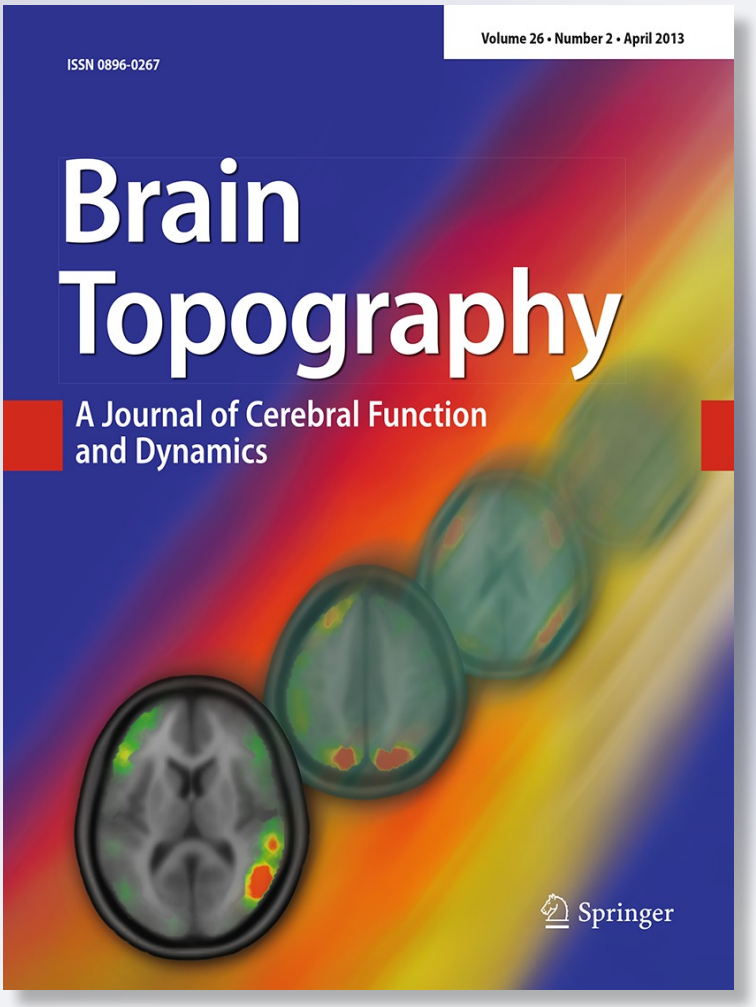
Spatio-temporal Regularization in Linear Distributed Source Reconstruction from EEG/MEG: A Critical Evaluation

Moritz Dannhauer, Eric Lämmel, Carsten H. Wolters & Thomas R. Knösche

Brain Topography
A Journal of Cerebral Function and Dynamics

ISSN 0896-0267
Volume 26
Number 2

Brain Topogr (2013) 26:229-246
DOI 10.1007/s10548-012-0263-9



Springer

Brain Topogr (2013) 26:229–246
DOI 10.1007/s10548-012-0263-9

ORIGINAL PAPER

Spatio-temporal Regularization in Linear Distributed Source Reconstruction from EEG/MEG: A Critical Evaluation

Moritz Dannhauer · Eric Lämmel · Carsten H. Wolters · Thomas R. Knösche

Received: 25 June 2012 / Accepted: 13 October 2012 / Published online: 31 October 2012

Abstract The high temporal resolution of EEG/MEG data offers a way to improve source reconstruction estimates which provide insight into the spatio-temporal involvement of neuronal sources in the human brain. In this work, we investigated the performance of spatio-temporal regularization (STR) in a current density approach using a systematic comparison to simple ad hoc or post hoc filtering of the data or of the reconstructed current density, respectively. For the used STR approach we implemented a frequency-specific constraint to penalize solutions outside a narrow frequency band of interest. The widely used sLORETA algorithm was adapted for STR and generally used for source reconstruction. STR and filtering approaches were evaluated with respect to spatial localization error and spatial dispersion, as well as to correlation of original and reconstructed source time courses in single source and two source scenarios with fixed source locations and oscillating source waveforms. We used extensive computer simulations and tested all algorithms with different parameter settings (noise levels and regularization parameters) for EEG data. To verify our results, we also used data from MEG phantom measurements. For the investigated scenarios, we did not find any evidence that STR-based methods outperform purely spatial algorithms applied to temporally filtered data. Furthermore, the results show very clearly that the performance of STR depends very much on the choice of regularization parameters.

M. Dannhauer · E. Lämmel · T. R. Knösche (✉)
Max Planck Institute for Human Cognitive and Brain Sciences,
P.O. Box 500355, 04303 Leipzig, Germany
e-mail: knoesche@cbs.mpg.de

C. H. Wolters
Institute for Biomagnetism and Biosignalanalysis,
University of Münster, 48149 Münster, Germany

Keywords Source localization · Spatio-temporal regularization · sLORETA · EEG

Introduction

There are many techniques for mapping neural function, which differ in numerous aspects including the physiological processes they observe, their degree of invasiveness as well as their spatial and temporal resolution. Among these, MEG and EEG are directly sensitive to electrophysiological neural mass action, are completely non-invasive and provide excellent temporal, and moderate spatial resolution. These properties make them attractive for studying the dynamics of neural networks in both neuroscientific experiments and clinical conditions, for example epileptic seizures. Therefore, the questions of spatial location and temporal evolvement of the underlying neuronal generators of EEG and MEG are of great interest.

In the last few decades, a considerable range of methodologies for the reconstruction of these generators, or sources, has emerged. All these methods are based on a *forward model*, which provides a physical description of the electromagnetic properties of the head tissues[1] and the sensors (usually their positions and, in case of MEG, additional properties, such as coil orientations) and thereby links the neuronal activity to the EEG/MEG measurements. Depending on the quality and quantity of the information available on head geometry and tissue conductivities, various head modeling techniques can be used, ranging from analytical equations for a simple

[1] As for the frequency ranges normally considered in EEG and MEG, the temporal derivatives in the Maxwell equations can be ignored (quasistatic assumption; Plonsey and Heppner 1967), these properties amount to the spatial distribution of the electrical conductivity.

spherical approximation of the head (de Munck and Peters 1993) to more complex and individually shaped head models, which need to be treated numerically (e.g., by using boundary element modeling (Fuchs et al. 2007; Huiskamp et al. 1997, 1999; Kybic et al. 2005; Zanow 1997) or finite element modeling techniques (Awada et al. 1997; Buchner et al. 1997; Güllmar et al. 2006, 2010; Lew et al. 2009; Marin et al. 1998; Vallaghe and Papadopoulo 2010; van den Broek et al. 1998; Wolters et al. 2007b). Subsequently, the *inverse problem* has to be solved. For this purpose, the neuronal sources are described by a *source model*, and the parameters of this model (e.g., locations, orientations and time course) are estimated from the measurements. The most universally used building block for such source models is the *current dipole*, which describes a quasi point-like current flow and is characterized by the location and orientation of this current as well as by the temporal evolvement of the current strength (de Munck et al. 1988; Lew et al. 2009; Wolters et al. 2007a, b). Combinations of multiple current dipoles can be used to describe virtually any possible source configuration underlying EEG and MEG.

In general, the solution of the inverse problem suffers from two fundamental problems: they are *non-unique* and *ill-conditioned*. The former describes the fact that, no matter how much data is acquired and with what accuracy this is done, there is no unique solution to the inverse problem. Instead, we have to deal with a multidimensional manifold of solutions and additional criteria must be used to select a representative solution from this manifold. The latter problem refers to the fact that the problem is often ill-conditioned and therefore noise in the data may have a large effect on the solution. Both non-uniqueness and bad condition are closely interrelated and refer to the eigenvalue spectrum of the problem. Solutions to these problems are commonly referred to as regularization.

Distributed Source Modeling

As already stated, solving the inverse problem requires specification of a *source model*, which, in most cases, is based upon current dipoles. There are two main strategies. The first is based on the optimization of a small number of source parameters (source locations, source orientations, and source time courses), which are considerably fewer in number than the measurement channels (e.g., Scherg and von Cramon 1986). The second main strategy involves the reconstruction of the strengths of a large number of dipolar sources, which densely cover the space of possible source locations and orientations. In this approach, which is known as *distributed source model*, there are usually many more source parameters than measurements. The resulting inverse problem is underdetermined and hence has no unique solution. This approach is the focus of the current paper. To make the solution unique, additional criteria for the selection of the solution have to be introduced.

A widely used criterion is the *minimum norm constraint* (Hämäläinen and Ilmoniemi 1994, see also "Method" section). Here, on the manifold of possible solutions, the one with the smallest L_2 norm is selected. This approach has been very popular since it leads to a linear formulation based on the Moore–Penrose pseudoinverse. However, the minimum norm does not treat all brain regions that might contribute to the observed signal equally. Instead, it suffers from a *depth bias*, which means that the algorithm preferentially composes the reconstructed activity of sources that have a major impact on the sensor outputs, like superficial rather than deep sources. This phenomenon can be partially compensated for by the introduction of certain weighting matrices, such as in weighted minimum norm estimation (WMN: Fuchs et al. 1999; Jeffs et al. 1987, see Eq. 1), which ensures that all dipoles have the same average impact on the sensors. Since the introduction of minimum norm estimation for EEG/MEG data, a whole range of linear distributed source reconstruction methods have been proposed to further improve localization performance and to avoid methodological biases by applying different a priori (Pascual-Marqui 2007; Pascual-Marqui et al. 1994) as well as a posteriori (Pascual-Marqui 2002) weighting schemes. Such weighting schemes can embody different kinds of additional information and may be formulated to select solutions with very specific properties, such as the spatially smoothest solution (LORETA, Pascual-Marqui et al. 1994). It is often possible to partially replace these rather arbitrary mathematical minimum norm constraints by using additional knowledge on the source parameters, drawn from alternative measurement modalities or from general anatomical or physiological knowledge (Ahlfors and Simpson 2004; Dale et al. 2000; Dale and Serano 1993; Liu et al. 1998).

As explained above, usable source reconstruction methods have to cope not only with the non-uniqueness problem, but also with the low condition (i.e., the instability of the solution towards uncertainties in the measurements). In reality, EEG and MEG signals originating from the brain processes targeted by the respective experiment are contaminated with large quantities of noise. In particular, source localization of single trials is not a simple matter, due to high levels of noise. In order to make the solutions unique and well-conditioned (see above), one has to apply appropriate constraints together with regularization methods. For example, a widely used technique involves the Moore–Penrose pseudo-inverse in combination with Tikhonov regularization.

Spatio-temporal Regularization (STR)

Most distributed source reconstruction methods solve the inverse problem separately for each time sample. These, therefore do not take advantage of any reasonable assumptions on the temporal evolvement of the activity.

However, in order to profit from the high temporal resolution of EEG/MEG, a considerable amount of research has been conducted in order to include temporal information about the source activity in discrete source models to further reduce the space of possible solutions (Baillet and Garnero 1997; Scherg and Berg 1991). Temporal information enhances the stability of these solutions (Zhang et al. 2005) against disturbances and provides insight into the dynamics of neuronal sources (Scherg 1990).

There also have been attempts to incorporate temporal constraints into distributed source reconstruction schemes. Brooks et al. (1999) proposed a method for electrocardiography, which introduces a temporal constraint to the minimum L_2 norm problem. The method works effectively as a low-pass filter, as a smooth time course is favored and high frequency components of the solution are suppressed. Both spatial and temporal regularization can be adjusted using separate regularization parameters. Brooks et al. (1999) also proposed various techniques to reduce the numerical costs by avoiding solving the problem directly and instead using block Jacobi iterative schemes or matrix diagonalization techniques. Based on the work of Brooks et al. (1999), Schmitt and colleagues (Darvas et al. 2001; Schmitt and Louis 2002; Schmitt et al. 2001, 2002) were able to improve this STR approach by further reducing the numerical costs using Sylvester equation solver techniques. Moreover, these authors extended the application of STR to EEG/MEG, computerized tomography and dynamic electrical impedance tomography (Schmitt et al. 2002). They also combined the STR approach with additional spatial constraints such as depth weighting (WMN, Fuchs et al. 1999; Jeffs et al. 1987) and spatial smoothing (LORETA, Pascual-Marqui et al. 1994) to reduce the solution space and the localization errors for deep sources in particular (Schmitt et al. 2001). This STR approach was also implemented in the SimBio software (Schmitt et al. 2004; SimBio 2011).

As these approaches seem to effectively apply a temporal filter to the solution during the reconstruction process, the question arises how the results obtained relate to noise reduction by simple temporal filtering of the data before source reconstruction (ad hoc) or filtering of the source time courses after reconstruction (post hoc). This question is of importance since the computationally more costly STR approach is only useful if it is significantly superior to these simple filtering techniques.

In this paper, we investigate the combination of the STR approach with a currently popular spatial reconstruction method called standardized LORETA (sLORETA) (Pascual-Marqui 2002). It has been shown that sLORETA performs with *zero localization error* (i.e., the peak of the current density coincides with the original dipole position within the accuracy given by the resolution of the reconstruction grid) in the theoretical case of one active source and no measurement noise (Pascual-Marqui 2007). Also, in comparison papers, Lucka et al. (2011) as well as Wagner et al. (2004) showed that the method performs well for single sources and Wagner et al. (2004) additionally reported a good performance for multiple sources, if they are distinct enough and of similar strength. In order to allow a direct comparison between STR and simple filtering, we modified the general approach put forward by Brooks et al. (1999) and Schmitt et al. (2001). These authors used a rather unspecific general smoothness assumption to constrain the solution. In contrast, we designed a specific temporal weighting matrix that embodies a finite impulse response (FIR) filter so that the results of the STR reconstruction can be directly compared to those of classical filtering applied ad hoc to the data or post hoc to the reconstructed time series. As it is well known that particular frequencies, such as alpha, beta or gamma oscillations, which are present in spontaneous brain activation, are relevant for brain function (e.g., Klimesch 1999), it can be sensible to use narrow band-pass filters. For brain activity that is expected to be restricted to certain frequency bands, in particular brain oscillations, such a technique is expected to greatly improve the SNR, thereby strongly regularizing the solution and rendering the technique suitable for the analysis of single trial data.

To deal with the theoretical and practical issues discussed above, we sought to answer the following questions: (1) Is it possible to combine the concept of sLORETA (Pascual-Marqui 2002) with the frequency specific STR framework? (2) Does STR perform better than simple ad hoc or post hoc data filtering? (3) How does STR perform in more realistic settings?

To answer these questions, we use both computer simulations of EEG data and MEG measurements using a physical phantom. Our investigations will include single source and two source scenarios with fixed source locations and oscillating source waveforms.

Methods

Source Analysis with Post-Weighted Minimum Norm Estimation (sLORETA)

Let $X \in \Re^{N \times T}$ denote a real-valued measurement matrix (N representing the number of EEG/MEG sensors, T being the number of samples in time). The spatio-temporal current density $J \in \Re^{M \times T}$ in the brain is modeled by M dipolar sources at K source positions and T samples. At each source position o_k $(k = 1 \ldots K)$, multiple dipoles with different directions can be modeled (usually, $S \in \{1, 2, 3\}$, $M = S \cdot K$) depending on the intended dimensionality of the

source space and the volume conductor properties. The forward solutions for all sources are combined in the leadfield matrix $L \in \Re^{N\times M}$, which serves as an essential ingredient for distributed source reconstruction methods. Under the quasistatic assumption (Plonsey and Heppner 1967), the relationship between source currents and resulting EEG/MEG measurements can be considered linear, hence $X = L{\cdot}J +$ noise. Since the leadfield matrix is generally not square (more specifically, there are more sources than measurement channels) the solution to the inverse problem is not unique (i.e., there is an entire manifold of possible solutions). The linear minimum norm estimation (MNE) approach (Hämäläinen and Ilmoniemi 1994) selects one solution out of this manifold, which not only fits the measurement data X, but simultaneously has a minimum L_2 norm of some weighted version of the current density:

$$\begin{aligned}\hat{j}^{MNE} &= \underset{j}{\operatorname{minarg}}(\|L\cdot j - x\|_2^2 + \lambda\cdot\|W\cdot j\|_2^2)\\ &= \left(W^T\cdot W\right)^{-1}{\cdot}L^T\cdot\left(L\cdot\left(W^T\cdot W\right)^{-1}{\cdot}L^T + \lambda\cdot I_N\right)^{-1}\cdot x\\ &= G\cdot x\end{aligned} \tag{1}$$

In Eq. (1), $W \in \Re^{M\times M}$ is a spatial weighting matrix, $I_N \in \Re^{N\times N}$ denotes the identity matrix with dimension N and $\langle\cdot\rangle^T$ indicates a matrix transpose. The spatial regularization parameter λ controls the relative influences of the *error term* (first summand, which measures how well the model $L{\cdot}j$ explains the data x) and the *model term* (second summand, which measures the norm of the weighted solution). Low λ reflects high confidence in the data and allows for a more detailed solution, while high λ produces solutions that are more governed by the smoothing influence of the model term. The regularization parameter can be derived from the Bayesian formulation of the problem and represents in that context the ratio between the variance of the data noise and the expected variance of the sources (which is usually not known). It is measured in units of $(\mathrm{V/Am})^2$ for EEG and $(\mathrm{T/Am})^2$ for MEG. As mentioned before, W can be specified to select a solution with specific properties, e.g., the smoothest solution (LORETA; Pascual-Marqui et al. 1994) or a solution that guarantees unbiased localization of single dipolar sources from noise-free data (eLORETA; Pascual-Marqui 2007). Hence, the spatial weighting matrix W represents prior knowledge or assumptions on the sources.

Besides the pre-weighting of the current density distribution, there have been some attempts to use post-weighting strategies. One popular example is the sLORETA method (Pascual-Marqui 2002), which estimates the current density distribution as a statistical map for each source k individually:

$$\hat{P}_k^{sLORETA} = \left(\hat{j}_k^{MNE}\right)^T\cdot R_k^{-1}\cdot\hat{j}_k^{MNE} \tag{2}$$

with the resolution kernel for a single source position $R_k \in \Re^{S\times S}$ and minimum norm estimate $\hat{j}_k^{MNE} \in \Re^{1\times S}$. Note that $\hat{j}^{MNE}$ denotes the minimum norm current density estimate and $\hat{P}^{sLORETA}$ stands for the estimated normalized current density power (sLORETA).

In the following, it will be described how the entries of $R_k \in \Re^{S\times S}$ are determined in the sLORETA approach.

Minimum norm reconstructions of focal sources are generally biased (i.e., blurred and often shifted, see, e.g., (Lucka et al. 2011)). The resolution kernel of the entire problem, $R \in \Re^{M\times M}$ which is the product of the inverse operator $G \in \Re^{M\times N}$ and the leadfield matrix $L \in \Re^{N\times M}$, reflects this bias. Since the resolution kernel is singular (Grave de Peralta Menendez et al. 2004), this bias cannot be completely corrected for. However, it is possible to reduce the bias by filtering the current density estimate through the inverse of the block-diagonalized resolution kernel $R^{BD} \in \Re^{M\times M}$, that is only composed of the diagonal block elements $R_k(k = 1\ldots K)$ (Pascual-Marqui 2002). Thus, this block-diagonalized resolution kernels computes as

$$\begin{aligned}R^{BD} &= R\circ(I_K\otimes 1_S) = G\cdot L\circ(I_K\otimes 1_S)\\ &= \begin{pmatrix} R_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & R_K\end{pmatrix}\end{aligned} \tag{3}$$

where $\circ$ is the element-wise product, $1_S \in \Re^{S\times S}$ is a matrix of ones and $\otimes$ is the Kronecker product, for $A \in \Re^{C\times D}$ defined as

$$A\otimes B = \begin{pmatrix} a_{11}B & \ldots & a_{1D}B\\ \vdots & \ddots & \vdots\\ a_{C1}B & \cdots & a_{CD}B\end{pmatrix}. \tag{4}$$

Source Analysis with STR and sLORETA

Brooks et al. (1999) as well as Schmitt and colleagues (Darvas et al. 2001; Schmitt and Louis 2002; Schmitt et al. 2001, 2002) proposed a spatio-temporal extension to the MNE approach, which additionally minimizes the temporal derivative $\left(\left\|\frac{d}{dt}j(t)\right\| \rightarrow \mathrm{Min}\right.$, approximated by a certain filter matrix F of the current density distribution.

While in (Schmitt et al. 2001), a filter matrix $F \in \Re^{(T-1)\times T}$ was used, we use $F \in \Re^{T\times T}$ here. In Schmitt et al. (2001), the authors showed, that the STR problem can be solved in sensor space as well as in source space, and that both representations are equivalent. Because of the much smaller computational costs we decided to solve the STR problem in sensor space:

$$\begin{aligned}\hat{j}_* &= \underset{j_*}{\operatorname{minarg}}(\|L_* \cdot j_* - x_*\|_2^2 + \lambda \cdot \|W_* \cdot j_*\|_2^2 \\ &\quad + \mu \cdot \|(F \otimes I_M) \cdot W_* \cdot j_*\|_2^2) \\ &= \left(W_*^T \cdot W_*\right)^{-1} \cdot L_*^T \cdot (L_* \cdot \left(W_*^T \cdot W_*\right)^{-1} \cdot L_*^T \\ &\quad + \lambda \cdot I_{N \cdot T} + \mu \cdot (F^T \cdot F) \otimes I_N)^{-1} \cdot x_* \\ &= \left(W_*^T \cdot W_*\right)^{-1} \cdot L_*^T \cdot A^{-1} \cdot x_* = G_* \cdot x_* \end{aligned} \quad (5)$$

In Eq. (5), the temporal filter matrix $F \in \Re^{T \times T}$ is introduced and its impact can be effectively adjusted by a second regularization parameter μ. Its function and meaning is analogous to the spatial regularization parameter λ, as explained below Eq. (1). All symbols using the subscript $[\cdot]_*$ denote spatio-temporal operator matrices. They are extended in size multiplicatively by T as compared to their purely spatial counterparts: $W_* = I_T \otimes W \in \Re^{T \cdot M \times T \cdot M}$, $G_* \in \Re^{T \cdot M \times T \cdot N}$, $j_* \in \Re^{M \cdot T \times 1}$, $x_* \in \Re^{N \cdot T \times 1}$. Since sLORETA (Pascual-Marqui 2002) is a purely unweighted minimum norm estimation, a spatial weighting matrix is not considered in any computation, hence $W = I$.

Within that approach, the highest computational costs (cubic complexity) occur while inverting the matrix $A_* \in \Re^{N \cdot T \times N \cdot T}$. Schmitt et al. (2001) suggested an elegant way to convert the STR problem into a so-called Sylvester equation problem, in which an approximate solution can be computed numerically in a very efficient way. Accordingly, we used this approach to implement a spatio-temporal version of sLORETA. The STR problem can be mathematically reformulated (described in Schmitt et al. 2001, Eq. 15) as the following

$$\begin{aligned}&(L_* \cdot L_*^T + \lambda \cdot I_{T \cdot N} + \mu \cdot (F^T \cdot F) \otimes I_N) \cdot u_* = x_*, \hat{j}_* \\ &\quad = L_*^T \cdot u_* \end{aligned} \quad (6)$$

$$(u_*, x_* \in \Re^{M \cdot T \times 1}, \hat{j}_* \in \Re^{M \cdot T \times 1}).$$

This is equivalent to the Sylvester equation system (Schmitt et al. 2001, Eq. 16):

$$(L \cdot L^T + \lambda \cdot I_N) \cdot U + U \cdot (\mu \cdot F^T \cdot F) = X \quad (7)$$

with $U \in \Re^{N \times T}$ and $J^{STR} = L^T \cdot U$, which can be solved using the Sylvester equation solver function "syl" (included in Octave™, which is based on LAPACK subroutines). This way, the spatio-temporal minimum norm solution can be computed very fast and with very low memory consumption, as only $(L \cdot L^T + \lambda \cdot I_N) \in \Re^{N \times N}$, $(F^T \cdot F) \in \Re^{T \times T}$, $L^T \in \Re^{M \times N}$ and $X \in \Re^{N \times T}$ are needed for one computation of J^{STR}. Therefore the usually very large inverse operator matrix G_* is not computed explicitly and also all needed values of the resolution kernels R_k can be computed this way for STR1 (R_k^{STR1}, see approach 3 below) and for STR2 (R_k^{STR2}, see approach 4 below).

Generally, sLORETA standardization is performed for each location and each time step separately, which is just a spatial weighting of the current density. This approach is represented by our method STR1 (see below). Alternatively, we explore a different method, STR2, where we perform a spatio-temporal weighting, which uses the whole reconstructed time series for each dipole activation estimate. For this purpose, the size of $R_k \in \Re^{S \times S}$ for each dipole k is enlarged from $R_k^{STR1} \in \Re^{S \times S}$ to $R_k^{STR2} \in \Re^{S \cdot T \times S \cdot T}$. Each column of the resolution kernel $R = G \cdot L$ represents the spatial blurring of one reconstructed unit dipole. For each column in R_k^{STR2} a single computation (using the Sylvester solver) is done to obtain the spatio-temporal spreading of a single active source (with one specific dipole orientation) at a certain time step as explained in more detail below. In the following computation, the data matrix $X \in \Re^{N \times T}$ contains only zeros except for the one column associated to the respective time step (forward solution of a unit dipole for a particular location pointing to a fixed direction is inserted there). Through the mapping of the resulting U onto the forward solution of all (e.g. $S = 3$) dipole directions the spatio-temporal spreading over all considered time steps for each dipole direction (x, y, z) can be computed:

$$L(k)^T \cdot U = \begin{bmatrix} \hat{j}_{t=1}^{STR\_x} & \cdots & \hat{j}_{t=T}^{STR\_x} \\ \hat{j}_{t=1}^{STR\_y} & \cdots & \hat{j}_{t=T}^{STR\_y} \\ \hat{j}_{t=1}^{STR\_z} & \cdots & \hat{j}_{t=T}^{STR\_z} \end{bmatrix} \in \Re^{S \times T} \quad (8)$$

$$L(k) = [L_x(k) L_y(k) L_z(k)] \in \Re^{N \times S}.$$

This spreading matrix, which represents the impact of the used filter, is then reshaped and stored as one column of R_k^{STR2}. The columns of R_k^{STR2} are sorted as follows, e.g. ith column of R_k^{STR2}, $S = 3$

$$\begin{aligned} R_k^{STR2}(i) &= \left[\hat{j}_{t=1}^{STR\_x} \hat{j}_{t=1}^{STR\_y} \hat{j}_{t=1}^{STR\_z} \ldots \hat{j}_{t=T}^{STR\_x} \hat{j}_{t=T}^{STR\_y} \hat{j}_{t=T}^{STR\_z}\right]^T \\ &\quad \in \Re^{S \cdot T \times 1}. \end{aligned} \quad (9)$$

This guarantees that the spatio-temporal spreading, which belongs to the particular dipole, can be computed using the *syl* function. To fill the spatio-temporal resolution matrix for dipole k, $S \cdot T$ of these computations have to be performed. Finally, this procedure must be performed for all M dipoles resulting in $R^{STR2} \in \Re^{M \cdot T \times M \cdot T}$. This means that the computational complexity increases quadratically with the used filter length. The procedure can be subdivided into small computational tasks and distributed over several computers or processors with standard memory requirements. In contrast, the direct solution of the problem is not feasible, even for comparatively small scenarios. Finally, note that the STR2 approach effectively computes a weighted average over all time samples for each dipole.

We used the spatial localization error to select the optimum regularization parameters. There are several other methods to find appropriate regularization parameters (e.g., generalized cross validation and L-curve/L-surface: Brooks et al. (1999); Schmitt et al. (2001)) but these have been criticized (e.g., Hansen and O'Leary 1993) for their poor performance. Please note, however, that in real situations the true source positions, and therefore also the true localization error, are not known and other methods, though imperfect, have to be utilized.

The temporal filter matrix F, proposed by Brooks et al. (1999) and similarly by Schmitt et al. (2001), performs high-pass filtering, which penalizes high frequencies and hence selects solutions that are smooth between successive samples.

We extend this concept to focusing on a specific frequency band by a simple exchange of the filter matrix. The temporal filter matrix F lets pass certain parts of the frequency content of the solution (see Eq. 5) and therefore causes the temporal regularization term to penalize exactly these frequencies. For example, if we want to reconstruct oscillatory activity with frequencies in the alpha band around 10 Hz, the filter should be a band-stop, which lets all frequencies except 10 Hz pass, so that they can be minimized in the current density estimate. The filter matrix F is a diagonal band matrix, each row of which contains the FIR filter coefficients, shifted by the respective time index. This, of course, generates problems for the first and last few time samples, referred to as edge effects. Brooks et al. (1999) used modifications (for the simple high-pass filter) to avoid edge effects, which are not easily transferable to our more complex band-stop filters. Therefore, at the upper left and at the lower right corner of the F matrix, some FIR coefficients were simply left out in such a way that at least half of the filter coefficients are used for each row of F. In this way, the matrix generated remains invertible. In simulations, it was shown that this simplification plays no role in the reconstruction result. The filter matrix scheme was used for all computations (see Tables 1 and 2). Based on simulations, we verified that the inverse of the band-stop filter matrix had exactly the opposite filtering effect and behaved as a band-pass filter. Because of the structure of the filter matrix, it is guaranteed that the condition of the problem is not affected and therefore stronger regularization is not needed.

All FIR filter coefficients were computed using the program *xfir* (EEProbe tools, ANT 2003).

Table 1 31 FIR coefficient from the band-stop filter used (EEG simulations)

$f_{n+15} = f_{n-15}$	−1.26e−02
$f_{n+14} = f_{n-14}$	5.46e−02
$f_{n+13} = f_{n-13}$	2.53e−02
$f_{n+12} = f_{n-12}$	4.41e−02
$f_{n+11} = f_{n-11}$	5.90e−02
$f_{n+10} = f_{n-10}$	6.77e−02
$f_{n+9} = f_{n-9}$	6.89e−02
$f_{n+8} = f_{n-8}$	6.20e−02
$f_{n+7} = f_{n-7}$	4.79e−02
$f_{n+6} = f_{n-6}$	2.83e−02
$f_{n+5} = f_{n-5}$	5.59e−02
$f_{n+4} = f_{n-4}$	−1.76e−02
$f_{n+3} = f_{n-3}$	−3.88e−02
$f_{n+2} = f_{n-2}$	−5.55e−02
$f_{n+1} = f_{n-1}$	−6.63e−02
f_n	9.30e−01

See "Methods" section for further explanation

Table 2 Band-stop filter schematics (EEG simulations, similar for MEG simulations)

f_n	f_{n+1}	…	f_{n+15}	0	0	0	0	0	0
f_{n-1}	f_n	f_{n+1}	…	f_{n+15}	0	0	0	0	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
0	f_{n-15}	…	f_{n-1}	f_n	f_{n+1}	…	f_{n+15}	0	0
0	0	f_{n-15}	…	f_{n-1}	f_n	f_{n+1}	…	f_{n+15}	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	f_{n-15}	…	f_{n-1}	f_n	f_{n+1}
.	.	.	.	.	.	f_{n-15}	…	f_{n-1}	f_n

Evaluation by Computer Simulations

As stated in the introduction, the source reconstruction techniques with spatio-temporal regularization need to be compared with their purely spatial counterparts (i.e., sLORETA) where the respective temporal filter is applied ad hoc to the data or post hoc to the reconstruction result. Hence, if G is the purely spatial inverse operator, G_* is the spatio-temporal inverse operator and F is the respective temporal filter matrix, we compare the following approaches:

1. *Spatial sLORETA (referring to as "Gx")*: Application of a purely spatial inverse operator without filtering: $\hat{J} = G \cdot X$, and followed by a sLORETA postweighting described in (Pascual-Marqui 2002).
2. *Simple filtering (referring to as "GFx")*: This is here realized as ad hoc filtering, i.e., application of spatial inverse operator (as in approach 1) with prior filtering in data space with the inverse of the filter used in the penalty term of Eq. (5): $\hat{J} = G \cdot (X \cdot (FF^T)^{-1})$, with $F \in \Re^{T \times T}$, and followed by a sLORETA postweighting described in (Pascual-Marqui 2002). Note that post hoc filtering, i.e., the application of spatial inverse operator with posterior filtering in source space: $\hat{J} = (G \cdot X) \cdot (FF^T)^{-1}$, is equivalent to ad hoc filtering, due to matrix associativity.

3. *STR + spatial sLORETA (referring to as "STR1")*: Application of the spatio-temporal inverse operator $\hat{j}_k^{STR} = G_* \cdot x_*$, $\hat{j}_*^{STR} \in \Re^{T \cdot M \times 1}$, and followed by a post-weighting [adapted from (Pascual-Marqui 2002)] using only the spatial weights from a spatio-temporal resolution matrix and reshaped to $\hat{J}_k^{STR1} \in \Re^{M \times T}$, see details below.
4. *STR + spatio-temporal sLORETA (referring to as "STR2")*: Using $\hat{j}_*^{STR}$ (from approach 3) followed by full spatio-temporal postweighting [adapted from (Pascual-Marqui 2002)] for each dipole resulting in $\hat{J}^{STR2} \in \Re^{M \times 1}$, see details below.

For purely spatially weighted source localization (approaches 1 and 2), the same inverse operator $G = L^T \cdot (L \cdot L^T + \lambda \cdot I)^{-1}$ and resolution kernel R is used. In order to compute $\hat{J}^{STR}$ (needed for approach 3 and 4), a single computation has to be performed using the Sylvester equation solver. All values for spatio-temporal standardization using the resolution kernel are efficiently computed (see above) in advance for all tested combinations of regularization parameters λ and μ. For STR1 (approach 3) all needed spatio-temporal weights are determined based on the previously computed STR2 weights. The current density reweighting (sLORETA) for STR1 (approach 3) and STR2 (approach 4) is performed as $\hat{P}_k^{STR1}(t) = \hat{j}_k^{STR}(t) \cdot (R_k^{STR1}(t))^{-1} \cdot \hat{j}_k^{STR}(t)^T$ and $\hat{P}_k^{STR2}(t) = reshape(\hat{j}_k^{STR}) \cdot (R^{STR2})^{-1} \cdot reshape(\hat{j}_k^{STR})^T$, respectively.

In the latter approach, $\hat{j}_k^{STR} \in \Re^{S \times T}$ is reshaped to a vector, $\hat{j}_k^{STR} \in \Re^{1 \cdot S \times T}$, to fit matrix dimensions used for STR2, $R_k^{STR2} \in \Re^{S \cdot T \times S \cdot T}$, by matching corresponding dipole directions and time steps. Therefore, as a final result, we get $\hat{P}_k^{STR1}(t) \quad (\forall t = 1, .., T, \forall k = 1, .., K)$ and $\hat{P}_k^{STR2} \quad (\forall k = 1, .., K)$. In order to evaluate the performance of STR2 (involving temporal weighting in both ad hoc and post hoc operators), we computed the time averages of the approaches 1 through 3. This is motivated by the fact that the STR2 approach effectively computes a weighted temporal average. Thus, for a fair comparison, the results of the other methods should be based on temporal averages as well.

In all simulations, we used uncorrelated normally distributed noise (*randn* function, *octave*), which was added to the simulated signal at each sensor. The noise level was described by the relative noise amplitude, defined as the standard deviation of noise divided by the standard deviation of the signal (over all sensors and time steps) $RNA = \frac{stddev(noise)}{stddev(signal)}$ (the proportion of noise in the simulations was not expected to be higher than the signal). For all single source simulations, we used the same 100 noise realizations, which were scaled for each defined noise level (RNA = 0.01, 0.1, 0.2, 0.3, 0.4 and 0.5). The resulting solutions were analyzed with respect to the localization errors and spatial dispersion (Molins et al. 2008), which were averaged over all time samples considered (approaches 1–3). In the case of comparing the proposed STR2 approach to its counterparts (approaches 1–3) the weighted current density is averaged before evaluating the spatial localization error. The averaged localization error over all noise realizations for each noise level was compared for all used (depending on the approach and the imaging modality) regularization parameter combinations. This means, for approach 3 (STR1) and approach 4 (STR2), all combinations λ and μ (for EEG: $\log_{10}\mu = \log_{10}\lambda = [-2, 0, 2, 4, 6, 8, 10, 12, 14, 16]$; for MEG: $\log_{10}\mu = \log_{10}\lambda = [-6, -4, -2, 0, 2, 4, 6, 8, 10, 12]$) and for approaches 1 and 2 (spatial sLORETA with or without simple filtering), only the defined values of λ were used. Large ranges of values for λ and μ were used to cover both over- and underregularization.

In the following, we will describe the simulation scenarios employed in greater detail.

Localization of Simulated EEG Measurements from Oscillating Sources

For volume conductor modeling, an individual 1 mm geometry-adapted hexahedral finite element model was used, which had been prepared for the same subject in Dannhauer et al. (2011) (subject 3, reference model). The skin and brain compartments were modeled homogeneously with conductivities commonly used in the literature, $\sigma_{Skin} = 0.43$ S/m and $\sigma_{Brain} = 0.33$ S/m (Ramon et al. 2006). From a T2-weighted magnetic resonance image, an estimation of soft and hard bone tissue distribution could be segmented and modeled separately, with isotropic conductivities $\sigma_{Soft\,Bone} = 0.02865$ S/m and $\sigma_{Hard\,Bone} = 0.0064$ S/m (Akhtari et al. 2002; Fuchs et al. 2007). The electrodes were placed on $N = 118$ positions on the scalp surface taken from a real experiment.

At each of the 3,724 source positions (regular hexahedral mesh, 7 mm edge length, $K = 3724$) in the brain compartment (minimum distance to brain surface 5 mm), three orthogonal dipoles ($S = 3$) were placed and a leadfield matrix ($L \in \Re^{118 \times 11172}$) for all dipoles and directions was computed with the SimBio-toolbox (SimBio 2011). Single dipolar probe sources were placed at 6 representative locations in the brain compartment: at somatosensory, frontal, temporal left/right, thalamus and occipital positions. For the simulations artificial Gaussian noise was added as described in the last paragraph. For each of these positions, a radial and a tangential (with respect to the inner skull surface) orientation were tested, which are known to produce very different topographies and might lead to different results. We investigated, both, the localization error (i.e., distance between the

true source position and the maximum of the estimate) and the spatial dispersion as proposed by Molins et al. (2008): $SD = \sqrt{\sum_{k=1}^{K}\left(\hat{j}_k \cdot \mathrm{d}_{ki}\right)} / \sqrt{\sum_{k=1}^{K}\left(\hat{j}_k\right)^2}$, where d_{ki} is the distance between source position k and the peak position in the estimate i. In addition, we studied the performance of the algorithms on a pair of equally strong dipoles located in both auditory cortices and oriented in anterior–superior direction (45°). This scenario was chosen because it is known that sLORETA works for multiple sources only if they are sufficiently distinct and of similar strength (Wagner et al. 2004). Two cases were tested: (i) both dipoles had the same sinusoidal activation time course (10 Hz), and (ii) there was a 120° phase shift between the time courses. For these configurations, besides the spatial localization error (averaged over left and right sources), we evaluated the reconstruction accuracy of the time courses (as correlation between original and reconstructed time courses, approaches 1–3). For the original time courses an oscillation of 10 Hz, mimicking alpha activity was modeled using 61 time samples and a sampling frequency of 200 Hz, yielding an epoch length of 305 ms. The source space was divided into the two brain hemispheres. For each time sample the local maximum of the sLORETA estimate (for all approaches 1–4) was evaluated in each hemisphere and the spatial localization error (Euclidian distance) with respect to the simulated source location was computed. Both localization errors (one per hemisphere) were averaged for visualization purposes. The time courses of both sources were sampled in a way to ensure that none of them was completely deactivated at any time.

The band-stop filter was created using xfir (ANT 2003), with a sample frequency of 200 Hz and 31 filter coefficients in total (see Table 1), whereas F_n represents the center element of the band pass filter which appears as the biggest filter coefficient value. The filter has a lower/upper cutoff frequency of 7.0/14.0 Hz. All other filter coefficients are mirrored to the center coefficient, which means that $F_{n-1} = F_{n+1}$, …, $F_{n-15} = F_{n+15}$.

Evaluation Using Measurements from MEG Phantom Single Trial Data

Using physical phantoms, it is possible to combine the main advantage of computer simulation, namely the precise knowledge of the sources to be reconstructed, with the presence of realistic noise conditions and measurement errors. We used an MEG multiple dipole phantom (Elekta Neuromag Oy, Helsinki, Finland) in combination with a 306-channel whole-head MEG system (Elekta Neuromag Oy). In this system, at each of the 102 sensor positions, two planar gradiometers (orthogonally oriented) and one magnetometer are installed. The MEG phantom, which is constructed in a way that it behaves like a spherically symmetric volume conductor, provides 3 current dipoles (called CD 1, 2, 3), each at a different depth from the surface of the phantom (compared to human anatomy, it covers cortical to subcortical source locations). Forward modeling and source reconstruction were performed using the measurements from all 306 MEG channels. We activated one dipole at a time and used two different dipole strengths (25, 50 nAm) and an oscillation frequency of 20 Hz, corresponding to the beta frequency band. All measurements were sampled at 1,000 Hz and internally band-pass filtered between 0.1 and 330 Hz. For each configuration (dipole number, dipole strength), 200 trials were recorded. Each trial comprised 750 ms of oscillatory data (15 cycles) and 50 ms baseline. For inverse analysis, we used a FIR filter (101 coefficients) and a sampling frequency of 500 Hz (data were downsampled) in order to get a narrow frequency response centered around 20 Hz (lower/upper cutoff frequency 11.4/29.0 Hz). Based on the 53 digitization points covering the phantom's spherical surface a sphere could be fitted and its parameters (sphere radius, sphere center) were used for specifying the forward model. For the source space, we used a 5 mm regular hexahedral grid ($M = 11{,}698$ locations) covering the entire hemisphere of the fitted sphere. At each source location, two tangentially ($S = 2$) oriented dipole directions were computed using the center of the spherical head model. As the volume conductor is well accounted for by a sphere, we used the analytical equation according to Sarvas (1987). We modeled each sensor accurately using integration points (8 for each gradiometer and 16 for each magnetometer) which were provided by the manufacturer. The analysis of the MEG measurement noise revealed a super-Gaussian distribution which differs from the artificial Gaussian noise used for EEG simulations. For the MEG measurement (for each dipole strength and each dipole location) an RNA estimate was computed based on the residuals from alternating averaging: for dipole positions CD 1, 2, 3 with dipole moment 25 nAm: RNA = [0.34, 0.44, 0.63]; with dipole moment 50 nAm: RNA = [0.62, 0.85, 1.11].

Results

Localization of Simulated EEG Measurements from Single Oscillating Sources

In Fig. 1, the source localization results for the EEG simulations are summarized. For each noise level, the best case (smallest averaged localization error) with respect to the regularization parameter(s) of the respective method is plotted. The benefit of using a filter can be clearly seen. The spatio-temporally regularized sLORETA method

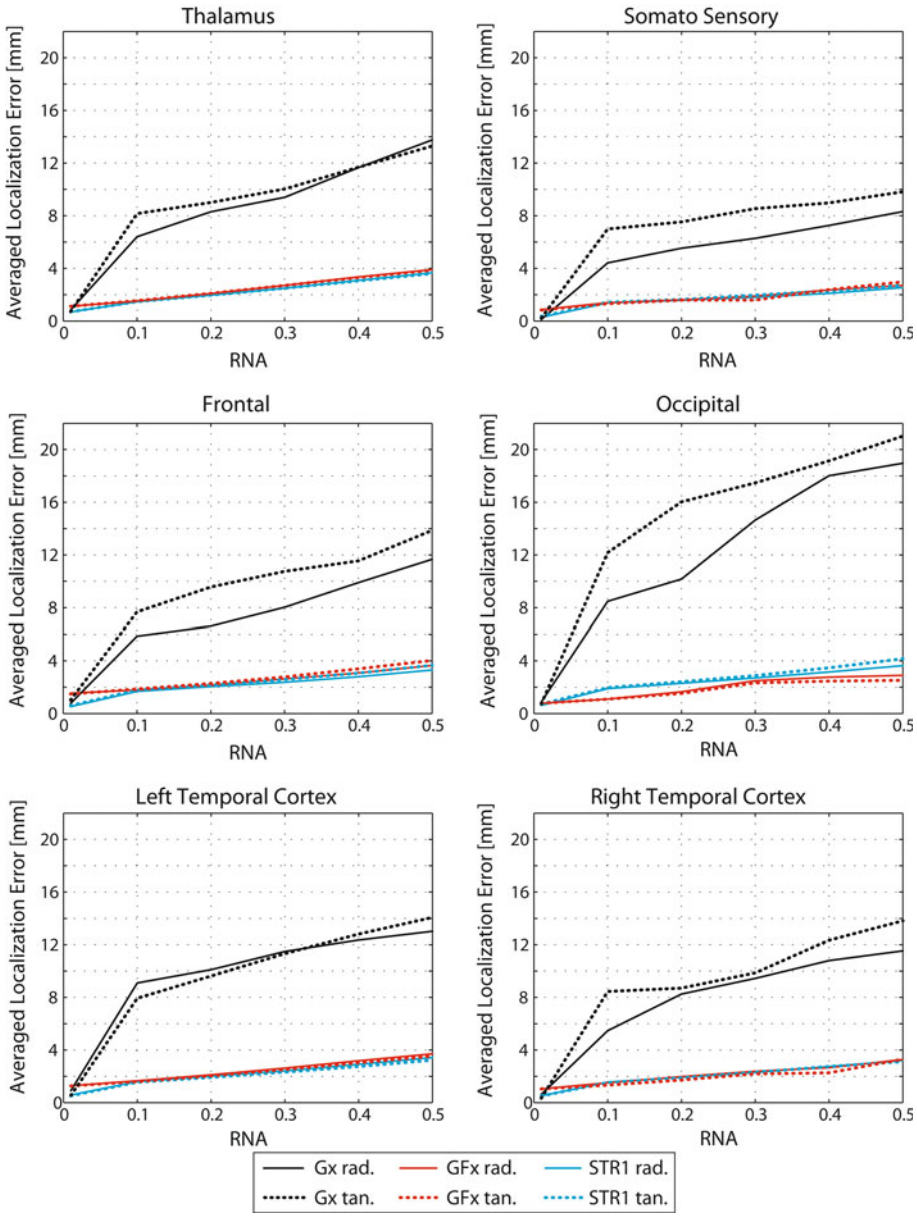


Fig. 1 EEG source localization versus noise contamination: smallest localization error over all regularization parameters per noise level, for sLORETA with spatio-temporal regularization (STR1), simple filtering (GFx), and classical spatial sLORETA (Gx). For each dipole location, a radial (rad.) and a tangential (tan.) dipole orientation were considered. The localization errors were averaged over 100 trials with independent noise realizations and 61 time steps per trial. All approaches, which are based on averaging of the reconstruction results (including STR2) perform with zero localization error for all considered noise levels for at least one regularization parameter (combination). *RNA* denotes the relative noise amplitude, see text for further explanation

(STR1) does not perform significantly better than the simple filtering approaches.

Figure 2 (upper row) shows, at the example of a thalamic source, how these results depend on the regularization parameter(s) chosen. For the spatial sLORETA methods, the localization error converges to small values for $\lambda > 10^4$ for the filtered versions (error less than 5 mm) and $\lambda > 10^6$ for the unfiltered version (error less than 20 mm). For STR1, the same magnitude of error as found for simple filtering is obtained for $\mu > 10^{10}$ and roughly $\lambda \leq \mu$. Clearly, STR1 does not outperform significantly any of the filtered sLORETA algorithms. Similar results were obtained for all other source locations and orientations tested.

As can be seen from Fig. 2 (middle row), the time-averaged current density computed with STR1 and GFx perform with zero localization error for wide ranges of the parameter space. The convergence limits for λ remain roughly the same for spatial sLORETA. STR2 (Fig. 2, lower row) also performs with zero localization error. It needs considerably lower μ compared to the pure temporally filtered spatial sLORETA to achieve a similar performance. This is expected, as temporal regularization affects both the inverse solution and post hoc weighting for STR2, while for STR1 the post hoc operator only contains the spatial regularization.

In Fig. 3 the regularization dependence of the spatial dispersion of the reconstructed current source density is shown in an analogous way. Interestingly, for the STR1 approach (with or without time averaging), the regularization scheme with the smallest spatial dispersion did not at all

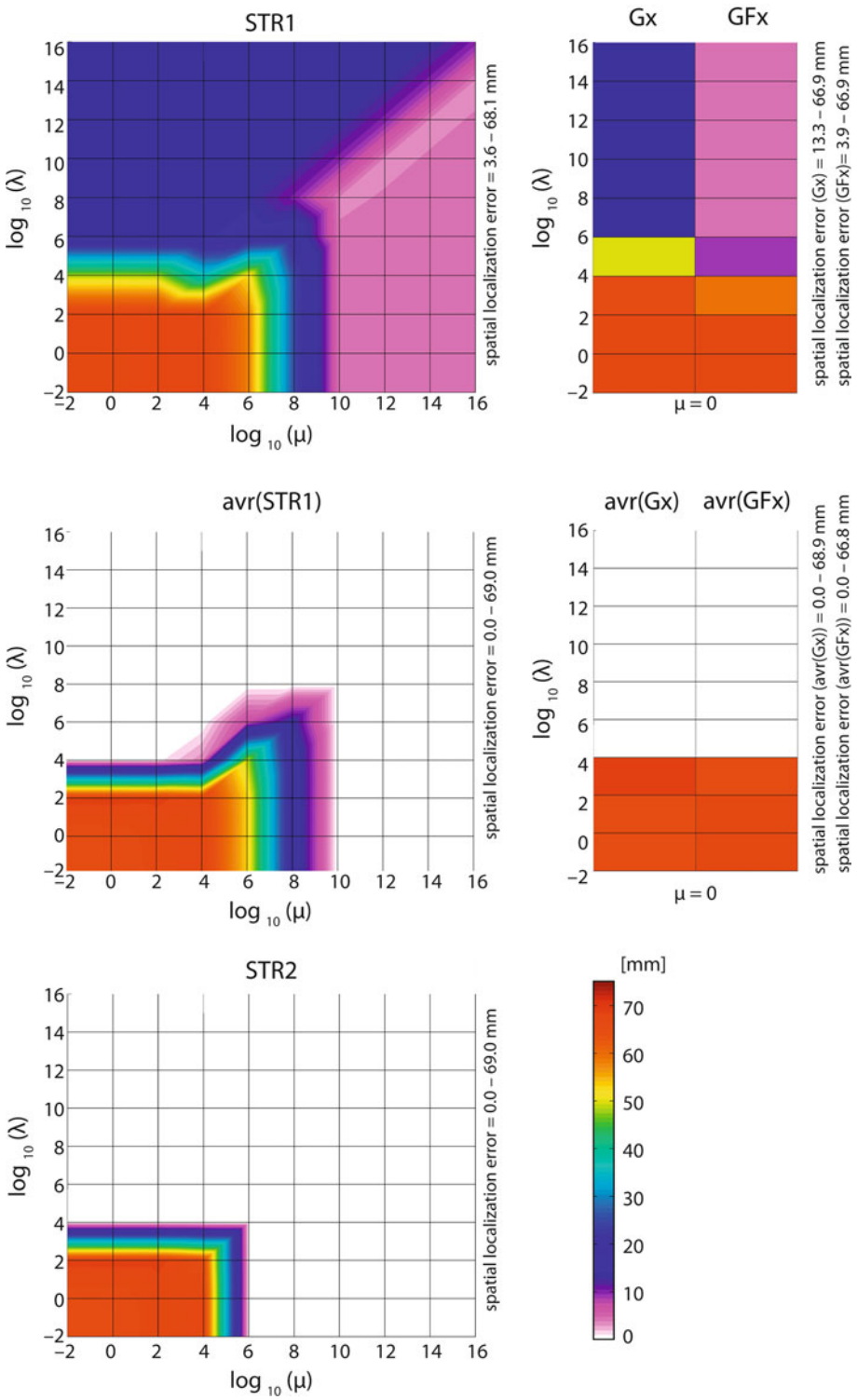


Fig. 2 Localization accuracy for a single thalamic source pointing in dorsal direction (relative noise amplitude $RNA = 0.5$). *Upper left*: Contour plot showing the dependence of the EEG source localization error (in mm) of the STR1 method on the regularization parameters. The minimum localization error is 3.6 mm. *Upper right*: Localization errors for the purely spatial methods. See legend of Fig. 1 for abbreviations. The plotted localization errors were averaged over 100 trials with independent noise realizations and 61 time steps per trial. *Middle row* same as *upper row* except that the current densities were here averaged over time, the localization errors determined on these averaged current densities and then averaged over 100 trials (with independent noise realizations). *Lower row*: Localization error for sLORETA with STR2 using, both, ad hoc and post hoc weighting

coincide with the one with the best localization accuracy. For example, the smallest spread for STR1 (without temporal averaging) amounts to SD $\approx$42 and occurs at regularizations ($\lambda \approx 10^{-2}\ldots10^{4}$, $\mu \approx 10^{6}$), where the localization error is about 50 mm (Fig. 2, upper left). Vice versa, where the localization error is fairly small (3.6 mm for $\mu > 10^{10}$), the spread is clearly larger (SD $\approx$60). In contrast, for the GFx method the smallest spatial dispersion is only slightly larger (SD $\approx$45), but almost coincides with the smallest localization error (3.9 mm for $\lambda \approx 10^{6}$). For larger regularization parameters, the performances of STR1 and GFx are nearly identical (SD $\approx$60 with localization error of about 4 mm, see

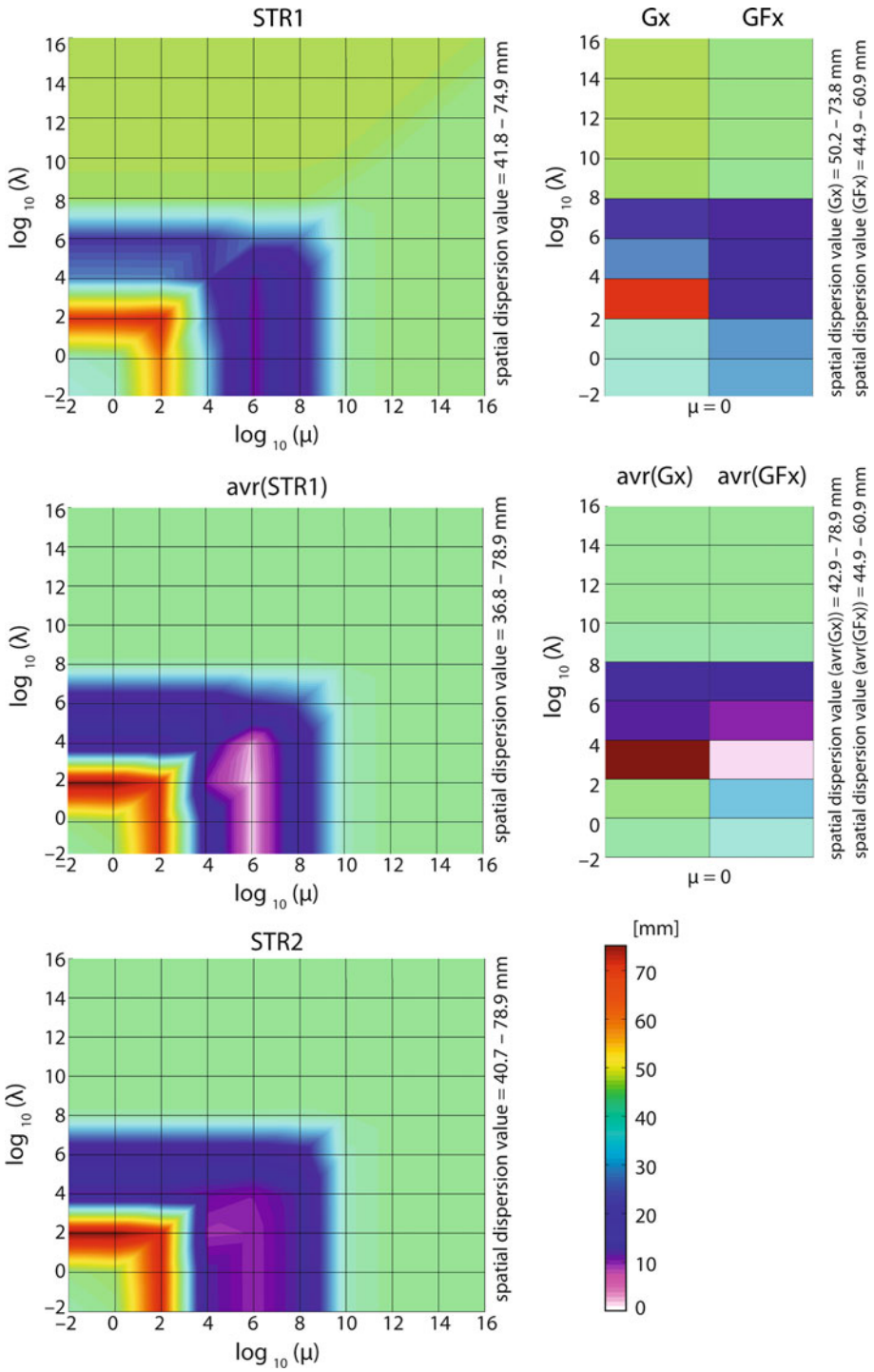


Fig. 3 Spatial dispersion for the same source as used in Fig. 2 (single thalamic source pointing in dorsal direction, RNA = 0.5). The arrangement is identical to Fig. 2

Figs. 2, 3). Comparing the time average variants of STR1 and GFx with STR2 leads to similar conclusions.

Localization of Simulated EEG Measurements from Two Simultaneously Oscillating Sources

In Fig. 4, the source localization results for the dipole pair are summarized. For each noise level, the best case (smallest averaged localization error) with respect to the regularization parameter(s) of the respective method is plotted. Again, it can be seen that for the simultaneously active sources (Fig. 4 right) filtering (approaches 2–4) enhances localization results. All averaging approaches (including STR2) come up with similar results. For the phase shifted time courses (Fig. 4 left) it is interesting to see that STR1 basically behaves as badly as using no filtering at all. It

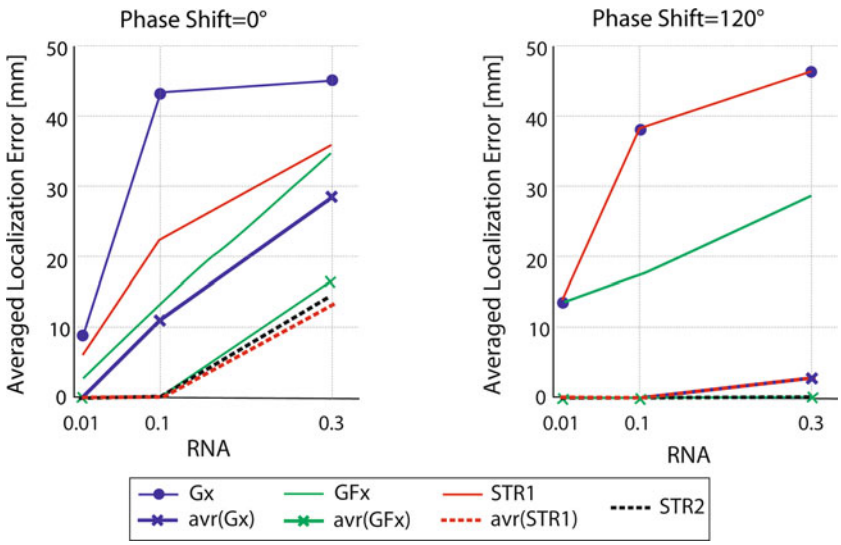


Fig. 4 Smallest localization errors for the tested regularization schemes for a pair of dipoles (averaged over hemispheres) in the auditory cortices, as function of noise level

shows that in such a case the spatio-temporal cross talk cannot be compensated for, simply by using spatial weights (STR1, R^{STR1}), since information about the total spatio-temporal spreading (R^{STR2}) would be needed. Moreover, even STR2 does not perform better than simple filtering.

Figure 5a (top row) and 5b present the reconstruction results for the case of non-synchronous time courses. For the spatial sLORETA methods (Gx and GFx), the localization error has an optimum for λ between 10^4 and 10^6 for the unfiltered and between 10^2 and 10^4 for the filtered version. However, the errors remain quite high (28 mm for the filtered and 46 mm for the unfiltered version). For STR1, the same magnitude of error (minimum 46 mm) is obtained for a strip roughly defined by $\lambda = 10^4 \ldots 10^6$ and $\mu = 10^{-2} \ldots 10^2$. Clearly, STR1 does not at all outperform any of the filtered sLORETA algorithms. However, while both STR1 and simple sLORETA (Gx) can, with proper regularization, reconstruct the time courses almost perfectly, simple filtering cannot (correlations <0.7). The reason for this is illustrated in Fig. 6—the filter applied produces edge effects, which have a substantial impact on the correlation for the low number of time steps in our simulations. For larger time series, these problems are likely to diminish.

If both time courses are synchronous (Fig. 7a, top row, 7b), the errors are generally higher. The time courses are of course much more accurate (with the same edge effect problem for GFx) because mixing between two different temporal patterns cannot occur. Inaccuracies are purely due to underregularization in the presence of noise.

The lower two rows of Figs. 5a and 7a show the results for the STR2 method in comparison to averaged results of the other algorithms. The dependence on the regularization parameters is quite similar to the ones described above.

Evaluation Using Measurements from MEG Phantom Data

The localization of dipoles in the physical phantom from MEG data yielded the following average localization errors for STR1: At source strength 25 nAm, the dipoles CD 1, 2, 3 were localized with minimum errors 0.9, 4.1 and 7.9 mm, respectively, while at source strength 50 nAm, the localization errors were 0.2, 1.6 and 4.1 mm. These localizations were performed on single trial data (see treatment of reconstruction of induced activity in the "Discussion" section), and the localized positions were averaged. Due to the high number of filter coefficients and data samples a fairly good localization performance for all methods could be achieved. There was no difference in localization error between filtered spatial sLORETA (GFx) and spatio-temporally regularized sLORETA (STR1).

The dependence of the localization error is plotted in Fig. 8. The pattern is qualitatively similar to the results from the simulated EEG data (Fig. 2). In any case, the best results obtained by STR1 (for higher μ) can be easily achieved by spatial sLORETA with simple filtering for a wide range of spatial regularization parameters. The fact remains that GFx performs equal to STR1 for a wide range of parameters. For STR2, the minimal localization error is in the same range compared to GFx. It also seems that smaller temporal regularization is needed—similar to the EEG simulation study. Nonetheless, there are only very small differences between the patterns of avr(STR1) and STR2 (with similar localization performance). We find no advantage in using current density averaging (approaches 1 and 2) compared to the performance of STR2.

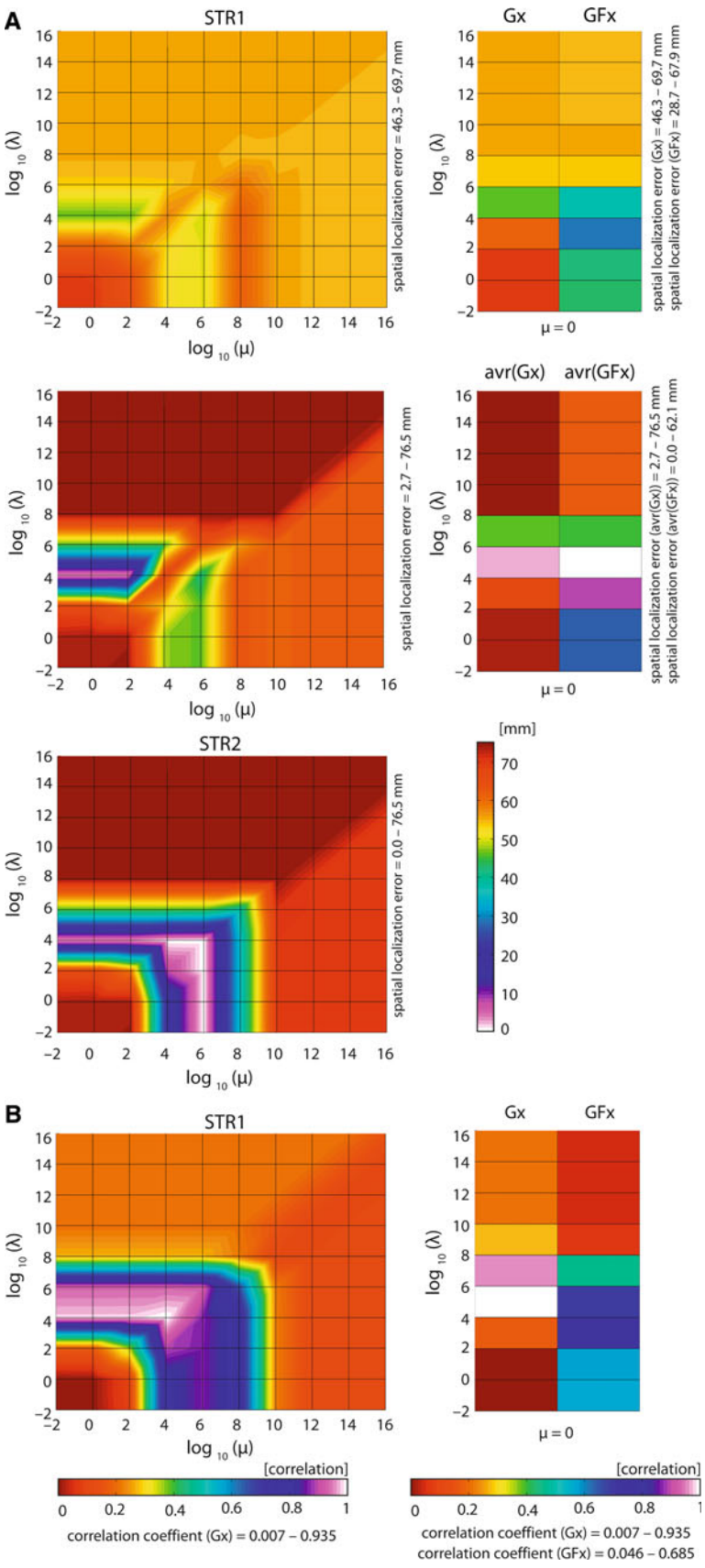


Fig. 5 Localization accuracy and quality of time course reconstruction (correlation) for a pair of simultaneously active sources in the auditory cortices with oscillatory activation time courses of 10 Hz, shifted by 120° towards each other (relative noise amplitude $RNA = 0.3$). **a** *Top left:* Contour plot showing the dependence of the EEG source localization error (in mm; *left*) of the STR1 method plotted versus the regularization parameters. The values are averaged over the two dipoles. *Top right:* Localization errors for the purely spatial methods. See legend of Fig. 1 for abbreviations. The plotted localization errors were averaged over 100 trials with independent noise realizations and 61 time steps per trial. *Middle*: Localization errors for STR1 as well as purely spatial methods, with current densities averaged over time. The localization errors were determined on these averaged current densities and then averaged over 100 trials (with independent noise realizations). *Bottom*: Localization errors for sLORETA with STR2 using, both, ad hoc and post hoc weighting. **b** Correlation between original and reconstructed time courses for STR1 (*left*) and the purely spatial methods (*right*)

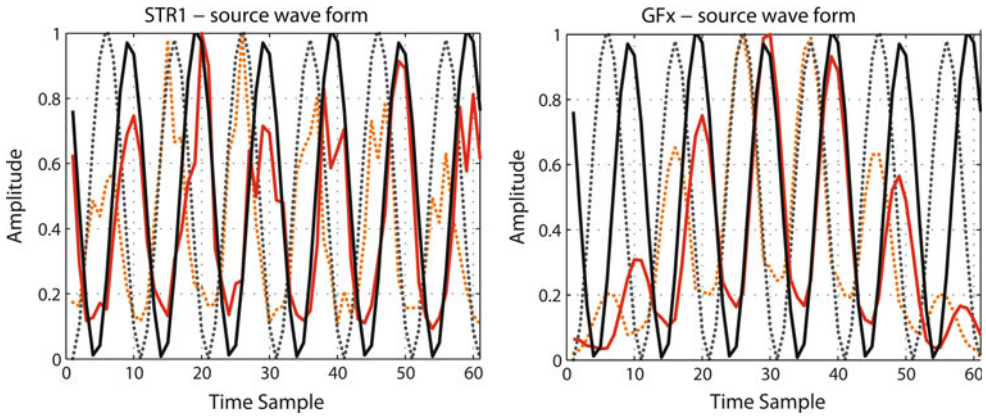


Fig. 6 Reconstructed time courses for reconstruction (relative noise amplitude $RNA = 0.3$) of dipole pair with 120° phase shift using STR1 (*left*) and GFx (*right*). *Red curves* represent the reconstructed, *black curves* the original time courses (amplitudes are squared). For both algorithms, regularization parameters were chosen for which both localization error and time course reconstruction accuracy are close to optimum (see Fig. 4). The figures demonstrate that the lower time course reconstruction accuracy for GFx seen in Fig. 4 is due to edge effects of the filter, which have less effect in the STR1 scheme

Discussion

STR of distributed source reconstruction algorithms has been proposed in the literature (Brooks et al. 1999; Darvas et al. 2001; Schmitt and Louis 2002; Schmitt et al. 2001, 2002) and discussed as one possibility to stabilize inverse solutions and fill the information gap caused by the non-uniqueness of the inverse problem. Effectively, these methods enrich the linear inverse operator by providing a temporal filter. This raises the question of whether simple ad hoc (to the data) or post hoc (to the reconstructed source time courses) temporal filters applied in combination with a purely spatial reconstruction algorithm could be used to achieve the same results. This is particularly important because the STR methods are computationally very expensive. In the initial studies by Brooks et al. (1999), no systematic investigation of this issue was offered. Indeed, Schmitt et al. (2001) did compare their results with ad hoc application of one particular temporal low-pass filter and demonstrated the superiority of STR in this particular case. However, there was no apparent correspondence between the definitions of their temporal regularization (i.e., the filter within the inverse operator) and this ad hoc filter. In other words, it cannot be ruled out that other definitions of the ad hoc filter would have performed better or at least equally well compared to the STR. Moreover, this comparison additionally depends on the proper choice of the regularization parameters. Finally, most of the previous results for EEG/MEG have been achieved with strongly simplified configurations of head models and sensor positions, with one exception in (Darvas et al. 2001), where a pair of dipoles was reconstructed from simulated data in a realistically shaped volume conductor model.

Here, we used realistic EEG configurations with a finite element volume conductor model to systematically evaluate the performance of STR-based source reconstruction methods. We used a number of practically relevant source configurations, including single and double dipole schemes. Moreover, we systematically varied both the spatial and temporal regularization parameters. Most importantly, we compared the STR (both variants STR1, STR2) method to simple ad hoc (and post hoc) filtering using identical filter definitions, thus allowing a precise assessment of the potential benefit of the STR method over the computationally much cheaper filtering methods. Finally, we applied our methods to MEG data from a phantom measurement in order to confirm the results under even more realistic conditions.

In contrast to previous studies, we used the sLORETA method, which is known to be able to reconstruct single dipoles without any depth bias (Pascual-Marqui 2002). Moreover, the method allowed us to assess the effect of temporal regularization incorporated in the post hoc weighting operator as well. Another difference to previous work (Brooks et al. 1999; Darvas et al. 2001; Schmitt and Loui 2002; Schmitt et al. 2001, 2002) is that we did not use a high-pass, but a band-stop filter in the penalty term because we consider single-trial reconstruction from induced oscillatory activity as one of the major potential applications of spatio-temporal regularization.

We found that STR based on sLORETA yielded improved reconstructions of single sources from realistic data, in comparison to sLORETA without filtering or spatial regularization, thus confirming the findings of previous researchers (Brooks et al. 1999; Darvas et al. 2001; Schmitt and Louis 2002; Schmitt et al. 2001, 2002). We also

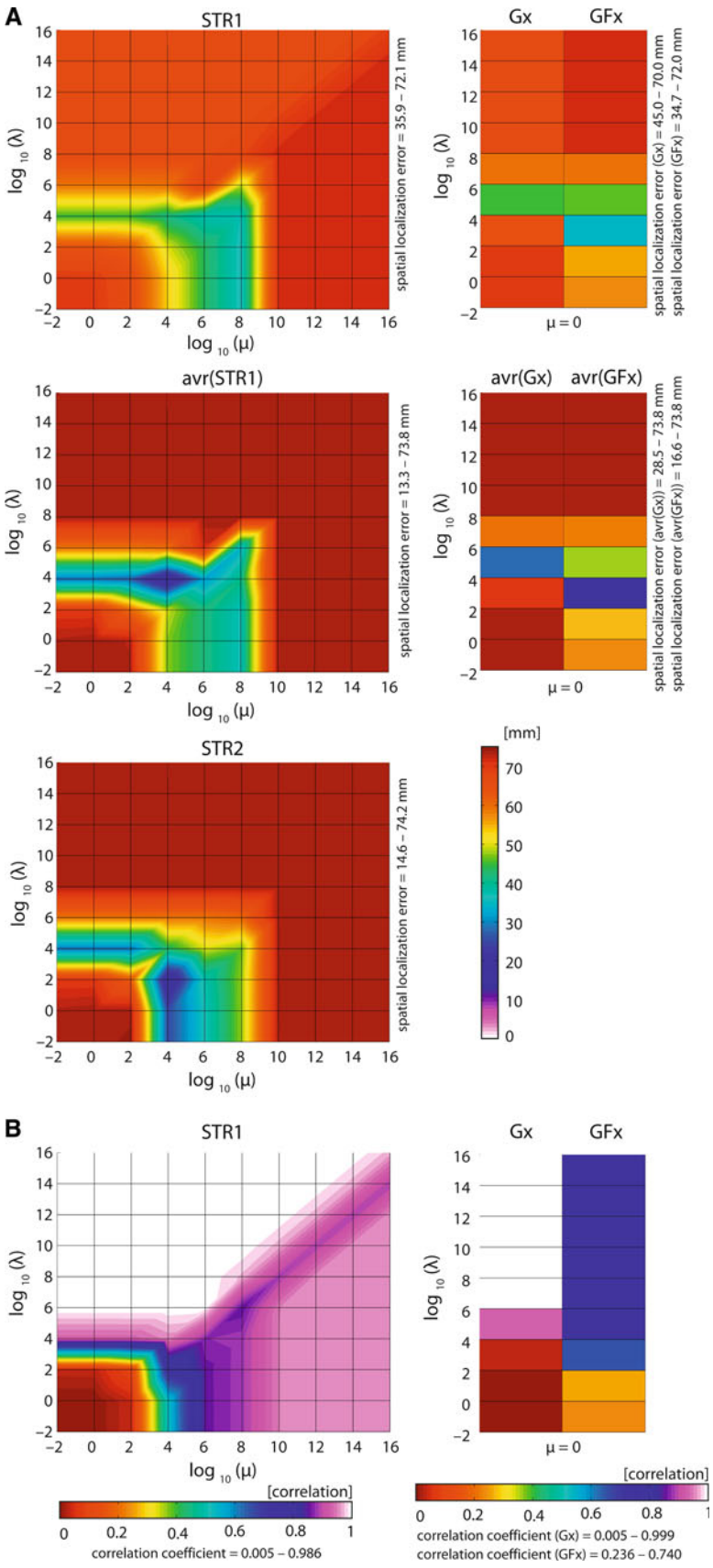


Fig. 7 Localization accuracy and quality of time course reconstruction (correlation) for a pair of simultaneously active sources in the auditory cortices with oscillatory activation time courses of 10 Hz, without mutual phase shift (relative noise amplitude $RNA = 0.3$). **a** Localization accuracies—for further details, see legend of Fig. 4a. **b** Correlation between original and reconstructed time courses for STR1 (*left*) and the purely spatial methods (*right*) (Color figure online)

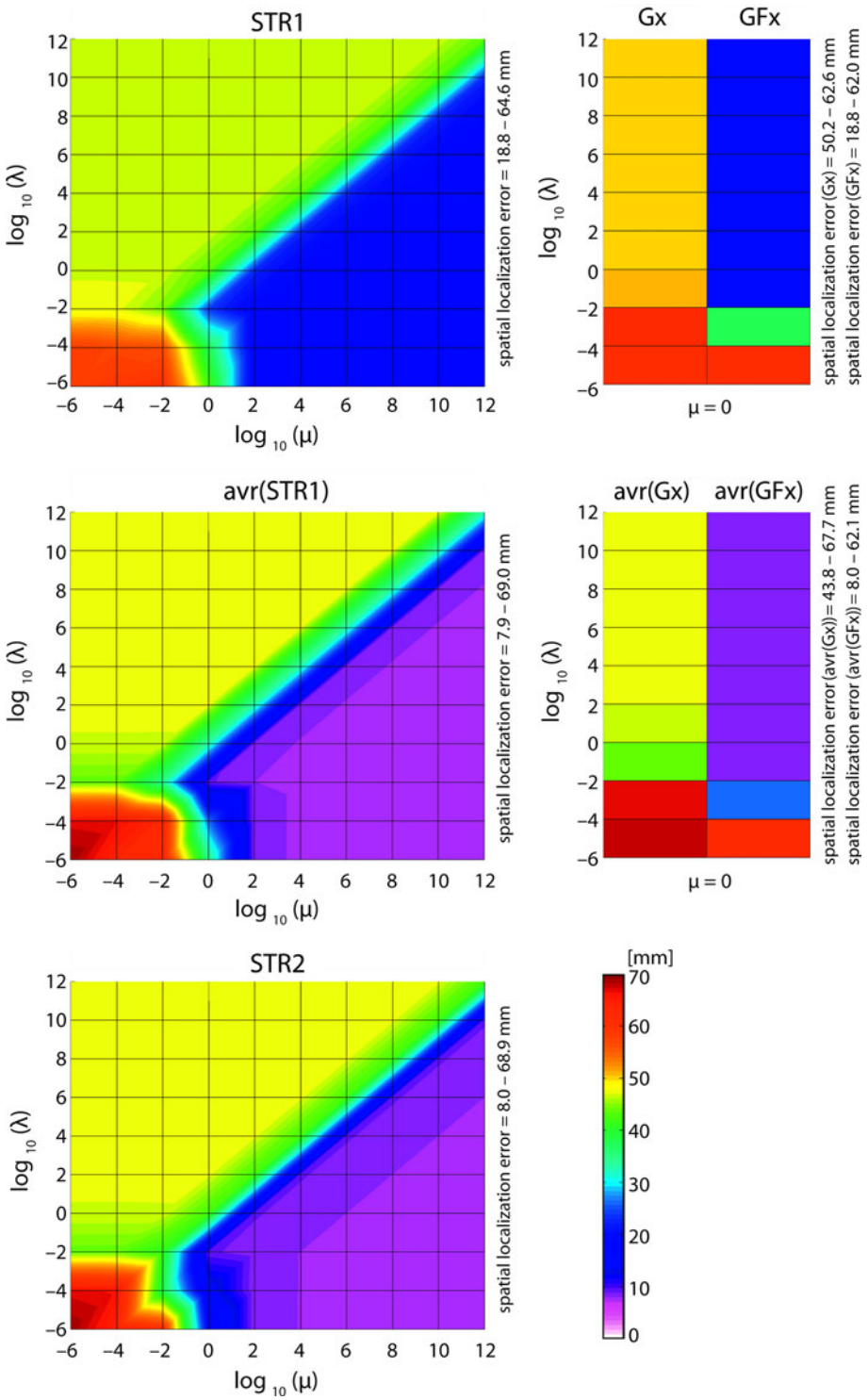


Fig. 8 Localization accuracy for reconstruction from MEG phantom data. *Upper left*: Contour plot showing the dependence of the MEG source localization error (in mm) of the STR1 method on the regularization parameters for the deepest dipole (CD3) at strength of 25 nAm. The minimum localization error is about 7.9 mm. *Upper right*: Localization errors for the purely spatial methods. See legend of Fig. 1 for abbreviations. The plotted localization errors were averaged over 200 trials with independent noise realizations and 151 time steps per trial. *Lower row* same as *upper row*; however, here the current densities were averaged over time, the localization errors determined on these averaged current densities and then averaged over 200 trials (with independent noise realizations)

demonstrated that the localization success hinges on properly chosen spatial and temporal regularization parameters. In practice, this could mean that a two-dimensional L-surface would have to be computed (Brooks et al. 1999), but our simulations suggest, that a rather high μ, and a somewhat smaller λ is a reasonable choice to achieve optimal results. Note that both parameters μ and λ also depend on spatial and temporal properties (spatiotemporal resolution, condition number, etc.), which are application specific.

However, we also found that using simple filtering (ad hoc or post hoc—both yield exactly the same results) performs equally well. The MEG phantom results

essentially confirm these findings. When looking at the spatial dispersion of the reconstructed solutions, they were found very similar between STR and simple filtering, in particular for those regularization parameter combinations, for which the localization accuracy reasonable was reasonably small.

Schmitt et al. (2001) compared a double dipole reconstruction with STR to one without temporal regularization, but with ad hoc filtering. They claim superiority for the STR, but, unfortunately, only report qualitative results. However, as they did not ensure correspondence between the ad hoc filter and the STR temporal weighting and they did not systematically vary regularization parameters, this claim is not in contradiction to our findings.

Another important criterion for the reconstruction performance is the accuracy of the time course reconstruction, in particular, when more than one source is active and mixing problems can occur. Schmidt and colleagues (Schmitt et al. 2001) have shown that, at least in their simplified configuration, STR reconstructs the original time series with high accuracy. Here, we confirm this finding. When using simple filtering, the time course reconstruction accuracy (as measured by correlation) is not as good. The reason for this appears to lie in the edge effects of the band-pass filter applied, which have a high impact on the overall correlation due to the short time series. Using longer time series will likely reduce this problem. Hence, if, for some reason, very short time series have to be used that cannot be filtered before epoching, the STR method may offer an advantage, in particular, if the temporal information is important (e.g., in connectivity analysis). It is, however, conceivable that alternative filter designs (less steep, different windowing) might reduce or eliminate this advantage of STR.

Although our results appear quite clear-cut and univocally suggest that, at least for the sLORETA method, STR does not offer a significant advantage over simple filtering, the limits of the generalizability should be discussed. First, we apply the STR concept to one particular linear inverse method, that is, sLORETA. Therefore, strictly speaking, our results are only valid for this method. However, it is likely that the results can be generalized to other linear methods, including (weighted) minimum norm and LORETA. We tested the main findings of this paper using the minimum norm method instead of sLORETA and found no major difference to the ones reported here. On the other hand, it is not as obvious, and remains to be investigated, how non-linear distributed source algorithms [e.g., L_1 norm-based, see, e.g., (Fuchs et al. 1999)] or hierarchical Bayesian modeling (see, e.g., Lucka et al. 2011) perform with spatio-temporal regularization.

Second, we used a band-stop filter in order to focus on narrow frequency bands. Although it seems plausible that the equivalence between simple filtering and STR using the inverse of the same filter as a penalty is universal, this cannot be definitely proven and remains to be investigated.

Conclusion

In this work, we investigated the concept of STR in linear estimation-based distributed source reconstruction. We used sLORETA and focused on oscillatory activity in narrow frequency bands. In particular, we sought to answer the question whether, in the context of the sLORETA framework, the computationally very costly STR outperforms the classical purely spatial method applied to temporally filtered data. We systematically varied source configurations and both spatial and temporal regularization parameters. Within the scenarios tested in this study, STR did not offer any significant advantage.

Acknowledgments This work was kindly supported by the Deutsche Forschungsgemeinschaft (contract Grant Numbers KN 588/2-1,4-1 and WO 1425/1-1,3-1). None of the authors has any conflict of interest, financial or otherwise, related to the submitted.

References

Ahlfors SP, Simpson GV (2004) Geometrical interpretation of fMRI-guided MEG/EEG inverse estimates. Neuroimage 22(1):323–332

Akhtari M, Bryant HC, Marnelak AN, Flynn ER, Heller L, Shih JJ, Mandelkern M, Matlachov A, Ranken DM, Best ED, DiMauro MA, Lee RR, Sutherling WW (2002) Conductivities of three-layer live human skull. Brain Topogr 14(3):151–167

ANT S.D.G. (2003) filter design tool: xfir. ANT B.V. Advanced Neuro Technology, Enschede, Netherlands

Awada KA, Jackson DR, Williams JT, Wilton DR, Baumann SB, Papanicolaou AC (1997) Computational aspects of finite element modeling in EEG source localization. IEEE Trans Biomed Eng 44(8):736–752

Baillet S, Garnero L (1997) A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans Biomed Eng 44(5):12

Brooks DH, Ahmad GF, MacLeod RS, Maratos GM (1999) Inverse electrocardiography by simultaneous imposition of multiple constraints. IEEE Trans Biomed Eng 46(1):3–18

Buchner H, Knoll G, Fuchs M, Rienacker A, Beckmann R, Wagner M, Silny J, Pesch J (1997) Inverse localization of electric dipole current sources in finite element models of the human head. Electroencephalogr Clin Neurophysiol 102(4):267–278

Dale M, Serano MI (1993) Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci 5(1):15

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26(1):55–67

Dannhauer M, Lanfer B, Wolters CH, Knösche TR (2011) Modeling of the human skull in EEG source analysis. Hum Brain Mapp 32(9):1383–1399

Darvas F, Schmitt U, Louis AK, Fuchs M, Knoll G, Buchner H (2001) Spatio-temporal current density reconstruction (srCDR) from EEG/MEG-data. Brain Topogr 10(3):13

de Munck J, Peters M (1993) A fast method to compute the potential in the multi sphere model. IEEE Trans Biomed Eng 40(11): 1166–1174

de Munck J, van Dijk BW, Spekreijse H (1988) Mathematical dipoles are adequate to describe realistic generators of human brain activity. IEEE Trans Biomed Eng 35(11):960–966

de Peralta Grave, Menendez R, Murray MM, Andino SLG (2004) Improving the performance of linear inverse solutions by inverting the resolution matrix. IEEE Trans Biomed Eng 51(9): 1680–1683

Fuchs M, Wagner M, Kohler T, Wischmann HA (1999) Linear and nonlinear current density reconstructions. J Clin Neurophysiol 16(3):267–295

Fuchs M, Wagner M, Kastner J (2007) Development of volume conductor and source models to localize epileptic foci. J Clin Neurophysiol 24(2):101–119

Güllmar D, Haueisen J, Eiselt M, Giessler F, Flemming L, Anwander A, Knösche TR, Wolters CH, Dümpelmann M, Tuch DS, Reichenbach JR (2006) Influence of anisotropic conductivity on EEG source reconstruction: investigations in a rabbit model. IEEE Trans Biomed Eng 53:1841–1850

Güllmar D, Haueisen J, Reichenbach JR (2010) Influence of anisotropic electrical conductivity in white matter tissue on the EEG/MEG forward and inverse solution. A high-resolution whole head simulation study. Neuroimage 51:145–163

Hämäläinen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. Med Biol Eng Comput 32(1):8

Hansen PC, O'Leary DP (1993) The use of the L-curve in the regularization of discrete ill-posed problems. SIAM 14(6):17

Huiskamp G, Maintz J, Wieneke G, Viergever M, van Huffelen AC (1997) The influence of the use of realistic head geometry in the dipole localization of interictal spike activity in MTLE patients. Biomed Tech 42:4

Huiskamp G, Vroeijenstijn M, van Dijk R, Wieneke G, van Huffelen AC (1999) The need for correct realistic geometry in the inverse EEG problem. IEEE Trans Biomed Eng 46(11):1281–1287

Jeffs B, Leahy R, Singh M (1987) An evaluation of methods for neuromagnetic image reconstruction. IEEE Trans Biomed Eng 34(11):713

Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Rev 29:169–195

Kybic J, Clerc M, Abboud T, Faugeras O, Keriven R, Papadopoulo T (2005) A common formalism for the integral formulations of the forward EEG problem. IEEE Trans Med Imaging 24(1):12–28

Lew S, Wolters CH, Anwander A, Makeig S, MacLeod RS (2009) Improved EEG source analysis using low-resolution conductivity estimation in a four-compartment finite element head model. Hum Brain Mapp 30(9):2862–2878

Liu AK, Belliveau JW, Dale AM (1998) Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. Neuroimage 23(14):582

Lucka F, Pursiainen S, Burger M, Wolters CH (2011) Hierarchical Bayesian models for EEG inversion: depth localization and source separation for focal sources in realistic FE head models. In: Biomedical engineering, vol 56. De Gruyter, Berlin, pp 939–4990

Marin G, Guerin C, Baillet S, Garnero L, Meunier G (1998) Influence of skull anisotropy for the forward and inverse problem in EEG: simulation studies using FEM on realistic head models. Hum Brain Mapp 6(4):250–269

Molins A, Stufflebeam SN, Brown EM, Hämäläinen MS (2008) Quantification of the benefit from integrating MEG and EEG data in minimum l(2)-norm estimation. NeuroImage 42(3): 1069–1077

Pascual-Marqui RD (2002) Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find Exp Clin Pharmacol 24(1):8

Pascual-Marqui RD (2007) Discrete, 3D distributed linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization. Technical report: arXiv:0710.3341v2

Pascual-Marqui RD, Michel CM, Lehmann D (1994) Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int J Psychophysiol 18(1):49–65

Plonsey R, Heppner DB (1967) Considerations of quasi-stationarity in electrophysiological systems. Bull Math Biophys 29(1):8

Ramon C, Schimpf PH, Haueisen J (2006) Influence of head models on EEG simulations and inverse source localizations. Biomed Eng Online 5:10

Sarvas J (1987) Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys Med Biol 32(1):11–22

Scherg M (1990) Fundamentals of dipole source potential analysis. In: auditory evoked magnetic fields and electric potentials. Adv Audiol 6(1):30

Scherg M, Berg P (1991) Use of prior knowledge in brain electromagnetic source analysis. Brain Topogr 4(2):8

Scherg M, von Cramon D (1986) Evoked dipole source potentials of the human auditory cortex. Electroencephalogr Clin Neurophysiol 65(5):16

Schmitt U, Louis AK (2002) Efficient algorithms for the regularization of dynamic inverse problems: I. Theory. Inverse Probl 18(1):14

Schmitt U, Louis AK, Darvas F, Buchner H, Fuchs M (2001) Numerical aspects of spatio-temporal current density reconstruction from EEG-/MEG-data. IEEE Trans Med Imaging 20(4):11

Schmitt U, Louis AK, Wolters CH, Vauhkonen M (2002) Efficient algorithms for the regularization of dynamic inverse problems: II. Applications. Inverse Probl 18(1):18

Schmitt U, Wolters CH, Anwander A, Knösche T (2004) STR: a new spatio-temporal approach for accurate and efficient current density reconstruction. In: Halgren E, Ahlfors S, Hämäläinen M, Cohen D (eds) BIOMAG 2004, proceedings of the 14th international conference on biomagnetism. Biomag, Boston, pp 591–592

Vallaghe S, Papadopoulo T (2010) A trilinear immersed finite element method for solving the electroencephalography forward problem. Siam J Sci Comput 32(4):2379–2394. doi:10.1137/09075038X

van den Broek SP, Reiders F, Donderwinkel M, Peters MJ (1998) Volume conduction effects in EEG and MEG. Electroencephalogr Clin Neurophysiol 106(6):13

Wagner M, Fuchs M, Kastner J (2004) Evaluation of sLORETA in the presence of noise and multiple sources. Brain Topogr 16(4): 277–280

Wolters CH, Anwander A, Berti G, Hartmann U (2007a) Geometry-adapted hexahedral meshes improve accuracy of finite-element-method-based EEG source analysis. IEEE Trans Biomed Eng 54(8):1446–1453

Wolters CH, Köstler H, Möller C, Härdtlein J, Grasedyck L, Hackbusch W (2007b) Numerical mathematics of the subtraction method for the modeling of a current dipole in EEG source reconstruction using finite element head models. Siam J Sci Comput 30(1):24–45.

SimBio (2011) SimBio: a generic environment for bio-numerical simulations. https://www.mrt.uni-jena.de/simbio. Accessed 2 July 2012.

Zanow F (1997) Realistically shaped models of the head and their application to EEG and MEG. PhD thesis. University of Twente, Enschede

Zhang YH, Ghodrati A, Brooks DH (2005) An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework. Inverse Probl 21(1):357–382