal element represents the sensor variances. Note that Eqs. (4) and (7) hold for EEG or MEG modalities, independently.

2.2. Bayesian estimation for a dynamic source model

2.2.1. Source model

A source model is defined by the source state \mathbf{x}_t with $(\mathbf{Q}, \{\mathbf{x}, \mathbf{P}\}_{0|0})$ parameters, where \mathbf{Q} denotes a given uncertainty for the dynamics of the sources contained in the model and $\{\mathbf{x}, \mathbf{P}\}_{0|0}$ denotes the initial source state and its associated covariance. Note that the dimension of the source space for the model is $\dim(\mathbf{x}_t)$ and the number of sources or dipoles that constitute the model is $N_t = \dim(\mathbf{x}_t)/6$.

2.2.2. Bayesian estimation framework

The objective is to estimate the state of sources \mathbf{x}_t , given the previous states $\mathbf{x}_{0:t-1}$ and the electric and magnetic observations $\mathbf{m}_{1:t}$. In the Bayesian framework, the objective is to compute recursively the posterior density $p(\mathbf{x}_t | \mathbf{m}_{1:t})$ that represents the probability distribution of the neural sources, given the electric and magnetic measurements. $\mathbf{x}_{0:t}$ and $\mathbf{m}_{1:t}$ are assumed to be stochastic and Markov processes. The Bayes theorem is applied,

$$p(\mathbf{x}_t | \mathbf{m}_{1:t}) = \frac{p(\mathbf{m}_t | \mathbf{x}_t, \mathbf{m}_{1:t-1}) p(\mathbf{x}_t | \mathbf{m}_{1:t-1})}{p(\mathbf{m}_t | \mathbf{m}_{1:t-1})} \quad (8)$$

as well as the Markovian properties, to the process and measurements:

$$p(\mathbf{x}_t | \mathbf{m}_{1:t}) = \eta \cdot p(\mathbf{m}_t | \mathbf{x}_t) \cdot \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{m}_{1:t-1}) d\mathbf{x}_{t-1} \quad (9)$$

which is called the Bayes filter (Arulampalam et al., 2002). In this equation, $\eta = 1/p(\mathbf{m}_t | \mathbf{m}_{1:t-1})$ is a normalizing factor, $p(\mathbf{m}_t | \mathbf{x}_t)$ is the likelihood function defined by the measurement model, $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ is the probabilistic transition model, $p(\mathbf{x}_{t-1} | \mathbf{m}_{1:t-1})$ is the prior density representing the prior information on neural sources and $p(\mathbf{x}_t | \mathbf{m}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{m}_{1:t-1}) d\mathbf{x}_{t-1}$ is the prediction density.

The solution of this probabilistic recursion is conditioned by the transition and measurement models and by the type of noise in the process and measurements. In the case of linear and Gaussian systems, the optimal solution is the Kalman Filter (Welch and Bishop, 2006). For nonlinear and/or non-Gaussian systems, the solution is computed through either, local linearizations of the non-linear functions, e.g., the Extended Kalman Filter (Welch and Bishop, 2006), or by numerical approximation techniques such as those based on sequential Monte Carlo methods (Arulampalam et al., 2002; Doucet et al., 2001, 2000).

2.2.3. Extended Kalman Filter (EKF) estimation framework for a source model

In the EKF framework the solution equations are as follows:

$$\mathbf{x}_{t|t-1} = \mathbf{x}_{t-1|t-1}; \mathbf{P}_{t|t-1} = \mathbf{P}_{t-1|t-1} + \mathbf{Q} \quad (10)$$

$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t-1} + \mathbf{K} (\mathbf{m}_t - h(\mathbf{x}_{t|t-1})); \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K} \mathbf{S} \mathbf{P}_{t|t-1} \quad (11)$$

where $\mathbf{x}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$ are the mean and covariance of the prediction density, $\mathbf{x}_{t|t}$ and $\mathbf{P}_{t|t}$ are the mean and covariance of the posterior density (i.e., the final solution), $\mathbf{K} = \mathbf{P}_{t|t-1} \mathbf{H}^T \mathbf{S}^{-1}$ is the Kalman gain, $\mathbf{S} = \mathbf{H} \mathbf{P}_{t|t-1} \mathbf{H}^T + \mathbf{R}$ is the innovations covariance, and $\mathbf{H} = \partial h(\mathbf{x}_t) / \partial \mathbf{x}_{t|t-1}$ is the first term of the Taylor expansion of the nonlinear measurement model:

$$\frac{\partial h^{eg}(\mathbf{r}_\phi, \mathbf{r}_q, \mathbf{q})}{\partial \mathbf{r}_{q(x,y,z)}} = \frac{\partial k_3^{eg}(\mathbf{r}_\phi, \mathbf{r}_q)}{\partial \mathbf{r}_{q(x,y,z)}} \cdot \mathbf{q} \quad (12)$$

$$\frac{\partial h^{eg}(\mathbf{r}_\phi, \mathbf{r}_q, \mathbf{q})}{\partial \mathbf{q}_{(x,y,z)}} = k_3^{eg}(\mathbf{r}_\phi, \mathbf{r}_q) \cdot \frac{\partial \mathbf{q}}{\partial \mathbf{q}_{(x,y,z)}} \quad (13)$$

$$\frac{\partial h^{meg}(\mathbf{r}_b, \mathbf{r}_q, \mathbf{q})}{\partial \mathbf{r}_{q(x,y,z)}} = \frac{\partial k_1^{meg}(\mathbf{r}_b, \mathbf{r}_q)}{\partial \mathbf{r}_{q(x,y,z)}} \cdot \mathbf{q} \quad (14)$$

$$\frac{\partial h^{meg}(\mathbf{r}_b, \mathbf{r}_q, \mathbf{q})}{\partial \mathbf{q}_{(x,y,z)}} = k_1^{meg}(\mathbf{r}_b, \mathbf{r}_q) \cdot \frac{\partial \mathbf{q}}{\partial \mathbf{q}_{(x,y,z)}} \quad (15)$$

This EKF solution recursively estimates the neural sources from EEG/MEG data of a source model (with parameters: number of sources N , the covariance \mathbf{Q} , and initial prior estimate $\{\mathbf{x}, \mathbf{P}\}_{0|0}$).

2.2.4. Limitations of the framework

Dipolar source localization methods (Mosher et al., 1992; Scherg, 1990; Antelis and Minguez, 2009a,b) among others) as well as other frameworks that consider a unique source model, share three intrinsic limitations. Firstly, the number of neural sources or dipoles N is fixed. Therefore the number of neural sources could be underestimated or overestimated, accompanying the temporal increase or decrease in the real number of active brain areas (time-varying dimension). This limitation is also common to all dipolar techniques where the number of neural sources has to be defined a priori, leading to a fixed-dimensional solution. Secondly, static and unique neural dynamics is taken into account (represented in covariance \mathbf{Q}). The properties of the estimation are good as long as the neural process obeys the neural dynamics considered, and the dynamics do not change with time. This limitation is present in the previous EKF framework and in all existing techniques. Thirdly, a prior estimation of the neural sources $\{\mathbf{x}, \mathbf{P}\}_{0|0}$ or initial condition has to be provided. This initial estimate could have a large effect when the dimension of the source space increases. Although the initial estimate could be pre-estimated with a standard dipolar solution, this is again restricted to the pre-selected number of sources and this would only reflect a unique hypothesis of the neural sources.

In order to overcome these three limitations, the next section describes the *Dynamo* method, which addresses the time-varying dimensionality, the consideration of different neural dynamics, and the selection of the initial condition, yielding a dynamic multi-model estimation framework.

2.3. Bayesian estimation for multiple dynamic source models

Dynamo is based on the Integrating Multiple Model (IMM) algorithm (Bar-Shalom et al., 2001) and on the Novel-Interacting Multiple Model (Novel-IMM) algorithm (Quan et al., 2009). The method consists of a bank of Extended Kalman Filters running in parallel, which operate over models with the same dimension of the source space (same number of dipoles) but different parameters

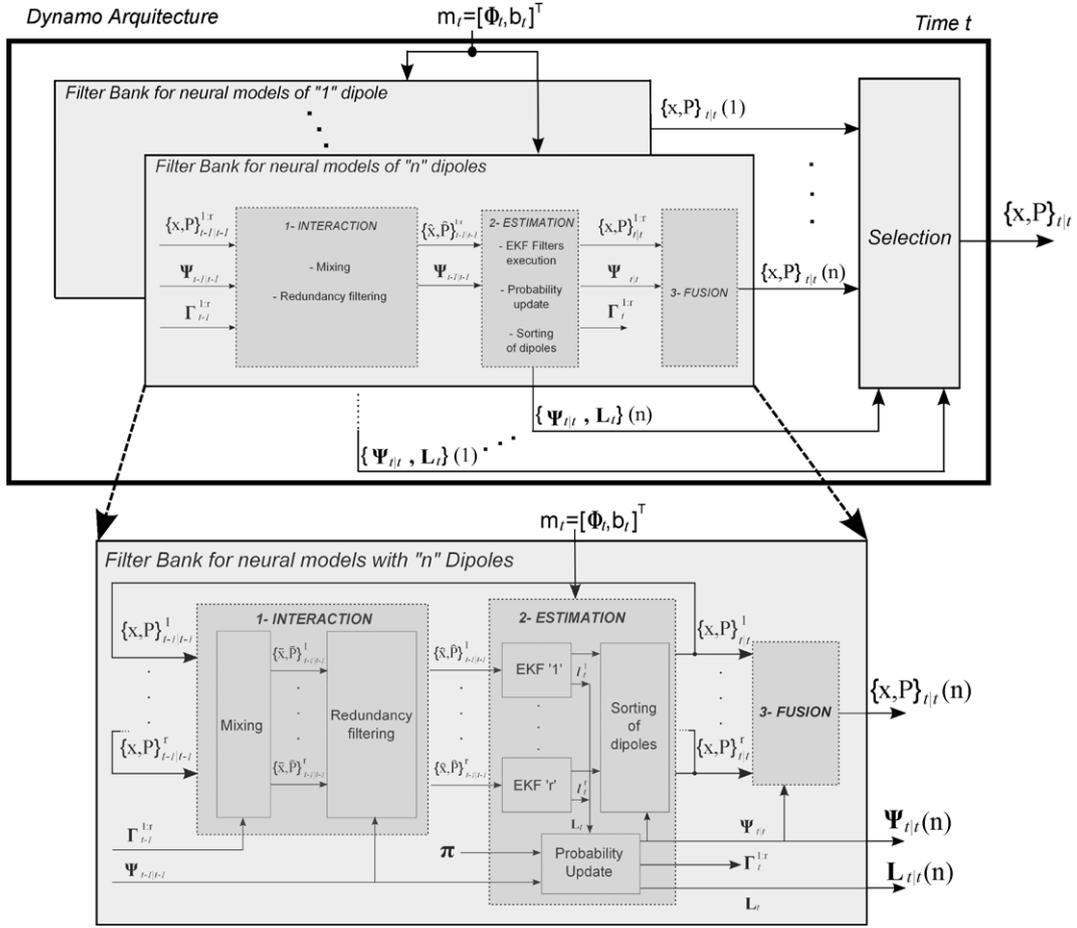


Fig. 1. Graphical representation of the dynamic multiple model estimation framework developed to solve the EEG/MEG source localization problem. The upper box describes the general architecture of the algorithm, composed of n filter banks and a selection block. The general input is the measurement vector \mathbf{m}_t and the general output is the source state vector with its associated covariance matrix $\{\mathbf{x}, \mathbf{P}\}_{t|t}$. The lower box represents the internal structure of a filter bank with its three steps: interaction, estimation, and fusion. Note that all filter banks have the same input observations. The output of a bank is the integrated source state vector with its associated covariance $\{\mathbf{x}, \mathbf{P}\}_{t|t}^1(n)$, posterior probability vector $\Psi_{t|t}^1(n)$, and likelihood vector $\mathbf{L}_{t|t}^1(n)$.

(dynamics and initial conditions). The state estimations of the EKF filters are probabilistically combined at each step to integrate the information of all filters and to build the bank estimation. Several banks of filters are running in parallel, but operating over models of different source space dimensions (different number of dipoles). The estimation of the method at each step is a probabilistic selection among the estimations of all banks. Fig. 1 shows the general architecture of the method.

2.3.1. Bank of EKF filters

Each bank is composed of r EKF filters and the estimation of the bank is computed in three steps: (i) interaction, (ii) estimation and (iii) fusion. Fig. 2 shows a graphical illustration of the execution of a one-time cycle of the bank of filters.

2.3.2. Interaction

The interaction step has two stages. Firstly, there is a probabilistic mixture (Mixing stage) of the prior estimates of all filters, where the estimations with poor performance are improved by the estimations that more likely explain the measurements. Secondly, there is a re-initialization of the filters that result redundant (redundancy filtering stage).

In the mixing stage, the mixed prior estimate $(\bar{\mathbf{x}}, \bar{\mathbf{P}})_{t-1|t-1}^i$ for each filter is computed as a weighted combination of the prior estimates of all filters:

$$\bar{\mathbf{x}}_{t-1|t-1}^i = \sum_{j=1}^r \mathbf{x}_{t-1|t-1}^j \gamma_{t-1}^{ji} \quad (16)$$

$$\bar{\mathbf{P}}_{t-1|t-1}^i = \sum_{j=1}^r \left[\mathbf{P}_{t-1|t-1}^j + \Delta_{t-1|t-1}^{ij} \cdot (\Delta_{t-1|t-1}^{ij})^T \right] \gamma_{t-1}^{ji} \quad (17)$$

where $\Delta_{t-1|t-1}^{ij} = (\bar{\mathbf{x}}_{t-1|t-1}^i - \mathbf{x}_{t-1|t-1}^j)$. The weights $\Gamma_{t-1}^i = [\gamma_{t-1}^{1i}, \dots, \gamma_{t-1}^{ri}]$ are the normalized mixing probabilities for filter i , which are computed in the $t-1$ estimation step and depend on the probability of filter i and on the probability transition between filter i and other filters.

In the next stage, filters with redundant estimations are re-initialized. This is motivated by the fact that during the estimation process some filters may tend to similar estimates, while the idea behind the bank of filters is to capture different solution areas of the state space (which is very appropriate in situations where the neural sources vanish and new ones appear in other brain regions). For each pair of filters, D_m and D_n , the redundancy between their

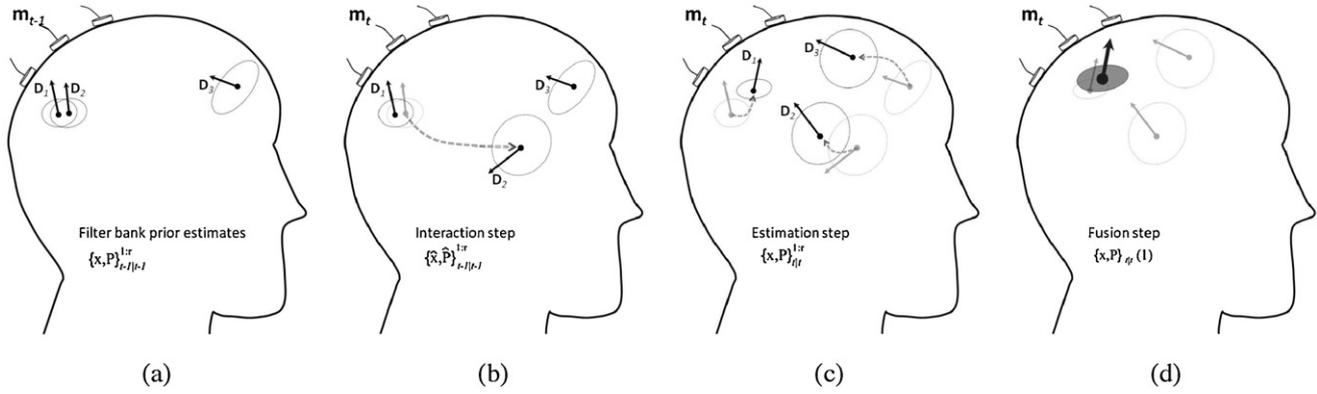


Fig. 2. Graphical illustration of a one-time cycle execution of a filter bank in the *Dynamo* algorithm. It is assumed a filter bank with source models of one dipole ($n=1$), which is composed of three ($r=3$) EKF filters (D_1 , D_2 and D_3). (a) Prior estimates given by the three EKF filters. Note that dipoles D_1 and D_2 seem to be equivalent in both position and moment. (b) Interaction step: firstly, prior estimates with poor performance are mixed with better prior estimates, and secondly, prior mixed estimates that are redundant are randomly re-initialized. Since D_1 and D_2 are redundant, and D_2 has the lower prior probability, D_2 will be randomly re-initialized. (c) Estimation step: firstly, each EKF filter produces its own source space estimate independently from its prior estimate and new available measurements \mathbf{m}_t , and secondly, the probability and likelihood are computed for each filter. (d) In the fusion step, all estimations are combined in terms of their posterior probabilities to produce the estimation of the bank of filters.

prior estimates $\{\bar{\mathbf{x}}, \bar{\mathbf{P}}\}_{t-1|t-1}^{D_m}$ and $\{\bar{\mathbf{x}}, \bar{\mathbf{P}}\}_{t-1|t-1}^{D_n}$ is given by means of their Mahalanobis distance:

$$d_{D_m, D_n}^2 = (\zeta_{t-1|t-1}^{D_m, D_n})^T (\bar{\mathbf{P}}_{t-1|t-1}^{D_m} + \bar{\mathbf{P}}_{t-1|t-1}^{D_n})^{-1} \zeta_{t-1|t-1}^{D_m, D_n} \quad (18)$$

where $\zeta_{t-1|t-1}^{D_m, D_n} = (\bar{\mathbf{x}}_{t-1|t-1}^{D_m} - \bar{\mathbf{x}}_{t-1|t-1}^{D_n})$. D_m and D_n are regarded statistically equivalent or redundant when $d_{D_m, D_n}^2 < \chi_{dim(\zeta), \alpha}^2$ (the distance is lower than the inverse chi squared cumulative distribution with a significance level of α). Thus, when they are redundant, the one with the lowest prior probability is randomly re-initialized and its probability is set to $\min(\Psi_{t-1|t-1})/1000$, while the other remains in the same estimate and its probability is set to $\psi_{t-1|t-1}^{D_m} + \psi_{t-1|t-1}^{D_n}$. Then $\Psi_{t-1|t-1}$ is normalized. The re-initialization of redundant filters allows to explore more efficiently the state space and does not modify the total mass probability of the bank.

2.3.3. Estimation

The estimation step consists of three stages. In the first step, each EKF filter computes its own source state vector with its associated covariance $\{\mathbf{x}, \mathbf{P}\}_{t|t}^i$, given the new measurements and the prior estimate (Eqs. (10) and (11)). In the second step, the posterior and mixing probabilities of all filters are computed. The posterior probability of the filters $\Psi_{t|t} = [\psi_{t|t}^1, \dots, \psi_{t|t}^r]^T$, is recursively updated by:

$$\Psi_{t|t} = \mathbf{L}_t \Psi_{t|t-1}^T \quad (19)$$

where $\mathbf{L}_t = [l_t^1, \dots, l_t^r]^T$ and $l_t^i = p(\mathbf{m}_t | \mathbf{x}_{t|t}^i) = \mathcal{N}(\mathbf{m}_t | \mathbf{x}_{t|t}^i, \mathbf{R})$ is the likelihood function of data, given that the estimation of the filter i is correct; and $\Psi_{t|t-1} = [\psi_{t|t-1}^1, \dots, \psi_{t|t-1}^r]^T$ is the predicted probability for all the filters given by:

$$\Psi_{t|t-1} = \pi \Psi_{t-1|t-1} \quad (20)$$

where $\Psi_{t-1|t-1} = [\psi_{t-1|t-1}^1, \dots, \psi_{t-1|t-1}^r]^T$ is the prior probability (defined initially equal for all the filters since all are randomly initialized and no prior information exists) and π is a mixing operator that accounts for transitions between models following a homogeneous constant Markov process. Every particular value π_{ji} represents the probability that, at time t , the selected filter is i , given that at time $t-1$ the selected filter was j . Given the prior probability vector (defined initially equal for all the filters since all filters are randomly initialized and no prior information exists) $\Psi_{t-1|t-1}$, and the probability transition (defined with the larger values on the

diagonal and then each row is normalized) π , each particular value of the mixing probability vector Γ_t^i for filter i is computed by:

$$\gamma_{t|t}^{ji} = \frac{\pi_{ji} \psi_{t-1|t-1}^j}{\psi_{t|t}^i}, \quad \forall j = 1, \dots, r \quad (21)$$

The final third stage is the sorting of the dipoles arrangement within the state vector and the covariance matrix $\mathbf{x}_{t|t}^i$ and $\mathbf{P}_{t|t}^i$. This is needed because filters with various dipoles (i.e., for banks of filters with two or more dipoles) may have multiple equivalent representations, which may cause trouble in the previous interaction step and the next fusion step when combining the estimations of all filters. The sorting is carried out for each filter by selecting the arrangement of dipoles in $\mathbf{x}_{t|t}^i$ and $\mathbf{P}_{t|t}^i$ (of all possible) that leads to the minimum Mahalanobis distance with respect to the filter with the highest probability, i.e., the state vector and the covariance matrix of each filter are reordered to match with the estimation given by the filter with the highest probability.

2.3.4. Fusion

In this step, all estimations are combined in terms of their posterior probabilities, to produce the resulting estimation of the bank of filters:

$$\mathbf{x}_{t|t}(n) = \sum_{i=1}^r \mathbf{x}_{t|t}^i \psi_{t|t}^i \quad (22)$$

$$\mathbf{P}_{t|t}(n) = \sum_{i=1}^r \left[\mathbf{P}_{t|t}^i + (\mathbf{x}_{t|t} - \mathbf{x}_{t|t}^i) (\mathbf{x}_{t|t} - \mathbf{x}_{t|t}^i)^T \right] \psi_{t|t}^i \quad (23)$$

The final estimation $\{\mathbf{x}, \mathbf{P}\}_{t|t}(n)$ is the weighted combination of the estimation given by all filters. The weight of each estimation depends on the posterior probability of the filter. The parameter n is the number of banks, provided that there are more banks running in parallel (see Fig. 1).

2.4. Selection of the estimation

Every bank of EKF filters estimates the neural sources of r source models with the same dimension of the source space (same number of dipoles) but different parameters (dynamics and initial conditions). The final step selects the final estimation from all banks. The selection algorithm firstly computes the overall bank

likelihood function value $LB_t(n)$ for every filter bank as a function of the posterior probability $\Psi_{t|t}(n)$ and likelihood vectors $\mathbf{L}_t(n)$.

$$LB_t(n) = \sum_{j=1}^r \mathbf{L}_t^j(n) \Psi_{t|t}^j(n) \quad (24)$$

Then, the bank likelihood and the number of sources associated with each bank are used to compute the Akaike information criteria (AIC):

$$AIC_t(n) = -2\ln(LB_t(n)) + 2N_s(n) \quad (25)$$

Subsequently, the filter bank with the minimum AIC value is selected:

$$n_m = \arg \min_n (AIC_t(n)) \quad (26)$$

The final estimate (state vector and covariance) at time t is the estimation of the n_m bank:

$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t}(n_m), \quad \mathbf{P}_{t|t} = \mathbf{P}_{t|t}(n_m) \quad (27)$$

In summary, *Dynamo* is a dynamic and multi-model estimation method for brain sources given EEG and/or MEG data. The estimation framework is based on the parallel execution of EKF filters operating over different source models, which probabilistically interact and fuse the information to obtain the best estimation at each time step.

2.5. Implementation of the *Dynamo* method

There are two main issues related to the implementation of the method: the sensor forward model and the parameters of *Dynamo*. The sensor forward model is based on the EEG/MEG head models (see Section 2.1.3) that were selected as follows. For EEG, a three shell spherical conductor model was used (radius fixed to 86, 96 and 100 mm for brain, skull and scalp respectively; conductivities fixed to 0.33 S/m for brain and scalp, and to 0.0042 S/m for skull). For MEG, a single-shell volume conductor model was used with a homogeneous and isotropic sphere of radius 100 mm (Grech et al., 2007). Both spherical head models were shaped to the realistic head model of the probabilistic magnetic resonance imaging atlas of the Montreal Neurological Institute (Evans et al., 1999). In addition, the observation noise variance matrices, \mathbf{R}^{eg} and \mathbf{R}^{meg} , are estimated as in Section 2.1.3.

The main parameters of *Dynamo* are the number of filter banks (given by the maximum number of dipoles), the number of EKFs per bank, the transition probability matrix, and the initial conditions and dynamics for each EKF. First, the *Dynamo* implementation used a maximum of three dipolar sources (i.e., three banks with models of 1, 2 and 3 dipoles), which is in line with studies supporting that a maximum number of three dipolar sources is sufficient to explain focalized brain activations (Plummer et al., 2008). Second, the number of EKF filters has an effect on *Dynamo* since the selection of a lower number of filters could omit valid hypotheses of neural sources. The selection of a higher number of filters would enhance the quality and accuracy of the estimation, but would lead to an unpractical application due to the computational burden. Simulations were carried out to study this effect, evaluating the estimation performance by varying the number of EKF filters per bank from 2 to 32. The best performance was found using 32 filters, as will be described in the next subsection. Third, the transition matrix was initialized with a higher probability to continue in the same model than to change to another one. This was implemented by assigning a probability of $r/(2r-1)$ to the diagonal elements and $1/(2r-1)$ to the off-diagonal ones, where r is the number of filters in the bank. Finally, for each EKF, the initial condition (for both, the first time step and for the redundancy filtering) was randomly

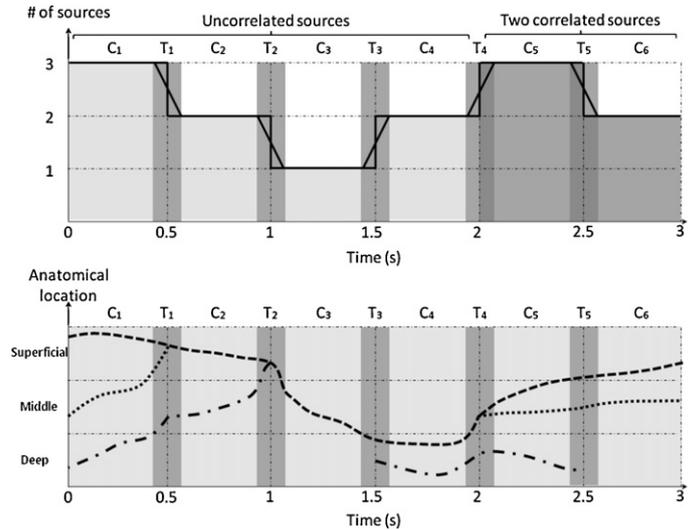


Fig. 3. Illustration of the time course of simulated neural sources. C_1, \dots, C_6 and T_1, \dots, T_5 are the conditions (stable behavior) and transitions (changes in characteristics) of the neural sources. Top: Illustration of the time course of the simulation versus the number of sources and temporal correlation (e.g., in C_2 there are two active and uncorrelated sources). Bottom: Illustration of the time course of the simulation versus anatomical location of the sources (e.g., in C_2 one source is cortical and the other is located in the middle brain).

located within the brain volume and within the allowable range of moment values (from -400 to 400 nA/m). The dynamics were classified as slow, moderate, and fast, which were implemented by the diagonal elements of \mathbf{Q} (these values were fixed to cover 10, 50 and 100% of brain volume in position and 10, 50 and 100% of a maximum possible value of moment components of 400nA/m). *Dynamo* was implemented with three banks of filters, 32 EKF filters per bank, with randomized initial conditions and slow, moderate, and fast neural dynamics on the EKF filters.

2.6. Evaluation in a simulated scenario

A key difficulty in the validation of source localization techniques is the lack of a *ground truth*, as the underlying neural sources are unknown. Therefore, the standard way of evaluating these techniques is to assess the ability to reconstruct known sources in a simulated scenario.

2.6.1. Design of neural sources and EEG/MEG signals

The neural sources were modeled by current dipoles with time-varying location within the brain volume and a time-dependent moment given by sinusoidal-damped functions. The neural sources were designed to address simultaneously three aspects that impose difficulties to existing source localization techniques: (i) variability in the number of sources; (ii) changes in the anatomical location of the sources, ranging from superficial (neocortical) to deep (subcortical); and (iii) changes in the temporal correlation of the sources. The simulation lasted 3 s, divided into C_i conditions with T_i transitions between each condition (Fig. 3).

The EEG/MEG head models were selected as described in the previous subsection to simulate sensor measurements given by previous neural sources. The EEG/MEG signals were sampled at 256 Hz and independently corrupted with a zero-mean Gaussian noise, with a standard deviation such that the signal-to-noise ratio (SNR) was 50 dB ($SNR = 20 \log(\sigma_M/\sigma_N)$, being σ_M the standard deviation of the obtained measurements and σ_N the standard deviation of the noise).

2.6.2. Evaluation design

As *Dynamo* is based on a stochastic process, the performance was evaluated by 10 Monte-Carlo simulations for each of the three modalities (EEG, MEG, and EEG + MEG). The following metrics were used for evaluation (Bar-Shalom et al., 2001; Peralta-Menendez and Gonzalez-Andino, 2002; Hesheng et al., 2004; Grova et al., 2006; Grech et al., 2007):

- *Error in the number of sources*: $\varepsilon(\mathbf{N}_s) = \text{abs}(\mathbf{N}_s - \tilde{\mathbf{N}}_s)$. Mismatch between the actual \mathbf{N}_s and the estimated number of sources $\tilde{\mathbf{N}}_s$.
- *Source position error*: Firstly, $\varepsilon(\mathbf{r}_q) = \sum_{i=1}^N \|\mathbf{r}_q^i - \tilde{\mathbf{r}}_q^i\|$ is the sum of the distances between the location of the actual sources and the location of the estimated sources, when the estimated and actual number of sources agree. Secondly, $\varepsilon'(\mathbf{r}_q) = \varepsilon(\mathbf{r}_q) + 1/M \sum_{j=1}^M \|R - \tilde{R}^j\|$ is the source position error plus the mean distance between the location of the M underestimated/overestimated sources (\tilde{R}) and the location of the estimated/real source (R), when the estimated and real number of sources do not agree. $\varepsilon(\mathbf{r}_q)$ and $\varepsilon'(\mathbf{r}_q)$ were normalized with respect to the radius of the brain in the head model.
- *Source moment error*: Firstly, $\varepsilon(\mathbf{q}) = \sum_{i=1}^N \|\mathbf{q}^i - \tilde{\mathbf{q}}^i\|$ is the sum of the distances between the moment of the actual sources and the moment of the estimated sources, when the estimated and actual number of sources agree. Secondly, $\varepsilon'(\mathbf{q}) = \varepsilon(\mathbf{q}) + 1/M \sum_{j=1}^M \|Q - \tilde{Q}^j\|$ is the source moment error plus the mean distance between the moment of the M underestimated/overestimated sources (\tilde{Q}) and the moment of the estimated/real source (Q), when the estimated and real number of sources do not agree. $\varepsilon(\mathbf{q})$ and $\varepsilon'(\mathbf{q})$ were normalized with respect to the maximum of the moment component.
- *Goodness-of-fit*: $\text{gof} = 1 - (\mathbf{m} - \tilde{\mathbf{m}})^T (\mathbf{m} - \tilde{\mathbf{m}}) / \mathbf{m}^T \mathbf{m}$. Degree of fitness between the measurements given by the estimated sources $\tilde{\mathbf{m}}$ and the real measurements \mathbf{m} .
- *Consistency index*: $ci = \text{nees} / \chi_{N_{dim}, \alpha}^2$, where $\text{nees} = (\mathbf{x} - \tilde{\mathbf{x}})^T \tilde{\mathbf{P}}^{-1} (\mathbf{x} - \tilde{\mathbf{x}})$ is the normalized estimation error square, α is the significance level and N_{dim} is the dimension of the state vector ($6 \cdot N$). When $ci < 1$ the estimation is consistent with the true value, otherwise the estimation may have diverged.

Based on the previous setup, four analysis were carried out: the first analysis addressed the effect of the number of filters on the estimation; the second verified the effect of the number of EEG sensors; the third assessed the accuracy of the estimation in different situations; and the final fourth analysis was a comparison against other techniques. In the first analysis, a 10/10 international system was adopted with 64 EEG sensors (assuming that no significant improvements would be obtained with more sensors, see Michel et al. (2004) – this assumption was positively confirmed later) and the number of filters ranged from 2 to 32. In the second analysis, the number of filters was set to 32 (as the best performance was achieved with 32 filters) and the number of sensors ranged from 16 to 96. In the third analysis, MEG data was included with 151 magnetic coils simulating a real acquisition CTF System and 32 filters were used. The accuracy of the estimation method was assessed by the three sensor modalities (EEG, MEG, and EEG + MEG) and several effects were studied (conditions C_i versus transitions T_i ; anatomical location of the sources –subcortical, middle and cortical–; correlated versus uncorrelated).

The fourth analysis was a comparison against widely spread methods: the Least-Squares *LS* method (Mosher et al., 1992), the minimum norm estimate *MNE* (Hämäläinen and Ilmoniemi, 1994), the linear-constrained minimum variance beamforming *LCMV* (Veen et al., 1997), and the low resolution electromagnetic tomography method *LORETA* (Pascual-Marqui et al., 1994). The *LS*

method is a parametric solution that employs a non-linear optimization process to estimate a predefined fixed number of current dipoles, which was implemented with one, two and three sources (referred as *LS1dip*, *LS2dip* and *LS3dip* respectively). The *MNE*, *LCMV* and *LORETA* methods are distributed solutions with a solution space implemented by a dense grid of fixed-positions spanning the entire brain volume. For comparison purposes, in the case of *LS* solutions the metrics $\varepsilon'(\mathbf{r}_q)$ and $\varepsilon'(\mathbf{q})$ were used as in some conditions C_i and transitions T_i the number of assumed dipoles did not match the real number of sources. In the case of methods *MNE*, *LCMV* and *LORETA*, equivalent current dipoles were calculated from the distributed solutions as follows (note that the number of sources is known). For one source, the equivalent dipolar position was the weighted sum of the grid positions, where the weights were the normalized power estimated by the method. For two or three sources, the equivalent dipolar positions were computed using an optimization process that minimized the distances of clusters in the distributed solution. Once these equivalent dipolar positions were obtained, a straightforward linear optimization was applied to estimate the equivalent dipolar moment ($\hat{\mathbf{q}}_t = K_t^\dagger \mathbf{m}_t$, where K_t is the lead field evaluated at the equivalent dipoles positions, \mathbf{m}_t is the recorded measurements at time t and K_t^\dagger is the pseudo inverse of K_t).

2.7. Evaluation in a real scenario

Real scenario evaluation is also a key difficulty to overcome, as the underlying neural sources are always unknown. Therefore, a usual way of evaluating source localization methods in this context is to assess the ability to reconstruct neural sources in a well-established and known scenario of brain activity. Note that the presented paradigm mainly serves as a test to understand the *Dynamo* technique with non-simulated EEG data.

2.7.1. Experimental paradigm and data collection

The experiment (Lopez-Larraz et al., 2010) replicated the study of (Miltner et al., 1997), which was based on a feedback stimuli presentation after a time-estimation task. The presentation of negative/positive feedback produces an event-related response mainly generated by Brodmann areas 24 and 32 in the anterior cingulate (Miltner et al., 1997; Carter et al., 1998; Bush et al., 2000; Holroyd and Coles, 2002), which was used as *ground truth*.

Five healthy male subjects participated in the experimental sessions. The mental task was to estimate the duration of one second. Each trial started with a visual cue to indicate that the subject had to press a button one second later. At the end of the mental count, the subject pressed the button and 0.6 seconds later he/she received a visual feedback indicating whether the estimated time was correct or incorrect. The EEG data was recorded from 32 electrodes uniformly distributed according to the 10/10 international system. The ground electrode was on FPz and the reference was placed on the left earlobe. The EEG was digitalized with a sampling frequency of 256 Hz, power-line notch-filtered and bandpass-filtered between 0.5 and 60 Hz. For each participant, approximately 150 error trials and 150 correct trials were recorded. After the experimental sessions, epochs were extracted starting from the visual cue presentation until one second later, then the EEG signals were average-referenced and filtered with a zero-phase shift, sixth-order, low-pass Butterworth filter with a cutoff frequency of 10 Hz.

2.7.2. Evaluation design

The neural sources were estimated through the execution of 10 Monte-Carlo runs. The performance of the method was evaluated according to the following metrics:

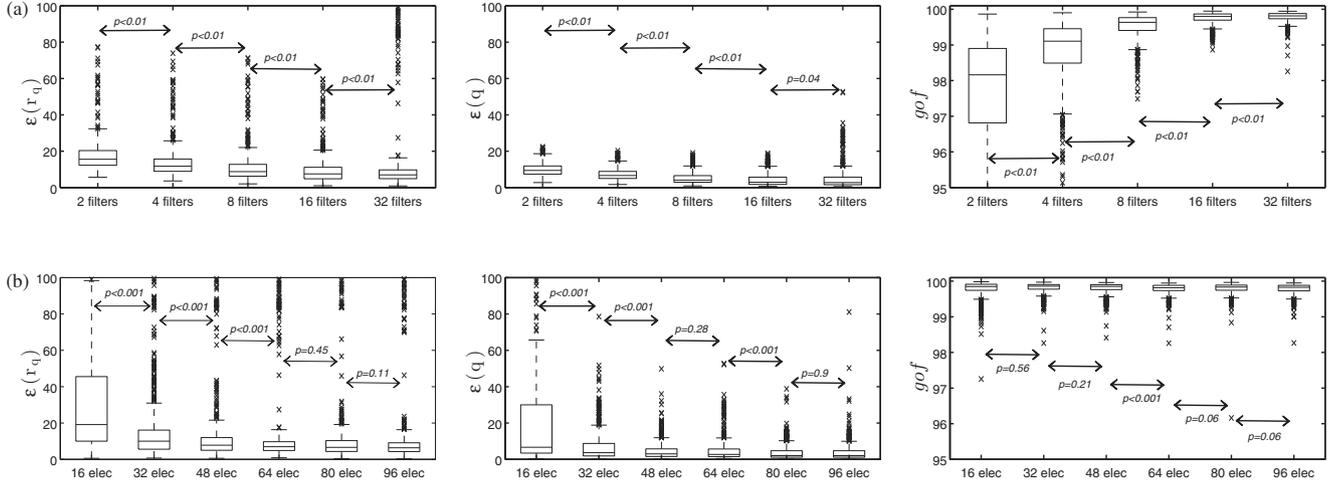


Fig. 4. (a) Distributions of $\varepsilon(\mathbf{r}_q)$, $\varepsilon(\mathbf{q})$ and gof for different configurations of *Dynamo* for the case of EEG with 64 electrodes. The results show that the estimation accuracy of *Dynamo* increased with the number of filters but computation time dramatically increased. (b) Distributions of $\varepsilon(\mathbf{r}_q)$, $\varepsilon(\mathbf{q})$ and gof for different numbers of EEG sensors for the case of *Dynamo* implemented with 32 filters per bank. These results show that the estimation accuracy of *Dynamo* increases up to 64 electrodes, and that after this, the improvement is not significant.

- **Brodmann area and brain structure:** BA and BS . Brodmann area and the brain structure to which the estimated source belongs.
- **Distance to the error-processing region:** DE_t . Minimum distance from the estimated source to the brain region involved in error-processing (i.e., BA 's 24 and 32, and BS anterior cingulate and cingulate gyrus) at a given time.

$$DE_t = \min \|\mathbf{r}_t - \mathbf{r}_t^i\|, \quad i = 1, \dots, N \quad (28)$$

where \mathbf{r}_t is the position of the estimated source and $\mathbf{r}_t^i, i=1, \dots, N$ are the position of N points located within the brain and labeled as either BA 24, BA 32, BS anterior cingulate or BS cingulate gyrus.

- **Activation of the Brodmann areas and brain structures:** $ABA(i)$ and $ABS(i)$. Total activation of the i th Brodmann area or i th brain structure in a time interval.

$$ABA'(i) = \sum_{t=1}^T \sum_{n=1}^{N_t} \|\mathbf{q}_t^n\|^2 \cdot \alpha_{t,n}^i, \quad ABA(i) = \frac{ABA'(i)}{\sum_j ABA'(j)} \quad (29)$$

$$ABS'(i) = \sum_{t=1}^T \sum_{n=1}^{N_t} \|\mathbf{q}_t^n\|^2 \cdot \alpha_{t,n}^i, \quad ABS(i) = \frac{ABS'(i)}{\sum_j ABS'(j)} \quad (30)$$

where $ABA(i)$ is the activation percentage of the i th Brodmann area ($i = 1, \dots, 52$ as there are 52 Brodmann areas labeled in the MNI space), $ABS(i)$ is the activation percentage of the i th brain structure ($i = 1, \dots, 32$ as there are 32 brain structures of the brain labeled in the MNI space, and for simplicity, $ABS(24)$ and $ABS(32)$ will refer to the activation of the BS 's anterior cingulate and cingulate gyrus, respectively), T is the number of time samples, N_t is the number of neural sources estimated at time t , $\alpha_{t,n}^i$ is a binary value equal to one if the actual source is located within the i th Brodmann area (i.e., $\mathbf{r}_t^n \in i^{th}$ Brodmann area) or zero otherwise, and \mathbf{r}_t^n and \mathbf{q}_t^n are the position and moment estimates of source n at time t . The full Gaussian-approximate distribution computed by *Dynamo* was taken into account by drawing $N = 100$ samples from $\mathcal{N}(\mathbf{x}_t, \mathbf{P}_t)$ (instead of the mean estimate \mathbf{x}_t). In distributed solutions, the maximum of the estimated activity at each time is taken into account to compute the ABA and ABS .

Based on these metrics, two analysis were carried out. The first analysis studied the estimation of the neural sources for the grand average difference between error and correct conditions for all trials and subjects. The second analysis studied the estimation of the

neural sources for the grand average difference, but separately for each subject. In both analysis the neural sources were studied at the occurrence of the most prominent peaks and during the time interval of the evoked activity, and a comparison among *Dynamo*, *LS1dip*, *LCMV* and *LORETA* methods was performed.

3. Results

3.1. Evaluation for the simulated scenario

3.1.1. Analysis 1: number of filters

Fig. 4a summarizes the results of the influence of the number of filters on the performance of the technique. Significant differences were found between distributions of $\varepsilon(\mathbf{r}_q)$ ($p < 0.01$, Wilcoxon rank-sum test) as the number of filters was progressively modified from 2 to 32. No significant differences were encountered between distributions of $\varepsilon(\mathbf{q})$ ($p > 0.01$, Wilcoxon rank-sum test) when the number of filters changed from 16 to 32. Regarding gof , as the number of filters changed successively from 2 to 32 significant differences were found ($p < 0.01$, Wilcoxon rank-sum test). These results show that the performance of the method improved as the number of filters also increased; however, the computation time rapidly increased along with the number of filters.

3.1.2. Analysis 2: number of EEG sensors

Fig. 4b summarizes the results of the analysis of the number of EEG sensors. Significant differences were found ($p < 0.01$, Wilcoxon rank-sum test) between successive distributions of $\varepsilon(\mathbf{r}_q)$, and the median of the distributions decreased as the number of electrodes progressively increased from 16 to 64. However, no significant differences were found ($p > 0.01$, Wilcoxon rank-sum test) when the number of electrodes changed from 64 to 80 or from 80 to 96. No significant differences were found between distributions of $\varepsilon(\mathbf{q})$ ($p > 0.01$, Wilcoxon rank-sum test) when the number of sensors changed from 48 to 64 or from 80 to 96. Regarding gof , significant differences between distributions were found ($p < 0.01$, Wilcoxon rank-sum test) when the number of sensors varied from 48 to 64, and no significant differences were found ($p > 0.01$, Wilcoxon rank-sum test) when the number of sensors changed either from 64 to 80 or from 80 to 96. These results revealed that the accuracy in the estimation of the sources increased until 64 electrodes; however, the increase is not as significant when using 80 or 96 electrodes. These results agreed with the results of Michel et al. (2004), where

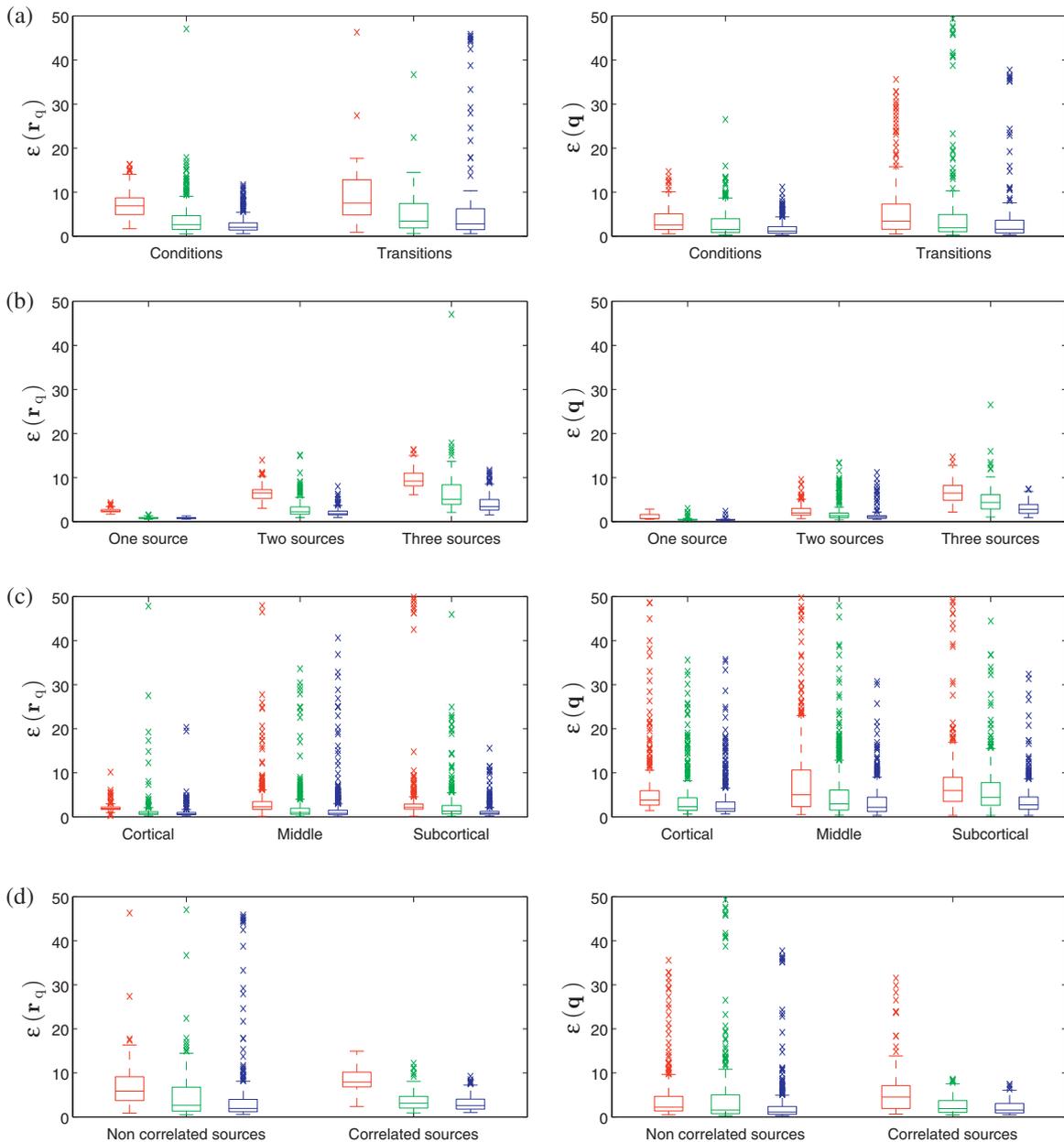


Fig. 5. Distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$: (a) during conditions and transitions, (b) considering the number of sources, (c) considering the anatomical location of sources and (d) considering the correlation between sources. The results are presented in red, green, and blue for the EEG, MEG and EEG + EMG sensor modalities, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

simulations and interictal epileptic data were used to show that the source localization accuracy increased with the number of electrodes, and that after 60 electrodes the improvement in localization is minimal.

3.1.3. Analysis 3: evaluation of estimation accuracy

The position accuracy $\varepsilon(\mathbf{r}_q)$ of the method was 12.06%, 8.23% and 5.43% for the EEG, MEG, and EEG + MEG sensor modality, respectively, and the moment accuracy $\varepsilon(\mathbf{q})$ was 4.55%, 4.23% and 2.53% for the EEG, MEG, and EEG + MEG sensor modality, respectively. These results are in line with Kiebel et al. (2008) and Sorrentino et al. (2009), which reported a position error of 8 mm (i.e., $\varepsilon(\mathbf{r}_q) = 11.63\%$) and of 6 mm (the $\varepsilon(\mathbf{r}_q)$ could not be computed as the head model was not reported). For the three sensor modalities, the estimation was consistent with the true values ($ci < 1$) at all times, which indicated that the estimation did not diverge. Note that high values (i.e., $ci \approx 1$) were obtained during the first time and during T_i

transitions, which are times when the uncertainty of the estimation increased as the characteristics of the neural sources were changing. Regardless of sensor modality, *Dynamo* was able to consistently and accurately estimate the neural sources in the experimental situations.

Effects of the sensor modality: The distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ were significantly different ($p < 0.01$, Wilcoxon rank-sum test) across sensor modalities, and the lower median values were achieved with EEG + MEG. No significant differences were found ($p > 0.01$, Wilcoxon rank-sum test) between the distributions of *gof*. These results indicate that the use of combined EEG + MEG offered better performance than using either EEG or MEG alone, which agrees with previous studies (Sharon et al., 2007; Molins et al., 2008; Henson et al., 2009).

Effects of estimation situations (C_i and T_i): For the three sensor modalities, the distributions of $\varepsilon(\mathbf{r}_q)$, $\varepsilon(\mathbf{q})$ and *gof* were significantly different ($p < 0.01$, Wilcoxon rank-sum test) across conditions and

transitions (Fig. 5a). The medians of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ were higher in the transitions than in the conditions, and the medians of gof were slightly lower in transitions than in conditions. Moreover, nonzero $\varepsilon(\mathbf{N}_s)$ was found in transitions but not in the conditions. The time course of these metrics also revealed poor performance during initial time instants (i.e., high $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$, nonzero $\varepsilon(\mathbf{N}_s)$, and low gof), which was due to the fact that all filters were randomly initialized and thus some iterations were needed to update the filters and produce good estimates. These results indicate that the estimation accuracy was better during C_i than during T_i . This is because during T_i the neural sources are changing and *Dynamo* requires some iterations to accommodate the estimation by updating the filters and selecting the appropriate bank.

Effects of the number of sources: For the three sensor modalities, significant differences were found ($p < 0.01$, Wilcoxon rank-sum test) in the distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$. Both median values increased as the number of sources required for estimation increased (Fig. 5b). These results indicate that the estimation accuracy was lower as the number of sources to be estimated increased. This occurs because the estimation problem becomes more challenging as the dimensionality of the state-space increases (i.e., more sources and thus more parameters to estimate).

Effects of the anatomical location of the sources: For the three sensor modalities, the distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ from subcortical sources were significantly different and higher ($p < 0.01$, Wilcoxon rank-sum test) than those from middle and cortical sources (Fig. 5c). No significant differences in the distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ were found ($p > 0.01$, Wilcoxon rank-sum test) between middle and cortical sources (Fig. 5c). These results show a lower performance in the estimation of neural sources located in deep regions of the brain, which agrees with other studies (Whittingstall et al., 2003). This occurs because EEG and MEG signals arising from subcortical sources (those located farther away from the sensors) have a smaller magnitude than middle or cortical sources (as in the mathematics of the forward model the EEG/MEG signals are inversely related to the cubic distance between sensors and sources), and therefore the signals are more affected by measurement noise and by EEG/MEG signals of other middle or cortical sources.

Effects of the correlation of the sources: For all sensor modalities, the distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ revealed no significant differences ($p > 0.01$, Wilcoxon rank-sum test) between uncorrelated and correlated sources (Fig. 5d). These results indicate that the technique had a similar performance irrespective of whether the sources were correlated or not. This is because the *Dynamo* solution relies on the forward solution for its implementation in a manner that is independent of the observations, whereas other methods such as adaptive distributed solutions (e.g., LCMV) relies on both the forward model and observations (Quraan and Cheyne, 2010), (i.e., the measurements influence the estimation of each source individually).

3.1.4. Analysis 4: comparison against other methods

For all conditions and transitions, the distributions of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ of *Dynamo* were significantly different ($p > 0.01$, Wilcoxon rank-sum test) and lower than those obtained with other techniques (Fig. 6). Indeed, the median values of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ obtained with *Dynamo* (12.06% and 4.55%) were lower than those obtained with *LS1dip* (40.65% and 20.08%), *LS2dip* (24.08% and 13.41%), *LS3dip* (39.80% and 9.82%), *MNE* (33.36% and 7.17%), *LCMV* (41.05% and 8.83%) and *LORETA* (34.19% and 7.35%). Regarding gof , significant differences were found ($p < 0.01$, Wilcoxon rank-sum test) between the distribution of *Dynamo* and the *LS1dip* solution, and no significant differences were found between the distribution of *Dynamo* and the other methods. These results show the high performance of *Dynamo* in this experimental setup.

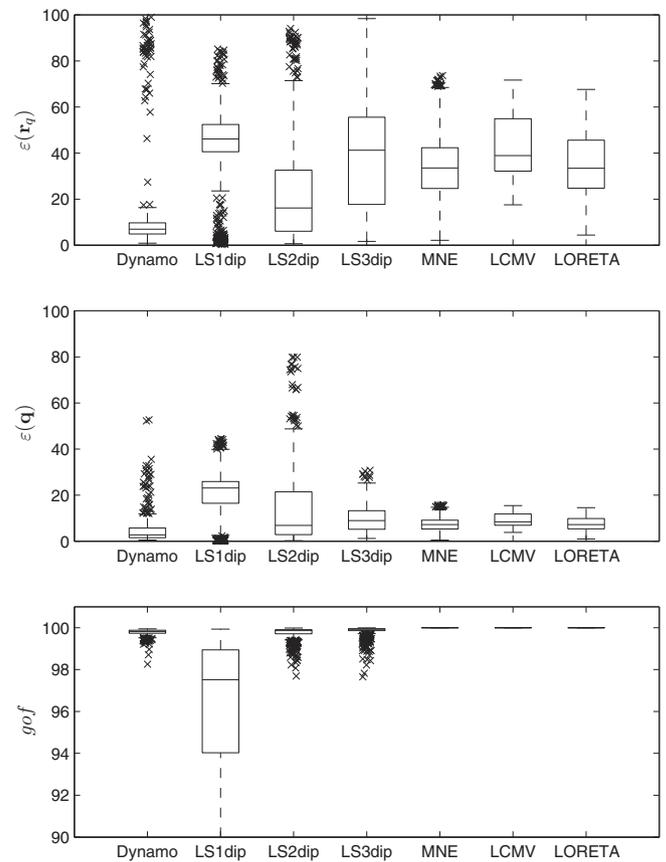


Fig. 6. Distributions of $\varepsilon(\mathbf{r}_q)$, $\varepsilon(\mathbf{q})$ and gof obtained with *Dynamo*, the least squares solutions (*LS1dip*, *LS2dip* and *LS3dip*) and the distributed solutions (*MNE*, *LCMV* and *LORETA*) for all the conditions and transitions.

In addition, the distributions of $\varepsilon(\mathbf{r}_q)$, $\varepsilon(\mathbf{q})$ and gof were obtained separately for each C_i condition for all methods. In all conditions *Dynamo* presented the best performance. The *LS* solutions presented good performance solely when the number of sources assumed by the method agreed with the number of sources of the condition. In condition C_2 for example, the median values of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ obtained with *Dynamo* (4.90% and 1.83%) were similar to those obtained with *LS2dip* (7.71% and 6.03%) and lower than those obtained with *LS1dip* (57.42% and 32.94%), *LS3dip* (43.96% and 13.53%), *MNE* (25.39% and 5.46%), *LCMV* (38.71% and 8.32%) and *LORETA* (27.42% and 5.90%). Even in this condition, which was the most favorable for *LS2dip* (as there are two active neural sources), the estimation accuracy is similar but still better for *Dynamo*.

3.2. Evaluation in the real scenario

3.2.1. Analysis 1: neural sources of the grand average difference between conditions, averaged for all subjects

Fig. 7a displays, for all trials and subjects, the grand average of feedback potentials in the error and correct conditions at channel FCz. In the error condition, the feedback potential has a negative deflection at ≈ 281 ms, a positive peak at ≈ 379 ms and a negative peak at ≈ 488 ms; while in the correct condition, the feedback potential has similar components but with noticeable smaller amplitudes. The evoked activity is observed between ≈ 200 and ≈ 500 ms. The topographical scalp maps of the grand average difference at the occurrence of those prominent peaks reveal focalized patterns on the fronto-central scalp (Fig. 7a). These results agree with Miltner et al. (1997) and Holroyd and Coles (2002) in the same experimental paradigm.

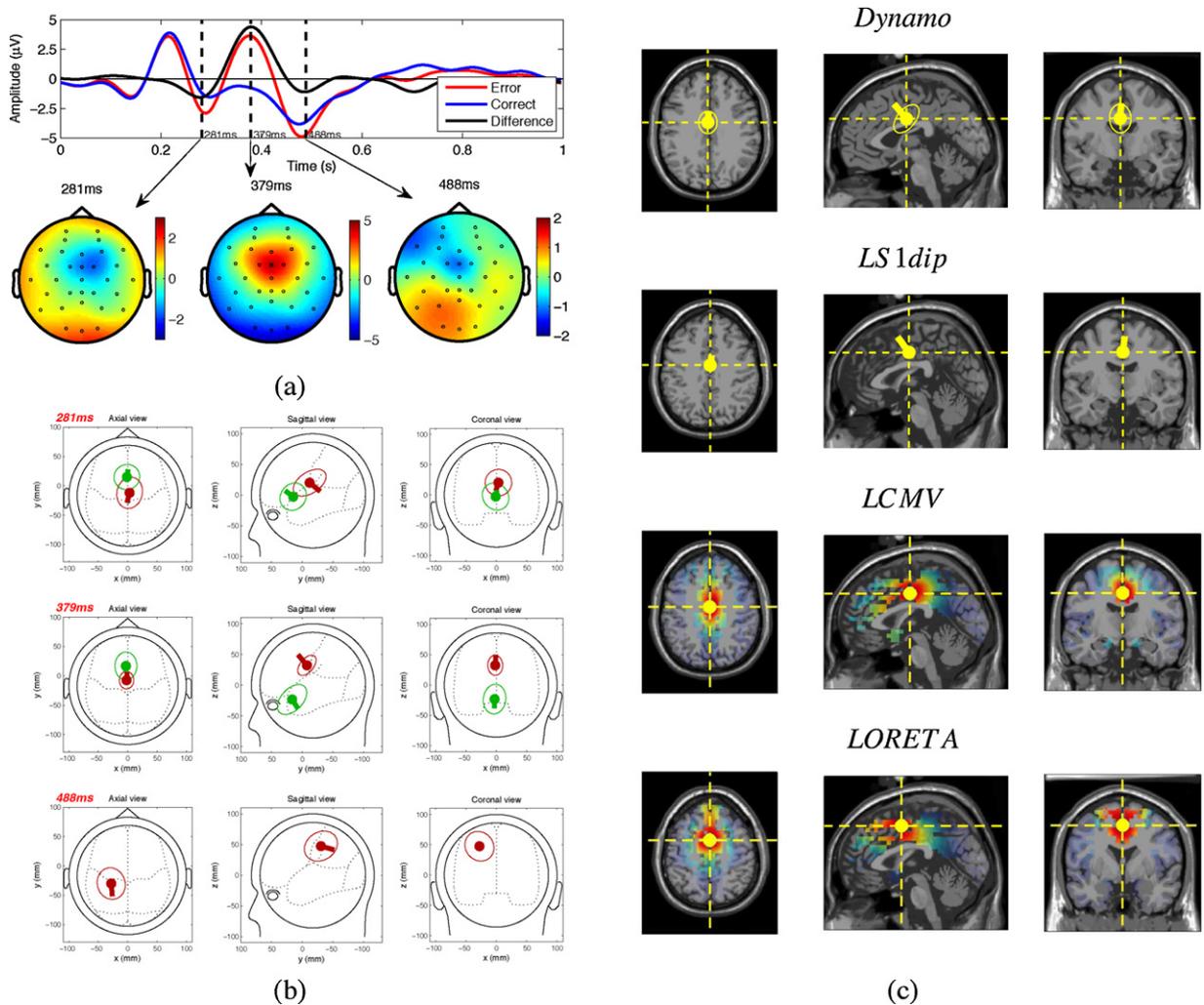


Fig. 7. (a) Waveforms of grand average potentials for all trials and subjects in the error condition (red), correct condition (blue) and difference between conditions (black) for channel FCz, and topographical scalp maps of the grand average difference between both conditions at time latencies of the more prominent peaks (dotted vertical lines). (b) Neural sources estimated with *Dynamo* projected onto spherical head slices for the time latencies of the three more prominent peaks (281, 379 and 488 ms). (c) Neural source of *Dynamo*, *LS1dip*, *LCMV*, and *LORETA* projected onto MRI slices for a time latency of 379 ms (second peak of the ErrP feedback showed in panel a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The neural sources behind the grand average difference between the two conditions were estimated with *Dynamo*. In the first negative and second positive peaks (≈ 281 and ≈ 379 ms, respectively), *Dynamo* estimated two neural sources (upper and middle plots in Fig. 7b). In both peaks, one of the sources lies on BA 24 in the BS cingulate gyrus. This result is consistent with (Miltner et al., 1997). The other source was located, for the first negative peak in BA 25 in the BS anterior cingulate, and for the second positive peak in BA 11 in the BS rectal gyrus (associated with reasoning and decision making, which might be related to the mental recognition of feedback cues and is also consistent with (Miltner et al., 1997). For the third peak (≈ 488 ms), one source located on BA 4 in the BS precentral gyrus was estimated (lower plot in Fig. 7b). These results show that, at the occurrence of the prominent peaks, the estimated sources were located in the brain regions related with error processing.

The neural sources was also estimated with other source localization techniques. Fig. 7c shows, at the time latency of the second positive peak (≈ 379 ms), the neural sources estimated with *Dynamo*, *LS1dip*, *LCMV*, and *LORETA* projected onto sagittal MRI slices (for *Dynamo*, only the closest source to the error-processing region is presented). Firstly, for this time latency, the estimated source of *Dynamo* and *LS1dip* solutions as well as the maximum of the estimated activity in the *LCMV* and *LORETA* solutions lied on BA

24 in the BS cingulate gyrus. Results with equivalent BA's and BS's were also obtained in the other two peaks. This indicates that in the prominent peaks, all methods estimated similar neural sources and in the brain regions that mediate in error-processing. Secondly, Fig. 8a shows, for all methods, the distance to the error-processing region, *DE*, averaged between 200 and 500 ms. Note that *DE* in the *Dynamo* solution was lower and significantly different ($p < 0.01$, Wilcoxon rank-sum test) to the other solutions. This indicates that for the time interval of the evoked activity, *Dynamo* was the closest solution to the error-processing region. Thirdly, the activation of the Brodmann areas and brain structures pertaining to the error-processing regions, *ABA*(24, 32) and *ABS*(24, 32), computed in the time interval of the evoked activity (200 to 500 ms), are displayed in Fig. 8b and c. Note that these activations are higher for *Dynamo* (*ABA*(24, 32)=50% and *ABS*(24, 32)=34%) than for the other techniques. These results show that in the time interval of the evoked activity, *Dynamo* led to higher activations of the Brodmann areas and brain structures that were involved in error-processing.

3.2.2. Analysis 2: neural sources of the grand average difference for each subject

The neural sources of the grand average difference were also estimated separately for each subject with *Dynamo*, *LS1dip*, *LCMV*

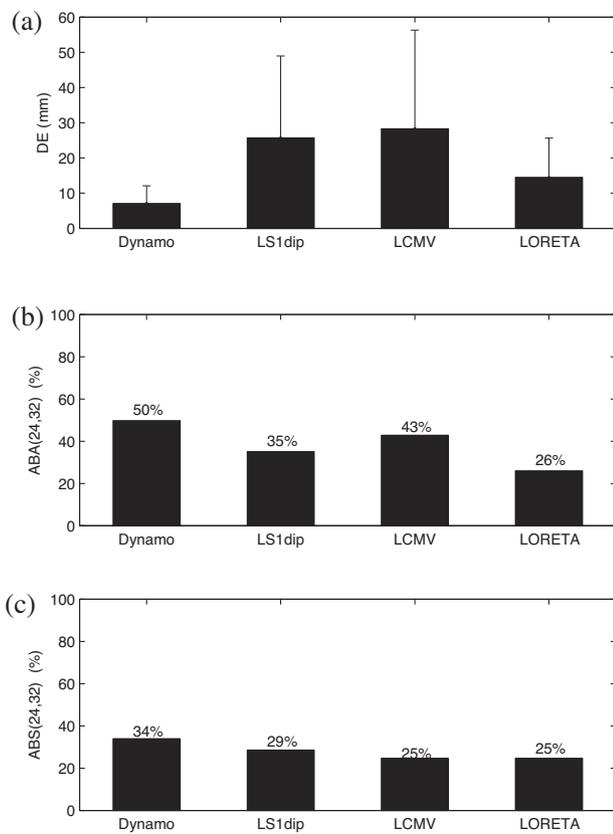


Fig. 8. For the neural sources of grand average difference between the error and correct conditions for all trials and subjects, and estimated with *Dynamo*, *LS1dip*, *LCMV*, and *LORETA*, (a) distance to the error-processing region (*DE*) averaged for the time interval from 200 to 500 ms, and activation of the error-related (b) Brodmann areas *ABA*(24, 32) and (c) brain structures *ABS*(24, 32), computed for the time interval from 200 to 500 ms.

and *LORETA* methods. At the time of the second positive peak (≈ 379 ms), the topographic scalp distribution revealed a noticeable positive focalized fronto-central scalp pattern in subjects 1, 2, 3 and 5, and also a positive but less focalized (more diffuse) scalp pattern in subject 4 (Fig. 9a). For this time latency, Fig. 9c–e display for all techniques the estimated neural source projected onto sagittal MRI slices, and the *BA*'s and *BS*'s associated with those estimated sources.

For subjects 1, 2, 3 and 5, the *Dynamo* estimate is located on *BA* 24 in the *BS* anterior cingulate; the *LS1dip* and *LCMV* estimates are located in *BA*'s 24, 32 or 23 in the *BS*'s anterior cingulate and cingulate gyrus; and the *LORETA* estimate is located on *BA*'s 6, 32, 31 and 24 in the *BS*'s superior frontal gyrus, cingulate gyrus and the paracentral lobule. For subject 4, more posterior but less similar sources are observed in all methods, which could be attributed to the more noisy and diffused scalp pattern as shown in Fig. 9a. Note that these results show that in subjects 1, 2, 3 and 5, the estimation of *Dynamo* is always in the error-related processing brain regions, while the estimation of *LS1dip*, *LCMV* and *LORETA* is not only in the error-related regions but also in other adjacent brain regions.

Fig. 10a displays for all subjects and methods, the distance to the error-processing region metric (*DE*), averaged between 200 and 500 ms. Note that in all subjects *DE* was lower and significantly different ($p < 0.01$, Wilcoxon rank-sum test) for *Dynamo* than for the other techniques. In addition, average *DE* was 8.95 mm, 27.07 mm, 32.53 mm and 24.07 mm, for *Dynamo*, *LS1dip*, *LCMV* and *LORETA*, respectively (Fig. 10a). These results indicate that for all subjects, *Dynamo* provided the closest solution to the error-processing region. Fig. 10b and c shows for all subjects and solutions, the

activation of the Brodmann areas 24 and 32, *ABA*(24, 32), and the activation of the anterior cingulate and cingulate gyrus, *ABS*(24, 32), both computed between 200 and 500 ms. For subjects 1, 2, 3 and 5, *ABA*(24, 32) and *ABS*(24, 32) are both greater for *Dynamo* than for other solutions. Note that *ABA*(24, 32) and *ABS*(24, 32) are both very low in subject 4 for all solutions. This could be attributed to the high diffuse scalp pattern (noisy EEG signals) as shown in Fig. 9a. In addition, the average of *ABA*(24, 32) for all subjects was 44.42%, 32.11%, 27.53% and 12.22% for *Dynamo*, *LS1dip*, *LCMV* and *LORETA*, respectively; while the average of *ABS*(24, 32) was 38.49%, 24.20%, 27.79% and 10.91%. These results show that *Dynamo* led to high activations of the Brodmann areas and brain structures involved in error-processing.

4. Discussion

This paper described *Dynamo*, a framework for the estimation of neural sources that relies on the probabilistic fusion of multiple dynamic dipolar source models. Every model is defined as a dynamic system, whose parameters are the number of sources (each source modeled with a current dipole with its position and moment), and the dynamics and priors of these sources. The estimation for each model is recursively computed within a Bayesian filter framework, and for all models, the estimations are probabilistically integrated to build the estimation solution. There are several aspects of this technique that require further discussion, such as the representation of the state space, the multi-model definition, the Bayesian tools, and the implementation.

In classical dipolar solutions, the number of sources must be defined a priori with their corresponding temporal dynamics and priors (e.g., techniques based on minimization (Scherg, 1990; Mosher et al., 1992) and on Bayesian estimation (Somersalo et al., 2003; Sorrentino et al., 2008)). This is a limitation for the estimation of a wide range of neural processes, where these parameters change in time. *Dynamo* addresses such limitations in a unified framework by means of probabilistic integration of the estimations of multiple models with different values for these parameters. This integration also allowed for the tracking of several hypotheses of neural sources, which is a key issue (the source localization problem is ill-posed and the vast majority of existing dipolar and distributed techniques do not address this). On one hand, some techniques use regularization schemes (Liu et al., 2002; Darvas et al., 2004) in this direction, however these are mathematical strategies to select a solution but not to represent different models. On the other hand, other techniques propose the use of different models in the form of different priors that yield different solutions (Phillips et al., 2005; Mattout et al., 2006), but still, the issues of time-varying dimensionality and neural dynamics remain.

In the *Dynamo* technique, the estimation of each model is carried out with the Extended Kalman Filter (EKF) algorithm. As the estimation assumes Markovian properties, the estimation of the sources is recursive and only requires the previous estimated state and the new available measurements. One advantage with respect to classical solutions such as the dipole fitting algorithm (Mosher et al., 1992) is that *Dynamo* avoids the execution of nonlinear optimization algorithms that require long execution times, and also that the solution embodies a smoothness in the temporal sequence of the estimation. The final estimation at each time is the probabilistic selection across estimations of all banks, which provides a straightforward way to deal with changes in the number of neural sources and/or its underlying dynamics (multiple hypotheses of neural sources are considered in an integrated way). In addition, the estimation is a probability distribution of the sources' position and moment (parameterized by a mean and its covariance) and not solely a discrete point estimate. This allows for a direct

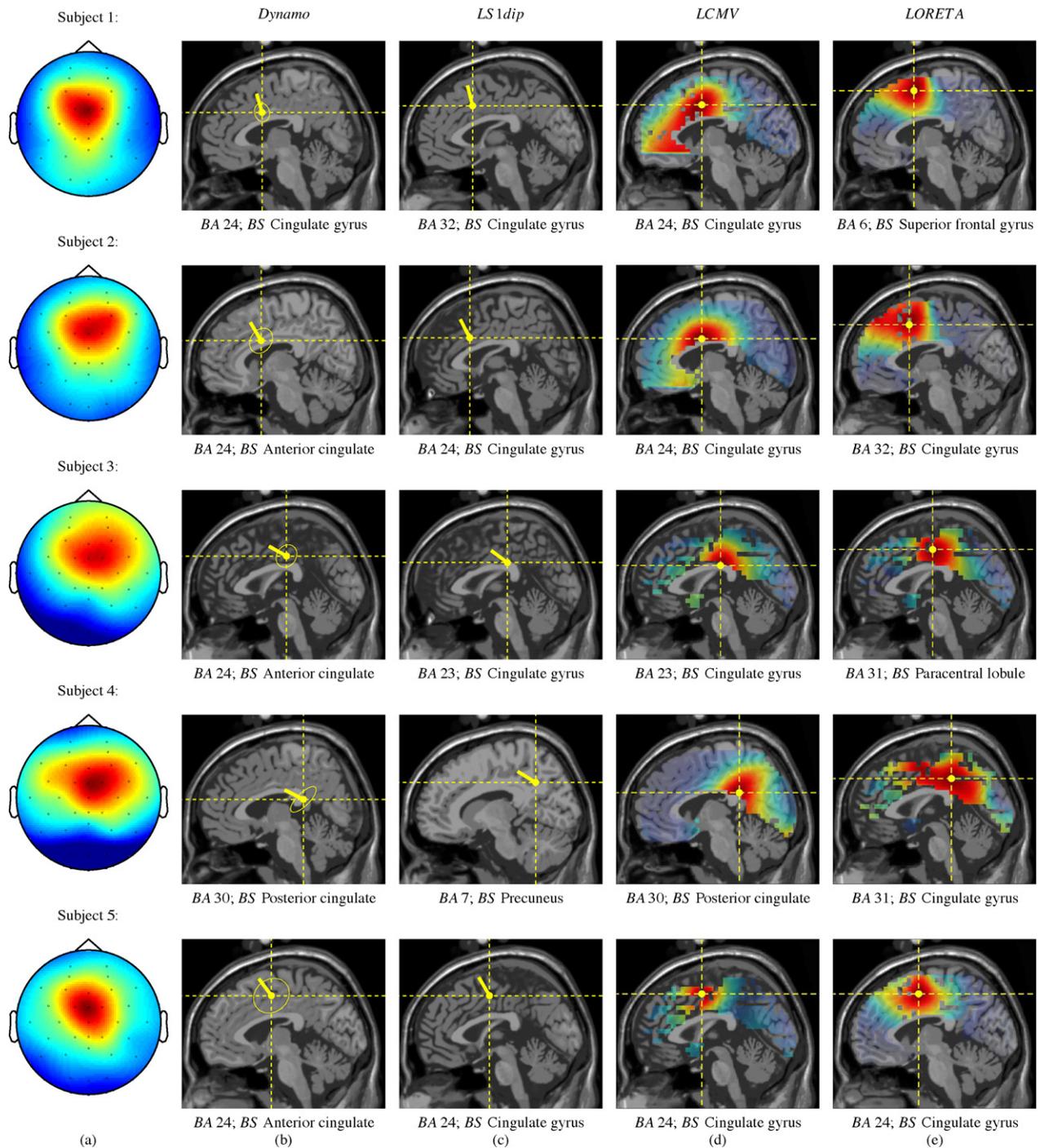


Fig. 9. (a) Topographic scalp maps at the time latency of the second positive peak (≈ 379 ms for all subjects). A positive focalized fronto-central scalp pattern is observed for subjects 1, 2, 3 and 5, and also a positive but less focalized scalp pattern for subject 4. (b, c, d, e) Neural sources for this time latency estimated with *Dynamo*, *LS1dip*, *LCMV*, and *LORETA* projected onto sagittal MRI slices.

representation of probability volumes or confidence intervals for the position and moment of the sources, as already reported to be necessary in dipolar solutions of the EEG/MEG source localization problem (Braun et al., 1997; Fuchs et al., 2004).

In the first iteration of *Dynamo*, all models were randomly initialized. Note that an initial estimate could be computed using the first EEG/MEG observation with a classical dipolar solution (Antelis and Mínguez, 2009a). Such computation leads to all the models with roughly the same initial prior, which would be naturally taken into account in the interaction step by a global re-initializing due to redundancy. The redundancy filter is an important part of the

algorithm, as it avoids the concentration of similar solutions in the same brain area by re-distributing redundant estimations in other brain areas for exploration. In this way, sudden changes in the location of the brain activity can be addressed, where a new brain area separated from the actual activity begins to be de/activated.

The implementation of *Dynamo* utilized three-shell (for the EEG case) and one-shell (for the MEG case) spherical head models to solve the EEG/MEG forward model. The spherical model is a sufficient description for MEG as these signals are unaffected by conductivity profiles of head tissues, but it is a rough simplification for EEG due to the inhomogeneous and anisotropic electric

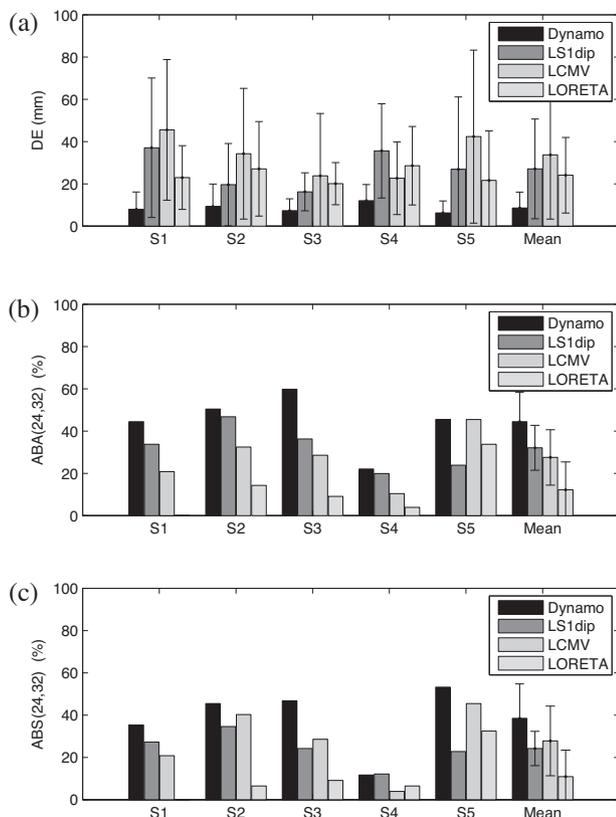


Fig. 10. For the neural sources of the grand average difference between the error and correct conditions estimated with *Dynamo*, *LS1dip*, *LCMV*, and *LORETA*. (a) distance to the error-processing regions *DE* averaged for time interval between 200 and 500 ms, and activation of the error-related (b) Brodmann areas *ABA*(24, 32) and (c) brain structures *ABS*(24, 32) computed for the time interval from 200 to 500 ms.

properties of the head. The rationale behind the use of these spherical models is that they are described by analytic equations expressed in closed form, which allows for the computation and evaluation of the first order derivatives as required by the EKF algorithm. Additionally, previous studies have demonstrated that a three-shell head model could be appropriate when the brain activity is focalized (Vanrumste et al., 2002; Vatta et al., 2010), which is the main context of applicability of *Dynamo*.

Dynamo has been tested in a simulated scenario, where the neural sources were designed to address simultaneously three aspects that impose difficulties to existing source localization techniques: (i) variability in the number of sources; (ii) changes in the anatomical location of the sources, ranging from superficial (neocortical) to deep (subcortical); and (iii) changes in the temporal correlation of sources. The results suggested that the best compromise was achieved with 32 filters and 64 EEG sensors, which were used in subsequent evaluations of the technique. The evaluation results showed a high accuracy in all estimation situations, irrespective to the sensor modality. In particular, the accuracy in position achieved with the EEG sensor modality ($\varepsilon(\mathbf{r}_q) = 12.06\%$) was in the same range of other studies (e.g., equivalent position error of 11.63% in Kiebel et al., 2008). Also note that the values of $\varepsilon(\mathbf{r}_q)$ were lower than $\varepsilon(\mathbf{q})$. This could be explained by the fact that the mathematical equations of the EEG/MEG forward model are linear functions with respect to the moment, but nonlinear with respect to the position, whereby the estimation of the position is prone to higher inaccuracies due to linearizations within the EKF algorithm.

The performance of *Dynamo* was compared with other widely accepted techniques: significant differences and lower values of $\varepsilon(\mathbf{r}_q)$ and $\varepsilon(\mathbf{q})$ were obtained, which supports the fact that the

proposed method was more accurate than other techniques for this evaluation scenario (note that the setup was selected to cover difficult estimation scenarios for existing techniques). The lower performance of *LS* dipolar solutions could be attributed to the fact that in some conditions, the number of sources disagreed with the existing number of sources. The lower performance of the distributed solutions could be due to the fact that these solutions are well conditioned to estimate widespread (and not focalized) brain activity.

The method was also tested in a real scenario with five subjects, where the presentation of negative/positive feedback produced an event-related response mainly generated in the Brodmann areas 24 and 32 in the anterior cingulate cortex (Miltner et al., 1997; Carter et al., 1998; Bush et al., 2000; Holroyd and Coles, 2002 among others), which was used as *ground truth*. The neural sources of the grand average difference for all subjects were estimated in the brain regions that mediate in error-processing, and were similar with the other techniques. In addition, the *Dynamo* estimation during the evoked activity was more accurate and led to significant higher degree of activation than the other methods, for all subjects and per subject. The results show that *Dynamo* has a high performance in the context of error-related potentials, although further investigation is required to fully understand the advantages/disadvantages of the proposed technique in other real contexts.

Finally, it is worth to mention that the method could be used, for example, in applications that require the study of the temporal evolution of neural sources, or in cases where statistics of sources, position, and moment should be computed over different conditions or over different subjects (as a small set of sources' parameters is more appropriate than a high dimensional space provided by distributed methods). *Dynamo* is well-suited for applications whether the brain activity is actually focalized. For cases where the brain activity is widespread, the estimation of the neural sources assuming a dipolar approach will not reflect a real brain activity, and therefore the use of a distributed method could be preferred.

References

- Antelis JM, Minguez J. Dynamic solution to the EEG source localization problem using Kalman filters and particle filters. In: Conf Proc IEEE Eng Med Biol Soc; 2009a. p. 77–80.
- Antelis JM, Minguez J. EEG source localization based on dynamic bayesian estimation techniques. Int J Bioelectromagn 2009b;11(4):179–84.
- Arulampalam S, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. IEEE Trans Sig Process 2002;50(2):174–88.
- Baillet S, Garnero L. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans Biomed Eng 1997;44(5):374–85.
- Baillet S, Mosher J, Leahy R. Electromagnetic brain mapping. IEEE Eng Med Biol Mag 2001;18(6):14–30.
- Bar-Shalom Y, Li XR, Kirubarajan T. Estimation with applications to tracking and navigation: theory, algorithms and software. Wiley; 2001.
- Barton M, Robinson P, Kumar S, Galka A, Durrant-Whyte H, Guivant J, et al. Evaluating the performance of kalman-filter-based EEG source localization. IEEE Trans Biomed Eng 2009;56(1):122–36.
- Bertrand C, Ohmi M, Suzuki R, Kado H. A probabilistic solution to the MEG inverse problem via MCMC methods: The reversible jump and parallel tempering algorithms. IEEE Trans Biomed Eng 2001;48(5):533–42.
- Braun C, Kaiser S, Kincses WE, Elbert T. Confidence interval of single dipole locations based on EEG data. Brain Topogr 1997;10(1):31–9.
- Bush G, Luu P, Posner M. Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 2000;4(6):215–22.
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 1998;280:747–9.
- Darvas F, Pantazis D, Kucukaltun-Yildirim E, Leahy R. Mapping human brain function with MEG and EEG: methods and validation. Neuroimage 2004;23(1):289–99.
- Doucet A, de Freitas N, Gordon N. Sequential Monte Carlo methods in practice. Springer-Verlag; 2001.
- Doucet A, Godsill S, Andrieu C. On sequential monte carlo sampling methods for bayesian filtering. Stat Comput 2000;10(3):197–208.
- Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM. 3D statistical neuroanatomical models from 305 MRI volumes. IEEE Nucl Sci Symp Conf Rec 1999;3:1813–7.

- Fuchs M, Wagner M, Kastner J. Confidence limits of dipole source reconstruction results. *Clin Neurophysiol* 2004;115(6):1442–51.
- Galka A, Yamashita O, Ozaki T, Biscay R, Valdes-Sosae P. A solution to the dynamical inverse problem of EEG generation using spatiotemporal kalman filtering. *Neuroimage* 2004;23:435–53.
- Grech R, Cassar T, Muscat J, Camilleri K, Abri S, Zerbakis M, et al. Review on solving the inverse problem in EEG source analysis. *J Neuroeng Rehabil* 2007;5(25).
- Grova C, Daunizeau J, Lina JM, Benar C, Benali H, Gotman J. Evaluation of EEG localization methods using realistic simulations of interictal spikes. *Neuroimage* 2006;29(3):734–53.
- Hämäläinen M, Ilmoniemi RJ. Interpreting magnetic fields of the brain: minimum norm estimates. *Med Biol Eng Comput* 1994;31(1):35–42.
- Henson R, Mouchlianitis E, Friston K. MEG and EEG data fusion: simultaneous localization of face-evoked responses. *Neuroimage* 2009;47(2):581–9.
- Hesheng L, Xiaorong G, Schimpf PH, Fusheng Y, Shangkaï G. A recursive algorithm for the three-dimensional imaging of brain electric activity: shrinking LORETA-FOCUSS. *IEEE Trans Biomed Eng* 2004;51(10):1794–802.
- Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002;109(4):679–709.
- Jun SC, George JS, Plis SM, Ranken DM, Schmidt DM, Wood CC. Improving source detection and separation in a spatiotemporal Bayesian inference dipole analysis. *Phys Med Biol* 2006;51(10):2395–414.
- Kiebel S, Daunizeau J, Phillips C, Friston K. Variational bayesian inversion of the equivalent current dipole model in EEG/MEG. *Neuroimage* 2008;39(2):728–41.
- Koles ZJ. Trends in EEG source localization. *Electroencephalogr Clin Neurophysiol* 1998;106(2):127–37.
- Liu AK, Dale AM, Belliveau JW. Monte Carlo simulation studies of EEG and MEG localization accuracy. *Hum Brain Mapp* 2002;16(1):47–62.
- Lopez-Larraz E, Iturrate I, Montesano L, Mínguez J. Real-time recognition of feedback error-related potentials during a time-estimation task. In: *Conf Proc IEEE Eng Med Biol Soc. Buenos Aires*; 2010. p. 2670–3.
- Mattout J, Phillips C, Penny W, Rugg M, Friston K. MEG source localization under multiple constraints: an extended Bayesian framework. *Neuroimage* 2006;30(3):753–67.
- Michel CM, Murraya MM, Lantza G, Gonzalez S, Spinellib L, de Peralta RG. EEG source imaging. *Clin Neurophysiol* 2004;115(10):2195–222.
- Miltner WH, Braun CH, Coles MGH. Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a generic neural system for error detection. *J Cogn Neurosci* 1997;9(6):788–98.
- Molins A, Stufflebeam S, Brown E, Hämäläinen M. Quantification of the benefit from integrating MEG and EEG data in minimum L2-norm estimation. *Neuroimage* 2008;42(3):1069–77.
- Mosher JC, Leahy RM, Lewis PS. EEG and MEG: Forward solutions for inverse methods. *IEEE Trans Biomed Eng* 1999;46(3):245–59.
- Mosher JC, Lewis PS, Leahy RM. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans Biomed Eng* 1992;39(6):541–57.
- Pascual-Marqui R, Esslen M, Kochia K, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J of psychophysiol* 1994;18(1):49–65.
- Peralta-Menendez RGD, Gonzalez-Andino S. Comparison of algorithms for the localization of focal sources: evaluation with simulated data and analysis of experimental data. *Int J Bioelectromagn* 2002;4(1).
- Phillips C, Mattout J, Rugg MD, Maquet P, Friston KJ. An empirical bayesian solution to the source reconstruction problem in EEG. *Neuroimage* 2005;24:997–1011.
- Plummer C, Harvey AS, Cook M. EEG source localization in focal epilepsy: Where are we now? *Epilepsia* 2008;49(2):201–18.
- Quan QH, Ping PL, Hong LS. A novel interacting multiple model algorithm. *Signal Process* 2009;89(11):2171–7.
- Quraan MA, Cheyne D. Reconstruction of correlated brain activity with adaptive spatial filters in MEG. *Neuroimage* 2010;49(3):2387–400.
- Scherg M. Auditory evoked magnetic fields and electric potentials. Karger; 1990, volume 6. p. 40–69.
- Schmidt DM, George JS, Wood C. Bayesian inference applied to the electromagnetic inverse problem. *Hum Brain Mapp* 1999;7:195–212.
- Sekihara K, Sahani M, Nagarajan SS. Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *Neuroimage* 2005;25:1056–67.
- Sharon D, Hämäläinen MS, Tootell RB, Halgren E, Belliveau JW. The advantage of combining MEG and EEG: Comparison to fMRI in focally stimulated visual cortex. *Neuroimage* 2007;36(4):1225–35.
- Somersalo E, Voutilainen A, Kaipio J. Non-stationary magnetoencephalography by bayesian filtering of dipole models. *Inverse Probl* 2003;19(17):1047–63.
- Sorrentino A, Parkkonen L, Pascarella A, Campi C, Piana M. Dynamical MEG source modeling with multi-target bayesian filtering. *Hum Brain Mapp* 2009;30(6):1911–21.
- Sorrentino A, Parkkonen L, Piana M. Particle filters: a new method for reconstructing multiple current dipoles from MEG data. In: *New Frontiers in Biomagnetism*, vol. 1300; 2007. p. 173–6, *New Frontiers in Biomagnetism. Proceedings of the 15th international conference on biomagnetism*, Vancouver, BC, Canada, August 21–25, 2006.
- Sorrentino A, Pascarella A, Campi C, Piana M. Particle filters for the magnetoencephalography inverse problem: increasing the efficiency through a semi-analytic approach (rao-blackwellization). *J Phys Conf Ser* 2008;124(1).
- Trujillo-Barreto NJ, Aubert-Vazquez E, Penny WD. Bayesian M/EEG source reconstruction with spatio-temporal priors. *Neuroimage* 2008;39:318–35.
- Trujillo-Barreto NJ, Aubert-Vazquez E, Valdes-Sosa PA. Bayesian model averaging in EEG/MEG imaging. *Neuroimage* 2004;21(4):1300–19.
- Vanrumste B, Hoey GV, de Walle RV, D'Havé MRP, Lemahieu IA, Boon PAJM. Comparison of performance of spherical and realistic head models in dipole localization from noisy eeg. *Med Eng Phys* 2002;24(6):403–18.
- Vatta F, Meneghini F, Esposito F, Mininell S, Salle FD. Realistic and spherical head modeling for EEG forward problem solution: a comparative cortex-based analysis. *Comput Intell Neurosci* 2010:2010.
- Veen BDV, van Dronglen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* 1997;44(9):867–80.
- Welch G, Bishop G. An introduction to the Kalman Filter. Technical report; UNC-Chapel Hill; 2006.
- Whittingstall K, Stroink G, Gates L, Connolly J, Finley A. Effects of dipole position, orientation and noise on the accuracy of EEG source localization. *Biomed Eng Online* 2003;2(14).
- Wipf D, Nagarajan S. A unified bayesian framework for MEG/EEG source imaging. *Neuroimage* 2009;14(3):947–66.
- Yamashita O, Galka A, Ozaki T, Biscay R, Valdes-Sosa P. Recursive penalized least squares solution for dynamical inverse problems of EEG generation. *Hum Brain Mapp* 2004;21:221–35.