MEG/EEG source reconstruction based on Gabor thresholding in the source space

Daniel Strohmeier[1], Alexandre Gramfort[234], Jens Haueisen[156], Matti Hämäläinen[2] and Matthieu Kowalski[7]
[1]Institute of Biomedical Engineering and Informatics, Ilmenau University of Technology, Ilmenau, Germany
[2]Martinos Center for Biomedical Imaging, Department of Radiology, MGH, Harvard Medical School, Boston, MA
[3]INRIA Parietal Project Team, Saclay, France
[4]LNAO/NeuroSpin, CEA Saclay, Bat. 145, Gif-sur-Yvette, France
[5]Biomagnetic Center, Department of Neurology, University Hospital Jena, Jena, Germany
[6]Department of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
[7]Laboratoire des Signaux et Systèmes (L2S), Université Paris-Sud, Orsay, France
Email: daniel.strohmeier@tu-ilmenau.de, gramfort@nmr.mgh.harvard.edu, jens.haueisen@tu-ilmenau.de,
msh@nmr.mgh.harvard.edu, matthieu.kowalski@lss.supelec.fr

***Abstract*—Thanks to their high temporal resolution, source reconstruction based on Magnetoencephalography (MEG) and/or Electroencephalography (EEG) is an important tool for noninvasive functional brain imaging. Since the MEG/EEG inverse problem is ill-posed, inverse solvers employ priors on the sources. While priors are generally applied in the time domain, the time-frequency (TF) characteristics of brain signals are rarely employed as a spatio-temporal prior. In this work, we present an inverse solver which employs a structured sparse prior formed by the sum of ℓ_{21} and ℓ_1 norms on the coefficients of the Gabor TF decomposition of the source activations. The resulting convex optimization problem is solved using a first-order scheme based on proximal operators. We provide empirical evidence based on EEG simulations that the proposed method is able to recover neural activations that are spatially sparse, temporally smooth and non-stationary. We compare our approach to alternative solvers based also on convex sparse priors, and demonstrate the benefit of promoting sparse Gabor decompositions via a mathematically principled iterative thresholding procedure.**

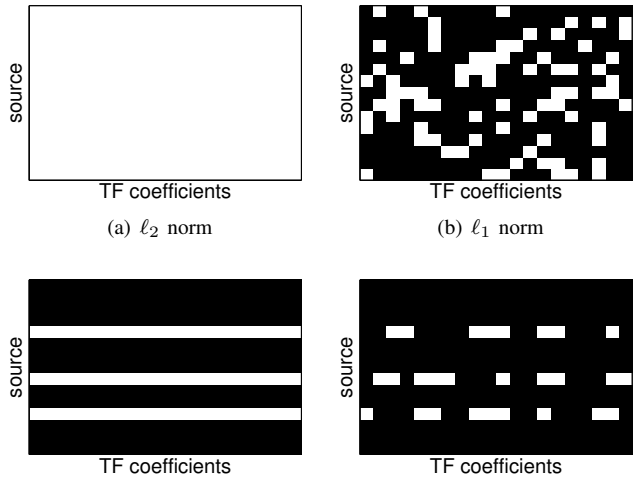
I. Introduction

For both neuroscience research and clinical diagnosis, noninvasive measurements of the brain function are of major importance. Solving the bioelectromagnetic inverse problem of Magnetoencephalography (MEG) and/or Electroencephalography (EEG) (collectively M/EEG) is a way to achieve this. Due to the ill-posed nature of the M/EEG inverse problem, it is mandatory to impose suitable constraints on the solution. These constraints, also called *priors*, should reflect neurophysiologically motivated assumptions on the sources. Priors are typically related to the number of active sources, their amplitudes or their spatio-temporal characteristics.

During the past few years, several source reconstruction methods based on spatially sparse priors have been introduced, sparsity-inducing Bayesian formulations [1]–[3], and convex mixed norms [4]–[6]. These approaches assume that only a limited number of focal sources is involved in a specific cognitive task. To do so, these imaging methods rely on sparsity inducing priors in the time domain. In contrast, the inverse solvers presented in [7], [8] impose structured sparsity on the time-frequency (TF) decomposition of the source signals. As will be illustrated in this work, this allows to perform source reconstruction and TF analysis in the source space simultaneously. While the TF decomposition in the sensor space is commonly applied as a preprocessing step to identify TF components, which can then be localized is a second step [9]–[11], the TF characteristics of the source signals are rarely employed directly as a prior to regularize the inverse problem. The relevant TF components are typically identified by looking at high coefficients in Wavelet or Gabor decompositions and denoising is then naturally done by thresholding [1]. This results in smooth time series which can also be obtained with an ℓ_1 regularization [12]. In order to obtain spatially sparse source estimates, Ou et al. [5] apply a ℓ_{21} mixed norm prior, which promotes a dense block row structure on the matrix of the source estimates. The reconstructed sources are spatially sparse, but their activation time series are likely to be noisy since they are not regularized. In order to have spatially sparse and temporally smooth source estimates, we present a composite prior combining ℓ_{21} mixed norm and ℓ_1 norm priors that is applied on the TF coefficients of the source signals. Such a prior promotes a block row structure with intra-row sparsity, see Fig. 1.

In a previous work [13], we described the mathematical framework which applies structured sparsity of the TF decomposition coefficients to recover the spatial sparsity, temporal smoothness and non-stationarity of neural activations. In this contribution, we start by presenting some aspects of the convex optimization procedure detailed in [13] and then extend this work by providing empirical evidence based on EEG simulations using isolated current dipoles and dipole patches that our method is able to recover the location as well as the TF characteristics of neuronal sources simultaneously.

This work was supported by a research grant offered by the German Research Foundation (Ha 2899/8-1).



(a) ℓ_2 norm (b) ℓ_1 norm

(c) ℓ_{21} mixed norm (d) ℓ_{21} mixed norm + ℓ_1 norm

Fig. 1. Sparsity patterns promoted by different priors in the TF domain: (a) ℓ_2 norm: no non-zero TF coefficients; (b) ℓ_1 norm: scattered and unstructured non-zero TF coefficients; (c) ℓ_{21} mixed norm: block row structured non-zero TF coefficients; and (d) ℓ_{21} mixed norm + ℓ_1 norm: block row structured non-zero TF coefficients with intra-row sparsity. In all panels, non-zero coefficients are shown in white.

We compare our method to other inverse solvers based on convex sparse priors and show that the thresholding of TF decomposition coefficients in the source space as performed by our structured sparse prior improves the solution of the M/EEG inverse problem.

II. Materials and Methods

A. Notation

We denote vectors with bold letters, $\mathbf{a} \in \mathbb{R}^N$ (resp. $\mathbb{C}^N$) and matrices with capital bold letters, $\mathbf{A} \in \mathbb{R}^{N\times M}$ (resp. $\mathbb{C}^{N\times M}$). $\mathbf{a}_i$ identifies the ith vector element and $\mathbf{A}_{i,j}$ the matrix element with row index i and column index j. We indicate the ℓ_1 norm as $\|\mathbf{A}\|_1 = \sum_{i,j=1}^{N} |\mathbf{A}_{i,j}|$, and the ℓ_{21} mixed norm as $\|\mathbf{A}\|_{21} = \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{N} |\mathbf{A}_{i,j}|^2}$. $\|\mathbf{A}\|_{\text{Fro}}$ denotes the Frobenius norm with $\|\mathbf{A}\|_{\text{Fro}}^2 = \sum_{i,j=1}^{N} |\mathbf{A}_{i,j}|^2$. $\mathbf{A}^T$ and $\mathbf{A}^{\mathcal{H}}$ denote a matrix transpose and a Hermitian transpose, respectively.

B. The inverse problem with sparse TF decompositions

Given a lead field matrix $\mathbf{G} \in \mathbb{R}^{N\times P}$, where N is the number of sensors and P the number of sources, the measurements $\mathbf{M} \in \mathbb{R}^{N\times T}$ (T number of time instants) are related to the source amplitudes $\mathbf{X} \in \mathbb{R}^{P\times T}$ by $\mathbf{M} = \mathbf{G}\mathbf{X}$. By applying a linear TF decomposition on $\mathbf{X}$ based on a dictionary of TF atoms $\mathbf{\Phi}^{\mathcal{H}} \in \mathbb{C}^{K\times T}$, the forward model can be expressed as:

$$\mathbf{M} = \mathbf{G}\mathbf{X} + \mathbf{E} = \mathbf{G}\mathbf{Z}\mathbf{\Phi}^{\mathcal{H}} + \mathbf{E} \ , \tag{1}$$

where $\mathbf{Z} \in \mathbb{C}^{P\times K}$ is the coefficient matrix of the TF decomposition, and $\mathbf{E}$ is an additive white noise, $\mathbf{E} \sim \mathcal{N}(0, \lambda\mathbf{I})$.

In order to solve the ill-posed M/EEG inverse problem, priors have to be imposed on the solution. As we assume that only a small number of focal sources with smooth activation time series are active during a cognitive task, we apply the composite prior proposed in [7] and [13]. It is formed by the sum of the ℓ_1 and the ℓ_{21} norms (cf. Equation (2)) applied on the coefficients of the TF decomposition $\mathbf{Z}$. While the ℓ_{21} mixed norm promotes spatially sparse solutions, the ℓ_1 norm denoises the reconstructed source signals by limiting the number of TF atoms used for the reconstruction. The maximum a posteriori estimate is obtained by solving:

$$\begin{aligned} \mathbf{Z}^\star &= \arg\min_{\mathbf{Z}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{Z}\mathbf{\Phi}^{\mathcal{H}}\|_{\text{Fro}}^2 + \lambda\Omega(\mathbf{Z}) \\ \text{with } \Omega(\mathbf{Z}) &= \rho\|\mathbf{Z}\|_1 + (1-\rho)\|\mathbf{Z}\|_{21} \end{aligned} \tag{2}$$

where $\lambda \in \mathbb{R}^+$ is the regularization parameter and $\rho \in [0\ \ 1]$ is the trade-off parameter between the ℓ_{21} and the ℓ_1 norms. Note that a Gabor transform as described here corresponds to a discrete short time Fourier transform (STFT) with a Gabor window.

For M/EEG inverse problems without orientation constraints, the composite prior needs to be adapted. We group the different orientations in a common ℓ_2 norm such as in [5]. Hence, assuming that each source is indexed by a spatial location i and an orientation $o \in \{1, 2, 3\}$ and the TF atoms in the TF dictionary are indexed by k, the ℓ_1 and ℓ_{21} norms read:

$$\begin{aligned} \|\mathbf{Z}\|_1 &= \sum_{ik} \sqrt{\sum_{o=1}^{3} |\mathbf{Z}_{(i,o),k}|^2} \\ \|\mathbf{Z}\|_{21} &= \sum_{i} \sqrt{\sum_{k}\sum_{o=1}^{3} |\mathbf{Z}_{(i,o),k}|^2} \ , \end{aligned} \tag{3}$$

1) Implementation: The sparse regression problem in Eq. (2) is a convex optimization problem. This implies that the optimization cannot be trapped in local minima and that improper initialization does not cause problems. The structure and convexity of the cost function enables us to apply a first-order scheme, known as FISTA, based on proximal operators [14]. As detailed in [13], the proximal operator of the composite prior $\mathbf{Z} = \text{prox}_{\lambda(\rho\|.\|_1 + (1-\rho)\|.\|_{21})}(\mathbf{Y}) \in \mathbb{C}^{P\times K}$ is given by

$$\begin{aligned} Z_{p,k} = &\frac{Y_{p,k}}{|Y_{p,k}|}\left(|Y_{p,k}| - \lambda\rho\right)^+ \cdot \\ &\cdot \left(1 - \frac{\lambda(1-\rho)}{\sqrt{\sum_k ((|Y_{p,k}| - \lambda\rho)^+)^2}}\right)^+ \ , \end{aligned} \tag{4}$$

where p and k index the TF decomposition matrix, $(x)^+ = \max(x, 0)$ for $x \in \mathbb{R}$, and by definition $\frac{0}{0} = 0$. This proximal operator is equivalent to applying the ℓ_1 and the

ℓ_{21} proximal operators successively [15], [16]. The pseudo-code for the optimization process is given in Algorithm 1. Note that contrary to previous contributions like [7] that use a truncated Newton method, which should be only applied to differentiable problems, this contribution provides a principled method to minimize the cost function in Equation (2).

Algorithm 1 FISTA with TF dictionaries

Require: Measurements $\mathbf{M}$, lead field matrix $\mathbf{G}$, regularization parameter $\lambda > 0$ and I the number of iterations.

Ensure: $\mathbf{Z}^\star$

1: Aux. variables: $\mathbf{Y}$ and $\mathbf{Z}_o \in \mathbb{R}^{P\times K}$, and τ and $\tau_o \in \mathbb{R}$.
2: Estimate the Lipschitz constant $\mathcal{L}$ with the power iteration method.
3: $\mathbf{Y} = \mathbf{Z}^\star = \mathbf{Z}$, $\tau = 1$, $0 < \mu < \mathcal{L}^{-1}$
4: **for** $i = 1$ to I **do**
5: $\quad \mathbf{Z}_o = \mathbf{Z}^\star$
6: $\quad \mathbf{Z}^\star = prox_{\mu\lambda\Omega}\left(Y + \mu\mathbf{G}^T(\mathbf{M} - \mathbf{G}\mathbf{Y}\mathbf{\Phi}^{\mathcal{H}})\mathbf{\Phi}\right)$
7: $\quad \tau_o = \tau$
8: $\quad \tau = \frac{1+\sqrt{1+4\tau^2}}{2}$
9: $\quad \mathbf{Y} = \mathbf{Z}^\star + \frac{\tau_o - 1}{\tau}(\mathbf{Z}^\star - \mathbf{Z}_o)$
10: **end for**

We apply a Gabor transform with a Gabor frame based on the LTFAT toolbox [17] to compute the TF decomposition. This allows to recover stationary as well as non-stationary source signals $\mathbf{x} \in \mathbb{R}^T$ from their Gabor coefficients $(\langle \mathbf{x}, \phi_{k,f} \rangle) = \mathbf{\Phi}^{\mathcal{H}}\mathbf{x}$. To reduce the computation time, we compute the Gabor transform only for active sources, i.e., sources with non-zero TF decomposition coefficients.

C. Simulation setup

In order to evaluate the performance of the sparse TF inverse algorithm, we performed simulations with isolated current dipoles and dipole patches as sources. The simulation setup consisted of a three shell boundary element model (BEM) which was computed with OpenMEEG [18] using skin, skull, and CSF surfaces segmented from a MRI scan using MNE[1]. The conductivities were chosen according to [19]. The source space was built from 5000 triplets of orthogonal current dipoles placed on the boundary between white and gray matter, which was segmented using FreeSurfer [20]. We used 128 EEG channels for spatially sampling of the simulated data. Finally, a common-average reference transform was applied to the leadfield matrix.

D. Simulation

1) Isolated current dipoles : We modeled 2 current dipoles located in Brodmann areas 3b and 1. The orientations are set perpendicular to the white matter surface. The geodesic distance between the simulated dipoles was $15.7\,\mathrm{mm}$, and the angle between dipole moment vectors $101.4°$. The activation curves were modeled using synthetic, non-stationary chirp signals with an exponential decay (cf. Fig. 2). To evaluate the performance with different signal to noise ratios (SNR) (-6db, 0dB, and 6dB), which we define here as $20\log_{10}(\|M\|_{\mathrm{Fro}}/\|E\|_{\mathrm{Fro}})$, white noise was added to the forward simulation. By adding white noise, we assume that the data were spatially whitened based on a noise covariance matrix estimated during baseline.

We applied the sparse TF inverse approach with orientation constraint and a tight Gabor frame with time shift $k = 4\,\mathrm{samples}$ and frequency shift $f = 4\,\mathrm{samples}$. In order to compensate the depth bias, the columns of the lead field matrix were normalized before solving the inverse problem. We compare our solution to inverse solvers based on ℓ_{21} mixed norm priors [5] and a ℓ_1 norm priors, respectively. We imposed the ℓ_1 norm prior both on the source activations in the time domain and on the coefficients of the TF decomposition. In contrast, as the ℓ_{21} mixed norm prior in the time and TF domain are mathematically equivalent when using tight Gabor frames, we chose to apply it in the time domain.

Although solving the inverse problem based on the composite prior is tractable for high dimensional data sets (see [13] for a real MEG data set), we reduced the source space with regard to the complexity of the simulations in order to decrease the computation time. Faster optimization relies on an active set approach [21] and is further discussed in [13]. As the application and evaluation of this method is beyond the scope of this contribution, we simulated the result of the active set approach by restricting the source space to patches with geodesic radii of $50\,\mathrm{mm}$, which include the simulated sources. The restricted source space was applied for all inverse solvers making results comparable.

2) Patches: The presented inverse solver is based on the assumption that the neural activity can be modeled by spatially sparse isolated current-dipole sources. In order to analyze the performance of the proposed method in the case that this model assumption is violated, we performed a simulation study with spatially extended sources. It is based on three source patches with geodesic radii of $10\,\mathrm{mm}$ located in the left and right auditory cortices and the left motor cortex. A similar simulation is described in [22]. The sources were perpendicular to the white matter surface. The source activations are simulated as non-stationary, synthetic chirp signals with an exponential decay. The simulated patches as well as their activation curves are displayed in Fig. 4. Different SNRs ($-6\,\mathrm{dB}$, $0\,\mathrm{dB}$, and $6\,\mathrm{dB}$) are generated by adding white noise. The inverse problem was solved without orientation constraints. The regularization parameter λ and the trade-off parameter ρ were set to $\lambda = 6 \cdot 10^{-5}$ and $\rho = 0.1$ respectively.

[1] http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofMNE.php

III. RESULTS

A. Isolated current dipoles

To compare the different priors, we calculate the root mean square error (RMSE) in the source space defined as $\text{RMSE} = \|\mathbf{X}_{\text{sim}} - \mathbf{X}_{\text{est}}\|^2_{\text{Fro}}$. For each prior and SNR, we analyze the RMSE as a function of the regularization parameter λ and determined the minimum. λ was chosen from 10^{-6} to 10^{-3} and ρ fixed to 0.1 for the composite prior. The results are presented in Table I. All values are normalized to the RMSE of the composite prior.

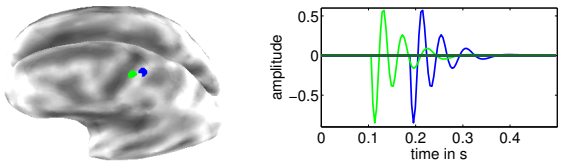
TABLE I
NORMALIZED RMSE IN THE SOURCE SPACE FOR DIFFERENT SPARSITY INDUCING PRIORS. THE BEST RMSE IS MARKED IN BOLD

prior	normalized RMSE in the source space		
	SNR $= -6$ dB	SNR $= 0$ dB	SNR $= 6$ dB
$\ell_{21} + \ell_1$ norm (TF domain)	**1**	**1**	**1**
ℓ_{21} norm (time domain)	1.66	1.32	1.22
ℓ_1 norm (time domain)	2.10	1.71	1.77
ℓ_1 norm (TF domain)	2.12	1.86	2.26

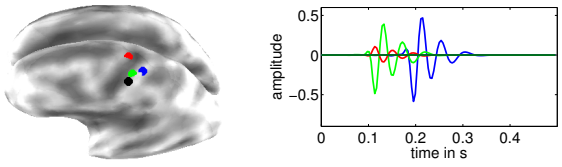
The ℓ_1 norm based approaches led to the highest RMSE followed by the ℓ_{21} mixed norm based method. The minimal RMSE was achieved by the composite prior based technique. Fig. 2 shows the simulated as well as the estimated source signals for SNR $= 0$ dB. Compared to the solutions obtained with the alternative priors, the composite prior led to the smoothest and spatially sparsest solution (cf. Fig. 2). This is particularly obvious in periods dominated by noise where the composite prior based approach benefits from the denoising based on Gabor TF thresholding. Besides the two simulated dipoles, at least two erroneous sources, located close to the simulated ones, are reconstructed with all priors which are located close to the simulated sources. Their amplitude is significantly smaller than the true sources and the time courses correlate primarily with the deep tangential source. This can be attributed to the applied depth compensation and the lower sensitivity of EEG to tangential sources.

Due to the TF resolution of the applied Gabor frame and the shape of the covered Gabor atoms, the composite prior based inverse solver was not able to reconstruct perfectly the abrupt changes at the signals onsets. This resulted in a transient oscillation at the signals onsets (cf. Fig. 2(b)). By decreasing the length of the Gabor atoms and increasing the TF resolution, the reconstruction of abrupt changes can be improved.

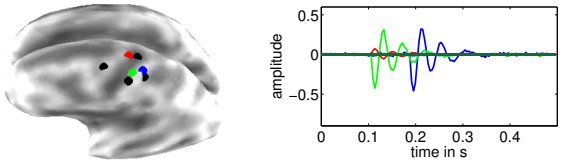
Besides the source location and activation time series, the proposed inverse solver also provides the time-frequency distribution (TFD) of the source signals. This is illustrated in Fig. 3. Thanks to the structured sparsity imposed by the composite prior, small TFD coefficients are discarded and the TF structure is well identified. The non-stationarity and the chirp appear clearly. Vertical components in the TFDs, at 100 ms and 200 ms respectively, reflect the steep slope at the onset of the chirps. In contrast, due to the unstructured sparsity pattern, the TFDs based on the ℓ_1 norm prior do not allow to identify correctly the TF characteristics. Note that the Gabor dictionary does not contain any chirp like atom, which suggests that the solver does not rely on a very accurate a priori knowledge of the temporal dynamics of the sources.



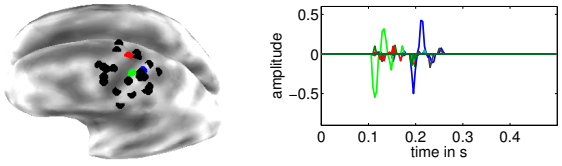
(a) simulation



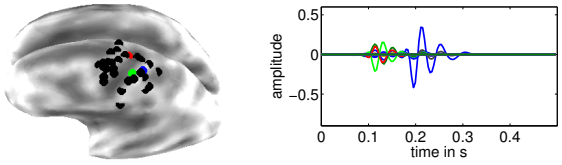
(b) $\ell_{21} + \ell_1$ in the TF domain



(c) ℓ_{21} in the time domain

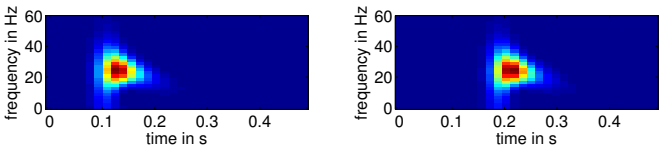


(d) ℓ_1 in the time domain

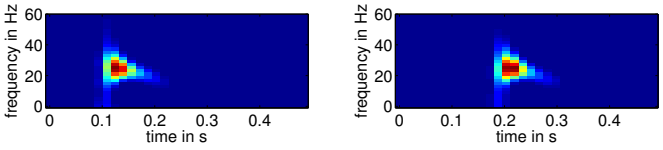


(e) ℓ_1 in the TF domain

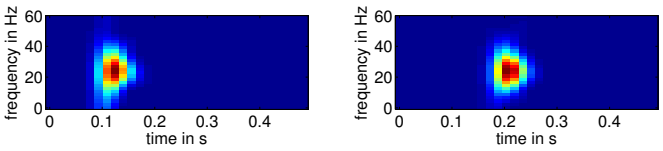
Fig. 2. Simulated and estimated source signals based on sparse priors. The simulated dipoles, which are reconstucted in Brodmann area 3b and 1, are marked in green and blue, the erroneous dipoles in black. The erroneous dipole with the highest amplitude is colored in red. Note the different number of erroneous dipoles for the different priors.



(a) simulation

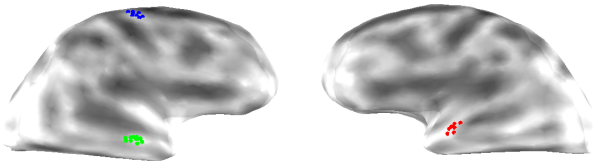


(b) $\ell_{21} + \ell_1$ in the TF domain

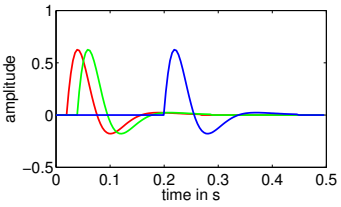


(c) ℓ_1 in the time domain

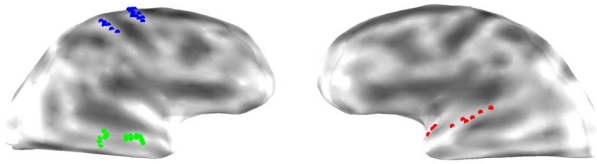
Fig. 3. Simulated and estimated source signals in Brodmann 3b (left column) and 1 (right column) in the TF domain for SNR = 0 dB. The TFDs of the simulated signals and the ℓ_1 norm prior solutions are computed using a Gabor transform. The TFDs of the composite prior solution are obtained from the inverse solution in the TF domain.



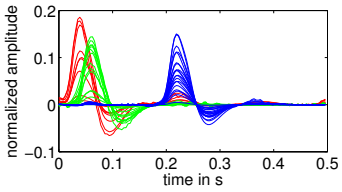
(a) simulated extended sources



(b) simulated source activations



(c) reconstructed dipolar sources



(d) reconstructed source activations

Fig. 4. Results of the patch simulation study for SNR = 0 dB. The reconstructed source activations were clustered in three groups based on the correlation of the simulated and reconstructed source activations. The activations within each group were normalized to have a unit Frobenius norm.

B. Patches

Fig. 4 shows the simulated and the reconstructed source activations for SNR $= 0$ dB. The final active set contains 45 active dipoles. For visualization purposes, the dipoles were clustered into three groups based on the correlation of the simulated and reconstructed source activations. In order to compare the contribution of each group to the source estimates, source amplitudes within each group were normalized to have a unit Frobenius norm. Both the smoothness and non-stationarity of the simulated source signals is recovered. The sites of the estimated sources agree with the simulation.

IV. Discussion and Conclusion

In this work, we presented an M/EEG inverse solver based on a composite prior, formed by the sum of a ℓ_{21} and a ℓ_1 norm priors, which is imposed on the TF decomposition of the source activations. We briefly described the mathematical framework, which is applied to formulate the convex optimization problem (for details see [13]), and provided further information on the implementation. We obtained empirical evidence by using isolated current dipoles and patches in EEG simulations that our approach is able to reconstruct the spatial sparsity, temporal smoothness, and non-stationarity of neural activations. We showed that the solver provides in one step, the locations and TF characteristics of the sources. The simulations indicated that the presented method is more robust to noise than alternative sparse inverse solvers due to the Gabor thresholding in the source space. The results evidenced that the proposed inverse solver is a promising approach for M/EEG source analysis.

Further investigations are needed to study the impact of the choice of the time-frequency dictionary. To improve the signal reconstruction, a union of different Gabor dictionaries can be used to allow for different time-frequency resolutions. In addition, the application of TF decomposition algorithms, which are robust to outliers such as artifacts, is of interest.

Acknowledgments

The authors would like to thank Alexander Hunold for providing the segmented MRI data.

REFERENCES

[1] N. J. Trujillo-Barreto, E. Aubert-Vázquez, and W. D. Penny, "Bayesian M/EEG source reconstruction with spatio-temporal priors," *Neuroimage*, vol. 39, no. 1, pp. 318–35, jan 2008.

[2] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *Neuroimage*, vol. 39, no. 3, pp. 1104–20, feb 2008.

[3] D. Wipf and S. Nagarajan, "A unified Bayesian framework for MEG/EEG source imaging," *Neuroimage*, vol. 44, no. 3, pp. 947–966, feb 2009.

[4] S. Haufe, V. V. Nikulin, A. Ziehe, K.-R. Müller, and G. Nolte, "Combining sparsity and rotational invariance in EEG/MEG source reconstruction," *NeuroImage*, vol. 42, no. 2, pp. 726–38, aug 2008.

[5] W. Ou, M. Hämaläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, feb 2009.

[6] A. Gramfort, "Multi-condition M/EEG Inverse Modeling with Sparsity Assumptions: How to Estimate What Is Common and What Is Specific in Multiple Experimental Conditions," in *Proc. of the 17th International Conference on Biomagnetism Advances in Biomagnetism (Biomag2010)*, Dubrovnik, , Croatia, apr 2010, pp. 124–127.

[7] A. Polonsky and M. Zibulevsky, "MEG/EEG source localization using spatio-temporal sparse representations," in *Proc. of the 5th Intern. Conference on Independent Component Analysis and Blind Signal Separation (ICA 2004)*, vol. 3195, Granada, Spain, sep 2004, pp. 1001–1008.

[8] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624 – 638, mar 2006.

[9] A. B. Geva, "Spatio-temporal matching pursuit (SToMP) for multiple source estimation of evoked potentials," in *Proc. of the 19th Convention of Electrical and Electronics Engineers in Israel*, Jerusalem, Israel, apr 1996, pp. 113 –116.

[10] P. J. Durka, A. Matysiak, E. M. Montes, P. Valdés-Sosa, and K. J. Blinowska, "Multichannel matching pursuit and EEG inverse solutions," *Journal of Neuroscience Methods*, vol. 148, no. 1, pp. 49 – 59, oct 2005.

[11] D. Lelic, M. Gratkowski, M. Valeriani, L. Arendt-Nielsen, and A. M. Drewes, "Inverse modeling on decomposed electroencephalographic data: A way forward?" *Journal of Clinical Neurophysiology*, vol. 26, no. 4, pp. 227–235, aug 2009.

[12] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, may 1995.

[13] A. Gramfort, D. Strohmeier, J. Haueisen, M. Hämäläinen, and M. Kowalski, "Functional brain imaging with m/eeg using structured sparsity in time-frequency dictionaries," in *Information Processing in Medical Imaging (to appear)*, Monastery Irsee, Germany, jul 2011.

[14] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, mar 2009.

[15] R. Jenatton, J. Mairal, G. Obozinski, and F. Bach, "Proximal methods for sparse hierarchical dictionary learning," in *Proc. of the 27th Intern. Conference on Machine Learning (ICML '10)*, Haifa, Israel, jun 2010.

[16] M. Kowalski, "Sparse regression using mixed norms," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 303–324, nov 2009.

[17] P. Soendergard, B. Torrésani, and P. Balazs, "The linear time frequency toolbox," Technical University of Denmark, Tech. Rep., 2009.

[18] A. Gramfort, T. Papadopoulo, E. Olivi, and M. Clerc, "OpenMEEG: opensource software for quasistatic bioelectromagnetics," *BioMedical Engineering OnLine*, vol. 9, no. 1, p. 45, sep 2010.

[19] L. Geddes and L. Baker, "The specific resistance of biological material - a compendium of data for the biomedical engineer and physiologist," *Medical and Biological Engineering and Computing*, vol. 5, pp. 271–293, may 1967.

[20] A. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. segmentation and surface reconstruction," *NeuroImage*, vol. 9, no. 2, pp. 179 – 194, feb 1999.

[21] V. Roth and B. Fischer, "The group-lasso for generalized linear models: uniqueness of solutions and efficient algorithms," in *Proc. of the 25th International Conference on Machine learning (ICML '08)*, Helsinki, Finland, jul 2008, pp. 848–855.

[22] A. Bolstad, B. V. Veen, and R. Nowak, "Space-time event sparse penalization for magneto-/electroencephalography," *NeuroImage*, vol. 46, no. 4, pp. 1066–81, jul 2009.