Brain activity mapping from MEG data via a hierarchical Bayesian algorithm with automatic depth weighting

Daniela Calvetti[*], Annalisa Pascarella[†], Francesca Pitolli[‡]
Erkki Somersalo[§], Barbara Vantaggi[¶]

June, 2018

Abstract

A recently proposed Iterated Alternating Sequential (IAS) MEG inverse solver algorithm, based on the coupling of a hierarchical Bayesian model with computationally efficient Krylov subspace linear solver, has been shown to perform well for both superficial and deep brain sources. However, a systematic study of its ability to correctly identify active brain regions is still missing. We propose novel statistical protocols to quantify the performance of MEG inverse solvers, focusing in particular on how their accuracy and precision at identifying active brain regions. We use these protocols for a systematic study of the performance of the IAS MEG inverse solver, comparing it with three standard inversion methods, wMNE, dSPM, and sLORETA. To avoid the bias of anecdotal tests towards a particular algorithm, the proposed protocols are Monte Carlo sampling based, generating an ensemble of activity patches in each brain region identified in a given atlas. The performance in correctly identifying the active areas is measured by how much, on average, the reconstructed activity is concentrated in the brain region of the simulated active patch. The analysis is based on Bayes factors, interpreting the estimated current activity as data for testing the hypothesis that the active brain region is correctly identified, vs. the hypothesis of any erroneous attribution. The methodology allows the presence of a single or several simultaneous activity regions, without assuming that the number of active regions is known. The testing protocols suggest that the IAS solver performs well with both with cortical and subcortical activity estimation.

Keywords: MEG inverse problem; Activity map; Brain region; Bayes factor; Deep sources

1 Introduction

In the ongoing quest for understanding brain functions and connectivity, during specific tasks and in resting states, with the hope of finding a functional fingerprint of autism, pathological depression or schizophrenia, the potential of MEG as a non-invasive and non-disturbing brain activity mapping modality cannot be emphasized enough. The high temporal resolution of MEG

---

[*]Case Western Reserve University, Department of Mathematics, Applied Mathematics and Statistics, 10900 Euclid Avenue, Cleveland, OH 44106, e-mail: `dxc57@case.edu`

[†]CNR - National Research Council, Istituto per le Applicazioni del Calcolo "Mauro Picone", Via dei Taurini 19, 00185 Rome, Italy, e-mail: `a.pascarella@iac.cnr.it`

[‡]University of Roma "La Sapienza", Department SBAI, Via A. Scarpa 16, 00161 Roma, Italy e-mail: `francesca.pitolli@sbai.uniroma1.it`

[§]Case Western Reserve University, Department of Mathematics, Applied Mathematics and Statistics, 10900 Euclid Avenue, Cleveland, OH 44106, e-mail: `ejs49@case.edu`

[¶]University of Roma "La Sapienza", Department SBAI, Via A. Scarpa 16, 00161 Roma, Italy, e-mail: `barbara.vantaggi@sbai.uniroma1.it`

offers enormous possibilities for understanding the fine details of brain dynamics, however the reliable reconstruction of activity patterns in cortical and deep brain regions relies on the capability of the underlying inverse solver to correctly identify active brain areas. In this work we propose a comprehensive, statistically sound methodology to assess the performance of different inverse solvers, and demonstrate the viability of the approach by applying it to four different algorithms.

A fundamental difficulty in the solution of the MEG inverse problem (Hämäläinen *el al.*, 1993, Baillet *et al.*, 2001, Brette *el al.*, 2012) comes from the non-uniqueness of the solution and the high sensitivity of the problem to noise in the data. To address the non-uniqueness, it is necessary to augment the data with additional information about the solution, which entails either the introduction of regularization techniques in a deterministic setting, or the introduction of prior models in the Bayesian framework. The sensitivity of the computed solution to noise, with low signal-to-noise ratio of data, in turn, requires a good understanding of the noise sources, which in the Bayesian framework is related to a design of a reliable likelihood. These considerations highlight the importance of being able to assess the reliability of a given inversion algorithm in the multitude of tasks it may be applied to, and to quantify the uncertainty in computed solutions.

In this article, we present a systematic study of the accuracy and precision of the Iterative Alternating Sequential (IAS) MEG inversion algorithm (Calvetti *et al.*, 2009, 2015), which pairs hierarchical Bayesian modeling with a computationally very efficient prior-based preconditioned iterative linear solver. Numerous algorithms based on hierarchical Bayesian modeling have been suggested in the literature, see, e.g., (Auranen *et al.*, 2005, Calvetti *et al.*, 2009, Henson *el al.*, 2009, 2010, Kiebel *el al.*, 2008, Lucka *et al.*, 2012, Lopez *et al.*, 2014, Mattout *et al.*, 2006, Nummenmaa *et al.*, 2007,b, Owen *et al.*, 2012, Sato *et al.*, 2004, Stephan *et al.*, 2009, Trujillo-Barreto *et al.*, 2004, Wipf and Nagarajan, 2009, Wipf *et al.*, 2010). A common feature of these methods is the use of parameter-dependent prior models, the model parameters representing the second layer of unknowns. A standard approach in hierarchical Bayesian modeling is to introduce a hyperprior model for the parameters, and to either marginalize, or model average, the parameters, or to estimate them by maximizing the evidence, see, e.g., Bernardo and Smith (2004) for general reference. A more time-consuming approach to the hierarchical modeling avoiding optimization techniques is to use Markov chain Monte Carlo (MCMC) methods.

The IAS MEG algorithm, too, is based on the use of hierarchical, conditionally Gaussian model for the prior. The hyperparameters are the variances of the elementary sources in the current density model, and are adjusted by learning focality from data. To reduce the computational complexity of the problem, especially in case of high dimensionality of the MEG data, the algorithm approximates the maximum of the posterior distribution instead of computing the posterior distribution itself. The iterative alternating sequential minimization procedure was shown to converge in Calvetti *et al.* (2015), where an easily implementable procedure for the estimation of the maximizer of the posterior with respect to both the unknown of primary interest and the hyperparameters was proposed. The maximization algorithm uses a dynamically preconditioned Krylov subspace iterative solver equipped with a stopping rule that monitors the fidelity of the solution to data and stops the iteration before amplified noise components begin to degrade the computed solution. The solution of the regularized subproblem does not rely on a Tikhonov-regularized linear solver, but on a very fast converging preconditioned Conjugate Gradient Least Squares (CGLS) algorithm, which makes the solution depend non-linearly on the data (Calvetti *et al.*, 2018). Moreover, the IAS MEG inverse method employs an anatomically justified prior: Based on the segmented subject-specific MRI image containing the information of the cortical surface orientation, for each dipole there is a direction which is favored, but not forced, by the prior model. Similar ideas with different implementation have been suggested in the literature, see. e.g. (Lin *el al.*, 2006b).

Some of the appeal of standard MEG solvers such as MNE (Hämäläinen MS and Ilmoniemi RJ, 1984, Lin *el al.*, 2006a) or LORETA (Pasqual-Marqui, 1999) is their simplicity of use, with minimal user intervention required, e.g., for setting parameters or controlling the optimization

process, which is not always the case with hierarchical models. One of the goals in developing the IAS algorithm was to have a robust and convergent method that depends on very few user-supplied parameters with an intuitive meaning. In Calvetti *et al.* (2009), the connection between the conditionally Gaussian hierarchical models and several sparsity-promoting methods (Gorodnitsky and Rao, 1997, Nagarajan *et al.*, 2006, Uutela *et al.*, 1999) was pointed out.

A novel contribution of this paper is to provide a physical interpretation of the hyperparameters and to establish a connection with sensitivity weighting. The IAS algorithm, as it was rigorously shown in Calvetti *et al.* (2015), contains only one user-supplied parameter that controls the sparsity of the solution and, as pointed out in Section 2.2, the other hyperparameters can be semi-automatically set based on the information about the signal-to-noise ratio. This formulation provides a proper Bayesian interpretation for the sensitivity weighting that is commonly used with, e.g., the Minimum Norm Estimate (MNE) and the Minimum Current Estimate (MCE) (Uutela *et al.*, 1999) algorithms to overcome the propensity of the methods to favor superficial sources over deep ones, thereby filling the gap between Bayesian modeling and traditional regularization.

The quest for MEG inverse solvers capable of identifying activities both on superficial cortex and in deeper brain regions continues to be of interest. In Attal *et al.* (2012), Attal and Schwartz (2013) the localization accuracy of subcortical generators was studied for three different algorithms: the weighted Minimum Norm Estimate (wMNE), dynamic Statistical Parametric Mapping (dSPM), and standardized Low Resolution Brain Electromagnetic Tomography (sLORETA). The procedure relies on a realistic electrophysiological model of deep brain activity and compares the performance of the three methods in recovering activity in the neocortex, hippocampus, amygdala and thalamus. In the present study the comparison is extended to include the IAS method using a model of the subcortical regions that considers their surface envelopes, adding basal ganglia, brainstem and cerebellum to the list of brain regions, and introducing a statistically justified metric to evaluate the performances of the different algorithms.

More specifically, our investigation addresses the following question: Given an atlas of anatomo-functional brain regions (BR), how well does the IAS method for the MEG inversion identify active regions, and to what extent non-active regions will be correctly deemed as inactive ones? With a slight abuse of terminology, we may refer to the former question as sensitivity of the MEG inverse solver, and to the latter as specificity, elaborating on the reasons for this choice further below. In line with the Bayesian paradigm that IAS is based on, we investigate these questions using statistical tools. For this purpose we introduce an Activity Level Indicator (ALI), measuring the mean activity level of each BR given a simulated patch activity in a selected BR, and compute the first and second order statistics of this indicator over a sample of random activation patches in the selected BR. To assess the relative performance of IAS, we compare these statistics with those of three other standard inversion methods, wMNE (Lin *el al.*, 2006a), dSPM (Dale *et al.*, 2000), and sLORETA (Pasqual-Marqui, 1999, Wagner *et al.*, 2004), using the implementations provided in Brainstorm (Tadel *et al.*, 2011). Furthermore, to test the relative performance of the four methods, we observe that a common approach in the Bayesian methodology to test hypotheses or models is to use Bayes factors of probabilities of the data, given the competing hypotheses or models (Bernardo and Smith, 2004, Kass RE and Raftery AE, 1995). In the current setting, we interpret the estimated activity as data, and test the probabilities of them under the competing hypotheses of having activity in the correct BR versus having it in another randomly chosen BR. We compare the specificity of the IAS algorithm to that of the other three standard solvers using simulated data under two different scenarios: when there is only one active patch in a cortical or subcortical region, and when there are multiple (from two to six) active patches in cortical and subcortical regions. In this way we can test not only the sensitivity and specificity of each inverse solver with respect to a single BR, but also their ability of recovering an activation pattern when other regions disturb the identification. We conclude the computed experiments by comparing the reconstructions obtained by the four inverse solvers in an example starting from real data, where

the underlying activation pattern is not known exactly but is inferred from the protocol used for the data collection.

2 Materials and Methods

We start by presenting a brief review of the *Iterative Alternating Sequential* (IAS) MEG inverse solver, described in detail in Calvetti *et al.* (2015). Subsequently, we introduce the computational tools that will be used for testing the algorithm's performance, focusing in particular on its capacity to identify active areas of interest from the MEG data and comparing it to three other mainstream MEG inverse solvers.

To discretize the problem, we approximate the cortical surface, as well as the surface of the subcortical structures, by a triangular mesh with n vertices, whose coordinates are denoted by v_j. We associate to each vertex a unit length vector $\vec{\nu}_j$ pointing in the direction normal to the local cortical/subcortical surface.

The forward model is of the form

$$b = \sum_{j=1}^{n} \mathsf{M}_j \vec{q}_j + \varepsilon, \tag{1}$$

where $b \in \mathbb{R}^m$ is the observation vector representing the measured magnetic field components, $\mathsf{M}_j \in \mathbb{R}^{m \times 3}$ is the lead field matrix associated with the jth dipole located at v_j, $\vec{q}_j$ is the jth dipole moment and $\varepsilon \in \mathbb{R}^m$ is additive observation noise. The lead field matrices take into account the conductivity structure of the head model. While changing the forward model does not affect the inversion algorithm, the results are to some extent sensitive to the model used (cf. Vorwerk *el al.* (2014)).

We model the noise term as a zero mean Gaussian random variable, $\varepsilon \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathbb{R}^{m \times m}$ is the covariance matrix. To model biological noise, the covariance needs to reflect the cross correlations between the channels, as will be discussed below. Therefore, the likelihood density of b conditioned on $\vec{q}_1, \ldots, \vec{q}_n$ can be written as

$$\pi(b \mid \vec{q}_1, \ldots, \vec{q}_n) \propto \exp\left(-\frac{1}{2} \| b - \sum_{j=1}^{n} \mathsf{M}_j \vec{q}_j \|_{\Sigma}^2 \right),$$

where $\| z \|_{\Sigma}^2 = z^{\mathsf{T}} \Sigma^{-1} z$ is the square of the Mahalanobis norm and "$\propto$" stands for "proportional up to a normalizing constant". In our computer simulation a small exogenous noise component is added to the biological noise.

2.1 The inverse solver: an overview

The IAS algorithm is based on a Bayesian hierarchical prior model of the activity of a single dipole of the form

$$\pi_{\text{prior}}(\vec{q}_j \mid \theta_j) \sim \mathcal{N}(0, \theta_j \mathsf{C}_j) \propto \frac{1}{\theta_j^{3/2}} \exp\left(-\frac{\vec{q}_j^{\mathsf{T}} \mathsf{C}_j^{-1} \vec{q}_j}{2\theta_j} \right) = \exp\left(-\frac{\| \vec{q}_j \|_{\mathsf{C}_j}^2}{2\theta_j} - \frac{3}{2} \log \theta_j \right);$$

here $\mathsf{C}_j \in \mathbb{R}^{3 \times 3}$ is the local anatomical prior matrix,

$$\mathsf{C}_j = \vec{\nu}_j \vec{\nu}_j^{\mathsf{T}} + \delta(\vec{\xi}_j \vec{\xi}_j^{\mathsf{T}} + \vec{\zeta}_j \vec{\zeta}_j^{\mathsf{T}}),$$

where $0 < \delta < 1$ and $(\vec{\xi}_j, \vec{\zeta}_j, \vec{\nu}_j)$ is a local orthonormal frame at the j-th vertex and $\vec{\nu}_j$ is orthogonal to the cortical/subcortical surface. The parameter $\theta_j > 0$, scaling the local prior covariance of the dipole $\vec{q}_j$, is modeled further as a random variable following the gamma distribution,

$$\theta_j \sim \Gamma(\theta_j^*, \beta_j) \propto \theta_j^{\beta_j - 1} \exp\left(-\frac{\theta_j}{\theta_j^*}\right).$$

The scaling parameter θ_j^* of the gamma density can be chosen to express our a priori belief about the order of magnitude of the expected value of the variance θ_j, while the value of the shape parameter β_j controls the sparsity of the solution. In all computed experiments the latter is held constant for all dipoles, that is, $\beta_j = \beta$, while θ_j^* is chosen by an empirical Bayes argument to correspond to a sensitivity scaling. The details and justification for the choice of the hyperparameters can be found in the Appendix (see also the original article Calvetti *et al.* (2015)). By Bayes' theorem the posterior model can be written as

$$\pi(\vec{q}_1, \ldots, \vec{q}_n, \theta_1, \ldots, \theta_n \mid b) \propto \prod_{j=1}^{n} \pi_{\text{prior}}(\vec{q}_j \mid \theta_j) \pi(\theta_j \mid \theta_j^*, \beta) \pi(b \mid \vec{q}_1, \ldots, \vec{q}_n)$$

$$\propto \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\frac{\|\vec{q}_j\|_{\mathsf{C}_j}^2}{\theta_j} + \sum_{j=1}^{n}\left[(\beta - \frac{5}{2})\log\theta_j - \frac{\theta_j}{\theta_j^*}\right] - \frac{1}{2}\|b - \sum_{j=1}^{n}\mathsf{M}_j\vec{q}_j\|_{\Sigma}^2\right) \quad (2)$$

The IAS algorithm for computing an estimate of the maximum a posteriori (MAP) solution of (2), is based on the following alternating iterative minimization scheme:

1. Initialize $\theta_j = \theta_j^*$, $1 \leq j \leq n$, and set $k = 0$;

2. Until the convergence criterion is met:

    (i) Update $\vec{q}_1, \ldots, \vec{q}_n$, setting
    $$(\vec{q}_1^{k+1}, \ldots, \vec{q}_n^{k+1}) = \text{argmax}\{\pi(\vec{q}_1, \ldots, \vec{q}_n, \theta_1^k, \ldots, \theta_n^k \mid b)\};$$

    (ii) Update $\theta_1, \ldots, \theta_n$, setting
    $$(\theta_1^{k+1}, \ldots, \theta_n^{k+1}) = \text{argmax}\{\pi(\vec{q}_1^{k+1}, \ldots, \vec{q}_n^{k+1}, \theta_1, \ldots, \theta_n \mid b)\};$$

    (iii) Increase $k \to k + 1$.

It was proved in Calvetti *et al.* (2015) that the above optimization algorithm converges to a unique local maximizer. Moreover, the iterations can be performed effectively by noticing that when minimizing the energy function defined as

$$\mathcal{E}(\vec{q}_1, \ldots, \vec{q}_n; \theta_1, \ldots, \theta_n) = \overbrace{\|b - \sum_{j=1}^{n}\mathsf{M}_j\vec{q}_j\|_{\Sigma}^2}^{(a)} + \underbrace{\sum_{j=1}^{n}\frac{\|\vec{q}_j\|_{\mathsf{C}_j}^2}{\theta_j} - 2\sum_{j=1}^{n}\left[(\beta - \frac{5}{2})\log\theta_j - \frac{\theta_j}{\theta_j^*}\right]}_{(b)}, \quad (3)$$

the optimization with respect to $\vec{q}$, affecting only part (a) of the above expression, reduces to a quadratic minimization problem, while the minimization with respect to the hyperparameters θ_j, depending only on part (b) of the energy function, can be done independently of other components, and admits a solution in closed form. Finally, we point out that an approximate solution of the quadratic minimization problem can be found very efficiently using a priorconditioned CGLS algorithm (see: Figure 1). For the omitted details about the algorithm and a discussion about the meaning of w in Figure 1, we refer to (Calvetti *et al.*, 2015).

Box 1: IAS algorithm

Variables:

- $q = [\vec{q}_1^{\mathsf{T}}, \ldots, \vec{q}_n^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^{3n}$

- $w = [w_1, \ldots, w_{3n}]^{\mathsf{T}} \in \mathbb{R}^{3n}$ (auxiliary variable)

- $\theta = [\theta_1, \ldots, \theta_n]^{\mathsf{T}} \in \mathbb{R}^n$

Given:

- $\theta^* = [\theta_1^*, \ldots, \theta_n^*]^{\mathsf{T}} \in \mathbb{R}^n$

- $\beta > 5/2$

- $\mathsf{M} = \begin{bmatrix} \mathsf{M}_1 & \cdots & \mathsf{M}_n \end{bmatrix} \in \mathbb{R}^{m \times 3n}$ (lead field matrix)

- $\mathsf{C} = \mathrm{diag}\left\{ \begin{matrix} \mathsf{C}_1 & \cdots & \mathsf{C}_n \end{matrix} \right\} \in \mathbb{R}^{3n \times 3n}$ (covariance of the anatomical prior)

Algorithm:

- Initialize $\theta \to \theta^*, \quad w \to 0$

- Until $\|\theta - \theta_{\mathrm{old}}\| < \tau \|\theta\|$, τ = given tolerance

  1. $\mathsf{C}_\theta = \mathrm{diag}\left\{ \begin{matrix} \theta_1 \mathsf{C}_1 & \cdots & \theta_n \mathsf{C}_n \end{matrix} \right\}$, $\mathsf{D}_\theta^{\mathsf{T}} \mathsf{D}_\theta = \mathsf{C}_\theta$, $\mathsf{A}_\theta = \mathsf{M} \mathsf{D}_\theta^{-\mathsf{T}}$

  2. Initialize: $d \to b - \mathsf{A}_\theta w, \quad r \to \mathsf{A}_\theta^{\mathsf{T}} d, \quad p \to r$

  3. Until $\|d\| < \sqrt{m}$ repeat:
  $$\begin{cases} y \to \mathsf{A}_\theta p \\ \alpha \to r^{\mathsf{T}} r / y^{\mathsf{T}} q \\ w \to w + \alpha\, p \\ d \to d - \alpha y \\ \tilde{r} \to \mathsf{A}_\theta^{\mathsf{T}} d \\ p \to r + (r^{\mathsf{T}} r)/\tilde{r}^{\mathsf{T}} \tilde{r}\, \tilde{r} \\ r \to \tilde{r} \end{cases}$$

  4. $q \to \mathsf{D}_\theta^{-\mathsf{T}} w, \quad \mathsf{V} = [\|\vec{q}_1\|_{\mathsf{C}_1}^2 / \theta_1^*, \ldots, \|\vec{q}_n\|_{\mathsf{C}_n}^2 / \theta_n^*]^{\mathsf{T}}$

  5. $\theta_{\mathrm{old}} = \theta$

  6. $\theta = \frac{1}{2} \theta^* \left(\beta - \frac{5}{2} + \sqrt{\left(\beta - \frac{5}{2}\right)^2 + 2\mathsf{V}} \right)$

Figure 1: An outline of the IAS algorithm.

2.2 Initializing the IAS algorithm

To initialize the IAS algorithm we have to choose the scaling parameters θ_j^*, $1 \le j \le n$. This can be done observing that θ_j^* can be viewed as a sensitivity weight that takes into account the distribution of the active sources and the signal-to-noise ratio SNR. If we assume

$$\varepsilon \sim \mathcal{N}(0, \Sigma),$$

where the covariance matrix Σ is symmetric positive definite, not necessarily diagonal, the SNR is defined as

$$\text{SNR} = \frac{\mathsf{E}\{\|b\|^2\}}{\mathsf{E}\{\|\varepsilon\|^2\}} = \frac{\mathsf{E}\{\|b_0\|^2\}}{\mathsf{E}\{\|\varepsilon\|^2\}} + 1, \quad \text{SNR}_{\text{dB}} = 10 \log_{10}(\text{SNR}), \tag{4}$$

and

$$\mathsf{E}\{\|\varepsilon\|^2\} = \text{trace}(\Sigma).$$

Furthermore, assuming that the a priori expectation of the number of active dipoles is $\bar{k}$, we get

$$\theta_j^* = \sum_{k=1}^{n} \frac{p_k}{k} \frac{(\text{SNR} - 1) \times \text{trace}(\Sigma)}{\beta \|\mathsf{M}_j \mathsf{C}_j^{1/2}\|_{\text{F}}^2}, \quad 1 \le j \le n, \quad p_k \sim \text{Poisson}(\bar{k}),$$

where the subscript F denotes the Frobenius norm (see the Appendix for details).

We point out that the expectation of the variance of a single dipole cannot exceed a physiologically meaningful upper bound $\theta_{\max}$, which in practice means that we need to choose θ_j^* as

$$\theta_j^* = \min\left\{\theta_{\max}, \frac{C}{\|\mathsf{M}_j \mathsf{C}_j^{1/2}\|_{\text{F}}^2}\right\},$$

where the constant C comprises the summation over k above. Thus, the analysis leaves only three parameters to be selected by the user: $(\beta, \bar{k}, \theta_{\max})$, each of which has a clear physiological interpretation.

2.3 Test protocols

The protocols for the validation of the IAS algorithm focus on the sensitivity of the method, by which we intend how well the method is able to identify an active area, and on its specificity, intended as how often the algorithm misidentifies a non-active area as active. The validation methodology discussed here is general and can be used for any MEG inversion algorithm, therefore providing a flexible platform for comparing the performance of different methods.

In the sequel, we assume that the source space representing the brain is divided into anatomo-functional BRs following a given atlas. We denote by L the number of BRs in the atlas, and assume that every vertex v_j can be uniquely attributed to a single BR.

2.3.1 Patch activity generation

We investigate the activity attribution with synthetic data produced by the activity of a patch $\mathcal{P}$ in a given anatomo-functional BR. The patch activity is generated with a Monte Carlo algorithm to account for the effect of random source-dependent variations in the data on the reconstructions. The activity patch $\mathcal{P}$ consists of a preselected number $N_\mathcal{P}$ of vertices and corresponding dipole moments. (see the Appendix for details on the construction of $\mathcal{P}$).

Let $\mathcal{P} = \{(v_1, \vec{q}_1), \ldots, (v_{N_\mathcal{P}}, \vec{q}_{N_\mathcal{P}})\}$ denote the coordinates of the vertices and the corresponding dipole moments on the patch $\mathcal{P}$. For later reference, we define the barycenter of the activity patch by the formula

$$m_\mathcal{P} = \frac{1}{\sum_{v_j \in \mathcal{P}} \|\vec{q}_j\|} \sum_{v_j \in \mathcal{P}} v_j \|\vec{q}_j\|, \tag{5}$$

where the summation is over all vertices in the patch. Observe that the barycenter of an activity patch need not coincide with any of the nodes; in fact, there is no guarantee that the barycenter is inside the BR of interest due to the lack of convexity of the brain regions.

When testing the performance of the algorithms, we consider two simulation protocols: A single active BR, or several active BRs. For both protocols, we generate a sample $\mathcal{B} = \{b^1, b^2, \ldots, b^K\}$ of data vectors as follows. In the single active BR case, we select one of the regions from the atlas, and generate K independent random activity patches in the selected BR. For each activity patch, we compute the magnetic field components at the sensor locations with the model (1), then add simulated brain noise. In the second protocol, we consider activity occurring simultaneously in N different BRs. For each of the N active regions, we generate independently an activity patch, repeating the process independently K times, and compute the magnetic field data at the sensors, corrupting the data with the additive simulated brain noise as in the single region case. The simulation of biologically justified brain noise is described in the following section.

2.3.2 Simulated biological noise

To acknowledge the fact that a significant portion of noise in MEG is not of exogenous ambient origin but due to the subject itself, we generate additive noise that has a correlation structure characteristic to biological noise. We follow here the random dipole brain noise model in de Munck *et al.* (1992), Huizenga *et al.* (2002). Since the simulations in this work are single time slice simulations, we do not include the temporal correlation component developed in the cited articles.

We generate a small number n_d of dipole sources, drawing the positions independently with uniform distribution over the source space, the dipole moments independently from normal distribution, and compute the magnetic responses at the magnetometers using the lead field matrix. The resulting data vector constitutes a single realization of unscaled brain noise. To compute the covariance, we generate a large number N_d of brain noise realizations, denoted m_{noise}^j, $1 \leq j \leq N_d$, and compute the empirical unscaled covariance Σ_d. To guarantee the positive definiteness of the matrix, and to address the ambient noise, we add a small diagonal contribution to the covariance, leading to the model

$$\Sigma = \alpha \left(\Sigma_d + \delta^2 \text{diag} \left(\Sigma_d \right) \right).$$

In our simulations, we used parameter values $n_d = 20$, $N_d = 5000$, $\delta = 0.01$ and $\alpha = 0.1$. Hence, we assume that the variance of the biological noise is around one tenth of the average power of the random dipoles generating the realizations of the noise. Finally, given the Cholesky decomposition of the covariance matrix, $\Sigma = \mathsf{B}^\mathsf{T}\mathsf{B}$, we may generate random realizations ε of the noise with the prescribed covariance through

$$\varepsilon = \mathsf{B}^\mathsf{T} w, \quad w \sim \mathcal{N}(0, \mathsf{I}),$$

that is, w is a realization of a random standard normal vector. By fixing the noise level, the SNR for the simulated data depends on the location of the simulated activity, SNR being significantly lower for deep sources than for superficial cortical sources. Since both the patch activity and the noise are random, the SNR_{dB}, estimated as the ratio of the squared norms of the signal and the noise, varies from a simulation to another, being around 13 dB for cortical sources, and 7 dB for subcortical deep sources, with a significant variation due, e.g., to the orientation of the activity patch.

2.3.3 Activity level indicator vectors

An important criterion for assessing the performance of an MEG inverse solver for analyzing the brain activity is how reliably it maps the observed data to the correct region of interest. Because it is unreasonable to expect a perfect performance due to the severe ill-posedness of the problem, we propose an evaluation tool to quantify how well the BR where the simulated activity takes place can be identified, and what the most likely confounding regions are. The correct identification of regions further away from the sensors, e.g., deep brain structures, is expected to be most challenging. In general, however, for a reliable solver, it is reasonable to expect that the confounding regions should be anatomically close to the one where the activity occurs. The procedure for quantifying the sensitivity of a MEG inverse solver to the BR of the activity is outlined below.

Let

$$
F : \mathbb{R}^m \to \mathbb{R}^{3n}, \quad b \mapsto \begin{bmatrix} \hat{\vec{q}}_1 \\ \vdots \\ \hat{\vec{q}}_n \end{bmatrix},
$$

denote the map defined by the MEG inversion algorithm, where $\hat{\vec{q}}_j$, $j = 1, \ldots, n$ are the estimated dipole moments at the n vertices. We introduce the estimated brain activity vector $a \in \mathbb{R}^n$, whose jth component is the norm of the jth current dipole found by the MEG inverse solver, that is, $a_j = \|\hat{\vec{q}}_j\|$. Since each entry of the vector a can be associated to one of the BRs in which the corresponding vertex is located, we may define the estimated BR Activity Level Indicator (BR-ALI) vector $\alpha \in \mathbb{R}^L$, corresponding to the brain activity estimate a as

$$
\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_L \end{bmatrix}, \quad \alpha_\ell = \text{mean}\, \{a_j, j \in \mathcal{R}_\ell\},
$$

where $\mathcal{R}_\ell$, $\ell = 1, \ldots, L$, is the set of the vertex indices corresponding to the ℓ-th brain region.

Assuming that the data b are generated by an active patch in the $\tilde{\ell}$th BR alone, if F were a *perfect* MEG inversion algorithm for identifying active regions, the estimated ALI vector would have only the $\tilde{\ell}$th component different from zero, while $\alpha_\ell = 0$ for all $\ell \neq \tilde{\ell}$. Based on this observation, we propose to measure the sensitivity of an MEG inversion algorithm to activity in a given BR by how many components of α corresponding to other BRs are different from zero and how large they are.

In our study of the IAS MEG inverse solver and of the other three algorithms considered, we assess the performance using the Monte Carlo simulations of K independent activity patches as described in Section 2.3.1. More specifically, for each BR, to evaluate the MEG inverse solver, we

i) apply the MEG inverse solver with the synthetic data sample $\mathcal{B} = \{b^1, \ldots, b^K\}$, to estimate the corresponding activity vectors, $\{a^1, \ldots, a^K\}$, $a^k \in \mathbb{R}^n$,

ii) compute the corresponding BR-ALI vectors $\{\alpha^1, \ldots, \alpha^K\}$, $\alpha^k \in \mathbb{R}^L$,

iii) compute the sample mean and the sample covariance of the K BR-ALI vectors,

$$
\overline{\alpha} = \frac{1}{K} \sum_{k=1}^K \alpha^k, \quad \Gamma_\alpha = \frac{1}{K-1} \sum_{k=1}^K (\alpha^k - \overline{\alpha})(\alpha^k - \overline{\alpha})^\mathsf{T}.
$$

The display of the mean BR-ALI vector $\overline{\alpha}$ in the form of a histogram provides a visual clue of the method's performance; the more tightly the histogram is concentrated on the active patch, the more reliable the estimate is. The diagonal entries of Γ_α, which are the variances of the estimated

activity levels in various BRs and measure the consistency of the mean BR-ALI vector retrieval, are indicated in the histograms displayed in Figures 3–5 as standard deviation whiskers. We point out a formal similarity of the analysis in pattern recognition: The computed sample of BR-ALI vectors can be seen as a library of patterns, and the histograms represent frequencies at which different individual patterns appear.

The analysis has a close similarity with the resolution matrix analysis (Molins *et al.*, 2008, Hedrich *et al.*, 2017), the vector $\overline{\alpha}$ being the average response to activity in a single brain region, and can therefore be thought as a column of the resolution matrix; however, here we consider noisy signals, random sources and nonlinear maps.

The proposed methodology generalizes naturally to several active regions of interest. In the case where there are N simultaneously active BRs, the statistics of the estimated BR-ALI vectors can be investigated as in the case of a single patch.

2.3.4 Bayes factors, sensitivity and specificity

To set up a protocol to assess an MEG inverse solver, consider the following simple elementary test that constitutes one of the basic building blocks of Bayes factor analysis, see e.g. Kass RE and Raftery AE (1995). In turn, there are aspects of Bayes factor analysis that share similarities with concepts of sensitivity and specificity used in the context of binary classifiers.

Given activity in a patch $\mathcal{P}$ with barycenter $m_{\mathcal{P}}$ defined by (5), consider two spheres with the same fixed radius R, $B_0 = B(m_{\mathcal{P}}, R)$ centered at the barycenter of the activity, and $B_1 = B(v_j, R)$, centered at a randomly selected vertex v_j in the brain, respectively. Let $a \in \mathbb{R}^n$ be the activity vector estimated from the data b generated by the activity in the patch $\mathcal{P}$, and formulate the following pair of hypotheses:

$$(H_0) \qquad \text{The activity is in the sphere } B_0,$$
$$(H_1) \qquad \text{The activity is in the sphere } B_1.$$

We consider the estimated activity vector a as data, and associate to the hypotheses H_0 and H_1 the probabilities

$$P(a \mid H_0) = \frac{1}{|a|} \sum_{v_j \in B_0} a_j \quad \text{and} \quad P(a \mid H_1) = \frac{1}{|a|} \sum_{v_j \in B_1} a_j,$$

where $|a| = \sum_{j=1}^{n} a_j$ is a normalizing factor. We quantify the strength of the hypothesis (H_0) against the hypothesis (H_1) in terms of the Bayes factor,

$$\text{BF}(H_0, H_1) = \frac{P(a \mid H_0)}{P(a \mid H_1)},$$

in the following manner. If $\text{BF}(H_0, H_1) < 1$ we argue that the data support the hypothesis H_1, while if $\text{BF}(H_0, H_1) > 1$, the evidence is in favor of the hypothesis H_0, and the support for H_0 is stronger the larger the value of the ratio. Following Kass RE and Raftery AE (1995), we classify the relative strength of the two hypotheses according to the following scale:

$$\begin{aligned}
\text{BF}(H_0, H_1) &\in (0, 1) &&\text{The evidence is against } (H_0), \\
\text{BF}(H_0, H_1) &\in [1, 3) &&\text{The evidence is weakly in favor of } (H_0), \\
\text{BF}(H_0, H_1) &\in [3, 10) &&\text{The evidence is strongly in favor of } (H_0), \\
\text{BF}(H_0, H_1) &\geq 10 &&\text{The evidence is overwhelmingly in favor of } (H_0).
\end{aligned} \qquad (6)$$

Clearly, the hypothesis (H_1) depends on the choice of the point v_j defining the sphere B. Because we are, in fact, interested in assessing whether the MEG inversion method is favoring the

correct activation area over *any* other area, we enrich the Bayes factor test by generating not one, but a family of M competing hypotheses,

$$(H_m) \quad \text{The activity is in the sphere } B_m = B(v_{j_m}, R), \quad 1 \le m \le M,$$

where the center points v_{j_m} are drawn from the set of vertices with uniform probability. This leads to a set of M Bayes factors, $\mathrm{BF}(H_0, H_m)$, $1 \le m \le M$, and we can compute the frequencies of occurrences in the scale defined above. A good solver should have most of the Bayes factors in the intervals $[3, 10)$ or above 10. Numerous occurrences of Bayes factors below 1 indicate that the solver tends to attribute the activity to an incorrect location.

It may be argued that the Bayes factor analysis as described above implicitly uses the information that the MEG data was generated by a single active patch, while in reality the number of active brain regions is unknown. To address this issue, we test the algorithm also with several simultaneous patch activities. More specifically, in the case where the activity occurs simultaneously in N different BRs, we select one of these active regions at a time and set the activity spheres B_0 centered at the barycenter of the selected patch, while the center of the competing activity spheres B_m are drawn randomly from any other BR. Subsequently we compute the Bayes factors supporting the hypothesis

$$(H_0): \quad \text{The activity is in the sphere } B_0 = B(m_{\mathcal{P}}, R).$$

versus the family of competing hypotheses,

$$(H_m) \quad \text{The activity is in the sphere } B_m = B(v_{j_m}, R), \quad 1 \le m \le M,$$

where the centers v_{j_m} are again drawn from the uniform distribution. Observe that if B_m overlaps with one of the active BRs other than the one defining the hypothesis (H_0), the Bayes factor may be low; thus, the other active regions can be seen as confounding factors in this case. However, this simulation is closer to the situation with the real data in which the number of active regions of interest is unknown.

While it is not one of the aims of this paper to propose a classifier based on the Bayes factors, to illustrate how elements of our analysis can be remindful of the classical concepts of sensitivity and specificity, consider a binary query to find whether a given algorithm correctly identifies an active area. Suppose that with normalized activity vector a, a sharp threshold β is introduced to distinguish whether an area is deemed to be active according to the rule

$$p(a \mid H_0) \ge \beta \text{ means that } B_0 \text{ is active,}$$
$$p(a \mid H_0) < \beta \text{ means that } B_0 \text{ is not active.}$$

Given a set of N simulated data with activity in B_0, an algorithm with high sensitivity correctly identifies the domain B_0 as active, and therefore the number of reconstructions for which $p(a \mid H_0) \ge \beta$ should be close to N. On the other hand, an algorithm with high specificity queried whether B_m is active when in reality the activity is in B_0, should frequently report a negative event, and therefore the number of cases in which $p(a \mid H_m) < \beta$ should again be close to N. In summary, for a method with high sensitivity *and* high specificity, the Bayes factor should frequently satisfy $BF(H_0, H_m) > \beta/\beta = 1$. It is in this spirit that histogram of Bayes factors heavily slanted towards high values can be seen as an indicator of high sensitivity and high specificity. Observe, however, that there is no established one-to-one correspondence between these indicators, although probabilistic bounds could be developed.

2.4   Model settings

2.4.1   Source space, brain regions, and simulated data

We generate the synthetic data using the anatomical and geometric information provided by the MEG-SIM web portal Aine *et al.* (2012). The MRI data provided in the portal were segmented with Freesurfer (Dale *et al.*, 1999) and imported in Brainstorm (Tadel *et al.*, 2011) to generate a source space including both the cortical surface and the substructure regions. The source space obtained in this manner consists of 27 000 nodes identifying possible dipole locations in the gray matter. A randomly generated patch of activity in the left lateral occipital cortex is shown in the top row of Figure 2 and a randomly generated patch of activity in the right cerebellum is shown in the middle row of Figure 2. To avoid the overly optimistic results that typically result from testing inversion algorithms using the same model for data generation and inverse solver, we solved the inverse problem with a sparser mesh with $n = 20\,000$ nodes obtained through independent sampling of the cortical and subcortical structure.

The labeling of the brain regions in the cortical surface follows the Desikan-Killiany Cortical Atlas (Desikan *et al.*, 2006) consisting of 34 cortical anatomical regions in each hemisphere, to which we added 8 bilateral structures (accumbens, amygdala, caudate, hippocampus, pallidum, putamen, thalamus, cerebellum) and 1 central structure (brainstem). Thus the atlas we used in the simulations consists of a total of 85 brain regions.

The data acquisition geometry corresponds to the 306-channel Elekta Neuromag device and the lead field matrix is computed using the single layer model implemented in the OpenMEEG (Kybic *el al.*, 2005, Gramfort *el al.*, 2010) software provided in the Brainstorm package.

FIGURE 2 HERE

2.4.2   Real and realistic data

The real data set used for our computed examples is the MEG sample data acquired with the Neuromag Vectorview system at MGH/HMS/MIT Athinoula A. Martinos Center Biomedical Imaging and made available, together with the MRI reconstructions created with FreeSurfer, in the MNE software package (Gramfort *el al.*, 2014). As part of the protocol for the data collection, checkerboard patterns were presented into the left and right visual field, interspersed by tones to the left or right ear. The interval between the stimuli was 750 ms. Occasionally a smiley face was presented at the center of the visual field. The subject was asked to press a key with the right index finger as soon as possible after the appearance of the face (Gramfort *el al.*, 2014). In our computed experiments we only consider the trials corresponding to the left-ear auditory stimulus and perform the averaging on these trials. To test how the inverse methods accurately recover activity in the deep regions, we added to the real data set described above the magnetic field evoked by a patch of 20 active dipoles in the brainstem, shown in the bottom row of Figure 2 at time $T = 5$ms, whose time series follow a Gaussian distribution peaked at $T = 5$ms with standard deviation $\sigma = 2$; see Parkkonen *et al.* (2009) for details. The source space containing both cortical regions and subcortical structures after discretization comprises 22019 vertices.

The peak amplitude of the simulated activity is about one fourth of the peak amplitude of the auditory response of the real data, the estimated SNR being about 5.3 dB. For comparison, the SNR of the auditory response in the real data is approximately 15 dB.

2.5   Setting of IAS parameters

As shown in Sections 2.1-2.2, in the IAS algorithm the user has to assign the values of the three parameters β, $\theta_{\max}$, $\overline{k}$ and of δ in the local anatomical prior matrix C_j. The choice of all the four parameters values can be driven by their physiological interpretation.

The hyperparameter $\beta > \frac{5}{2}$ control the focality of the reconstructed sources: when β approaches the limit value $\frac{5}{2}$, the reconstructed brain activity becomes more and more focal. Let us define $\eta = \beta - \frac{5}{2}$. Several numerical simulations suggest that a reasonable values of η should range between 0.1 and 0.001 (cf. Calvetti *et al.* (2015)), depending on whether we aspect less or more focal sources: in the following tests, where we assume focal sources, we set $\eta = 0.001$.

The parameter $\theta_{\max}$ is used to avoid unrealistic high values of the dipole variances. Reliable estimates for the maximum dipole variance have been reported in the literature, see, e.g., (Mosher *el al.*, 1993); alternatively, a heuristic criterion is to truncate the values of the variance retaining just those values below the threshold $\theta_{\max} = 0.9 \max_j(\theta_j)$. The goodness of this choice has been confirmed with extensive simulations (cf. Calvetti *et al.* (2015)).

The parameter $\overline{k}$ is related to the number of sources we expect to be active. A direct computation shows that $\xi = \sum_{j=1}^{n} p_k/k$ assumes values between 0.5 and 0.2 when $\overline{k}$ varies from 1 to 10, therefore a value of ξ in this interval is a good choice if we do not have any prior information on the number of active sources. If we assume that only one source is active, we can set $\xi = 1$ (see the Appendix). Although when testing with synthetic data we know exactly the number of active sources, we set $\xi = 1$ in all the tests we performed to show that this choice is a good initial guess for both single and multiple source distributions.

The last parameter to be set in IAS is δ that accounts for possible inaccuracy in the segmentation of the cortical/subcortical surface normals. If the directions of the normals would be exact, δ could be set equal to zero; in the realistic case the anatomical data extracted from the MRI are not fully accurate so that the normal should be considered a preferred direction. To allow the dipoles to follow quasi-normal orientation, we choose a small value of δ. Several numerical tests show that a good choice is $\delta = 0.05$ (cf. (Calvetti *et al.*, 2015)).

The tests in the following section show the effectiveness of our settings.

3   Results

3.1   Synthetic data tests

In this section, we systematically test the sensitivity and specificity, in the sense explained earlier, of the IAS MEG inverse solver algorithm using the validation tools described in the previous section with synthetic data sets. For the sake of comparison, we run the same tests with other three standard MEG solvers, the weighted Minimum Norm Estimate (wMNE) (Lin *el al.*, 2006a, Gramfort *el al.*, 2014), the dynamic Statistical Parametric Mapping (dSPM) (Dale *et al.*, 2000) and sLORETA (Pasqual-Marqui, 1999).

3.2   Performance via BR-ALI maps: synthetic data, single activity patch

The first suite of computed experiments is designed to assess whether and how the location of the BR where the activity occurs affects the quality of the IAS reconstruction. In order to provide a measure that is robust over patches with different anatomical characteristics, for each one of the 85 BRs specified by the selected atlas, we generate a sample of $K = 100$ patch activities by the random process described in Section 2.3.1, calculate the magnetic field measured by the sensors, and add to the latter synthetic brain noise generated as described in Section 2.3.2. For each sample problem, we reconstruct the activity a with the IAS algorithm described in Section 2.1, as well as

with the wMNE, DSPM and sLORETA algorithms as implemented in Brainstorm. For the last three algorithms we used default parameters, i.e. $SNR_{db} = 3$ and loose orientation. The results of the reconstructions are summarized in the average BR-ALI vectors and visualized in the form of histograms with whiskers indicating the marginal standard deviation around the mean. An example of activity reconstructions provided by the different methods for the patch activities in the lateral occipital region and in the cerebellum shown in Figure 2 can be found in Figures 4 and 6, respectively.

Figure 3 shows the average BR-ALI vectors when the active patch is in the left lateral occipital cortex and the MEG inverse problem is solved with the IAS algorithm (top left), wMNE (top right), dSPM (bottom left) and sLORETA (bottom right), respectively. The results for the BRs in the left hemisphere are shown in red, those for the BRs in the right hemisphere in black. Figure S1 displays the histograms relative to the case where the active patch is in the frontal pole cortical region in the right hemisphere.

<div align="center">
FIGURE 3 HERE
FIGURE 4 HERE
FIGURE 5 HERE
FIGURE 6 HERE
</div>

The deeper into the brain the active patch is, the more difficult we expect the mapping from the MEG data to the current dipoles to be. Figure 5 displays the histograms relative to 100 simulations with the active patch located in the right portion of the cerebellum. Compared to the previous tests, the real challenge here arises from the distance of the activity region from the receivers.

The amygdala, a subcortical structure which is part of the limbic system, is believed to be involved in attentional and emotional processes and in the formation of memories. Its remote location and small dimensions add to the challenge of localizing activity confined into this region from MEG data. Figure S2 displays the histograms for the four inverse methods relative to the case where the active patch is confined to the left amygdala.

The four reported results are a representative subset of the performance of the four different inverse solvers in cortical and subcortical brain regions.

3.3   Bayes factor analysis: synthetic data, single activity patch

To assess the IAS algorithm and to compare it to that of the other three MEG inverse solvers considered in this study, we begin by performing a Bayes factor analysis with MEG data generated according to the procedure described in Section 2.3.1 with a single activity patch. After generating an ensemble of $K = 20$ randomly generated patches of activity restricted to a given BR, we compute the corresponding low noise synthetic data set, estimate the activity pattern with each inversion algorithm and compute the corresponding Bayes factors, thus testing the evidence supporting the correct identification of the active area in the reconstruction. For each of the 20 active patches in the sample, we draw 100 random spheres as competing alternative to the hypothesis that the activity is in the BR where it actually occurs, for a total of 2 000 Bayes factors per BR.

The summary of the Bayes factor analysis for each active BR can be represented graphically in the form of a histogram with four bins, recording the number of occurrence of the four levels of evidence (6) supporting the correct identification of the activity. The more occurrences there are in the two top categories, the more reliable the algorithm is at correctly identifying the area of activity. For a easier visual assessment, we color coded the bars indicating the numbers of times a Bayes factor falls into one of the four categories, using green for Bayes factors greater than 10 (overwhelming support of the hypothesis that the active patch is indeed in that BR), blue for Bayes factors between 3 and 10 (strong support), red for Bayes factors between 1 and 3 (weak support) and black for Bayes factor smaller than 1, in which case the support is for the hypothesis that the active patch is not in that BR. A prevalence of green and blue indicates a strong support

of the hypothesis that the activity has been detected in the correct region, while a dominance of black and red suggests poor identification of active areas. Figure 7, 8 show the results for all cortical and subcortical BRs included in the atlas when performing the inverse mapping with the four inverse solvers.

<div align="center">FIGURE 7 HERE</div>
<div align="center">FIGURE 8 HERE</div>

3.4   Bayes factor analysis: Multiple activity patches

We extend the Bayes factor analysis to the case where several patches are active simultaneously in different BRs by performing a suite of four different tests. In the first two tests, we generate two randomly determined activity patches: In the first one, the activity patches are in the left cerebellum and the left prefrontal cortex, and in the second one, in the left amygdala and the left prefrontal cortex. In both cases one of the patches is cortical and the other one deep in the brain. In the third set of tests, the simulated data arises from three activated patches, located in the left precuneous, right precuneous and left inferior parietal cortex. In the last protocol we generate activity patches in the six different regions comprising the default mode network (DMN), namely in the left precuneus, right precuneus, left inferior parietal cortex, right inferior parietal cortex, left caudal anterior cingulate and right caudal anterior cingulate regions.

As in the case of a single active BR, we generate $K = 20$ independent activity patterns in the selected BRs, compute the corresponding low noise data, solve the inverse problem with the four different algorithms and test the support of the hypothesis H_0 for each of the activated regions individually with $M = 100$ independently drawn competing spheres B_m. Observe that when we test the hypothesis H_0 for a single selected brain region while the data are generated by several active sources, the magnetic field due to the additional sources represent high amplitude brain noise masking the signal from the source of interest.

The top row of Figure 9 shows the histograms of the Bayes factors, binned and color coded according to the strength of evidence (6), corresponding to the four different inverse solvers in the case when the active patches are in the left amygdala and in the left frontal pole. The second row of Figure 9 shows the histograms of the Bayes factors for the four different inverse solvers in the case when the active patches are in the left cerebellum and in the left precentral cortex.

The histograms of the Bayes factors for the four inverse solvers when there are three active patches located in the left hippocampus, precentral gyrus and right thalamus are displayed in the third row of Figure 9. Finally, the histograms for the Bayes factors corresponding to the different solvers in the case where there are active patches in the six different BRs in the DMN are displayed in the two bottom rows of Figure 9.

<div align="center">FIGURE 9 HERE</div>

3.5   Reconstruction of the brain activity from real data

Finally, we apply the different inverse solvers to the real data, augmented with realistic activity in the brainstem as explained in Section 2.4.2. Figures 10 and 11 show respectively the average BR-ALI vectors at times $T = 5$ ms and $T = 100$ ms for the four inverse solvers. In Figures 12 and 15 we show ten axial slices of the two reconstructions of the brain activity at times $T = 5$ ms and $T = 100$ ms obtained using the IAS (left) and the wMNE (right) inverse solvers. In Figure 13 and 16 we show the corresponding reconstructions obtained by dSPM (left) and sLORETA (right). In Figures 14, 17 we show the 3D reconstructions obtained at $T = 5$ ms and $T = 100$ ms by the four inverse solvers.

<div align="center">FIGURE 10 HERE</div>
<div align="center">FIGURE 11 HERE</div>

FIGURE 12 HERE
FIGURE 13 HERE
FIGURE 14 HERE
FIGURE 16 HERE
FIGURE 16 HERE
FIGURE 17 HERE

4 Discussion

4.1 BR-ALI maps

The sparsity of the BR-ALI map is a good indicator for the sensitivity and specificity of an inverse solver: The more concentrated the histograms are on the active areas, the more likely it is that the activity is correctly identified without too much of confounding. In general, we expect a higher sensitivity and specificity for BRs closer to the sensors than in BRs deep in the brain, as confirmed by the computed experiments.

4.1.1 Single patch in cortical BRs

It emerges clearly from the panels in the top row of Figure 3 that the IAS and wMNE inverse solvers have high sensitivity to activation restricted to the left lateral occipital cortex, indicated by the outstanding histogram bar, and that the confounding is confined mostly to anatomically proximal regions, most notably the nearby pericalcarine fissure in the same hemisphere. Neither algorithm suggest any significant activity in the right hemisphere, or in the deep brain structures. In the BR-ALI vector computed by the dSPM or sLORETA algorithms, on the other hand, the left lateral occipital BR is not as clearly identifiable, and the likely locus of activity is attributed to several nearby BRs in the same hemisphere, e.g., the nearby pericalcarine fissure, lingual, and cuneous BRs. Some of the brain activity is suggested also in the right hemisphere and in the deep brain structures. Overall, the performance of the IAS algorithm appears to be significantly more similar to that of wMNE than to dSPM or sLORETA, the latter ones producing more confounding activity.

The suite of simulations with the activity confined to the right frontal pole region confirms the conclusions of the relative sensitivity of the four different approaches. The corresponding BR-ALI vectors, visualized in the form of histograms in Figure S1, show that both the IAS and wMNE algorithms reconstruct a substantial portion of the dipole activity in the right frontal pole, while finding some activity also in the nearby left frontal pole and, to a lesser extend, to the medial orbitofrontal cortex, also anatomically close. In neither case, any significant leakage of the reconstructed activity to BRs in the deep brain occurs. The localization of the active BR is much weaker for the dSPM and sLORETA methods, whose BR-ALI vectors indicate a smearing of the reconstructed activity over several BRs, including subcortical ones. In this test sLORETA appeared to recognize the BR where the signal generating the data came from better than dSPM which attributes the highest average activity to the rostral anterior cingulate cortex.

The first two tests are representative for the four methods when the data correspond to activity in a restricted patch located in a cortical BR, with IAS and wMNE performing quite satisfactorily when it comes to identifying the active BR, while dSPM and sLORETA have the tendency to favor activity in deeper regions of the brain.

4.1.2 Single patch in subcortical BRs

The sensitivities of the four solvers when the active patch is confined to the right cerebellum are summarized in Figure 5. As in the case of active patches in cortical regions, the attribution of the brain activity to the correct BR is less confounded when the inverse problem is solved using the IAS or wMNE algorithms. The BR-ALI vectors corresponding to these two algorithms point more clearly to the right cerebellum, with a little activity smeared to the left cerebellum and brainstem. The BR-ALI vectors for the dSPM and sLORETA inversion methods, on the other hand, tend to distribute the activity in the deeper regions of the brain, without the cerebellum standing out clearly among them. Moreover, a substantial fraction of the activity is mapped to cortical regions, e.g., lingual, parahippocampal and pericalcarine gyrus, suggesting potential problems when it comes to specificity.

The same patterns are observed when the active patch is in the left amygdala. As shown in Figure S2, the IAS inverse solver is quite effective at identifying the presence of activity in the left amygdala, with some leakage to the anatomically adjacent left temporal pole and left entorhynal cortex. The wMNE BR-ALI histogram does not show a localization to the left amygdala of the activity; instead, the activity is distributed over the left hemisphere, with a slight preference for the BRs closer to the left amygdala. The BR-ALI vectors relative to dSPM and sLORETA, on the other hand, map the reconstructed activity without much distinction onto all BRs corresponding to internal structures, and a few cortical BRs in the proximity of the left amygdala, suggesting a bias towards regions of the brain away from the surface.

4.2 Bayes factor analysis

The Bayes factor analysis is based on the counts of how often, and at which level, out of the 2 000 tests per brain region, the reconstruction of the brain activity by a given algorithm supports the hypothesis that the activity is indeed in the correct region. Due to our choice of color coding, we can be confident that the activity within a BR when the corresponding histogram has taller green or blue bars can be identified by the inverse solver rather unequivocally, while taller black or red bars are an indication that activity in that BR is likely to be erroneously attributed to activity in another part of the brain.

4.2.1 Single active patch

The overwhelming predominance of green in the left panels of Figure 7, displaying the histograms of the Bayes factors for the cortical BRs, is a strong indicator that the solution of the MEG inverse problem produced by the IAS algorithm is much more likely to find active current dipoles in the correct BR than in other regions. A slightly lesser dominance of the green and blue bars for the insula and isthmus cingulate BR suggests that activity concentrated in either of these two regions may have a slightly larger tendency to be incorrectly attributed elsewhere. In most of the Bayes factors histograms for the subcortical BRs the green and blue bars dominate over the red and black ones, with the exception of the thalamus and, to a lesser extent, the caudate gyrus, where the red and black take over, suggesting that the IAS has difficulties to identify unequivocally activity localized in this BR.

In the histograms of the classification of the Bayes factors for the wMNE algorithm, shown in the right panels of Figure 7, the green and blue bars tend to be more prominent than the red and black ones in the top row, although not as markedly as for IAS. This is an indication that when the measured signal comes from a patch in a cortical BR, the solution computed by wMNE is likely to be concentrated in the correct cortical region, with the exceptions for the caudal anterior cingulate, insula, isthmus cingulate, parahippocampal, posterior cingulate and rostral anterior cingulate regions.

The situation is quite different when it comes to correctly attributing activity generated in a subcortical region, as shown in the two panels in the bottom row of the same figure. Here the taller bars are red or black, with the exception of the cerebellum and, to some extent, the brainstem, pointing to a low specificity of wMNE in subcortical regions.

The Bayes factor analysis for dSPM is summarized by the histograms in the left panels of Figure 8. In the histograms pertaining the cortical BRs the blue and green bars dominate the red and black ones, although the green presence is not as massive as for the IAS or wMNE inverse solvers, indicating that the dSPM inversion is still capable of identifying active cortical BR, but with lower specificity than IAS and wMNE. In many of the cortical BRs the height of the blue bars varies between 800 and 1000, but there is a strong presence of red bars, with height between 200 and 700, pointing to the fact that in some cases, dSPM has difficulties with recovering the activity in the correct region and attributing it instead, to a different region. The dominance of blue and red in the histograms for the subcortical BRs, together with a green presence, suggests that the localization of the activity in a patch deeper in the brain by dSPM is more accurate than with wMNE, but not as precise as with IAS.

The conclusions of our study with Bayes factors for sLORETA, whose histograms are displayed in the right panels of Figure 8 are very similar to those for dSPM, with a better performance than wMNE, but not as good as IAS, in subcortical BRs and not as specific when it comes to recovering activity in cortical BRs.

4.2.2 Multiple activity patches

The histograms of the Bayes factor for the IAS and the three other reference methods, in case where there are two patches of activity, one in the left amygdala and the other in the frontal pole, displayed in Figure 9, show that the reconstructions computed with the IAS suggest very strongly the presence of activity in the left amygdala, as shown by the tall green bar in the corresponding histogram as well as that in the cortical frontal pole. In this experiment, sLORETA outperforms both wMNE and dSPM, the latter showing the least precision in mapping the activity in the frontal pole area.

The second test with two patches of activity, in the left cerebellum and in the left precentral region, respectively, confirm that the IAS is the most precise of the four inverse solvers when it comes to finding the location of both active patches, as shown in the histograms in the second row of Figure 9. The performance of wMNE in this case is very good for both the cortical and subcortical BRs, while in this case the specificity of sLORETA is not as good as that of dSPM.

The Bayes factors for the case where there are three active patches, located in right precentral gyrus, left hippocampus and right thalamus, are summarized in the third row of Figure 9. The IAS algorithm identifies with least confounding the activity in the precentral region and hippocampus, but is not as accurate when it comes to the thalamic patch. Both interior patches turn out to be very challenging for wMNE, which has no problems with accurately recovering the patch in the precental region. In this case, dSPM is best at localizing the activity in the thalamus and in the hippocampus.

In the last protocol with Bayes factors, the MEG signal came from six active patches located in the six BRs involved in the DMN. The histograms for the six different BRs with the four inverse solvers, displayed in the two bottom rows of Figure 9, indicate that the localization of the activity in the caudal anterior cingulate region in either hemisphere is challenging for all methods, and particularly so for wMNE whose histograms show the black bars to be tallest. The localization of the activity in the inferior parietal BRs is resolved very well by IAS and wMNE, and satisfactorily by dSPM and sLORETA.

4.3   Real data: IAS favors sparse solution

Figures 10 and 11 show the BR-ALI vectors for the four inverse methods at times T=5 ms and T=100 ms, respectively. At T=5 ms all inverse methods find activity also but not only in the brainstem region: the estimated activities using sLORETA and dSPM are mutually very similar, suggesting activation also in the amygdala and in the parahippocampal cortex; the deep activity is dominating over cortical activity in the reconstructions. The reconstructed activity by wMNE, compared to IAS, is more smooth, while IAS finds activity also in cerebellum. At T=100 ms, IAS suggests activity in the temporal region, mainly in the upper bank of the superior temporal cortex and in the transverse temporal region, that are part of the primary auditory cortex. The IAS MEG solution shows activity in the superior temporal regions of both hemispheres, while when using other methods such bilaterality is not evident. In general, the dSPM and sLORETA algorithms seem to be in favor of activity deeper in brain, and the spatial smoothing characteristics of the algorithms is clearly visible.

4.4   Methodological issues and future work

As shown in Section 2.5, the choice of the parameters in the IAS algorithm can be driven by their physiological meaning. Several tests we performed show that the reconstructed activity map is not very sensitive to the values of the parameters δ, $\theta_{\max}$ and ξ. The values we used in the tests are reasonable for most applications in neuroscience. Thus, the parameter that the user has really to select is $\eta = \beta - \frac{5}{2}$. Since η controls the sparsity of the reconstruction, its choice is related to the a priori knowledge we have on the protocol of the neuroscience experiment under study. Values of η in the order of 0.1 favor more spread reconstruction and for these values the IAS algorithm produces activity maps very similar to the maps obtained by the wMNE algorithm. On the other hand, values in the order of 0.001 favor focal reconstructions, which is the more interesting case in neuroscience studies, especially when deep brain sources are involved.

Further studies of the performance of the IAS algorithm on the localization of deep brain activity will be done in the future, including the use of a more accurate model for describing deep brain sources, e.g., the model proposed in Attal *et al.* (2012), Attal and Schwartz (2013) or the mixed source space model available in the MNE software[1]. In this case, different values of the parameter η for cortical and subcortical regions can be used in order to increase the sensitivity of the algorithm to deep sources. Finally, to take into account the imprecision due the use of an averaged physiological atlas that cannot reproduce exactly the individual anatomy, one can use some mathematical methods related to fuzzy logic (see, e.g., (Algorri and Flores-Mangas, 2004, Ciofolo and Barillot, 2009)).

The IAS algorithm described in this paper is designed for a single-time slice analysis: the algorithm is re-initialized each time and does not retain any previous information. Actually, one of the main advantage of the MEG devices is in that they can measure the neuromagnetic field with a high temporal resolution. In order to deal with MEG time series some preliminary results show that when the IAS algorithm starts from a θ^* that is related to the values of θ_j at the previous time step, the rate of convergence of the algorithm is increased. Then, for a further analysis we can model θ as a Markov process, so the estimation of θ at time t depends just on the value of θ at time $t - \delta t$.

---

[1]http://martinos.org/mne

5   Conclusions

Finding a robust metric for assessing the performance of an inverse solver in MEG is not a simple task. Algorithms that are based on the goal of localizing single dipoles may be judged according to the precision of the localization, but such metric may not be a judicious one for methods that estimate distributed sources. Vice versa, single dipole methods may have a limited success when distributed activities are to be estimated. In this paper we propose a metric for the algorithm assessment based on the reliability of an algorithm to identify active brain regions, superficial and deep, regardless of whether a single or several active regions occur. As pointed out, the BR-ALI analysis has a formal similarity with the resolution matrix analysis which is widely used as a basis of performance analysis, and allows the computation of simple measures for performance: Resolution index (RI), dipole resolution error (DLE), and spatial dispersion (SD), see Molins *et al.* (2008) and Attal and Schwartz (2013), Hedrich *et al.* (2017). The precision of finding a single dipole is not of concern here, although the focality of the reconstructions helps discerning between brain regions that are anatomically close. One of the messages of the analysis is that an algorithm such as IAS favoring focal solutions produces less confounding reconstructed activity in the anatomically close regions, reducing the ambiguity in the interpretation of the reconstruction. To avoid the pitfall of anecdotal successes, the methodology is based on extensive independent Monte Carlo sampling, and the results are processed into a form of first and second order statistics, and Bayes factor analysis. The long-term goal of this work is to build reliable uncertainty quantification tools to analyze extensive data sets with non-event based brain data, e.g., various resting states or states of consciousness, brain connectivity, or fingerprinting of diseased or abnormal states. Assessment of success rates of algorithms as presented here, as opposed to precision case studies with activity localization, is in line with the current algorithm testing paradigm in data mining and big data analysis. The proposed methodology was tested with four inverse solvers, one of which is the recently developed IAS algorithm, and three standard methods. Our conclusion is that the IAS method seems to perform relatively consistently in the tasks that it is originally designed for, that is to identify both active cortical and deep brain regions without a significant confounding beyond the inevitable leakage of the estimated activity to anatomically close regions, which is due to the inherent ill-posedness of the problem. The comparison of the algorithm was done against three standard MEG algorithms without detailed tuning of the default settings. A more comprehensive meta-analysis should include the optimization of the model parameters as well as the inclusion of other algorithms favoring focal solutions, but is beyond the scope of this work.

Appendix: Interpretation of hyperparameters

In Calvetti *et al.* (2015), the interpretation of the hyperparameters $\beta \in \mathbb{R}$ and $\theta^* \in \mathbb{R}^N$ was discussed: It was shown that β allows the user to control the sparsity of the IAS solution, while the empirical Bayesian approach provided a way to relate θ^* to the sensitivity scaling. We summarize here the analysis on hyperparameters, developing the discussion of θ^* further, so that the parameter tuning can be done easily and semi-automatically.

Parameter β and control of sparsity

In Calvetti *et al.* (2015), it was proved (Theorem 2.1) that the sequential minimization that constitutes the core of the IAS algorithm can be interpreted as a fixed point iteration to find a minimizer $\hat{Q} = [\tilde{\hat{q}}_1, \tilde{\hat{q}}_2, \dots, \tilde{\hat{q}}_n]^\mathsf{T} \in \mathbb{R}^{3n}$ of the energy functional (3),

$$\hat{Q} = \mathrm{argmin}\{\mathcal{E}(Q, S(Q))\}, \quad \hat{\Theta} = S(\hat{Q}),$$

where $\hat{\Theta} = [\hat{\theta}_1; \hat{\theta}_2; \ldots, \hat{\theta}_n] \in \mathbb{R}^n$, and $S : \mathbb{R}^{3n} \to \mathbb{R}^n$ is defined componentwise as

$$\theta_j = S_j(\vec{q}_j) = \theta_j^* \left(\frac{\eta_j}{2} + \sqrt{\frac{\eta_j^2}{4} + \frac{\|\vec{q}_j\|_{\mathsf{C}_j}^2}{2\theta_j^*}} \right), \quad \eta_j = \beta_j - \frac{5}{2}, \quad 1 \le j \le n.$$

Furthermore, it was shown that if we write $\beta_j = 5/2 + \eta$, $1 \le j \le n$, then, as $\eta \to 0^+$, we have the asymptotic expression

$$\mathcal{E}(Q, S(Q)) = \frac{1}{2} \|b - \sum_{j=1}^{n} \mathsf{M}_j \vec{q}_j\|_{\Sigma}^2 + \sqrt{2} \sum_{j=1}^{n} \frac{\|\vec{q}_j\|_{\mathsf{C}_j}}{\sqrt{\theta_j^*}} + \mathcal{O}(\eta). \tag{7}$$

In particular, the penalty term in the above expression is a weighted ℓ^1-norm for the dipole amplitudes that are measured in the metric defined by the anatomical prior matrices C_j. This argument demonstrates that at the limit, the IAS algorithm provides an effective algorithm for finding a weighted Minimum Current Estimate (MCE), with the modification given by the anatomical prior (Uutela *et al.*, 1999). In conclusion, we see that the role of the hyperparameter β is to control the sparsity of the IAS estimate. In Calvetti *et al.* (2015), this effect was demonstrated using numerical simulations.

Parameter θ^* and sensitivity

The asymptotic expression (7) is indicative also from the point of view of the interpretation of θ_j^*. It is well-known that MEG algorithms based on penalized minimization of the fidelity to data tend to favor superficial sources, and to compensate this effect, a sensitivity weight is often introduced, see, e.g. Lin *el al.* (2006a). From the Bayesian point of view, sensitivity weighting is a problematic practice, since traditionally the prior should be independent of the observation model, a condition that the sensitivity weight does not satisfy. However, it is possible to find a satisfactory Bayesian interpretation for θ_j^* so that it effectively works as a sensitivity weight. The connection between sensitivity and hypermodels is built through the analysis of the signal-to-noise ratio as follows. Consider the linear forward model

$$b = \mathsf{M}Q + \varepsilon = \sum_{j=1}^{n} \mathsf{M}_j \vec{q}_j + \varepsilon = b_0 + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Sigma).$$

To estimate the expected power of the noiseless signal appearing in the signal-to-noise ratio defined in (4), observe that from the prior model, conditional on $\Theta \in \mathbb{R}^n$, we have

$$\mathsf{E}\{\|b_0\|^2 \mid \Theta\} = \sum_{j=1}^{n} \theta_j \mathrm{trace}\big(\mathsf{M}_j \mathsf{C}_j \mathsf{M}_j^{\mathsf{T}}\big) = \sum_{j=1}^{n} \theta_j \|\mathsf{M}_j \mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2,$$

where the subscript refers to the Frobenius norm of the matrix. Furthermore, by using the gamma hyperprior model $\theta_j \sim \Gamma(\beta, \theta_j^*)$ for the vector Θ, we arrive at

$$\mathsf{E}\{\|b_0\|^2\} = \sum_{j=1}^{n} \mathsf{E}\{\theta_j\} \|\mathsf{M}_j \mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2 = \sum_{j=1}^{n} \beta \theta_j^* \|\mathsf{M}_j \mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2.$$

The choice of the hyperparameters θ_j^* must therefore be compatible of what we a priori assume about the distribution of the activity and the resulting SNR. To begin with, assume that we have

a reason to believe that only one source is active, but we do not know which one. Then, by the definition (4) of SNR, the active source j_1 must satisfy

$$\beta\theta_{j_1}^*\|\mathsf{M}_{j_1}\mathsf{C}_{j_1}^{1/2}\|_{\mathrm{F}}^2 = (\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma), \quad\text{or}\quad \theta_{j_1}^* = \frac{(\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)}{\beta\|\mathsf{M}_{j_1}\mathsf{C}_{j_1}^{1/2}\|_{\mathrm{F}}^2}.$$

This must be true, whichever the active source is, and if each source has equal probability to be active, the exchangeability argument yields the scaling law

$$\theta_j^* = \frac{(\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)}{\beta\|\mathsf{M}_j\mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2}, \quad 1\le j\le n.$$

This argument can be generalized to several, and unknown, number of active sources. Assume that we believe that k of the sources are non-zero, but we do not know which ones. Denoting the indices to the active sources by $j_1, j_2, \ldots, j_k$, we must have

$$\sum_{\ell=1}^k \beta\theta_{j_\ell}^*\|\mathsf{M}_{j_\ell}\mathsf{C}_{j_\ell}^{1/2}\|_{\mathrm{F}}^2 = (\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma). \tag{8}$$

For the exchangeability argument, let us denote by $\gamma\in\mathbb{R}^n$ the vector with entries $\gamma_j = \beta\theta_j^*\|\mathsf{M}_j\mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2$, and by $\mathsf{P}_k\in\mathbb{R}^{n_k\times n}$ the matrix such that the pth row contains exactly k entries equal to one, other entries being zero, and each permutation appears only once in the matrix. Therefore, the number of rows in P_k is $n_k = n!/(k!(n-k)!)$. Since we assume that (8) holds regardless of the selection of the active sources, the vector γ must satisfy the linear system

$$\mathsf{P}_k\gamma = (\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)\mathbf{1}_{n_k},$$

where $\mathbf{1}_{n_k}\in\mathbb{R}^{n_k}$ denotes a vector with unit entries. The only possible solution of this system is

$$\gamma_j = \frac{(\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)}{k},$$

and therefore, we arrive at the scaling law

$$\theta_j^* = \frac{1}{k}\frac{(\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)}{\beta\|\mathsf{M}_j\mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2}, \quad 1\le j\le n.$$

Finally, assume that we only have a prior idea of how many non-zero sources may be active, and we express this belief in a form of a probability density

$$\mathsf{P}\{\#\text{ of active sources} = k\} = p_k, 1\le k\le n,$$

where $\sum_k p_k = 1$. If we expect that out of the n dipoles, it is reasonable to expect that $\overline{k} = sn$ are active, a binomial distribution can be used for p_k, $k_k \sim \mathrm{Binom}(n, s)$, $0 < s < 1$; in practice the binomial can be approximated by a Poisson distribution, $p_k \sim \mathrm{Poisson}(\overline{k})$ with mean $\overline{k}$ provided by the user. Using the previous result, conditioned on k, we arrive at the scaling law

$$\theta_j^* = \frac{C}{\|\mathsf{M}_j\mathsf{C}_j^{1/2}\|_{\mathrm{F}}^2}, \quad 1\le j\le n,$$

with

$$C = (\mathrm{SNR}-1)\times\mathrm{trace}(\Sigma)\sum_{k=1}^n \frac{p_k}{k}.$$

This argument confirms that, in order to match the model with the SNR, the parameters θ_j^* should indeed be chosen to be inversely proportional to the sensitivity.

Appendix: Construction of the activity patch

To select the vertices in the activity patch $\mathcal{P}$, we first pick randomly a seed vertex from the BR of interest, then grow the patch by adding iteratively the neighboring vertices, pruning off at each step those that fall outsize the pertinent BR, and stopping the process as soon as the desired number $N_{\mathcal{P}}$ of vertices have been included. The selected nodes along with the edges of the triangular mesh form a local graph. To generate the activity in the patch, we start by computing a positive graph Laplacian of the patch, which is the matrix $\mathsf{L} \in \mathbb{R}^{N_{\mathcal{P}} \times N_{\mathcal{P}}}$ with entries

$$
L_{i,j} = \left\{ \begin{array}{ll} -\deg(v_i) & \text{if } i = j, \\ 1 & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j, \\ 0 & \text{otherwise,} \end{array} \right.
$$

where $\deg(v_i)$ is the number of the edges that terminate at the vertex v_i.

After defining a correlation length λ, given in units of the number of steps, we draw a random amplitude vector by setting

$$
Q = (\mathsf{L} + \lambda^2 \mathsf{I}_{N_{\mathcal{P}}})^{-1} W,
$$

where $\mathsf{I}_{N_{\mathcal{P}}} \in \mathbb{R}^{N_{\mathcal{P}} \times N_{\mathcal{P}}}$ is the unit matrix and $W \in \mathbb{R}^{N_{\mathcal{P}}}$ is a standard normal Gaussian random vector, that is, $W \sim \mathcal{N}(0, \mathsf{I}_{N_{\mathcal{P}}})$. The amplitudes are scaled so that the amplitude of the dipole at the seed vertex is one. Finally, we draw the dipole moment directions from the anatomical prior, making sure that adjacent dipoles are not pointing in the opposite sides of the cortex patch.

Acknowledgements

References

Aine CJ, Sanfratello L, Ranken D, Best E, MacArthur JA, Wallace T, Gilliam K, Donahue CH, Montano, R, Bryant, JE and others (2012) MEG-SIM: a web portal for testing MEG analysis methods using realistic simulated and empirical data. Neuroinformatics 10(2): 141–158

Algorri ME and Flores-Mangas F (2004) Classification of anatomical structures in MR brain images using fuzzy parameters. IEEE Trans Biomed Engineering 51: 1599–1608

Auranen T, Nummenmaa A, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Vehtari A, Sams M (2005) Bayesian analysis of the neuromagnetic inverse problem with ℓ_p-norm priors. NeuroImage 26: 870–884

Attal, Y, Maess B, Friederici A and David, O (2012) Head models and dynamic causal modeling of subcortical activity using magnetoencephalographic/electroencephalographic data. Rev Neurosci 23: 141–158

Attal Y, Schwartz D (2013) Assessment of subcortical source localization using deep brain activity imaging model with minimum norm operators: a MEG study. PLoS One 8: e59856

Baillet S, Garnero L (1997) A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans BME 44:374–385

Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. IEEE Signal Proc Mag 18:14–30

Bernardo JM, Smith AFM (2004) Bayesian theory. John Wiley & Sons

Brette R and Destexhe A (2012) Handbook of neural activity measurement. Cambridge University Press

Calvetti D, Pascarella A, Pitolli F, Somersalo E, Vantaggi B (2015) A hierarchical Krylov–Bayes iterative inverse solver for MEG with physiological preconditioning. Inverse Problems 31:125005

Calvetti D, Pitolli F, Somersalo E, Vantaggi B (2018) Bayes meets Krylov: preconditioning CGLS for underdetermined systems. SIAM Rev 60: DOI 10.1137/15M1055061

Calvetti D, Hakula H, Pursiainen S, Somersalo E (2009) Conditionally Gaussian hypermodels for cerebral source localization. SIAM J Imag Sci 2: 879–909

Ciofolo, C and Barillot C (2009) Atlas-based segmentation of 3D cerebral structures with competitive level sets and fuzzy control. Medical Image Analysis 13: 456–470

Collins, D Louis and Zijdenbos, Alex P and Kollokian, Vasken and Sled, John G and Kabani, Noor J and Holmes, Colin J and Evans, Alan C (1998) Design and construction of a realistic digital brain phantom. IEEE Trans BME 17: 463–468

24

Dale AM, Fischl B and Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. NeuroImage 9: 179–194

Dale MA, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD and Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26: 55–67

de Munck JC, Vijn PCM and da Silva FH Lopes (1992) A random dipole model for spontaneous brain activity. IEEE Trans BME 39: 791–804

Desikan RS , Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale, AM, Maguire RP, Hyman BT and others (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31: 968–980

Gorodnitsky IF and Rao BD (1997) Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. IEEE Trans. on Signal Processing 45: 600–616

Gramfort A, Papadopoulos T, Olivi E and Clerc M (2010) Open MEEG: open source software for quasistatic bioelectromagnetics. Biomedical Engineering Online 9

Gramfort A, Luessi, M, Larson E, Engemann, DA, Strohmeier D, Brodbeck C, Parkkonen L and Hämäläinen MS (2014) NeuroImage 31: 446–460

Hedrich T, Pellegrino G, Kobayashi E, Lina JM and Grova G (2017) Comparison of the spatial resolution of source imaging techniques in high-density EEG and MEG. NeuroImage 175:531–544.

Henson RN, Mattout J, Phillips C and Friston KJ (2009) Selecting forward models for MEG source-reconstruction using model-evidence. NeuroImage 46: 168–176

Henson, RN, Flandin G, Friston KJ and Mattout J (2010) A parametric empirical Bayesian framework for fMRI-constrained MEG/EEG source reconstruction. Human Brain Mapping 31: 1512–1531

Hämäläinen MS, Hari R, Ilmoniemi RJ, Knuutila, J and Lounasmaa, OV (1993) Magnetoencephalography–theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys. 65: 413–498

Hämäläinen MS and Ilmoniemi RJ (1984) Interpreting measured magnetic fields of the brain: estimates of current distributions. Report TKK-F-A559

Huizenga HM, De Munck JaC, Waldorp LJ and Grasman RPPP (2002) Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. IEEE Trans BME 49: 533–539

Kass RE and Raftery AE (1995) Bayes factors. JASA 90: 773–795

Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. Kiebel SJ, Daunizeau J, Phillips, C and Friston KJ (2008) NeuroImage 39: 728–741

Kybic, J, Clerc M, Abboud T, Faugeras O, Keriven R and Papadopoulos T (2005) A common formalism for the integral formulations of the forward EEG problem. IEEE Trans Med Imaging 24: 12–28

Lin FH, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW and Hämäläinen, MS (2006) Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. NeuroImage 31: 160–171

Lin FH, Belliveau JW, Dale AM and Hämäläinen MS (2006) Distributed current estimates using cortical orientation constraints. Human Brain Mapping 27: 1–13

Lopez,JD, Litvak V, Espinosa JJ, Friston K, Barnes GR (2014) Algorithmic procedures for Bayesian MEG/EEG source reconstruction in SPM. NeuroImage 84: 476–487

Lucka F, Pursiainen S, Burger M and Wolters CH (2012) Hierarchical Bayesian inference for the EEG inverse problem using realistic FE head models: depth localization and source separation for focal primary currents. NeuroImage 61: 1364–1382

Mattout J, Phillips C, Penny WD, Rugg MD, Friston KJ (2006) MEG source localization under multiple constraints: an extended Bayesian framework. NeuroImage 30: 753–767

Molins A, Stufflebeam SM, Brown EN and Hämáläinen MS (2008) Quantification of the benefit from integrating MEG and EEG data in minimum ?2-norm estimation. Neuroimage, 42(3):1069-1077.

Mosher JC, Spencer ME, Leahy RM and Lewis PS (1993) Error bounds for EEG and MEG dipole source localization. Electroenceph clin Neurophysiol 86: 303–321

Nagarajan SS, Portniaguine O, Hwang D, Johnson C, Sekihara K (2006) Controlled support MEG imaging. NeuroImage 33: 878–885

Nummenmaa A, Auranen T, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Sams M, Vehtari A (2007) Hierarchical Bayesian estimates of distributed MEG sources: theoretical aspects and comparison of variational and MCMC methods. NeuroImage 35: 669–685

Nummenmaa A, Auranen T, Hämäläinen MS, Jääskeläinen IP, Lampinen J, Sams M, Vehtari A (2007) Automatic relevance determination based hierarchical Bayesian MEG inversion in practice. NeuroImage 3: 876–889

Owen JP, Wipf DP, Attias HT, Sekihara K, Nagarajan SS (2012) Performance evaluation of the Champagne source reconstruction algorithm on simulated and real M/EEG data. NeuroImage 60: 305–323

Parkkonen L, Fujiki N, Mäkelä JP (2009) Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. Human Brain Mapping 30:1772–1782.

Pascual-Marqui RD (1999) Review of methods for solving the EEG inverse problem. Int J Bioelectromagn 1:75–86

Sato M, Yoshioka T, Kajihara S, Toyama K, Goda N, Doya K, Kawato M (2004) Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage 23:806–826

Bayesian model selection for group studies. Stephan KE, Penny WD, Daunizeau J, Moran RJ and Friston KJ (2009) NeuroImage 46: 1004–1017

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. Comp Intell Neurosci 2011:879716.

Trujillo-Barreto NJ, Aubert-Vázquez E, Valdés-Sosa PA (2004) Bayesian model averaging in EEG/MEG imaging. NeuroImage 21:1300–1319

Uutela K, Hämäläinen MS, Somersalo E (1999) Visualization of magnetoencephalographic data using minimum current estimates. NeuroImage 10:173–180

Wagner M, Fuchs M, Kastner J (2004) Evaluation of sLORETA in the presence of noise and multiple sources. Brain Topography 16:277–280

Wipf D, Nagarajan S (2009) A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44:947–966

Wipf DP, Owen JP, Attias HT, Sekihara K, Nagarajan SS (2010) Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. NeuroImage 49:641–655

Vorwerk J, Cho JH, Rampp S, Hamer H, Knösche TR and Wolters CH (2014) A guideline for head volume conductor modeling in EEG and MEG. NeuroImage 100: 590–607

Activity Patch

Activity Patch
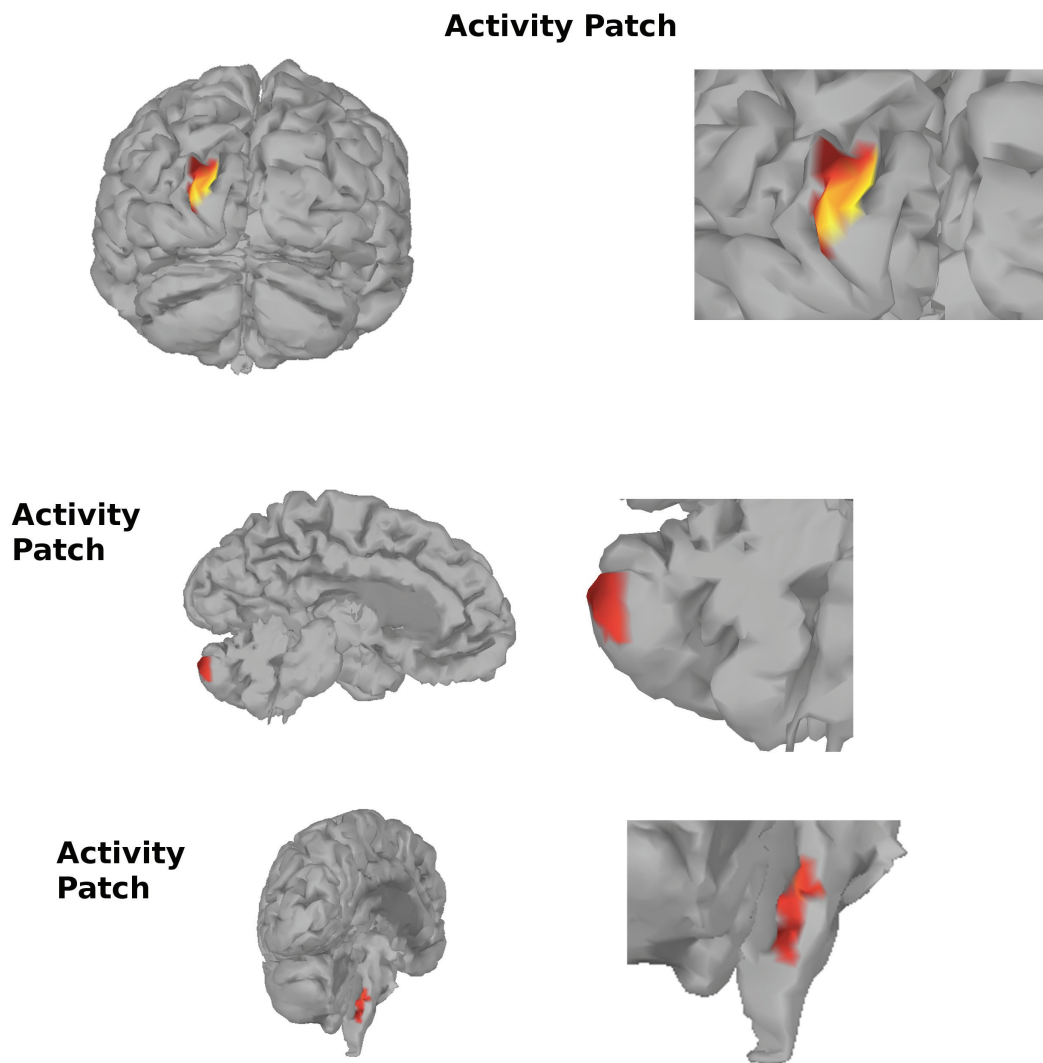
Activity Patch

Figure 2: Top row: A randomly generated patch of activity in the left lateral occipital cortex region of the brain. Middle row: A randomly generated patch of activity in right cerebellum. Bottom row: A randomly generated patch of activity in the brainstem.
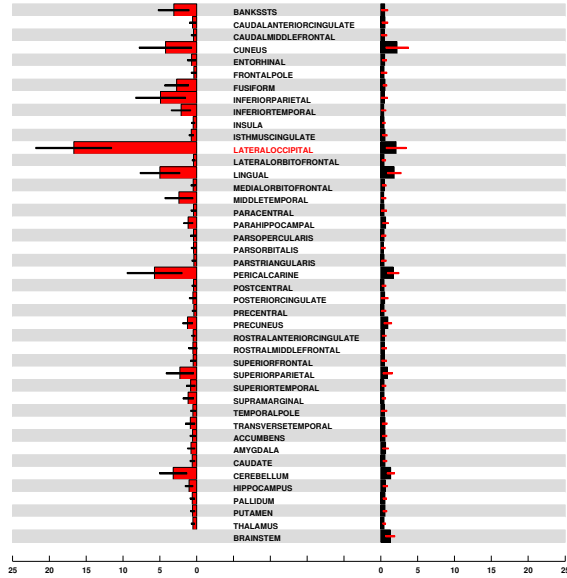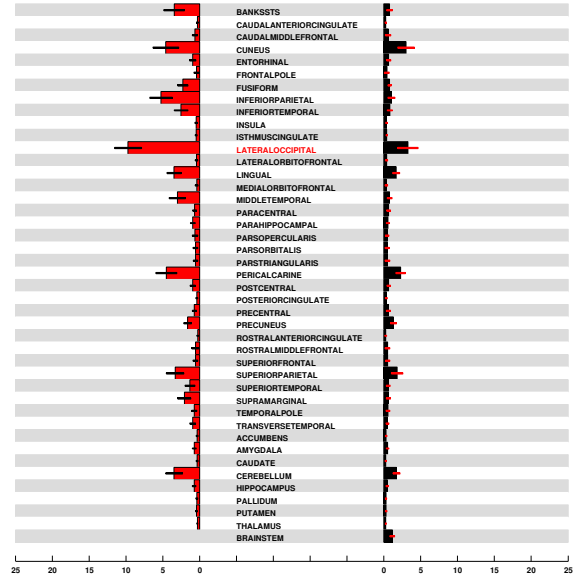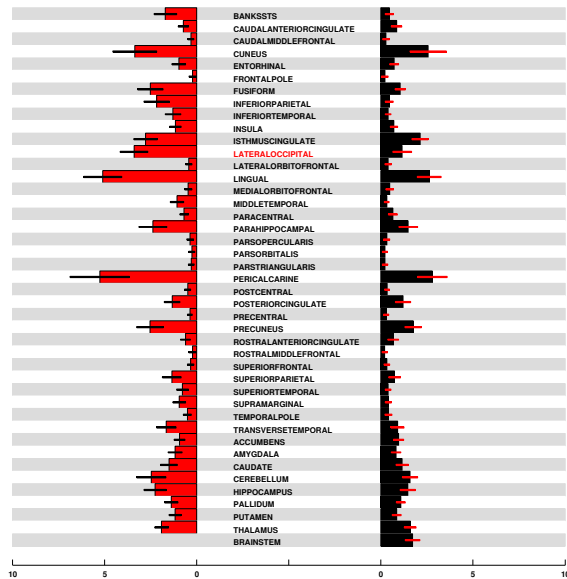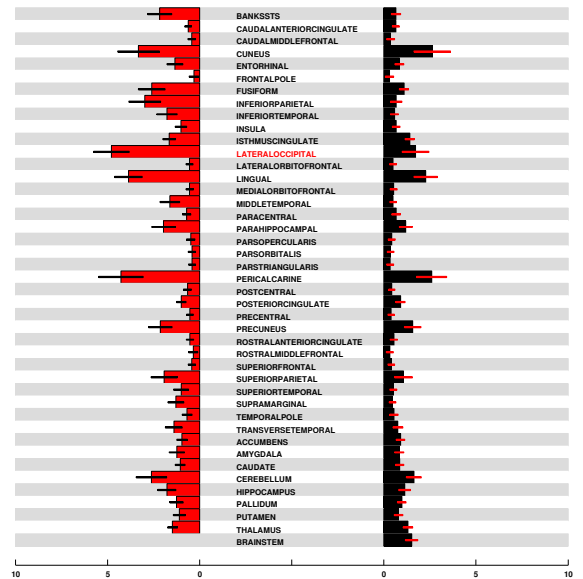
Figure 3: Mapping of the brain activity to 85 different BRs over 100 simulations using synthetic data corresponding to randomly generated activity patches in the *left lateral occipital cortex*, indicated in red in the list of the BRs reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.

IAS
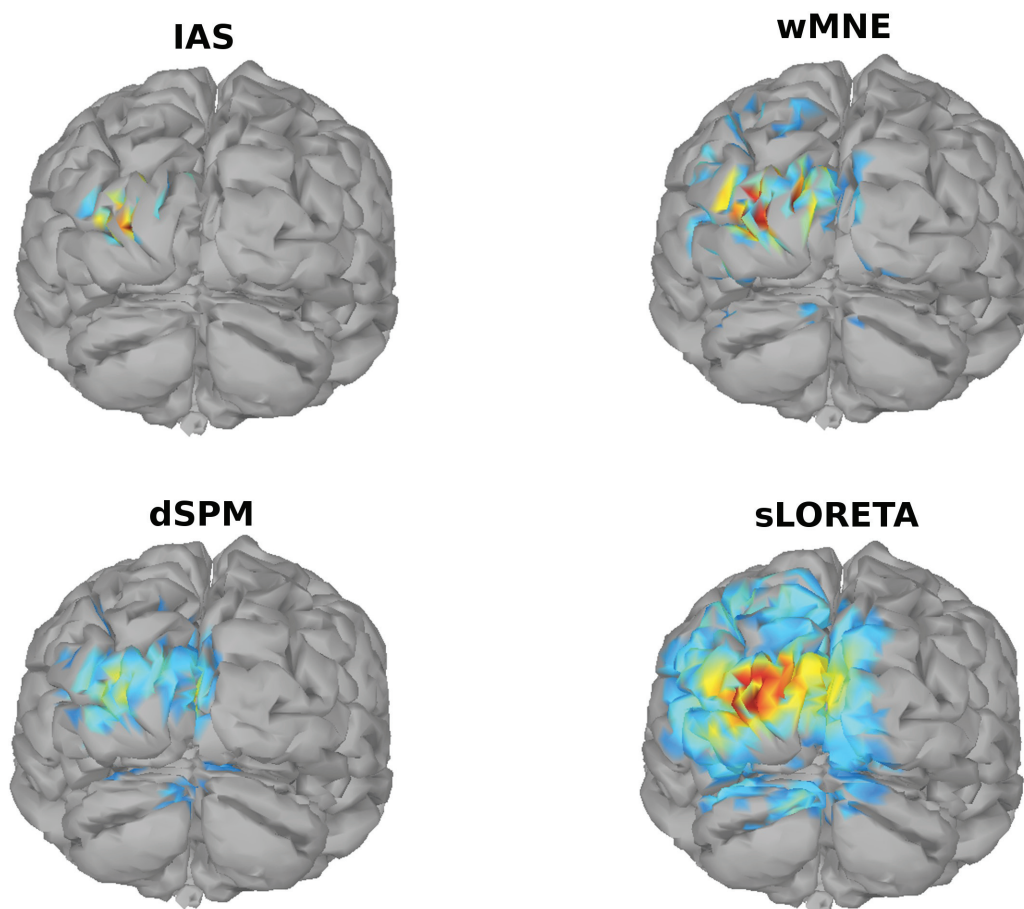
wMNE

dSPM

sLORETA

Figure 4: Reconstructions obtained with, clockwise from the top left, IAS, wMNE, dSPM and sLORETA, respectively, from the signal generated by the activated patch in the left lateral occipital region shown in the top row of Figure 2.
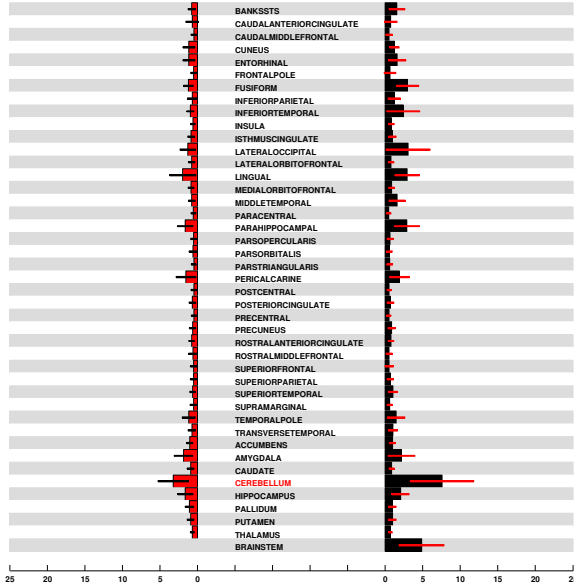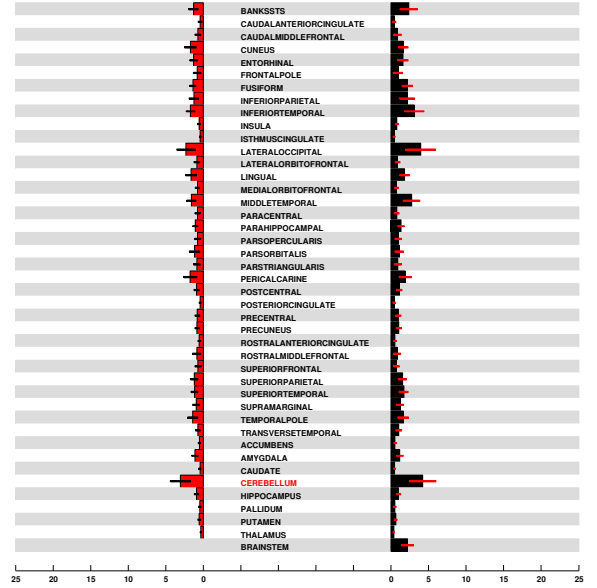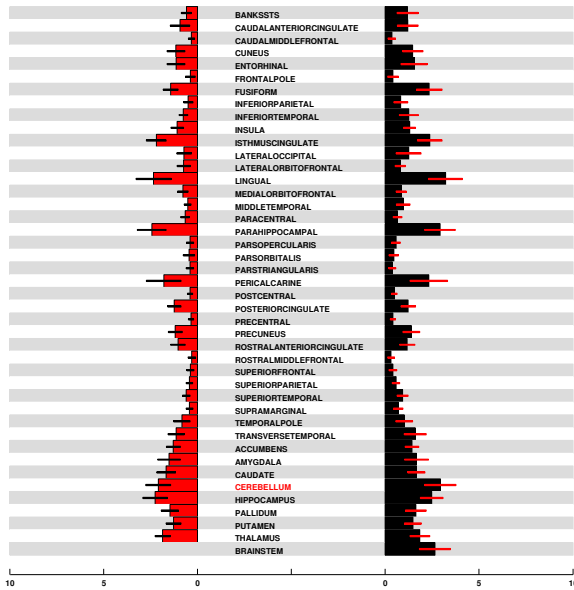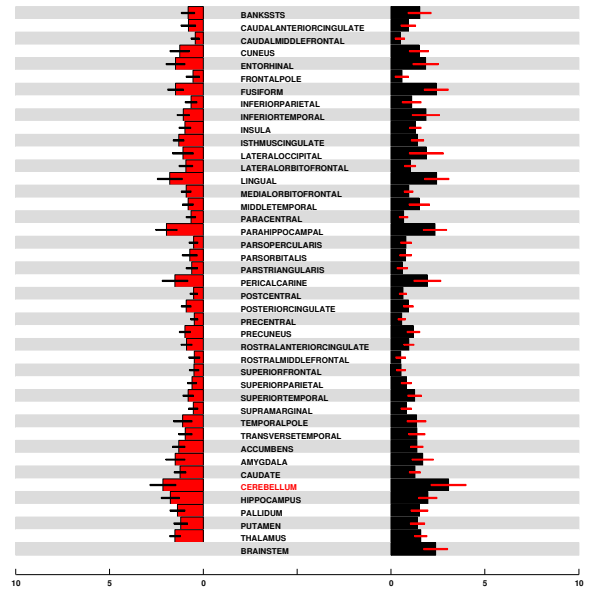
Figure 5: Mapping of the brain activity to 85 different BRs over 100 simulations using synthetic data corresponding to randomly generated activity patches in the *right cerebellum*, indicated in red in the list of the BRs reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.

Left

Right

IAS
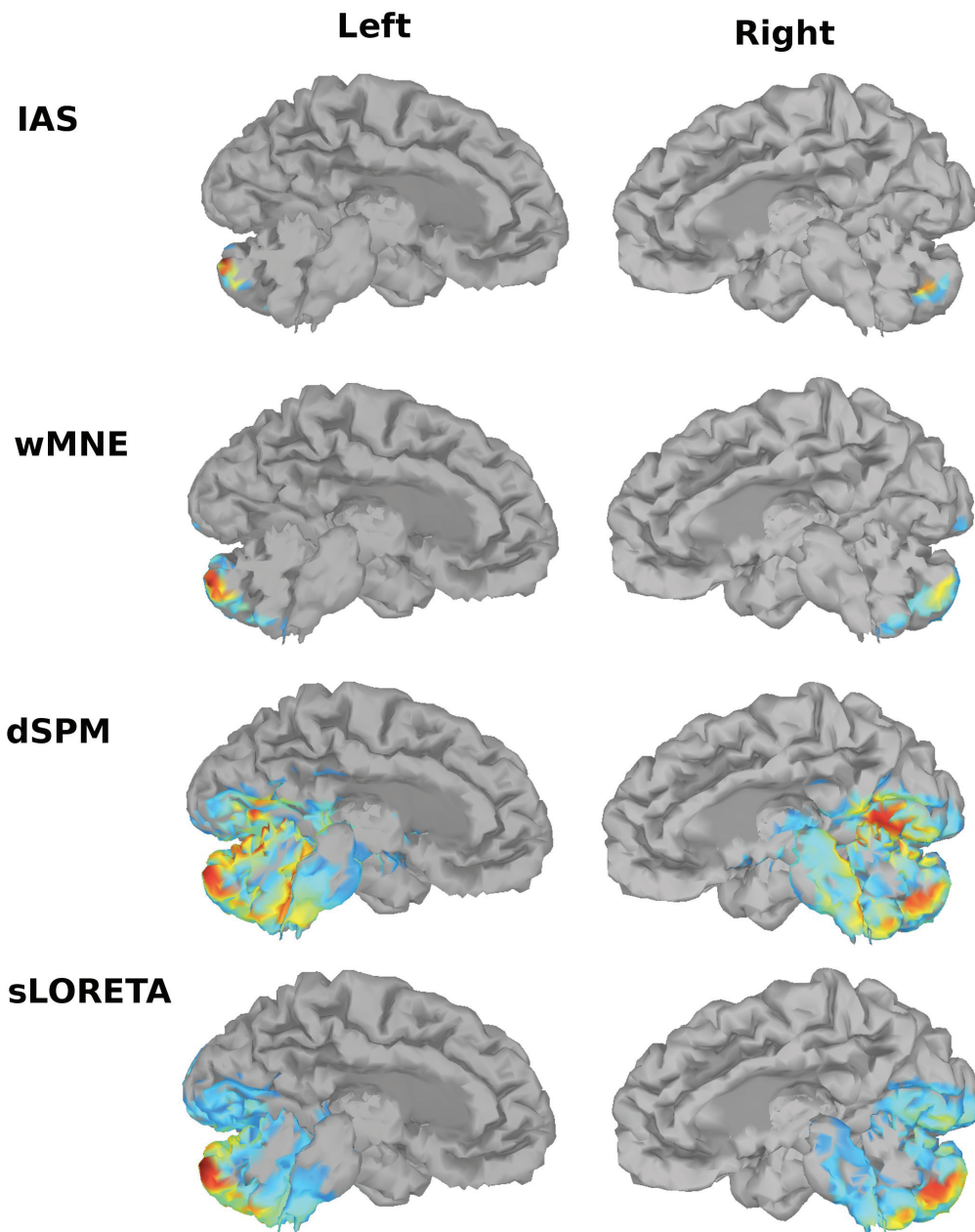
wMNE

dSPM

sLORETA

Figure 6: Reconstructions obtained with, clockwise from the top left, IAS, wMNE, dSPM and sLORETA, respectively, from the signal generated by the activated patch in the right cerebellum shown in the middle row of Figure 2.
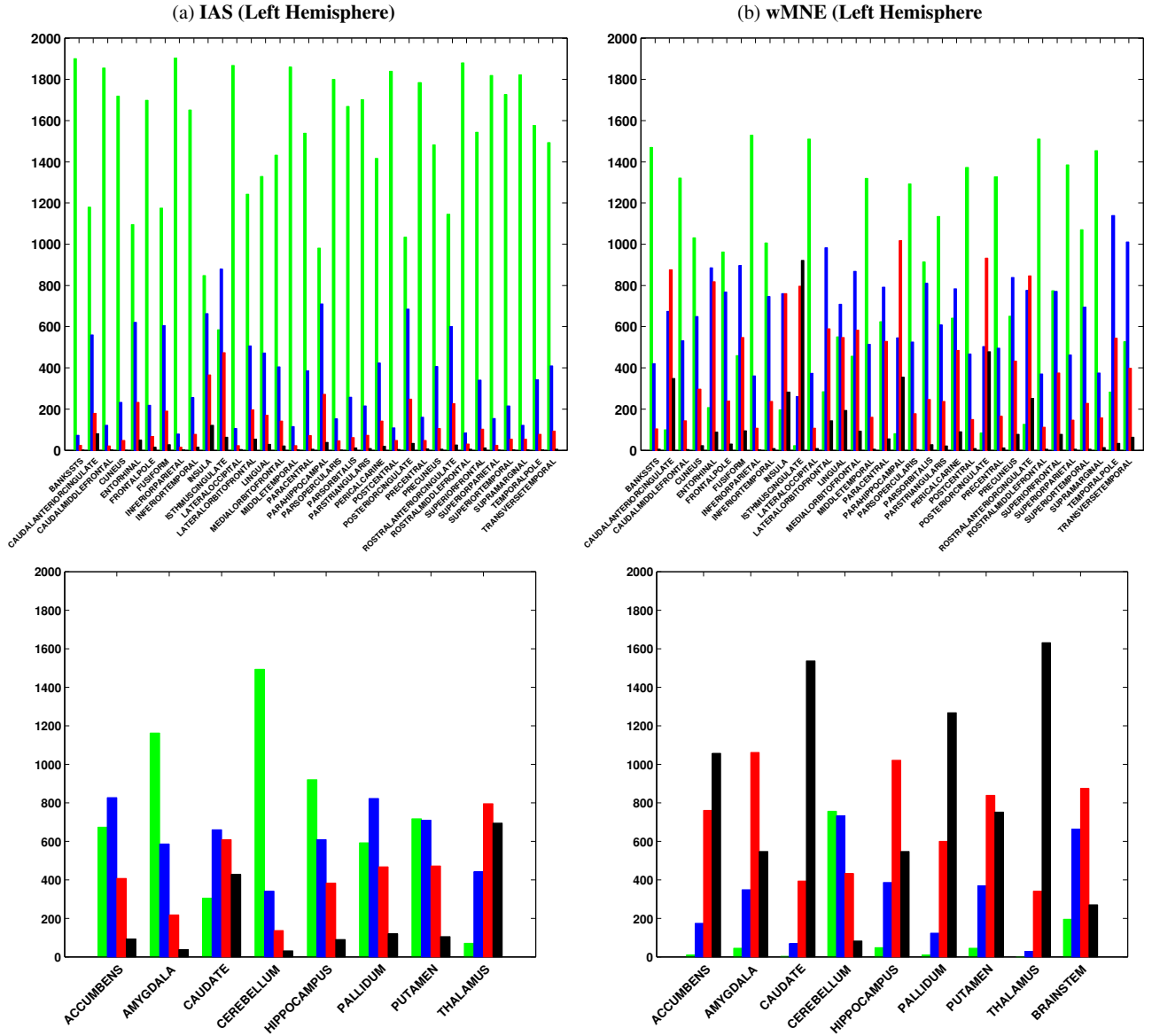
Figure 7: Histograms of the Bayes factors for the cortical (top) and subcortical (bottom) BRs when using the IAS algorithm (left panel) and the wMNE algorithm (right panel) to map the MEG data to brain activity in the left hemisphere. Each histogram for each BR is the summary of 2000 Bayes factor computations color coded to indicate the different level of support of the hypothesis: green for overwhelming support, blue for strong support, red for weak support, black no support.
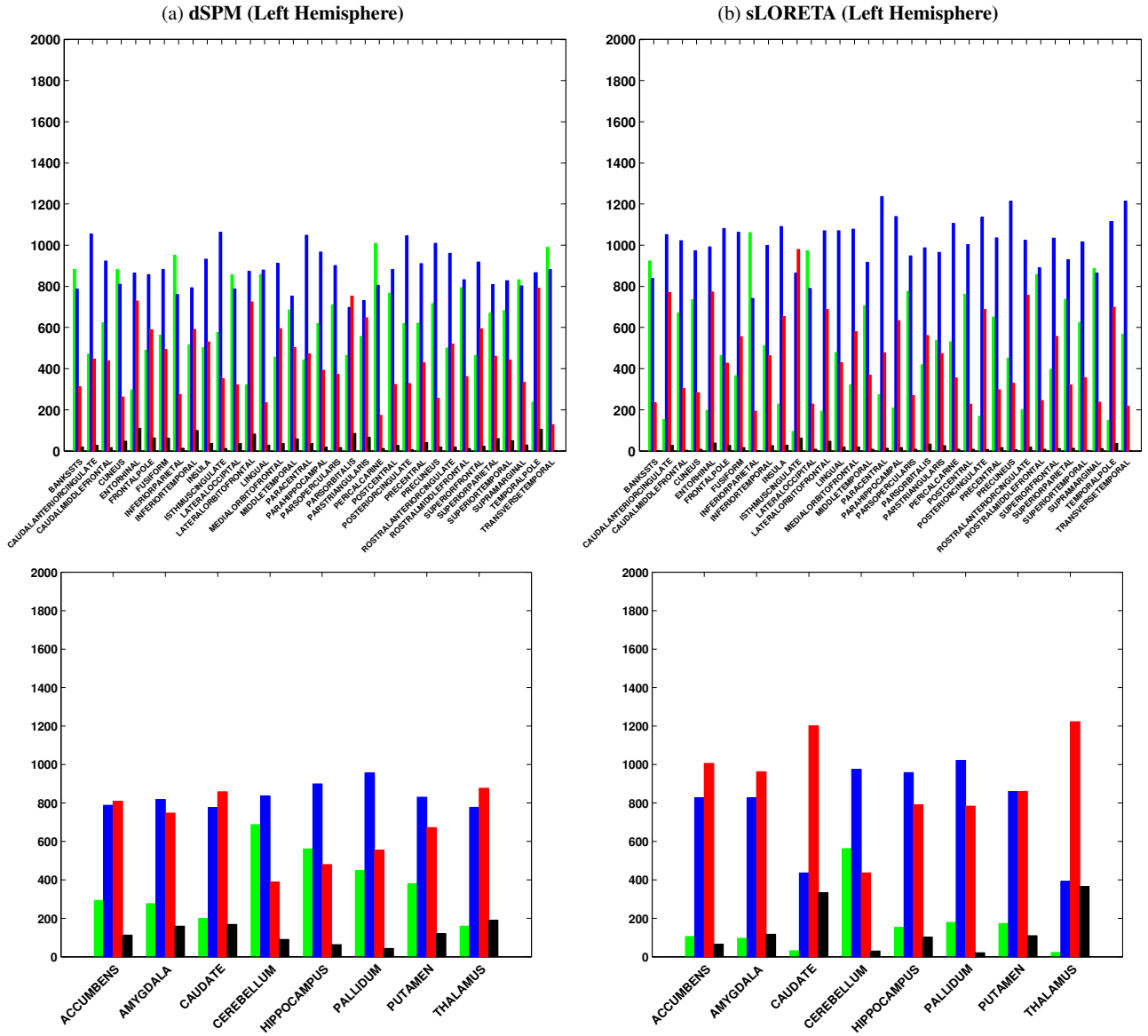
Figure 8: Histograms of the Bayes factors for the cortical (top) and subcortical (bottom) BRs when using the dSPM algorithm (left panel) and the sLORETA algorithm (right panel)to map the MEG data to brain activity in the left hemisphere. Each histogram for each BR is the summary of 2000 Bayes factor computations and the color coding is as explained in Figure 7.
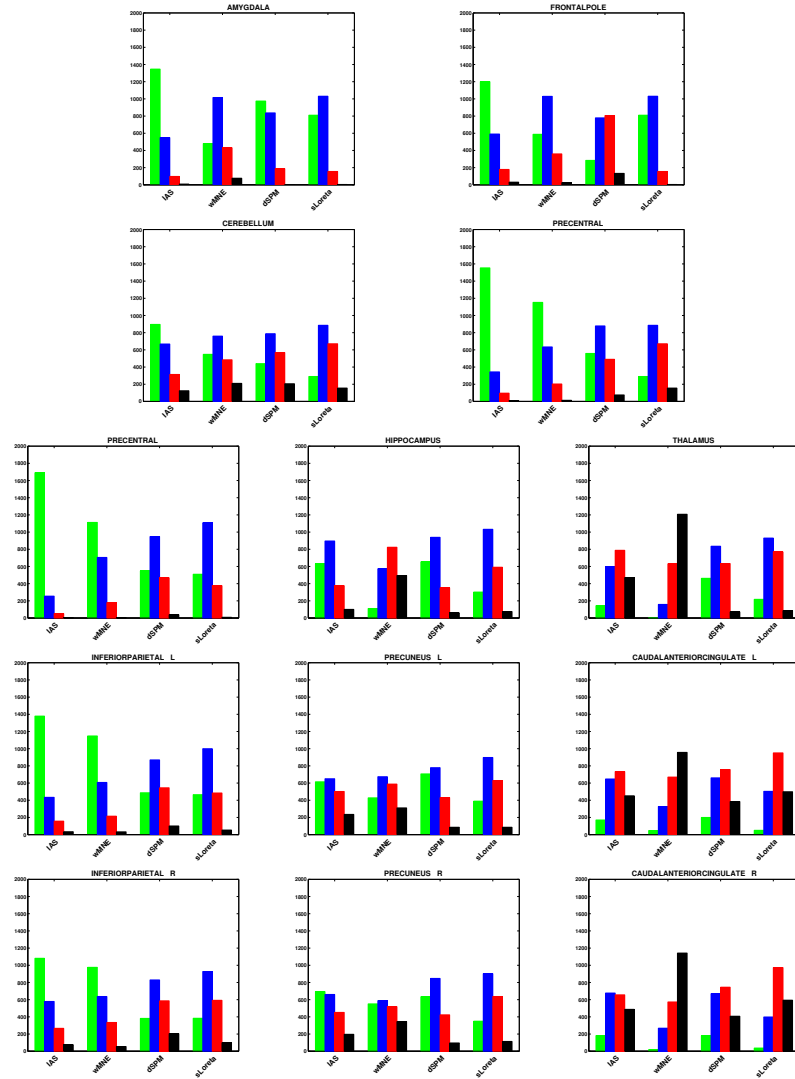
Figure 9: Row-wise from top to bottom: Row 1: Bayes factors in the case of two active patches. Left panel: Bayes factors in support of the hypothesis that the activity is in the *left amygdala* when the data comes from two active patches, one in the left amygdala and the other in the left frontal pole, and the MEG inverse problem is solved with the IAS, wMNE, dSPM and sLORETA algorithms, respectively. The histograms where computed from 2000 realizations and the color coding is as explained in Figure 7. Right panel: Bayes factors in support of activity in the *left frontal pole* computed from the same 2000 realizations. Row 2: Bayes factors in the case of two active patches. Left panel: Bayes factors in support of the hypothesis that the activity is in the *left cerebellum* when the data comes from two active patches, one in the left cerebellum and the other in the left precentral cortex, and the MEG inverse problem is solved with the IAS, wMNE, dSPM and sLORETA algorithms, respectively. The histograms were computed from 2000 realizations Right panel: Bayes factors in support of activity in the *left precentral cortex* computed from the same 2000 realizations. Row 3: Bayes factors in the case of active patches in three different BRs, the right precentral gyrus (left), and the left hippocampus (center) and the right thalamus (right), and the MEG inverse problem is solved with the IAS, wMNE, dSPM and sLORETA algorithms, respectively. The histograms were computed from 2000 realizations. Rows 5-6: Bayes factors for the case where there are active patches in all six BRs comprising the Default Mode Network: inferior parietal gyrus, precuneus and caudal anterior cingulate gyrus in the left (top) and right (bottom) hemispheres to test the specificity of the IAS inverse solver and the three reference inverse solvers.
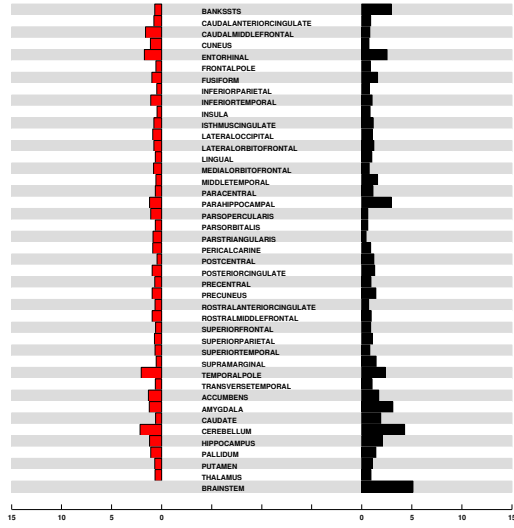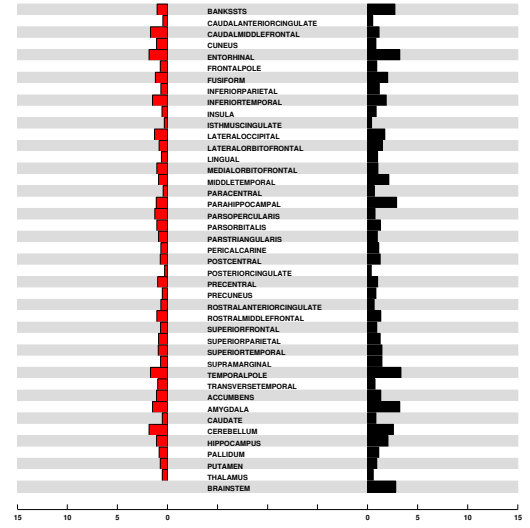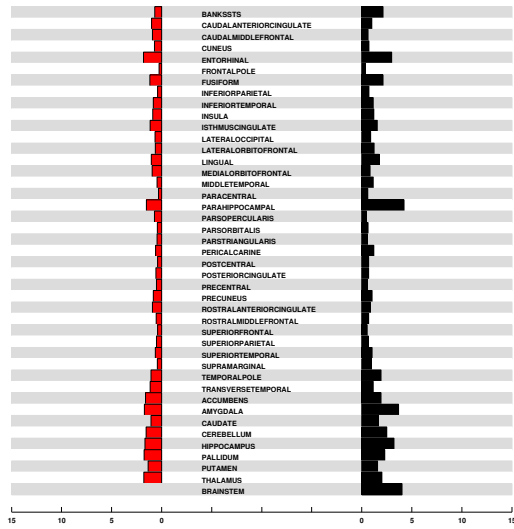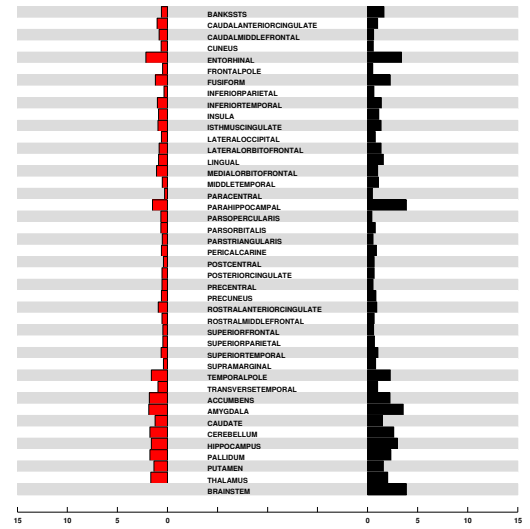
Figure 10: Mapping of the brain activity to 85 different BRs for the time $T = 5$ ms with real/realistic data example reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.
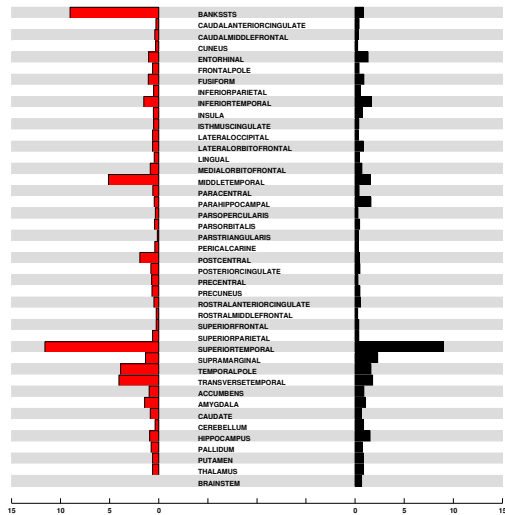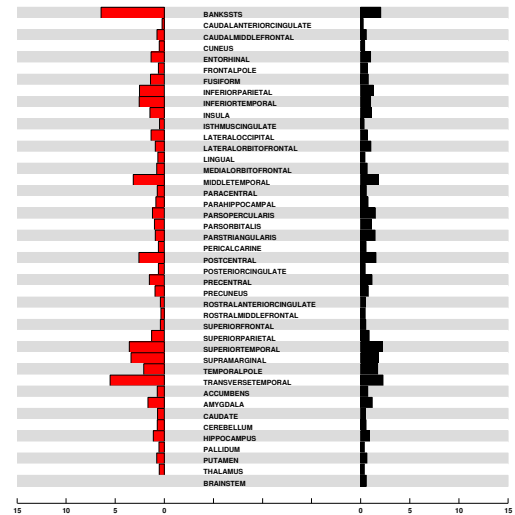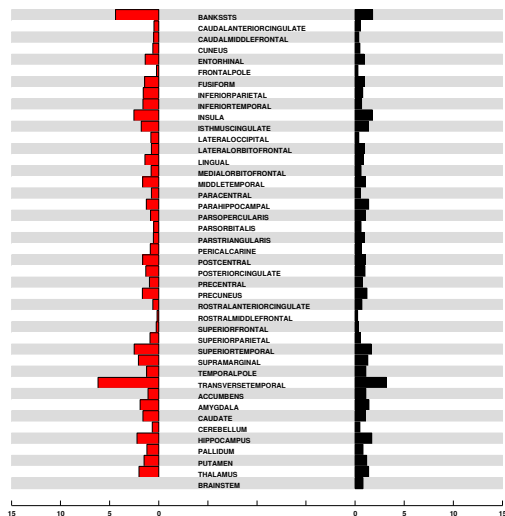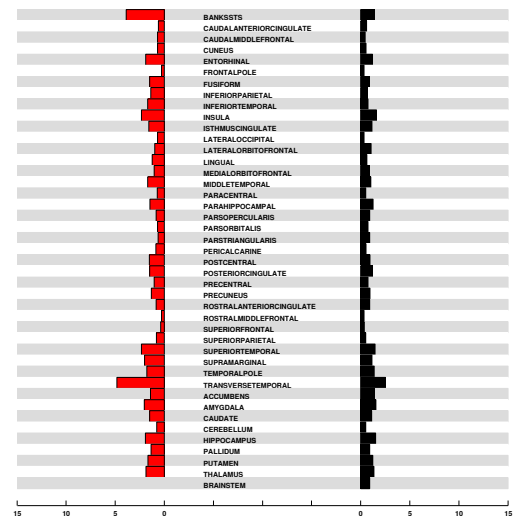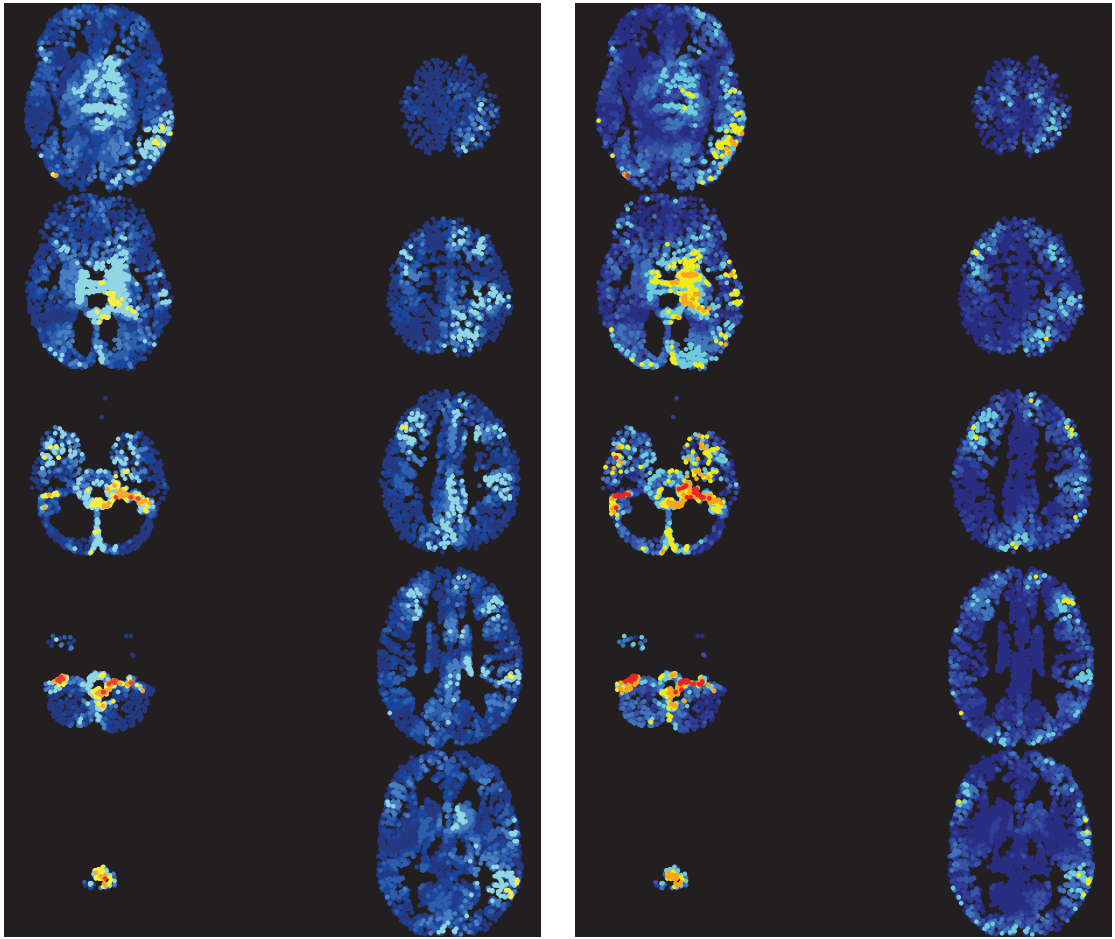
Figure 11: Mapping of the brain activity to 85 different BRs for the time $T = 100$ ms with real/realistic data example reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.
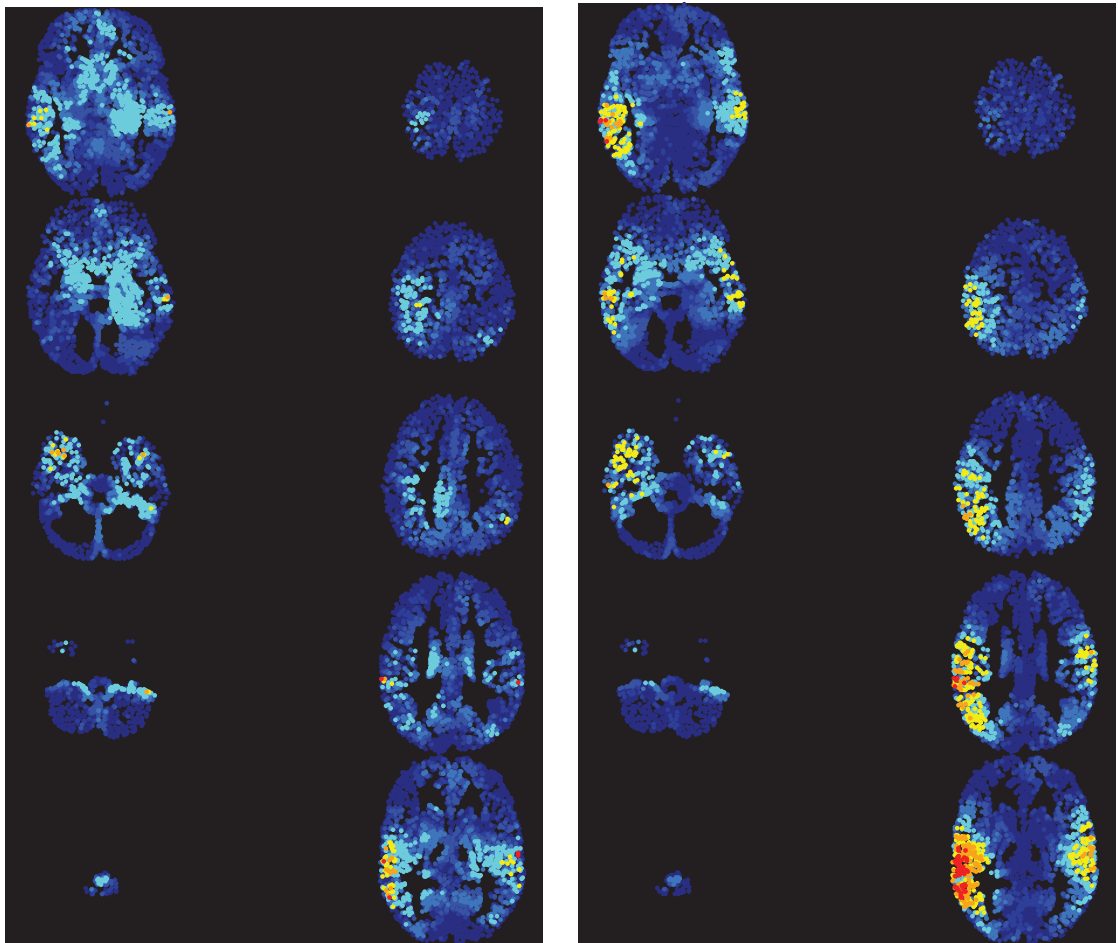
Figure 12: Visualization of the axial view of the brain activity at $T = 5$ ms as reconstructed by IAS (left) and wMNE (right).
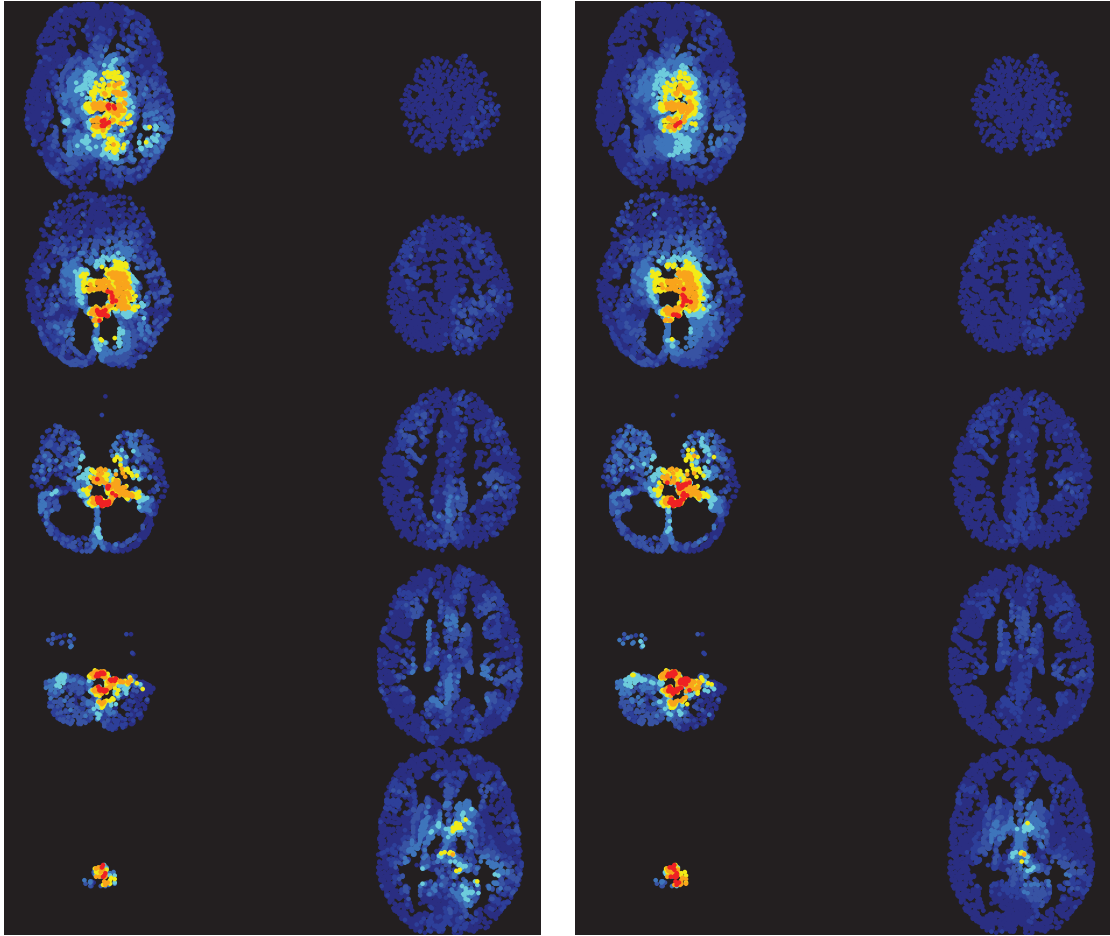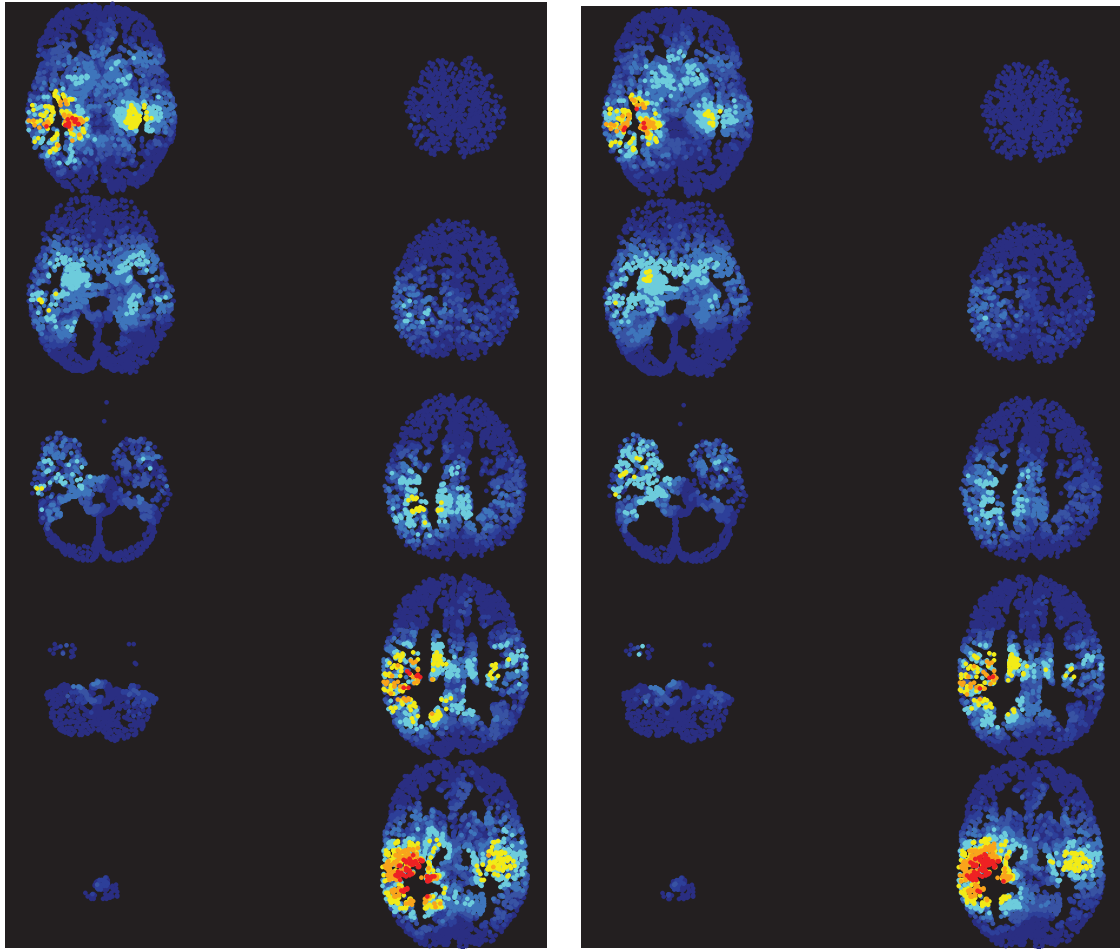
Figure 13: Visualization of the axial view of the brain activity at $T = 5$ ms as reconstructed by dSPM (left) and sLORETA (right).
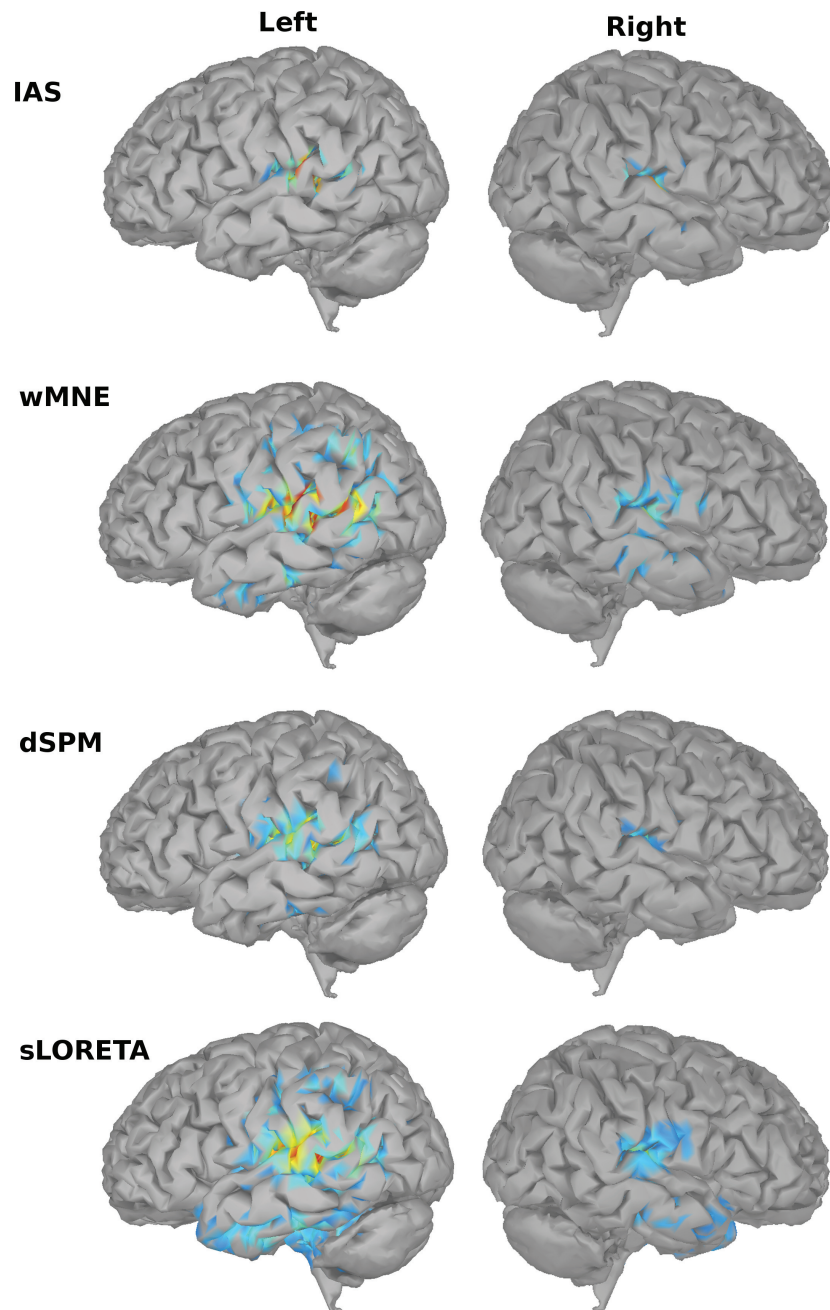
IAS
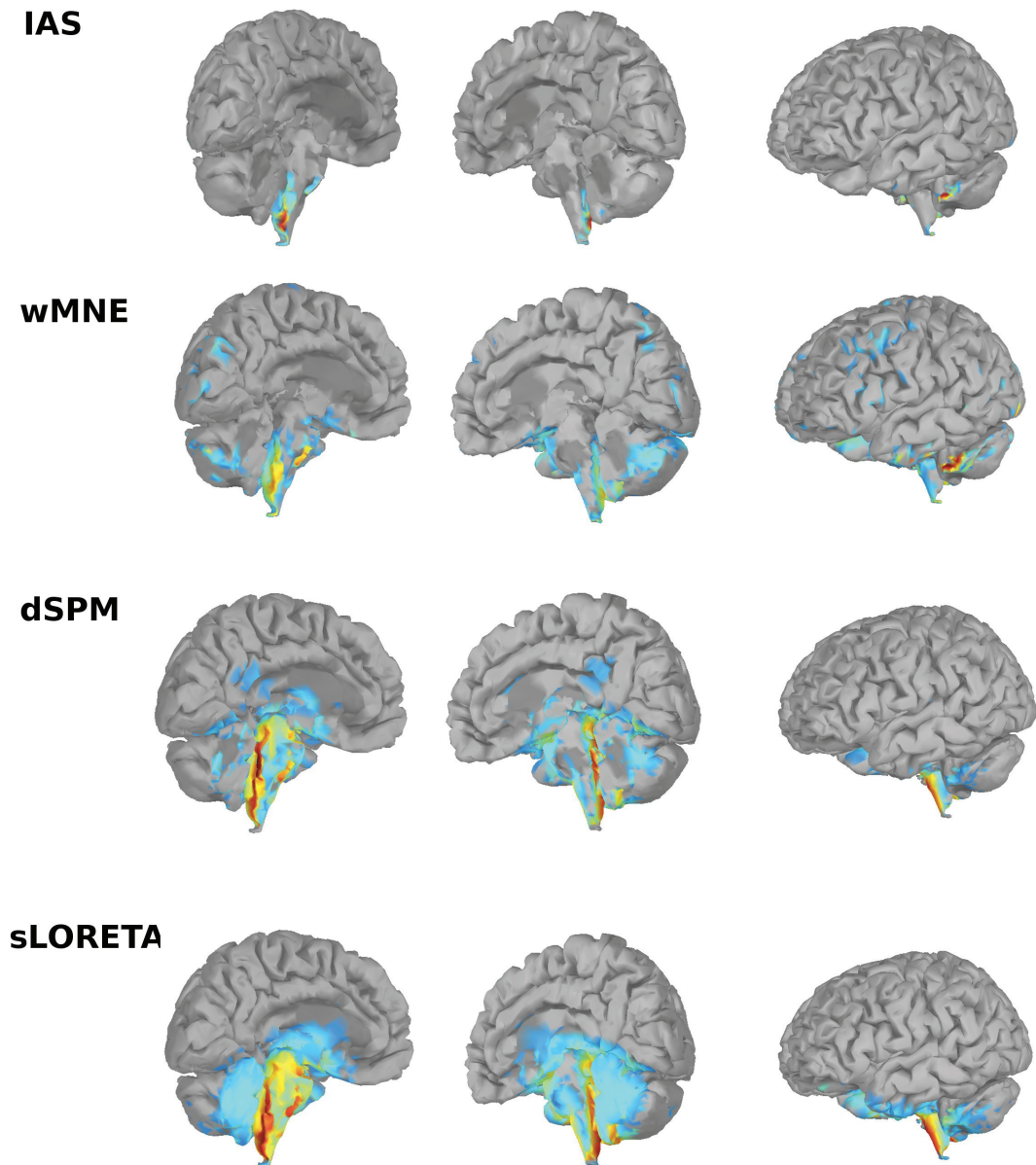
wMNE

dSPM

sLORETA

Figure 14: Reconstructions obtained from real data at time $T = 5$ ms with, top to bottom, IAS, wMNE, dSPM and sLORETA, respectively. At this time instant, most of the signal is generated by the patch activity in the brainstem shown in the bottom row of Figure 2.

Figure 15: Visualization of the axial view of the brain activity at $T = 100$ ms as reconstructed by IAS (left) and wMNE (right).

Figure 16: Visualization of the axial view of the brain activity at $T = 100$ ms as reconstructed by dSPM (left) and sLORETA (right).
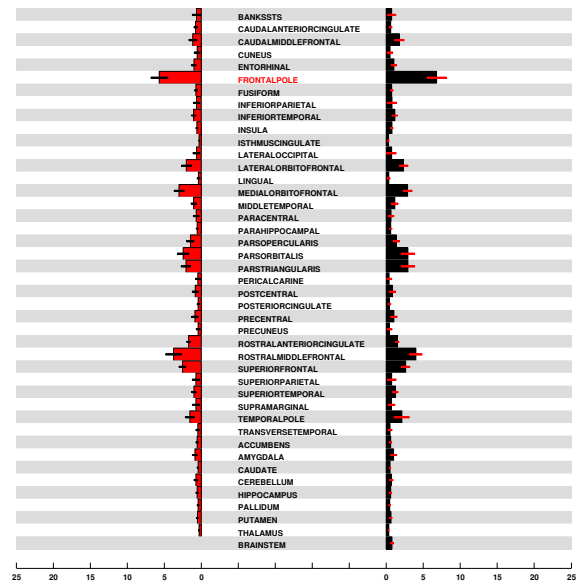
Figure 17: Reconstructions obtained from real data at time $T = 100$ ms with, top to bottom, IAS, wMNE, dSPM and sLORETA, respectively.

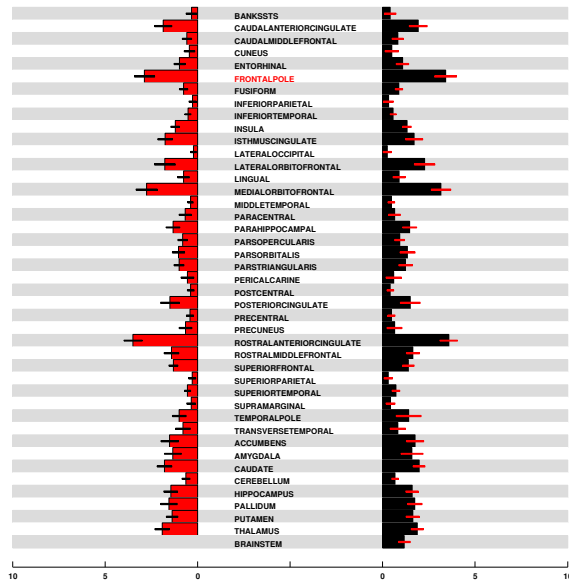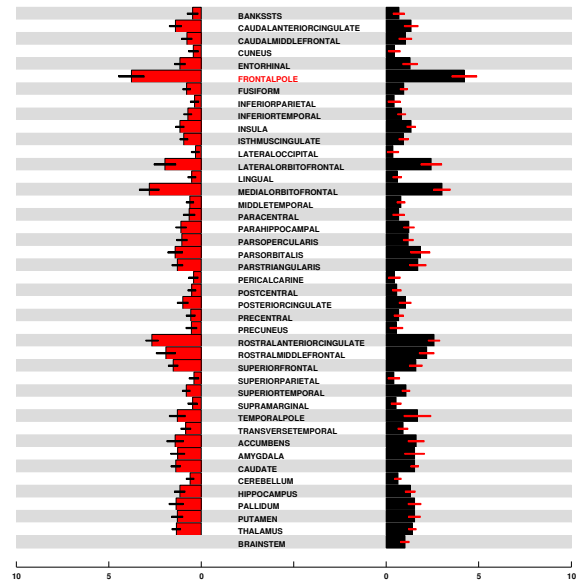Supplementary material
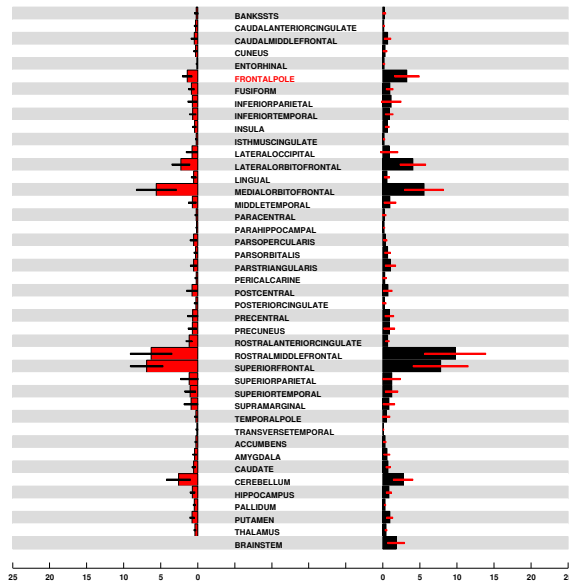
Figure S1: Mapping of the brain activity to 85 different BRs over 100 simulations using synthetic data corresponding to randomly generated activity patches in the *right frontal pole*, indicated in red in the list of the BRs reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.

Figure S2: Mapping of the brain activity to 85 different BRs over 100 simulations using synthetic data corresponding to randomly generated activity patches in the *left amygdala*, indicated in red in the list of the BRs reconstructed with, respectively, IAS (a), wMNE (b), dSPM (c) and sLORETA (d). The histograms bin the average activity in each BR: in red the BRs of the left hemisphere and in black the ones of the right hemisphere.