Probabilistic Algorithm for Electromagnetic Brain Imaging with Spatio-Temporal and Forward Model Priors

Carsten Stahlhut[1], David Wipf[2], Hagai T. Attias[3],
Lars Kai Hansen[1], Srikantan S. Nagarajan[2]

[1]Technical University of Denmark
Department of Informatics and Mathematical Modelling
Richard Petersens Plads, DK-2800 Kgs. Lyngby, Denmark

[2]University of California, San Francisco
Department of Radiology and Biomedical Imaging
513 Parnassus Avenue, S362, San Francisco, CA 94143, USA

[3]Convex Imaging
650 California Street, San Francisco, CA 94108, USA

Technical report: June 2, 2010, revised April 6, 2011

Abstract

In this paper we present a novel spatio-temporal inverse method for solving the inverse M/EEG problem. The contribution is two-folded; firstly, the proposed model allows for a sparse spatial and temporal source representation of the M/EEG by applying an automatic relevance determination type prior. The utility of a sparse spatio-temporal representation is based on the assumption that the underlying source activity is indeed sparse and smooth in time. Secondly, we seek to reduce the influence of forward model errors on the source estimates, by applying a stochastic forward model. Applying a stochastic forward model is motivated by the random noise contributions such as the geometry of the cortical surface and the electrode positions. Simulated data provide evidence that the spatio-temporal model leads to improved source estimates, especially at low signal-to-noise ratios, which is often the case in M/EEG.

1 Introduction

In order to understand the human brain many experimental and modeling techniques have been invoked, including brain imaging by functional magnetic resonance (fMRI), positron emission tomography (PET), electro- and magnetoencephalography (EEG, MEG). EEG and MEG are both very promising modalities with their excellent temporal resolution, since they measure the neural current activity in the brain.

The relation between the measured M/EEG signal and the brain's current sources can be expressed as a linear instantaneous form in the sources. The forward relation can be written as [1]

$$\mathbf{M} = \mathbf{A}\mathbf{S} + \mathcal{E}, \tag{1}$$

where the noise $\mathcal{E}$ is assumed additive and the measured M/EEG signal is denoted $\mathbf{M} \in \Re^{N_c \times N_t}$, the current sources $\mathbf{S} \in \Re^{N_d \times N_t}$, and with N_c, N_d, and N_t being the number of channels, dipoles, and time samples, respectively. The coupling of sensors and the current sources is expressed through the lead field matrix/forward model $\mathbf{A} \in \Re^{N_c \times N_d}$ with the rows referred to as the lead fields for the sensors and the columns as the forward fields for the sources.

The existing inverse method algorithms face the challenge of the many sources of noise that interfere with the true signals in the MEG/EEG data. Electrical, thermal and biological noise as well as background room interference can be present. As a consequence of the many noisy contributions and the highly ill-posed nature of the electromagnetic source imaging (ESI) this leads to high requirements on robust inverse methods. In this paper we pursue a spatio-temporal method.

General approaches to the M/EEG inverse problem may be categorized as parametric or imaging methods [1]. In a parametric setting, or sometimes referred to as scanning methods, the M/EEG is described by a small number of dipoles [2]. The locations of these sources are found by scanning over all possible locations in order to find the best set of sources to represent the data. Beamforming [3] and MUSIC [4] are examples of scanning methods. Choosing the number of sources in the solution is the key problem for scanning methods, since this needs to be specified a priori. In contrast, imaging methods reconstruct a spatial distribution of the current sources.

The use of both spatial and temporal information to constrain the source estimates have been applied for decades. In [5] they proposed temporal priors, which operate on penalizing differences in neighboring time points. [6] incorporates temporal smoothness priors based on second derivatives.

To describe the temporal dynamics of the sources wavelet temporal basis functions have received much attention in the M/EEG community. In [7] the focus is to represent event related potential (ERPs) with the use of a small set of wavelet bases. Additionally, [8] presented an Variational Bayes approach which tries to represent the M/EEG signal by a sparse set of coefficients by applying a wavelet shrinkage procedure.

In [9] they apply event sparse penalty (ESP) regularization which is closely related to l1 regularization. The ESP approach seeks a solution composed of a small number of space-time events (STEs). Each STE is a spatio-temporal signal defined in terms of a group of basis functions. Similarly, we pursue ESI solutions with sparse spatio-temporal representation. In contrast to [9] approach with we will apply an automatic relevance determination (ARD) like prior, see Section 2.

This paper is divided two parts - firstly we present a model that takes the spatio-temporal information into consideration when solving the electromagnetic source imaging problem in order to obtain robust source estimates given a fixed forward model. Secondly, we extend this spatio-temporal model to include a stochastic forward model.

Applying a stochastic forward model is motivated by the many noise processes that contribute to the forward model. These noise processes include the conductivity distribution, electrode positions, the geometrical representation of the cortical surface, and head movements when performing MEG scannings. The geometry of the head model is influenced by the resolution and tissue segmentation errors when 'realistic head models' are constructed from tissue segmentation based on an MRI.

Several approaches have been proposed to model the forward process [10, 11, 12] yet from quite a different viewpoint than ours. In [10] the modeling of the uncertainty associated with the forward model is performed by parametric model which also includes estimation of the skull-brain conductivity ratio. Similarly, [11] apply a probabilistic distributed model that also accounts for uncertainties involved in the skull conductivity. In [12] they apply a stochastic forward model where a standard forward model is used as a prior mean. In this paper we will apply a similar stochastic forward model. However, none of these methods explores the potential of applying spatio-temporal priors.

2 The Akvavit Algorithm

We here propose a spatio-temporal model for solving the ill-posed EEG/MEG source localization problem and while also take the uncertainty of the forward model into consideration. Our model has its origin in the Champagne algorithm [13], in which each source has an associated hyperparameter to control its relevance, better known as automatic relevance determination (ARD) type prior. However, in contrast to [13], which assumes that a given current source is independent over time, we here restrict the current sources to be correlated over time by imposing a fixed temporal basis set $\mathbf{\Phi} = \left[\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{N_k}\right]^T$ consisting of N_k temporal basis functions of lengths N_t. These basis functions can either be specified based on some prior expectation or learned directly from the data [14], when pre-stimulus data is available. Similarly, we will assume the noise covariance matrix as fixed, since this as well can be learned (such as e.g. a variational Bayesian factor model proposed in [15]) and effectively suppress noisy factors.

In terms of the stochastic forward model we apply a quite similar approach taken in [12], i.e. the lead fields are modeled as independent multivariate Gaussian distributions with the conventional forward propagation model $\mathbf{A}^{(0)}$ used as prior mean and a hyperparameter associated to each of the forward fields due to an expectation that forward fields from different regions in the brain will be corrupted differently. Given the temporal basis functions we reformulate Eq. (1)

$$\mathbf{Y} = \mathbf{AG\Phi} + \mathbf{E} \tag{2}$$

with $\mathbf{G}$ being the spatio-temporal maps that we are interested to find. For notation convencience we will make use of $\mathbf{m} = vec\left(\mathbf{Y}\right)$, $\mathbf{x} = vec\left(\mathbf{G}\right)$, $\mathbf{B} = \left(\mathbf{\Phi}^T \otimes \mathbf{A}\right)$ interchangeable. The Kronecker-product[1] is denoted $\otimes$. Specification of model

$$\begin{aligned}
p\left(\mathbf{Y}\left|\mathbf{G}\right.\right) &= \prod_{n=1}^{N_t} \mathcal{N}\left(\mathbf{y}_n\left|\mathbf{AG}\boldsymbol{\varphi}_n\right.\right) = p\left(\mathbf{m}\left|\mathbf{x}\right.\right) \\
&= \prod_{n=1}^{N_t} \mathcal{N}\left(\mathbf{y}_n\left|\left(\boldsymbol{\varphi}_n^T \otimes \mathbf{A}\right)\mathbf{x}\right., \mathbf{\Sigma}_\varepsilon\right) \\
&= \mathcal{N}\left(\mathbf{m}\left|\left(\mathbf{\Phi}^T \otimes \mathbf{A}\right)\mathbf{x}\right., \mathbf{I}_{N_t} \otimes \mathbf{\Sigma}_\varepsilon\right) \qquad (3) \\
p\left(\mathbf{G}\left|\mathbf{\Gamma}\right.\right) &= p\left(\mathbf{x}\left|\mathbf{\Gamma}\right.\right) = \prod_{k=1}^{N_k} \mathcal{N}\left(\mathbf{g}_k\left|\mathbf{0}, \mathbf{\Gamma}_k^{-1}\right.\right) = \mathcal{N}\left(\mathbf{x}\left|\mathbf{0}, \mathbf{\Gamma}^{-1}\right.\right) \qquad (4) \\
p\left(\tilde{\mathbf{A}}\left|\boldsymbol{\lambda}, \boldsymbol{\alpha}\right.\right) &= p\left(\mathbf{z}\left|\boldsymbol{\lambda}, \boldsymbol{\alpha}\right.\right) = N\left(\mathbf{z}\left|\mathbf{z}^{(0)}\right., \mathbf{\Omega}^{-1}\right). \qquad (5)
\end{aligned}$$

Here $\mathbf{g}_k$ denotes the k'th column in $\mathbf{G}$ and $\mathbf{\Gamma}$ is as a blockdiagonal matrix given by $\mathbf{\Gamma} = bdiag\left(\mathbf{\Gamma}_1, ..., \mathbf{\Gamma}_{N_k}\right)$ and $\tilde{\mathbf{A}} = \mathbf{\Sigma}_\varepsilon^{-1/2}\mathbf{A}$. Additionally we have $\mathbf{z} = vec(\tilde{\mathbf{A}})$ and $\mathbf{\Omega} = \mathrm{diag}(\mathrm{vec}(\boldsymbol{\lambda}\alpha^T))$. We split the derivation of the updates of $\mathbf{G}$ and $\mathbf{A}$ into two separate problems, such that we assume the forward model to be known and fixed when calculating the posterior distribution of the spatio-temporal maps $\mathbf{G}$ and vice versa when updating the forward model $\mathbf{A}$. Alternatively, we can adopt a Variational Bayesian (VB) approach [17, 18], in which we maximize a lower bound of the log marginal likelihood. However, this may be very cumbersome and memory intensive if the number of sources N_d and number of basis functions N_k are quite large, since in a standard VB framework each variational posterior distribution needs to be evaluated in turn with respect to the all others. In contrast by assuming $\mathbf{G}$ and $\mathbf{A}$ for fixed and known when deriving $\mathbf{A}$ and $\mathbf{G}$, we obtain computationally easier update rules at the expense of an iterative approximation. This basically forces the variational posterior distribution of one the variables to be delta-functions centered at their mean values, when the other variational posterior distribution is evaluated.

Estimation of G: When $\mathbf{A}$ is assumed fixed, maximization of the posterior

[1] A discussion of the vec-operator and Kronecker product can be found in e.g. [16].

distribution for $\mathbf{G}$ leads to

$$\begin{aligned}
q(\mathbf{x}\,|\mathbf{Y}) &\propto \exp\left\{-\frac{1}{2}\Big((\mathbf{m}-\mathbf{B}\mathbf{x})^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)(\mathbf{m}-\mathbf{B}\mathbf{x})+\mathbf{x}^T\boldsymbol{\Gamma}\mathbf{x}\Big)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\mathbf{x}^T\left(\mathbf{B}^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{B}+\boldsymbol{\Gamma}\right)\mathbf{x}-2\mathbf{x}^T\mathbf{B}^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{m}\right)\right\} \\
&= \mathcal{N}(\mathbf{x}\,|\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x) \qquad (6)
\end{aligned}$$

where the covariance and mean is given by

$$\begin{aligned}
\boldsymbol{\Sigma}_x &= \left(\mathbf{B}^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{B}+\boldsymbol{\Gamma}\right)^{-1} \\
&= \Big(\left(\boldsymbol{\Phi}^T\otimes\mathbf{A}\right)^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\left(\boldsymbol{\Phi}^T\otimes\mathbf{A}\right)+\boldsymbol{\Gamma}\Big)^{-1} \\
&= \Big(\boldsymbol{\Phi}\boldsymbol{\Phi}^T\otimes\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}+\boldsymbol{\Gamma}\Big)^{-1} \qquad (7) \\
\boldsymbol{\mu}_x &= \boldsymbol{\Sigma}_x\mathbf{B}^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{m}=\boldsymbol{\Sigma}_x\left(\boldsymbol{\Phi}^T\otimes\mathbf{A}\right)^T\left(\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{m} \\
&= \boldsymbol{\Sigma}_x\left(\boldsymbol{\Phi}\otimes\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{m}. \qquad (8)
\end{aligned}$$

We note that by selecting the temporal basis functions to be orthonormal, the estimation of the posterior distribution $q(\mathbf{x}\,|\mathbf{Y})$ reduces significantly since the covariance now becomes a block-diagonal matrix which can be inverted in sub-blocks. With $\boldsymbol{\Phi}\boldsymbol{\Phi}^T=\mathbf{I}_{N_k}$ the covariance is given

$$\boldsymbol{\Sigma}_x=\left(\mathbf{I}_{N_k}\otimes\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}+\boldsymbol{\Gamma}\right)^{-1}=bdiag\left(\boldsymbol{\Sigma}_{\mathbf{g}_1},...,\boldsymbol{\Sigma}_{\mathbf{g}_{N_k}}\right) \qquad (9)$$

where

$$\boldsymbol{\Sigma}_{\mathbf{g}_k}=\left(\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{A}+\boldsymbol{\Gamma}_k\right)^{-1}. \qquad (10)$$

Thus, the posterior distribution is a product of N_k normal distributions such that we have

$$q(\mathbf{x}\,|\mathbf{Y})=q(\mathbf{G}\,|\mathbf{Y})=\prod_{k=1}^{N_k}\mathcal{N}\left(\mathbf{g}_k\,|\boldsymbol{\mu}_{g_k},\boldsymbol{\Sigma}_{g_k}\right) \qquad (11)$$

$$\begin{aligned}
\boldsymbol{\mu}_{g_k} &= \boldsymbol{\Sigma}_{g_k}\left(\boldsymbol{\phi}_k^T\otimes\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\right)\mathbf{m}=\boldsymbol{\Sigma}_{g_k}\mathbf{A}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{Y}\boldsymbol{\phi}_k \\
&= \boldsymbol{\Gamma}_k^{-1}\mathbf{A}^T\left(\mathbf{A}\boldsymbol{\Gamma}_k^{-1}\mathbf{A}^T+\boldsymbol{\Sigma}_\varepsilon\right)^{-1}\mathbf{Y}\boldsymbol{\phi}_k. \qquad (12)
\end{aligned}$$

Since the model is specified given a set of hyperparameters $\boldsymbol{\Gamma}$, we need to optimize these as well in order to obtain the most likely posterior distribution. This is performed by maximization of the marginal likelihood or sometimes referred to as the model evidence. The marginal likelihood is given by

$$\begin{aligned}
p(\mathbf{Y}\,|\boldsymbol{\Gamma}) &= \int p(\mathbf{Y}\,|\mathbf{x})\,p(\mathbf{x}\,|\boldsymbol{\Gamma})\,d\mathbf{x} \\
&= N\left(\mathbf{m}\,\middle|\,\mathbf{0},\left(\boldsymbol{\Phi}^T\otimes\mathbf{A}\right)\boldsymbol{\Gamma}^{-1}\left(\boldsymbol{\Phi}^T\otimes\mathbf{A}\right)^T+\mathbf{I}_{N_t}\otimes\boldsymbol{\Sigma}_e\right) \\
&= N(\mathbf{m}\,|\mathbf{0},\boldsymbol{\Sigma}_m)\propto|\boldsymbol{\Sigma}_m|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\mathbf{m}^T\boldsymbol{\Sigma}_m^{-1}\mathbf{m}\right) \qquad (13)
\end{aligned}$$

Mathematically, maximizing the marginal likelihood is equivalent to minimizing

$$\mathcal{L}(\boldsymbol{\Gamma}) = -2\log p(\mathbf{Y}|\boldsymbol{\Gamma}) = -\frac{1}{2}\log|\boldsymbol{\Sigma}_m| - \mathbf{m}^T\boldsymbol{\Sigma}_m^{-1}\mathbf{m} \tag{14}$$

Optimization of the hyperparameters can conviently be done in expected-maximization (EM) framework [18], in which the hyperparameters is updated by taking the derivative of the expected complete data log likelihood w.r.t. the hyperparameters in the M-step.

$$\frac{\partial}{\partial\gamma_i}\langle\log p(\mathbf{m},\mathbf{x})\rangle_{q(\mathbf{x})} \Rightarrow \gamma_i^{-1} \quad = \quad (\boldsymbol{\Sigma}_x)_{ii} + (\boldsymbol{\mu}_x)_i^2\,. \tag{15}$$

However, the convergence can be extremely slow, and thus we apply the MacKay update rule, which are significant faster to converge [19]. The updates are obtained by taking the derivative of Eq. (14) and equate to zero, which leads to the following fixed-point update

$$\gamma_i = (1-\gamma_i(\boldsymbol{\Sigma}_x)_{ii})\Big/(\mu_x)_i^2 \tag{16}$$

From Eq. (16) it is noted that it is only the contribution from the statistics for the i'th element in $\mathbf{G}$ that contributes to the update. In our formulation of the model we specified a hyperparameter per element in $\mathbf{G}$. In relation to the approach proposed in [9], where they seek sparse sets of space-time subspaces, the proposed model in this paper can find sparse representation of temporal subspaces to the same extent. This situation is actually a special case of the formulated model, since it would correspond to restricting a group of γ_i^{-1} values to their mean value, i.e. we have $\gamma_{i\subset W}^{-1} = N_w^{-1}\sum_{w=1}^{N_w}\gamma_w^{-1}$ with W denoting the group of spatio-temporal events with the same hyperparameter.

Estimation of A We now assume the spatio-temporal map $\mathbf{G}$ to be constant when deriving the updates for the forward model. For notational convenience we remove the noise covariance matrix by reformulating the likelihood in a equivalent form

$$p\left(\tilde{\mathbf{Y}}\middle|\mathbf{G},\tilde{\mathbf{A}}\right) \quad = \quad \prod_{n=1}^{N_t}\mathcal{N}\left(\tilde{\mathbf{y}}_n\middle|\tilde{\mathbf{A}}\mathbf{G}\varphi_n,\mathbf{I}_{N_c}\right) \tag{17}$$

$$\begin{aligned} p\left(\tilde{\mathbf{m}}\middle|\tilde{\mathbf{A}},\mathbf{x}\right) &= \mathcal{N}\left(\tilde{\mathbf{m}}\middle|\left(\boldsymbol{\Phi}^T\otimes\tilde{\mathbf{A}}\right)\mathbf{x},\mathbf{I}_{N_cN_t}\right) \\ &= N\left(\tilde{\mathbf{m}}\middle|\left((\mathbf{G}\boldsymbol{\Phi})^T\otimes\mathbf{I}_{N_c}\right)\mathbf{z},\mathbf{I}_{N_cN_t}\right), \end{aligned} \tag{18}$$

where we have made use of $\tilde{\mathbf{Y}} = \boldsymbol{\Sigma}_\varepsilon^{-1/2}\mathbf{Y}$, $\tilde{\mathbf{A}} = \boldsymbol{\Sigma}_\varepsilon^{-1/2}\mathbf{A}$, $\tilde{\mathbf{E}} = \boldsymbol{\Sigma}_\varepsilon^{-1/2}\mathbf{E}$, $\tilde{\mathbf{m}} = \mathrm{vec}(\tilde{\mathbf{Y}})$, $\mathbf{x} = \mathrm{vec}(\mathbf{G})$, $\mathbf{z} = \mathrm{vec}(\tilde{\mathbf{A}})$, and the vec-operator and Kronecker product[2]. To obtain the posterior distribution of the forward model we now assume the

[2]A discussion of the vec-operator and Kronecker product can be found in e.g. [16].

source-temporal maps $\mathbf{G}$ for known and thus the posterior distribution is given by

$$
\begin{aligned}
q\left(\mathbf{z}\,|\tilde{\mathbf{m}}\right) &= \exp\left\{-\frac{1}{2}\left\|\tilde{\mathbf{m}}-\left((\mathbf{G}\boldsymbol{\Phi})^T\otimes\mathbf{I}_{N_c}\right)\mathbf{z}\right\|_2^2-\frac{1}{2}\left\|\mathbf{z}^T\boldsymbol{\Omega}\mathbf{z}\right\|_2^2\right\} \\
&= \exp\left\{-\frac{1}{2}\left(\mathbf{z}^T\left(\mathbf{G}\boldsymbol{\Phi}\boldsymbol{\Phi}^T\mathbf{G}^T\otimes\mathbf{I}_{N_c}+\boldsymbol{\Omega}\right)\mathbf{z}-2\mathbf{z}^T\left((\mathbf{G}\boldsymbol{\Phi})^T\otimes\mathbf{I}_{N_c}\right)^T\tilde{\mathbf{m}}\right)\right\} \\
&= \mathcal{N}\left(\mathbf{z}\,|\mu_z,\boldsymbol{\Sigma}_z\right) \qquad (19)
\end{aligned}
$$

in which

$$
\mu_z=\boldsymbol{\Sigma}_z\left((\mathbf{G}\boldsymbol{\Phi})^T\otimes\mathbf{I}_{N_c}\right)^T\tilde{\mathbf{m}} \quad \text{and} \quad \boldsymbol{\Sigma}_z=\left(\mathbf{G}\boldsymbol{\Phi}\boldsymbol{\Phi}^T\mathbf{G}^T\otimes\mathbf{I}_{N_c}+\boldsymbol{\Omega}\right)^{-1}. \qquad (20)
$$

Given the posterior distribution of the forward model $\mathbf{A}$ and the source-temporal maps $\mathbf{G}$, we can perform an iterative update of each of the variable with the assumption that the other variable is fixed and known. Since we assume $\mathbf{A}$ is fixed and known when updating the source-temporal maps $\mathbf{G}$ the posterior distribution is given in Eq. 6-7. Similar to the Akvavit algorithm the expression for the mean and covariance for the forward model simplifies significantly when the temporal basis functions are chosen to be orthonormal. For orthonormal temporal basis functions the mean and covariance of $\mathbf{G}$ is given by Eq. 9-12 and for $\mathbf{A}$ we have

$$
\begin{aligned}
\boldsymbol{\Sigma}_z &= \left(\mathbf{G}\mathbf{G}^T\otimes\mathbf{I}_{N_c}+\boldsymbol{\Omega}\right)^{-1}=bdiag\left(\boldsymbol{\Sigma}_{\tilde{\mathbf{l}}_1},\boldsymbol{\Sigma}_{\tilde{\mathbf{l}}_2},...,\boldsymbol{\Sigma}_{\tilde{\mathbf{l}}_{Nc}}\right) \qquad (21)\\
\boldsymbol{\Sigma}_{\tilde{\mathbf{l}}_j} &= \left(\mathbf{G}\mathbf{G}^T+\boldsymbol{\Omega}_j\right)^{-1} \qquad (22)\\
\mu_{\tilde{\mathbf{l}}_j} &= \boldsymbol{\Sigma}_{\mathbf{l}_j}\mathbf{G}\boldsymbol{\Phi}\mathbf{Y}_{j\cdot}^T=\left(\mathbf{G}\mathbf{G}^T+\boldsymbol{\Omega}_j\right)^{-1}\mathbf{G}\boldsymbol{\Phi}\tilde{\mathbf{Y}}_{j\cdot}^T\\
&= \boldsymbol{\Omega}_j^{-1}\mathbf{G}\left(\mathbf{G}^T\boldsymbol{\Omega}_j^{-1}\mathbf{G}+\mathbf{I}_{N_k}\right)^{-1}\boldsymbol{\Phi}\tilde{\mathbf{Y}}_{j\cdot}^T \qquad (23)
\end{aligned}
$$

Similar to the hyperparameters associated with $\mathbf{G}$ we also need to the hyperparameters associated with $\mathbf{A}$. The MacKay update is given by $\Omega_{ii}=\left(1-\Omega_{ii}\left(\boldsymbol{\Sigma}_z\right)_{ii}\right)\big/\left(\mu_z\right)_i^2$.

3 Empirical Evaluation

In this section we test the performance of our algorithm on simulated data and a fixed forward model. For validation purposes we use three widely used metrics: mean square error (MSE) between the simulated and estimated source distribution, variance explained (VE) of data, and area under the receiver-characteristic curve (AUC). More comprehensive experiments with focus on the forward modeling as well as performance on real M/EEG data will be presented in forthcoming papers. In our tests we evaluate two extremes of the Akvavit algorithm in terms of number of hyperparameters associated to the temporal basis. We will refer to the two situations as Akvavit1 and Akvavit2. Akvavit1 includes a hyperparameter per source candidate in (i.e. a parameter per row in $\mathbf{G}$), whereas Akvavit2 includes a hyperparameter per source candidate per temporal basis function.

For the sake of simulation, we apply the models to a relatively low cortical resolution ($\sim$500 vertices) obtained by subsampling a more dense set. However, such low resolution or even lower resolution, might actually be the case if applied in a context similar to [9], in which vertices are clustered such that they form a number of patches based on anatomical and functional information. As head model we use a Boundary Element Method (BEM) forward model from SPM8[3].

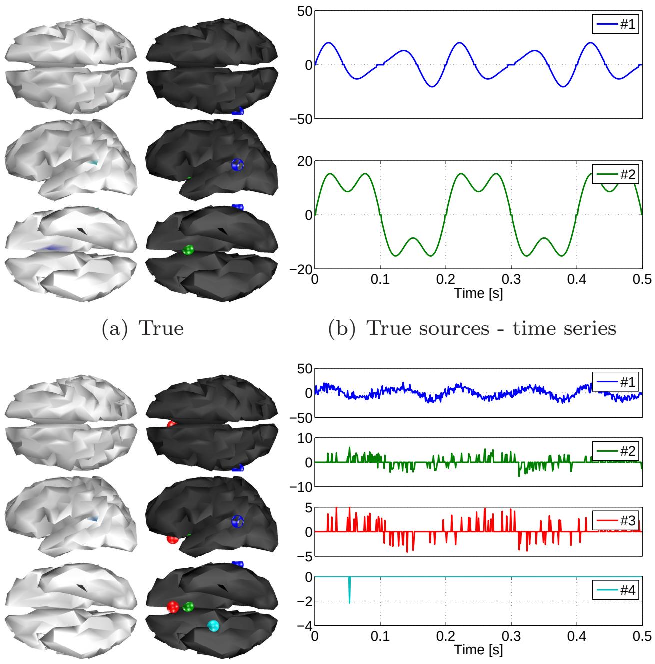
With simulated data evaluate how well the different methods reconstruct two simulated current source generated from a mixture of two sines-functions with different frequencies. We examine the performance of the methods at different signal-to-noise ratios (SNR). We inspect two scenarios: the simulated sources time series are within the temporal subspaces that are used by the our algorithm or outside the subspace. However, note that even though the sources time series are within the temporal subspaces this does not mean that a temporal basis directly represents a source time series but in stead that simulated time series can be constructed from the temporal basis set. We compare our algorithm with the champagne (CH) algorithm [20] and model proposed in [9] for finding sparse representation of small number of space-time events. The latter one we will denote ESP.

Due to the limited space in the paper we only include the results with the simulated sources lying within the temporal subspace $\mathbf{\Phi}$, since the $\mathbf{\Phi}$ may be directly learned from the data as in [14]. However, not surprisingly the performance with source time course outside the subspace deteriorate the ESI estimates of the methods relying on the temporal prior to perform equal and sometimes even worse than methods not incorporating the temporal information. Thus, the learning or specifying the temporal dictionary is crucial in order to obtain robust ESI methods.

In Fig. 1(a) an example of two simulated sources and their associated time courses are given with the white brain illustrating the source activity at time point t = X. On the black brain we have added colored spheres, such that it is possible to see which source is associated with the time courses given in the subfigure to the right. Figure 1(c) illustrates the source estimates obtained using the champagne algorithm. It is noted that the algorithm leads to a relative sparse solution only with 4 remaining source in the estimate. Of the estimated sources the position of both source #1 and #2 are correctly reconstructed. However, their time courses on the other hand are highly corrupted by noise. Especially the time series for source #2 is very difficult to recognize. Source reconstruction using the [9] approach with event sparse penalty (ESP) are given in Fig. 2(a) and Fig. 2(c) corresponding to ESP1 and ESP2, respectively. In ESP1 we use the same approach as we do with our Akvavit1 approach, i.e. we

[3]The forward model was constructed using the SPM8 academic software (`http://www.fil.ion.ucl.ac.uk/spm/`) based on routines from FieldTrip (`http://fieldtrip.fcdonders.nl/`).

only seek a sparse solution in the spatial domain. In contrast ESP2 will favor sparsity in both space and temporal basis functions. Unfortunately both models are only able to localize the position of the simulated source #1 correctly and with ESP1 slightly better in capturing the overall temporal evolution of the source. Note that source that correspond most best to the simulated source #1 in ESP1 appears as source #3 (red). However, all the source estimates are very small, which may indicate too much regularization found by the heuristic approach for selecting the regularization parameter in [9]. Indeed the source reconstruction with Akvavit algorith leads to improved source estimates for both Akvavit1 and Akvavit2 (see Fig. 2(e) and 2(g)), with Akvavit1 as the winner. The reconstructed time courses for Akvavit1 are very similar to the simulated once and more importantly also correctly associated the location in the brain. Slightly reduced estimation of time series of source #2 is obtained with Akvavit2.



(a) True (b) True sources - time series

(c) Champagne (d) Champagne - time series

Figure 1: White brain: Activity at a snap shot. Black brain: illustrates the five sources with the largest variances. The sources are color-coded such that their corresponding time courses can be seen in the right most plot.

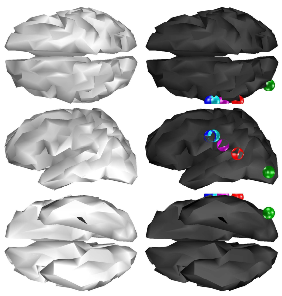
In Fig. 3 we examine the models performance on the validation metrics (MSE, VE, and AUC) as function of different SNRs. A similar overall picture as the results reported in Fig. 1 and Fig. 2 are found here, i.e. with the Akvavit1 as the best model and Akvavit2 as the runner up. The champagne is also performing quite well especially a higher SNRs. Due to the down weighting of the source estimates found by ESP1 and ESP2 these models both perform relative poorly in the metrics.

4 Discussion

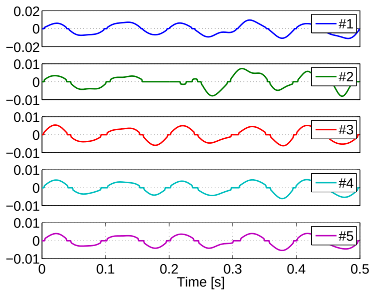
This paper presented a novel spatio-temporal model, which describes the underlying current activity by a spatially and temporally sparse representation. Moreover, we derived the model to be able to model noisy contributions in the forward model that may corrupt the source estimates. It should be noted that we applied a stochastic forward model, and thus deterministic errors which are likely to occur as well will not be modeled in present algorithm.

We presented the model in the extreme case in which a hyperparameter is assigned per spatial-temporal component and similarly for the forward model. This way we can easily reduce the model to a simpler one (in terms of a reduction in number of parameters), if e.g. one believes that the sources are better described by a group of spatio-temporal subspaces. In such a setting one can group the hyperparameters that are part of that specific group and assign them to one common value for the subspace. A serious concern when modeling the forward model is the risk of overfitting the data. By applying an ARD type prior on both the forward model and the matrix including the spatio-temporal maps, we sought solutions that prune most of the parameters according to their relevance or not.

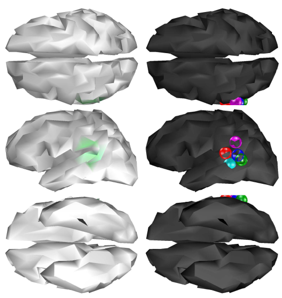
Simulations with a fixed forward model demonstrated that the proposed model lead to improved sources estimates at low SNR compared to some of the state-of-the-art methods; the event sparse penalty model proposed by [9] and the champagne algorithm, which does not include temporal information. In the derivation of the model we applied an iterative approximation when calculating the posterior distribution of the spatio-temporal matrix $\mathbf{G}$ and the posterior distribution of the forward model $\mathbf{A}$ in order to minimize the computational complexity of the updates by basically discarding the second moment.



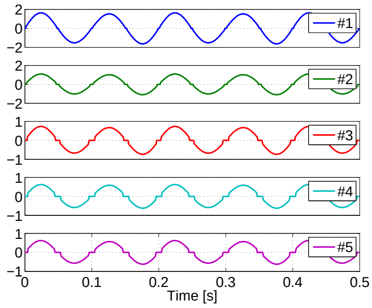
(a) ESP1



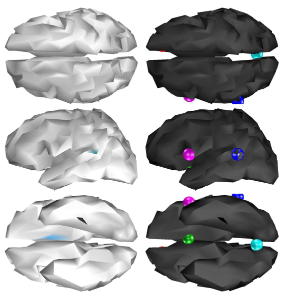
(b) ESP1 - time series



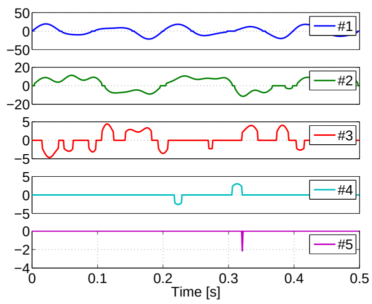
(c) ESP2



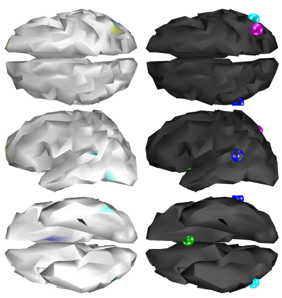
(d) ESP2 - time series



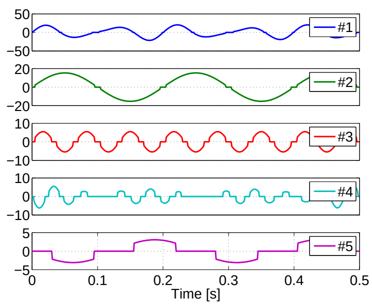
(e) Akvavit1



(f) Akvavit1 - time series

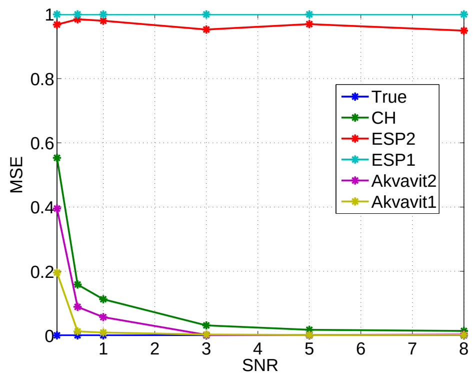


(g) Akvavit2

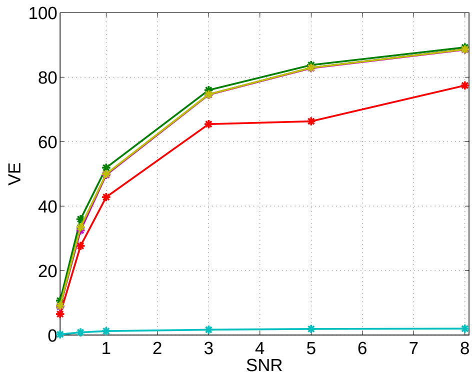


(h) Akvavit2 - time series

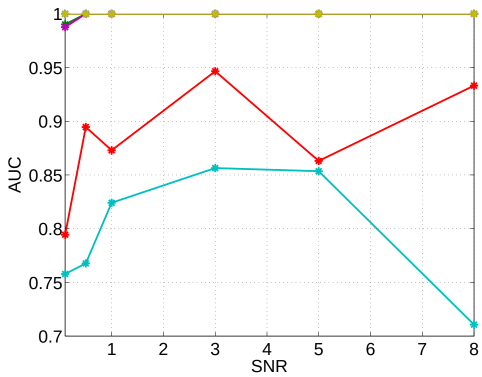
Figure 2: White brain: Activity at a snap shot. Black brain: illustrates the five sources with the largest variances. The sources are color-coded such that their corresponding time courses can be seen in the right most plot.



(a) MSE



(b) VE



(c) AUC

Figure 3: Mean square error, degree of focality, variance explained, area under ROC-curve, correlaiton coefficient, A-prime. The graphs are the mean of 5 repetitions at each SNR.

References

[1] S. Baillet, J.C. Mosher, and R.M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18:14–30, 2001.

[2] M. Scherg, T. Bast, and P. Berg. Multiple source analysis of interictal spikes: Goals, requirements, and clinical value. *J. Clin. Neurophysiol.*, 16:214–224, 1999.

[3] BD Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE transactions on biomedical engineering*, 44(9):867–880, 1997.

[4] J. C. Mosher, P. S. Lewis, and R. M. Leahy. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. on Biomedical Engineering*, 39(6):541–557, 1992.

[5] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans. on Biomedical Engineering*, 44(5):374–385, 1997.

[6] J. Daunizeau, J. Mattout, D. Clonda, B. Goulard, H. Benali, and J. M. Lina. Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. *Biomedical Engineering, IEEE Transactions on*, 53(3):503–516, 2006.

[7] A. Effern, K. Lehnertz, G. Fernandez, T. Grunwald, P. David, and CE Elger. Single trial analysis of event related potentials: non-linear de-noising with wavelets. *Clinical neurophysiology*, 111(12):2255–2263, 2000.

[8] N.J. Trujillo-Barreto, E. Aubert-Vazquez, and W.D. Penny. Bayesian M/EEG source reconstruction with spatio-temporal priors. *NeuroImage*, 39:318–335, 2008.

[9] A. Bolstad, B. Van Veen, and R. Nowak. Spacetime event sparse penalization for magneto-/electroencephalography. *NeuroImage*, 46(4):10661081, 2009.

[10] S. Lew, C. Wolters, A. Anwander, S. Makeig, and R.S. MacLeod. Low resolution conductivity estimation to improve source localization. In *New Frontiers in Biomagnetism. Proc. of the 15th Int. Conf. on Biomag.*, volume 1300 of *Int. Congress Series*, pages 149–152, 2007.

[11] S.M. Plis, J.S. George, S.C. Jun, D.M. Ranken, P.L. Volegov, and D.M. Schmidt. Probabilistic forward model for electroencephalography source analysis. *Physics in Medicine and Biology*, 52(17):5309–5328, 2007.

[12] C. Stahlhut, M. Mørup, O. Winther, and L. K. Hansen. Hierarchical bayesian model for simultaneous eeg source and forward model reconstruction (SOFOMORE). In *Machine Learning for Signal Processing, 2009. Proceedings of the 2009 19th IEEE Signal Processing Society Workshop on*, 2009.

[13] D. P. Wipf, J. P. Owen, H. T. Attias, K. Sekihara, and S. S. Nagarajan. Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *Neuroimage*, 49(1):641–655, 2010.

[14] J. M. Zumer, H. T. Attias, K. Sekihara, and S. S. Nagarajan. Probabilistic algorithms for meg/eeg source reconstruction using temporal basis functions learned from data. *NeuroImage*, 41:924–940, 2008.

[15] S. S. Nagarajan, H. T. Attias, K. E. Hild, and K. Sekihara. A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data. *Statistics in Medicine*, 26(21):3886–3910, 2007.

[16] R.A. Horn and C.R. Johnson. *Topics in matrix analysis*. Cambridge Univ Pr, 1994.

[17] H. Attias. A variational Bayesian framework for graphical models. *Advances in neural information processing systems*, 12(1-2):209–215, 2000.

[18] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, NY 10013 (USA), 2006.

[19] M.E Tipping. Sparse Bayesian learning and the relevance vector machine. *J. of Machine Learning Research 1*, 1:211–244, 2001.

[20] D. Wipf, J. Owen, H. T Attias, K. Sekihara, and S. S. Nagarajan. Estimating the location and orientation of complex, correlated neural activity using meg. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1777–1784. 2009.