Probabilistic M/EEG source imaging from sparse spatio-temporal event structure

Carsten Stahlhut[a], Hagai T. Attias[b], David Wipf[c],
Lars K. Hansen[a], and Srikantan S. Nagarajan[d]

[a] Technical University of Denmark, DTU Informatics, Lyngby, Denmark
[b]Convex Imaging, San Francisco, CA, USA
[c]Visual Computing Group, Microsoft Research Asia, Beijing, China
[d]UCSF, Biomagnetic Imaging Laboratory, San Francisco, CA, USA

Abstract. While MEG and EEG source imaging methods have to tackle a severely ill-posed problem their success can be stated as their ability to constrain the solutions using appropriate priors. In this paper we propose a hierarchical Bayesian model facilitating spatio-temporal patterns through the use of both spatial and temporal basis functions. We demonstrate the efficacy of the model on both artificial data and real EEG data.

Keywords: M/EEG, spatio-temporal patterns, variational Bayes

1 Introduction

We are interested in imaging of spatio-temporal brain events using magneto- and electro-encephaloghy (M/EEG), known as an inverse problem of estimating the current sources, $\mathbf{S}$, given the observed M/EEG signal, $\mathbf{Y}$. The forward relation between the measured M/EEG signal and the brain's current sources is linear

$$\mathbf{Y} = \mathbf{AS} + \mathbf{E}, \tag{1}$$

where the $N_c \times N_t$ matrix $\mathbf{E}$ holds assumed additive noise in a short time window of length N_t. The matrix $\mathbf{A} \in \Re^{N_c \times N_d}$, for which the rows are denoted as 'the lead fields' for the sensors and the columns as the 'forward fields' for the sources, is considered fixed in the present work. A reasonable imaging resolution, i.e., number of dipole/source locations is on the order of $N_d = 5,000 - 80,000$, which outnumbers the number of measured channels, $N_c = 32 - 300$, dramatically. Consequently, the inverse problem is severely ill-posed and success will depend on our ability to constrain the solutions using relevant priors. Multiple methods based in brain bio-physics are available for estimating $\mathbf{A}$, typically they lead to severely ill-conditioned matrices, which may lead to numerical instability of the problem. On top, imaging algorithms also need to address numerous sources of noise that interfere with the true signals, including thermal and neurophysiological noise as well as many sources of confounding quasi-static electrical fields. To overcome this both spatial and temporal assumptions have been pursued by

several authors. Among others, [13] has proposed to integrate structural priors in terms of spatial basis funcitons. Of more recent atempts [9] has demonstated spatial flexibility using spherical Gaussians with different spatial standard deviations and an $\ell_{1,2}$-term for promoting source density represented by a small number of spatial basis fields.

Similar, a number of papers have proposed temporal assumptions in order to promote temporal smoothness, e.g. by penalizing differences in neighboring time points such as [2]. In [4] temporal smoothness is incorporated by the second derivative. Wavelets have also received much attention in the M/EEG community, as this allow the temporal dynamics of the sources to be described more flexibly. In [5] the focus is to represent event related potential (ERPs) with the use of a small set of wavelet bases. Additionally, [15] presents a variational Bayesian approach which tries to represent the M/EEG signal by a small set of coefficients using a wavelet shrinkage procedure.

Alternatively, regularization directly in the time-frequency domain has also been pursued in e.g. [8] in which a $\ell_{1,2}$ regulizer is used to promote sparsity in space and smoothness in time. Smoothness in time is obtained through the ℓ_2-regularization term on frequency bands.

In this contribution we analyze new spatio-temporal representations designed to facilitate plausible source dynamics of brain networks. In contrast to most related work invoking spatio-temporal basis functions we seek source configurations that are expected to have their support in the spatio-temporal space spanned by the basis functions, however, *importantly not limited to this space*. Whether the sources are within the subspace given by spatial and temporal basis functions is determined by the data itself by trading off prior information and data driven likelihood. In contrast to, e.g., [3] we here explore the trade-off in a hierarchical Bayesian model borrowing from probabilistic graphical modeling which allows us to implement tractable posterior distributions on the sought imaging representations.

2 Methods

To solve the ill-posed M/EEG source localization problem we here propose a hierarchical Bayesian spatio-temporal model, which we will refer to as Aquavit. The proposed model decomposes $\mathbf{S}$ by both spatial and temporal basis functions in a similar form as [3]'s space-time event (STE) representation. However, in contrast to [3] we regard the spatio-temporal space spanned by the basis functions as potentially incomplete and thus we incorporate an additional source term, the spatio-temporal correction term, $\mathbf{V}$, such that

$$\mathbf{S} = \mathbf{\Psi W \Phi} + \mathbf{V}. \tag{2}$$

Here, $\mathbf{\Psi}$ is the spatial basis functions, $\mathbf{W}$ the spatio-temporal map to be estimated, and $\mathbf{\Phi}$ the temporal basis functions. While $\mathbf{W}$ is constrained to the space spanned by $\mathbf{\Psi}$ and $\mathbf{\Phi}$, the spatio-temporal correction term operates on the dipole level and independent across time, as specified below.

An important difference between [3] and our Aquavit model is that while [3] only computes point estimates by solving a penalized generalized ℓ_1 norm regression problem, the Aquavit model uses ideas from probabilistic graphical modeling, by allowing us to compute a tractable posterior distribution on the unknown sources.

Model specification:

$$p\left(\mathbf{Y}\,|\mathbf{W},\mathbf{V}\right) = \prod_{n=1}^{N_t} \mathcal{N}\left(\mathbf{y}_n\,|\mathbf{A}\boldsymbol{\Psi}\mathbf{W}\boldsymbol{\varphi}_n + \mathbf{A}\mathbf{v}_n, \boldsymbol{\Sigma}_\varepsilon\right)$$

$$p\left(\mathbf{V}\right) = \prod_{n=1}^{N_t} \mathcal{N}\left(\mathbf{v}_n\,\middle|\mathbf{0}, \boldsymbol{\Gamma}^{-1}\right), \quad p\left(\mathbf{W}\right) = \prod_{k=1}^{N_k} \mathcal{N}\left(\mathbf{w}_{k\cdot}\,\middle|\mathbf{0}, \mathbf{D}_k^{-1}\right), \tag{3}$$

with $\boldsymbol{\Sigma}_\varepsilon$ denoting the noise covariance matrix, $\boldsymbol{\Gamma} = \mathrm{diag}\left(\boldsymbol{\gamma}\right)$ the precision matrix for the sources, and $\boldsymbol{\Phi} = \left[\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_{N_t}\right]$. $\mathbf{D}_k = \mathrm{diag}\left(\boldsymbol{\alpha}_k\right)$ is the precision matrix for the spatio-temporal weights associated with the k'th temporal basis function and all spatial basis functions, which leaves us with a total of $N_d + N_k N_l$ hyperparameters ($\boldsymbol{\gamma}$ and N_k of $\boldsymbol{\alpha}_k$'s) to be estimated for the model. N_d, N_l, and N_k, denote the number of dipoles, spatial basis functions, and temporal basis functions, respectively.

Ideally, we would like to integrate out all unknowns (including hyperparameters) and then compute the posterior over sources $p\left(\mathbf{S}|\mathbf{Y}\right) = p\left(\mathbf{V},\mathbf{W}|\mathbf{Y}\right)$, which contains all possible information about $\mathbf{S}$ conditioned on the observed data $\mathbf{Y}$. However, the exact posterior in our model is computationally intractable. Here we choose to infer the hyperparameters using the maximum a-posteriori (MAP) point estimate. Given this, the joint posterior over the source 'noise' and STE maps $p\left(\mathbf{V},\mathbf{W}|\,\mathbf{Y}\right)$ is Gaussian, and can be computed analytically. Nevertheless, here we apply a variational Bayesian (VB) approximation $p\left(\mathbf{V},\mathbf{W}|\,\mathbf{Y}\right) \approx q\left(\mathbf{V},\mathbf{W}\right)$ to speed up the computation. In VB we maximize a lower bound of the log marginal likelihood, $\mathcal{L} \geq \left\langle \log p\left(\mathbf{Y},\mathbf{W},\mathbf{V}\right) - \log q\left(\mathbf{W},\mathbf{V}\right)\right\rangle_{q(\mathbf{W},\mathbf{V})} = \mathcal{F}$ where $q\left(\mathbf{W},\mathbf{V}\right)$ is the joint variational distribution and $\langle\cdot\rangle$ denotes the expected value. Other than the factorial assumption $q\left(\mathbf{W},\mathbf{V}\right) = q\left(\mathbf{W}\right)q\left(\mathbf{V}\right)$, we place no constraints on this distribution. We then iteratively maximize the objective function $\mathcal{F}$ with respect to $q\left(\mathbf{W}\right)$, $q\left(\mathbf{V}\right)$, and any hyperparameters (i.e., $\mathbf{D}_k$, $\boldsymbol{\Gamma}$) we choose to update. Upon convergence, it can be shown that $q\left(\mathbf{V}\right)$ becomes a Gaussian approximation to $p\left(\mathbf{V}|\mathbf{Y}\right)$ with analytically computable moments.

The model specification is actually related to our previous inverse method incorporating basis functions [14], however, that work does not include spatial clustering functions and apply a different strategy on which terms hyperparameters should be introduced. In contrast to [14] our new Aquavit model has seperate hyperparameters on each of the two source terms (spatio-temporal maps and the spatio-temporal correction term), which seems more in line with the factorial assumption applied by the variational framework.

Variational distributions: In turn each of the marginal posterior distributions can be maximized as follows [1]. Kronecker product is denoted $\otimes$.

$$q\left(\mathbf{V}\right) \propto \exp\Big(\left\langle \log p\left(\mathbf{Y},\mathbf{W},\mathbf{V}\right)\right\rangle_{q(\mathbf{W})}\Big) = \prod_{n} \mathcal{N}\left(\mathbf{v}_n \left|\bar{\mathbf{v}}_n, \boldsymbol{\Sigma}_v\right.\right) \tag{4}$$

$$\bar{\mathbf{v}}_n = \boldsymbol{\Sigma}_v \mathbf{A}^T \boldsymbol{\Sigma}_\varepsilon^{-1}\left(\mathbf{y}_n - \mathbf{A}\boldsymbol{\Psi}\bar{\mathbf{W}}\boldsymbol{\varphi}_n\right), \qquad \boldsymbol{\Sigma}_v^{-1} = \mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A} + \boldsymbol{\Gamma}. \tag{5}$$

Similarly, we maximize the variational posterior distribution of $q\left(\mathbf{W}\right)$ by

$$q\left(\mathbf{W}\right) \propto \exp\Big(\left\langle \log p\left(\mathbf{Y},\mathbf{W},\mathbf{V}\right)\right\rangle_{q(\mathbf{V})}\Big) = \mathcal{N}\left(\mathrm{vec}\left(\mathbf{W}\right)\Big|\mathrm{vec}\left(\bar{\mathbf{W}}\right), \widetilde{\boldsymbol{\Omega}}\right) \tag{6}$$

$$\widetilde{\boldsymbol{\Omega}}^{-1} = \boldsymbol{\Phi}\boldsymbol{\Phi}^T \otimes \boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}\boldsymbol{\Psi} + \widetilde{\mathbf{D}} \tag{7}$$

$$\mathrm{vec}\big(\bar{\mathbf{W}}\big) = \widetilde{\boldsymbol{\Omega}} \cdot \mathrm{vec}\left(\boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\left(\mathbf{Y} - \mathbf{A}\bar{\mathbf{V}}\right)\boldsymbol{\Phi}^T\right). \tag{8}$$

An interesting observation of $\widetilde{\boldsymbol{\Omega}}$ is that the computation simplifies significantly for orthonormal temporal basis functions, such that $\widetilde{\boldsymbol{\Omega}}^{-1} = \mathbf{I}_{N_k} \otimes \boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}\boldsymbol{\Psi} + \widetilde{\mathbf{D}} = \mathrm{blkdiag}\left(\boldsymbol{\Omega}_1^{-1}, \boldsymbol{\Omega}_2^{-1}, \ldots, \boldsymbol{\Omega}_{N_k}^{-1}\right)$ with $\boldsymbol{\Omega}_k^{-1} = \boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}\boldsymbol{\Psi} + \mathbf{D}_k$. With this block diagonal structure of $\widetilde{\boldsymbol{\Omega}}$ the variational posterior $q\left(\mathbf{W}\right)$ reduces to a set of N_k multivariate normal distributions,

$$q\left(\mathbf{W}\right) = \prod_{k=1}^{N_k} \mathcal{N}\left(\mathbf{w}_k \left|\bar{\mathbf{w}}_k, \boldsymbol{\Omega}_k\right.\right), \qquad \text{for} \quad \boldsymbol{\Phi}\boldsymbol{\Phi}^T = \mathbf{I}_{N_k} \tag{9}$$

$$\bar{\mathbf{w}}_k = \boldsymbol{\Omega}_k\boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\left(\mathbf{Y} - \mathbf{A}\bar{\mathbf{V}}\right)\boldsymbol{\phi}_k, \tag{10}$$

where $\boldsymbol{\phi}_k$ is the k'th temporal basis function, i.e. k'th row of $\boldsymbol{\Phi}$.

Hyperparameters: To obtain optimal variational posterior distributions the hyperparameters need to be maximized as well. Standard VB-EM pursuits hyperparameter maximization through the derivative of the complete data log likelihood w.r.t. the hyperparameters. In terms of convergency the VB-EM can be extremely slow in practice, and thus, we here pursuit update techniques by Mackay [12] resulting in generally significantly faster convergence. Given the variational posterior distributions our objective function $\mathcal{F}$ spells within a constant

$$\begin{aligned} 2\mathcal{F} = {} & N_t \log\left|\boldsymbol{\Sigma}_\varepsilon^{-1}\right| + N_t \log\left|\boldsymbol{\Gamma}\right| + \sum_{k=1} \log\left|\mathbf{D}_k\right| - N_t \log\left|\boldsymbol{\Sigma}_v^{-1}\right| - \log\left|\widetilde{\boldsymbol{\Omega}}^{-1}\right| \\ & - \sum_n \left\|\mathbf{y}_n - \mathbf{A}\left(\boldsymbol{\Psi}\bar{\mathbf{W}}\boldsymbol{\varphi}_n + \bar{\mathbf{v}}_n\right)\right\|^2_{\boldsymbol{\Sigma}_\varepsilon^{-1}} - \sum_n \left\|\bar{\mathbf{v}}_n\right\|^2_{\boldsymbol{\Gamma}} - \sum_{lk} \alpha_{lk}\bar{\mathbf{w}}_{lk}^2, \end{aligned} \tag{11}$$

where we have made used of $\left\|\mathbf{x}\right\|^2_{\Delta} = \mathbf{x}^T\Delta\mathbf{x}$. Mackay updates for $\boldsymbol{\Gamma}$ and $\mathbf{D}_k$ are

$$2\partial\mathcal{F}/\partial\gamma_i = \gamma_i^{-1} - \left(\boldsymbol{\Sigma}_v\right)_{ii} - r_i = 0, \quad r_i = \frac{1}{N_t}\sum_n \left\langle v_{in}\right\rangle^2 \tag{12}$$

$$\gamma_i = \mathbf{H}_{ii}/r_i, \quad \mathbf{H} = \mathbf{I} - \boldsymbol{\Gamma}\boldsymbol{\Sigma}_v = \mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}\boldsymbol{\Sigma}_v \tag{13}$$

We define $\mathbf{H}$ to express $\boldsymbol{\Sigma}_v$ in (12). Similarly, the update of the hyperparameter α_{lk} (l'th spatial basis function and k'th temporal basis function) is given by

$$2\partial\mathcal{F}/\partial\alpha_{lk} = \alpha_{lk}^{-1} - \widetilde{\boldsymbol{\Omega}}_{jj} - \langle w_{lk}\rangle^2 = 0 \tag{14}$$

$$\alpha_{lk} = \mathbf{G}_{jj} \Big/ \langle w_{lk}\rangle^2, \quad \mathbf{G} = \mathbf{I} - \widetilde{\mathbf{D}}\widetilde{\boldsymbol{\Omega}} = \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^T \otimes \boldsymbol{\Psi}^T\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}\boldsymbol{\Psi}\right)\widetilde{\boldsymbol{\Omega}}. \tag{15}$$

where we have defined j such that $\widetilde{\mathbf{D}}_{jj} = \alpha_{lk}$ and defined $\mathbf{G}$ to express $\widetilde{\boldsymbol{\Omega}}$.

3 Experiments

While validation of inverse methods is generally difficult to conduct due to the missing ground truth, we examine the algorithms performance on simulated data in which the ground truth is known and secondly on a real EEG dataset. We benchmark the Aquavit model with [3]'s STE model and Champagne [16]. As forward model we use a BEM obtained from OpenMEEG toolbox [7], with a cortical resolution of 8,124 vertices. In both simulations and real EEG data we apply anatomical basis functions corresponding to Brodmann's areas.

Simulations: As M/EEG recordings are known to consist of both transient and oscillatory behaviour, we follow the [11] approach of simulating both these patterns. Similar as [11] we examine temporal decompostion obtained through stationary wavelet transform with Symlets wavelets as our temporal basis functions. To mimic some of the challenges an inverse solver has to deal with, we select 3 patches on the cortical surface randomly as having same transient signatures. The oscillatory part of the artificial signal is made up of 20 Hz oscillations and three different anatomical regions being active simultaneously with random sampled phase difference, Fig. 1. All methods were tested at a signal-to-noise-ratio SNR=0dB using white noise and evaluated with the AUC measure. AUC: 0.985 (Aquavit), 0.857 ([3]'s STE), and 0.488 (Champagne [16]).

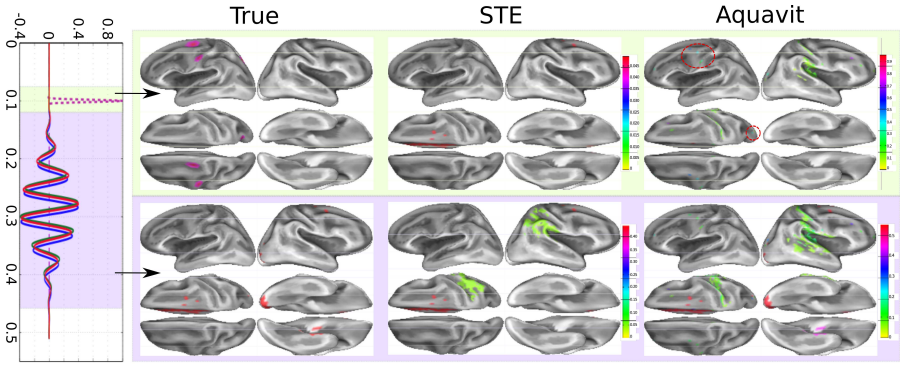
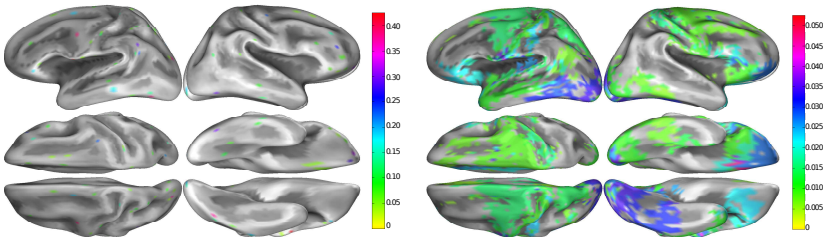


Fig. 1. Simulated sources time series and locations. Reconstructed RMS power within the window of interest using [3] and the proposed model Aquavit.

In Fig. 1 root-mean-square (RMS) activation maps during the transient (upper row) and oscillatory period (lower row) for the two spatio-temporal inverse methods are given. It seems as Aquavit detects one strong dipole located correctly and two other dipoles close by to a second simulated area in the left hemisphere, while the STE method in [3] fails to detect any of the true sources in the transient period. In fact the falsely detected area (part of medial longitudinal fissure - right hemisphere) in the transient period is a consequence of this region being active in the oscillation period. Here, it is correctly detected while STE fails in detecting the remaining two regions. In contrast, the Aquavit method detects all the regions correctly. However, it also reconstructs activity with minor amplitude in the right parietal/temporal region wrongly.

Real EEG Data: Of real EEG data we focus on a multimodal dataset studying face perception [10]. The data set includes trials with either real or scrambled faces as stimuli measured using 128 channels. In this contribution we reconstruct the average event related potential (ERP) of trials involving real faces as stimuli. Fig. 2(a) illustrates the total source contribution $\mathbf{S}$ and seems to be driven by the source noise term with a few spikes. However, the amount of energy that the STE-term account for of the total source energy is actually 34.5%. Fig. 2(b) shows the STE source term and reveals ventral occipito-temporal responses in accordance with previous face studies, [6].



(a) $\mathbf{S} = \mathbf{\Psi W \Phi} + \mathbf{V}$ (b) $\mathbf{\Psi W \Phi}$

Fig. 2. Source reconstruction using the Aquavit model on face perception data [10]. RMS-power within 160-180ms poststim.

4 Conclusion

We have proposed a hierarchical Bayesian model incorporating spatio-temporal event structure priors. The model provides tractable variational posterior distributions, which is a significant benefit compared to related state-of-the-art spatio-temporal models such as [3] as we naturally can provide estimates of the uncertainties associated with the source estimates. We have demonstrated that the model is capable in balancing spatio-temporal prior guidance and i.i.d. description of the current sources to obtain a superior estimation of the source distribution compared to current inverse methods.

Acknowledgments. This work was partly supported by the Danish Lundbeck Foundation / The Danish Council for Independent Research (CS) and R01 DC004855 / R21 NS76171 (SSN).

References

1. Attias, H.: A variational Bayesian framework for graphical models. Advances in neural information processing systems 12(1-2), 209–215 (2000)
2. Baillet, S., Garnero, L.: A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Tr. Bio.Eng. 44, 374–385 (1997)
3. Bolstad, A., Veen, B.V., Nowak, R.: Space-time event sparse penalization for magneto-/electroencephalography. NeuroImage 46(4), 1066–1081 (2009)
4. Daunizeau, J., Mattout, J., Clonda, D., Goulard, B., Benali, H., Lina, J.M.: Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. Biomedical Engineering, IEEE Transactions on 53(3), 503–516 (2006)
5. Effern, A., Lehnertz, K., Fernandez, G., Grunwald, T., David, P., Elger, C.: Single trial analysis of event related potentials: non-linear de-noising with wavelets. Clinical neurophysiology 111(12), 2255–2263 (2000)
6. Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J.: Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39, 1104–1120 (2008)
7. Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., et al.: Openmeeg: opensource software for quasistatic bioelectromagnetics. Biomed. Eng. Online 9(1), 45 (2010)
8. Gramfort, A., Strohmeier, D., Haueisen, J., Hamalainen, M., Kowalski, M.: Functional brain imaging with m/eeg using structured sparsity in time-frequency dictionaries. In: Information Processing in Medical Imaging, Lecture Notes in Computer Science, vol. 6801, pp. 600–611. Springer Berlin / Heidelberg (2011)
9. Haufe, S., Tomioka, R., Dickhaus, T., Sannelli, C., Blankertz, B., Nolte, G., Müller, K.: Large-scale eeg/meg source localization with spatial flexibility. NeuroImage 54(2), 851–859 (2011)
10. Henson, R., Goshen-Gottstein, Y., Ganel, T., Otten, L., Quayle, A., Rugg, M.: Electrophysiological and hemodynamic correlates of face perception, recognition and priming. Cerebral Cortex 13, 793–805 (2003)
11. Jmail, N., Gavaret, M., Wendling, F., Kachouri, A., Hamadi, G., Badier, J.M., Bnar, C.G.: A comparison of methods for separation of transient and oscillatory signals in eeg. Journal of Neuroscience Methods 199(2), 273 – 289 (2011)
12. MacKay, D.: Bayesian interpolation. Neural computation 4(3), 415–447 (1992)
13. Phillips, C., Rugg, M., Friston, K.: Anatomically Informed Basis Functions for EEG Source Localisation: Combining Functional and Anatomical Constraints. NeuroImage 16(3), 678–695 (2002)
14. Stahlhut, C., Attias, H.T., Wipf, D., Hansen, L.K., Nagarajan, S.S.: Sparse spatio-temporal inference of electromagnetic brain sources. In: Machine Learning in Medical Imaging, LNCS, vol. 6357, pp. 157–164. Springer Berlin / Heidelberg (2010)
15. Trujillo-Barreto, N., Aubert-Vazquez, E., Penny, W.: Bayesian M/EEG source reconstruction with spatio-temporal priors. NeuroImage 39, 318–335 (2008)
16. Wipf, D.P., Owen, J.P., Attias, H.T., Sekihara, K., Nagarajan, S.S.: Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. Neuroimage 49(1), 641–655 (2010)