Combining spatial wavelets and sparse Bayesian learning for extended brain sources reconstruction

Samy Mokhtari, Jean-Michel Badier, Christian G. Bénar, Bruno Torrésani

Combining spatial wavelets and sparse Bayesian learning for extended brain sources reconstruction

Samy Mokhtari, Jean-Michel Badier, Christian G. Bénar, Bruno Torrésani

Abstract*—*Objective: **the accurate reconstruction of extended cortical activity from M/EEG data is a difficult, ill-conditioned problem. This work proposes to model unknown sources as expansions on a wavelet system defined on the cortical surface, and addresses resulting numerical optimization problems. The objective is to obtain accurate source localization, together with quantitatively relevant amplitude and time course.**

Approach: **Unknown sources are expanded on a system of spectral graph wavelets (SGW) defined on the cortical surface. Unknown wavelet coefficients are estimated using either variational or Bayesian formulations, involving priors that favor extended source through sparsity in the wavelet domain: sparsity-inducing regularization, or sparse Bayesian learning (SBL). These approaches are tested and compared with concurrent approaches on numerical simulations. The quality of reconstructions is assessed using a set of complementary metrics.**

Results: **SGW-based approaches are able to identify accurately extended sources. The combination with SBL is particularly attractive, as it doesn't involve hyperparameter tuning. It yields accurate and robust results with respect to all considered metrics, and performs remarkably well in terms of depth bias.**

Conclusion: **This paper demonstrates the usefulness of cortical wavelets for reconstructing cortical activity from M/EEG data, and the impact of sparse Bayesian learning in this context.**

Significance: **Being able to identify localization, depth and time course of brain activity from M/EEG data is important in clinical applications such as epilepsy, as it can improve the detection of potential sources of seizures.**

***Index Terms*—M/EEG inverse problem, extended sources, spectral graph wavelets, Sparse Bayesian Learning.**

I. Introduction

Magneto- and Electro-encephalography (M/EEG) are the main (non-invasive) imagery modalities which allow, thanks to their high temporal resolution, physicians and clinicians to detect and localize cortical activities. From a given set of sensors and measurements (which contain cortical activity, brain background activity and sensor noise), the aim is to reconstruct the sources activity. This is done by solving the M/EEG inverse problem [1], [2], [3]. To address this challenge, multiple methods have been designed, generally classified in parametric versus imaging/distributed methods. While the former considers a few equivalent current dipoles of unknown locations, the latter considers a fine spatial discretization of the brain cortical surface, with $\sim 10^4$ unknown cortical sources with known and fixed locations. While inherently leading to very under-determined inverse problems, distributed methods can be expected to be better suited to large cortical activities. However, results tend to depend strongly on prior information, which is injected into the chosen approach through regularization (in variational methods) or prior (in Bayesian approaches).

The goal of this paper is to introduce and study variational and Bayesian approaches which rely on a spatial transformation, and aim at promoting extended solutions compressed in a relevant transformed domain. We particularly focus on a spatial wavelet transform which, combined with sparsity-promoting regularizations/priors is indeed able to produce extended solutions. While variational approaches suffer, as often, from the difficulty of choosing appropriate values for hyperparameters, we show that the integration of wavelet-based regularization into the sparse Bayesian learning (SBL) paradigm yields more robust solutions. Wavelet-based solutions had already been studied in [4], where spherical wavelet bases were used in the framework of variational approaches. We rather rely here on spectral graph wavelets [5], which provide expansions on an overcomplete spatial wavelet family, and investigate corresponding variational and SBL solutions. These solutions are compared with reference solutions using extensive numerical simulations designed to mimic real evoked potentials. Performance assessment is done through a set of metrics, which include standard ones (implemented in specialized computer packages such as MNE-Python [6], which we therefore use) as well as new ones, which we found convenient in the context of sparse solutions.

Besides methodological contributions (in particular the use of cortical wavelet frames and additional comparison metrics), the main outcome of this paper is the fact that the combination of cortical wavelet expansions for the brain activity with the sparse Bayesian learning framework, provides relevant and robust source estimates in the case of extended activity, which do not rely on difficult parameter choices, and outperform significantly concurrent approaches of the same family (variational and Bayesian, in distributed source models).

The paper is organized as follows. After the current introduction, Section II presents reference solution that will be used in this paper. The construction of wavelets and the corresponding solvers are described in Section III, and evaluation metrics are discussed in Section IV. Section V presents the simulation protocol, and discusses and compares results obtained using various solvers. Finally, Section VI is devoted to discussion and conclusions. Preliminary results of this work were presented in the conference paper [7].

S. Mokhtari and B. Torrésani are with Aix Marseille Univ, CNRS, I2M, Marseille, France. email: samy.mokhtari@univ-amu.fr, bruno.torresani@univ-amu.fr

J.M. Badier and C. Bénar are with Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France. email: jean-michel.badier@univ-amu.fr, christian.benar@univ-amu.fr

This work has benefited from support from the French government through the BMWs project (ANR-20-CE45-0018), and the *Institut Convergence* ILCB (ANR-16-CONV-0002).

II. State of the art and reference solutions

Suppose M/EEG data are measured at J_0 sensors during L time steps, and the cortical surface is discretized on N source points. The M/EEG inverse problem starts with the observation equation connecting cortical sources to measurements :

$$\mathbf{Z_0} = \mathbf{G_0}\mathbf{S} + \mathbf{B_0} \ , \tag{1}$$

where $\mathbf{Z_0} \in \mathbb{R}^{J_0 \times L}$ denotes space-time measurements, $\mathbf{S} \in \mathbb{R}^{N \times L}$ the unknown sources activity, $\mathbf{B_0} \in \mathbb{R}^{J_0 \times L}$ the space-time matrix of both observation noise and background brain activity (zero-mean, with known covariance matrix $\mathbf{\Sigma_{B_0}}$), and $\mathbf{G_0} \in \mathbb{R}^{J_0 \times N}$ is the leadfield matrix which encodes electromagnetic propagation from cortex to sensors.

Given $\mathbf{G_0}$ and $\mathbf{\Sigma_{B_0}}$ (and some additional assumptions), M/EEG inverse problem aims to estimate sources $\mathbf{S}$ from observations $\mathbf{Z_0}$. Before explicitly addressing observation equation (1), data are often pre-processed to correct for noise correlations between sensors; this can be done with a whitening and dimension reduction of the noise covariance matrix. The resulting matrix $\mathbf{B} = \mathbf{\Upsilon}\mathbf{B_0} \in \mathbb{R}^{J \times L}$ (where $\mathbf{\Upsilon} \in \mathbb{R}^{J \times J_0}$ denotes the whitener) is thus white with zero mean and unit variance. For the sake of simplicity, we will hereafter consider whitened noise $\mathbf{B}$, leadfield $\mathbf{G}$ and data $\mathbf{Z}$ matrices.

Several approaches are routinely used in literature to tackle observation equation (1), which include among others dipole localization [8], [9] or various forms of beamformers [10], [11]. This article will be mainly concerned with variational and Bayesian approaches for distributed source models (see [1]). Most variational methods boil down to the following generic optimization problem (2), which involves a quadratic data fidelity term and a penalty term $f(\mathbf{S})$:

$$\mathbf{S}_* = \underset{\mathbf{S} \in \mathbb{R}^{N \times L}}{\arg\min} \left[\frac{1}{2} \left\| \mathbf{Z} - \mathbf{G}\mathbf{S} \right\|_F^2 + f(\mathbf{S}) \right] \ , \tag{2}$$

where $\| \cdot \|_F$ stands for the Frobenius norm, i.e. the square root of the sum of squared matrix coefficients.

A wide variety of penalizations/regularizations can be found in literature. While the current paper does not aim to provide a comprehensive review, we nevertheless recall the penalizations that will be of most interest for the present study.

A. Closed-form solutions with quadratic regularizations

The most commonly used penalization is quadratic :

$$f(\mathbf{S}) = \|\mathbf{D}\mathbf{S}\|_F^2 \ , \tag{3}$$

for some symmetric, generally diagonal, invertible matrix $\mathbf{D}$. With the above penalization, the solution of the optimization problem (2) admits a closed-form expression:

$$\mathbf{S}_* = \mathbf{D}^{-2}\mathbf{G}^T \left(\mathsf{I}_\mathbf{J} + \mathbf{G}\mathbf{D}^{-2}\mathbf{G}^T \right)^{-1} \mathbf{Z} \ . \tag{4}$$

Special cases of (3) include the *MNE* (minimal norm estimate) solution [12] ($\mathbf{D} = \sqrt{\lambda}\mathsf{I}_N$), and the *wMNE* (weighted minimal norm estimate) solution [13], [14], where $\mathbf{D}$ is a weight matrix intended to correct MNE depth bias. In e-LORETA [15], an iterative algorithm yields a data driven estimate for $\mathbf{D}$ and therefore a corresponding source estimate.

While closed form solutions are very convenient, MNE and wMNE are however known to overestimate the spatial extent of sources $\mathbf{S}$, and underestimate corresponding amplitudes. In addition, the parameter λ, linked to the source spatial covariance matrix $\mathbf{\Gamma} = \mathbf{D}^{-2}$, is an extra parameter whose fine tuning is not obvious. In many situations, the interest is more focused on the spatial extent, namely source localization. A number of post-processings are used to yield more accurate support estimates. dSPM [16] and s-LORETA [17] are among the most popular.

B. Sparse solutions through non-differentiable penalizations

To reduce the blurring effect, other penalizations f have been proposed in literature, often promoting some form of sparsity. There is now a vast literature on the subject, we refer to [18] for a review and a comparison of some recent contributions in the context of M/EEG inverse problem. In these approaches, no closed form solution exists as the penalization is no longer differentiable, and the optimization problem (2) has to be solved numerically. One may first think of the *MCE* (minimal current estimate) solution [19], [20], where $f(\mathbf{S}) = \lambda \|\mathbf{S}\|_1$ tends to yield very "spiky" source estimates, with quite irregular time courses. One may also cite the VB-SCCD approach [21], where $f(\mathbf{S}) = \lambda \|\mathbf{S}\|_{\mathsf{TV}} = \lambda \|\nabla \mathbf{S}\|_1$. The latter is based on total variation (TV), where ∇ is a discrete gradient operator which has to be properly defined on the surface mesh. TV-based solutions tend to enforce sparsity in a transformed domain, here the gradient. A slightly modified and sparser version, called sVB-SCCD [21], [22] is based upon a combined ℓ^1-TV penalization $f(\mathbf{S}) = \lambda \left(\alpha \|\mathbf{S}\|_1 + \|\mathbf{S}\|_{TV} \right)$. Hereafter, we shall rather consider the following form:

$$f(\mathbf{S}) = \lambda \left(\gamma \|\mathbf{S}\|_1 + (1-\gamma) \|\mathbf{S}\|_{\mathsf{TV}} \right), \ \gamma \in [0,1] \tag{5}$$

In [22], $\alpha \in [0.01, 1]$, and $\alpha = 0.67$ is advocated as a reasonable choice for the tested scenarios. For both MCE and sVB-SCCD, numerical algorithms have to be carefully chosen, [22] advocates for ADMM. Regularization (5) allows for an easier interpretation of the hyperparameters, where λ plays the same role as for the MCE solver, while γ drives the convex combination of ℓ^1 and TV-penalizations. Generalized TV penalization have been studied in [23].

Lastly, mixed norms (MxNE) combining sparsity in space with persistence in time were proposed and studied in [24], [25]. The regularization term is of the form $f(\mathbf{S}) = \lambda \|\mathbf{S}\|_{21} = \lambda \sum_k \|\mathbf{S}[:,k]\|_2$. Mixed norms were also used in combination with time-frequency dictionaries [26], and combined with TV spatial regularization [22].

C. Bayesian point of view

A major weakness of the above mentioned approaches is the difficulty of tuning the hyperparameters involved in the formulations (e.g. regularization parameter λ). In simple situations, Bayesian interpretations yield a relationship between these parameters and quantities such as signal to noise ratio (SNR), which can be guessed or measured from sensor signals. For instance, wMNE may be interpreted in terms of a *Maximum A Posteriori* (MAP) estimation, modeling columns of $\mathbf{B}$ (baseline) as iid samples from a normal distribution $\mathcal{N}(0, \mathbf{I})$,

and columns of $\mathbf{S}$ as iid samples from a normal distribution $\mathcal{N}(0, \mathbf{\Gamma})$, with $\mathbf{\Gamma} = \lambda^{-1}\mathsf{I}_N$ in the case of MNE. The posterior distribution of sources $p(\underline{s}(l)|\underline{z}(l))$ is multivariate normal, with covariance matrix

$$\mathbf{\Sigma_S} = \mathbf{\Gamma} - \mathbf{\Gamma}\mathbf{G}^T\left(\mathbf{\Sigma_B} + \mathbf{G}\mathbf{\Gamma}\mathbf{G}^T\right)^{-1}\mathbf{G}\mathbf{\Gamma}\ . \tag{6}$$

In the MCE case, modeling columns of $\mathbf{S}$ as iid samples from a product of 1D Laplacian distributions with parameter λ^{-1}, the sources covariance matrix is given by $\mathbf{\Gamma} = 2\lambda^{-2}\mathsf{I}_N$. For both MNE and MCE solutions, heuristics can be derived to connect the regularization parameter λ to the square signal to noise ratio $\rho^2 = \mathsf{Tr}\left(\mathbf{\Sigma_{Z_0}}\right)/\mathsf{Tr}\left(\mathbf{\Sigma_{B_0}}\right)$ (signal and noise levels in the sensor domain are defined by the traces of the covariance matrices $\mathbf{\Sigma_{Z_0}}$ and $\mathbf{\Sigma_{B_0}}$) :

$$\lambda_{\mathsf{MNE}} = \frac{\|\mathbf{G}_0\|_F^2}{(\rho^2-1)\mathsf{Tr}\left(\mathbf{\Sigma_{B_0}}\right)}\ ; \quad \lambda_{\mathsf{MCE}} = \sqrt{2\lambda_{\mathsf{MNE}}} \tag{7}$$

Such heuristics are more difficult to obtain for penalization terms $f(\mathbf{S})$ that involve TV or mixed norms, and even worse for combinations of norms that involve several parameters.

Remark 1: The above heuristics differ slightly from expressions routinely used in reference softwares, which perform similar calculations from whitened observations $\mathbf{Z}$ instead of original ones $\mathbf{Z}_0$ as we do. We believe the expressions we use are more relevant when interpreted in a Bayesian context.

D. Sparse Bayesian Learning

Beside MAP approaches, *Sparse Bayesian Learning* (SBL, see for example [27], [28]) allows estimating simultaneously the sources amplitudes and their covariance matrices, assumed to be diagonal $\mathbf{\Gamma} = \mathsf{diag}(\gamma_1, \ldots, \gamma_N)$. SBL-based inverse problems have recently been put into a unifying context [29], for which most algorithms fall into the class of *Majorization-Minimization* algorithms (MM, see [30] for a review). They essentially boil down to solving the following non-convex optimization problem, with respect to the source variances $\underline{\gamma} = (\gamma_1, \ldots, \gamma_N) \in (\mathbb{R}_+)^N$

$$\underline{\gamma_*} = \underset{\underline{\gamma}\in(\mathbb{R}_+)^N}{\mathsf{arg\,min}}\left[\frac{1}{L}\mathsf{Tr}\left(\mathbf{Z}^T{\mathbf{\Sigma_Z}}^{-1}\mathbf{Z}\right) + \ln\det(\mathbf{\Sigma_Z})\right]\ , \tag{8}$$

where $\mathbf{\Sigma_Z} = \left(\mathsf{I_J} + \mathbf{G}\mathbf{\Gamma}\mathbf{G}^T\right)$ is the posterior covariance matrix. Once $\mathbf{\Gamma}$ has been estimated, its value can be plugged in the (wMNE closed-form) expression of $\mathbf{S}$.

For the record, let us mention the alternative Bayesian approach, based on the MEM (maximum entropy on the mean) principle [31], [32], [33]. Remarkably, the MEM approach leads to a smooth optimization problem in the domain of sensors (instead of sources), hence of much lower complexity, and for which simple and faster differentiable optimization tools can be used. The wavelet techniques to be explained below would easily be included in this approach.

III. Use of a family of cortical wavelets

Approaches that exploit sparsity in the spatial domain, such as dipole models or MCE/MxNE are well adapted to identify focal activity. However, when the activity of interest has a significant spatial extent, they produce estimates that are often too sparse. In such situations, it is wise to search for solutions that are sparse in a transformed domain. As an example, VB-SCCD [18] optimizes the sparsity of a spatial gradient of the solution. Similar approaches based on Laplacian and spherical wavelet transform have been considered in [4], where authors report improved results for localizing extended activity.

A. Problem formulation

Bases or redundant systems of waveforms have often been used to describe time dependence and perform dimension reduction in the time domain. Examples include wavelets [32], [33], time-frequency systems, or waveforms estimated from data. Such approaches do not seem to have been considered for the spatial dependence of sources (except [4], and approaches based upon cortex parcellation [32]), due to the lack of such simple waveform families. In this section, we introduce and use spatial wavelets for the M/EEG inverse problem.

We define wavelets as functions $\underline{w}_k \in \mathbb{R}^N$ on the cortex, labelled by a vertex and a scale index. Concatenating these vectors into a matrix $\mathbf{W} = \left(\underline{w}_1 \quad \ldots \quad \underline{w}_{N_w}\right)^T \in \mathbb{R}^{N_w\times N}$, the family of wavelets forms a frame for functions on the surface if $\mathbf{W}$ is left invertible (in other words the family of wavelets is a complete set in $\mathbb{R}^N$). The construction of such a frame will be detailed in the next section. $\mathbf{W}$ is the analysis operator of the wavelet frame. A space-time matrix $\mathbf{S} \in \mathbb{R}^{N\times L}$ can be written as $\mathbf{S} = \mathbf{W}^T\mathbf{A}$, in terms of a wavelet-time coefficient matrix $\mathbf{A} \in \mathbb{R}^{N_w\times L}$. Putting the focus on the matrix $\mathbf{A}$ is called a synthesis-based approach, as opposed to the analysis-based approach which rests on the matrix $\mathbf{WS} \in \mathbb{R}^{N_w\times L}$. In this synthesis-based setting, the generic inverse problem (2) may be written as

$$\mathbf{A}_* = \underset{\mathbf{A}\in\mathbb{R}^{N_w\times L}}{\mathsf{arg\,min}}\left[\frac{1}{2}\|\mathbf{Z} - \mathbf{G}_W\mathbf{A}\|_F^2 + g(\mathbf{A})\right] \tag{9}$$

with $\mathbf{G}_W = \mathbf{G}\mathbf{W}^T$ the transformed leadfield matrix. This formulation is very close to the usual one, the difference being that the optimization is done in the synthesis coefficients domain. Notice that when solving (9), the waveform system is used only twice, for the computation of the transformed leadfield matrix $\mathbf{G}_W$ and for the synthesis of the solution $\mathbf{S}_*$. All reference solvers for the M/EEG inverse problem can be transposed *mutatis mutandis* to this new setting. We stress that the dimension of the transformed space $(N_w \times L)$ may be significantly larger than the one of the original space $(N \times L)$, as the number N_w of waveform atoms used to decompose the solution may be, in the case of a redundant wavelet family, several times larger (depending on the number of wavelet scales considered) than the number of sources N.

B. Signals, Fourier transform and wavelets on graphs

We now detail the construction of the wavelet frame $\mathbf{W}$ for the graph describing cortical sources. In the present work, we decided to follow the spectral graph wavelets (SGW) framework developed in [5].

Consider the cortical mesh graph $(\mathcal{V}, \mathcal{E}, \omega)$ (with vertex set $\mathcal{V}$, edge set $\mathcal{E}$ and weight function ω, the latter allowing to define the graph adjacency matrix $\mathcal{A}$). The graph Laplacian is given by $\mathcal{L} = \mathcal{D} - \mathcal{A}$, where $\mathcal{D}$ is the diagonal degree

matrix, whose diagonal elements d_i equal the number of edges connected to edge i. $\mathcal{L}$ is a real symmetric matrix (for undirected graphs), and therefore possesses a complete set of orthonormal eigenvectors $\{\chi_1, \ldots, \chi_N\}$, associated to real, non-negative and increasing eigenvalues $\{\lambda_1, \ldots, \lambda_N\}$. The minimal eigenvalue λ_1 equals zero, with a multiplicity equal to the number of connected components of the graph. The linear transform $\mathcal{F}: \ X \in \mathbb{R}^N \longmapsto \{\mathcal{F}[X](n) = \langle \chi_n, X \rangle\}_{1 \leq n \leq N}$ may be interpreted as the graph analogue of the Fourier transform. The orthonormality of the basis $\{\chi_1, \ldots \chi_N\}$ gives the Parseval formula and the inverse transform: for any $X \in \mathbb{R}^N$, $\|\mathcal{F}[X]\| = \|X\|$ and $X = \sum_{n=1}^{N} \mathcal{F}[X](n)\chi_n$.

Following [5], a graph analogue of the wavelet transform can be derived from the above graph Fourier transform. Given a dilation parameter $s > 0$ and a vertex n, the graph wavelet transform of a signal $X \in \mathbb{R}^N$ may be defined as follows :

$$\mathcal{W}[X](s, n) = \langle \Psi_{s,n}, X \rangle \ , \tag{10}$$

with $\Psi_{s,n}$ the wavelet atom associated to the vertex n and dilation parameter s. In [5], these atoms $\Psi_{s,n}$ are generated in the spectral domain, as the graph Fourier transforms of Dirac functions centered on each vertex, multiplied by a band-pass positive kernel g : for any $l \in \{1, \ldots, N\}$:

$$\mathcal{F}\left[\Psi_{s,n}\right](l) = g\left(s\lambda_l\right) \langle \chi_l, \delta_n \rangle \tag{11}$$

As outlined in [5], such a construction may lead to an ill-conditioned matrix for the synthesis operator (frame vectors associated to small scales s may become numerically parallel). To bypass this difficulty, [5] suggests to encapsulate the low-frequency component of signals through an ad hoc "scaling function" Φ_n (labelled by a vertex index), defined independently, in the spectral domain, via a low-pass kernel h :

$$\mathcal{F}\left[\Phi_n\right](l) = h(\lambda_l) \langle \chi_l, \delta_n \rangle \tag{12}$$

For the specific design of both low-pass h and band-pass g kernels, we refer the reader to [5].

C. Wavelet-sparse solutions

We now consider the inverse problem formulated in the wavelet domain, using the synthesis point of view. This requires computing the transform-domain leadfield matrix $\mathbf{G_W} = \mathbf{G}\mathbf{W}^T$. This is the most costly part of the approach.

1) Variational formulations: Denote by $\mathbf{A} = [\mathbf{A}_\Psi, \mathbf{A}_\Phi]$ the vector of multiscale (wavelet+scaling) synthesis coefficients:

$$\mathbf{S} = \mathbf{W}^T\mathbf{A} = \mathbf{\Psi}^T\mathbf{A}_\Psi + \mathbf{\Phi}^T\mathbf{A}_\Phi \ .$$

In addition to classical regularizations in the sources or gradient domains, the wavelet domain optimization problem (9) will hereafter be solved using two different penalizations :

$$g_{\mathsf{MNE}}(\mathbf{A}) = \|\mathbf{D}\mathbf{A}\|_F^2 \ , \qquad g_{\mathsf{MCE}}(\mathbf{A}) = \lambda_{\mathsf{MCE}}\|\mathbf{A}\|_1 \tag{13}$$

The first one will be coupled with a wavelet-SBL approach to estimate the weight matrix $\mathbf{D}$, while the second aims at promoting sparsity in the wavelet coefficient domain.

2) Sparse Bayesian Learning approaches: The wavelet version of SBL is also a simple transposition of classical SBL approach, and yields the non convex problem

$$\underline{\gamma_A}^* = \underset{\underline{\gamma_A} \in (\mathbb{R}_+)^{N_w}}{\mathsf{arg\,min}} \left[\mathsf{Tr}\big(\mathbf{C_Z}\mathbf{\Sigma_Z}(\underline{\gamma_A})^{-1}\big) + \ln\det(\mathbf{\Sigma_Z}(\underline{\gamma_A}))\right] \tag{14}$$

where $\mathbf{\Gamma}_A = \mathsf{diag}(\underline{\gamma_A})$ is the prior covariance matrix of wavelet synthesis coefficients, $\mathbf{\Sigma_Z}(\underline{\gamma_A}) = \mathsf{I}_J + \mathbf{G_W}\mathbf{\Gamma}_A\mathbf{G_W}^T$ is the posterior covariance matrix of observations, and $\mathbf{C_Z} = \frac{1}{T}\mathbf{Z}\mathbf{Z}^T$ is the sample covariance matrix. Once $\underline{\gamma_A}^*$ has been obtained, estimated sources are obtained via

$$\mathbf{A}_* = \mathbf{\Gamma}_A^*\mathbf{G_W}^T\left[\mathbf{\Sigma}_Z(\underline{\gamma_A}^*)\right]^{-1}\mathbf{Z} \ , \qquad \mathbf{S}_* = \mathbf{W}^T\mathbf{A}_* \ . \tag{15}$$

This approach will be called sgw-SBL or w-SBL.

IV. Analysis of evaluation metrics

To assess, qualitatively and quantitatively, the performances of any inverse problem model for the distributed sources framework, evaluation metrics are mandatory. Given a simulation case where the ground truth is known, one may try to encapsulate, with well-chosen metrics :

- purely-distance based information (spatial extent of true/estimated sources, spatial deviation or discrepancy between these two sets...);
- information related to the number of true/estimated active sources, the number of true positives/negatives and false positives/negatives ;
- amplitude-based information ;
- or a mix of these features.

Beside these various choices and associated metrics, one shall first state which physical value will be studied under these scopes. Indeed, one could study straightforwardly the sources time-dependent amplitudes, their time-dependent absolute or squared value, or rather a time-average of these amplitudes on a given time segment of interest. One may also consider either normalized amplitudes or not. In the following, to facilitate interpretation and comparisons among all considered inverse problem solvers, all chosen metrics/scopes will be applied on time-averaged values, as hereafter detailed : given a space-time matrix $\mathbf{S}_* \in \mathbb{R}^{N \times L}$ gathering the amplitude maps for all time snapshots, a time-average $\langle \mathbf{S}_*^2 \rangle$ is computed on a prescribed *time segment of interest* denoted by soi, thus leading to an amplitude map $\mathcal{A} \in \mathbb{R}^N$ describing the amount of energy contained in the *segment of interest* :

$$\mathcal{A} = \sqrt{\langle \mathbf{S}_*^2 \rangle_{\mathsf{soi}}} \tag{16}$$

In this section, we lay down and describe the several evaluation metrics that will be used to assess the models performances. To this end, let us first detail some notations.

Notations: from now on,

- I denotes the index set of true active sources; its cardinality is denoted by $\sharp I$.
- $\widehat{I}_\alpha$ denotes the index set of estimated active sources, defined with respect to a threshold parameter $\alpha > 0$;
- $|\widetilde{s}_i| = \frac{1}{\sum_{k=1}^{N}|s_k|}|s_i|$ denotes the normalized (positive, sum to 1) sources amplitudes;
- $\|\cdot\|$ denotes the 3D Euclidean norm;
- r_i denotes the 3D vector position of source # i.

A. Region Localization Error : purely distance-based metrics

We start with the classical Region Localization Error (RLE). Given a thresholding parameter $\alpha > 0$ (by default a percentage of the maximum value), and assuming $I \cap \widehat{I}_\alpha \neq \emptyset$ and $\widehat{I}_\alpha \not\equiv I$, the RLE metrics is computed as follows :

$$RLE(\alpha) = \frac{1}{2\,\sharp I} \sum_{k \in I \setminus \left(I \cap \widehat{I}_\alpha\right)} \min_{l \in \widehat{I}_\alpha} \|r_k - r_l\| + \frac{1}{2\,\sharp \widehat{I}_\alpha} \sum_{l \in \widehat{I}_\alpha \setminus \left(I \cap \widehat{I}_\alpha\right)} \min_{k \in I} \|r_k - r_l\| \quad (17)$$

RLE combines equi-balanced information related to, on the one hand, the average smallest distance between the set of unexplained true sources/false negatives $\left(I \setminus \left(I \cap \widehat{I}_\alpha\right)\right)$ (and its relative population) and the estimated sources $\widehat{I}_\alpha$, and on the other hand, the average smallest distance between the set of "false" estimated sources/false positives $\left(\widehat{I}_\alpha \setminus \left(I \cap \widehat{I}_\alpha\right)\right)$ (and its relative population) and the true sources.

Notice that RLE does not measure any "distance" between true and estimated sources amplitudes, and is of course sensitive to the amplitude thresholding parameter α.

B. Spatial Deviation (SD) error

Spatial Deviation (SD) is another distance-based metrics, which interestingly involves an amplitude-weighting. Given a true active source in r_i, SD_i measures the spatial spread of the estimated sources around r_i, following

$$SD_i = \sum_{p=1}^{N} |\widetilde{s}_p| \, \|r_p - r_i\|^2, \quad (18)$$

Despite amplitude-weighting, SD is, like RLE, blind to a possible discrepancy between reference and estimated amplitudes.

In the case of a single true source, this metrics reaches its optimal value (zero, without any thresholding) when the estimated sources are concentrated in the true location r_i.

In the case of spatially-extended sources, the optimal value of the SD metrics is no longer zero. Indeed, for all true point sources $i_k \in I$ ($\sharp I \geq 2$), one has, for an optimal reconstruction (with respect to localization) :

$$SD_{i_k}^{opt} = \sum_{p \in I} |\widetilde{s}_p| \, \|r_p - r_{i_k}\|^2 \quad (19)$$

One may derive lower and upper bounds for this optimal value $SD_{i_k}^{opt}$ from the term $\frac{1}{\sharp I} \sum_{p \in I} \|r_p - r_{i_k}\|^2 = SD_{i_k}^{ref}$, which corresponds to the SD error for spatially-extended sources with identical amplitude. For a sub-optimal reconstruction with respect to localization, one may show that :

$$SD_{i_k} = (1-\beta) SD_{i_k}^{opt} + \beta \; SD_{i_k|I^c} \quad (20)$$

where $\beta = \frac{(N - \sharp I) S_{ave|I^c}}{N S_{ave}}$, and $S_{ave|I^c}$ refers to the average absolute amplitude on the complementary index set I^c. Thus, the SD metrics for each point source $i_k \in I$ of a spatially-extended activity is the convex combination of :

- an optimal SD error $SD_{i_k}^{opt}$, where only true sources locations are reconstructed by the model (i.e. no false positives);
- an SD error $SD_{i_k|I^c}$ computed by only taking into account false positives sources $i \notin I$.

The weight β in (20) takes into account the relative population of "false positives" $(N - \sharp I)$, and their average amplitude (relatively to the global average). We stress that, unlike the RLE metrics, SD is blind to the position of false negatives, as their zero-amplitude do not affect the SD error value.

To derive a global metric from the local SD value, we shall hereafter consider the maximum difference :

$$\delta_{SD} = \max_i |SD_i^{est} - SD_i^{ref}| \quad (21)$$

C. Depth metrics and distance between centers of mass

To investigate the possible depth bias phenomenon, well-known for the MNE solution, one may compare the local orthogonal distance to the inner-skull for both estimated and reference sources centers of mass, assuming that the brain curvature has a minimal impact on these measurements.

We stress that these centers of mass may either be restricted or not to the graph describing cortical sources, thus affecting the interpretation of the depth bias. In addition to depth, it is also interesting to keep track of the 3D Euclidean distance between estimated and reference centers of mass, as a large distance may affect the interpretation of the depth metrics. Notice that, even if these metrics take into account the respective estimated and reference amplitudes to compute centers of mass, they do not measure any "distance" between the two amplitudes.

D. Entropy, number of active sources, and spatial extent

Shannon entropy has been used in signal processing literature as a measure of dispersion. Unlike SD, entropy does not measure dispersion away from a reference location, and rather focuses on spreading/concentration. It is well suited to deal with a discrete set of sources amplitudes and possibly not-connected active cortical regions. Indeed, let us consider the discrete set of probabilities $\left\{\mathbb{P}\left(X^\Delta \equiv \text{source}_i\right) = |\widetilde{s}_i|\right\}_{1 \leq i \leq N}$. Its entropy is given by:

$$H_{X^\Delta} = -\sum_{i=1}^{N} |\widetilde{s}_i| \ln |\widetilde{s}_i| \quad (22)$$

and its link with the number of active sources N can be made explicit (using the normalized amplitudes $\widetilde{s}_i$) as:

$$H_{X^\Delta} = \ln(N) + \frac{1}{N S_{ave}} \sum_{i=1}^{N} \left[S_{ave} \ln\left(S_{ave}\right) - |s_i| \ln |s_i|\right] \quad (23)$$

Thus, Shannon entropy mixes information related to the number N of considered active sources, their average amplitude value S_{ave}, and the deviation of each amplitude term $|s_i| \ln |s_i|$ from its average value. Shannon entropy may then be compared to its "reference" counterpart, where N_{ref} active sources are for instance known to share the same amplitude. Indeed, in such a case, one has (with $H_{X^\Delta}^{ref} = \ln(N_{ref})$):

$$H_{X^\Delta} = H_{X^\Delta}^{ref} + \ln\left(\frac{N}{N_{ref}}\right) + \frac{\ln(S_{ave})}{N}\sum_{i=1}^{N}\left(1 - \frac{|s_i|\ln(|s_i|)}{S_{ave}\ln(S_{ave})}\right) \quad (24)$$

For connected active cortical regions, to link the Shannon entropy to an estimate of the sources spatial extent, one may consider the previous discrete set of probabilities as sorted with respect to the increasing Euclidean distance to the sources center of mass. Then, considering that the spatial discretization steps Δx_i between two consecutive sources are uniform, denote by $d_N = N\Delta x$ the distance of the furthest estimated active source with respect to the center of mass. Thus, one may re-write equation (24) as follows :

$$H_{X^\Delta} = H_{X^\Delta}^{ref} + \ln\left(\frac{d_N}{d_{ref}}\right) + \frac{\Delta x\ln(S_{ave})}{d_N}\sum_{i=1}^{N}\left(1 - \frac{|s_i|\ln(|s_i|)}{S_{ave}\ln(S_{ave})}\right) \quad (25)$$

Shannon entropy thus encapsulates both population-based (or distance-based) and amplitude-based information. Notice, however, that it does not measure any "distance" between S_{ave} and its reference counterpart S_{ave}^{ref}, which does not take part in $H_{X^\Delta}^{ref}$ computation for identical sources.

E. Binary classifier metrics : Precision, Recall, F1 score

The models performances may also be assessed using binarization of amplitude maps, through classical scores such as: the precision score, which measures the ability of a classifier not to label as false positive (FP) a sample that is truly negative (TN); the recall, which measures the ability of the classifier to find all the positive samples :

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \quad (26)$$

Precision and Recall range between 0 and 1. The F1 score is the harmonic mean of Precision and Recall, and thus deteriorates as soon as one of them decreases :

$$F1 = 2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \in [0,1] \quad (27)$$

These scores are known to be more robust to unbalanced sets than Receiver Operator Curves (and their associated Area Under the Curve).

F. Wasserstein distance : an Optimal Transport metrics measuring a distance between estimated and reference amplitudes

To take into account a "distance" between estimated and reference sources amplitudes, one may turn towards an Optimal Transport metrics, such as Wasserstein distance [34]. Given :

- a reference distribution $\mu_a = \sum_{i=1}^{N} a_i\delta_{x_i}$, with weights a_i (positive, sum to 1) associated with true sources amplitudes;
- an estimated distribution $\mu_b = \sum_{i=1}^{N} b_i\delta_{y_i}$ with weights b_i (positive, sum to 1) associated with estimated amplitudes;

one may compute the minimum cost (with respect to a distance d) to move the estimated distribution μ_b such that it coincides with the reference one μ_a. This minimum cost is given by the following optimization problem. Given $p \in [1, +\infty[$, set

$$W_p(d;a,b) = \min_{\gamma\in\mathbb{R}^{N\times N}}\left(\sum_{i,j}\gamma_{i,j}d(x_i,y_j)^p\right)^{\frac{1}{p}} \quad (28)$$

under the constraints that $\gamma.1_{\mathbb{R}^N} = a$ and $\gamma^T.1_{\mathbb{R}^N} = b$; $\gamma \geq 0$. The solution of the above optimization problem (28) results in the Wasserstein distance $W_p(d;a,b)$. In the above definition, γ refers to any transportation map from $\mathbb{R}^N$ to $\mathbb{R}^N$ which preserves mass between the distributions μ_a and μ_b. The "distance" between estimated and reference amplitudes is thus encapsulated within the constraints on γ, while the geometrical distance between estimated and reference sources is tracked by d. No thresholding is required in the definition of the Wasserstein distance, only a normalization of reference and estimated amplitudes. In the following numerical experiments, the 1-Wasserstein distance W_1 is used, with respect to the Euclidean distance, to assess the model performances in terms of both amplitude and localization.

V. Numerical experiments

A. Wavelet Frame design

For the current paper and the following performance evaluation section, we hereafter consider up to 3 wavelet scales, in addition to the low-frequency component recovered by the scaling function. Following [5], wavelets resulting from the band-pass kernel g are defined in a piecewise manner : cubic spline within the bandwidth, and 2nd degree polynomial increase and decrease outside. This results in a common quality factor ($Q \approx 1.38$), and rather close frame bounds 0.71 and 1.41. As can be seen on Fig. 1, the finest wavelet is designed to that its bandwidth catches the largest eigenvalue λ_{max}. All frequencies below the green bandwidth are then either encapsulated with the low-frequency component, or taken into account with additional wavelet scales, whose dilation parameters s are spread along a logarithmic scale.

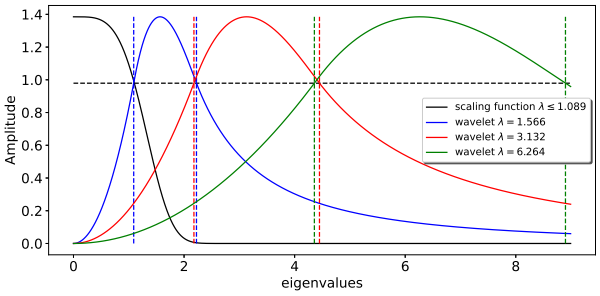

Fig. 1. Distribution of spectral graph wavelets along the graph Laplacian spectrum - vertical lines indicate the bandwidth

B. Experimental setup

In the current paper, we focus on simulated data, and for the sake of reproducibility and future comparisons, we decided to stick as much as possible to datasets and resources included in the MNE-Python software package [6].

1) Simulation protocol: Brain geometry, leadfield and baseline covariance matrices were imported from the "audiovisual" dataset, which contains EEG/MEG evoked data from one subject performing an audiovisual experiment. MNE-Python was also used to simulate a uniform brain activity over patches of two different sizes, generated using parcels from Destrieux atlas (74×2 parcels). Examples are given in Fig. 3. Time courses for source activity (see Fig. 2) were given an evoked response shape, still using MNE-python resources. A real spatial covariance matrix was imported from the MNE-python database to simulate baseline data, and added to simulated evoked sensor data. To reproduce the classical $1/f$ behavior of the noisy signal Power Spectral Density, an AR model was fitted on raw experimental data, and simulated evoked time series were filtered using the corresponding IIR filter. For each case, 320 epochs were simulated, and the trial average was used for the inverse problem.

2) Compared solutions: In the results section below, we compared a number of solutions described in Section II together with wavelet based solutions proposed in this paper. The variational reference solutions include MNE (quadratic penalization $f(\mathbf{S}) = \lambda_{\text{MNE}}\|\mathbf{S}\|_F^2$, closed form solution), MCE ($f(\mathbf{S}) = \lambda_{\text{MCE}}\|\mathbf{S}\|_1$, solved using the FISTA algorithm), sVB-SCCD ($f(\mathbf{S}) = \lambda_{\text{MCE}}\left(\gamma\|\mathbf{S}\|_1 + (1-\gamma)\|\mathbf{S}\|_{TV}\right)$, solved using the ADMM algorithm), with $\gamma = 0.40$ (L1-TV1) and $\gamma = 0.60$ (L1-TV2). These are compared with (sg)w-MCE 3 ($g(\mathbf{A}) = \lambda_{\text{MCE}}\|\mathbf{A}\|_1$, 3 wavelet scales, solved using FISTA) and (sg)w-SBL 1/2/3 (see equations (14-15)), using 1 to 3 wavelet scales and Champagne algorithm (100 iteration steps).

Remark 2 (Regularization of the noise spatial covariance matrix): despite a whitening and dimension reduction process on the noise spatial covariance matrix, most inverse problem solvers require a regularization step of the covariance matrix before solving the inverse problem ; interestingly, that is not the case for solvers using Sparse Bayesian Learning (SBL) approaches, in which the cortical source variances are "learned" through an optimization algorithm.

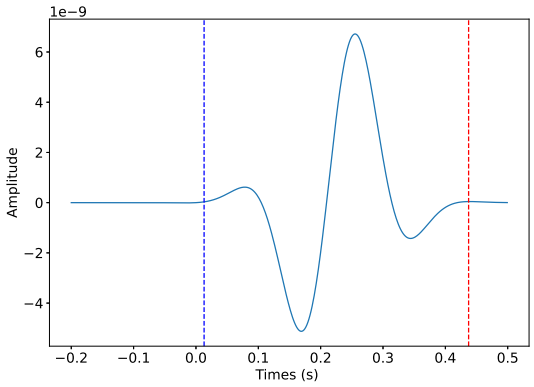

Fig. 2. Time course signal designed to mimick an evoked response

For the following model performance evaluation, all EEG and MEG channels were considered as input data for the inverse problem. Given two specific spatial extents (10 and 20 mm from each parcel center, see Fig. 3) for the simulated activity, noisy evoked-like response activities were generated

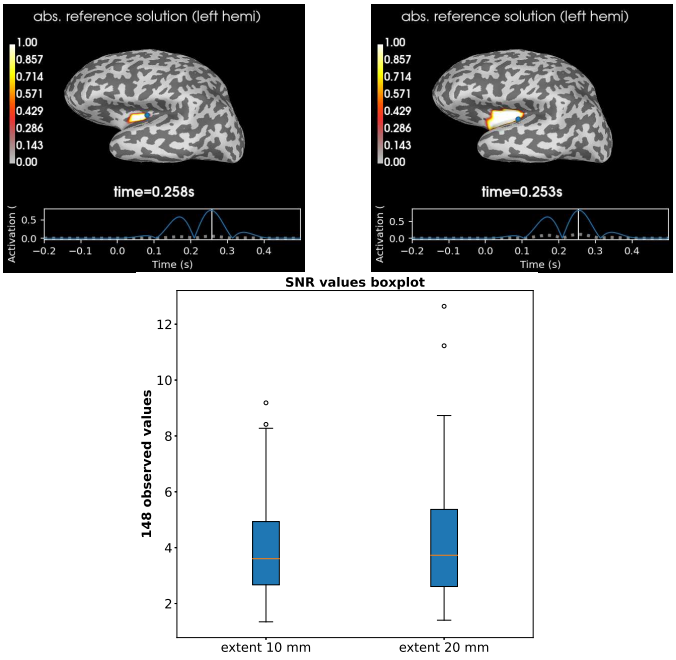

Fig. 3. Reference cortical activity # 1 (over 148) simulated from Destrieux's brain parcellation - extent 10 mm (upper left) and 20 mm (upper right) and SNR values boxplot (lower)

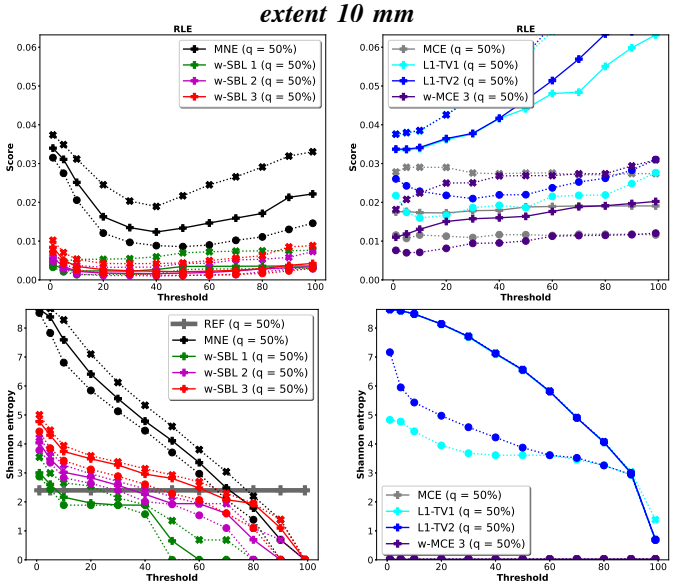

Fig. 4. RLE (upper) and entropy (lower) for the 10 mm case - quartiles q_1 (dashed), q_2 and q_3 (dashed) are represented for each solver

for each of the 148 cortical regions of Destrieux's brain parcellation, with rather similar SNR, as displayed on Fig. 3.

Given this simulation protocol, all the considered solvers were compared and challenged with respect to the previously detailed metrics. To facilitate interpretation and comparison, let us start with all threshold-based metrics.

Threshold-based metrics (without binarisation)

For all threshold-based metrics (RLE, Shannon entropy, nb. active sources and spatial extent), (sg)w-SBL outperforms significantly the other solvers, from low threshold values, as can be seen on Fig. 4 and 5 for the 10 mm case : (sg)w-SBL

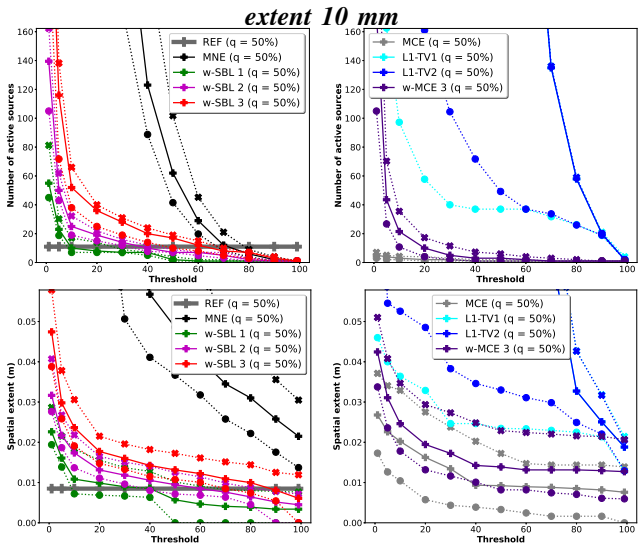

Fig. 5. Nb. active sources (upper) and spatial extent (lower) - quartiles q_1 (dashed), q_2 and q_3 (dashed) are represented for each solver

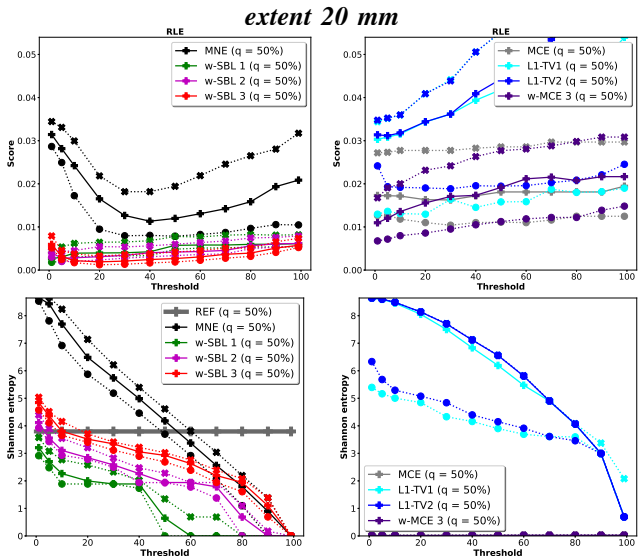

Fig. 6. RLE (upper) and entropy (lower) - quartiles q_1 (dashed), q_2 and q_3 (dashed) are represented for each solver

minimizes (with its tight q_1, q_2, q_3 quartiles) the RLE metrics, and reaches the fastest the reference value for the other metrics. Notice, however, that (sg)w-SBL 1 deviates from the reference solution when amplitude thresholding exceeds 40%. Similar results are obtained for the 20 mm case (see Fig. 6-7), where (sg)w-SBL reaches the reference extended solution even quicker. As for w-MCE 3, which promotes sparsity in the synthesis wavelet coefficients domain, its outperforms L1-TV solvers with respect to RLE and spatial extent, and MCE with respect to the number of active sources.

Geometrical metrics with amplitude weighting

Fig. 8 displays, for the 20 mm case, the boxplots obtained for two geometrical metrics with amplitude weighting : the maximum difference between reference and estimated SD vectors (δ_{SD}) and Euclidean distance between 3D centers of mass. Again (sg)w-SBL clearly minimizes these metrics.

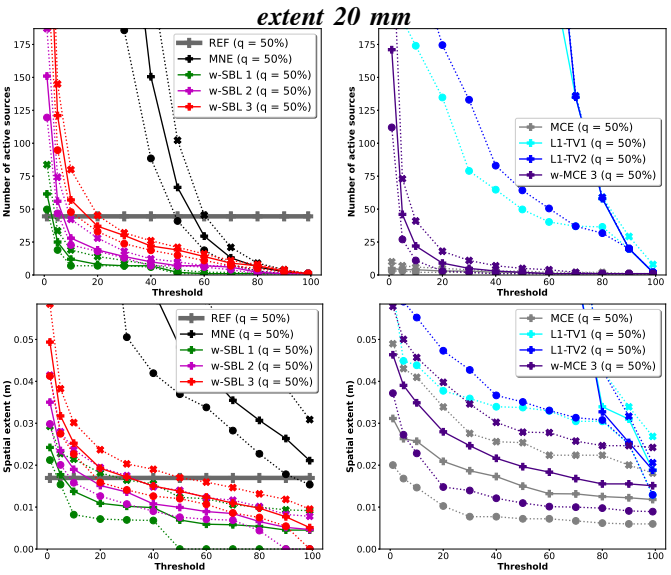

Fig. 7. Nb. active sources (upper) and spatial extent (lower) - quartiles q_1 (dashed), q_2 and q_3 (dashed) are represented for each solver

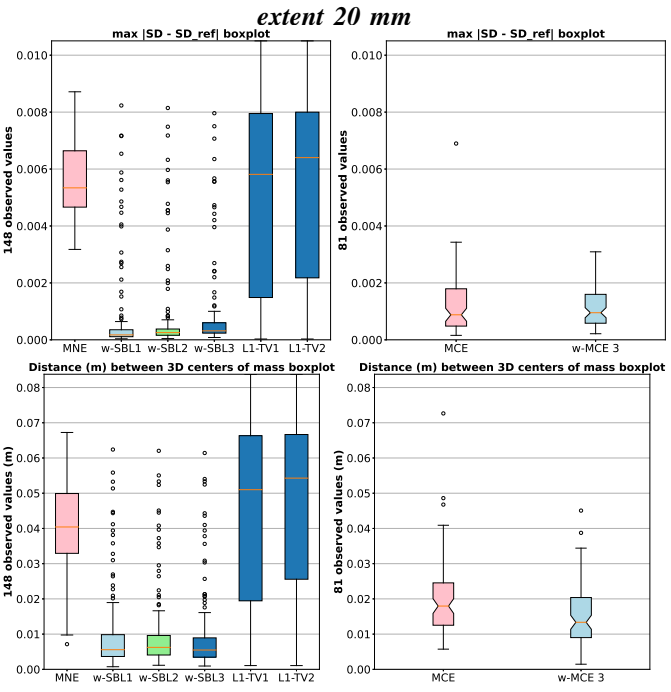

Fig. 8. Boxplots for geometrical metrics with amplitude-weighting : δ_{SD} (upper) and distance between 3D centers of mass (lower)

Furthermore, given the small distance between centers of mass obtained with (sg)w-SBL, the depth metrics is less affected by the skull curvature. w-MCE 3 solver comes in second position behind (sg)w-SBL, improving a bit MCE results.

Fig. 9 focuses on the depth metrics for the estimated centers of mass (here restricted to the cortex graph). Again, (sg)w-SBL exhibits the best performances, in both cases. Note that one retrieves the classical MNE depth bias towards superficial solutions. L1-TV (1 and 2) solvers also exhibit very unsatisfactory results, with an even stronger bias.

Fig. 10 displays the evolution of the difference between reference and estimated depth (for a c.o.m. restricted to the

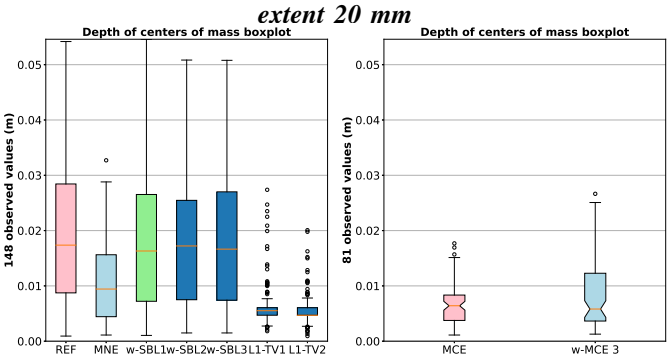

Fig. 9. Boxplots for geometrical metrics with amplitude-weighting : depth of centers of mass (restricted to the graph)

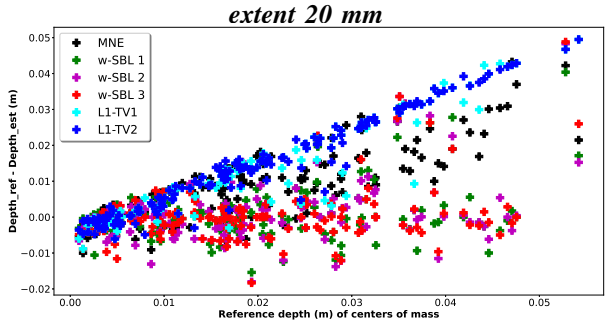

Fig. 10. Depths of centers of mass (restricted to the graph) with respect to reference depth

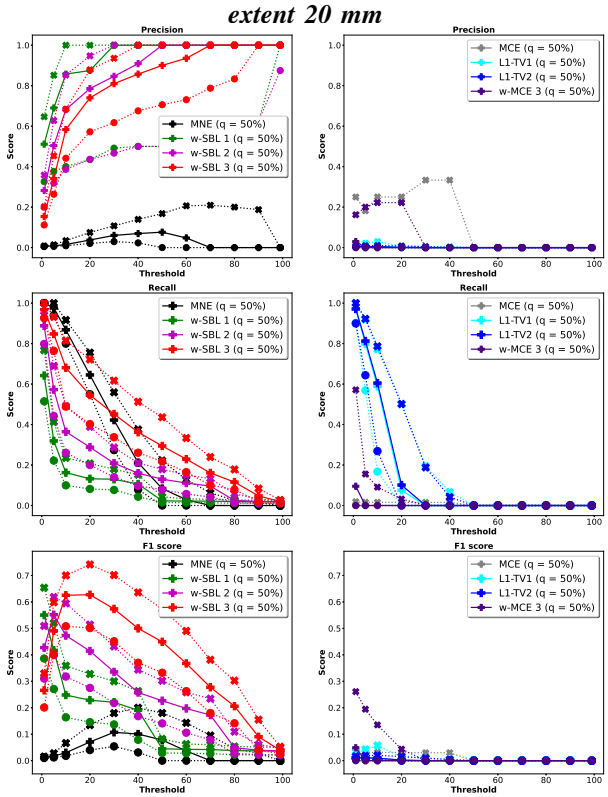

Fig. 11. Statistical scores with respect to threshold

graph) with respect to the reference depth. While MNE and L1-TV solvers clearly exhibit a linearly evolving bias with the reference depth, (sg)w-SBL shows very good and stationary performances throughout all the tested samples and depths. For clarity, MCE and w-MCE 3, which did not exhibit good results in Fig. 9, were not included in Fig. 10.

Binary classifier metrics

Now considering binary statistical maps built from reference and estimated energy maps, Fig. 11 displays the Precision, Recall and F1 score of the multiple tested solvers. (sg)w-SBL clearly exhibits, by far, and from low threshold values, the best performances with respect to Precision and F1 score. Both scores are rapidly increasing with the threshold value. The recall decay is also more contained with (sg)w-SBL.

1-Wasserstein distance between amplitudes maps

To assess solvers performances with respect to amplitude reconstruction, Fig. 12 displays, for the 20 mm case, boxplots for the 1-Wasserstein distance between normalized reference and estimated amplitudes maps. Once again, (sg)w-SBL minimizes by far this amplitude and geometry-based metrics. Note that L1-TV best results with respect to the 1-Wasserstein distance are better than (sg)w-SBL, but once again the huge results dispersion undermines the robustness of this solver.

l^2-norm conservation between amplitudes maps

To further evaluate solvers quantitative performance with respect to amplitude, Fig. 13 displays the ratio of l^2 norms between (un-normalized) reference and estimated amplitudes maps averaged on a time segment of interest $\mathcal{A} = \sqrt{\langle \mathbf{S}_*^2 \rangle_{soi}}$.

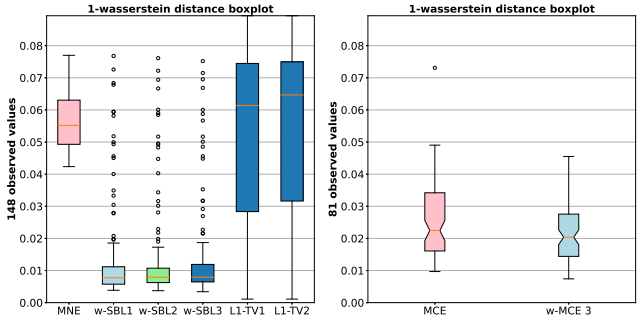

Fig. 12. 1-Wasserstein distance boxplot - extent 20 mm

Similarly to the Wasserstein metrics, (sg)w-SBL (especially with 3 wavelet scales) again outperforms the other solvers, with tight q_1, q_2, q_3 quartiles very close to the ideal unit value. w-MCE 3 then yields the second best results after (sg)w-SBL with respect to this purely amplitude-based metrics.

The impact of SNR on L1-TV solvers and hyperparameters

Finally, in order to better comprehend the surprisingly unsatisfactory results obtained with L1-TV solvers, one may investigate the impact of SNR on the solvers performances. Fig. 14 displays SNR impact on the 1-Wasserstein distance between normalized amplitudes maps. Thus, given the considered heuristics for the hyperparameters involved in L1-TV solvers, higher SNR values ($\geq$ 4) appear necessary to see results improve from the MNE solution, while (sg)w-SBL performances take the lead from low SNR values.

Visualization of a reconstructed activity

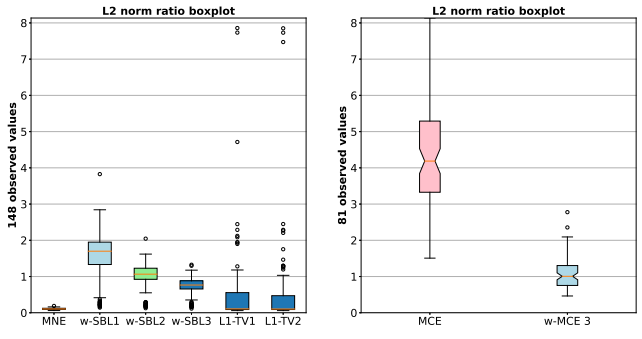

Fig. 13. l^2-norm ratio boxplot on amplitudes $\mathcal{A} = \sqrt{\langle \mathbf{S}_*^2 \rangle_{soi}}$ - extent 20 mm

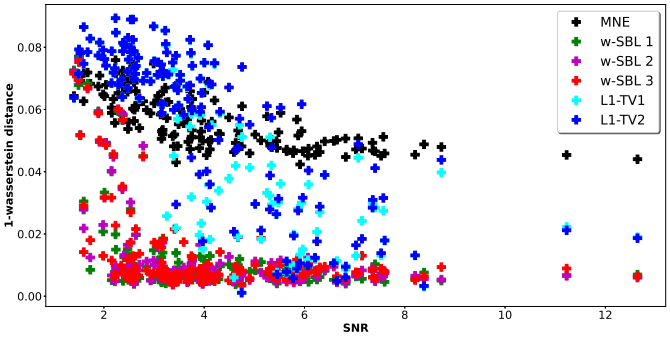

Fig. 14. 1-Wasserstein distance with respect to SNR - extent 20 mm

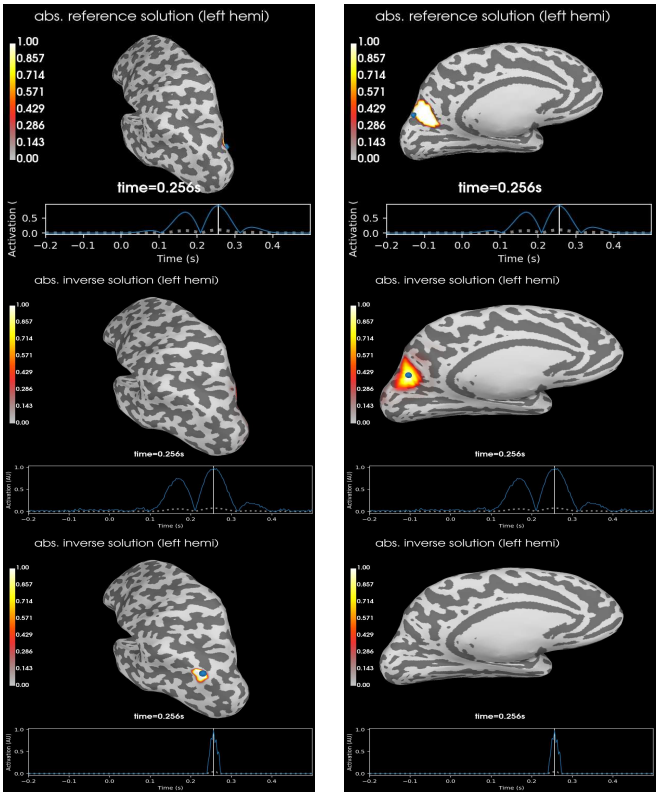

Fig. 15. Reference (upper) and reconstructed (lower : w-SBL 3 then L1-TV2) activity - patch 128 - parietal and medial views - extent 20 mm

Finally, Fig. 15 displays the visualization of a reference and reconstructed activity (time-dependent absolute amplitude), for patch #128 (SNR = 4.14, 20 mm extent case), which corresponds to the first quartile for the w-SBL 3 solver with respect to the 1-Wasserstein distance. For snapshots taken at peak-amplitude time, (sg)w-SBL recovers the true source location, with a good spatial extent and a time course close to the reference one. L1-TV solution deviates notably from the true location (from medial to parietal view), and is not able to reproduce the reference time course.

C. Discussion

From these extensive simulations of extended (non-focal) sources, one may note that, as expected, (w)MNE tends to produce source estimates that are too diffuse, and require post-processing (not addressed here) to yield fairly reliable source localization. MCE produces source estimates that are extremely focal, and is therefore not suitable for extended sources. Surprizingly enough, L1-TV approaches such as sVB-SCCD turn out to be extremely unstable, and often yield source estimates that are too extended and superficial. The simulation protocol was expected to favor these approaches. Our interpretation is that the two hyper-parameters involved there are difficult to tune, and the heuristics used in this work is not sufficient (to our knowledge, no systematic approach has been proposed in the literature). Wavelet-based solvers are indeed able to reconstruct such sources, more accurately and robustly than the other tested approaches. More specifically, (sg)w-MCE (ℓ^1 penalization on wavelet coefficients) yields fairly good estimates, that compare favourably to the state of the art for most evaluation metrics tested here. It however requires choosing an hyperparameter value, for which the heuristics derived in this paper seems relevant. The sparse Bayesian learning (sgw-SBL) approach, which provides an estimate for the wavelet covariance matrix together with wavelet coefficients, turns out to outperform consistently and significantly other approaches, with respect to a set of metrics designed to capture various aspects and possible artifacts. Quite remarkably, the sgw-SBL approach tends to perform quite well in terms of depth bias, which is a major difficulty in the M/EEG inverse problem.

We also plan to study similar approaches that use (orthonormal) wavelet bases on the cortical surface, instead of overcomplete wavelet frames as we do here. In the sgw case, scaling coefficients (low spatial frequencies) turn out to be almost systematically set to zero or very small values by the solvers (sgw-MCE or sgw-SBL). While relevant in practice, we conjecture that this could be a side effect of the redundancy of the wavelet frame, which we would like to confirm. We also plan to study the problem of hyperparameters tuning, in particular for L1-TV models.

VI. Conclusions

This paper proposed a new approach for the M/EEG inverse problem, in the framework of the distributed sources model, which exploits sparsity in the spectral graph wavelet (spatial) domain. Such wavelets forming an overcomplete system, wavelet expansion of the solution is not unique. We have chosen to focus on the so-called synthesis approach, where

optimization is performed with respect to wavelet coefficients. In variational approaches, the unknown (solutions, or wavelet coefficients) are obtained as solutions of an optimization problem, for which hyperparameters have to be chosen. The sparse Bayesian learning paradigm was adapted to the considered wavelet setting, leading to an optimization problem where hyperparameters are also optimized. These approaches were compared with standard ones using numerical simulations of joint EEG/MEG data provided by the MNE-python software. These extensive simulations and the numerous and discussed evaluation metrics show that the combination of spatial wavelets with sparse Bayesian learning brings a real improvement for extended brain sources reconstruction, both in terms of localization and quantitative estimation.

In a forthcoming paper, we will exploit the proposed solutions on real data for specific applications, and investigate evaluation of M/EEG source localization using companion intracerebral data.

Being able to identify localization, depth and time course of brain activity is important in clinical applications for all diseases that require source localization from MEG or EEG measurements. Our results show that the proposed approach improves significantly the state of the art for the reconstruction of extended brain sources.

REFERENCES

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

[2] S. Baillet, *Encyclopedia of Computational Neuroscience*. New York, NY: Springer New York, 2014, ch. Forward and Inverse Problems of MEG/EEG, pp. 1–8. [Online]. Available: https://doi.org/10.1007/978-1-4614-7320-6_529-1

[3] A. Sorrentino and M. Piana, "Inverse modeling for MEG/EEG data," *bioRxiv*, 2017. [Online]. Available: https://www.biorxiv.org/content/early/2017/05/08/135269

[4] W.-T. Chang *et al.*, "Spatially sparse source cluster modeling by compressive neuromagnetic tomography," *NeuroImage*, vol. 53, no. 1, pp. 146–160, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811910007196

[5] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129 – 150, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1063520310000552

[6] A. Gramfort *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, p. 267, 2013. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2013.00267

[7] S. Mokhtari *et al.*, "Sparse wavelet-based solutions for the M/EEG inverse problem," in *SampTA, International Conference on Sampling Theory and Applications*, 2023. [Online]. Available: https://arxiv.org/abs/2306.15262

[8] J. Mosher, P. Lewis, and R. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 6, pp. 541–557, June 1992.

[9] K. Jerbi *et al.*, "Localization of realistic cortical activity in MEG using current multipoles," *NeuroImage*, vol. 22, no. 2, pp. 779 – 793, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811904001028

[10] B. Van Veen *et al.*, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, pp. 867–880, 1997.

[11] K. Sekihara and S. S. Nagarajan, *Adaptive Spatial Filters for Electromagnetic Brain Imaging*, ser. Series in Biomedical Engineering, J. H. Nagel, Ed. Springer-Verlag Berlin Heidelberg, 2008.

[12] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical and Biological Engineering & Computing*, vol. 32, no. 1, pp. 35–42, 1994.

[13] M. Fuchs *et al.*, "Linear and nonlinear current density reconstructions," *Journal of clinical Neurophysiology*, vol. 16, no. 3, pp. 267–295, 1999.

[14] F.-H. Lin *et al.*, "Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates," *NeuroImage*, vol. 31, no. 1, pp. 160 – 171, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811905024973

[15] R. D. Pascual-Marqui, "Discrete, 3d distributed, linear imaging methods of electric neuronal activity. part 1: exact, zero error localization," 2007.

[16] A. M. Dale *et al.*, "Dynamic statistical parametric mapping: Combining fmri and meg for high-resolution imaging of cortical activity," *Neuron*, vol. 26, no. 1, pp. 55–67, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627300811381

[17] R. Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (SLORETA): Technical details," *Methods and findings in experimental and clinical pharmacology*, vol. 24 Suppl D, pp. 5–12, 2002.

[18] H. Becker *et al.*, "A performance study of various brain source imaging approaches," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 5869–5873.

[19] K. Matsuura and Y. Okabe, "Selective minimum-norm solution of the biomagnetic inverse problem," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 6, pp. 608–615, June 1995.

[20] K. Uutela, M. Hämäläinen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, no. 2, pp. 173–180, 1999.

[21] L. Ding, "Reconstructing cortical current density by exploring sparseness in the transform domain," *Physics in Medicine and Biology*, vol. 54, no. 9, pp. 2683–2697, apr 2009. [Online]. Available: https://doi.org/10.1088/0031-9155/54/9/006

[22] H. Becker *et al.*, "Fast, variation-based methods for the analysis of extended brain sources," in *22nd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 41–45.

[23] Y. Li *et al.*, "s-smooth: Sparsity and smoothness enhanced EEG brain tomography," *Frontiers in Neuroscience*, vol. 10, p. 543, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2016.00543

[24] W. Ou, M. S. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932 – 946, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811908007155

[25] A. Gramfort, M. Kowalski, and M. Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in Medicine and Biology*, vol. 57, no. 7, pp. 1937–1961, mar 2012. [Online]. Available: https://doi.org/10.1088/0031-9155/57/7/1937

[26] A. Gramfort *et al.*, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410 – 422, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811912012372

[27] D. Wipf and S. Nagarajan, "A unified bayesian framework for MEG/EEG source imaging," *NeuroImage*, vol. 44, no. 3, pp. 947–966, Feb. 2009. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/18602278

[28] D. P. Wipf, B. D. Rao, and S. Nagarajan, "Latent variable bayesian models for promoting sparsity," *IEEE Trans. Inf. Theor.*, vol. 57, no. 9, p. 6236–6255, sep 2011. [Online]. Available: https://doi.org/10.1109/TIT.2011.2162174

[29] A. Hashemi *et al.*, "Unification of sparse bayesian learning algorithms for electromagnetic brain imaging with the majorization minimization framework," *NeuroImage*, vol. 239, p. 118309, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811921005851

[30] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004. [Online]. Available: https://doi.org/10.1198/0003130042836

[31] C. Amblard, E. Lapalme, and J.-M. Lina, "Biomagnetic source detection by maximum entropy and graphical models," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 3, pp. 427–442, March 2004.

[32] J.-M. Lina *et al.*, "Wavelet-based localization of oscillatory sources from magnetoencephalography data," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 8, pp. 2350–2364, 2014.

[33] M.-C. Roubaud *et al.*, "Space-Time Extension of the MEM Approach for Electromagnetic Neuroimaging," in *IEEE International Workshop on Machine Learning for Signal Processing*, Aalborg, Denmark, Sep. 2018. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01843557

[34] S. Kolouri *et al.*, "Optimal mass transport: Signal processing and machine-learning applications," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 43–59, 2017.