NeuroImage 118 (2015) 598–612

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# Solving the EEG inverse problem based on space–time–frequency structured sparsity constraints

Sebastián Castaño-Candamil [a,c,*], Johannes Höhne [b], Juan-David Martínez-Vargas [c], Xing-Wei An [d], German Castellanos-Domínguez [c], Stefan Haufe [e,f,**]

[a] BrainLinks-BrainTools, Albert-Ludwig Universität Freiburg, Germany
[b] Neurotechnology Group, Technische Universität Berlin, Germany
[c] Signal Processing and Recognition Group, Universidad Nacional de Colombia, Colombia
[d] Department of Biomedical Engineering, Tianjin University, PR China
[e] Laboratory for Intelligent Imaging and Neural Computing, Columbia University in the City of New York, USA
[f] Machine Learning Group, Technische Universität Berlin, Germany

ARTICLE INFO

*Article history:*
Received 28 December 2014
Accepted 7 May 2015
Available online 3 June 2015

*Keywords:*
EEG
MEG
Inverse problem
Spatio-temporal priors
Structured sparsity
Non-stationarity

ABSTRACT

We introduce STOUT (spatio-temporal unifying tomography), a novel method for the source analysis of electroencephalograpic (EEG) recordings, which is based on a physiologically-motivated source representation. Our method assumes that only a small number of brain sources are active throughout a measurement, where each of the sources exhibits focal (smooth but localized) characteristics in space, time and frequency. This structure is enforced through an expansion of the source current density into appropriate spatio-temporal basis functions in combination with sparsity constraints. This approach combines the main strengths of two existing methods, namely Sparse Basis Field Expansions (Haufe et al., 2011) and Time–Frequency Mixed-Norm Estimates (Gramfort et al., 2013). By adjusting the ratio between two regularization terms, STOUT is capable of trading temporal for spatial reconstruction accuracy and vice versa, depending on the requirements of specific analyses and the provided data. Due to allowing for non-stationary source activations, STOUT is particularly suited for the localization of event-related potentials (ERP) and other evoked brain activity. We demonstrate its performance on simulated ERP data for varying signal-to-noise ratios and numbers of active sources. Our analysis of the generators of visual and auditory evoked N200 potentials reveals that the most active sources originate in the temporal and occipital lobes, in line with the literature on sensory processing.

## Introduction

In recent years, the advances in Neuroscience have led to a better understanding of cognitive processes in the human brain. One general goal is to identify brain areas related to certain cognitive processes or pathologies through measurements. Methods allowing for such kind of analyses are called neuroimaging techniques.

At the same time, it is also of importance to identify and characterize temporal brain activation patterns related to the cognitive phenomena under study. Magneto- and electroencephalography (MEG and EEG) have been widely used to study brain dynamics by identifying and analyzing temporal activation patterns, e.g., neural rhythms, event-related potentials (ERP), epileptic spikes, among others (e. g., Michel et al., 2004; Galka et al., 2004; Blankertz et al., 2011). M/EEG recordings also contain spatial information, because they are usually measured over the entire scalp using up to a few hundred sensors. Consequently, EEG and MEG have also been used as neuroimaging techniques (e. g., Zwoliński et al., 2010; Baillet et al., 2001; Toga and Mazziotta, 2002). While most of the considerations made here equally apply to MEG, we restrict the discussion to EEG in the following.

The pyramidal neurons believed to account for most of the EEG signal populate the entire cortical gray matter, and outnumber the available sensors by several orders of magnitude. Methods for estimating the generators of EEG activity therefore need to consider at least a few thousand potentially contributing brain sites as potential sources, which may be distributed evenly across the brain, or restricted to the cortical gray matter. Estimating the source distribution of brain electrical activity based on EEG measurements therefore amounts to solving

[*] Correspondence to: S. Castaño-Candamil, BrainLinks-BrainTools, Albert-Ludwig Universität Freiburg, Germany.

[**] Correspondence to: S. Haufe, Laboratory for Intelligent Imaging and Neural Computing, Columbia University in the City of New York, USA.

*E-mail addresses:* sebastian.castano@blbt.uni-freiburg.de (S. Castaño-Candamil), sh3397@columbia.edu (S. Haufe).

http://dx.doi.org/10.1016/j.neuroimage.2015.05.052
1053-8119/© 2015 Elsevier Inc. All rights reserved.

an ill-posed and mathematically underdetermined inverse problem, where a unique solution can only be obtained by making additional assumptions (Baillet et al., 2001; Galka et al., 2004; Grech et al., 2008; Babadi et al., 2014). This can for example be done by way means of introducing prior beliefs on the structure of possible source configurations in a Bayesian inference framework (Nummenmaa et al., 2007).

With respect to neurophysiological plausibility, it has been argued that solutions with a simple spatial structure may be favored. There are various algorithmic approaches to enforce simplicity. The minimum norm estimate (MNE, Hämäläinen and Ilmoniemi, 1994), for example, minimizes the overall power of the sources, whereas the Low Resolution Tomography estimate (LORETA, Pascual-Marqui et al., 1994) explicitly enforces spatial smoothness of the sources based on the argument that neighboring voxels should be similarly active. Technically, both approaches can be implemented using $\ell_2$-norm penalties. On the other hand, it has also been argued that, in event-related experimental designs, only a small fraction of the brain should be consistently activated. Consequently, methods assuming sparsity in the spatial domain have been proposed (Gorodnitsky et al., 1995; Matsuura and Okabe, 1995; Grech et al., 2008; Bolstad et al., 2009; Wipf and Nagarajan, 2009; Ou et al., 2008). Sparse methods are often based on the minimization of $\ell_1$-norm regularization terms or, in a more general sense, on the minimization of the volume spanned by the active coefficients of the sources.

While being physiologically motivated, all these solutions practically suffer from undesired properties, which include spatial blurring—and the resulting inability to spatially separate multiple sources—, the presence of so-called ghost sources for minimum $\ell_2$-norm solutions, as well as instability and spatial scattering for minimum $\ell_1$-norm solutions (Haufe et al., 2008b, 2011; Grech et al., 2008; Tibshirani, 1994). To overcome these issues, several authors have proposed to combine spatial smoothness and sparsity to obtain *focal* source activations, be it through a combination of penalty terms (see Haufe et al., 2008a,b; Vega-Hernández et al., 2008), or through representing brain activity as the sum of a small number of spatial basis functions describing smooth localized patches of potentially active brain regions (Friston et al., 2008; Haufe et al., 2008a, 2011).

Besides enforcing a preferred spatial structure, prior information may also be included in the form of temporal constraints describing dynamics of neural activity. Specifically, it has been shown that time-frequency representations provide insightful information about the dynamics of neural processes (Miwakeichi et al., 2004; Durka et al., 2005; Trujillo-Barreto et al., 2008; Gramfort et al., 2013). Generally, brain activity may be non-stationary (e. g., event-related), which is not taken into account by classical methods. In contrast, Gramfort et al. (2013) address the non-stationarity issue by representing brain activity through a sparse set of time-frequency basis functions (atoms).

The vast majority of inverse methods for neuroimaging employ constraints either in the spatial or temporal domain, but not simultaneously in both domains. Thus, some methods are able to accurately describe non-stationary brain activations (e. g., Gramfort et al., 2013, TF-MxNE), but their solutions may be too focal; that is, solutions are not composed of smooth activation patches, but of non-contiguous spikes of activation. The opposite holds for other methods (e. g., Haufe et al., 2011, S-FLEX) that enforce spatial focality while being unable to describe non-stationary brain activations. Here, we propose to fill this gap by enforcing neurophysiologically motivated structure both in time and space, and thereby to unify the advantages of S-FLEX and TF-MxNE. Precisely, we propose a spatio-temporal decomposition of source activations, which depends on three components: (1) a predefined dictionary of spatial basis fields, (2) a predefined dictionary of temporal basis functions, and (3) a matrix of spatio-temporal coefficients that needs to be estimated. By adopting spatial and temporal "dictionaries" from Gramfort et al. (2013) and Haufe et al. (2011), our method — termed spatio-temporal unifying tomography (STOUT) — is able to reconstruct the time courses of potentially non-stationary source activations with focal spatial topographies. Moreover, by enforcing sparse structure through a weighted combination of spatial and temporal penalty terms, our method is able to "trade" spatial focality for a simpler time-frequency representation, and vice versa. This tradeoff is quantified by a single hyperparameter that allows to access to an entire spectrum of solutions ranging between S-FLEX and TF-MxNE.

The present manuscript is organized as follows. In the Methods section, we give an introduction to the EEG inverse problem and present existing solutions as well as our novel source imaging method STOUT. In the Experiments and Results sections, we assess the reconstruction of simulated ERP activity using STOUT as compared to state-of-the-art source imaging approaches. We also apply STOUT to real EEG data, where the task is to localize the generators of auditory and visual evoked potentials recorded during an oddball experiment. Then, we discuss the properties of our method in the Discussion section, and conclude our contributions in the Conclusion section.

## Methods

### *EEG forward and inverse problem*

The electromagnetic field measured by EEG may be represented by the following linear model (Baillet et al., 2001; Grech et al., 2008):

$$\boldsymbol{Y} = \boldsymbol{L}\boldsymbol{J} + \epsilon. \tag{1}$$

Here, $\boldsymbol{Y} \in \mathbb{R}^{N_c \times N_t}$ is the EEG data measured at a set of $N_c$ sensors at $N_t$ time points, $\boldsymbol{J} \in \mathbb{R}^{3N_d \times N_t}$ (termed the *current density*) is the corresponding brain source activity matrix holding the 3D current vectors of $N_d$ dipolar electrical brain sources at the $N_t$ time points, and $\boldsymbol{L} \in \mathbb{R}^{N_c \times 3N_d}$ (the *lead field*) is a gain matrix representing the relationship between the current sources $\boldsymbol{J}$ and the measured EEG data $\boldsymbol{Y}$, composed as $\boldsymbol{L} = [\boldsymbol{L}_x, \boldsymbol{L}_y, \boldsymbol{L}_z]$, where the matrices $\boldsymbol{L}_{x/y/z}$ are the lead fields of the current sources in each direction $x$, $y$ and $z$, respectively. We also assume that $\boldsymbol{Y}$ is affected by Gaussian distributed noise $\epsilon \in \mathbb{R}^{N_c \times N_t}$ with covariance $cov(\epsilon) = \boldsymbol{Q}_\epsilon \in \mathbb{R}^{N_c \times N_c}$, where $\boldsymbol{Q}_\epsilon$ is the noise covariance matrix. In practice, $\boldsymbol{Q}_\epsilon$ can be estimated from data using baseline measurements (Nagarajan et al., 2007), be derived from the lead field (assuming i.i.d. source activations), or simply set to the identity matrix. The latter approach is applied in the present work. Under this model, the maximum a-posteriori (MAP) estimate of $\boldsymbol{J}$ can be found as the minimizer of the following cost function, which is composed of a quadratic error term and a regularization term (Grech et al., 2008):

$$\underset{\boldsymbol{J}}{\operatorname{argmin}} \left\{ ||\boldsymbol{Y} - \boldsymbol{L}\boldsymbol{J}||^2_{\boldsymbol{Q}_\epsilon} + \lambda \Theta(\boldsymbol{J}) \right\}. \tag{2}$$

Here, $\|\boldsymbol{P}\|_{\boldsymbol{Q}_\epsilon} = \sqrt{\operatorname{tr}\left\{\boldsymbol{P}^T \boldsymbol{Q}_\epsilon{}^{-1} \boldsymbol{P}\right\}}$ denotes the Mahalanobis distance, $\lambda \in \mathbb{R}^+$ is a regularization constant, and $\Theta(\boldsymbol{J}) \in \mathbb{R}^+$ is a function which formalizes the constraints that are imposed upon the source activity.

### *Existing inverse solutions*

The penalty function $\Theta(\boldsymbol{J})$ is commonly used to promote solutions with a certain spatial or temporal structure. Solution with purely smooth as well as purely sparse source activations have been argued to be neurophysiologically plausible (Hämäläinen and Ilmoniemi, 1994; Pascual-Marqui et al., 1994; Gorodnitsky et al., 1995; Matsuura and Okabe, 1995). An example of a spatially smooth method is the Low Resolution Tomography (LORETA) estimate (Pascual-Marqui et al., 1994), while an example of a spatially sparse method is the minimum-current or least absolute shrinkage and selection operator (LASSO) estimate (Matsuura and Okabe, 1995).

### *Sparse Basis Field Expansions*

In practice, methods encouraging *focal* solutions (which are smooth and focal at the same time) have been found to yield better source reconstructions than purely smooth or sparse approaches (Haufe et al., 2008b, 2011). One such method, termed Sparse Basis Field Expansions (S-FLEX), expresses the current density as a linear combination of locally smooth but spatially confined *spatial basis functions* (such as Gaussian curves) via $\boldsymbol{J} = (\boldsymbol{\Phi}_s \otimes \boldsymbol{I}_3)\boldsymbol{C}_s$ (Haufe et al., 2011). Here, the matrix $\boldsymbol{\Phi}_s \in \mathbb{R}^{N_d \times N_s}$ holds $N_s$ spatial basis functions, $\boldsymbol{I}_3$ is the 3 × 3 identity matrix, ⊗ denotes the Kronecker product, and $\boldsymbol{C}_s \in \mathbb{R}^{3N_s \times N_t}$ is a matrix of weighting coefficients to be estimated. The assumption of S-FLEX is that the current density can be represented using only a few spatial basis functions. Therefore, $\boldsymbol{C}_s$ is estimated under sparsity constraints using the following objective function

$$\arg\min_{\boldsymbol{C}_s} \left\{ ||\boldsymbol{Y} - \boldsymbol{L}\boldsymbol{\Phi}_s\boldsymbol{C}_s||^2_{\boldsymbol{Q}_\epsilon} + \lambda||\boldsymbol{C}_s||_{1,2} \right\}, \tag{3}$$

where the notation $|| \cdot ||_{1,2}$ stands for the so-called $\ell_{1,2}$ or group-LASSO norm and it is defined as $||\boldsymbol{C}_s||_{1,2} = \sum_{i=1}^{N_s} \left\|\boldsymbol{C}_s^{(i,\cdot)}\right\|_F$, where $\|\boldsymbol{C}_s^{(i,\cdot)}\|_F$ is the Frobenius norm of the $3 \times N_t$ matrix defining the coefficients of the i-th dipole at each time point for each of 3 spatial coordinates.

Thus, S-FLEX groups the $3N_t$ coefficients related to each of the $N_s$ basis functions under a common non-sparse $\ell_2$-norm, whereas the sparsity-inducing $\ell_1$-norm is used across basis functions in order to select a subset of these functions.

Importantly, by using the non-sparse $\ell_2$-norm penalty along the temporal direction, S-FLEX encourages sources that are consistently activated throughout the analyzed time window. In other words, it assumes stationarity of the source activations.

### *Time–Frequency Mixed Norm Estimates*

When the spatial profile of the source activity of interest changes within the analysis window (as expected in event-related experimental designs), the penalty function used in S-FLEX may be inappropriate. To model non-stationary source activations, Gramfort et al. (2013) propose the Time–Frequency Mixed Norm Estimate (TF-MxNE), which is obtained by decomposing the current density into *temporal basis functions* $\boldsymbol{\Phi}_t \in \mathbb{C}^{N_f \times N_t}$ and corresponding coefficients $\boldsymbol{C}_t \in \mathbb{C}^{3N_d \times N_f}$ as $\boldsymbol{J} = \boldsymbol{C}_t\boldsymbol{\Phi}_t$. The basis functions are so-called time-frequency atoms, which are bounded both in their temporal and spectral extent. Possible choices include wavelets and Fourier basis functions. To promote spatial sparsity on the level of individual dipoles while simultaneously promoting the selection of only a subset of the time-frequency atoms, the following constraint is used: $\Theta(\boldsymbol{C}_t) = \lambda_s\|\boldsymbol{C}_t\|_{1,2} + \lambda_t\|\boldsymbol{C}_t\|_1$, where $||\boldsymbol{C}_t||_1 = \sum_{i=1}^{3N_d}\sum_{j=1}^{N_f}|\boldsymbol{C}_t^{(i,j)}|$ is the $\ell_1$-norm of matrix $\boldsymbol{C}_t$, and $\lambda_s \in \mathbb{R}^+$ and $\lambda_t \in \mathbb{R}^+$ are regularization parameters adjusting the sparsity along the spatial and temporal directions. Thus, the objective function takes the form

$$\arg\min_{\boldsymbol{C}_t} \left\{ ||\boldsymbol{Y} - \boldsymbol{L}\boldsymbol{C}_t\boldsymbol{\Phi}_t||^2_{\boldsymbol{Q}_\epsilon} + \lambda_s||\boldsymbol{C}_t||_{1,2} + \lambda_t||\boldsymbol{C}_t||_1 \right\}. \tag{4}$$

The TF-MxNE estimate is capable of representing non-stationary source activity at each voxel in the brain as a linear combination of a small number of time-frequency atoms using a combination of $\ell_2$- and $\ell_1$-norm penalties along the temporal domain. However, by using the $\ell_1$-norm along the spatial dimension, it may exhibit the same spatial instability and scattering observed for other estimators based on spatial sparsity.

### *Spatio-temporal unifying tomography (STOUT)*

Here, we propose to overcome the spatial instability of TF-MxNE in the same way S-FLEX does for traditional approaches involving spatial sparsity, namely by means of an expansion into focal spatial basis functions. Our method, spatio-temporal unifying tomography (STOUT) thereby inherits the favorable spatial structure of S-FLEX estimates, while being able to represent and estimate brain activity with non-stationary temporal activity profiles.

### *Source model and cost function*

In analogy to the representations used in S-FLEX and TF-MxNE, STOUT models the current density as a linear combination of *spatio-temporal basis functions*, each of which is the outer product of a smooth spatially localized Gaussian radial basis function and a temporally and spectrally localized time-frequency atom. The expansion can be written as

$$\boldsymbol{J} = (\boldsymbol{\Phi}_s \otimes \boldsymbol{I}_3)\boldsymbol{C}_{st}\boldsymbol{\Phi}_t, \tag{5}$$

where $\boldsymbol{C}_{st} \in \mathbb{C}^{3N_s \times N_f}$ is a matrix holding three coefficients (one for each current orientation) for each of the spatio-temporal dictionary elements. As in TF-MxNE, a selection of a small number of these elements is achieved through the following penalized likelihood function including sparsity constraints

$$\arg\min_{\boldsymbol{C}_{st}} \left\{ ||\boldsymbol{Y} - \boldsymbol{L}(\boldsymbol{\Phi}_s \otimes \boldsymbol{I}_3)\boldsymbol{C}_{st}\boldsymbol{\Phi}_t||^2_{\boldsymbol{Q}_\epsilon} + \lambda_s||\boldsymbol{C}_{st}||_{1,2} + \lambda_t||\boldsymbol{C}_{st}||_1 \right\}. \tag{6}$$

As a result of selecting a small subset of the spatio-temporal atoms, STOUT fosters the reconstruction of brain activity, which is smooth but well localized in space, time and frequency as illustrated in Fig. 1.

The objective function (Eq. (6)) can be optimized using the Fast Shrinkage Thresholding Algorithm (FISTA, Beck and Teboulle, 2009; Gramfort et al., 2012), as shown in Appendix A.

### *Construction of the spatio-temporal basis functions*

For the definition of the spatial and temporal dictionaries $\boldsymbol{\Phi}_s$ and $\boldsymbol{\Phi}_t$, we closely follow Haufe et al. (2011) and Gramfort et al. (2013). Since we are working with discrete source spaces, spatial basis functions only need to be evaluated at each modeled dipole location. We here consider Gaussians centered at each dipole location. Denoting the $i$-th location by $\boldsymbol{x}_{(i)}$, the basis function centered at $\boldsymbol{x}_{(i)}$ and evaluated at $\boldsymbol{x}_{(j)}$ is given by

$$\boldsymbol{\Phi}_s(i,j) = exp\left\{ -d_g\{\boldsymbol{x}_{(i)}, \boldsymbol{x}_{(j)}\}^2 / \sigma^2 \right\}.$$

Here, we consider all dipoles to be located on the folded cortical surface. The metric $d_g\{\boldsymbol{x}_{(i)}, \boldsymbol{x}_{(j)}\} \in \mathbb{R}^+$ therefore refers to the geodesic distance between the $i$-th and the $j$-th dipole along the cortical surface. For discrete cortical meshes, as considered here, the geodesic distance can be computed using algorithms for finding the shortest paths between all pairs of nodes in a graph as implemented by the freely available MatlabBGL toolbox.[1] If dipoles are modeled to populate the entire brain volume, $d_g\{\cdot,\cdot\}$ can be replaced by the Euclidean distance. The parameter $\sigma$ refers to the spatial width of the Gaussians. Throughout this paper, it is set to $\sigma = 1.5$ cm, which ensures a compromise between sparsity and smoothness.

For the temporal dictionary $\boldsymbol{\Phi}_t$, we use time-frequency basis functions as defined by the Short Time Fourier Transform (STFT) and

[1] http://www.mathworks.com/matlabcentral/fileexchange/10922-matlabbgl.

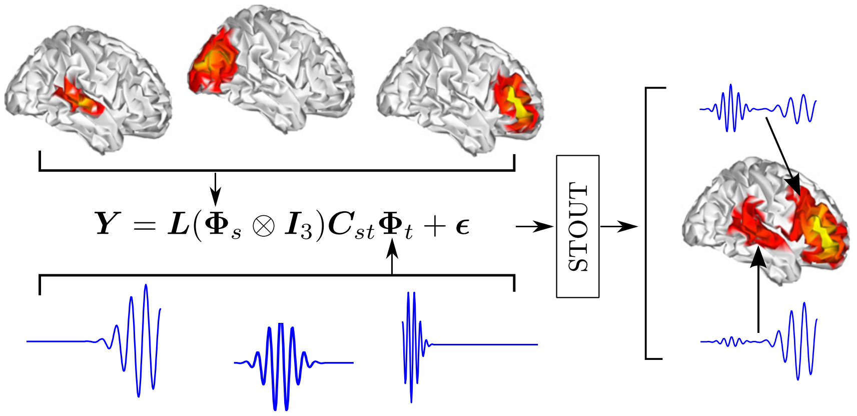


**Fig. 1.** Illustration of STOUT's source model.

implemented in Søndergaard et al. (2012). The coefficients of time-frequency transformation are defined as

$$C_t(m,n) = \sum_l f(l)e^{-2\pi imb(l-an)/L}g(l-an),$$

where $f(l)$ is the value of the input signal at time instant $l$, $g(\cdot)$ is the modulated window, $b$ is the length of the frequency shift, and $a$ is the length of the time shift. Throughout the present work, we use a frequency resolution of 1.2 Hz, a Gaussian window of length 80 ms, and time shift of length 41 ms.

### *Depth compensation*

Distributed inverse solutions based on norm penalties suffer from a depth bias in the sense that deep sources are mislocalized to superficial areas. This is because deeper sources need to be stronger than superficial sources in order to be observed at the sensor level with similar strength, which means that they cause larger penalties as measured by a norm function. To compensate for this localization bias, various depth-weighting schemes have been proposed (Köhler et al., 1996; Lin et al., 2006; Haufe et al., 2008b; Gramfort et al., 2013, see Appendix B). All of them can be implemented by introducing a diagonal matrix $\boldsymbol{W}$, which normalizes the source strength w. r. t. source location before calculating any penalty. Using such a weight matrix, the EEG model Eq. (1) becomes $\boldsymbol{Y} = \boldsymbol{LW}\tilde{\boldsymbol{J}} + \epsilon$, where $\boldsymbol{J} = \boldsymbol{W}\tilde{\boldsymbol{J}}$ is the actual source activity, and where constraints are however only imposed on the weighted activity $\tilde{\boldsymbol{J}}$. Note that, in analogy to the set of basis functions $\boldsymbol{\Phi}_s$, the weight matrix $\boldsymbol{W}$ can always be absorbed into the lead field and thereby does not change the mathematical structure of the cost function of any of the methods discussed. Here, we use a diagonal weight matrix $\boldsymbol{W} \in \mathbb{R}^{3N_d \times 3N_d}$ that is derived from the column $\ell_2$-norms of the lead field (Gramfort et al., 2013), and is defined as $\boldsymbol{W} = \boldsymbol{\Psi} \otimes \boldsymbol{I}_3$. Let $\boldsymbol{L}(\cdot, i)$ denote the $i$-th column of the lead field $\boldsymbol{L}$, then, the $i$-th entry of $\boldsymbol{\Psi}$ is defined as

$$\boldsymbol{\Psi}(i,i) = \sqrt{\left(||\boldsymbol{L}_x(\cdot,i||_2^2 + ||\boldsymbol{L}_y(\cdot,i||_2^2 + ||\boldsymbol{L}_z(\cdot,i||_2^2\right)^{\zeta}}, \quad (7)$$

where matrices $\boldsymbol{L}(\cdot, i)_{x/y/z}$ correspond to the lead field of the $i$-th dipole in the directions $x$, $y$ and $z$, respectively. The parameter $0 < \zeta < 1$ determines how strong the depth compensation is. When $\zeta = 0$, there is no depth bias compensation, while $\zeta = 1$ leads to full compensation. For the present work, we set $\zeta = 0.3$ as in Gramfort et al. (2013). Note that, while depth compensation in general is crucial for reducing location biases of the estimated sources, there are other compensation schemes achieving a very similar effect as the one considered here (see Appendix B).

## Experiments

### *Illustration*

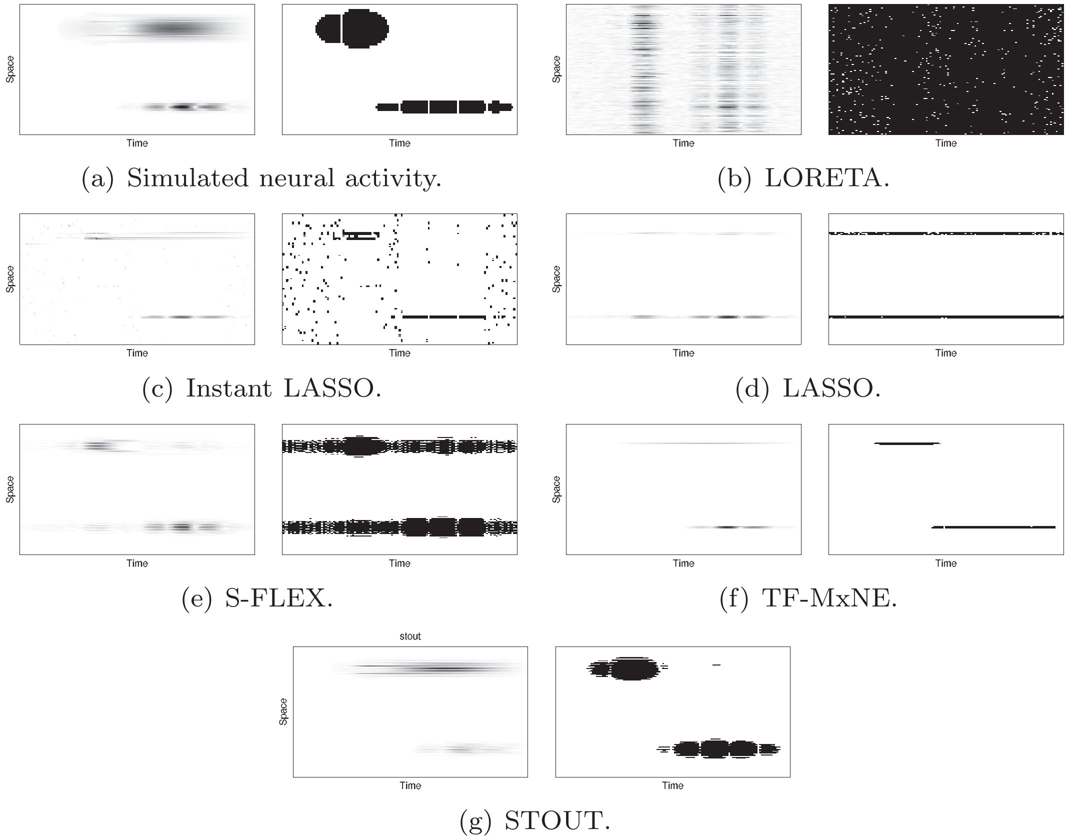
Fig. 2 illustrates typical amplitude and sparsity patterns achieved by different inverse methods in a simulated example comprising randomly generated data at 20 virtual electrodes, 50 virtual dipolar sources, and a randomly generated lead field matrix. Here, we define sparsity patterns as the spatio-temporal representation of all the non-zero coefficients of the achieved reconstruction. As can be seen from the illustration, LORETA does not foster any spatio-temporal sparsity pattern. Two versions of LASSO (with and without grouping of variables along the temporal dimension) encourage spatial sparsity, but they do not correctly reconstruct the simulated spatio-temporal structure (Subfigures 2(a)–2(c)–2(d)). The S-FLEX approach achieves spatial sparsity patterns coherent with the simulation, although it does not adequately recover the corresponding temporal sparsity patterns (Subfigure 2(e)). In contrast, TF-MxNE convincingly reconstructs temporal patterns but lacks accuracy in the spatial domain (Subfigure 2(f)). Lastly, only STOUT accurately reconstructs both the spatial and temporal structure of the simulated activity (Subfigure 2(g)). The results obtained with LORETA are omitted hereinafter for the sake of clarity, given that the accuracy achieved both in the spatial and the temporal domains is significantly lower compared to the other methods.

### *Simulated ERP data*

The common approach to assessing the quality of inverse solutions is to use simulated EEG recordings, for which the underlying brain activity is known. Here, we simulate time-locked brain activity as usually observed in ERP studies involving either one, three, or five active sources.

### *Data generation*

We generated source time series of 1.5 s length sampled at 120 Hz. The time-locked activation was modeled by the real part of Morlet wavelets. The central frequency of each wavelet was randomly sampled from a Gaussian distribution with mean 9 Hz and standard deviation 2 Hz. Each wavelet moreover had a random time shift, which was drawn from a Gaussian distribution with standard deviation of 0.05 s. The mean values were selected depending on the number of active sources: 0.75 s for one active source; 0.375 s, 0.75 s, and 1.25 s for three active sources; 0.25 s, 0.5 s, 0.75 s, 1 s, and 1.25 s for five active sources. A fourth simulation scenario was considered in order to show that the proposed algorithm can cope with more natural situations. This fourth scenario comprised three active sources where the corresponding time series were built using the principal components of a real ERP. The ERPs from which such components were extracted, corresponded to the average responses to target visual stimuli for



(a) Simulated neural activity.

(b) LORETA.

(c) Instant LASSO.

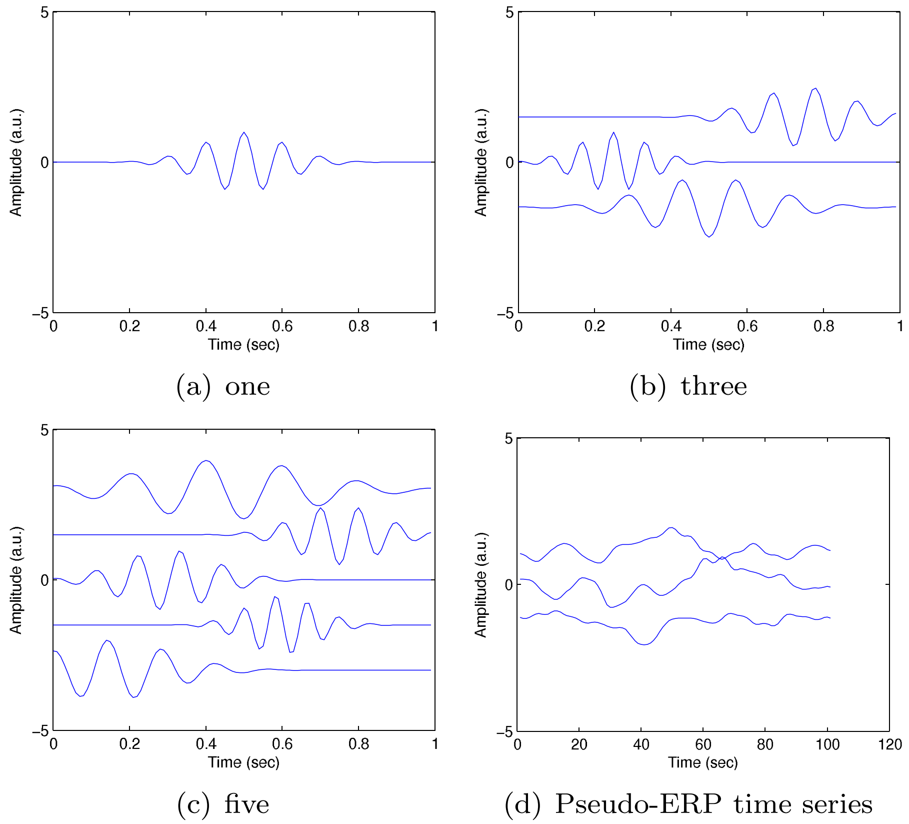
(d) LASSO.

(e) S-FLEX.

(f) TF-MxNE.

(g) STOUT.

**Fig. 2.** Time–space representation of the amplitude (left panel of each subfigure) and the sparsity pattern (right panel of each subfigure) of reconstruction accomplished by various inverse methods compared to the simulated ground truth.



(a) one

(b) three

(c) five

(d) Pseudo-ERP time series

**Fig. 3.** Examples of the generated source activation time series for different numbers of active sources.

three of the subjects of the dataset described in the Localization of auditory and visual evoked potentials in real EEG section. In the spatial domain, this simulation included sources with extended active areas with varying shapes and sizes instead of single dipoles. Specifically, for each reconstruction, three activity clusters of random size (between 1 cm and 5 cm) were generated, then, from each of the clusters, five individual dipoles where randomly selected, and afterwards a spatial low pass filter was applied in the source space; thus, obtaining activity clusters of varying shape and size. The time series within the clusters was the same. Fig. 3 shows examples of the generated source activation time series.

The simulated source activity was mapped to EEG sensor space using a realistic volume conductor model of the human head. The model was based on the nonlinearly averaged anatomy of 152 subjects as acquired by magnetic resonance imaging (Fonov et al., 2009, 2011). Each active source was modeled as a dipolar current source with random spatial orientation and location. Only locations on the cortical surface were allowed. The mapping from sources to sensors was computed using quasi-analytic expansions of the electric lead fields (Nolte and Dassios, 2005). We modeled 64 channels, which simulated the pseudo-EEG at positions defined by the extended international 10–20 system (Oostenveld and Praamstra, 2001).

To obtain realistic pseudo-EEG measurements, the sensor-space ERP signal was superimposed by simulated biological and measurement noise. The former represents non-timelocked (background) brain activity, and was modeled by 500 independent random time series with $1/f$-shaped (pink noise) spectrum originating from dipoles with randomly drawn location and orientation parameters. The activity at these 500 dipoles was mapped to EEG sensor space and superimposed by temporally and spatially independent Gaussian sensor noise. The biological-to-sensor noise ratio was $-5$ dB. Signal-to-noise ratio (SNR) is defined as $SNR_{[dB]} = 10log(P(\mathbf{X})/P(\Upsilon))$ where $P(\boldsymbol{X})$ and $P(\ \Upsilon)$ denote the power of the signal and the noise, respectively, and are defined as the mean variance across channels. The overall noise was then added to the simulated ERP signal to obtained pseudo-EEG measurements with a SNR of 0 dB at the single-trial level.

For each experiment, we generated multiple trials, for each of which we sampled new noise realizations. We generated five different datasets with the number of trials set to 5, 20, 50, 100 and 250. By averaging the pseudo-EEG data across trials, we obtained SNR values of $-5$, 0, 7, 12 and 14 dB, respectively, where the final SNR in a single trial after adding biological and measurement noise, was $-10$ dB. For each condition (one, three and five time-locked sources), we performed 30 experiments. The noise as well as the signal parameters (i.e. location and orientation of the source dipoles, frequency and time shift of the Morlet wavelets) were drawn randomly for each experiment.

Note that our simulation scenario mimics an event-related experimental design, in which all source activity of interest is time-locked to a stimulus (and thus non-stationary), while all noise sources are unrelated to the stimulus and thereby stationary. This implies that for increasing signal-to-noise ratio (as measured by the number of trials in the average as well as the number of ERP sources), the overall level of non-stationarity in the data rises (see Appendix C).

### *Application of source reconstruction algorithms*

We benchmarked several inverse methods regarding their ability to reconstruct the locations and time courses of the simulated time-locked activity. The methods considered were the minimum current estimate (here referred to as LASSO, Matsuura and Okabe, 1995), S-FLEX (Haufe et al., 2011), TF-MxNE (Gramfort et al., 2013) and STOUT. The tolerance used as a stopping criteria in the FISTA algorithm was set to $1 \times 10^{-3}$.

All methods were implemented within our own Matlab toolbox. The localization was carried out using the same head model in which the data were generated. The source space was defined to be the tessellated cortical surface, on which 4000 dipoles were placed. We here assumed that the orientations of the current sources are unknown, which is the case when no individual anatomy is available and generic head models (atlases) are used — see the related discussion in the Dipole orientation modeling section. Therefore, STOUT has to infer $3 \times 4000 \times 528 = 6{,}336{,}000$ source variables (spatial directions × elements of the spatial dictionary × elements of the time–frequency dictionary) from $118 \times 180 = 21{,}240$ (sensors × samples) noisy measurements.

For all methods considered, nonlinear optimization has to be carried out either separately for each sample or jointly for all samples. The LASSO has classically not been formulated as a spatio-temporal solver. However, to be able to reconstruct meaningful time series with this method, we here run it with $\ell_2$-norm constraints along the temporal dimension as suggested in Ou et al. (2008), Haufe et al. (2011). For S-FLEX, TF-MxNE and STOUT, we employ the spatial and temporal basis function dictionaries and the depth weighting matrix described in the Spatio-temporal unifying tomography (STOUT) section. All methods were run with $\boldsymbol{Q}_\epsilon = \boldsymbol{I}_{N_c}$, where $\boldsymbol{I}_{N_c}$ denotes the $N_c \times N_c$ identity matrix.

### *Tuning of the regularization parameters*

A challenge in practice is the selection of the regularization parameters of all methods governing the tradeoff between the likelihood term $||\boldsymbol{Y}-\boldsymbol{LJ}||^2_{\boldsymbol{Q}_\epsilon}$ and the penalty terms in $\Theta(\boldsymbol{J})$, as well the relative importance of the individual penalties.

In order to investigate to which extent STOUT is capable to interpolate between the purely spatial regularization of S-FLEX and the purely temporal structured sparsity of TF-MxNE, we applied STOUT with several fixed ratios $\lambda_s : \lambda_t = 90 : 0.5$, 90:10, 90:30, 90:90. TF-MxNE was tested with the same set of ratios. However, we here present only the results obtained for $\lambda_s : \lambda_t = 90 : 30$, which is also the ratio used in Gramfort et al. (2013). Notably, this ratio yielded the best accuracy in the temporal domain, while no significant difference in spatial accuracy was observed across the tested ratios.

For adjusting the tradeoff between data fidelity and complexity terms in the cost functions, we make use of the fact that the ground truth about the underlying source activations is known in our simulation setting. We therefore tune the regularization parameters of all methods such that the residual (noise) variance $||\boldsymbol{Y}-\boldsymbol{L}\hat{\boldsymbol{J}}||^2_{\boldsymbol{Q}_\epsilon}$ due to the estimated current density $\hat{\boldsymbol{J}}$ matches the residual variance $||\boldsymbol{Y}-\boldsymbol{LJ}||^2_{\boldsymbol{Q}_\epsilon}$ of the actual simulated sources $\boldsymbol{J}$. This is achieved by starting with very strong regularization, and iteratively adjusting the regularization parameter of each method by factors of 0.7 until the achieved residual norm falls below the ideal value. Note that, while this procedure ensures a fair comparison of all methods regardless of their parametrization, it cannot be used for real data, for which the true residual variance is unknown. Here, the regularization strength has to be chosen based on criteria derived from the data (see the Localization of auditory and visual evoked potentials in real EEG section).

### *Evaluation of source reconstruction accuracy*

The source reconstruction accuracy needs to be determined with respect to both the spatial and temporal domain. To assess the spatial accuracy of the estimated neural activity, we compare the spatial distribution of the dipole-wise power of the estimated source activity with the true power of the simulated sources using the Earth Mover's Distance (EMD, Haufe et al., 2008b). The dipole-wise power is computed as the sum of the squared currents across the three spatial directions and across time. The EMD, denoted by $m_e \in \mathbb{R}^+$, measures the effort it takes to transform the estimated power distribution into the true distribution by “transporting” probability mass. Here, lower value of $m_e$ indicate better reconstruction quality in the spatial domain.

To assess the reconstruction quality in the temporal domain, we measure the correlation between each of the simulated ERP time series and the reconstructed signals at all dipoles. For each simulated source, the maximum correlation across all dipoles and spatial orientations is computed. The maximum values are then averaged across the simulated sources to give an average maximum correlation $m_c \in \mathbb{R}^+$, which we report. Here, higher values of $m_c$ generally correspond to better reconstruction in the temporal domain.

### *Localization of auditory and visual evoked potentials in real EEG*

We analyzed EEG recordings from fifteen healthy subjects that participated in an ERP study involving auditory and visual stimulation. The study was carried out in a brain-computer interface (BCI) context, and is described in detail in An et al. (2014). We here iterate the aspects of the study relevant for the present analysis.

#### *Experimental paradigm*

All subjects perceived sequences of either auditory (condition A) or visual (condition V) stimuli. In both conditions, each stimulus lasted 130 ms, while the time delay between the onsets of two consecutive stimuli was 200 ms. For each trial (stimulation sequence), six physically different stimuli were repetitively presented in a pseudo-random sequence. This design resulted in 36 stimuli per trial. During each trial, subjects had to focus their attention to one predefined stimulus (target), and to count its occurrences while ignoring the presentation of the remaining five stimuli (non-targets). In the study of An et al. (2014), the conditions A and V only differed with respect to the type of stimulus presentation, so that the ERP responses of each type of sensory stimulus could be compared for each subject. For condition V, six visual stimuli differing in color and shape were presented in the center of a 19" screen, while for condition A, six naturally recorded auditory stimuli (spoken syllables recorded from several voices) differing in pitch and spatial direction were delivered via headphones. Trials of condition A and V were alternated. For both conditions, there were 24 trials resulting in a total number of 144 target stimuli and 720 non-target stimuli acquired per subject.

#### *Data acquisition and preprocessing*

EEG data were recorded by using 63 Ag/AgCl electrodes symmetrically placed at the standard positions of the international 10–20 system (Oostenveld and Praamstra, 2001). The EEG data were down-sampled to 100 Hz, band-pass filtered between 0.4 and 23 Hz, and epoched around each stimulus presentation (from 150 ms before stimulus onset to 800 ms after stimulus onset). Baseline subtraction was performed for each trial by using the interval −150 ms–0 ms. EEG epochs containing eye and muscle artifacts were excluded from further analysis. The remaining epochs were averaged separately for each subject, stimulation condition, and for targets and non-targets.

#### *Source localization*

We applied STOUT to the averaged ERP time series of each subject for each of the four stimulus conditions visual target, visual non-target, auditory target and auditory non-target. The spatial-to-temporal regularization ratio $\lambda_s : \lambda_t$ was set to 90:30. Since no ground truth about the noise level was available, the overall regularization strength was set according to a data-driven criterion. Here we used cross-validation as in Haufe et al. (2008a, 2011), Habermehl et al. (2014). To this end, the set of electrodes was split into five sets, where each of the sets served as the holdout set in one fold of the cross-validation. The source reconstruction was carried out on the remaining four electrode sets, while the estimated activity was projected to the held-out electrodes using the lead field, and the difference between the measured and estimated activity at the held-out electrodes $\boldsymbol{Y}_{\text{test}}$ and $\tilde{\boldsymbol{Y}}_{\text{test}}$ was measured using the criterion $||\boldsymbol{Y}_{\text{test}} - \tilde{\boldsymbol{Y}}_{\text{test}}||^2_{\boldsymbol{Q}_\epsilon}$. To save computation time, the cross-validation was carried out using LORETA, where the regularization parameter was calculated using the generalized crossvalidation criterion as described in Grech et al. (2008). The regularization parameter of STOUT was then adjusted to achieve the same average residual variance of the cross-validated LORETA estimate. This is done to save computation time, given that the time needed for optimization of the hyperparemeters using STOUT itself is prohibitively long. In order to enable the reader to replicate our analysis, we provide the ERP data, the head model and the Matlab source code for STOUT.[2]

#### *Statistical analysis*

For each subject and experimental condition, we computed the dipole-wise source power as in the simulation setting. Grand-average source power was then calculated by taking the mean over the subject-wise values. We moreover compared two conditions across subjects, using a two-sided pairwise Student-t test. Therefore, the difference of the log-power was evaluated between several conditions: visual target vs. visual non-target, auditory target vs. auditory non-target, as well as visual target vs. auditory target. Significant differences in mean power were assumed for brain areas featuring t-scores with absolute values greater than 2.1 (alpha level $p < 0.05$, uncorrected).

## Results

### *Simulated ERP data*

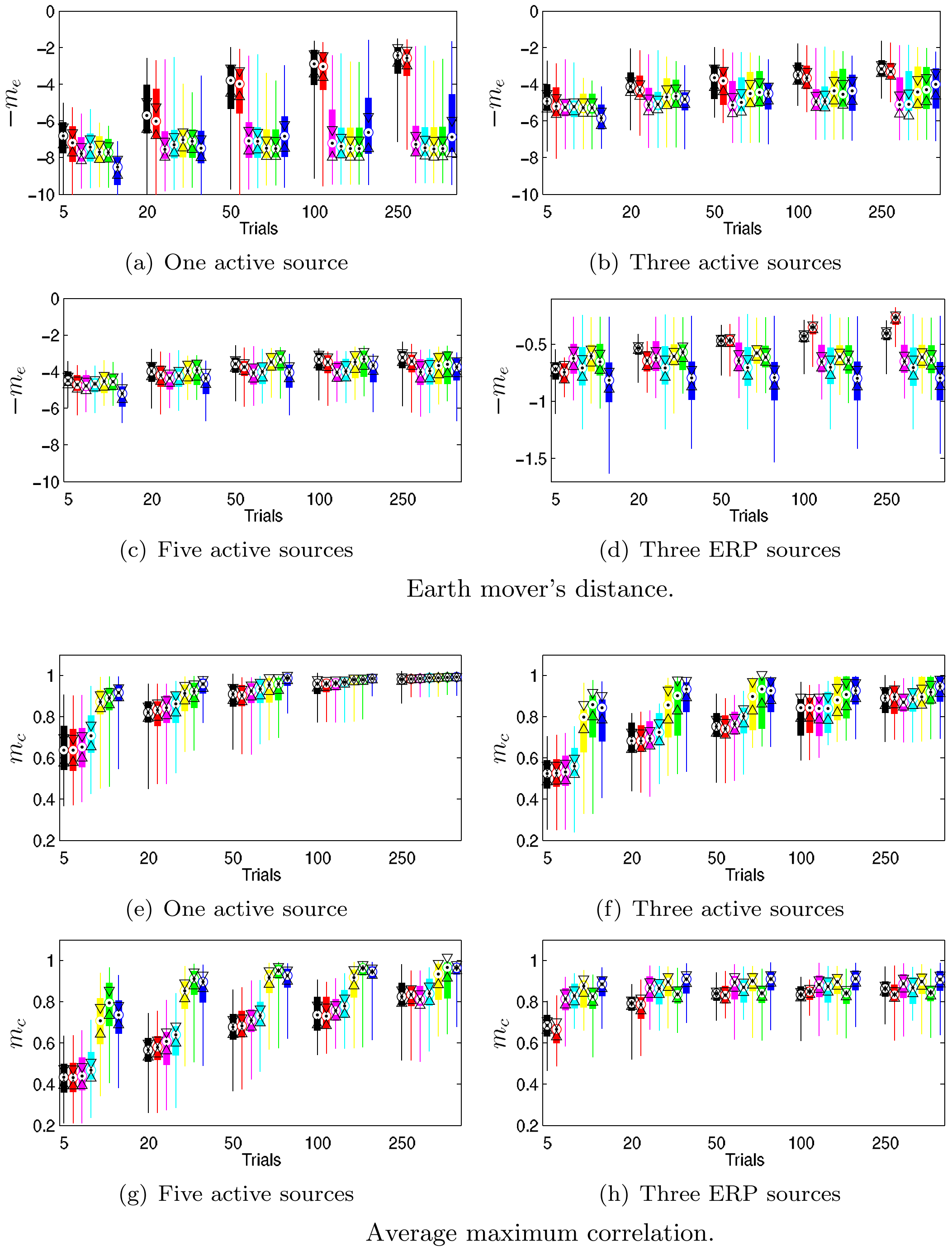
The spatial and temporal reconstruction performance of LASSO, S-FLEX, TF-MxNE, and STOUT on simulated ERP data is depicted in Fig. 4. We observe that S-FLEX yields higher spatial accuracy than TF-MxNE, while TF-MxNE yields higher temporal accuracy (higher average maximal correlation) than S-FLEX across all SNR values as well as for one, three and five non-stationary sources. Such trend is also observable in the simulation containing realistic sources. LASSO performs worse than the other methods, however, providing a spatial accuracy close to TF-MxNE and a temporal accuracy close to S-FLEX.

STOUT's performance lies strictly in between those of S-FLEX and TF-MxNE for both the spatial and the temporal domain. Thus, in terms of temporal reconstruction STOUT outperforms methods which neglect the possibility of non-stationary activations (such as S-FLEX). In terms of spatial accuracy, STOUT outperforms methods which fail to impose a realistic spatial structure on the sources (such as TF-MxNE). Here, the chosen ratio of the parameters weighting the spatial and temporal regularization terms defines whether STOUT behavior is closer to that of S-FLEX or to that of TF-MxNE. That is, by adjusting the spatial-to-temporal regularization ratio $\lambda_s : \lambda_t$, STOUT can be tuned to put more emphasis on spatial or temporal reconstruction depending on the requirements for the specific data at hand.

### *Auditory and visual evoked potentials*

Fig. 5 shows the sensor-space data, as well as the results of the source reconstruction using STOUT for the non-target ERPs in the visual (see Subfigures 5(a), 5(b), 5(c), 5(d)) and auditory (Subfigures 5(a), 5(b), 5(c), 5(d)) conditions for representative subjects, respectively. Likewise, Fig. 6 shows the corresponding results for the target stimuli. In the first subfigures, the trial-wise stimulus-locked EEG time series are shown, while in the second, time series of the reconstructed activity. Likewise, in the third subfigures the average scalp topography at $t = 200$ ms (marked with a red line in the trace in the first panel) is depicted. The last subfigure shows the corresponding source reconstruction (dipole-wise power) according to STOUT at the same time point. The time point of $t = 200$ ms was selected for presentation, since it is known that the negative component of the ERPs found between $t = 100$ ms and $t = 200$ ms — typically termed N200 component — contain specific brain responses related to the processing of visual and auditory stimuli (Folstein and Van Petten, 2008).

[2] https://github.com/jscastanoc/stout.



(a) One active source

(b) Three active sources

(c) Five active sources

(d) Three ERP sources

Earth mover's distance.

(e) One active source

(f) Three active sources

(g) Five active sources

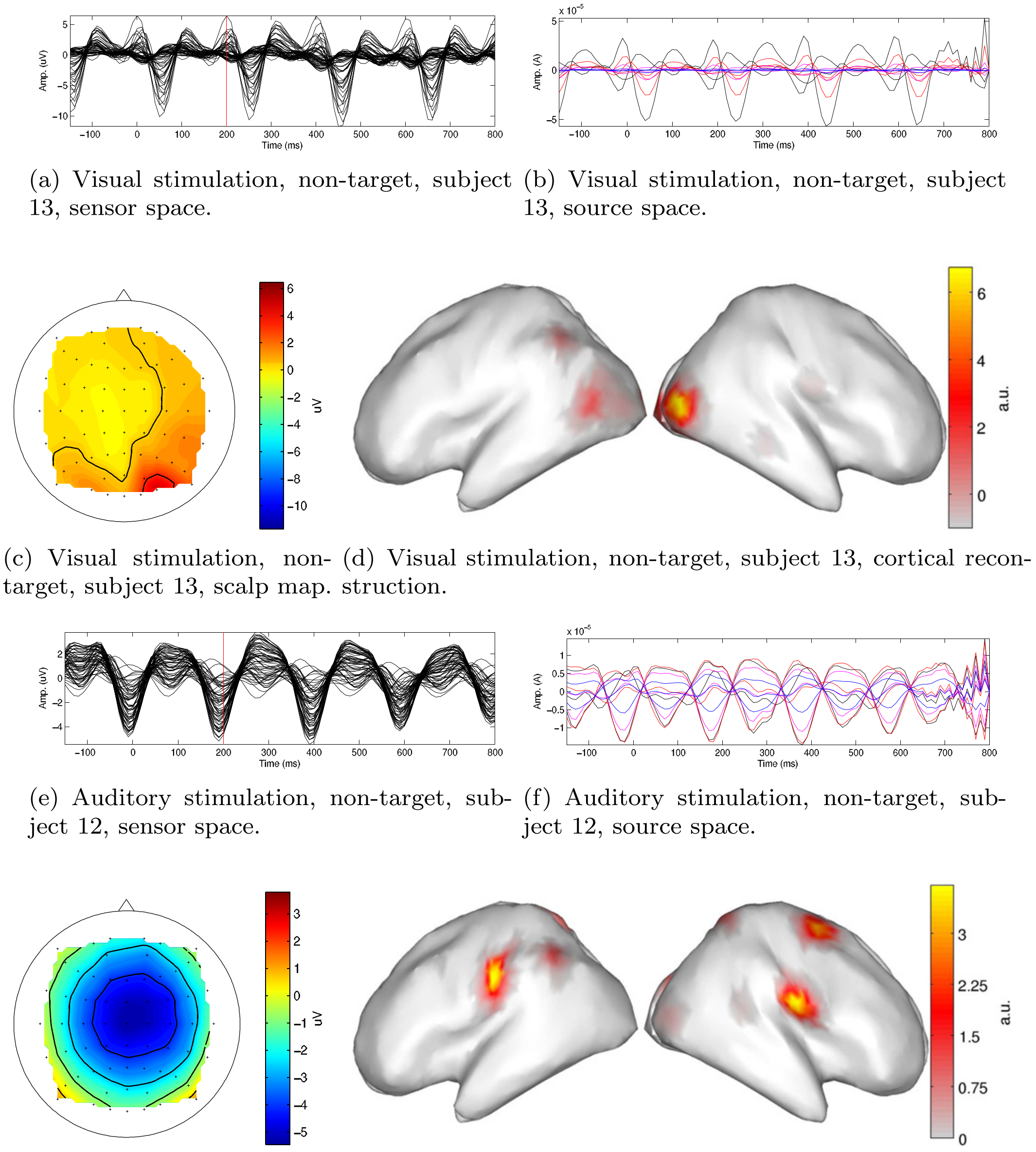
(h) Three ERP sources

Average maximum correlation.

**Fig. 4.** Spatial (a–d) and temporal (e–h) reconstruction accuracy according to the earth mover's distance, $m_e$, and average maximum correlation, $m_c$, of ⬛LASSO 🟥S-FLEX, 🟪STOUT 90:0.5, 🟦STOUT 90:10, 🟨STOUT 90:30, 🟩STOUT 90:90, 🟦TF-MxNE.

The ERP waveforms show a prominent fronto-central peak at around 200 ms in the auditory modality and a prominent temporo-occipital negative peak at around 180 ms in the visual modality.

For both visual and auditory stimulation, STOUT localizes the N200 component in the vicinity of brain areas associated with the corresponding sensory stimulus processing. In the case of auditory stimulation, these are areas slightly above the left and right auditory cortices in the temporal lobes, while in the case of visual stimulation, these are predominantly areas in the occipital lobe covering the visual cortices. In both cases, the reconstructed time courses show a clean representation of the dynamics found in the original ERP responses. The spatial error in the localization of the auditory and visual cortices can be attributed to the use of an average head model that does not reflect the exact anatomy of each individual subject. Generally, we observe that non-target stimuli lead to more confined regions of estimated brain activity, while for target stimuli additional activity appears in the central and frontal areas. The activities in additional brain areas are even more pronounced if we look at grand-average source reconstructions instead of the sources estimates of individual representative subjects (see Fig. 7). We hypothesize that these additional activations may be related to conscious stimulus processing (numeration of the stimuli) carried out by the subjects. However, such activation patterns may also be associated with noise caused by the lower number of trials (five times less) available for the target condition compared to the non-



(a) Visual stimulation, non-target, subject 13, sensor space.

(b) Visual stimulation, non-target, subject 13, source space.

(c) Visual stimulation, non-target, subject 13, scalp map.

(d) Visual stimulation, non-target, subject 13, cortical reconstruction.

(e) Auditory stimulation, non-target, subject 12, sensor space.

(f) Auditory stimulation, non-target, subject 12, source space.

(g) Auditory stimulation, non-target, subject 12, scalp map.

(h) Auditory stimulation, non-target, subject 12, cortical reconstruction.

**Fig. 5.** Sensor-space EEG data and STOUT source reconstruction of visual and auditory evoked potentials elicited by *non-target* stimuli for representative subjects. Subfigures 4(a) and 4(e): averaged stimulus-locked EEG signal, where $t = 0$ ms refers to the stimulus onset. The red vertical line at $t = 200$ ms post-stimulus marks the N200 component analyzed below. Subfigures 4(c) and 4(g): scalp topography at $t = 200$ ms. Subfigures 4(d) and 4(h): source power at $t = 200$ ms as estimated by STOUT. Subfigures 4(b) and 4(f): Time series of the reconstructed sources.

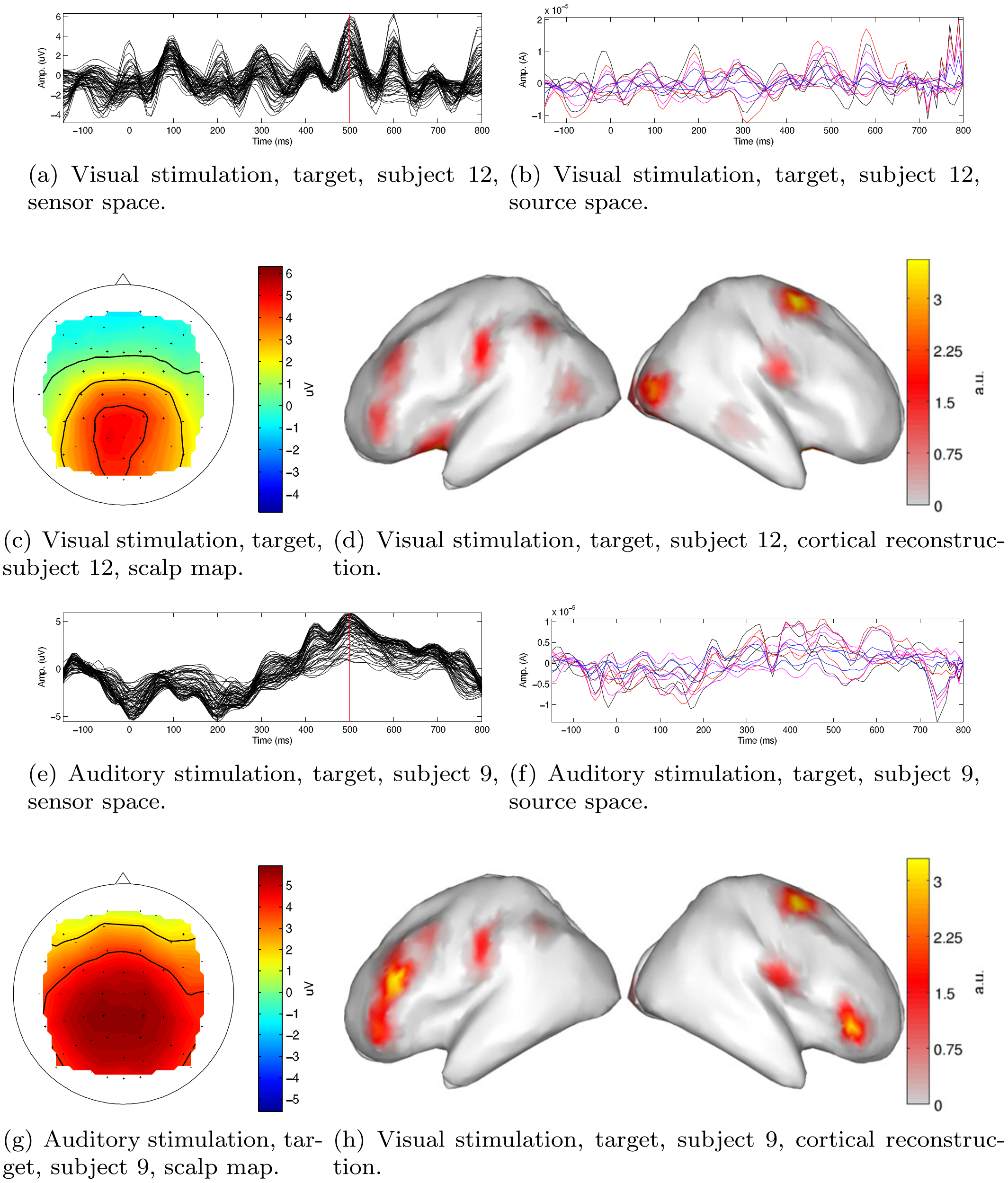
target condition. In the grand-average, we also notice broader activations for auditory stimulation in general as compared to visual stimulation.

Fig. 8 shows the results of dipole-wise Student-t tests for differences in log-power between conditions. Here, red color denotes higher activity in the condition mentioned first, while blue color denotes higher activity in the condition mentioned second. When comparing non-target stimuli across sensory modalities (Subfigure 8(a)), we observe significantly higher activity in the occipital lobes for visual stimulation, and significantly higher activity in temporal lobes for auditory stimulation, which is in line with the functional relevance of these brain areas for processing visual and auditory stimuli Linden et al. (1999). We also observe that activations in central and frontal areas (see also Fig. 7) are stronger for auditory than for visual stimulation.

Comparing responses to non-targets and targets (Subfigures 8(b) and 8(c)), we note that for both stimulus modalities, non-targets lead to focal activation in sensory related brain areas. In the case of visual stimulation (Subfigure 8(b)), those are the occipital lobes, while in case of auditory stimulation (Subfigure 8(c)), those are the temporal regions. In both cases, we observe higher activity in additional frontal and central brain areas for target stimuli. We hypothesize that these results are related to a larger degree of high-level conscious processing in target stimuli as compared to non-target stimuli.

## Discussion

Our results demonstrate that STOUT is able of reconstructing the spatio-temporal profile of brain activity underlying EEG measurements



(a) Visual stimulation, target, subject 12, sensor space.

(b) Visual stimulation, target, subject 12, source space.

(c) Visual stimulation, target, subject 12, scalp map.

(d) Visual stimulation, target, subject 12, cortical reconstruction.

(e) Auditory stimulation, target, subject 9, sensor space.

(f) Auditory stimulation, target, subject 9, source space.

(g) Auditory stimulation, target, subject 9, scalp map.

(h) Visual stimulation, target, subject 9, cortical reconstruction.

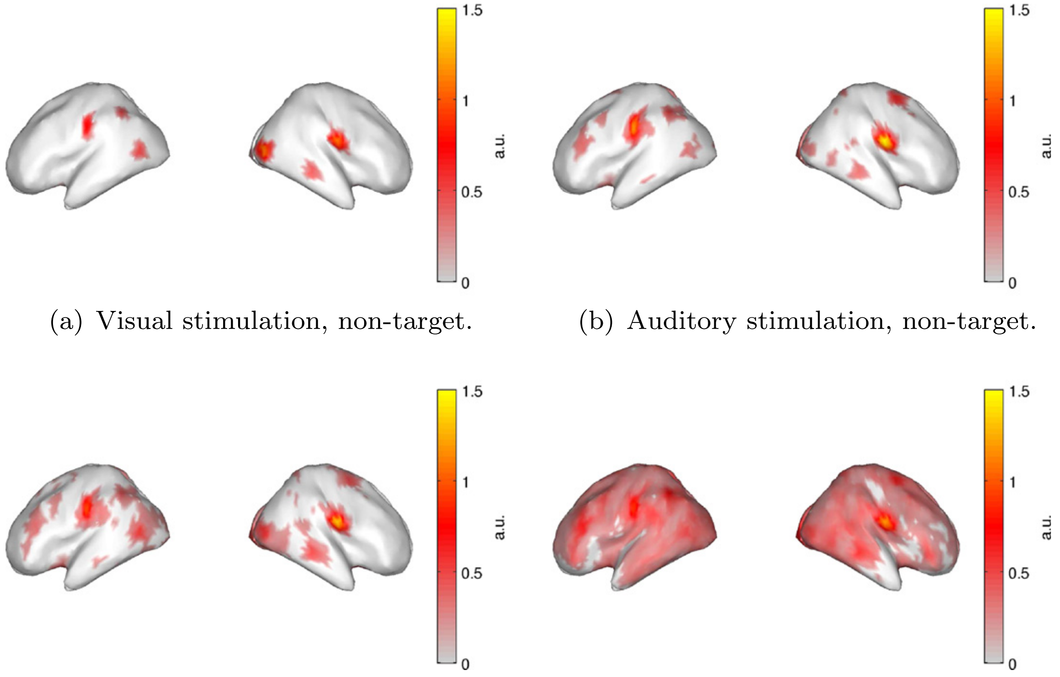
**Fig. 6.** Sensor-space EEG data and STOUT source reconstruction of visual and auditory evoked potentials elicited by *target* stimuli for representative subjects. Subfigures 5 (a) and 5(e): averaged stimulus-locked EEG signal, where $t = 0$ ms refers to the stimulus onset. The red vertical line at $t = 500$ ms post-stimulus marks the P3 component analyzed below. Subfigures 5(c) and 5(g): scalp topography at $t = 500$ ms. Subfigures 5(d) and 5(h): source power at $t = 500$ ms as estimated by STOUT. Subfigures 5(b) and 5(f): Time series of the reconstructed sources.

in realistic simulations as well as from real data. It does so through imposing physiologically-motivated constraints based on the assumption that the EEG measurements are generated by a small number of brain sources each having a focal (smooth but localized) structure in space, time and frequency. In simulations, we have shown that STOUT is able to reconstruct both the time courses and the spatial profiles of the sources, where it typically outperforms methods neglecting the possibility of non-stationary activations (such as S-FLEX) in terms of temporal reconstruction and methods failing to impose a realistic spatial structure on the sources (such as methods assuming sparsity on the voxel level) in terms of spatial accuracy. Our method can be regarded as a fusion of S-FLEX and TF-MxNE. Alternatively, STOUT can be regarded as an extension of S-FLEX enabling the reconstruction of non-stationarity source activations, or as an extension of TF-MxNE enabling the recovery of sources with focal but non-sparse spatial structure. By adjusting the ratio between two regularization terms, STOUT is also capable of trading temporal for spatial reconstruction accuracy and vice versa, and thereby in a sense to interpolate between the S-FLEX and TF-MxNE solutions.

### *Modeling source non-stationarity and realistic spatial extent*

To the best of our knowledge, STOUT is the first method to incorporate prior knowledge both on spatial and temporal properties of the sources in a single model using basis function expansions. Like other approaches employing spatio-temporal regularization (Dannhauer et al., 2013), our model is motivated by the assumption that spatial and temporal properties are equally important to provide a holistic characterization of physiologically plausible source activity.

Spatial focality of the sources is not only essential for the neurophysiological interpretation of source reconstruction results. It is also required for group-level statistical analyses as described in the



(a) Visual stimulation, non-target.

(b) Auditory stimulation, non-target.

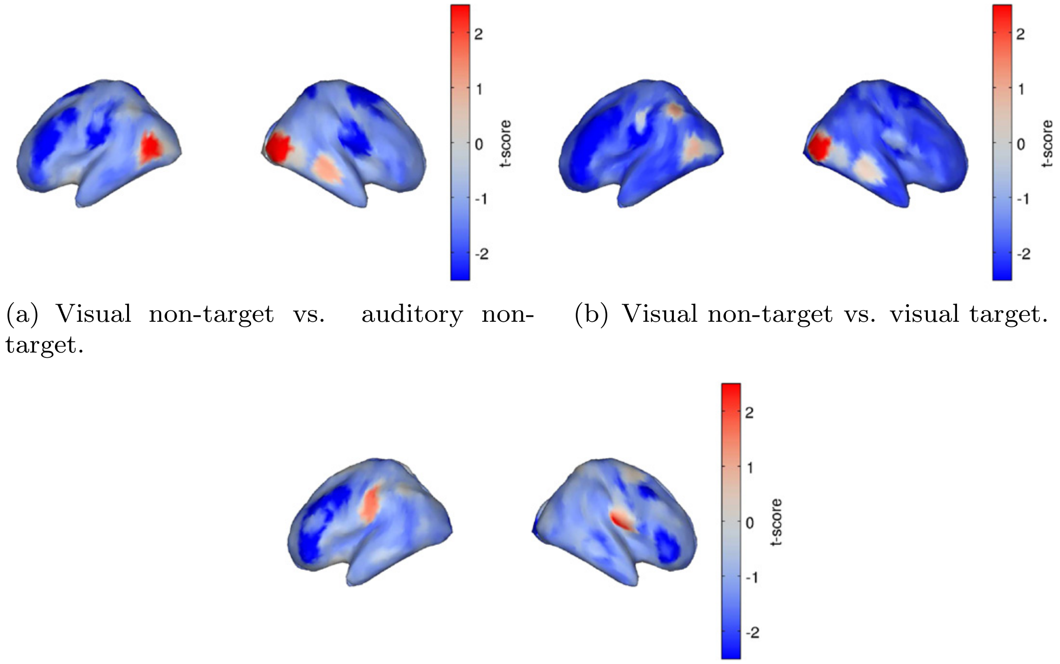
(c) Visual stimulation, target.

(d) Auditory stimulation, target.

**Fig. 7.** Grand-average source power at $t = 200$ms after visual/auditory stimulation as estimated by STOUT for target and non-target trials.

Localization of auditory and visual evoked potentials in real EEG section. Performing the same analyses on source estimates that are sparse on the level of individual dipoles would not be meaningful due to the differing sparsity patterns per subject.

On the other hand, non-stationarity is an intrinsic property of many neurophysiological recordings (Woolrich et al., 2013). Non-stationary source activations arise from many factors including event-related experimental designs. Therefore, the source model of STOUT allows for changes in neural activation patterns over time. This makes our method particularly suited for the analysis of ERPs, as well as of event-related synchronization and desynchronization (ERS/D) phenomena (Neuper and Klimesch, 2006). Note, however, that the fact that STOUT models non-stationary source activations does not necessarily imply that it is necessarily robust to non-stationary noise contributions originating from outside the brain. This issue remains a subject of further study. Measures based on the Kullback–Leibler divergence as discussed in



(a) Visual non-target vs. auditory non-target.

(b) Visual non-target vs. visual target.

(c) Auditory non-target vs. auditory target.

**Fig. 8.** Differences in log source power at $t = 200$ms as estimated by STOUT for different experimental conditions. (a) visual non-target vs. auditory non-target. (b) visual non-target vs. visual target. (c) auditory non-target vs. auditory target. Red color means higher activity in the condition mentioned first, while blue color means higher activity in the condition mentioned second. Differences were quantified using a two-sided pairwise Student-t test. T-scores with absolute values greater than 2.1 are significant at alpha level $p < 0.05$ (uncorrected).

Appendix C are capable of quantifying the overall degree of stationarity in the data. However, they cannot distinguish whether non-stationary behavior is caused by brain or noise components.

*Amplitude bias*

Belonging to the broad class of sparse estimators relying on $\ell_1$-norm penalties, STOUT is susceptible to a certain bias in the sense that the amplitude of the estimated currents is consistently underestimated. If the global scale of the estimated coefficients is of interest, this bias can be removed by using STOUT only for selecting the subset of nonzero coefficients, while the actual value of these nonzero coefficients is determined in a subsequent ordinary-least-squares (OLS) regression. This concept behind least-angle-regression-OLS (LARS-OLS) as well as the 'relaxed LASSO' Efron et al., 2004, Meinshausen 2007. However, since for our purposes of source localization and time series reconstruction we are only interested in relative amplitude differences between coefficients/current sources, which should be preserved by using the same amount of regularization for all coefficients, we do not use STOUT in combination with a technique like the relaxed LASSO.

*Tuning of the hyperparameters*

Besides the actual source coefficients $\boldsymbol{C}_{st}$ estimated from data through convex optimization, STOUT is implicitly characterized by a number of *hyperparameters*. Most importantly, the shape of the spatial and temporal basis functions (atoms) crucially affects the space–time–frequency structure and thereby the neurophysiological plausibility of the obtained inverse solutions. Here, we use Gaussian curves defined on the cortical manifold to encode the assumption that brain sources have sparse but locally smooth (that is, focal) spatial activation profiles. For the temporal domain, we use windowed Fourier basis functions to encode that the time courses of the source activations are focal in both time and frequency. Given these building blocks, we construct dictionaries consisting of scaled and translated versions of the individual spatial and temporal atoms. By restricting these dictionaries to specific ranges of scales and translations, it is generally possible to introduce further prior knowledge about the location of the activities of interest in space, time and frequency, as well as their spatial extent, temporal activation window, and bandwidth. An alternative strategy would consist in using massively *overcomplete* dictionaries covering broad ranges of scales and locations, and using the sparsity mechanism of STOUT to select all source parameters including the appropriate spatial and time-frequency scales. Here, it should be kept in mind, however, that any inflation of the basis function dictionaries comes at the expense of higher computational complexity of the source reconstruction.

Another set of hyperparameters are the regularization constants $\lambda_s$ and $\lambda_t$ governing the influence of the terms measuring spatial and temporal sparsity in the cost function (see Eq. (6)) relative to the data fidelity term and to each other. Here, we fix the ratio $\lambda_s : \lambda_t$ to a value allowing for a balanced reconstruction of the temporal and spatial properties of the source activations. The tradeoff between data fidelity and the spatio-temporal constrains of the model is specified by the common scale of $\lambda_s$ and $\lambda_t$, which is adjusted using a data-driven criterion based on cross-validation. The use of other heuristics as well Bayesian model selection criteria is also conceivable (e. g., Hansen, 1992; Nummenmaa et al., 2007).

Finally, we are required to decide for a depth compensation scheme to prevent location biases arising from the penalization of the source activity using norm constraints. Based on empirical results indicating that weightings calculated using the column-norms of the lead field and the estimates of dipole-wise variance achieve similar performances (see Appendix B), we here employ column-norm weights as in Gramfort et al. (2013).

*Dipole orientation modeling*

Another aspect in the design of an inverse method is the spatial orientation of the dipolar current sources. It has been widely acknowledged that cortical gray matter is the main (although not single) contributor to the EEG signal. This is due to the fact that the dendritic trunks of the pyramidal neurons predominantly found in cortical structures are locally co-oriented (perpendicular to the cortical surface), allowing postsynaptic potentials to add up to a degree measurable from outside the head (e. g., Baillet et al., 2001; Nunez and Srinivasan, 2005). For a known cortical anatomy, it is therefore possible to constrain each current source to be oriented along the normal vector of the local cortical patch, and to estimate current amplitudes only. When using spatial basis functions, it is moreover possible to extend this idea in a way such that each basis function takes into account the local orientation at each of cortical locations it covers. This gives rise to complex but physiologically plausible scalp potentials as building blocks for the approximation of the measured EEG signal.

When precise information about the local cortical folding is unavailable (such as when using generic head models based on average anatomies), approaches relying on predefined orientations, however, become error-prone, and free-orientation models are more appropriate. Therefore, we here formulate STOUT in a way that assigns a single but flexible orientation to each spatial basis function. Note that, by using the $\ell_{1,2}$-norm penalty as in Haufe et al. (2008b), Ding and He (2008), Haufe et al. (2011), we avoid biases in the estimation of these orientations, which are otherwise common for methods based on spatial sparsity.

## Conclusion

We have introduced STOUT, a novel method for the reconstruction and localization of brain activity from measured EEG time series. Our method is particularly suited for the localization of non-stationary data such as ERPs, as was shown in simulations. Our STOUT analysis of the generators of EEG visual and auditory evoked N200 potentials reveals that most active sources originate in the temporal and occipital lobes, which stands in line with the literature on sensory processing. Summarizing, STOUT improves upon previous characterizations of EEG source activity by introducing a physiologically-motivated space-time-frequency representation, which should be the stepping stone towards obtaining more accurate source reconstructions in practice.

## Acknowledgments

This research was supported by BrainLinks-BrainTools Excellence cluster initiative at the University of Freiburg, funded by the German Research Foundation (DFG), grant number EXC 1086. Universidad Nacional de Colombia with the scholarship program "Beca de estudiantes sobresalientes de posgrado". Colciencias with the *Programa Nacional de Formacion de Investigadores GENERACION DEL BICENTENARIO — 2011*. COLCIENCIAS, project *Evaluación asistida de potenciales evocados cognitivos como marcador del transtorno por déficit de atensión e hiperactividad* (*TDAH*). It was also funded by a Marie Curie International Outgoing Fellowship (grant No. PIOF-GA-2013-625991) within the 7th European Community Framework Programme.

## Appendix A. Optimizing STOUT using the FISTA algorithm

Note that the spatial basis functions $\boldsymbol{\Phi}_s$ can be absorbed into the lead field by means of the transformation $\boldsymbol{L}' = \boldsymbol{L}(\Phi_s \otimes \boldsymbol{I}_3)$. With this modified lead field $\boldsymbol{L}'$, the objective function Eq. (6) exactly resembles Eq. (4). Therefore, the STOUT estimate inherits all properties of the TF-MxNE solution including most importantly convexity, and can be obtained in the same way using the Fast Shrinkage Thresholding Algorithm (FISTA, Beck and Teboulle, 2009; Gramfort et al., 2012) employed in Gramfort et al.

(2013). For clarity, we summarize the optimization procedure in Algorithm 1. The proximity operator of the $(\ell_{1,2} + \ell_1)$-norm penalty used by FISTA is defined as

$$\mathrm{prox}_{\ell_1+\ell_{1,2}}(\boldsymbol{X}, \lambda_s, \lambda_t) = \frac{\boldsymbol{X}_{(i,j)}}{\left|\boldsymbol{X}_{(i,j)}\right|}\{|\boldsymbol{X}_{(i,j)}| - \lambda_t\}_+ \left(1 - \frac{\lambda_s}{\sqrt{\sum_{\forall i \in N_f}\{|\boldsymbol{X}_{(i,j)}| - \lambda_t\}_+^2}}\right),$$

where $\{\cdot\}_+$ yields the maximum value between the argument and 0, and $|\cdot|$ is the absolute value of the argument. The value $\boldsymbol{X}_{(i,j)}$ corresponds to the $i$-th row and $j$-th column of the matrix $\boldsymbol{X}$ which is the argument of the proximity operator,

$$\boldsymbol{X} := \lambda_s(\Phi_s \otimes \boldsymbol{I}_3)^T \boldsymbol{L}^T(\boldsymbol{Y} - \boldsymbol{L}(\Phi_s \otimes \boldsymbol{I}_3)\boldsymbol{C}_{st}\Phi_t).$$

Here, $\Phi_t^{\mathbb{H}} \in \mathbb{R}^{3N_d \times N_f}$ denotes the Hermitian transpose of $\Phi_t$ (Gramfort et al., 2013).

**Algorithm 1.** FISTA algorithm used to compute the STOUT estimate.

```
Require: $\boldsymbol{Z} = \boldsymbol{0}$, with all-zeros matrix $\boldsymbol{Z} \in \mathbb{C}^{3N_d \times N_f}$. $\lambda_s, \lambda_t \in \mathbb{R}^+$
Ensure: $\boldsymbol{X} = \lambda_s(\boldsymbol{\Phi}_s \otimes \boldsymbol{I}_3)^\top \boldsymbol{L}^\top(\boldsymbol{Y} - \boldsymbol{L}(\boldsymbol{\Phi}_s \otimes \boldsymbol{I}_3)\boldsymbol{C}_{st}\boldsymbol{\Phi}_t)\boldsymbol{\Phi}_t^{\mathbb{H}}$,
Ensure: $\tau = 1$
  while $||\boldsymbol{Z} - \boldsymbol{Z}_0||_F / ||\boldsymbol{Z}_0||_F >$ tolerance do
    $\boldsymbol{Z}_0 \Leftarrow \boldsymbol{Z}$
    $\boldsymbol{Z} \Leftarrow \mathrm{prox}_{\ell_1+\ell_{1,2}}(\boldsymbol{C}_{st} + \boldsymbol{X}, \lambda_s, \lambda_t)$
    $\tau_0 \Leftarrow \tau$
    $\tau \Leftarrow 1 + \sqrt{1 + 4\tau_0^2}/2$
    $\boldsymbol{C}_{st} \Leftarrow \boldsymbol{Z} + \boldsymbol{Z}_0(\tau_0 - 1)/\tau$
  end while
```

## Appendix B. Depth compensation

It is widely accepted that depth compensation (see the Depth compensation section) is essential for reducing the location bias of the estimated sources when using general norm penalized inverse imaging approaches. However, the efficacy of the various weighting matrices proposed in the literature has rarely been studied. Here we compare the column-norm weighting used in Gramfort et al. (2013) with the sLORETA-based weighting proposed in Haufe et al. (2008b). The latter is based on the estimated dipole-wise variance of the sources used in Pascual-Marqui (2002) to standardize the dipole-wise source power. The variance is estimated from the lead field as $\hat{\Upsilon} \in \boldsymbol{L}^{\mathrm{T}}\left(\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}\right)^{-1}\boldsymbol{L}$. Then the depth compensation matrix is defined as

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{W}_1 & \dots & 0 \\ \vdots & \boldsymbol{W}_2 & \vdots \\ 0 & \dots & \boldsymbol{W}_3 \end{pmatrix}, \tag{B.1}$$

where the $i$th block element $\boldsymbol{W}_i \in \mathbb{R}^{N_d \times N_d}$ of $\boldsymbol{W}$ corresponds to the matrix square root of the $i$th block element of $\hat{\Upsilon}$, which corresponds to the directions $x$, $y$ and $z$ (Haufe et al. (2008b, 2011)).

We generated 100 trials of noisy pseudo-EEG data with a single active source according to the simulation scheme explained in the Simulated ERP data section. Fig. B.9 shows the earth mover's distance between the true and the estimated current densities for the column-norm based depth weighting according to Eq. (7), the sLORETA-based depth weighting according to Eq. (B.1), and without depth weighting. As can be seen, adopting either depth weighting strategy similarly improves the reconstruction accuracy as compared to not adopting any weighting scheme. Note that the equivalence of column-norm and sLORETA-based weighting stands in contrast to the results of Haufe et al. (2008b), which show a general superiority of the sLORETA-based weighting scheme. This discrepancy might have two reasons. First, the $\zeta$ parameter in Eq. (7) was set to $\zeta = 0.3$ here, while it was set to $\zeta = 1$ in Haufe et al. (2008b). Second, we here consider only cortical sources, while the study of Haufe et al. (2008b) places sources in the full 3D brain volume. The average depth of cortical sources is lower than for sources in the full volume, which could be the reason for the column-norm weighting achieving good performance here.

## Appendix C. Measuring stationarity of EEG data

The Kullback–Leibler (KL) divergence $d_{KL}(P||Q) = \sum_i P(i)\frac{P(i)}{Q(i)}$ is a measure of dissimilarity between probability distributions $P$ and $Q$ that can be used to quantify the degree of (non-) stationarity of a multivariate time series, based on the second order stationary definition. Restricting the analysis to the first two statistical moments, a measure $\xi \in \mathbb{R}$ of non-stationarity can be derived, which makes use of an analytic expression for the KL divergence between multivariate Gaussian distributions (note that this does not imply that the data is assumed to be Gaussian distributed):

$$\xi = \sum_{i=1}^{Ne} d_{KL}\left\{\mathcal{N}(\mu_i, \boldsymbol{\Sigma}_i), \mathcal{N}(\overline{\mu}, \overline{\boldsymbol{\Sigma}}\right\}.$$

Here, the data are supposed to be split into $N_e$ segments, where $\mu_i \in \mathbb{R}^{N_c}$ and $\boldsymbol{\Sigma}_i \in \mathbb{R}^{N_c \times N_c}$ denote the empirical mean and covariance of each segment, respectively. Analogously, $\overline{\mu}$ and $\overline{\boldsymbol{\Sigma}}$ denote the mean and covariance averaged across the entire EEG recording. Generally, higher values of $\xi$ indicate a higher degree of non-stationarity of the analyzed recording.

We use the non-stationarity index $\xi$ to verify that the simulation protocol described in the Simulated ERP data section indeed induces non-stationary behavior. To this end, we partition the simulated pseudo-EEG data into $N_e = 5$ segments each containing 40 samples approximately. As the number of EEG channels is higher than the number of samples per segment, which leads the KL divergence to be undefined, we use an upper bound of the non-stationary measure, computed as Hara et al. (2012):

$$\xi = Tr\left\{\frac{1}{N_e}\sum_{i=1}^{N_e}\left\{\mu_i\mu_i^T + 2\boldsymbol{\Sigma}_i\overline{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}_i\right\} - \overline{\mu}\overline{\mu}^T - 2\overline{\boldsymbol{\Sigma}}\right\}.$$

Finally, we set $\overline{\boldsymbol{\Sigma}}$ to be an identity matrix through a whitening procedure. As $\overline{\boldsymbol{\Sigma}}^{-1}$ is computationally infeasible, we use a Moore–Penrose pseudoinverse $\overline{\boldsymbol{\Sigma}}^{\dagger}$. Thus, the proposed measure can be interpreted as the variance of the mean and covariance across all epoch, which is directly related with the second order stationarity definition, thus, the greater the $\xi$ value, the more (non-)stationary the analyzed multichannel time–series.

Fig. C.10 depicts the degree of (non-) stationarity for varying SNR (as defined by the numbers of trials in the average) as well as for varying numbers of simulated ERP sources for a fixed SNR of 5 dB. As expected, higher SNR values and larger numbers of sources lead to increased non-stationarity at the sensor level as indicated by larger $\xi$ values.

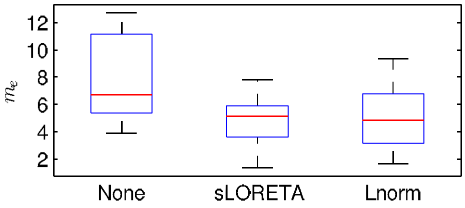


**Fig. B.9** Earth mover's distance achieved for different depth compensation methods and without depth compensation.

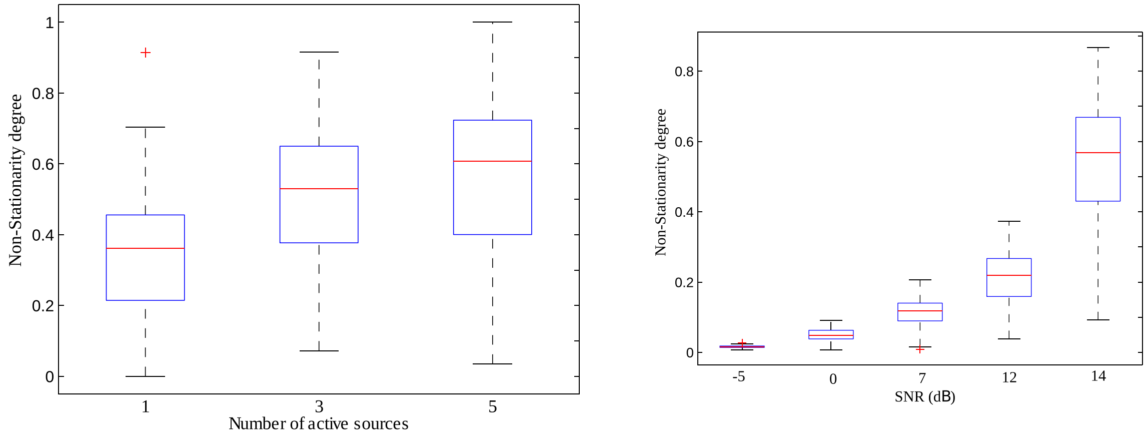


**Fig. C.10.** Degree of non-stationarity ξ in the simulated event-related potential (ERP) data as a function of the number of simulated ERP sources (left part) and as a function of the signal-to-noise ratio (SNR) as defined by the numbers of trials in the average (right part). Larger values of the non-stationarity index indicate a higher degree of non-stationarity.

## References

An, X.-W., Höhne, J., Ming, D., Blankertz, B., 2014. Exploring combinations of auditory and visual stimuli for gaze-independent brain-computer interfaces. PLoS One 9 (10), e111070.

Babadi, B., Obregon-Henao, G., Lamus, C., Hämäläinen, M.S., Brown, E.N., Purdon, P.L., 2014. A subspace pursuit-based iterative greedy hierarchical solution to the neuromagnetic inverse problem. NeuroImage 87, 427–443.

Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Proc. Mag. 18, 14–30.

Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. 2 (1), 183–202.

Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.-R., 2011. Single-trial analysis and classification of ERP components — a tutorial. NeuroImage 56 (2), 814–825 (multivariate Decoding and Brain Reading).

Bolstad, A., Veen, B.V., Nowak, R., 2009. Space–time event sparse penalization for magneto-/electroencephalography. NeuroImage 46 (4), 1066–1081.

Dannhauer, M., Lämmel, E., Wolters, C.H., Knösche, T.R., 2013. Spatio-temporal regularization in linear distributed source reconstruction from EEG/MEG: a critical evaluation. Brain Topogr. 26 (2), 229–246.

Ding, L., He, B., 2008. Sparse source imaging in EEG with accurate field modeling. Hum. Brain Mapp. 29 (9), 1053–1067.

Durka, P.J., Matysiak, A., Montes, E.M., Sosa, P.V., Blinowska, K.J., 2005. Multichannel matching pursuit and EEG inverse solutions. J. Neurosci. Methods 148 (1), 49–59.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. Ann. Stat. 32 (2), 407–451.

Folstein, J.R., Van Petten, C., 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. Psychophysiology 45 (1), 152–170.

Fonov, V., Evans, A., McKinstry, R., Almli, C., Collins, D., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. NeuroImage 47.

Fonov, V., Evans, A., Botteron, K., Almli, C., McKinstry, R., Collins, D., BDCG, 2011. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 54.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39 (3), 1104–1120.

Galka, A., Yamashita, O., Ozaki, T., Biscay, R., Valdés-Sosa, P., 2004. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. NeuroImage 23 (2), 435–453.

Gorodnitsky, I.F., George, J.S., Rao, B.D., 1995. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. Electroencephalogr. Clin. Neurophysiol. 95 (21), 231–251.

Gramfort, A., Kowalski, M., Hämäläinen, M., 2012. Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. Phys. Med. Biol. 57 (7), 1937–1961 (Mar.).

Gramfort, A., Strohmeier, D., Haueisen, J., Hämäläinen, M., Kowalski, M., 2013. Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activations. NeuroImage 70, 410–422.

Grech, R., Cassar, T., Muscat, J., Camilleri, K., Fabri, S., Zervakis, M., Xanthopoulos, P., Sakkalis, V., Vanrumste, B., 2008. Review on solving the inverse problem in EEG source analysis. J. NeuroEng. Rehabil. 5 (25), 792–800.

Habermehl, C., Steinbrink, J., Müller, K.-R., Haufe, S., 2014. Optimizing the regularization for image reconstruction of cerebral diffuse optical tomography. J. Biomed. Opt. 19 (9) (096006).

Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32, 35–42.

Hansen, P.C., 1992. Analysis of discrete ill-posed problems by means of the L-curve. SIAM Rev. 34 (4), 561–580.

Hara, S., Kawahara, Y., Washio, T., von Bünau, P., Tokunaga, T., Yumoto, K., 2012. Separation of stationary and non-stationary sources with a generalized eigenvalue problem. Neural Netw. 33, 7–20.

Haufe, S., Nikulin, V.V., Ziehe, A., Müller, K.-R., Nolte, G., 2008a. Estimating vector fields using sparse basis field expansions. In: Koller, D., Schuurmans, D., Bengio, Y., Bottou, L. (Eds.), Advances in Neural Information Processing Systems 21. MIT Press, Cambridge, MA, pp. 617–624.

Haufe, S., Nikulin, V.V., Ziehe, A., Müller, K.-R., Nolte, G., 2008b. Combining sparsity and rotational invariance in EEG/MEG source reconstruction. NeuroImage 42 (2), 726–738.

Haufe, S., Tomioka, R., Dickhaus, T., Sannelli, C., Blankertz, B., Nolte, G., Müller, K.-R., 2011. Large-scale EEG/MEG source localization with spatial flexibility. NeuroImage 54 (2), 851–859.

Köhler, T., Wagner, M., Fuchs, M., Wischmann, H.-A., Drenckhahn, R., Theißen, A., 1996. Depth normalization in MEG/EEG current density imaging. 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 812–813.

Lin, F.-H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hämäläinen, M.S., 2006. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. NeuroImage 31 (1), 160–171.

Linden, D.E., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F.E., Goebel, R., Dierks, T., 1999. The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. Cereb. Cortex 9 (8), 815–823.

Matsuura, K., Okabe, Y., 1995. Selective minimum-norm solution of the biomagnetic inverse problem. IEEE Trans. Biomed. Eng. 42, 608–615.

Meinshausen, N., 2007. Relaxed lasso. Comput. Stat. Data Anal. 52 (1), 374–393.

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115 (10), 2195–2222.

Miwakeichi, F., Martinez-Montes, E., Valdés-Sosa, P.A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y., 2004. Decomposing EEG data into space–time–frequency components using parallel factor analysis. NeuroImage 22 (3), 1035–1045.

Nagarajan, S.S., Attias, H.T., Hild, K.E., Sekihara, K., Sep. 2007. A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data. Stat. Med. 26 (21), 3886–3910.

Neuper, C., Klimesch, W., 2006. Event-related dynamics of brain oscillations. Event-Related Dynamics of Brain Oscillations. Elsevier Science.

Nolte, G., Dassios, G., 2005. Analytic expansion of the EEG lead field for realistic volume conductors. Phys. Med. Biol. 50 (16), 3807.

Nummenmaa, A., Auranen, T., Hämäläinen, M.S., Jääskeläinen, I.P., Lampinen, J., Sams, M., Vehtari, A., 2007. Hierarchical bayesian estimates of distributed MEG sources: theoretical aspects and comparison of variational and MCMC methods. NeuroImage 35 (2), 669–685.

Nunez, P.L., Srinivasan, R., 2005. Electric Fields of the Brain: The Neurophysics of EEG. Oxford University Press, Oxford.

Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin. Neurophysiol. 112, 713–719.

Ou, W., Hämäläinen, M.S., Golland, P., 2008. A distributed spatio-temporal EEG/MEG inverse solver. NeuroImage 44, 932–946.

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24, 5–12 (Suppl. D).

Pascual-Marqui, R., Michel, C.M., Lehman, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18 (18), 49–65.

Søndergaard, P.L., Torrésani, B., Balazs, P., 2012. The linear time frequency analysis toolbox. Int. J. Wavelets Multiresolut. Anal. Inf. Process. 10 (4).

Tibshirani, R., 1994. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B 58, 267–288.

Toga, A., Mazziotta, J., 2002. Brain Mapping: The Methods. 2nd Edition. Elsevier Science.

Trujillo-Barreto, N.J., Aubert-Vázquez, E., Penny, W.D., 2008. Bayesian M/EEG source reconstruction with spatio-temporal priors. NeuroImage 39 (1), 318–335.

Vega-Hernández, M., Martínez-Montez, E., Sánchez-Bornot, J.M., Lage-Castellanos, A., 2008. Penalized least squares methods for solving the EEG inverse problem. Stat. Sin. 18 (18), 1535–1551.

Wipf, D., Nagarajan, S., 2009. A unified bayesian framework for MEG/EEG source imaging. NeuroImage 44 (3), 947–966.

Woolrich, M.W., Baker, A., Luckhoo, H., Mohseni, H., Barnes, G., Brookes, M., Rezek, I., 2013. Dynamic state allocation for MEG source reconstruction. NeuroImage 77, 77–92.

Zwoliński, P., Roszkowski, M., Żygierewicz, J., Haufe, S., Nolte, G., Durka, P.J., 2010. Open database of epileptic EEG with MRI and postoperational assessment of foci—a real world verification for the EEG inverse solutions. Neuroinformatics 8, 285–299.