Combining sparsity and rotational invariance in EEG/MEG source reconstruction

Stefan Haufe[a,b,*], Vadim V. Nikulin[c,d], Andreas Ziehe[a,b], Klaus-Robert Müller[a,b,d] and Guido Nolte[b]

[a]*Machine Learning Group, Department of Computer Science, TU Berlin, Franklinstr. 28/29, D-10587 Berlin, Germany*

[b]*Intelligent Data Analysis Group, Fraunhofer Institute FIRST, Kekuléstr. 7, D-12489 Berlin, Germany*

[c]*Neurophysics Group, Department of Neurology, Campus Benjamin Franklin, Charité University Medicine Berlin, D-12203 Berlin, Germany*

[d]*Bernstein Center for Computational Neuroscience, Berlin, Germany*

Abstract

We introduce Focal Vector Field Reconstruction (FVR), a novel technique for the inverse imaging of vector fields. The method was designed to simultaneously achieve two goals: a) invariance with respect to the orientation of the coordinate system, and b) a preference for sparsity of the solutions and their spatial derivatives. This was achieved by defining the regulating penalty function, which renders the solutions unique, as a global ℓ_1-norm of local ℓ_2-norms. We show that the method can be successfully used for solving the EEG inverse problem. In the joint localization of 2-3 simulated dipoles, FVR always reliably recovers the true sources. The competing methods have limitations in distinguishing close sources because their estimates are either too smooth (LORETA, Minimum ℓ_2-norm) or too scattered (Minimum ℓ_1-

Preprint submitted to NeuroImage *13 June 2008*

norm). In both noiseless and noisy simulations, FVR has the smallest localization error according to the Earth Mover's Distance (EMD), which is introduced here as a meaningful measure to compare arbitrary source distributions. We also apply the method to the simultaneous localization of left and right somatosensory N20 generators from real EEG recordings. Compared to its peers FVR was the only method that delivered correct location of the source in the somatosensory area of each hemisphere in accordance with neurophysiological prior knowledge.

Key words: EEG/MEG, Inverse Problem, Source Localization, Second-Order Cone Programming, ℓ_1-norm Regularization, Sparsity, Vector Fields, Rotational Invariance

1 Introduction

Precise localization of neuronal activity is an important aspect for a better understanding of brain functioning. Several functional imaging methods have been developed for investigating this issue, including Single Photon Emission Computed Tomography (SPECT), Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI). These techniques provide high spatial resolution of brain activity using metabolic indicators such as blood oxygenation level (fMRI) or the concentration of radioactively marked substances (SPECT/PET) in the tissue. Due to the slow response of the metabolism, however, these measures cannot be used to assess rapidly varying neuronal activity in a range of few milliseconds. Apart from measuring direct neuronal activity, Electroencephalography (EEG) and Magnetoencephalogra-

* Corresponding author.
Email address: haufe@cs.tu-berlin.de (Stefan Haufe).

phy (MEG) allow very accurate reconstruction of the time course of neuronal signals with a microsecond precision (Nunez and Srinivasan, 2005). Importantly both techniques are noninvasive, and do not interfere with neuronal activities. However, the signal arriving at the sensors contains contributions from all areas of the brain, as well as external noise. The forward mapping from cerebral sources to sensors is well-defined and can be described mathematically with the help of a suitable model of the head. Inferring the sources that lead to a certain measurement, on the other hand, is impossible, as infinitely many source configurations will fulfill the forward equation. In other words, the inverse problem is ill-posed.

One strategy to still obtain a unique solution to the inverse problem is to regularize, i.e. to restrict the search space to a sufficiently simple class of sources. A common approach is to assume that the measured scalp pattern has been generated by dipolar (point-like) sources (Scherg and von Cramon, 1986; Mauguière et al., 1997; Komssi et al., 2004; Huttunen et al., 2006). Respective approaches model a small number of dipoles, where the optimal number has to be known in advance. The inversion is carried out by solving an overdetermined nonlinear system in a least-squares sense. Unfortunately, the cost function of dipole fits is highly nonconvex and the obtained solution depends heavily on the initialization. Additionally, dipolar sources can be a poor approximation if, e.g., the true sources are spatially extended and oriented normal to a folded cortical surface.

An approach related to dipole fitting is dipole imaging. Imaging methods model a large but fixed number of dipoles. These are arranged in a regular grid covering the whole brain (or optionally just the cortical areas). Inferring the dipole current vectors requires solving a heavily underdetermined system, in which the fulfillment of the forward equation constitutes only a constraint. Several methods tackle the underdetermined nature of the problem by incorporating additional information. Very often temporal structure in the signal is used, as for example in beamforming (Veen and Buckley, 1988), subspace methods like MUSIC and (RAP)MUSIC (Schmidt, 1986; Mosher and Leahy, 1999) and the methods proposed in (Baillet and Garnero, 1997; Huang et al., 2006; Malioutov et al., 2005; Cotter et al., 2005; Polonsky and Zibulevsky, 2004). The approach of Dale and Sereno (1993) imposes anatomical constraints obtained from MRI on the sources. A general overview on inverse methods for EEG and MEG is given by Baillet et al. (2001).

In this paper we focus on the situation in which only the scalp pattern at one time point is available. In this case, imaging methods have to define an additional quality criterion in order to obtain a unique solution. Ideally, this regularizing criterion should encode prior knowledge on how a "good" solution looks like. We here assume that a) brain sources are focal and we request b) invariance with respect to rotations of the coordinate system. Standard Minimum ℓ_p-norm solutions, weighted or not, are either rotationally invariant but highly non-focal (p=2) or focal but violating rotational invariance (p=1). We will propose an alternative consisting of a global ℓ_1-norm of local ℓ_2-norms which fulfills both goals simultaneously. Local ℓ_2-norms can be calculated both of the sources (as in "standard" Minimum ℓ_2-norm solutions and of their second order spatial derivatives (as in LORETA). We here suggest to use a specific combination of the two, relaxing the strict focality requirement in favor of a more robust "simplicity" requirement.

This paper is organized as follows. In section 2 we will first give an overview of existing methods and then we will present the mathematical details of our method. In section 3 we show illustrative examples in a simple constructed one-dimensional scenario, followed by detailed simulated examples of EEG inverse calculations and a case study using real EEG data from electric stimulation of left and right median nerve. We finally discuss the results and give a conclusion in sections 4 and 5, respectively.

2 Materials and methods

2.1 Inverse imaging

Let $\mathbf{x} \in \mathbb{R}^M$ denote a scalp pattern measured at M EEG or MEG sensors. The current density in the brain is modeled by N dipolar sources $\mathbf{d}_i = (\mathbf{r}_i^T, \mathbf{s}_i^T)^T,\ i \in \{1, \ldots, N\}$. The locations $\mathbf{r}_i \in \mathbb{R}^3$ are kept fixed, so that the quantities to be inferred are the dipole moment vectors $\mathbf{s}_i = (s_{i,x}, s_{i,y}, s_{i,z})^T,\ i \in \{1, \ldots, N\}$. Let $\mathbf{s} \in \mathbb{R}^{3N} = (\mathbf{s}_1^T, \ldots, \mathbf{s}_N^T)^T$ be the vector containing the stacked moments. As the relationships between source currents and EEG/MEG measurements are linear, the forward equation just reads $\mathbf{x} = L\mathbf{s}$ in both cases. The matrix L is called lead field matrix. It comprises information about geometric and conductive properties of the tissue. We will assume L to have maximal rank, that is for EEG the reference electrode is not included in $\mathbf{x}$ and L. If we require that the estimated solution explains the data exactly, the inverse solution for an imaging methods can be cast as

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} f(\mathbf{s}) \quad \text{s.t.} \quad \mathbf{x} = L\mathbf{s}, \tag{1}$$

where f defines the imaging method. The choice of f choice crucially affects the shape of the estimated source distribution, as there are much less constraints on $\mathbf{s}$ than degrees of freedom.

For practical reasons it is desirable to choose f to be convex, as only then numerics guarantees to find the best solution. Important convex functions in this context are ℓ_p-norms. Minimizing a norm of $\mathbf{s}$ is reasonable, since unnecessarily complicated source configurations (e.g. sources with opposite moment at nearby locations) are avoided. The first approach along these lines used the ℓ_2-norm and its solution is traditionally called Minimum Norm Estimate (MNE, Hämäläinen and Ilmoniemi, 1994, extending their 1984 technical report). However, signal attenuation in the tissue causes this method to underestimate deep sources. A method known as sLORETA (Pascual-Marqui, 2002) overcomes the problem by standardizing the MNE and is proven to recover the location of a single point source exactly in the absence of measurement noise. A variety of so-called Weighted Minimum Norm Estimate (WMNE) approaches employ weighting matrices for depth compensation (Jeffs et al., 1987; Köhler et al., 1996). One particular method is LORETA (Pascual-Marqui et al., 1994), which searches for the smoothest current density explaining the data. Minimum ℓ_2-norm methods have the desirable property that they are linear, i.e. their solutions are obtained by simply multiplying a precalculated pseudoinverse matrix to the measurement vector. These solutions, however, tend to be smeared, making it difficult to separate distinct close sources. The occurrence of spurious "ghost sources" is another problem of linear methods.

Smoothness related problems are addressed by Minimum ℓ_1-norm solutions, also referred to as Minimum Current Estimates (MCE, Matsuura and Okabe, 1995). These solutions are sparse, which seems to be congruent with the assumption that only a few narrow regions of the brain are active in a certain experimental condition. This argument has also been used to justify the FO-

CUSS algorithm (Gorodnitsky et al., 1995), which provides even sparser solutions by implicitly minimizing ℓ_p-quasinorms ($p < 1$). Sparse imaging methods usually do not model any spatial relation between dipoles, which causes their solutions to be scattered. For example, such a method may explain a single dipolar source located off-grid by several disconnected dipoles (see Figs. 3 and 4). The spatial scattering problem can be alleviated by averaging the sparse inverse solutions at different time points (Uutela et al., 1999), assuming that the source configuration is stable over time.

Another issue with many sparse approaches is that they do not take into account the vectorial nature of currents. As a result, the orientation of the estimated dipoles are often axes-parallel, as one or two. Several techniques are used to alleviate this problem. One possibility is to a-priori fix the orientations in a meaningful way. In Uutela et al. (1999) the dipole orientations are taken from MNE, while dipole amplitudes are minimized using ℓ_1-norm. A much more complicated approach is suggested in Huang et al. (2006), where activity in a voxel is discouraged, if the orientation of the MNE solution in the respective voxel is close to one of the coordinate axes. In cortically-constrained approaches (Dale and Sereno, 1993; Kincses et al., 2003) dipoles are usually oriented perpendicular to the cortical surface, modeling the apical dendrites of pyramidal neurons, which are known as the main generators of cortical EEG/MEG. This approach, however, requires very precise knowledge of the cortical geometry, as small changes of the normal vector can already lead to considerably different forward equations.

2.2 Focal Vector Field Reconstruction (FVR)

Many real-world signals possess a sparse structure, i.e. they can be expressed by a linear combination of a few basis functions. This concept has been utilized in Basis Pursuit (BP, Chen et al., 1998), where a time series is approximately represented by a small number of Gabor functions from an overcomplete dictionary. Other authors have used sparsity for image denoising (Rudin et al., 1992) and reconstruction (Compressive Sensing, Candes et al., 2006). In the regression and classification context, ℓ_1-norm regularization leads to sparse coefficients (Bennett and Mangasarian, 1992; Tibshirani, 1996; Graepel et al., 1999). In inverse imaging, predominantly sparsity in the "natural" basis of unit impulses has been considered so far, although other bases may be as well useful. We reason, that plausible source estimates should have a relatively simple structure. This is the case for functions with sparse second derivatives, which, in one dimension, are just the piecewise linear functions. In our proposed approach, we impose sparsity of the current density as well as sparsity of its second spatial derivatives. Sources fulfilling both our criteria will be mainly zero, except for a minimal number of continuous patches. Interestingly, this approach has structural similarity to the fused lasso algorithm recently proposed in statistics (Tibshirani et al., 2005), which also considers a joint regularization in two bases.

Discrete Laplace operator

For calculating discrete second derivatives, we consider the Laplacian rather than the full 3×3 Hessian. The Laplacian has the advantage of rotational invariance, compared to other local operators. Assume the brain to be segmented into voxels of size h, the activity in each is represented by a dipole in the center. The $N \times N$ operator is given by

$$D_{i,j}^{(N\times N)} = \frac{1}{h^2}\begin{cases} -\left|\{k \mid \|\mathbf{r}_i - \mathbf{r}_k\|_2 = h\}\right| & i = j \\ 1 & \|\mathbf{r}_i - \mathbf{r}_j\|_2 = h \\ 0 & \text{else}\ , \end{cases} \tag{2}$$

i.e. each diagonal entry $D_{i,i}^{(N\times N)}$ is equal to the number of voxels adjacent to voxel i. With this definition nonzero currents at the boundary are not necessarily penalized. This is important, as cerebral activity measured by EEG/MEG can often be expected to originate from cortical structures. Note that, in contrast to FVR, LORETA (Pascual-Marqui et al., 1994) uses a definition with -6 on the diagonal, regardless of the number of adjacent voxels. While this choice makes the Laplacian non-singular, which is a prerequisite for the analytical inversion carried out by LORETA, it also practically prohibits the correct localization of superficial sources.

Laplace-filtering is done separately for each moment of the current density. Hence, the full $3N \times 3N$ operator can be written as $D = D^{(N\times N)} \otimes I^{(3\times 3)}$, with $I^{(K\times K)}$ being the $K \times K$ identity matrix and $\otimes$ denoting the Kronecker product.

Depth compensation

We conduct a depth compensation, that is inspired by the post-hoc current standardization of sLORETA (Pascual-Marqui, 2002). More precisely, we make use of the source covariance estimate derived in there, which is defined by $\hat{S} = \bar{L}^T \left(\bar{L}\bar{L}^T\right)^{-1} \bar{L} \in \mathbb{R}^{3N\times 3N}$, with $\bar{L} = L$ for MEG and $\bar{L} = HL$ for EEG, where $H = I - \mathbf{1}\mathbf{1}^T/\mathbf{1}^T\mathbf{1} \in \mathbb{R}^{M\times M}$ is the centering matrix and $\mathbf{1}$ a column vector of ones. Let W_i denote the 3×3 matrix square root of the part of $\hat{S}$ belonging to the ith dipole. We include the W_i as penalties in the cost function of FVR, i.e. we penalize large currents at positions with high a-priori uncertainty. This approach differs from the one used in Pascual-Marqui (2002) in that allows to standardize not only current power, but vectorial currents.

Cost function of FVR

A central aspect of our method is the way sparsity of the current density is enforced. We propose to minimize the ℓ_1-norm of the current amplitudes, rather than the individual moments of the current vectors. With this choice, rotational invariance of the FVR solution is guaranteed. Let $\mathbf{s}_i \in \mathbb{R}^3$ denote the dipole moment at the ith voxel such that $\mathbf{s} = (\mathbf{s}_1^T, \ldots, \mathbf{s}_N^T)^T$. Similarly, let $\mathbf{t}_i = \mathbf{t}_i(\mathbf{s}) \in \mathbb{R}^3$ denote the moment of the Laplacian of the source field at the ith voxel. Then the FVR optimization problem takes the following form

$$\begin{aligned} \hat{\mathbf{s}}^{\text{FVR}} &= \arg\min_{\mathbf{s}} \sum_{i=1}^{N} \|W_i\mathbf{s}_i\|_2 + \alpha \sum_{i=1}^{N} \|W_i\mathbf{t}_i\|_2 \\ \text{s.t.} \qquad & \mathbf{x} = L\mathbf{s} \;. \end{aligned} \tag{3}$$

Formulations like Eq. (3), which involve sums of ℓ_2-norms generally arise whenever joint sparsity of groups of variables is desired [1]. While we arrive at the

[1] Note that in Eq. (3) it is not possible to replace the inner ℓ_2-norms by their squared counterparts without losing the sparsity property. This is easily understood, as a sum of squared ℓ_2-norms is nothing but a global ℓ_2-norm, which is known for

FVR objective from a rotational invariance requirement, this concept has also been used for regression (Yuan and Lin, 2006) and sparse spatio-temporal decompositions (Malioutov et al., 2005). Minimizing sums of ℓ_2-norms is harder than traditional ℓ_1-norm or ℓ_2-norm minimization, although not substantially. While the latter problems are solved by linear and quadratic programs, respectively, Eq. (3) can be cast as an instance of Second-Order Cone Programming, (SOCP, see e.g. Lobo et al., 1998). SOCP problems are also convex and thus unambigously solvable (Boyd and Vandenberghe, 2004). By introducing auxiliary variables $\mathbf{u}$ and $\mathbf{v}$, Eq. (3) can be rewritten using SOC constraints

$$\begin{aligned} \hat{\mathbf{s}}^{\mathrm{FVR}} = \arg\min_{\mathbf{s},\mathbf{u},\mathbf{v}} \quad & \sum_{i=1}^{N} u_i + \alpha \sum_{i=1}^{N} v_i \\ \text{s.t.} \quad & \|W_i \mathbf{s}_i\|_2 \leq u_i \ , \qquad i = 1, \ldots, N \\ & \|W_i \mathbf{t}_i\|_2 \leq v_i \ , \qquad i = 1, \ldots, N \\ & \mathbf{x} = L\mathbf{s} \ . \end{aligned} \tag{4}$$

Rotational invariance

If the coordinate system is rotated by an orthogonal matrix U, the lead field $\bar{L}$ and the sources $\mathbf{s}$ are transformed as

$$\bar{L} \longrightarrow \bar{L}\hat{U}^T \equiv \bar{L}_U \tag{5}$$
$$\mathbf{s} \longrightarrow \hat{U}\mathbf{s} \equiv \mathbf{s}_U \tag{6}$$

where

$$\hat{U} \equiv I^{(N \times N)} \otimes U \tag{7}$$

producing nonsparse estimates (as in WMNE and LORETA).

is the rotation operator for all voxels. Then

$$\hat{S} = \bar{L}^T \left(\bar{L}\bar{L}^T\right)^{-1} \bar{L} \longrightarrow \bar{L}_U^T \left(\bar{L}_U \bar{L}_U^T\right)^{-1} \bar{L}_U = \hat{U}\hat{S}\hat{U}^T \tag{8}$$

and the block diagonal entries of the square root of $\hat{S}$ transform as

$$W_i \longrightarrow \left(US_iU^T\right)^{1/2} = U\sqrt{S}_iU^T = UW_iU^T. \tag{9}$$

Now, Eq. 3 is rotationally invariant since a)

$$\|W_i\mathbf{s}_i\| \longrightarrow \|UW_iU^TU\mathbf{s}_i\| = \|W_i\mathbf{s}_i\| \tag{10}$$

and b) the same holds for $\|W_i\mathbf{t}_i\|$ because the Laplacian is a scalar differential operator and the rotation is independent of space (i.e. the moment of each voxel is rotated identically).

Note that a rotation of the coordinate system must be distinguished from a rotation of the grid. Invariance with respect to the former is exactly fulfilled implying that the method itself does not prefer specific source orientations. Rotational invariance of the latter is an approximation limited by the discrete approximation of the Laplacian, however with negligible impact for small voxel distances.

Computational cost

At present, the computational requirements of FVR are quite high. For the 7mm grid used in our experiments (amounting to 6249 dipoles), an inverse calculation ($M \approx 100$) took approximately 45 min on a single-processor computer (2 GHz clock rate, 2 GB memory), compared to 3 min for MCE. For a coarser grid with 1 cm inter-voxel distances (2142 dipoles), the time required by FVR was only 4 min on the same machine.

Several options for accelerating the computation exist. A considerable speedup may already be achieved by reducing the number of constraints using Truncated Singular Value Decomposition (TSVD, see below) of the lead field. Furthermore, the currently used generic solver (Sturm, 1999) could be replaced by a specialized algorithm. Parallel implementations as the one described in Nakata et al. (2006) are also conceivable; this, however, goes beyond the scope of this contribution.

Measuring accuracy of reconstruction results

In order to assess the quality of source reconstructions, we here propose to measure the disagreement of the simulated and the estimated dipole amplitudes by means of the Earth Mover's Distance (EMD, Rubner et al., 2000). This quantity is suitable for comparing distributions with possibly nonoverlapping support, for which a distance measure in the domain space is available. In the case of EEG/MEG inverse solutions the Euclidean distance between dipoles provides such a measure.

To understand the Earth Mover's Distance, consider that for a given source distribution the amplitude at each voxel is divided into a huge number of units [2] with tiny and fixed amplitude. Two source distributions have the same total number of units. One can now transform the first source into the second source by moving the units of the first source to match those of the second

[2] For a formal definition of the Earth Mover's Distance, no division into units is necessary. This was just introduced here to give the reader a better intuition. The exact definition of EMD along with an efficient algorithm for its computation is provided in Rubner et al. (2000).

source. The average distance the units have to be transported depends on the specific transformation we choose. The minimum average distance (averaged over all units and minimized over all possible transformations) defines the EMD. The idea of using this metric in this context is that it provides a meaningful measure for arbitrary types of source distributions. We can, e.g., compare a few dipole solution with highly distributed sources without having to worry which local maximum corresponds to which dipole, or we can compare a 3-dipole solution with a 2-dipole solution in a meaningful way.

3 Results

3.1 Illustration

[Fig. 1 about here.]

Fig. 1 illustrates the main properties of the inverse methods LORETA and MCE compared to that of FVR. The current density domain was defined to be a straight line of 300 scalar sources. Three source configurations, consisting of either three Hanning windows, two boxcar windows or a single sine wave, were simulated. Source reconstruction was performed based on noise-free "measurements", which were obtained by smoothing and subsampling the sources. Apparently, only FVR is able to recover the exact number of sources in all three cases. LORETA is not able to distinguish all three sources in the Hanning example. Instead, one estimated source is placed exactly in between two true sources. MCE estimates consist of spikes, the number and locations not always being in line with the true source configuration.

[Fig. 2 about here.]

In Fig. 2 the effect of enforcing sparse current amplitudes is illustrated on the basis of two simulations. The sources were modeled as a straight line of 100 two-dimensional vectors. In one case we simulated two sources with Hanning window envelopes. In the other example, two boxcar windows were used. All vectors belonging to the same source had equal orientation. Ten pseudo measurements were constructed from the source vectors by means of lowpass-filtering. Note, that for this example the "forward solution" has no physical origin. It was just constructed to contain essential features of real EEG/MEG forward mappings in a simple one-dimensional case and for illustration purposes only. In the examples shown, MCE source estimates according to Matsuura and Okabe (1995) are all parallel to one of the two axes. In contrast to that, the modified version minimizing the ℓ_1-norm of vector amplitudes recovers the original orientations very well, while being even sparser. Finally, the additional sparsity of the amplitudes of the Laplacian removes the problem of source scattering.

3.2 Simulated dipoles

We conducted simulations in a realistic volume conductor using the publicly available Montreal head (Holmes et al., 1998) with three shells (brain, skull, skin). Grids with 7mm and 10mm voxel distances were constructed fully inside the inner shell. The forward problem in a realistic head model was solved using semi-analytic expansions of the electric lead fields (Nolte and Dassios, 2005). We simulated four source configurations, consisting of either two or three dipoles located at random off-grid positions. In each example, either the sagittal or the axial coordinate was the same for all dipoles. The dipoles moment vectors had random orientation and unit amplitude. Hypothetical EEG patterns were constructed by carrying out the forward calculation for 118 standard electrode positions. We investigated the noiseless case as well as the case in which the pattern was superimposed by Gaussian white noise. In each example, the signal-to-noise ratio, defined by signal strength and noise standard deviation averaged over all channels, was set to 5. Inverse imaging solutions were computed according to WMNE, LORETA, MCE (according to Matsuura and Okabe (1995)) and FVR. For WMNE and MCE the sLORETA-based weighting, as well as the standard approach of weighting each moment with the ℓ_2-norm of the corresponding column of L was used.

In the presence of noise, a relaxation of the hard constraint $\mathbf{x} = L\mathbf{s}$ is advisable. Most commonly, Truncated Singular Value Decomposition is used for doing so, while a different option may be given by quadratic constraint

$$\|L\mathbf{s} - \mathbf{x}\|_2 \leq \epsilon. \tag{11}$$

In TSVD, perfect reconstruction is requested only in the space spanned by the right-singular vectors of L belonging to the k largest singular values. This has the consequence that only the low-frequency components of the scalp pattern have to be explained, as these contain the most variance. In contrast, the constraint (11) equips the inverse method with maximal flexibility to "smoothen" the pattern. We therefore adopted this approach and minimized the cost function of each inverse method subject to Eq. (11). This resulted in all cases in convex problems, which were solved exactly using an iterative algorithm (Sturm, 1999). We set ϵ based on our prior knowledge of the noise level, i.e. $\epsilon = \|\mathbf{x}\|_2/5$. In other words, the deviation of the model and the measured electric potential was adjusted to be consisted with statistical expectations.

The tradeoff between sparsity and simplicity of the FVR solution is controlled by means of the model parameter α. The choice of α does not affect the quality of fit of the solutions. For the experiments reported in this paper α was set to 10^{-2} cm^2, i.e. we regarded sparsity more important than simplicity.

[Fig. 3 about here.]

[Fig. 4 about here.]

3.3 Somatosensory Evoked N20

To provide a real world example as a proof of concept, we recorded 113-channel EEG of one male subject (26 years) during electrical median nerve stimulation. EEG electrodes were positioned according to the international 10-20 system and their spatial position was obtained using a 3D digitizer. The electrode positions were mapped later onto the surface of the Montreal head and forward calculations were performed. EEG data were recorded with sampling frequency of 2500 Hz, and digitally bandpass-filtered between 15 Hz and 450 Hz. For the following analysis the data was decimated to 1250 Hz. Left and right median nerves were stimulated in separate blocks by constant square 0.2 ms current pulses with intensities of approx. 9 mA (above motor threshold). The inter-stimulus interval varied randomly between 500 and 700 ms. About 1100 trials were recorded for each hand. The study was approved by the local Ethics Committee of the Charité, University Medicine Berlin.

Electrodes were excluded from the analysis if standard deviation at these electrodes exceeded 50 µV. The remaining 106 channels were segmented into epochs in a time interval from -100 ms to 70 ms relative to the stimulus onset. Baseline correction was based on the mean amplitude in the prestimulus interval (-100 ms to -10 ms). An epoch was rejected from the averaging, if its amplitude was more than 100 µV in either the prestimulus or poststimulus interval (10 ms to 70 ms). After this, at least $l = 973$ epochs remained in each class. These epochs were averaged separately for the left and right median nerve stimulation. Visual inspection of the averaged signals revealed that the peak time of the N20 deflection was approximately at 21 ms. Fig. 5 shows both the average time courses of both conditions, as well as the average potential patterns at this time. A combined pattern was created by arithmetic summation of the patterns related to left and right N20.

[Fig. 5 about here.]

We inverted the single left and single right as well as the summed pattern, amounting to a (joint) localization of the left and right N20 generators. The methods tested were LORETA, MCE and FVR. For MCE, the sLORETA-based depth compensation was employed. All three methods were required to provide the same quality of fit. The regularization parameter ϵ was set to $\|\mathrm{SE}(X)\|_2 / \|x\|_2$, where X is the $m \times 973$ matrix containing the left, right and summed trials, respectively, and $\mathrm{SE}(X)$ is the $m \times 1$ vector of electrode standard errors.

[Fig. 6 about here.]

In cases like above, where the presence of more than one source is indicated, an automatic decomposition of the estimated current density is desirable. In the case of sparse solutions, such a decomposition is easily obtained by computing the connected components (with respect to the grid neighborhood relation) of the set of dipoles having nonzero estimated amplitude.

[Fig. 7 about here.]

3.4 Quantitative performance analysis

We also performed a quantitative comparison of the inverse solutions. For that purpose, source localization was repeatedly performed within a 5×5 crossvalidation, i.e. for each localization task the following procedure was carried out five times. The channels were randomly divided into five sets of equal size. For each union of four sets, inverse solutions were computed.

The patterns to be inverted were grouped into the N20 evoked potential, the noiseless and the noisy simulated patterns. For the simulations, localization in terms of the Earth Movers distance was considered the ultimate performance measure. Apart from localization, we defined a number of performance criteria that do not rely on explicit knowledge of the true sources. These include sparsity, defined as the fraction of dipoles with (close to) zero amplitude. The generalization error of a crossvalidation run was defined as the mean squared difference of the measurement at those channels, which were taken out for the source estimate, and the prediction for these channels based on the estimated source distribution. Finally, stability of the solution was assessed as the sum of the variances of the dipole moments over the 25 crossvalidation runs. Table 1 lists the results of the numerical analysis. Mean and standard errors (SE) were computed across the experiments of a group and the crossvalidation runs. As stability aggregates information of all runs, mean and SE were taken across experiments only for this measure.

[Table 1 about here.]

4 Discussion

In the simulations, FVR outperformed the other methods with significantly better localization in both the noiseless and the noisy case. In the noiseless simulations, FVR also had by far the highest stability and generalization performance. In the presence of noise (simulations and N20 localization), this advantage became insignificant for generalization (on par with sLORETA weighted MCE) and vanished for stability. Here, WMNE and LORETA achieved the best scores, followed by FVR. High stability seems, however, less valuable in conjunction with a large localization bias, as it is indicated for these methods in Table 1. The sLORETA weighted MCE outperformed all other methods significantly in terms of sparsity, which was above 99 % on average. Column-norm weighted MCE and FVR were, however, also very sparse (above 97 %).

The good localization of FVR becomes apparent also in Figs. 3 and 4, which show inverse solutions based on the whole set of 118 channels. FVR was the only method that had exactly as many distinct active patches as there were true sources. In the noiseless setting, the centers of gravity of these patches were located on top of the simulated dipoles (at the closest gridpoints). With noise added to the pattern, only a small offset was observed for some sources. For WMNE and MCE the column-norm weighting was not a sufficient depth compensation. These methods became comparative to FVR only when the sLORETA based weighting was used.

The sources estimated by LORETA and WMNE are typically distributed over the whole brain. For this reason, it is common in smooth inverse methods to analyze only maximal values of the source distribution (although this is in disagreement to the model the estimation was based on). However, for some of the simulations in Figs. 3 and 4 this analysis does not yield acceptable results, as the local maxima of the LORETA/WMNE estimates were not even close to some of the true sources. For example, LORETA estimated a spurious ghost source even in a noisefree simulation (SAG3), while at the same time two real sources were spuriously merged into one local maximum in the middle. Also thresholding, which is another popular way to preprocess smooth estimates, would not alleviate this problem.

The difference in the estimates obtained from noisy and noisefree patterns were relatively small for WMNE, LORETA and FVR. For MCE, on the other hand, virtually disjoint sets of dipoles were predicted to be active. MCE solutions usually featured several spikes, that were scattered around a true source in the noisefree case. Due to the scatter, the distinction of sources being than a few centimeters apart was hardly possible (see e.g. example AX3).

In the localization of the single left and single right N20 component, LORETA and FVR detected strongest currents in the contralateral somatosensory cortex (inverse solutions are not shown here). This is in good agreement with the localization of the hand areas reported in the literature (Huttunen et al., 2006; Komssi et al., 2004). Both methods estimated one source centered in the respective somatosensory area. The extension of this source, however, was too large to be realistic for LORETA, whereas it was much smaller for FVR [3].

[3] Of course, a realistic spatial extent of the sources does not imply, that the exact shape and size of the true sources is always recovered. This cannot be achieved by

MCE revealed several activation spots (mainly single dipoles) in the contralateral sensorimotor cortex.

Simultaneous localizations of left and right N20 generators performed with LORETA, sLORETA weighted MCE and FVR are shown in Fig. 6. The FVR solution showed two major patches, which closely match the estimates from single pattern localization. LORETA, on the other hand, estimated only one large active region over the central area, with the maximum lying exactly in between the two individually estimated sources. The MCE solution consisted of several small patches scattered across the whole somatosensory area in the proximity of activation spot obtained in the single pattern localizations.

Due to the linearity of the forward equation, the estimates from the simultaneous localizations should ideally be just the sum of activations obtained from inverting the single left and single right N20 patterns. As mentioned above, this was approximately the case for FVR, but dids not hold for LORETA. The better ability of FVR to recover the same sources in both cases manifests also in the low EMD between its joint source reconstruction and the sum of its single-pattern reconstructions, which was only 0.76 compared to 1.20 for LORETA and 3.29 for MCE.

We conducted a connected components analysis of the FVR inverse solution for the summed N20 pattern. For LORETA and MCE this did not seem helpful, as the analysis returned either too few (one) or too many (more than 20) components, which were considered unrealistic. The FVR source distribution revealed five distinct sources. They are shown in Fig. 7, along with the individual scalp patterns obtained from forward calculations. For each component

any method, due to the genuine ambiguity of inverse problems.

a dipole having the mean orientation was drawn at the center of gravity. The components were named C1 to C5 according to the decreasing ℓ_2-norm of their EEG patterns. The two strongest sources according to this criterion are the N20 generators. Their patterns resemble very much the single-component patterns shown in Fig. 5, except that the latter ones are a bit more central and even contralateral. It seems that the residuals were combined in a third, more central component with questionable physiological relevance. However, this component is already three times less pronounced (in terms of the norm of its pattern) than the first two. Judging by its characteristic frontal pattern, component C4 seems to contain an eyeblink artifact. This is in line with the location of C4, which is maximally close to the eyes. Component C5 has a rather negligible influence on the EEG measurement and may be considered biological noise.

Using FVR and performing the connected components analysis we could clearly separate task related activity of interest from artifictual activity. Thus, as a side effect, the method can be used to reject artifacts based on a purely spatial criterion, given that brain regions likely to pick artifacts are known. At this point, it should also be mentioned that task related and artifactual components were inseparably mixed in the LORETA estimate. A similar problem can be expected to occur in cortically constrained approaches, where deep artifactual sources can only be modeled using large parts of the cortex.

5 Conclusion and Outlook

In this paper we have introduced Focal Vector Field Reconstruction as a new method for localizing generators of human brain activity on the basis of an

EEG or MEG measurement. So far, the method was formulated to analyze an evoked potential at a single time point. For the future, we intend to generalize this to more complex cases such as time-dependent evoked potential, cross-spectral matrices, and subspaces defined by the latter. We expect that in all these cases the problem will ultimately lead to a Second-Order Cone Problem. This emerged here naturally from the requirement of rotational invariance, but always follows whenever sparsity makes sense e.g. in spatial dimension but not in other dimensions like dipole moment, time or frequency.

Future studies will also apply our novel Focal Vector Field Reconstruction to complex imaging paradigms and use it for contributing to the analysis of interactivity and causality of neural information processing. Note, however, that the numerics of solving a SOCP will need further improvements as the ultimate goal would be to have a fast solver that is able to model complex brain signals in real-time.

Acknowledgements

This work was supported in part by the Bundesministerium für Bildung und Forschung (16SV2234, 01GQ0415), the Deutsche Forschungsgemeinschaft (MU 987/3-1) and the IST Programme of the European Community, under the PASCAL Network of Excellence (IST-2002-506778). We thank Friederike Hohlefeld and Monika Weber for help in preparing the experiment, and Ryota Tomioka for fruitful discussions.

References

Baillet, S., Garnero, L., 1997. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans. Biomed. Eng. 44, 374–385.

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

Bennett, K. P., Mangasarian, O. L., 1992. Robust linear programming discrimination of two linearly inseparable sets. Optim. Method. Softw. 1, 23–34.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press, Cambridge.

Candes, E., Romberg, J., Tao, T., 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory 52, 489–509.

Chen, S. S., Donoho, D. L., Saunders, M. A., 1998. Atomic decomposition by basis pursuit. SIAM J. Sci. Comput. 20, 33–61.

Cotter, S. F., Rao, B. D., Engan, K., Kreutz-Delgado, K., 2005. Sparse solutions to linear inverse problems with multiple measurement vectors. IEEE Trans. Signal Proces. 53, 2477–2488.

Dale, A. M., Sereno, M. I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. J. Cognitive Neurosci. 5, 162–176.

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.

Graepel, T., Herbrich, R., Schölkopf, B., Smola, A., Bartlett, P., Müller, K.-R., Obermayer, K., Williamson, R., 1999. Classification on proximity data with LP–machines. In: Proceedings of the Ninth International Conference on Artificial Neural Networks. pp. 304–309.

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

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., Evans, A. C., 1998. Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. 22, 324–333.

Huang, M.-X., Dale, A. M., Song, T., Halgren, E., Harrington, D. L., Podgorny, I., Canive, J. M., Lewise, S., Lee, R. R., 2006. Vector-based spatial-temporal minimum L1-norm solution for MEG. NeuroImage 31, 1025–1037.

Huttunen, J., Komssi, S., Lauronen, L., 2006. Spatial dynamics of population activities at S1 after median and ulnar nerve stimulation revisited: An MEG study. NeuroImage 32, 1024–1031.

Jeffs, B., Leahy, R., Singh, M., 1987. An evaluation of methods for neuromagnetic image reconstruction. IEEE Trans. Biomed. Eng. 34, 713–723.

Kincses, W. E., Braun, C., Kaiser, S., Grodd, W., Ackermann, H., Mathiak, K., 2003. Reconstruction of extended cortical sources for EEG and MEG based on a monte-carlo-markov-chain estimator. Human Brain Mapping 18, 100–110.

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

Komssi, S., Huttunen, J., Aronen, H. J., Ilmoniemi, R. J., 2004. EEG minimum-norm estimation compared with MEG dipole fitting in the localization of somatosensory sources at S1. Clin. Neurophysiol. 115, 534–542.

Lobo, M. S., Vandenberghe, L., Boyd, S., Lebret, H., 1998. Applications of second-order cone programming. Lin. Alg. Appl. 284, 193–228.

Malioutov, D., Çetin, M., Willsky, A. S., 2005. A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Signal Proces. 55, 3010–3022.

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

Mauguière, F., Merlet, I., Forss, N., Vanni, V. Jousmäki, P. A., Hari, R., 1997. Activation of a distributed somatosensory cortical network in the human brain. a dipole modelling study of magnetic fields evoked by median nerve stimulation. part I: location and activation timing of SEF sources. Electroenceph. Clin. Neurophysiol. 104, 281–298.

Mosher, J. C., Leahy, R. M., 1999. Source localization using recursively applied and projected (RAP) MUSIC. IEEE Trans. Signal Proces. 47, 332–340.

Nakata, K., Yamashita, M., Fujisawa, K., Kojima, M., 2006. A parallel primal-dual interior-point method for semidefinite programs using positive definite matrix completion. Parallel Comput. 32, 24–43.

Nolte, G., Dassios, G., 2005. Analytic expansion of the EEG lead field for realistic volume conductors. Phys. Med. Biol. 50, 3807–3823.

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

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

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

Polonsky, A., Zibulevsky, M., 2004. MEG/EEG source localization using spatio-temporal sparse representations. In: Puntonet, C. G., Prieto, A. (Eds.), Independent Component Analysis and Blind Signal Separation, Fifth International Conference. Lecture Notes in Computer Science. Springer, pp. 1001–1008.

Rubner, Y., Tomasi, C., Guibas, L. J., 2000. The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vision 40, 99–121.

Rudin, L. I., Osher, S., Fatemi, E., 1992. Nonlinear total variations based noise removal algorithms. Physica D 60, 259–268.

Scherg, M., von Cramon, D., 1986. Evoked dipole source potentials of the human auditory cortex. Electroenceph. Clin. Neurophysiol. 65, 344–360.

Schmidt, R. O., 1986. Multiple emitter location and signal parameter estimation. IEEE Trans. Antenn. Propag. 43, 276–280.

Sturm, J. F., 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Method. Softw. 11–12, 625–653.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. Roy. Stat. Soc. B Meth. 58, 267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. J. Roy. Stat. Soc. B Meth. 67, 91–108.

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

Veen, B. D. V., Buckley, K. M., 1988. Beamforming: a versatile approach to spatial filtering. IEEE ASSP Magazine 5, 4–24.

Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. J. Roy. Stat. Soc. B Meth. 68, 49–67.

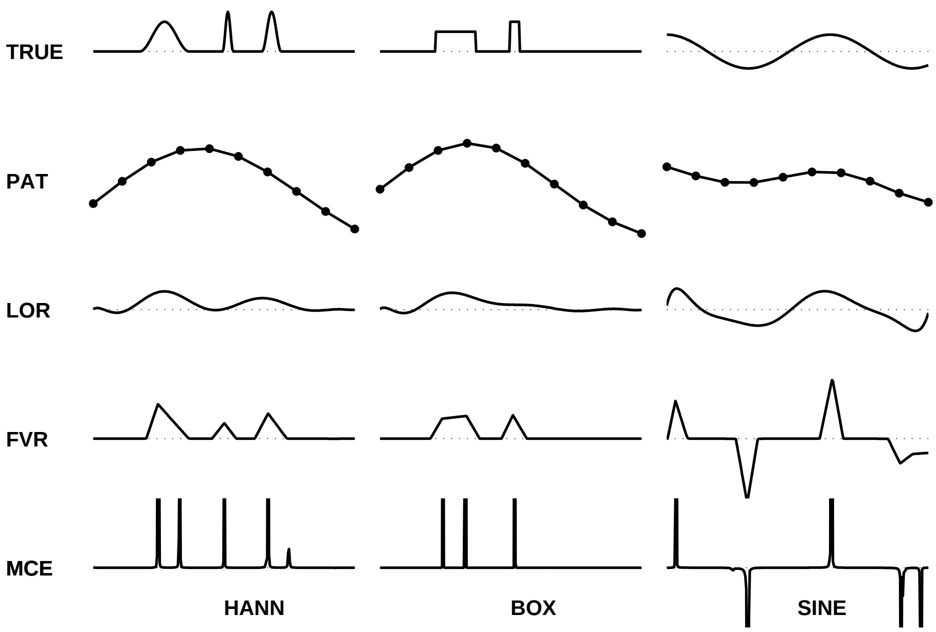

Fig. 1. One-dimensional simulations illustrating the characteristics of standard inverse solutions and FVR. Simulated source configurations (TRUE) include three hanning windows (HANN), two boxcar windows (BOX) and a sine wave (SINE). Hypothetical measurements (PAT) were obtained by smoothing and subsampling the sources. Source reconstructions according to one-dimensional versions of LORETA (LOR), FVR and MCE are shown in the three lower panels.

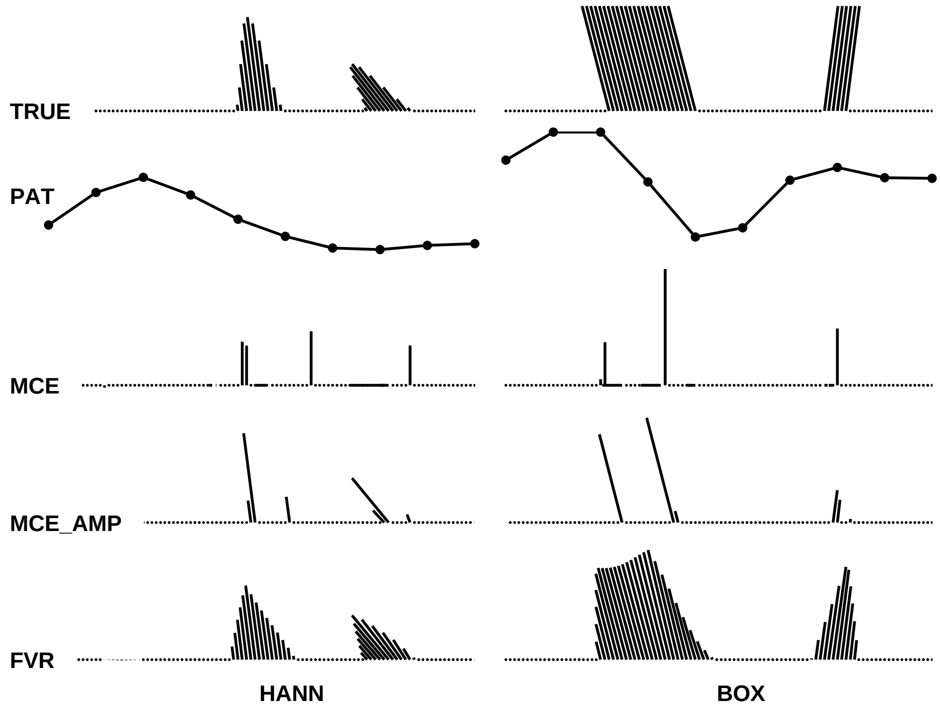

Fig. 2. Simulations, illustrating the approaches of MCE and FVR to achieve sparsity of a vector field. A straight line of two-dimensional vectors models the current density. The vector envelopes were taken to be combinations of either two Hanning (HANN) or two boxcar windows (BOX). Vector orientations were fixed within sources. True sources (TRUE) and pseudo patterns are shown in the upper panels of the plot. The inverse solutions of MCE, the corrected version of MCE working on amplitudes (MCE_AMP) and FVR are shown below.

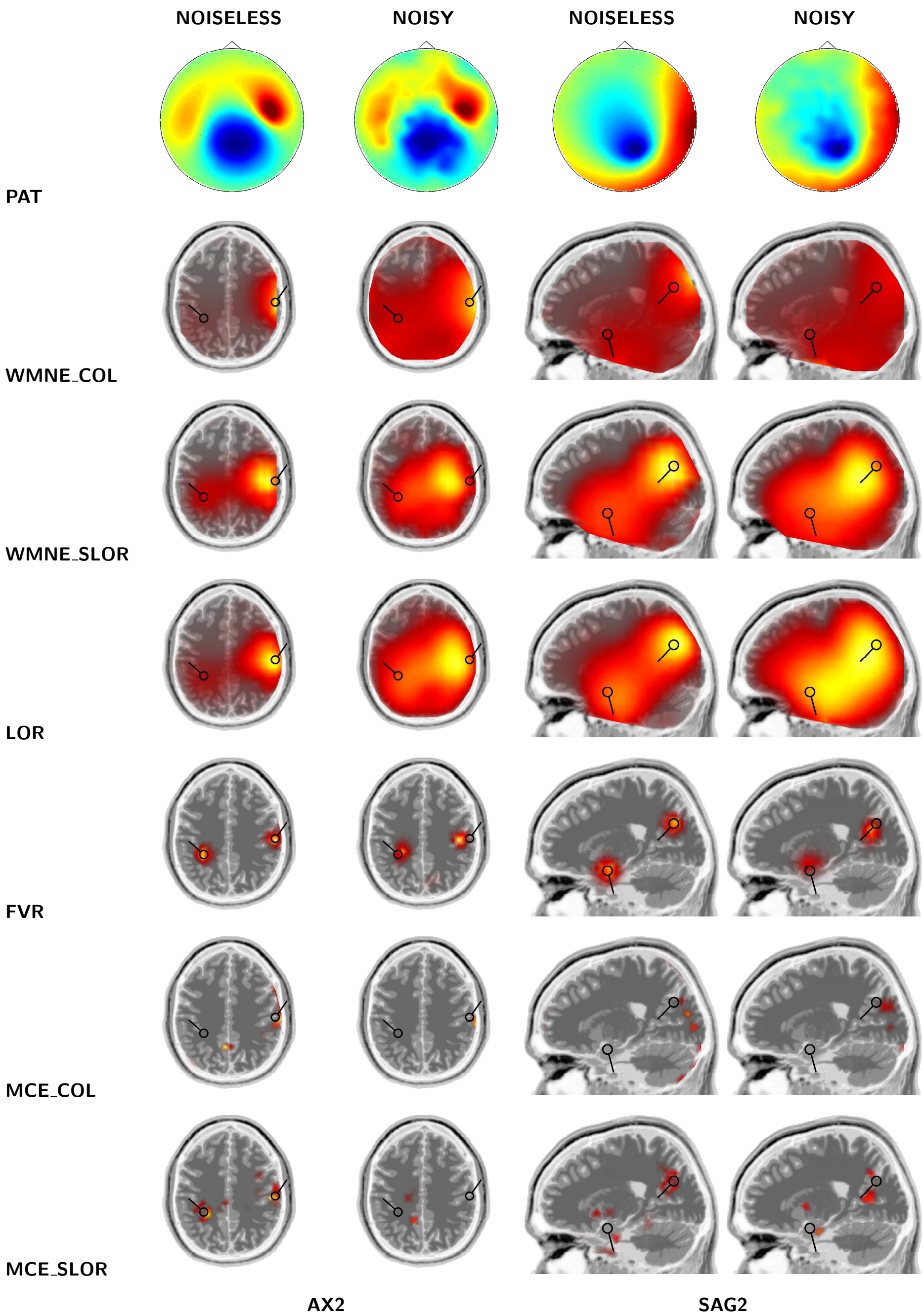

Fig. 3. Comparison of inverse solutions in two realistic examples. In each, two dipoles (black color) were put into an either axial (AX2) or sagittal (SAG2) slice of the brain. The resulting scalp patterns with noise (NOISY) or without it (NOISELESS) are shown in the top panel. Panels below show current densities reconstructed from these patterns by means of column-norm weighted Minimum Norm (WMNE_COL), sLORETA weighted MNE (WMNE_SLOR), LORETA (LOR), FVR, column-norm weighted MCE (MCE_COL) and sLORETA weighted MCE (MCE_SLOR). Dipole amplitudes are shown color-coded (red = low, yellow = high) with scales adjusted to the range of the individual solution. In each plot the average activity within 1 cm is shown. Mean amplitudes exceeding 7 % of the individual scale are shown opaque. Between 0 % and 7 % opacity is linearly scaled between 0 and 1. Spaces between grid dipoles are interpolated.

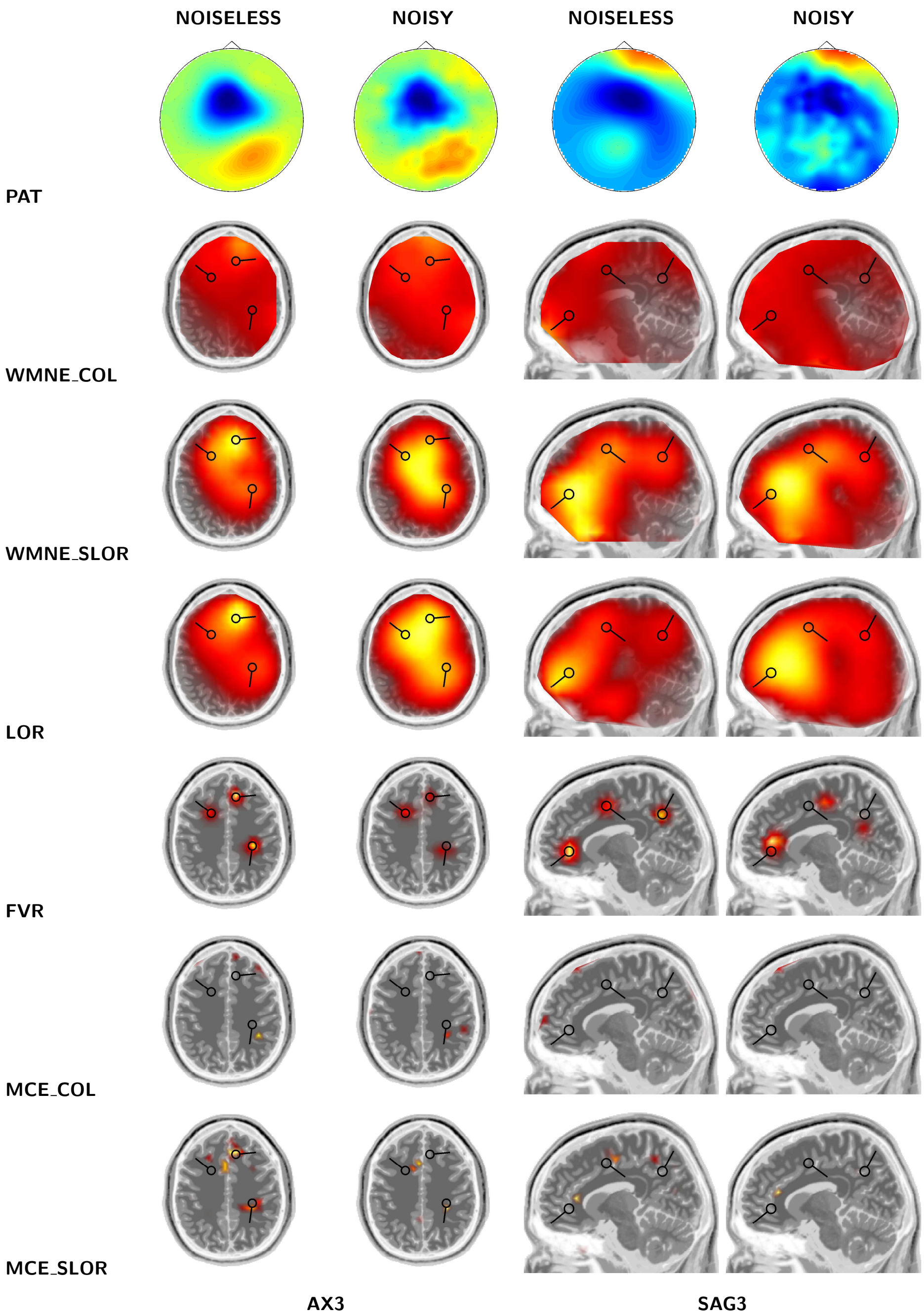

Fig. 4. Comparison of inverse solutions in examples with three simulated dipoles.

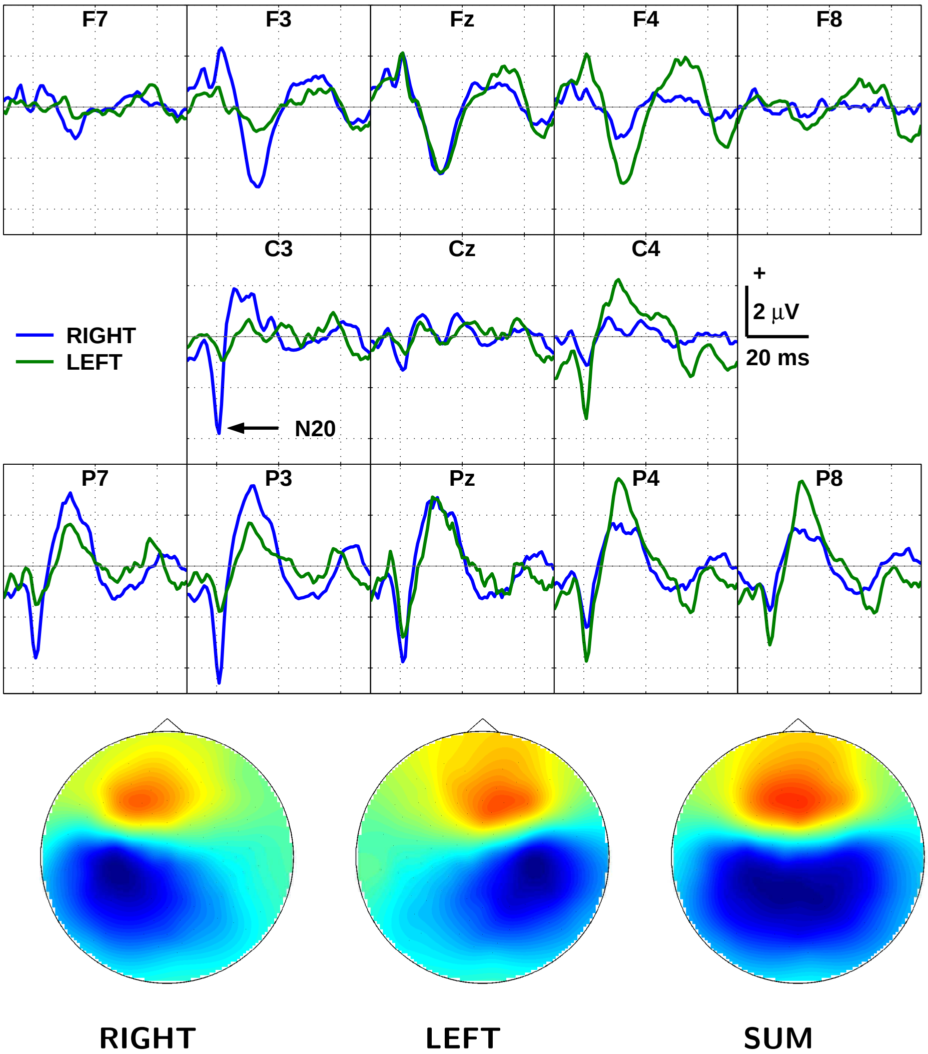

Fig. 5. Somatosensory Evoked N20 after left and right median nerve stimulation. Upper part: Averaged time series between 10 ms and 70 ms after stimulus onset. Lower part: Averaged scalp patterns at 21 ms and sum of left and right pattern.

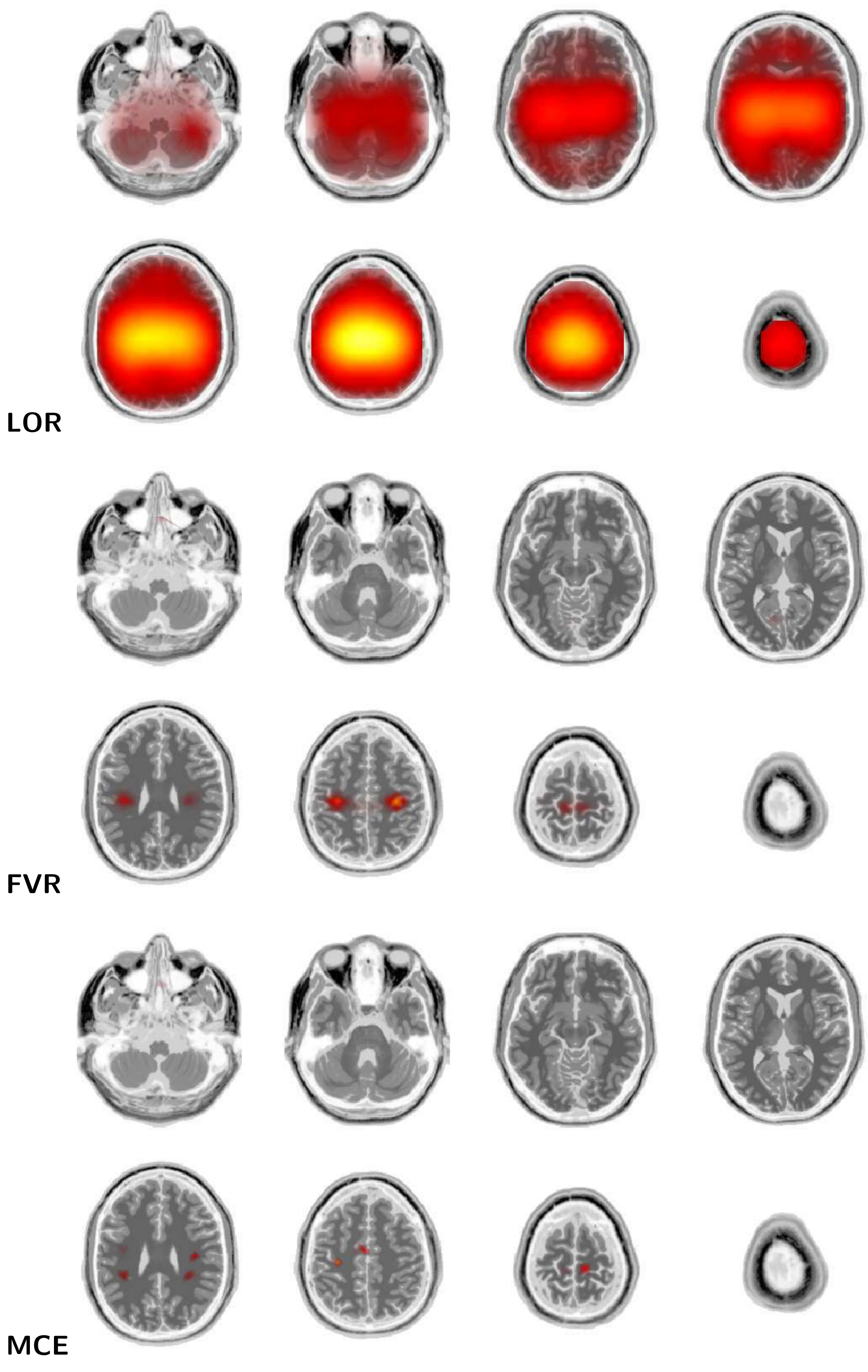

Fig. 6. Source localization of summed left and right N20 component. The average estimated dipole amplitudes of eight consecutive axial slices (thickness 2 cm) of the brain is shown for the inverse solutions of LORETA (LOR), FVR and MCE.

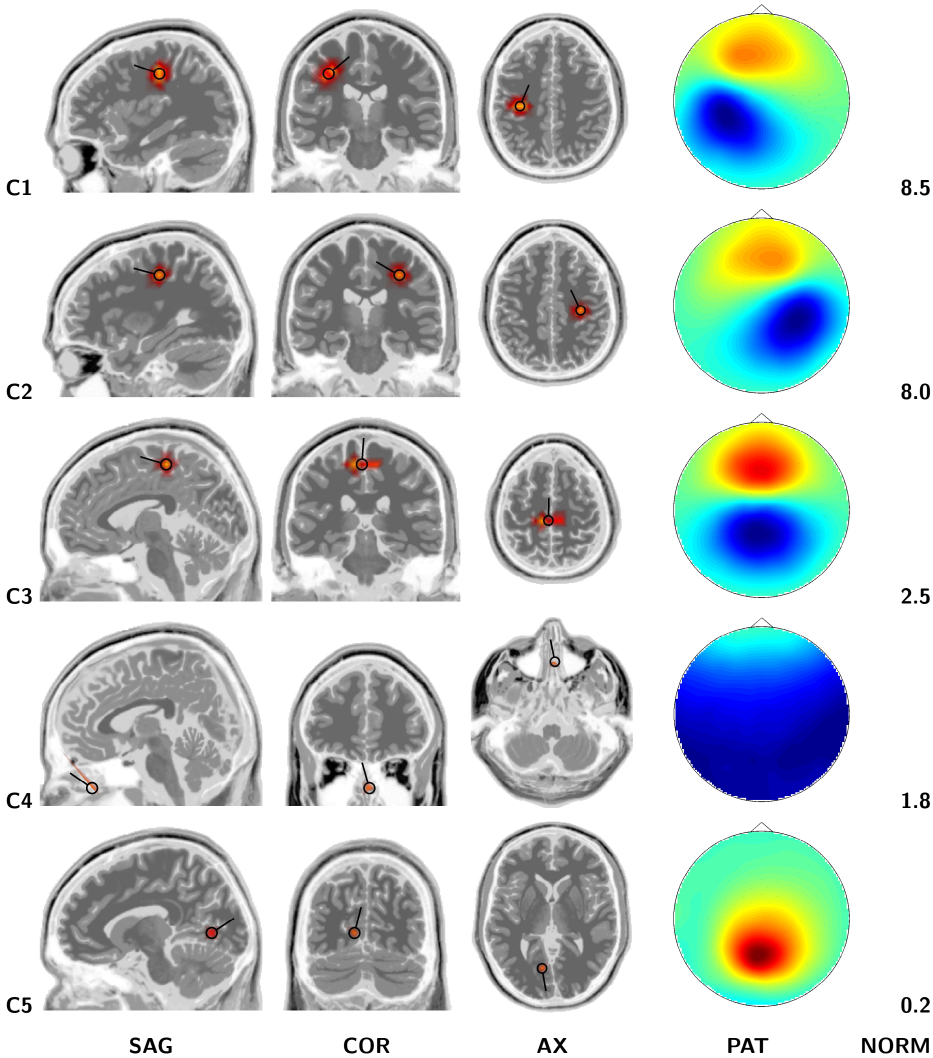

Fig. 7. Connected component analysis of the FVR inverse solution for the summed N20 pattern. Components were sorted from top to bottom according to the ℓ_2-norm (NORM) of their scalp patterns (PAT). For each source component, the average dipole amplitudes of 1 cm sagittal (SAG), coronal (COR) and axial (AX) slices around the source gravity center are shown (red and yellow color, different scale for each component). Additionally, a single dipole (black) having the mean orientation of the source and unit amplitude is drawn at the gravity center.

	LOCALIZATION	STABILITY $\times 10^5$	GENERALIZATION $\times 10^2$	SPARSITY $\times 10^2$
SIM_NOISELESS				
WMNE_COL	4.70 ± 0.03	0.31 ± 0.18	1.28 ± 0.15	09.1 ± 0.38
WMNE_SLOR	4.36 ± 0.06	0.38 ± 0.24	1.31 ± 0.15	19.6 ± 0.56
LOR	4.35 ± 0.06	0.31 ± 0.23	1.31 ± 0.16	14.1 ± 0.55
FVR	***2.11 ± 0.10***	***0.03 ± 0.02***	***0.03 ± 0.00***	97.5 ± 0.11
MCE_COL	4.56 ± 0.10	3.27 ± 0.29	3.23 ± 0.17	98.8 ± 0.00
MCE_SLOR	2.56 ± 0.12	1.25 ± 0.16	0.15 ± 0.01	***99.2 ± 0.03***
SIM_NOISY				
WMNE_COL	5.17 ± 0.03	**0.20 ± 0.04**	9.45 ± 0.23	19.7 ± 0.39
WMNE_SLOR	4.66 ± 0.04	***0.18 ± 0.04***	9.31 ± 0.21	29.0 ± 0.30
LOR	4.66 ± 0.04	***0.18 ± 0.04***	9.29 ± 0.20	26.9 ± 0.47
FVR	***2.40 ± 0.07***	1.34 ± 0.13	***8.58 ± 0.17***	97.4 ± 0.05
MCE_COL	5.16 ± 0.10	2.90 ± 0.43	11.0 ± 0.23	99.6 ± 0.00
MCE_SLOR	2.79 ± 0.07	2.76 ± 0.16	**8.80 ± 0.19**	***99.8 ± 0.00***
N20				
WMNE_COL		0.08	**4.03 ± 0.18**	14.3 ± 0.15
WMNE_SLOR		0.08	**3.97 ± 0.16**	33.2 ± 0.03
LOR		0.06	**4.02 ± 0.15**	30.6 ± 0.47
FVR		1.29	***3.87 ± 0.13***	98.1 ± 0.05
MCE_COL		2.61	4.59 ± 0.16	99.7 ± 0.00
MCE_SLOR		2.53	**3.91 ± 0.16**	***99.8 ± 0.00***

Table 1
Performance of column-norm weighted minimum norm (WMNE_COL), sLORETA weighted MNE (WMNE_SLOR), LORETA (LOR), FVR, column-norm weighted MCE (MCE_COL) and sLORETA weighted MCE (MCE_SLOR) in noiseless simulations (SIM_NOISELESS), noisy simulations (SIM_NOISY) and somatosensory evoked N20 localization (N20). Winning entries in each category are shown in slanted font. Entries being within a confidence interval of three standard errors around the winner are shown in bold.