A Distributed Spatio-temporal EEG/MEG Inverse Solver

Wanmei Ou[1], Polina Golland[1], and Matti Hämäläinen[2]

[1] Computer Science and Artificial Intelligence Laboratory, MIT, USA
[2] Athinoula A. Martinos Center for Biomedical Imaging, MGH, USA
wanmei@csail.mit.edu

Abstract. We propose a novel $\ell_1\ell_2$-norm inverse solver for estimating the sources of EEG/MEG signals. Based on the standard ℓ_1-norm inverse solver, the proposed sparse distributed inverse solver integrates the ℓ_1-norm spatial model with a temporal model of the source signals in order to avoid unstable activation patterns and "spiky" reconstructed signals often produced by the original solvers. The joint spatio-temporal model leads to a cost function with an $\ell_1\ell_2$-norm regularizer whose minimization can be reduced to a convex second-order cone programming problem and efficiently solved using the interior-point method. Validation with simulated and real MEG data shows that the proposed solver yields source time course estimates qualitatively similar to those obtained through dipole fitting, but without the need to specify the number of dipole sources in advance. Furthermore, the $\ell_1\ell_2$-norm solver achieves fewer false positives and a better representation of the source locations than the conventional ℓ_2 minimum-norm estimates.

1 Introduction

Localizing activated regions from Electroencephalography (EEG) or Magnetoencephalography (MEG) data involves solving an ill-posed inverse problem. This paper introduces an integrated spatio-temporal regularizer to overcome the instabilities in the standard sparse solutions.

There are two main types of inverse solvers in EEG/MEG applications: discrete parametric solvers, also known as dipole fitting [13], and distributed inverse solvers [7,17]. The standard dipole fitting algorithms estimate the location, orientation, and amplitudes of a fixed number of current dipoles. While dipole fitting often provides robust estimates for activation signals, localization is challenging when several sources are active. The quality of the results degrades substantially when the assumed number of dipoles differs from the true number [9].

In contrast, distributed solvers discretize the source space into locations on the cortical surface or in the brain volume without fixing the number of current dipoles. The solution is computed by minimizing a cost function that depends on all sources in the source space. The widely used minimum norm estimate (MNE) [7] recovers a source distribution with minimum overall energy (ℓ_2-norm) that produces data consistent with the measurements. Although MNE leads to

a linear inverse operator, its solutions are often too diffuse to be biologically plausible, especially in early sensory activations. Other regularizers based on a norm penalty were proposed to model sparsity in such applications. Among them, the minimum current estimate (MCE) (ℓ_1-norm) is the most popular [17].

One of the drawbacks of the conventional MCE is its sensitivity to noise. Similar to other distributed solvers, the conventional MCE is computed at each time point separately. The solver's sensitivity to noise causes the estimated activations to "jump" among neighboring spatial locations from one time instant to another. Equivalently, the time course at a particular location can show substantially "spiky" discontinuities. To address this problem, the vector-based spatio-temporal minimum ℓ_1-norm solver (VESTAL) [10] projects the point-wise ℓ_1-norm estimates onto a set of temporal basis functions estimated from the data. Some other wavelet-based methods perform reconstruction of the coefficients of each basis function separately [6,16].

Similar to MCE, we employ the ℓ_1-norm regularizer to encourage spatial sparsity. Furthermore, we incorporate a temporal model of the source signals into the regularizer. Specifically, we assume that the source signals are linear combinations of *multiple* temporal basis functions, and apply the distributed inverse solver to the coefficients of all basis functions simultaneously, in contrast to the two-step approach proposed in [10]. This combined spatio-temporal regularizer is at the core of our $\ell_1\ell_2$-norm inverse solver, and it is motivated by a previously demonstrated method in farfield narrowband sensor array applications [12]. We construct the temporal basis functions through singular-value decomposition (SVD) of the measurements, since it compactly represents the signal subspace. Although we focus on the EEG/MEG application, the proposed framework is applicable to computed tomography reconstruction, with modifications on the spatial model so as to encourage piece-wise constant solutions (i.e., ℓ_1-norm on spatial derivatives of the source).

To summarize, the proposed solver imposes ℓ_1-norm regularization in the spatial domain and ℓ_2-norm regularization in the temporal domain. The resulting inverse problem can be formulated as a second-order cone programming (SOCP) problem and solved efficiently using the interior-point method [1]. Thanks to the integrated spatio-temporal model, our method achieves accurate reconstruction results and outperforms other solvers including MNE [7], MCE [17], and VESTAL [10].

2   Method

Under the *quasi-static* approximation of Maxwell's equations, the observed EEG/MEG signal $\mathbf{y}(t)$ at time t is a linear function of the current sources $\mathbf{s}(t)$:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t), \tag{1}$$

where $\mathbf{A}$ is the $N \times M$ lead-field matrix. $\mathbf{e}(t) \sim \mathcal{N}(\mathbf{0}, \Sigma)$ is the measurement noise, where the noise covariance Σ can be estimated from the pre-stimulus data. $\mathbf{s}(t)$, $N \times 1$, and $\mathbf{y}(t)$, $M \times 1$, are vectors in the source space and the signal space,

respectively. The number of sources N $(\sim 10^3 - 10^4)$ is much larger than the number of measurements M $(\sim 10^2)$, leading to an infinite number of solutions satisfying Eq. (1) even for $\mathbf{e}(t) = \mathbf{0}$. Without loss of generality, we can apply spatial whitening based on the estimated noise covariance Σ to both the data and the lead-field matrix, leading to $\mathbf{e}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ in the derivations.

Spatio-Temporal Model. The quasi-static assumption allows inverse estimation for each time instant independently. However, this often results in highly variable source time courses. To mitigate this problem, we utilize the knowledge of the temporal properties of the source signals to further constrain the solution. To this end, we express the data model in Eq. (1) for all time instants:

$$\mathbf{Y} = \mathbf{AS} + \mathbf{E}, \tag{2}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \cdots, \mathbf{y}(T)]$ is an $M \times T$ matrix that contains EEG/MEG measurements for all T temporal samples, and $\mathbf{S}$ is an $N \times T$ matrix that represents the source signals. Noise $\mathbf{E}$ is assumed to be independent in time.

The underlying source signals vary smoothly but with sharp changes at certain deflections [9]. To model the time-varying frequency content of the signals, we assume that the source signals are linear combinations of multiple orthonormal temporal basis functions $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_K]$ that collectively capture the temporal properties of the source signals. $\mathbf{v}_k$, $T \times 1$, denotes the kth basis function. We can therefore transform the problem into a much lower dimensional space of projection coefficients $\widetilde{\mathbf{Y}} = \mathbf{A}\widetilde{\mathbf{S}} + \widetilde{\mathbf{E}}$, where $\widetilde{\mathbf{Y}} = \mathbf{YV}$, $\widetilde{\mathbf{S}} = \mathbf{SV}$, and $\widetilde{\mathbf{E}} = \mathbf{EV}$. The (n, k) element of $\widetilde{\mathbf{S}}$, $\widetilde{s}_{nk}$, indicates the kth coefficient for the source signal at location n. Denoting the kth column of $\widetilde{\mathbf{E}}$ by $\widetilde{\mathbf{e}}_k$, the temporal independence assumption of $\mathbf{E}$ and orthonormality of $\mathbf{V}$ imply that $\widetilde{\mathbf{e}}_k \sim \mathcal{N}(\mathbf{0}; \mathbf{I})$, and that $\widetilde{\mathbf{e}}_k$ and $\widetilde{\mathbf{e}}_{k'}$ are independent for $k \neq k'$.

To compute the inverse solutions for all K basis functions simultaneously, we use the signal magnitude in the subspace spanned by $\mathbf{V}$, $\sqrt{\sum_{k=1}^{K} \widetilde{s}_{nk}^2}$, as an indicator of the activation status at location n. In other words, we apply ℓ_2-norm regularization to the K coefficients for each source location. Because we choose to work with orthonormal basis functions, the ℓ_2-norm of the reconstructed source signal in the temporal domain is equal to the ℓ_2-norm in the transformed domain. However, we find it more intuitive to present the model in the transformed domain. Furthermore, if two bases span the same subspace, the reconstruction results using these bases are guaranteed to be identical.

With the ℓ_1-norm regularization in the spatial domain and the ℓ_2-norm regularization in the temporal domain, we incorporate the integrated spatio-temporal $\ell_1\ell_2$-norm regularizer

$$|\widetilde{\mathbf{S}}|_{\ell_1}^{\ell_2} = \sum_{n=1}^{N} \sqrt{\sum_{k=1}^{K} \widetilde{s}_{nk}^2} \tag{3}$$

into the estimation problem:

$$\widetilde{\mathbf{S}}^* = \arg\min_{\widetilde{\mathbf{S}}} \sum_{k=1}^{K} \|\widetilde{\mathbf{y}}_k - \mathbf{A}\widetilde{\mathbf{s}}_k\|_{\ell_2}^2 + \lambda|\widetilde{\mathbf{S}}|_{\ell_1}^{\ell_2}, \tag{4}$$

where $\widetilde{\mathbf{s}}_k$ and $\widetilde{\mathbf{y}}_k$ are the kth column vectors in $\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{Y}}$. λ controls the regularization strength; $\|\mathbf{x}\|_{\ell_2} = \sqrt{\mathbf{x}^{\text{T}}\mathbf{x}}$. Since the ℓ_2-norm does not encourage sparsity,

many coefficients for an active location are usually non-zero in the estimates. The reconstructed source signals are then obtained as linear combinations of the basis functions: $\mathbf{S}^* = \widetilde{\mathbf{S}}^* \mathbf{V}^T$. Here we focus on fixed-orientation source models, but the model can also be easily extended to the free-orientation formulation as proposed in [4,11].

The above derivations are independent of the selected basis $\mathbf{V}$, but a compact representation of the signals can substantially reduce computation. We choose to use the K largest SVD components of the measurements, which can often compactly capture the time-varying frequency content and differences in source signals from different regions.

From $\ell_1\ell_2$-Norm Regularizer to SOCP. We cannot directly apply gradient based methods to the optimization problem specified by Eq. (4) since the $\ell_1\ell_2$-norm penalty term is not differentiable at zero. However, Eq. (4) can be reduced to the second-order cone programming (SOCP) problem [1]:

$$< \widetilde{\mathbf{S}}^*, q^*, z^*, \mathbf{w}^*, \mathbf{r}^* >= \arg\min_{<\widetilde{\mathbf{S}}, q, z, \mathbf{w}, \mathbf{r}>} (q + \lambda z) \tag{5}$$

$$\text{s.t.} \quad \|\widetilde{\mathbf{y}}_k - \mathbf{A}\widetilde{\mathbf{s}}_k\|_{\ell_2}^2 \le w_k \quad \forall\, k = 1, \cdots, K \tag{6}$$

$$\sum_{k=1}^{K} w_k \le q; \quad \sqrt{\sum_{k=1}^{K} \widetilde{s}_{nk}^2} \le r_n \quad \forall\, n = 1, \cdots, N; \quad \sum_{n=1}^{N} r_n \le z \tag{7}$$

The conversion introduces new variables, q, z, $\{w_k\}_{k=1}^K$, and $\{r_n\}_{n=1}^N$. w_k is an upper bound on the discrepancy between the measurements and the signals induced by the estimated sources along $\mathbf{v}_k$. q is an upper bound on all w_k's. r_n is an upper bound on the activation strength for location n. z is an upper bound on the ℓ_1-norm of the activation strength of all N locations. At the minimum, the inequality constraints in Eq. (6-7) are satisfied with equality; otherwise, the objective function can be further reduced.

SOCP is a convex optimization problem. Linear programs and quadratically constrained quadratic programs are subsets of the SOCP problems. Furthermore, the SOCP problem is a special case of a semi-definite program, and therefore SOCP can be solved efficiently using the interior-point method [1].

3   Results

Due to the lack of ground truth in real experiments, we first study the behavior of the method and its sensitivity to noise and parameter settings using simulated data. We then compare the method to standard inverse solvers using real MEG data from a somatosensory study.

3.1   Simulation Studies

To simulate MEG measurements, we created active vertices A, B, and C (Fig. 1 top) on the cortical sheet at source spacing of 20 mm, with current source orientation along the normal to the cortex. The simulated time courses (black, Fig. 1 bottom) exhibit temporal characteristics similar to those of the auditory evoked
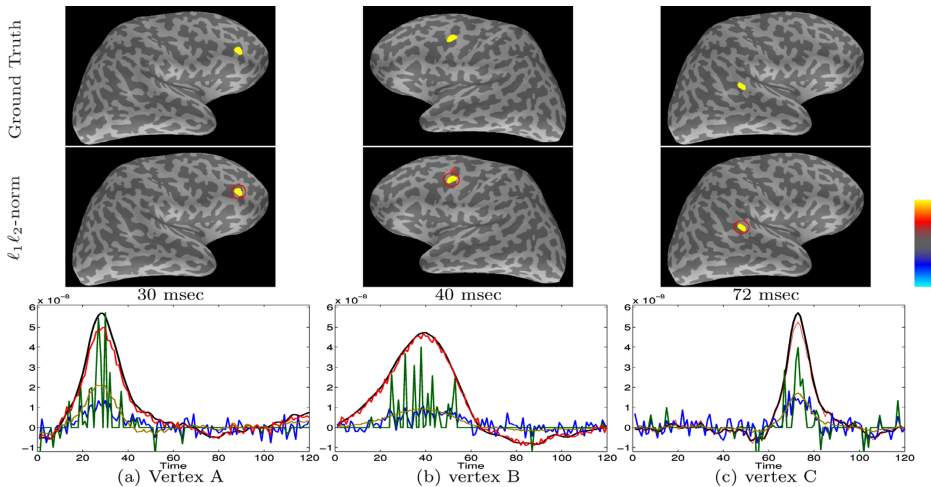
Fig. 1. Simulated study. Top: ground truth activation maps at peak response time for three sources. Middle: spatial maps estimated using $\ell_1\ell_2$-norm inverse solver for the same time points. Bottom: time courses from the three active vertices: ground truth (black), MNE (blue), MCE (green), VESTAL (yellow), and $\ell_1\ell_2$-norm (red).

responses, but with additional temporal translation and scaling. The source signals at vertices A and B show activation during overlapping time intervals, which makes the inverse problem difficult. For the forward calculations, we employed the sensor configuration of the 306-channel MEG system used in our real studies and added Gaussian noise to the signals. The resulting signals have an SNR of 3 dB, within the typically SNR range of real MEG data.

Fig. 1 shows the inverse solutions at three time frames obtained by the proposed method using three basis functions and $\lambda = 10^9$. These parameters were selected based on validation experiments presented later in this section. The red curves in Fig. 1 correspond to source signals estimated by the method. The resulting spatial maps and source time courses match well with the ground truth.

Comparison with MNE, MCE, and VESTAL. We compared the proposed method with the standard MNE, MCE, and VESTAL (Fig. 1 bottom). The estimates from the standard MNE are smaller and more diffuse when compared with the simulated signals. The estimated time courses from MCE exhibit "spiky" discontinuities due to the solver's sensitivity to noise. Projecting MCE results into the signal subspace spanned by $\mathbf{V}$ (VESTAL) removes the discontinuities, but the estimation accuracy is significantly worse than that of $\ell_1\ell_2$-norm since the two-step procedure cannot fully compensate for the errors in the original MCE solutions.

Sensitivity to Noise, Basis Selection, and Regularization. To examine the sensitivity of the method to noise and basis selection, we computed inverse solutions for 100 independently generated data sets for each noise setting (SNR 1-8 dB) and basis selection cutoff (K=1-6). Fig. 2a shows the relative mean
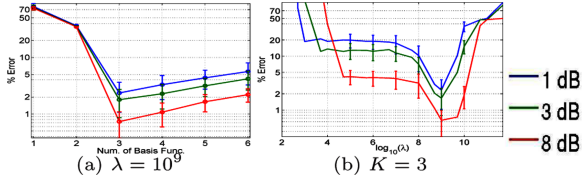
Fig. 2. Relative mean squares error for different values of basis size K and regularization parameter λ under three different SNR settings. The error bars close to the bottom of the figures appear large due to the logarithmic scale.

squares error[1] of the $\ell_1\ell_2$-norm inverse solutions. The method achieves the best performance for $K{=}3$ basis functions. If the chosen number of basis functions is too high, some basis functions represent noise, resulting in slight degradation of the result quality as reflected by the gentle slope on the logarithmic scale. Including too few basis functions leads to significant signal loss: the solver fails to recover the missing signals.

We also investigated the method's sensitivity to the value of the regularization parameter λ. Large λ corresponds to a high penalty on the strength of the current sources; small λ emphasizes the data fidelity term. Due to whitening, the first term in Eq. (4) is on the order of MK ($M \sim 10^2$). For an activated vertex in our experiment, $\widetilde{s}_{nk}$ is on the order of 10^{-8}. Hence, $|\widetilde{\mathbf{S}}|_{\ell_1}^{\ell_2}$ is approximately $10^{-7}K$. Therefore, $\lambda = 10^9$ balances between the data fidelity and the regularization terms in Eq. (4). Fig. 2b confirms that λ around 10^9 provides the most accurate reconstruction results. The regularization shows no effect for $\lambda < 10^3$; when $\lambda > 10^{10}$, the data fidelity term is effectively ignored in the optimization process. For $\lambda = 10^9$, the standard deviation of the 100 simulated data sets is less than 1%. In the experiments using real MEG data, we set $\lambda = 10^9$.

3.2   Real MEG Experiments

We compared the performance of the proposed solver to MNE and dipole fitting using real MEG data from median-nerve stimulation experiments. The measurements were acquired using a 306-channel Neuromag VectorView MEG system. Anatomical images, collected with a 1.5 T scanner, were used to construct the source space and the forward model [3,8].

In this study, the median nerve was stimulated at the left wrist according to an event-related protocol, with a random inter-stimulus-interval ranging from 1.5 to 2 seconds. A 200-msec baseline before the stimulus was used to estimate the noise covariance matrix. The average signal, computed from approximately 300 trials, was used as the input to the inverse solver.

The median-nerve stimulation activates a complex cortical network [9]. The first activation of the contralateral primary somatosensory cortex (cSI) peaks around

---

[1]   Defined as $\dfrac{\|\text{reconstructed signals–ground truth signals}\|_{\ell_2}^2}{\|\text{ground truth signals}\|_{\ell_2}^2}$.
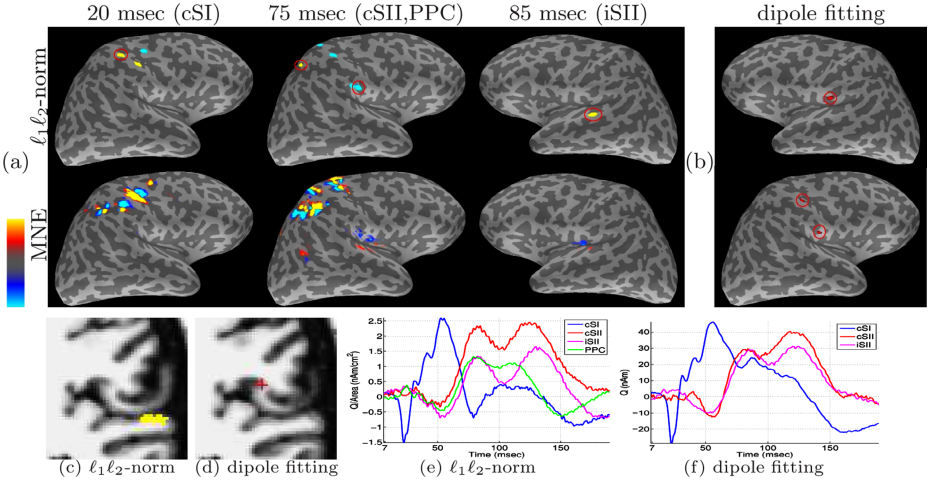
Fig. 3. Real MEG study. (a) Significance statistics of the $\ell_1\ell_2$-norm inverse solver and MNE for the median-nerve experiment. Hot/cold colors correspond to outward/inward current flow. The most active areas in the $\ell_1\ell_2$-norm solutions are highlined. (b) Dipole fitting with three sources. (c-d) Zoomed-in coronal slices for the detected iSII activations. (e-f) Reconstructed time courses for the highlined areas in (a-b).

20 msec and continues over 100 msec; then the secondary somatosensory cortex (SII) activates bilaterally between 70 and 200 msec. The posterior parietal cortex (PPC), located on the wall of the post-central sulcus, activates at 70-110 msec.

Employing six basis functions, which explain 80% variance of the data, is sufficient to capture the responses. We adopted a standard multi-resolution scheme, corresponding to source spacing of 20 and 10 mm. The efficient interior-point optimizer [15] allows us to estimate statistical significance of the resulting activations via permutation testing [14]. We controlled the false discovery rate (FDR) [5] at 0.05, computed from 5000 permutations (run time of 60 seconds per permutation). We also compared our results to MNE and dipole fitting computed using standard software packages [8,18]. In practice, experts often interpret MNE through its statistics, the dynamic statistical parameter maps (dSPM) [2], with thresholds adjusted by an expert. For the purpose of comparison, we selected the threshold for dSPM so that all four regions of interest, cSI, cSII, iSII, and PPC, were included in the dSPM.

Fig. 3a presents the activation maps obtained using $\ell_1\ell_2$-norm and MNE for one subject in the study. At 20 msec, the $\ell_1\ell_2$-norm pinpoints cSI on the postal wall of the central sulcus. MNE produces a more diffuse solution leading to false positives in the post-central sulcus. At 75 msec, both MNE and $\ell_1\ell_2$-norm capture signals from cSII. $\ell_1\ell_2$-norm successfully localizes PPC at the post-central sulcus; the location of PPC is ambiguous in the MNE results. The $\ell_1\ell_2$-norm detects iSII, but places it at the superior temporal lobe instead of the inferior parietal lobe. As shown in the volumetric display (Fig. 3cd), these two regions are very close, making the inverse problem particularly challenging. MNE

also detects weak iSII signals; the location is ambiguously spread between the iSII region and the superior temporal lobe.

Dipole fitting results, shown in Fig. 3b, did not correctly localize PPC in this experiment because PPC is very close to cSI. The locations for cSI and cSII estimated by our solver match with the dipole fitting results. The correct localization of iSII using dipole fitting required manual intervention in selecting appropriate channels, in contrast to the automatic $\ell_1\ell_2$-norm solver.

Our method yields stable time courses that capture the main deflections precisely (Fig. 3e). The reconstructed time courses are similar to those estimated through dipole fitting (Fig. 3f), except for the cSI activation between 70 and 150 msec. This is due to the PPC activation missed by dipole fitting. This comparison demonstrates the ability of the $\ell_1\ell_2$-norm regularization to achieve high-quality reconstructions of source signals.

4   Conclusions

The proposed model takes advantage of the relatively smooth nature of the underlying EEG/MEG source signals via an ℓ_2-norm regularizer on the projection coefficients of the temporal basis functions. The $\ell_1\ell_2$-norm inverse solver is formulated as an SOCP problem. Performing reconstruction in the signal subspace while jointly considering the coefficients for all selected basis functions leads to stable estimates with a reduced number of false positives as confirmed by our experiments using simulated and real MEG data.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Technical Report. RRR Report number 51-2001, RUTCOR, Rutgers University (2001)
2. Dale, A.M., Sereno, M.: Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J. Cogn. Neurosci. 5, 162–176 (1993)
3. Dale, A.M., et al.: Cortical surface-based analysis: I. Segmentation and surface reconstruction. NeuroImage 9, 179–194 (1999)
4. Ding, L., He, B.: Sparse source imaging in electroencephalography with accurate field modeling. Human Brain Mapping (in press)
5. Genovese, C.R., et al.: Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15, 870–878 (2002)
6. Geva, A.B.: Bioelectric sources estimation using spatio-temporal matching pursuit. Applied Sig. Proc. 5, 195–208 (1998)

7. Hämäläinen, M.S., et al.: Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 65, 413–497 (1993)
8. Hämäläinen, M.S.: MNE software user's guide. NMR Center, Mass General Hospital
9. Hari, R., Forss, N.: Magnetoencephalography in the study of human somatosensory cortical processing. Philos. Trans. R. Soc. Lond. B 354, 1145–1154 (1999)
10. Huang, M.X., et al.: Vector-based spatial-temporal minimum L1-norm solution for MEG. NeuroImage 31, 1025–1037 (2006)
11. Lin, F.-H., et al.: Distributed current estimates using cortical orientation constraints. Human Brain Mapping 27, 1–13 (2006)
12. Malioutov, M., et al.: A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Sig. Proc. 53, 3010–3022 (2005)
13. Mosher, J.C., et al.: Multiple dipole modeling and localization from spatiotemporal MEG data. IEEE Trans. Biomed. Engr. 39, 541–557 (1992)
14. Pantazis, D., et al.: A comparison of random field theory and permutation methods for the statistical analysis of MEG data. NeuroImage 25, 383–394 (2005)
15. Sturm, J.F., et al.: Self-Dual-Minimization (SeDuMi): solver for optimization problems over symmetric cones, http://sedumi.mcmaster.ca/
16. Trujillo-Barreto, N., et al.: Bayesian M/EEG source reconstruction with spatio-temporal priors. NeuroImage 39, 318–335 (2008)
17. Uutela, K., et al.: Visualization of magnetoencephalographic data using minimum current estimates. NeuroImage 10, 173–180 (1999)
18. Source modeling software (xfit), Elekta-Neuromag Oy, Helsinki, Finland