Open access • Journal Article • DOI:10.1109/MSP.2015.2413711

Brain-Source Imaging: From sparse to tensor models — Source link

Hanna Becker, Laurent Albera, Pierre Comon, Rémi Gribonval ...+2 more authors

Institutions: University of Rennes, French Institute for Research in Computer Science and Automation

Published on: 14 Oct 2015 - IEEE Signal Processing Magazine (IEEE)

Topics: Identifiability and Inverse problem

Brain source imaging: from sparse to tensor models

Hanna Becker, Laurent Albera, Pierre Comon, Rémi Gribonval, Fabrice Wendling, Isabelle Merlet

Brain source imaging: from sparse to tensor models

Hanna Becker[1,2], **Laurent Albera**[3,4,5], **Pierre Comon**[2], **Rémi Gribonval**[5], **Fabrice Wendling**[3,4] **and Isabelle Merlet**[3,4]

[1] I3S, Université de Nice Sophia-Antipolis, CNRS, F-06903, France;

[2] GIPSA-Lab, CNRS UMR5216, Grenoble Campus, St Martin d'Heres, F-38402;

[3] INSERM, U1099, Rennes, F-35000, France;

[4] Université de Rennes 1, LTSI, Rennes, F-35000, France;

[5] INRIA, Centre Inria Rennes - Bretagne Atlantique, France.

For correspondence: Laurent Albera, LTSI, Campus de Beaulieu, Université de Rennes 1, 263 Avenue du General Leclerc - CS 74205 - 35042 Rennes Cedex, France.

Tel: (33) - 2 23 23 50 58, E-Mail: laurent.albera@univ-rennes1.fr

Abstract

A number of application areas such as biomedical engineering require solving an underdetermined linear inverse problem. In such a case, it is necessary to make assumptions on the sources to restore identifiability. This problem is encountered in brain source imaging when identifying the source signals from noisy electroencephalographic or magnetoencephalographic measurements. This inverse problem has been widely studied during the last decades, giving rise to an impressive number of methods using different priors. Nevertheless, a thorough study of the latter, including especially sparse and tensor-based approaches, is still missing. In this paper, we propose i) a taxonomy of the algorithms based on methodological considerations, ii) a discussion of identifiability and convergence properties, advantages, drawbacks, and application domains of various techniques, and iii) an illustration of the performance of selected methods on identical data sets. Directions for future research in the area of biomedical imaging are eventually provided.

I. Introduction

In brain source imaging, one is confronted with the analysis of a linear static system - the head volume conductor - that relates the electromagnetic activity originating from a number of sources located inside the brain to the surface of the head, where it can be measured with an array of electric or magnetic

sensors using Electroencephalography (EEG) or Magnetoencephalography (MEG). The source signals and locations contain valuable information about the activity of the brain, which is crucial for the diagnosis and management of some diseases such as epilepsy or for the understanding of the brain functions in neuroscience research. Albeit, without surgical intervention, the source signals cannot be observed directly and have to be identified from the noisy mixture of signals originating from all over the brain, which is recorded by the EEG/MEG sensors at the surface of the head. This is known as the inverse problem. On the other hand, deriving the EEG/MEG signals for a known source configuration is referred to as the forward problem (cf. Fig. 1). Thanks to refined models of head geometry and advanced mathematical tools that permit to compute the so-called lead-field matrix (referred to as the mixing matrix in other domains), solving the forward problem has become straightforward, whereas finding a solution to the inverse problem is still a challenging task.

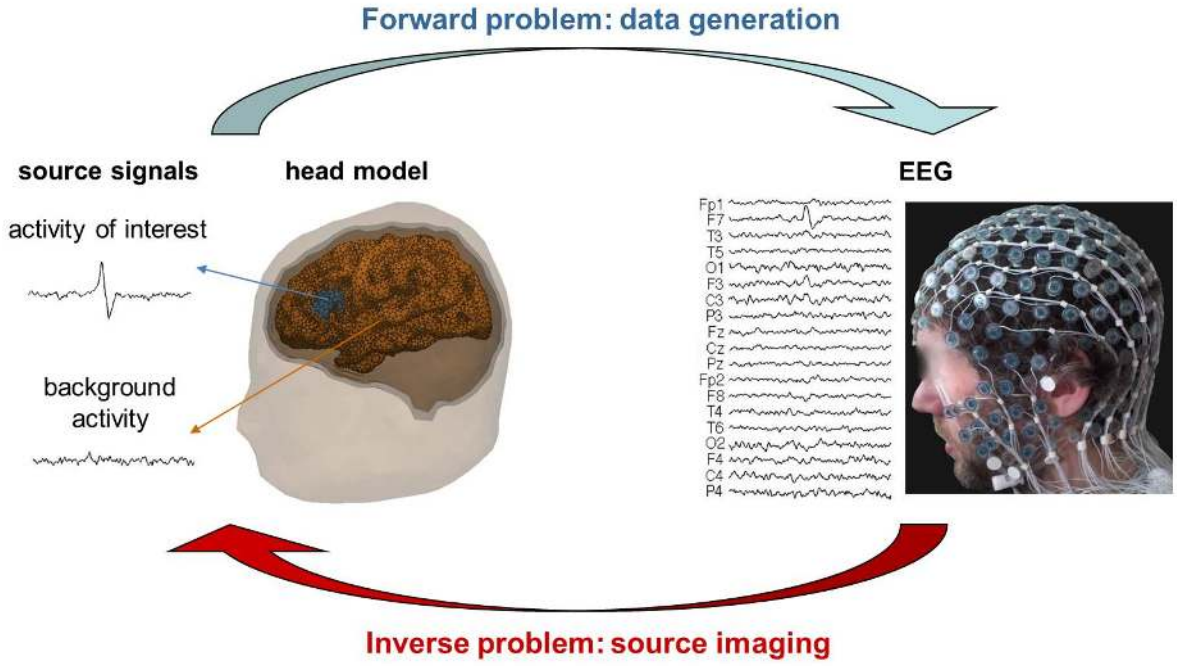

Fig. 1. Illustration of the forward and inverse problems in the context of EEG.

The methods that are currently available for solving the inverse problem of the brain can be broadly classified into two types of approaches that are based on different source models: the equivalent current dipole and the distributed source [26]. Each equivalent current dipole describes the activity within a spatially extended brain region, leading to a small number of active sources with free orientations and

positions anywhere within the brain. The lead-field matrix is hence not known, but parameterized by the source positions and orientations. Equivalent current dipole methods also include the well-known Multiple Signal Classification (MUSIC) algorithm [42], [1] and beamforming techniques (see [48] and references therein). These methods are based on a fixed source space with a large number of dipoles, from which a small number of equivalent current dipoles are identified. On the other hand, the distributed source approaches aim at identifying spatially extended source regions, which are characterized by a high number of dipoles (largely exceeding the number of sensors) with fixed locations. As the positions of the source dipoles are fixed, the lead-field matrix can be computed and is thus known.

In this paper, we concentrate on the solution of the inverse problem for the case where the lead-field matrix is known, and focus on the distributed source model. This inverse problem is one of the main topics in biomedical engineering [2], [39], [26], [54] and has been widely studied in the signal processing community, giving rise to an impressive number of methods. Our objective is to provide an overview of the currently available source imaging methods that takes into account recent advances in the field.

II. Data model and hypotheses

EEG and MEG are multi-channel systems that record brain activity over a certain time interval with a number of sensors covering a large part of the head. The two-dimensional measurements are stored into a data matrix $\mathbf{X} \in \mathbb{R}^{N \times T}$ where N denotes the number of EEG/MEG sensors and T the number of recorded time samples. The brain electric and magnetic fields are known to be generated by a number of current sources within the brain, which can be modeled by current dipoles [43]. In this paper, we assume that the latter correspond to the dipoles of a predefined source space, which can be derived from structural Magnetic Resonance Imaging (MRI). Furthermore, different hypotheses on the location and orientation of the sources can be incorporated by considering either a volume grid of source dipoles with free orientations or a surface grid of source dipoles with fixed orientations. Indeed, most of the activity recorded at the surface of the head is known to originate from pyramidal cells located in the gray matter and oriented perpendicular to the cortical surface [16].

Assuming a source space with *free orientation dipoles* and denoting $\mathbf{S} \in \mathbb{R}^{3D \times T}$ the signal matrix that contains the temporal activity with which each of the $3D$ dipole components of the D sources contributes to the signals of interest, the measurements at the surface constitute a linear combination of the source signals:

$$\mathbf{X} = \mathbf{GS} + \mathbf{N} = \mathbf{GS} + \mathbf{X}_i + \mathbf{X}_b \tag{1}$$

in the presence of noise $\mathbf{N}$. The noise is composed of two parts: instrumentation noise $\mathbf{X}_i$ introduced by the measurement system and background activity $\mathbf{X}_b = \mathbf{G}\mathbf{S}_b$, which originates from all dipoles of the source space that do not contribute to the signals of interest, but emit perturbing signals $\mathbf{S}_b \in \mathbb{R}^{3D\times T}$. The matrix $\mathbf{G} \in \mathbb{R}^{N\times 3D}$ is generally referred to as the *lead-field matrix* in the EEG/MEG context. For each dipole component of the source space, it characterizes the propagation of the source signal to the sensors at the surface.

In the case of *dipoles with fixed orientations*, the signal matrices $\mathbf{S}$ and $\mathbf{S}_b$ are replaced by the matrices $\tilde{\mathbf{S}} \in \mathbb{R}^{D\times T}$ and $\tilde{\mathbf{S}}_b \in \mathbb{R}^{D\times T}$, which characterize the brain activity of the D dipoles. Furthermore, the lead-field matrix $\mathbf{G}$ is replaced by the matrix $\tilde{\mathbf{G}} \in \mathbb{R}^{N\times D}$, which is given by $\tilde{\mathbf{G}} = \mathbf{G}\mathbf{\Theta}$ where $\mathbf{\Theta} \in \mathbb{R}^{3D\times D}$ contains the fixed orientations of the dipoles. The lead-field matrix $\mathbf{G}$ can be computed numerically based on Maxwell's equations. Several methods have been developed to this end and various software packages are available [23].

In this paper, we assume that the lead-field matrix is known and consider the EEG/MEG *inverse problem* that consists in estimating the unknown sources $\mathbf{S}$ or $\tilde{\mathbf{S}}$ (depending on the source model) from the measurements $\mathbf{X}$. As the number of source dipoles D (several thousands) is much higher than the number of sensors (several hundreds), the lead-field matrix is severely underdetermined, making the inverse problem ill-posed. In order to restore identifiability of the underdetermined source reconstruction problem, it is necessary to make assumptions on the sources. We discuss a large number of hypotheses that have been introduced in the context of the EEG/MEG inverse problem. In the following, we distinguish between three categories of assumptions depending on whether the hypotheses apply to the spatial, temporal, or spatio-temporal (deterministic or statistical) distribution of the sources. Subsequently, we give a short description of possible hypotheses.

A. Hypotheses on the spatial distribution of the sources

Sp1) Minimum energy: The power of the sources is physiologically limited. A popular approach thus consists in identifying the spatial distribution of minimum energy.

Sp2) Minimum energy in a transformed domain: Due to a certain synchronization of adjacent neuronal populations, the spatial distribution of the sources is unlikely to contain abrupt changes and can therefore be assumed to be smooth. This hypothesis is generally enforced by constraining the Laplacian of the source spatial distribution to be of minimum energy.

Sp3) Sparsity: In practice, it is often reasonable to assume that only a small fraction of the source dipoles contributes to the measured signals of interest in a significant way. For example, audio or visual

stimuli lead to characteristic brain signals in certain functional areas of the brain only. The signals of the other source dipoles are thus expected to be zero. This leads to the concept of sparsity.

Sp4) Sparsity in a transformed domain: If the number of active dipoles exceeds the number of sensors, which is generally the case for spatially extended sources, the source distribution is not sufficiently sparse for standard methods based on sparsity in the spatial domain to yield accurate results, leading to too focused source estimates. In this context, another idea consists in transforming the sources into a domain where their distribution is sparser than in the original source space and imposing sparsity in the transformed domain. The applied transform may be redundant, including a large number of basis functions or atoms, and is not necessarily invertible.

Sp5) Separability in space and wave-vector domains: For each distributed source, one can assume that the space-wave-vector matrix at each time point, which is obtained by computing a local spatial Fourier transform of the measurements, can be factorized into a function that depends on the space variable only and a function that depends on the wave-vector variable only. The space and wave-vector variables are thus said to be separable. In the context of brain source imaging, this is approximately the case for superficial sources.

Sp6) Gaussian joint probability density function with parameterized spatial covariance: For this prior, the source signals are assumed to be random variables that follow a Gaussian distribution with a spatial covariance matrix that can be described by a linear combination of a certain number of basis covariance functions. This combination is characterized by so-called hyperparameters, which have to be identified in the source imaging process.

B. Hypotheses on the temporal distribution of the sources

Te1) Smoothness: Since the autocorrelation function of the sources of interest usually has a full width at half maximum of several samples, the source time distribution should be smooth. This is, for example, the case for interictal epileptic signals or event-related potentials.

Te2) Sparsity in a transformed domain: Similar to hypothesis Sp4), this assumption implies that the source signals admit a sparse representation in a domain that is different from the original time domain. This can, for instance, be achieved by applying a wavelet transform or a redundant transformation such as the Gabor transform to the time dimension of the data. The transformed signals can then be modeled using a small number of basis functions or atoms, which are determined by the source imaging algorithm.

Te3) Pseudo-periodicity with variations in amplitude: If the recorded data comprise recurrent events such as a repeated time pattern that can be associated with the sources of interest, one can exploit

the repetitions as an additional diversity. This does not necessarily require periodic or quasi-periodic signals. Indeed, the intervals between the characteristic time patterns may differ, as may the amplitudes of different repetitions. Examples of signals with repeated time patterns include interictal epileptic spikes and Event-Related Potentials (ERP).

Te4) Separability in time and frequency domains: This hypothesis is the equivalent of hypothesis Sp5) and assumes that the time and frequency variables of data transformed into the time-frequency domain (for example by applying a Short Time Fourier Transform (STFT) or a Wavelet transform to the measurements) separate. This is approximately the case for oscillatory signals as encountered, for example, in epileptic brain activity.

Te5) Non-zero higher-order marginal cumulants: Regarding the measurements as realizations of an N-dimensional vector of random variables, this assumption is required when resorting to statistics of order higher than two, that offer a better performance and identifiability than approaches based on second order statistics. It is generally verified in practice, as the signals of interest usually do not follow a Gaussian distribution.

C. Hypotheses on the spatio-temporal distribution of the sources

SpTe) Synchronous dipoles: Contrary to point sources, which can be modeled by a single dipole, in practice, one is often confronted with so-called distributed sources. A distributed source is composed of a certain number of grid dipoles, which can be assumed to transmit synchronous signals. This hypothesis concerns both the spatial and the temporal distributions of the sources and is generally made in the context of dipoles with fixed orientations. In this case, it permits to separate the matrix $\tilde{\mathbf{S}}_{\mathcal{I}_r}$, which contains the signals of all synchronous dipoles of the r-th distributed source, indicated by the set $\mathcal{I}_r$, into the coefficient vector $\boldsymbol{\psi}_r$ that characterizes the amplitudes of the synchronous dipoles and thereby the spatial distribution of the r-th distributed source and the signal vector $\overline{\mathbf{s}}$ that contains the temporal distribution of the distributed source. This gives rise to a new data model:

$$\mathbf{X} = \mathbf{H}\overline{\mathbf{S}} + \mathbf{N} \tag{2}$$

where the matrix $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_R]$ contains the lead-field vectors for R distributed sources and the matrix $\overline{\mathbf{S}} \in \mathbb{R}^{R \times T}$ characterizes the associated distributed source signals. Each distributed source lead-field vector $\mathbf{h}_r$ corresponds to a linear combination of the lead-field vectors of all grid dipoles belonging to the distributed source: $\mathbf{h}_r = \tilde{\mathbf{G}}\boldsymbol{\psi}_r$. The distributed source lead-field vectors can be used as inputs for source imaging algorithms, simplifying the inverse problem by allowing for a separate localization of each source.

D. Hypotheses on the noise

While both the instrumentation noise and the background activity are often assumed to be Gaussian, the instrumentation noise can be further assumed to be spatially white, whereas the background activity is spatially correlated due to the fact that signals are mixed. To meet the assumption of spatially white Gaussian noise made by many algorithms, the data can be prewhitened based on an estimate of the noise covariance matrix $\mathbf{C}_n$. More precisely, the prewhitening matrix is computed as the inverse of the square root of the estimated noise covariance matrix. To achieve prewhitening, the data and the lead-field matrices are multiplied from the left by the prewhitening matrix.

III. Algorithms

In this section, we provide an overview of various source imaging methods that have been developed in the context of the EEG/MEG inverse problem. Based on methodological considerations, we distinguish four main families of techniques: regularized least squares approaches, tensor-based approaches, Bayesian approaches, and extended source scanning approaches. Each class of methods is associated with a certain number of hypotheses that are exploited by the algorithms. The knowledge of these hypotheses leads to a better understanding of the functioning of the source imaging techniques.

A. Regularized least squares methods

A natural approach to solve the ill-posed EEG/MEG inverse problem consists in finding the solution that best describes the measurements in a least squares sense. In the presence of noise, this is generally achieved by solving an optimization problem with a cost function of the form:

$$L(\mathbf{S}) = ||\mathbf{X} - \mathbf{G}\mathbf{S}||_{\mathrm{F}}^2 + \lambda f(\mathbf{S}). \tag{3}$$

For methods that do not consider the temporal structure of the data, but work on a time sample by sample basis, the data matrix $\mathbf{X}$ and the source matrix $\mathbf{S}$ are replaced by the column vectors $\mathbf{x}$ and $\mathbf{s}$, respectively.

The first term on the right-hand side of (3) is generally referred to as the data fit term and characterizes the difference between the measurements and the surface data reconstructed from given sources. The second is a regularization term and incorporates additional constraints on the sources according to the *a priori* information. The regularization parameter λ is used to manage a trade-off between data fit and a priori knowledge and depends on the noise level, since the gap between measured and reconstructed data is expected to become larger as the SNR decreases. Fig. 2 provides an overview of the regularized least squares algorithms with different regularization terms that are discussed in the following sections.

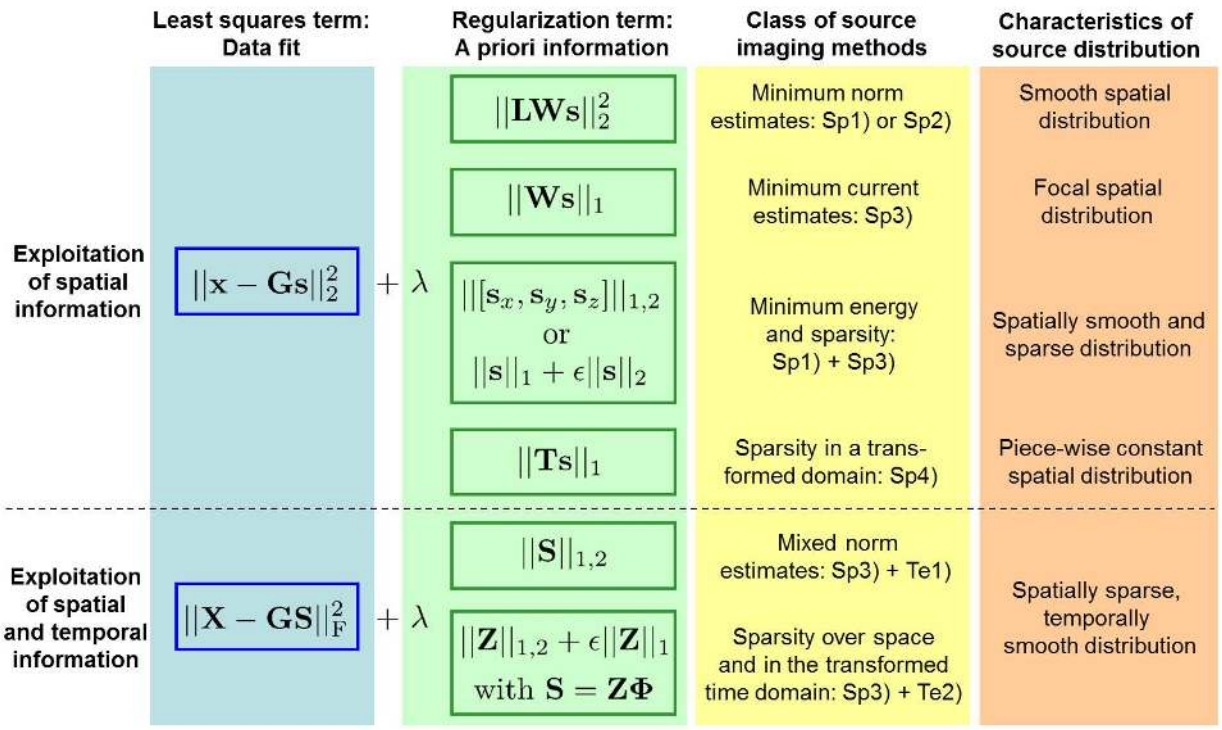

Fig. 2. Overview of regularized least squares algorithms (for an explanation of the employed notations for the different algorithms see the text in the associated sections).

1) Minimum Norm Estimates (MNE) – assumption Sp1) or Sp2) : The minimum norm solution is obtained by employing a prior, which imposes a minimal signal power according to hypothesis Sp1), leading to a regularization term that is based on the L_2-norm of the signal vector: $f(\mathbf{s}) = ||\mathbf{Ws}||_2^2$. To compensate for the depth bias, the diagonal matrix $\mathbf{W} \in \mathbb{R}_+^{3D\times 3D}$ containing fixed weights was introduced in the weighted MNE (WMNE) methods. Furthermore, one can consider the variance of the noise or the sources, leading to normalized estimates. This approach is pursued by the dSPM [15] algorithm, which takes into account the noise level, and sLORETA [45], which standardizes the source estimates with respect to the variance of the sources.

The MNEs generally yield smooth source distributions. Nevertheless, spatial smoothness can also be more explicitly promoted by applying a Laplacian operator $\mathbf{L}$ to the source vector in the regularization term, leading to the popular LORETA method [46], which is based on assumption Sp2). In this case, the L_2-norm constraint is imposed on the transformed signals, yielding a regularization term of the form $f(\mathbf{s}) = ||\mathbf{LWs}||_2^2$. More generally, the matrix $\mathbf{L}$ can be used to implement a linear operator that is applied to the sources.

The original MNEs have been developed for sources with free orientations. Modifications of the algorithms to account for orientation constraints can, for example, be found in [53], [34].

2) Methods based on sparsity – assumption Sp3) or Sp4): As the minimum norm estimates generally lead to blurred source localization results, as widely described in the literature (see for example [56]), source imaging methods based on hypothesis Sp3), which promote sparsity, were developed to obtain more focused source estimates. One of the first algorithms proposed in this field was FOCUSS [22], which iteratively updates the minimum norm solution using an L_0 "norm". This gradually shrinks the source spatial distribution, resulting in a sparse solution. Around the same time, source imaging techniques based on an L_p-norm $(0 \leq p \leq 1)$ regularization term of the form $f(\mathbf{s}) = ||\mathbf{W}\mathbf{s}||_p$, where $\mathbf{W}$ is a diagonal matrix of weights, were put forward [36]. The parameter p is generally chosen to be equal to 1, leading to a convex optimization problem[1]. However, by treating the dipole components independently in the regularization term, the estimated source orientations are biased. To overcome this problem, Uutela et al. [50] proposed to use fixed orientations determined either from the surface normals or estimated using a preliminary minimum norm solution. This gave rise to the MCE algorithm. Extensions of this approach, which require only the knowledge of the signs of the dipole components, or which permit to incorporate loose orientation constraints, have been treated in [29], [34]. Another solution to the problem of orientation bias of the sparse source estimates consists in imposing sparsity dipole-wise instead of component-wise [20]. In [56], a combination of the ideas of FOCUSS and L_p-norm $(p \leq 1)$ regularization was implemented in an iterative scheme.

To find a compromise between smoothness and sparsity of the spatial distribution, the use of a prior that is composed of both an L_1-norm and an L_2-norm regularization term was proposed in [52].

Another idea consists in imposing sparsity in a transformed domain. This is generally achieved by employing a regularization term of the form $||\mathbf{T}\tilde{\mathbf{s}}||_1$ where $\mathbf{T}$ is a transformation matrix. In the literature, different transformations have been considered. The authors of [10] have used a surface Laplacian, thus imposing sparsity on the second order spatial derivatives of the source distribution, in combination with classical L_1-norm regularization. Another way to promote a piece-wise constant spatial distribution was proposed by Ding, giving rise to the VB-SCCD method [19], which is closely related to the Total Variation (TV) approach. A third approach that makes use of sparsity in a transformed domain considers a spatial wavelet transform that permits to compress the signals through a sparse representation of the sources in

[1]Note that the minimization of this cost function is closely related to the optimization problem $\min ||\mathbf{W}\mathbf{s}||_p \quad \text{s. t.} \quad ||\mathbf{x} - \mathbf{G}\mathbf{s}||_2^2 \leq \delta$ with regularization parameter δ, on which the algorithm proposed in [36] is based.

the wavelet domain [31], [10].

3) Mixed norm estimates – assumption Sp3) or Sp4) and assumption Te1) or Te2): To impose hypotheses simultaneously in several domains, *e.g.*, the space-time plane, one can resort to mixed norms. Efficient algorithms that have been developed to deal with the resulting optimization problem are presented in [24]. In [44], a source imaging method, called MxNE, that imposes sparsity over space (hypothesis Sp3)) and smoothness over time (assumption Te1)) using a mixed $L_{1,2}$-norm regularization has been proposed.

An approach that imposes sparsity over space (hypothesis Sp3)) as well as in the transformed time domain (assumption Te2)) is taken in the TF-MxNE method. This technique makes use of a dictionary, $\mathbf{\Phi}$, from which a small number of temporal basis functions are selected to characterize the source signals. In [25], Gabor basis functions were considered, whereas the authors of [49] employed a data-dependent temporal basis obtained using an SVD of the measurements and a data-independent temporal basis that is given by Natural Cubic Splines (NCS). The method is based on mixed norms and uses a composite prior of two regularization terms similar to [52].

Furthermore, in [28], one can find an approach that imposes sparsity in a spatial transform domain similar to [10], but which is based on a mixed $L_{1,2}$-norm to take into account temporal smoothness of the source distribution. Finally, let us point out that it is also possible to consider both temporal and spatial basis functions (assumptions Sp4) and Te2)) as suggested in [7] for the ESP algorithm.

B. Tensor-based source localization – assumption SpTe), assumption Sp5), Te3), or Te4), assumptions Sp4) and Sp3)

The objective of tensor-based methods consists in identifying the lead-field vectors and the signals of distributed sources, *i.e.*, matrices $\mathbf{H}$ and $\overline{\mathbf{S}}$ in data model (2), from measurements. To separate R simultaneously active distributed sources, tensor-based methods exploit multi-dimensional data (at least one dimension in addition to space and time) and assume a certain structure underlying the measurements. The multi-dimensional data is then approximated by a model that reflects the assumed structure and comprises a number of components which can be associated with the sources. A popular tensor model is the rank-R Canonical Polyadic (CP) decomposition [14], which imposes a multilinear structure on the data. This means that each element of a third order tensor $\boldsymbol{X}$ can be written as a sum of R components, each being a product of three univariate functions, a_r, b_r, d_r:

$$\boldsymbol{X}(\alpha_k, \beta_\ell, \gamma_m) = \sum_{r=1}^{R} a_r(\alpha_k) b_r(\beta_\ell) d_r(\gamma_m). \tag{4}$$

The samples of functions a_r, b_r, d_r can be stored into three loading matrices $\mathbf{A} \in \mathbb{C}^{K \times R} = [\mathbf{a}_1, \ldots, \mathbf{a}_R]$, $\mathbf{B} \in \mathbb{C}^{L \times R} = [\mathbf{b}_1, \ldots, \mathbf{b}_R]$, and $\mathbf{D} \in \mathbb{C}^{M \times R} = [\mathbf{d}_1, \ldots, \mathbf{d}_R]$ that characterize the tensor $\boldsymbol{X} \in \mathbb{C}^{K \times L \times M}$.

In the literature, a certain number of tensor methods based on the CP decomposition have been proposed in the context of EEG/MEG data analysis. These methods differ in the dimension(s) which is (are) exploited in addition to space and time. In this paper, we focus on third order tensors. Here, a first distinction can be made between approaches that collect an additional diversity directly from the measurements, for instance, by taking different realizations of a repetitive event (see [40]), or methods that create a third dimension by applying a transform which preserves the two original dimensions, such as the STFT or wavelet transform. This transform can be applied either over time or over space, leading to Space-Time-Frequency (STF) data (see, *e.g.*, [17] and references therein) or Space-Time-Wave-Vector (STWV) data [5]. Depending on the dimensions of the tensor, the CP decomposition involves different multilinearity assumptions: for Space-Time-Realization (STR) data, hypothesis Te3) is required, for STF data, hypothesis Te4) is involved, and for STWV data, we resort to hypothesis Sp5).

Once several simultaneously active distributed sources have been separated, using the tensor decomposition, and estimates for the distributed source lead-field vectors have been derived, the latter can be used for source localization. The source localization is then performed separately for each distributed source. To this end, a dictionary of potential elementary distributed sources is defined by a number of circular-shaped cortical areas of different centers and sizes, subsequently called disks. Each disk describes a source region with constant amplitudes, leading to a sparse, piecewise constant source distribution, which can be attributed to hypotheses Sp3) and Sp4). For each source, a small number of disks that correspond best to the estimated distributed source lead-field vector are then identified based on a metric and are merged to reconstruct the distributed source. The steps of the algorithm based on STWV data and referred to as STWV-DA [5] are schematically summarized in Fig. 3.

C. Bayesian approaches – assumption Sp6)

Bayesian approaches are based on a probabilistic model of the data and treat the measurements, the sources, and the noise as realizations of random variables. In this context, the reconstruction of the sources corresponds to obtaining an estimate of their posterior distribution, which is given by:

$$p(\mathbf{s}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{x})} \tag{5}$$

where $p(\mathbf{x}|\mathbf{s})$ is the likelihood of the data, $p(\mathbf{s})$ is the source distribution, and $p(\mathbf{x})$ is the model evidence. The crucial point consists in finding an appropriate prior distribution $p(\mathbf{s})$ for the sources, which in the

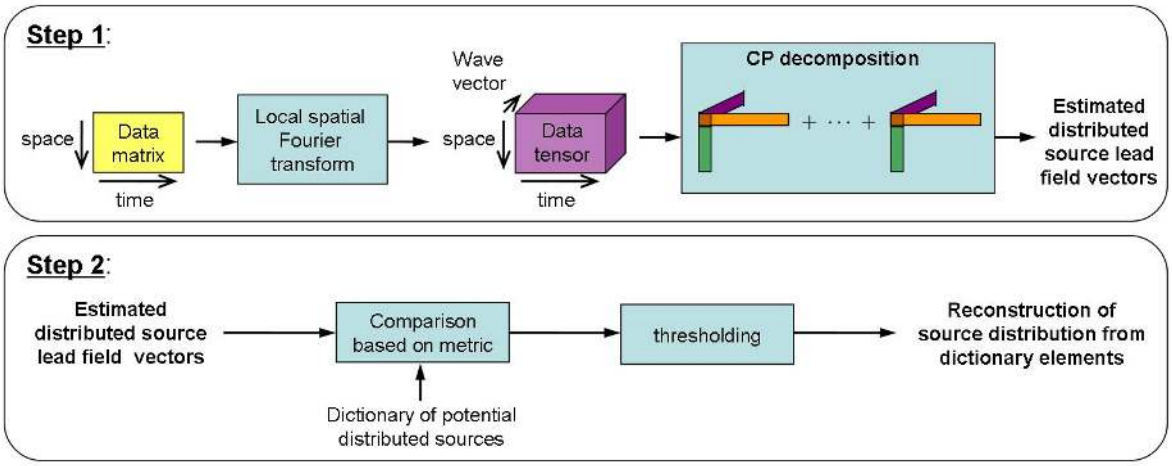

Fig. 3. Schematic representation of the STWV-DA algorithm.

Bayesian framework incorporates the hypotheses that regularize the ill-posed inverse problem. We can distinguish three classes of Bayesian approaches [54]: Maximum A Posteriori (MAP) estimation for the sources, variational Bayes, and empirical Bayes. The first approach employs a fixed prior $p(\mathbf{s})$ leading to MNE, MCE, and MxNE solutions, which have been addressed in previous sections. In this section, we focus on variational and empirical Bayesian approaches, which use a flexible, parameterized prior $p(\mathbf{s}|\boldsymbol{\gamma})$, which is modulated by the hyper-parameter vector $\boldsymbol{\gamma} \in \mathbb{R}^L$. More particularly, in the EEG/MEG context, the source distribution is generally assumed to be zero-mean Gaussian with a covariance matrix $\mathbf{C}_s$ that depends on hyper-parameters, such that:

$$p(\mathbf{s}|\boldsymbol{\gamma}) \propto \exp(-\frac{1}{2}\mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1}(\boldsymbol{\gamma})\mathbf{S}). \tag{6}$$

The hyper-parameters can either directly correspond to the elements of $\mathbf{C}_s$ (as in the Champagne algorithm [55]) or parameterize the covariance matrix such that $\mathbf{C}_s = \sum_{i=1}^{I} \gamma_i \mathbf{C}_i$. Here, $\mathbf{C}_i$, $i = 1, \ldots, I$, are predefined covariance components. The hyper-parameters are then learned from the data to perform some kind of model selection by choosing appropriate components.

1) Variational Bayesian approaches : The variational Bayesian methods (see [21] and references therein) try to obtain estimates of the posterior distributions of the hyper-parameters $\hat{p}(\boldsymbol{\gamma}|\mathbf{x})$. To this end, additional assumptions are required, such as (i) statistical independence of the hyper-parameters (a.k.a. mean-field approximation), or (ii) a Gaussian posterior distribution of the hyper-parameters (a.k.a. Laplace approximation). This permits not only to approximate the distribution $p(\mathbf{s}|\mathbf{x})$ and thereby to estimate the sources, but also to provide an estimate of the model evidence $p(\mathbf{x})$, which can be used to compare different models (*e.g.*, for different sets of covariance components).

2) Empirical Bayesian approaches : The empirical Bayesian approaches (see, e.g., [37], [55] and references therein) on the other hand are concerned with finding a point estimate of the hyper-parameters, which is obtained by marginalization over the unknown sources $\mathbf{s}$:

$$\hat{\gamma} = \arg\max_{\boldsymbol{\gamma}} \int p(\mathbf{x}|\mathbf{s})p(\mathbf{s}|\boldsymbol{\gamma})p(\boldsymbol{\gamma})d\mathbf{s}. \tag{7}$$

For known hyper-parameters, the conditional distribution $p(\mathbf{s}|\mathbf{x}, \boldsymbol{\gamma})$ can be determined. To obtain a suitable estimate of the sources, one can for instance apply the Expectation Maximization (EM) algorithm [18], which alternates between two steps: the M-step in which the maximum likelihood estimates of the hyper-parameters are updated for fixed $\mathbf{s}$, and the E-step in which the conditional expectation of the sources is determined based on the hyper-parameters obtained in the M-step. An example for an empirical Bayesian algorithm is the Champagne algorithm introduced in [55].

D. Extended source scanning methods

Here, the idea is to identify active sources from a dictionary of potential distributed sources. To this end, a metric is computed for each element of the dictionary. The source estimates are then obtained from the elementary source distributions that are associated with the maxima of the metric. Based on the employed metric, we subsequently distinguish two types of scanning methods that correspond to spatial filtering, a.k.a. beamforming, and subspace-based approaches.

1) Beamforming approaches – assumptions Sp3) and Sp4): Beamforming techniques have originally been proposed in the context of equivalent current dipole localization from MEG measurements [51]. The basic approach employs the Linearly Constrained Minimum Variance (LCMV) filter, that is based on the data covariance matrix and which is derived for each dipole of the source space to reconstruct its temporal activity while suppressing contributions from other sources. The filter output is then used to compute a metric which serves to identify the active dipole sources. The LCMV beamformer was shown to yield unbiased solutions in the case of a single dipole source [48], but leads to source localization errors in the presence of correlated sources. To overcome this problem, extensions of the beamforming approach to multiple, potentially correlated (dipole) sources have been considered (see [41] and references therein). Furthermore, in [33], the beamforming approach has been extended to the localization of distributed sources. This is achieved by deriving spatial filters for all elements of a dictionary of potential source regions, also called patches. The source imaging solution is then obtained from dictionary elements associated with the maxima of the metric which is derived from the filter outputs, resulting in a spatially

sparse source distribution with a small number of active source regions according to hypotheses Sp3) and Sp4).

2) Subspace-based approaches – assumptions SpTe), Te5), Sp3) and Sp4): Similar to Bayesian approaches, subspace-based methods also treat the measurements made by several sensors as realizations of a random vector. They then exploit the symmetric $2q$-th $(q \geq 1)$ order cumulant matrix $\mathbf{C}_{2q,x}$ of this random vector from which the signal and noise subspaces are identified by means of an EigenValue Decomposition (EVD). For source imaging purposes, one then exploits the fact that the higher-order lead-field vector $\tilde{\mathbf{g}}_r^{\otimes q}$, $r = 1, \ldots, R$, where $\tilde{\mathbf{g}}_r^{\otimes q}$ is a shorthand notation for $\tilde{\mathbf{g}}_r \otimes \tilde{\mathbf{g}}_r \otimes \cdots \otimes \tilde{\mathbf{g}}_r$ with $q-1$ Kronecker products (denoted by $\otimes$), must lie in the $2q$-th order signal subspace and be orthogonal to the noise subspace. To this end, MUSIC-like algorithms can be employed, which have first been used in the context of equivalent current dipole localization [42], [1]. Recently, the $2q$-MUSIC algorithm [12] has been adapted to the identification of distributed sources [6], then referred to as $2q$-ExSo-MUSIC. In analogy to the classical MUSIC algorithm, the $2q$-ExSo-MUSIC spectrum is computed for a number of predefined parameter vectors $\boldsymbol{\psi}$. To this end, one defines a dictionary of disks as described in Section III-B, assuming a sparse, piece-wise constant source distribution (corresponding to hypotheses Sp3) and Sp4)) similar to VB-SCCD and STWV-DA. The spectrum is then thresholded and all coefficient vectors $\boldsymbol{\psi}$ for which the spectrum exceeds a fixed threshold are retained and united to model distributed sources. An advantage of subspace-based techniques exploiting the $2q$-th order statistics with $q > 1$ over other source imaging algorithms lies in their asymptotic robustness to Gaussian noise, because cumulants of order higher than 2 of a Gaussian random variable are null.

IV. Discussion

In this section, we discuss several aspects of the brain source imaging methods described in the previous section, including identifiability and convergence issues, advantages and drawbacks of representative algorithms, and application domains. Table I lists several source imaging methods mentioned in the previous section and summarizes the exploited hypotheses.

A. Identifiability

For methods that solve the inverse problem by exploiting sparsity, the uniqueness of the solution depends on the conditioning of the lead-field matrix. More particularly, sufficient conditions that are based on the mutual or cumulative coherence of the lead-field matrix are available in the literature [11] and can easily be verified for a given lead-field matrix. However, in brain source imaging, these

TABLE I

CLASSIFICATION OF DIFFERENT ALGORITHMS MENTIONED IN SECTION III ACCORDING TO THE EXPLOITED HYPOTHESES.

BRAIN SOURCE IMAGING	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	Te1	Te2	Te3	Te4	Te5	SpTe
Regularized least squares algorithms												
sLORETA [45]	X											
LORETA [46]		X										
MCE [50]			X									
VB-SCCD [19]				X								
MxNE [44]			X				X					
TF-MxNE [25]			X					X				
Bayesian approaches												
Champagne [55]						X						
Extended source scanning methods												
$2q$-ExSo-MUSIC [6]				X							X	X
Tensor-based methods												
STR-DA [5]				X					X			X
STF-DA [5]				X						X		X
STWV-DA [5]				X	X							X

Hypotheses on the spatial distribution

Sp1) Minimum energy

Sp2) Minimum energy in a transformed domain

Sp3) Sparsity

Sp4) Sparsity in a transformed domain

Sp5) Separability in the space-wave-vector domain

Sp6) Parameterized spatial covariance

Hypotheses on the temporal distribution

Te1) Smoothness

Te2) Sparsity in a transformed domain

Te3) Pseudo-periodicity

Te4) Separability in the time-frequency domain

Te5) Non-zero higher order marginal cumulants

Hypotheses on the spatio-temporal distribution

SpTe) Synchronous dipoles

conditions are generally not fulfilled because the lead-field vectors of adjacent grid dipoles are often highly correlated, making the lead-field matrix ill-conditioned.

A strong motivation for the use of tensor-based methods is the fact that the CP decomposition is essentially unique under mild conditions on the tensor rank [32]. These conditions are generally verified in brain source imaging because the rank R of the noiseless tensor corresponds to the number of distributed sources, which is usually small (less than 10) compared to the tensor dimensions. The limitations of the tensor-based approach thus arise from the approximations that are made when imposing a certain structure

on the data and not from the identifiability conditions. Note though, that these identifiability conditions only concern the CP decomposition, which separates the distributed sources. Additional conditions are indeed required for the uniqueness of the results of the subsequent source localization step that is applied for each distributed source separately. Nevertheless, the separation of the distributed sources facilitates their identification and may alleviate the identifiability conditions for the source localization step.

Finally, for subspace-based approaches, the number of sources that can be identified depends on the dimensions of the signal and noise subspaces of the cumulant matrix. In the best case, one can identify at most $N_{2q} - 1$ statistically independent distributed sources, where $N_{2q} \leq N^q$ denotes the maximal rank that can be attained by the $2q$-th order distributed source lead-field matrix and N is the number of sensors, while in the worst case, when all distributed sources are correlated, one can identify up to $N - 1$ sources. In the context of brain source imaging, these identifiability conditions are usually not very restrictive.

B. Convergence

The source imaging methods exploiting sparsity may be implemented using two types of convex optimization algorithms: interior point methods such as SOCP [9] and proximal splitting methods like FISTA [3] or ADMM [8]. Both types of solvers are known to converge to the global solution of a convex optimization problem. However, the interior point methods are computationally too expensive to solve large-scale problems as encountered in brain source imaging and the simpler and more efficient proximal splitting methods are to be preferred in this case.

To solve the optimization problem associated with the CP decomposition, a wide panel of algorithms, including alternating methods like Alternating Least Squares (ALS), derivative-based techniques such as Gradient Descent (GD), or Levenberg-Marquardt (LM) [14], and direct techniques (see, e.g., [47], [35] and references therein), have been used. Even if local convergence properties hold for most of these methods, there is no guarantee that they will converge to the global minimum because the cost function generally features a large number of local minima. However, in practical situations, it has been observed [30] that good results can, for example, be achieved by combining a direct method such as the DIAG algorithm described in [35] with a derivative-based technique like GD.

Similar to the tensor decomposition algorithm, there is no guarantee of global convergence for the EM algorithm, which is popular in empirical Bayesian approaches, or for the alternating optimization method employed by the Champagne algorithm.

C. Advantages and drawbacks

Since strengths and weaknesses are often specific to a given source imaging method and cannot be generalized to other techniques of the same family of approaches, we subsequently focus on seven representative algorithms. Table II lists the advantages and drawbacks of each of these methods. On the one hand, the regularized least squares techniques sLORETA, MCE, and MxNE are simple and computationally efficient, but the source estimates obtained by these algorithms tend to be very focal (for MCE and MxNE) or blurred (for sLORETA). On the other hand, VB-SCCD, STWV-DA, and 4-ExSo-MUSIC, which permit to identify spatially extended sources, feature a higher computational complexity. Furthermore, STWV-DA and 4-ExSo-MUSIC have additional requirements such as knowledge of the number of sources or the signal subspace dimension, a certain structure of the data (for STWV-DA) or a sufficiently high number of time samples (for 4-ExSo-MUSIC). While all these methods require adjusting certain parameters, which are tedious to tune in practice, the main advantage of the Champagne algorithm consists in the fact that there is no parameter to adjust. However, this method also has a high computational complexity and leads to very sparse source estimates.

D. Application domains

Brain source imaging finds application both in the clinical domain and in cognitive neuroscience. The most frequent clinical application is in epilepsy, where the objective consists in delineating the regions from where interictal spikes or ictal discharges arise [38]. To this end, brain source imaging methods such as VB-SCCD, STWV-DA, or 4-ExSo-MUSIC, which can identify both the spatial extent and the shape of a small number of distributed sources, are well suited. In cognitive neuroscience, multiple brain structures are often simultaneously activated, particularly when the subjects are asked to perform complex cognitive tasks during the experimental sessions [2]. The source imaging methods employed for the analysis of these data should thus be able to deal with multiple correlated sources. This is, for example, the case for VB-SCCD and other regularized least squares techniques, but not for STWV-DA or 4-ExSo-MUSIC. On the other hand, during simple tasks like those related to perceptual processes, the analysis of EEG signals of ERPs can also aim at identifying focal sources, in which case methods such as MCE, MxNE, or Champagne are to be preferred. Finally, there is a rising interest in the analysis of source connectivity [27]. While sLORETA, MCE, MxNE, or Champagne can be employed for this purpose, VB-SCCD, STWV-DA, and 4-ExSo-MUSIC, which enforce identical signals for dipoles belonging to the same patch, would theoretically be less suited, especially for the analysis of very local cortical networks. Nevertheless, at a

TABLE II

ADVANTAGES AND DRAWBACKS OF SOURCE IMAGING ALGORITHMS

algorithm	advantages	disadvantages
sLORETA [45]	• simple to implement • computationally efficient • no localization error for a single dipole source in the absence of noise • works on a single time sample	• blurred results • assumes independent dipole sources
MCE [50]	• simple • can localize correlated sources • works on a single time sample • low computational cost for small numbers of time samples	• very focal source estimates
VB-SCCD [19]	• identifies spatially extended sources • flexible with respect to the patch shape • permits to localize multiple simultaneously active (and correlated) patches • works on a single time sample	• overestimates size of small patches • computationally expensive • systematic error on estimated amplitudes
MxNE [44]	• exploits the temporal structure of the data • extracts smooth time signals • small computational cost	• very focal source estimates
Champagne [55]	• no parameter to adjust manually • easy to implement • permits perfect source reconstruction under certain conditions • works on a single time sample	• very sparse source estimates • assumes independent dipole signals • high computational complexity
STWV-DA [5]	• separates (correlated) sources • identifies extended sources • does not require spatial prewhitening to yield accurate results	• makes strong assumptions on data structure that are difficult to verify in practice • requires knowledge of the number of sources to separate • computationally expensive for long data lengths
4-ExSo-MUSIC [6]	• identifies extended sources • robust to Gaussian noise	• high computational complexity • requires knowledge of the signal subspace dimension • requires a sufficiently large number of time samples (> 500) to estimate the data statistics • difficulties in localizing highly correlated sources

macroscopic level, these algorithms may be employed to identify cortical networks that characterize the connectivity between distinct brain regions.

V. Results

To give the reader an idea of the kind of source imaging results that can be obtained with different types of algorithms, in this section, we illustrate and compare the performance of seven representative algorithms on simulated data for an example of epileptic EEG activity. To this end, we consider two or three quasi-simultaneous active patches and model epileptiform spike-like signals that spread from one brain region to another. The sources are localized using the sLORETA, MCE, MxNE, VB-SCCD, STWV-DA, Champagne, and 4-ExSo-MUSIC algorithms. To quantitatively evaluate the performance of the different methods, we use a measure called the Distance of Localization Error (DLE) [13], which characterizes the difference between the original and the estimated source configuration. The DLE is averaged over 50 realizations of EEG data with different epileptiform signals and background activity. For detailed descriptions of the data generation process, the implementation of the source imaging methods, and the evaluation criterion, the reader is referred to [4].

The CPU runtimes that are required for the application of the different source imaging methods, implemented in Matlab and run on a machine with a 2.7 GHz processor and 8 GB of RAM, are listed in Table III. Note that the runtime of 4-ExSo-MUSIC cannot be compared to that of the other algorithms because this method is partly implemented in C.

TABLE III

Average CPU runtime of the different source imaging algorithms for the considered three-patch scenarios.

	sLORETA	VB-SCCD	MxNE	MCE	Champagne	STWV-DA	4-ExSo-MUSIC
CPU runtime in s	0.18	120	5.9	2.2	233	156	58

We first consider two scenarios with two patches of medium distance composed of a patch in the inferior frontal region (InfFr) combined once with a patch in the inferior parietal region (InfPa) and once with a patch in the middle posterior temporal gyrus (MidTe). The patches are all located on the lateral aspect of the left hemisphere, but the patch MidTe is partly located in a sulcus, leading to weaker surface signals than the patches InfFr and InfPa, which are mostly on a gyral convexity. This has an immediate influence on the performance of all source imaging algorithms except for Champagne. For the first scenario, the algorithms exhibit high dipole amplitudes for dipoles belonging to each of the

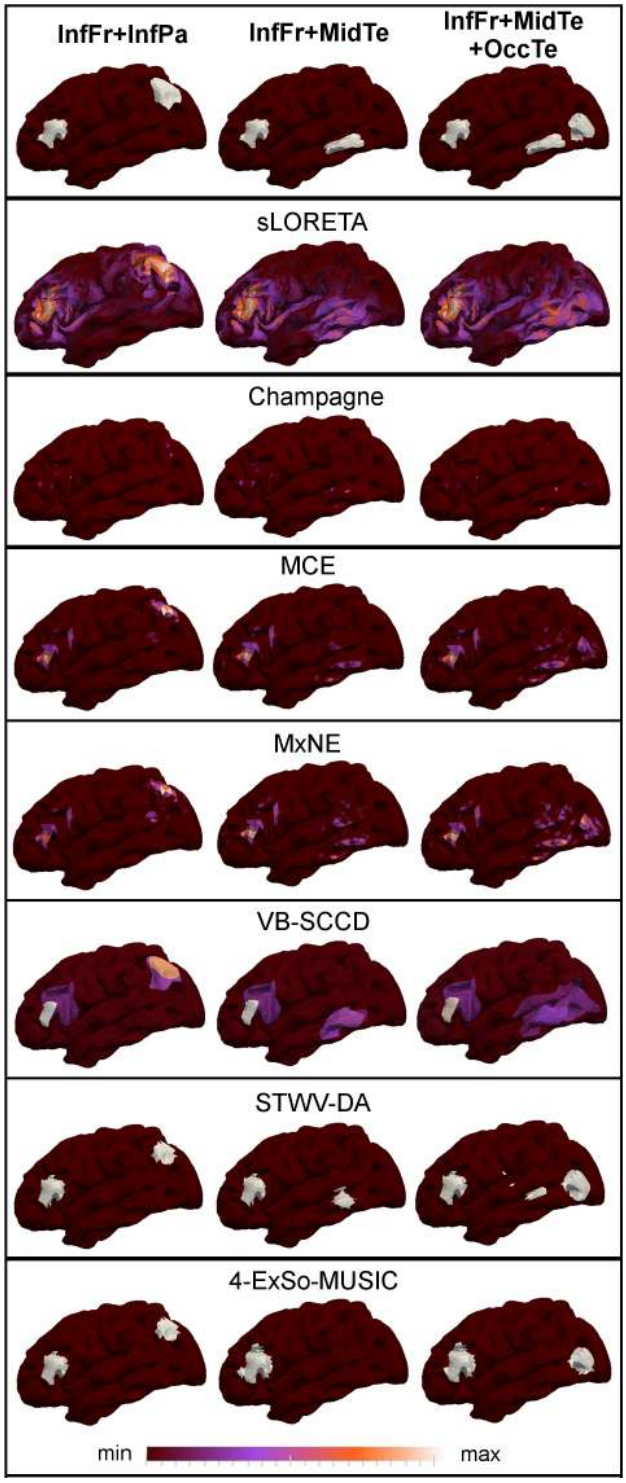

Fig. 4. Original patches and source reconstructions of different source imaging algorithms for the scenarios InfFr+InfPa, InfFr+MidTe, and InfFr+MidTe+OccTe.

true patches. For the second scenario on the other hand, the weak patch is difficult to make out on the estimated source distributions of sLORETA, slightly better visible on the MCE and MxNE solutions, but completely missing for 4-ExSo-MUSIC. VB-SCCD and STWV-DA both recover the second patch, but with smaller amplitude in case of VB-SCCD and smaller size for STWV-DA. According to the DLE, MCE leads to the best results among the focal source imaging algorithms while STWV-DA outperforms the other distributed source localization methods.

In the third scenario, we add a patch at the temporo-occipital function (OccTe) to the InfFr and MidTe patches, which further complicates the correct recovery of the active grid dipoles. The best result in terms of DLE (see Fig. 7 and lower part of Table IV) is achieved by VB-SCCD. Even though this method mostly identifies the brain regions that correspond to the active patches, it does not permit to distinguish the patches MidTe and OccTe into two separate active sources. STWV-DA, on the other hand, identifies all three patches, even though the extent of the estimated active source region that can be associated to the patch MidTe is too small. Albeit, this method also identifies several spurious source regions of small size located between the patches MidTe and InfFr. 4-ExSo-MUSIC and Champagne recover only one of the two patches located in the temporal lobe. Similar to VB-SCCD, sLORETA does not permit to distinguish patches MidTe and OccTe. This distinction is better performed by MCE and especially by MxNE, which displays three foci of brain activity.

TABLE IV

DLE (IN CM) OF SOURCE IMAGING ALGORITHMS FOR DIFFERENT SCENARIOS

scenario	sLORETA	Champagne	MCE	MxNE	VB-SCCD	STWV-DA	ExSo-MUSIC
InfFr+InfPa	2.97	4.03	3.51	3.52	1.23	0.59	0.61
InfFr+MidTe	6.13	4.34	4.40	4.50	1.51	1.17	14.90
InfFr+MidTe+OccTe	5.88	4.83	4.59	4.51	2.54	5.99	4.30

VI. CONCLUSIONS AND PERSPECTIVES

In this paper, we classified existing source imaging algorithms based on methodological considerations. Furthermore, we discussed the different techniques, both under theoretical and practical considerations, by addressing questions of identifiability and convergence, advantages and drawbacks of certain algorithms as well as application domains, and by illustrating the performance of representative source imaging algorithms through a simulation study.

While uniqueness conditions are available for both tensor-based and sparsity-based techniques, in the context of brain source imaging, these conditions are generally only fulfilled for tensor-based approaches, which exploit the concept of distributed sources, whereas the bad conditioning of the lead-field matrix practically prohibits the unique identification of a sparse source distribution. On the other hand, while convex optimization algorithms used for sparse approaches usually converge to the global minimum, such algorithms are not available for tensor decompositions, which suffer from multiple local minima, making it almost impossible to find the global optimum. *In practice, despite the limitations concerning identifiability and convergence, both tensor-based and sparse approaches often yield good source reconstruction.*

Since the various source localization algorithms have different advantages, drawbacks and requirements, source imaging solutions may vary depending on the application. As discussed in this paper, for each problem, an appropriate source imaging technique has to be chosen depending on the desired properties of the solution, on the characteristics of the algorithm, and on the validity of the hypotheses employed by the method. Furthermore, it is advisable to compare the results of different methods for confirmation of the identified source region(s).

To summarize the findings of the simulation study, we can say that *sLORETA, Champagne, MCE, and MxNE recover well the source positions, though not their spatial extent as they are conceived for focal sources, while ExSo-MUSIC, STWV-DA, and VB-SCCD also permit to obtain an accurate estimate of the source size.* We noticed that most of the methods except for ExSo-MUSIC and STWV-DA require pre-whitening of the data or a good estimate of the noise covariance matrix (in case of Champagne) in order to yield accurate results. On the one hand, this can be explained by the hypothesis of spatially white Gaussian noise made by some approaches, while on the other hand, the prewhitening also leads to a decorrelation of the lead-field vectors and therefore to a better conditioning of the lead-field matrix, which consequently facilitates the correct identification of active grid dipoles. Furthermore, the source imaging algorithms generally have some difficulties in identifying mesial sources, located close to the midline, as well as multiple quasi-simultaneously active sources. On the whole, for the situations addressed in our simulation study, *STWV-DA seems to be the most promising algorithm for distributed source localization, both in terms of robustness and source reconstruction quality.* However, more detailed studies are required to confirm the observed performances of the tested algorithms before drawing further conclusions.

Based on these results, we can identify several promising directions for future research. As the VB-SCCD algorithm demonstrates, imposing sparsity in a suitable spatial transform domain may work better than applying sparsity constraints directly to the signal matrix. This type of approach should thus be further developed. Another track for future research consists in further exploring different combinations

of a priori information, for example by merging successful strategies of different recently established source imaging approaches, such as tensor-based or subspace-based approaches and sparsity. In a similar way, one could integrate the steps of two-step procedures such as STWV-DA into one single step in order to process all the available information and constraints at the same time.

ACKNOWLEDGMENTS

H. Becker was supported by Conseil Régional PACA and by CNRS France. The work of P. Comon was funded by the FP7 European Research Council Programme, DECODA project, under grant ERC-AdG-2013-320594. The work of R. Gribonval was funded by the FP7 European Research Council Programme, PLEASE project, under grant ERC-StG-2011-277906. Furthermore, we acknowledge the support of Programme ANR 2010 BLAN 0309 01 (project MULTIMODEL).

REFERENCES

[1] L. Albera, A. Ferréol, D. Cosandier-Rimélé, I. Merlet, and F. Wendling, "Brain source localization using a fourth-order deflation scheme," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 490–501, 2008.

[2] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, Nov. 2001.

[3] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal of Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[4] H. Becker, "Denoising, separation and localization of EEG sources in the context of epilepsy," Ph.D. dissertation, University of Nice-Sophia Antipolis, 2014.

[5] H. Becker, L. Albera, P. Comon, M. Haardt, G. Birot, F. Wendling, M. Gavaret, C. G. Bénar, and I. Merlet, "EEG extended source localization: tensor-based vs. conventional methods," *NeuroImage*, vol. 96, pp. 143–157, Aug. 2014.

[6] G. Birot, L. Albera, F. Wendling, and I. Merlet, "Localisation of extended brain sources from EEG/MEG: the ExSo-MUSIC approach," *NeuroImage*, vol. 56, pp. 102 – 113, 2011.

[7] A. Bolstad, B. Van Veen, and R. Nowak, "Space-time event sparse penalization for magneto-/electroencephalography," *NeuroImage*, vol. 46, pp. 1066 – 1081, 2009.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1 – 122, 2010.

[9] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[10] W. Chang, a. Nummenmaa, J. Hsieh, and F. Lin, "Spatially sparse source cluster modeling by compressive neuromagnetic tomography," *NeuroImage*, vol. 53, no. 1, pp. 146–160, Oct. 2010.

[11] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4634 – 4643, 2006.

[12] P. Chevalier, A. Ferréol, and L. Albera, "High-resolution direction finding from higher order statistics: the 2q-MUSIC algorithm," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 2986–2997, 2006.

[13] J.-H. Cho, S. B. Hong, Y.-J. Jung, H.-C. Kang, H. D. Kim, M. Suh, K.-Y. Jung, and C.-H. Im, "Evaluation of algorithms for intracranial EEG (iEEG) source imaging of extended sources: feasibility of using (iEEG) source imaging for localizing epileptogenic zones in secondary generalized epilepsy," *Brain Topography*, no. 24, pp. 91–104, 2011.

[14] P. Comon, L. Luciani, and A. L. F. D. Almeida, "Tensor decompositions, alternating least squares and other tales," *Journal of Chemometrics*, vol. 23, pp. 393–405, 2009.

[15] A. M. Dale, A. K. Liu, B. R. Fischl, R. L. Buckner, J. W. Belliveau, J. D. Lewine, and E. Halgren, "Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity," *Neuron*, vol. 26, no. 1, pp. 55–67, 2000.

[16] A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach," *Journal of Cognitive Neuroscience*, vol. 5, no. 2, pp. 162–176, 1993.

[17] W. Deburchgraeve, P. J. Cherian, M. De Vos, R. M. Swarte, J. H. Blok, G. H. Visser, and P. Govaert, "Neonatal seizure localization using parafac decomposition," *Clinical Neurophysiology*, vol. 120, pp. 1787–1796, 2009.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statisitcal Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[19] L. Ding, "Reconstructing cortical current density by exploring sparseness in the transform domain," *Physics in Medicine and Biology*, vol. 54, pp. 2683 – 2697, 2009.

[20] L. Ding and B. He, "Sparse source imaging in EEG with accurate field modeling," *Human Brain Mapping*, vol. 19, September 2008.

[21] K. J. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, no. 1, pp. 1104–1120, 2008.

[22] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *Electroencephalography and Clinical Neurophysiology*, pp. 231–251, 1995.

[23] A. Gramfort, "Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision," Ph.D. dissertation, Telecom ParisTech, 2009.

[24] A. Gramfort, M. Kowalski, and M. Hamalainen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in Medicine and Biology*, vol. 57, pp. 1937 – 1961, 2012.

[25] A. Gramfort, D. Strohmeier, J. Haueisen, M. Hamalainen, and M. Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410 – 422, 2013.

[26] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, Nov. 2008.

[27] J. Gross, J. Kujala, M. Hämäläinen, L. Timmermann, A. Schnitzler, and R. Salmelin, "Dynamic imaging of coherent sources: Studying neural interactions in the human brain," *PNAS*, vol. 98, no. 2, pp. 694 – 699, 2001.

[28] S. Haufe, V. Nikulin, A. Ziehe, K.-R. Mueller, and G. Nolte, "Combining sparsity and rotational invariance in EEG/MEG source reconstruction," *NeuroImage*, vol. 42, 2008.

[29] M.-X. Huang, A. M. Dale, T. Song, E. Halgren, D. L. Harrington, I. Podgorny, J. M. Canive, S. Lewis, and R. R. Lee, "Vector-based spatial-temporal minimum l1-norm solution for meg," *NeuroImage*, vol. 31, pp. 1025–1037, 2006.

[30] A. Karfoul, L. Albera, and P. Comon, "Canonical decomposition of even order Hermitian positive semi-definite arrays," in *Proc. on EUSIPCO*, Glasgow, Scotland, 2009, pp. 515–519.

[31] K. Liao, M. Zhu, L. Ding, S. Valette, W. Zhang, and D. Dickens, "Sparse imaging of cortical electrical current densities using wavelet transforms," *Physics in Medicine and Biology*, vol. 57, pp. 6881 – 6901, 2012.

[32] L.-H. Lim and P. Comon, "Blind multilinear identification," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1260–1280, 2014.

[33] T. Limpiti, B. D. Van Veen, and R. T. Wakai, "Cortical patch basis model for spatially extended neural activity," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 9, pp. 1740 – 1754, 2006.

[34] F. Lin, J. W. Belliveau, A. M. Dale, and M. S. Hamalainen, "Distributed current estimates using cortical orientation constraints," *Human Brain Mapping*, vol. 27, pp. 1 – 13, 2006.

[35] X. Luciani and L. Albera, "Canonical polyadic decomposition based on joint eigenvalue decomposition," *Chemometrics and Intelligent Laboratory Systems*, vol. 132, pp. 152–167, Mar. 2014.

[36] K. Matsuura and Y. Okabe, "A robust reconstruction of sparse biomagnetic sources," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 8, pp. 720 – 726, Aug. 1997.

[37] J. Mattout, C. Phillips, W. Penny, M. Rugg, and K. Friston, "MEG source localization under multiple constraints: an extended Bayesian framework," *NeuroImage*, vol. 30, pp. 753–767, 2006.

[38] I. Merlet, "Dipole modeling of interictal and ictal EEG and MEG," *Epileptic Disord Special Issue*, pp. 11 – 36, July 2001.

[39] C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. G. D. Peralta, "EEG source imaging," *Clinical Neurophysiology*, vol. 115, no. 10, pp. 2195–2222, Oct. 2004.

[40] J. Möcks, "Decomposing event-related potentials: a new topographic components model," *Biological Psychology*, vol. 26, pp. 199–215, 1988.

[41] A. Moiseev, J. M. Gaspar, J. A. Schneider, and A. T. Herdman, "Application of multi-source minimum variance beamformers for reconstruction of correlated neural activity," *NeuroImage*, vol. 58, pp. 481–496, 2011.

[42] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Transactions on Biomedical Engineering*, vol. 39, pp. 541–557, June 1992.

[43] P. L. Nunez and R. Srinivasan, *Electric fields of the brain*, 2nd ed. New York: Oxford University Press, 2006.

[44] E. Ou, M. Hamalainen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, 2009.

[45] R. D. Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (sLORETA): technical details," *Methods and Findings in Experimental and Clinical Pharmacology*, 2002.

[46] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomograph: A new method for localizing elecrical activity in the brain," *Int. Journal of Psychophysiology*, vol. 18, pp. 49 – 65, 1994.

[47] F. Römer and M. Haardt, "A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalization (SECSI)," *Signal Processing*, vol. 93, pp. 2462–2473, 2013.

[48] K. Sekihara, M. Sahani, and S. S. Nagarajan, "Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction," *NeuroImage*, vol. 25, pp. 1056 – 1067, 2005.

[49] T. S. Tian and Z. Li, "A spatio-temporal solution for the EEG/MEG inverse problem using group penalization methods," *Statistics and Its Interface*, vol. 4, no. 4, pp. 521–533, 2011.

[50] K. Uutela, M. Hamalainen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, pp. 173 – 180, 1999.

[51] B. C. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, Sep. 1997.

[52] M. Vega-Hernández, E. Martínes-Montes, J. M. Sánchez-Bornot, A. Lage-Castellanos, and P. A. Valdés-Sosa, "Penalized least squares methods for solving the EEG inverse problem," *Statistics Sinica*, vol. 18, pp. 1535 – 1551, 2008.

[53] M. Wagner, M. Fuchs, H. A. Wischmann, and R. Drenckhahn, "Smooth reconstruction of cortical sources from EEG and MEG recordings," *NeuroImage*, vol. 3, no. 3, p. S168, 1996.

[54] D. Wipf and S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, pp. 947 – 966, 2009.

[55] D. Wipf, J. Owen, H. Attias, K. Sekihara, and S. Nagarajan, "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG," *NeuroImage*, vol. 49, pp. 641 – 655, 2010.

[56] P. Xu, Y. Tian, H. Chen, and D. Yao, "Lp norm iterative sparse solution for EEG source localization," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 3, March 2007.

Hanna Becker (hanna.becker@technicolor.com) received the B. Sc. and M. Sc. degrees in Electrical Engineering and Information Technology from the Ilmenau University of Technology, Germany, in 2010 and 2011, respectively, and the Ph. D. degree of the University of Nice-Sophia Antipolis in 2014. In 2010, she received the EUSIPCO Best Student Paper Award and in 2015, she has been nominated for the thesis award of the Société Francaise de Génie Biologique et Médical and the French IEEE EMBS section. She is currently working as a post-doctoral researcher at Technicolor R&D France. Her research interests include blind source separation, tensor decompositions, and source localization.

Laurent Albera (laurent.albera@univ-rennes1.fr) was born in Massy, France, in 1976. He received in 2003 the Ph.D. degree in Sciences from the University of Nice Sophia-Antipolis, France. He is now Assistant Professor at the University of Rennes 1 and is affiliated with the INSERM research group LTSI (Laboratoire Traitement du Signal et de l'Image). He was member of the Scientific Committee of the University of Rennes 1 between 2008 and 2010. He received in 2010 the H.D.R. (Habilitation to Lead Researches) degree in Sciences from the University of Rennes 1, France. His research interests include human electrophysiological inverse problems based on high-order statistics, sparsity and multidimensional algebra.

Pierre Comon (pierre.comon@gipsa-lab.fr) (M'87 - SM'95 - F'07) is research director with CNRS since 1998, now at Gipsa-Lab, Grenoble, France. Before 1998, he has been employed by various private companies, including the Thales group. His research interests include High-Order Statistics, Blind Source Separation, Statistical Signal and Array Processing, Tensor decompositions, Multi-Way Factor Analysis and Data Mining, with applications to health end environment. Dr. Comon was in the Editorial Boards of several international journals including the IEEE Transactions on Signal Processing, the Eurasip journal Signal Processing, and the IEEE Transactions on Circuits and Systems I. He is currently associate editor with the SIAM Journal on Matrix Analysis and Applications.

Rémi Gribonval (remi.gribonval@irisa.fr) (FM'14) is a Senior Researcher with Inria. A former student at École Normale Supérieure (Paris, France), he received the Ph. D. degree in applied mathematics from Université Paris-IX Dauphine in 1999. His research focuses on mathematical signal processing and machine learning, with an emphasis on sparse approximation, inverse problems, and dictionary learning. He founded the series of international workshops SPARS on Signal Processing with Adaptive/Sparse Representations. In 2011, he was awarded the Blaise Pascal Award in Applied Mathematics and Scientific Engineering from the SMAI by the French National Academy of Sciences, and a starting investigator grant from the European Research Council. He is the leader of the PANAMA research group on sparse audio processing.

Fabrice Wendling (fabrice.wendling@univ-rennes1.fr) is Director of Research at INSERM. He is heading the team SESAME: "Epileptogenic Systems: Signals and Models" at LTSI, Rennes, France. He got the Biomedical Engineering Diploma (1989) from the University of Technology of Compiègne (France), the Master of Science (1991) from Georgia Tech (Atlanta, USA) and the PhD (1996) from the University of Rennes (France). He has been working on brain signal processing and modeling for more than 20 years, in close collaboration with clinicians. In 2012, he received the award (Prix Michel Montpetit) from the French Academy of Science. He co-authored about 110 peer-reviewed articles.

Isabelle Merlet (isabelle.merlet@univ-rennes1.fr) is a full time research scientist at INSERM. She received the PhD degree in Neurosciences from Lyon 1 University in 1997. She has been working at the Epilepsy department of the Montreal Neurological Institute (1997-2000), at the Neurological Hospital of Lyon (2000-2005) and in the SESAME team of the Signal and Image Processing Laboratory of Rennes (2005 -). Since 1993, her work is devoted to the validation source localization methods and to their application to EEG or MEG signals. She is involved in the transfer of these methods to the clinical ground particularly in the field of epilepsy research.