Improving EEG Source Localization through Spatio-temporal Sparse Bayesian Learning

Ali Hashemi
Institut für Mathematik
Technische Universität Berlin
Strasse des 17. Juni 136, Berlin, Germany
Email:hashemi@math.tu-berlin.de

Stefan Haufe
Machine Learning Group
Technische Universität Berlin
Marchstr. 23, Berlin, Germany
Email:stefan.haufe@tu-berlin.de

***Abstract*—Sparse Bayesian Learning (SBL) approaches to the EEG inverse problem such as Champagne have been shown to outperform traditional ℓ_1-norm based methods in terms of reconstructing sparse source configurations. Current approaches are however sensitive to strong noise contributions and assume independent samples, whereas neurophysiological time series are strongly auto-correlated. Here we present extensions, backed by compressive sensing theory, to the Champagne algorithm that improve the reconstruction performance in low-SNR settings as well as in the presence of correlated measurements. Our numerical simulations using a realistic EEG forward model confirm the efficacy of our approaches.**

I. Introduction

Electroencephalography (EEG) is a non-invasive brain imaging technique that allows one to monitor brain activity with high temporal resolution. EEG activity is measured by a small number of sensors located outside the head. The EEG *measurement matrix*, $\mathbf{Y} \in \mathbb{R}^{M\times T}$, consists of T column vectors, $\mathbf{y}(t) \in \mathbb{R}^{M\times 1}, t = 1,\dots,T$ representing T samples of the EEG measurement at M sensors. It can be written as $\mathbf{Y} = \mathbf{L}\mathbf{X} + \mathbf{E}$, where the brain activity at N brain locations, $\mathbf{x}(t) \in \mathbb{R}^{N\times 1}, t = 1,\dots,T$, is stacked into the *source matrix*, $\mathbf{X} \in \mathbb{R}^{N\times T}$. The *lead field* matrix, $\mathbf{L} \in \mathbb{R}^{M\times N}$, maps the activity of the brain sources to the sensors and can be obtained for a given head geometry and estimates of the electrical conductivities of the main tissue types using the quasi-static approximation of Maxwell's equations to model the flow of extracellular neuronal currents in the head. It is typically computed using discretization methods [1], [2]. The matrix $\mathbf{E} \in \mathbb{R}^{M\times T}$ denotes white Gaussian noise assumed to be independent of the source activations. The EEG inverse problem is to infer the brain activity $\mathbf{X}$ from the EEG measurement $\mathbf{Y}$ for a given lead field matrix $\mathbf{L}$ (in other words to localize EEG measurements to anatomical brain structures). It is a highly ill-posed problem, since the number of brain sources ($10^4 \sim 10^5$) is much larger than the number of EEG sensors ($32 \sim 256$). As the lead field matrix is heavily ill-conditioned, the solution to the inverse problem is moreover highly sensitive to small changes in the measurement as well as to noise. The accuracy with which the locations of the brain sources can be reconstructed from EEG (in other words, the spatial resolution) is, therefore, relatively poor.

A common approach to deal with the ill-posedness of the EEG inverse problem is to employ regularized maximum-likelihood approaches minimizing the cost function $\mathcal{L}(\mathbf{X}) = ||\mathbf{Y} - \mathbf{L}\mathbf{X}||_Q^2 + \mathcal{R}(\mathbf{X})$. Here, the first term represents the *data fidelity* and the second term is a *penalty* encoding prior belief on spatial or temporal properties of the sources to be reconstructed. Common penalties are ℓ_1-norms, inducing sparsity [3], ℓ_2-norms, inducing smoothness [4], mixed norms [5]–[7], and penalties enforcing sparsity in a different domain using dictionaries such as Gabor frames or cortical patches [8]–[10]. Note that regularized maximum-likelihood approaches can also be interpreted in a Bayesian sense as maximum a-posteriori (MAP) estimators assuming pre-defined prior distributions for the brain sources (e.g., [11]). A powerful advancement of these approaches are so-called type-II maximum-likelihood (ML-II, also known as empirical Bayes) methods, in which the parameters of the prior distribution (i.e., the *hyperparameters*) are learned from the data along with the model parameters [12].

A desirable property of source reconstruction algorithms is the ability to correctly determine the set of active brain sites (in other words, the spatial support of $\mathbf{X}$). ML-II algorithms using *sparse Bayesian learning* (SBL, [13]) such as *Champagne* [14, Chapter 4] [15] have been shown to outperform MAP approaches using sparsity-inducing ℓ_1-norms in this regard. Current SBL algorithms for brain source localization are, however, characterized by high computational cost, and may however suffer from reduced performance in low-SNR (signal-to-noise ratio) regimes. Moreover, these algorithms typically assume independent samples. Consequently, they are unable to exploit temporal correlations in the data, and may even be negatively affected by such correlations. Finally, there is no theory yet in this field that would provide a non-asymptotic bound linking the amount of available data to the accuracy with which the active brain regions can be determined.

In this paper, we address these limitations building on recent results from the signal processing literature. First, we provide a novel algorithm for solving the SBL problem using a convex upper bound of the cost function that is tight in low-SNR regimes. Moreover, we present a theoretical result that expresses the spatial source reconstruction accuracy as a function of the number of samples. To comply with the

ISBN 978-90-827970-1-5

assumption of *independent and identically distributed* (i.i.d.) samples made by the theorem, we introduce a data-dependent temporal whitening scheme. Numerical simulations are carried out to assess the efficacy of the proposed approach for i.i.d. as well as auto-correlated pseudo-EEG data. As part of these simulations, we compare realistic EEG lead fields to artificial matrices exhibiting the restricted isometry property (RIP) required by our theory.

II. Sparse Bayesian Learning/Champagne

In the sparse Bayesian learning framework, a Gaussian prior distribution is assumed for the underlying brain sources, where it is assumed that the activities at different brain locations are independent, and where the variance at each location, γ_n, is an unknown parameter: $x_n(t) \sim \mathcal{N}(0, \gamma_n), n = 1, \ldots, N$. Then, the probability distributions of the sources and measurements can be derived as follows:

$$\mathbb{P}(\mathbf{X}|\boldsymbol{\gamma}) = \prod_{t=1}^{T} \mathbb{P}(\mathbf{x}(t)|\boldsymbol{\gamma}) = \prod_{t=1}^{T} \mathcal{N}(0, \boldsymbol{\Gamma}) \tag{1}$$

$$\mathbb{P}(\mathbf{Y}|\mathbf{X}) = \prod_{t=1}^{T} \mathbb{P}(\mathbf{y}(t)|\mathbf{x}(t)) = \prod_{t=1}^{T} \mathcal{N}(\mathbf{L}\mathbf{x}(t), \sigma^2\mathbf{I}) \ , \tag{2}$$

where $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_N]^\top$, $\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$, and where σ^2 denotes the measurement noise variance, which is considered known throughout this setting. This assumption is reasonable in the typical setting in which task-related brain activity is analysed. Here, the noise variance can be estimated from a preceding resting state measurement.

Instead of computing the full posterior distribution, $\mathbb{P}(\mathbf{X}, \boldsymbol{\gamma}|\mathbf{Y})$, ML-II algorithms replace the hyperparameters $\boldsymbol{\gamma}$ by their MAP estimates, $\hat{\boldsymbol{\gamma}}$, and focus on estimating the posterior distribution $\mathbb{P}(\mathbf{X}|\mathbf{Y}, \hat{\boldsymbol{\gamma}})$. In the context of EEG and MEG (magnetoencephalography), this approach has been introduced as the *Champagne* algorithm [14, Chapter 4]. It can be shown that for given $\boldsymbol{\gamma}$, this posterior is Gaussian [14]:

$$\mathbb{P}(\mathbf{X}|\mathbf{Y}, \boldsymbol{\gamma}) = \prod_{t=1}^{T} \mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}}(t), \boldsymbol{\Sigma}_{\mathbf{x}}) \ , \text{where} \tag{3}$$

$$\boldsymbol{\mu}_{\mathbf{x}}(t) = \boldsymbol{\Gamma}\mathbf{L}^\top(\boldsymbol{\Sigma}_{\mathbf{y}})^{-1}\mathbf{y}(t) \tag{4}$$

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{L}^\top(\boldsymbol{\Sigma}_{\mathbf{y}})^{-1}\mathbf{L}\boldsymbol{\Gamma} \tag{5}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}} = \sigma^2\mathbf{I} + \mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top \ . \tag{6}$$

For given $\boldsymbol{\mu}_{\mathbf{x}}(t)$ and $\boldsymbol{\Sigma}_{\mathbf{x}}$, the log of the marginal likelihood $\mathbb{P}(\mathbf{Y}|\boldsymbol{\gamma})$ is given by

$$\mathcal{L}(\boldsymbol{\gamma}) = \log|\boldsymbol{\Sigma}_{\mathbf{y}}| + \frac{1}{T}\sum_{t=1}^{T} \mathbf{y}(t)^\top \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}(t) \ . \tag{7}$$

Optimizing (7) w.r.t. $\boldsymbol{\gamma}$ leads to the update rule

$$\gamma_n = [\Sigma_{\mathbf{x}}]_{n,n} + \left[\frac{1}{T}\sum_{t=1}^{T}(\mu_{\mathbf{x}}(t))_n^2\right] \ \text{for } n = 1, \ldots, N \ . \tag{8}$$

Final estimates of the model parameters $\boldsymbol{\mu}_{\mathbf{x}}(t)$, $\boldsymbol{\Sigma}_{\mathbf{x}}$ and $\boldsymbol{\gamma}$ are obtained by iterating the update rules (4)–(6) and (8) until convergence [16]. The resulting algorithm is called expectation maximization (EM)-Champagne. Note that, in the final solution, many individual variances γ_n will be zero, leading to a sparse solution $\mathbf{X}$. In practice, values exactly equal to zero may not be obtained. Therefore, a those variances, for which $\gamma_n < \gamma_{\text{thresh}}, \ n = 1, \ldots, N$ holds, are set to zero in each iteration of the algorithm, where γ_{thresh} is a threshold.

An alternative algorithm can be derived by replacing the log-likelihood (7) with the convex upper bound

$$\tilde{\mathcal{L}}(\boldsymbol{\gamma}, \tilde{\mathbf{X}}, \mathbf{z}) = \mathbf{z}^\top\boldsymbol{\gamma} - \mathbf{z}_0 \tag{9}$$
$$+ \min_{\tilde{\mathbf{X}}} \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{\sigma^2}||\mathbf{y}(t) - \mathbf{L}\tilde{\mathbf{x}}(t)||^2 + \tilde{\mathbf{x}}(t)^\top\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{x}}(t)\right] \ ,$$

using auxiliary variables $\tilde{\mathbf{X}}$ and $\mathbf{z}$ [15]. Minimizing (9) w.r.t. $\boldsymbol{\gamma}$, $\tilde{\mathbf{X}}$, and $\mathbf{z}$ yields the update rules (4),

$$z_n = \mathbf{L}_n^\top(\boldsymbol{\Sigma}_{\mathbf{y}})^{-1}\mathbf{L}_n \ \text{for } n = 1, \ldots, N \ , \text{and} \tag{10}$$

$$\gamma_n = \sqrt{\frac{\left[\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\mu}_{\mathbf{x}}(t))_n^2\right]}{z_n}} \ \text{for } n = 1, \ldots, N \ , \tag{11}$$

where $\mathbf{L}_n$ in (10) is the n-th column of the lead field matrix. Final estimates of the variances $\boldsymbol{\gamma}$ and the posterior mean of the brain sources, $\bar{\mathbf{X}}$, are obtained by iterating the updates (10)–(11) until convergence. The resulting algorithm is called majorization-minimization (MM)-Champagne. For a general introduction to MM algorithms see [17].

III. Tight Upper Bound for Low-SNR Regimes

While the convexity-based MM-Champagne algorithm overcomes the high computational cost of the EM-Champagne algorithm, it has been noted that the performance of both algorithms is negatively affected in low-SNR regimes [18]. Here, we define the SNR as $\frac{\mathbb{E}\{||x(t)||^2\}}{\sigma^2}$, where $\mathbb{E}\{||x(t)||^2\}$ and σ^2 denote the power of signal and noise, respectively. In the following, we use that the non-convex log-likelihood cost function (7) has a tight upper bound in low-SNR regimes [19].

Proposition 1. *Eq.* (7) *is the sum of the concave term* $\log|\boldsymbol{\Sigma}_{\mathbf{y}}|$ *and the convex term* $\frac{1}{T}\sum_{t=1}^{\top}\mathbf{y}(t)^\top\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}(t)$*, and the best convex upper bound for* (7) *in low-SNR regimes can be written as*

$$\mathcal{L}_{\text{conv}}(\boldsymbol{\gamma}) = \mathrm{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top) + \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}(t)^\top\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}(t) \ . \tag{12}$$

Proof. The proof follows [19, Proposition 4]. Using that $\mathrm{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top)$ is the best linear approximation for $\log|\boldsymbol{\Sigma}_{\mathbf{y}}|$, we have that, for low SNR regimes, $\log|\boldsymbol{\Sigma}_{\mathbf{y}}| = \mathrm{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top) + \mathcal{O}(\mathrm{SNR})$. This means that (12) is tight when $\mathrm{SNR} \to 0$. □

The derivative of (12) w.r.t. to $\boldsymbol{\gamma}$ is

$$\frac{\partial}{\partial\gamma_n}\mathcal{L}_{\text{conv}}(\boldsymbol{\gamma}) = \frac{\partial}{\partial\gamma_n}\left[\mathrm{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top) + \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}(t)^\top\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{y}(t)\right]$$
$$= \mathbf{L}_n^\top\mathbf{L}_n + \left(-\frac{1}{\gamma_n^2}\right)\left[\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\mu}_{\mathbf{x}}(t))_n^2\right] \ . \tag{13}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

Setting (13) to zero leads to a new update rule for $\boldsymbol{\gamma}$:

$$\gamma_n = \sqrt{\frac{\left[\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\mu}_{\mathbf{x}}(t))_n^2\right]}{\mathbf{L}_n^\top \mathbf{L}_n}} \text{ for } n = 1, \ldots, N\,. \tag{14}$$

We call the algorithm obtained by iterating (4)–(6) and (14), which is derived by replacing the tangent plane $\mathbf{z}^\top\boldsymbol{\gamma} - \mathbf{z}_0$ in (9) by the proposed upper bound $\text{tr}(\mathbf{L}\boldsymbol{\Gamma}\mathbf{L}^\top)$, *LowSNR-MM-Champagne*.

IV. Spatio-temporal Trade-off Theorem

In this section, we mathematically show that the achievable spatial reconstruction performance (in terms of correctly estimating the support of the active brain locations) increases with sample size; thus, there exists a trade-off between the spatial reconstruction performance and the number of measurements. Here we built on recent theory developed for the multiple measurement vector (MMV) case by [18]. Since the EEG inverse problem can also be considered as an MMV problem by assuming samples as multiple measurement vectors, we can exploit the theory of [18] to formulate an error bound for the reconstruction of the spatial support of the active brain sources. We make the following assumptions.

Assumption I (independence): The sources $\mathbf{x}(t), t = 1, \ldots, T$ are *independent and identically distributed* (i.i.d.) zero-mean Gaussian vectors; thus, the prior probability can be written as $\mathbb{P}(\mathbf{X}|\boldsymbol{\gamma}) = \prod_{t=1}^{T}\mathcal{N}(0, \boldsymbol{\Gamma})$. The non-zero variances are bounded as $\gamma_n \in [\gamma_{\min}, \gamma_{\max}], n = 1, \ldots, N$.

Assumption II (common sparsity): The sources $\mathbf{x}(t), t = 1, \ldots, T$ are jointly K-sparse, that is, form a row-sparse matrix, $\mathbf{X}$, with support set $|\mathcal{S}^*| \leq K$. In other words, the locations of the active sources does not change across samples.

Proposition 2. *Perfect spatial reconstruction is achieved, if the algorithm can exactly recover the true support, $\mathcal{S}^*$. Mathematically speaking, a spatial reconstruction error occurs whenever the recovered support does not exactly match to the original support set, $\mathbb{P}(\textit{Support}(\hat{\gamma}) \neq \mathcal{S}^*)$.*

The following theorem holds for the general MMV case [18]:

Theorem 3 ([18])**.** *Assuming the preceding EEG source localization setting and assumptions (I-II) and let $\hat{\boldsymbol{\gamma}}$ be the sparse recovery solution of* EM-Champagne *method. Then, the spatial reconstruction error decreases exponentially with the number of samples available:*

$$\mathbb{P}(\text{Support}(\hat{\boldsymbol{\gamma}}) = \mathcal{S}^*) \geq 1 - \exp(-\frac{\eta}{4}T)\,,$$

provided that the following two conditions are satisfied:

Condition I: The self Khatri-Rao product of the lead-field matrix, $\mathbf{L} \odot \mathbf{L}$, satisfies *the restricted isometry property* (RIP) of order $2k$ with restricted isometry constant $\delta_{2k}^{\odot}$, where the self Khatri-Rao product is defined as $(\mathbf{L} \odot \mathbf{L}) \in \mathbb{R}^{M^2 \times N} = [\ldots, \mathbf{L}_i \otimes \mathbf{L}_i, \ldots]$, for $i = 1, \ldots, N$, and where $\mathbf{L}_i$ denotes the i-th column of $\mathbf{L}$.

Condition II: The number of time samples satisfies $T \geq \left(\frac{c_1 k \log(N)}{\eta}\right)$, where $\eta = c_2(\frac{\gamma_{\min}}{\sigma^2+\gamma_{\max}})^2(1-\delta_{2k}^{\odot})$, where c_1 and c_2 are two universal constants. For a lead field matrix $\mathbf{L}$ with columns normalized to unit-norm (as is a common pre-processing in practice), this leads to the following lower bound for the number of time samples: $T \geq \mathcal{O}(K^{\frac{5}{2}}\log K/M)$.

In practice, both conditions may be violated. Therefore, it is crucial to empirically study the effect of such violations using simulated data. In Section VI-A, we empirically assess the RIP property of the Kathri-Rao product of typical lead field matrices. To deal with the assumption of i.i.d. samples imposed by the theory, we introduce an adaptive whitening scheme for correlated samples (e.g. time-series data), which motivates two additional algorithm variants (see Section V). The reconstruction performance of all algorithms is compared in Section VI-B.

V. Adaptive Temporal Whitening Scheme

Brain processes change rapidly over time. Therefore, in order to take advantage of an increased number of samples while simultaneously satisfying the common sparsity assumption, we need to increase the number of time samples within a fixed duration of time. This, however, will increase the correlation between adjacent measurements and thus violate the i.i.d. assumption. To address this issue, we consider SBL algorithms that can deal with correlated measurements. It can be shown [20] that, by incorporating time correlation, the prior distribution of the sources can be written as $\mathbb{P}(\mathbf{x}^{\text{temp}}|\boldsymbol{\gamma}, \mathbf{B}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_0)$, where $\mathbf{x}^{\text{temp}} = \text{vec}(\mathbf{X}) \in \mathbb{R}^{NT\times 1}$ and $\boldsymbol{\Sigma}_0 = \boldsymbol{\Gamma} \otimes \mathbf{B}$ models the covariance matrix of time-correlated sources, $\mathbf{B} \in \mathbb{R}^{T\times T}$ is a positive definite (P.D.) matrix modeling the (same) temporal correlation structure for all sources, and $\otimes$ denotes the Kronecker product. $\mathbf{B}$ can either be a general P.D. matrix or have a Toeplitz structure, $\mathbf{B}_{i,j} = \beta^{|i-j|}$. The latter case corresponds to modeling the sources as first order auto-regressive (AR(1)) time series: $x_n(t+1) = \beta x_n(t) + \sqrt{1-\beta^2}\xi_n(t),\ n = 1, \ldots, N; t = 1, \ldots, T$ with AR coefficient $\beta \in (-1, 1)$ and innovation noise $\xi_n(t)$.

The log likelihood function can be formulated as [20]:

$$\mathcal{L}(\boldsymbol{\gamma}, \mathbf{B}) = \log|\boldsymbol{\Sigma}_{\mathbf{Y}}^{\text{temp}}| + \left[\left(\mathbf{y}^{\text{temp}}\right)^\top \left(\boldsymbol{\Sigma}_{\mathbf{Y}}^{\text{temp}}\right)^{-1}\mathbf{y}^{\text{temp}}\right] \tag{15}$$

$$\boldsymbol{\Sigma}_{\mathbf{Y}}^{\text{temp}} = \lambda\mathbf{I} + \mathbf{D}\boldsymbol{\Sigma}_0\mathbf{D}^\top\,, \tag{16}$$

where $\mathbf{y}^{\text{temp}} = \text{vec}(\mathbf{Y}) \in \mathbb{R}^{MT\times 1}, \mathbf{D} = \mathbf{L} \otimes \mathbf{I}_T$ and $\boldsymbol{\Sigma}_{\mathbf{Y}}^{\text{temp}}$ is the spatio-temporal version of the model covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}}$ defined in (6). Zhang et. al [21] derived the following update rules to minimize (15) using the EM method:

$$\gamma_n = [\boldsymbol{\Sigma}_{\mathbf{x}}]_{n,n} + \left[(\boldsymbol{\mu}_{\mathbf{x}})_n\mathbf{B}^{-1}(\boldsymbol{\mu}_{\mathbf{x}})_n^\top\right],\ n = 1, \ldots, N \tag{17}$$

$$\hat{\mathbf{B}} = \begin{cases} \sum_{n=1}^{N}\frac{(\boldsymbol{\mu}_{\mathbf{x}})_n^\top(\boldsymbol{\mu}_{\mathbf{x}})_n}{\gamma_n} & \text{for high SNR} \\ \sum_{n=1}^{N}\frac{(\boldsymbol{\mu}_{\mathbf{x}})_n^\top(\boldsymbol{\mu}_{\mathbf{x}})_n}{\gamma_n} + \theta I & \text{for low SNR} \end{cases} \tag{18}$$

$$\mathbf{B} = \hat{\mathbf{B}}/||\hat{\mathbf{B}}||_F\,, \tag{19}$$

where $||\hat{\mathbf{B}}||_F$ denotes the Frobenius-norm of the matrix $\hat{\mathbf{B}}$ and where θ is a positive scalar regularization term ensuring that $\hat{\mathbf{B}}$ is P.D. . Notice that $(\boldsymbol{\mu}_{\mathbf{x}})_n \in \mathbb{R}^T$ is a vector comprising the

ISBN 978-90-827970-1-5 © EURASIP 2018

entire T time series of the n-th brain source, $[\boldsymbol{\mu}_{\mathbf{x}}(t)]_n$, for $t = 1, \ldots, T$. In other words, if $\bar{\mathbf{X}} = \boldsymbol{\Gamma}\mathbf{L}^\top(\boldsymbol{\Sigma}_{\mathbf{y}})^{-1}\mathbf{Y}$ denotes the posterior mean of the sources recovered by EM-Champagne as defined in (4), then $(\boldsymbol{\mu}_{\mathbf{x}})_n$ is the n-th row of $\bar{\mathbf{X}}$. The algorithm that is obtained by iterating (17)-(19) until convergence is called *T-EM-Champagne*.

In the AR setting, $\boldsymbol{\gamma}$ is updated as in (17) and the AR-coefficients are optimized using steepest descent (SD) [22]:

$$\beta^{(new)} = \beta + \alpha \mathrm{tr}[(\boldsymbol{\Gamma} \otimes (\mathbf{BFB}))\, \boldsymbol{\Sigma}_{\mathbf{x}} - N(\mathbf{BF})]\,, \qquad (20)$$

where α denotes the optimization step size and $\mathbf{F} = \frac{\partial(\mathbf{B}^{-1})}{\partial\beta}$ is obtained numerically throughout the algorithm [23].

Note that the *Euclidean distance* $\left[(\boldsymbol{\mu}_{\mathbf{x}})_n(\boldsymbol{\mu}_{\mathbf{x}})_n^\top\right]$ in the independent setting (8) changes into a *Mahalanobis distance* $\left[(\boldsymbol{\mu}_{\mathbf{x}})_n\mathbf{B}^{-1}(\boldsymbol{\mu}_{\mathbf{x}})_n^\top\right]$ in the correlated setting (17). This can be shown to be equivalent to performing a temporal whitening of the sources in each iteration of the algorithm using the current estimate of the matrix $\mathbf{B}^{-\frac{1}{2}}$ [20, Page 74]. The resulting quasi-i.i.d measurements, which satisfy Condition II above, are then treated according to the standard Champagne algorithm.

VI. Numerical Simulations

A. Influence of the RIP of the lead field

Since the RIP of a general matrix cannot be computed in polynomial time [24], we compared the recovery performance using a realistic EEG lead field likely not exhibiting RIP to the setting in which random Gaussian matrices (satisfying RIP) are used as forward models. In each repetition of the experiment, we generated a random Gaussian matrix, $\mathbf{X} \in \mathbb{R}^{2004\times T}$ representing the brain activity of 2004 brain sources for different sizes $T = 10, 20, 50, 100$ and 200 i.i.d. samples. Apart from $K = 3$ randomly selected rows representing the active brain sources all entries of $\mathbf{X}$ were set to zero. The brain sources $\mathbf{X}$ were mapped to pseudo-EEG measurements at 58 sensors, $\mathbf{Y}$, using either a random Gaussian matrix $\boldsymbol{\Phi} \in \mathbb{R}^{58\times 2004}$ or a realistic lead field matrix, $\mathbf{L} \in \mathbb{R}^{58\times 2004}$. This lead field was generated using the New York Head model [2] taking into account the realistic anatomy and electrical tissue conductivities of an average human head. Finally, i.i.d. Gaussian noise $\mathbf{E}$ was added to the measurements, so that the overall SNR (as defined in Section III) was 0 dB. See [25] for a description of a general simulation framework. We applied EM-Champagne and measured its performance using three different measures, the *mean squared error* (MSE), the *earth mover's distance* (EMD), and the correlation between the original and reconstructed sources, $\hat{\mathbf{X}}$ and $\mathbf{X}$. The MSE was defined as $\mathrm{MSE} = \frac{||\hat{\mathbf{X}}-\mathbf{X}||_F^2}{||\mathbf{X}||}$. The EMD measures the cost needed to map a two probability distributions defined on the same metric domain (in this case, the power of the true and estimated source activations defined defined on the cortical surface of the brain) into each other, see [5], [26]. EMD was normalized to $[0, 1]$. The correlation between simulated and reconstructed source time courses was assessed as the mean of the absolute correlations obtained for each source, after optimally matching simulated and reconstructed sources. Each simulation was carried out 100 times using different instances of $\mathbf{X}$ and $\mathbf{E}$, and the mean and standard error of the mean (SEM) of each performance measure across repetitions was calculated.

TABLE I
Source reconstruction performance (mean ± SEM) of EM-Champagne for data generated by a realistic lead field matrix as well as a random Gaussian matrix under variation of the number of time samples. Performance was measured in terms of the mean squared error (MSE), earth mover's distance (EMD) and time-course correlation (Corr).

		MSE	EMD	Corr
Random matrix	$T = 10$	0.17 ± 0.004	0.201 ± 0.0029	0.970 ± 0.0011
	$T = 20$	0.11 ± 0.002	0.181 ± 0.0021	0.971 ± 0.0006
	$T = 50$	0.07 ± 0.001	0.159 ± 0.0019	0.973 ± 0.0004
	$T = 100$	0.06 ± 0.001	0.143 ± 0.0018	0.973 ± 0.0003
	$T = 200$	0.05 ± 0.000	0.125 ± 0.0013	0.974 ± 0.0002
Lead field	$T = 10$	0.77 ± 0.037	0.177 ± 0.0041	0.936 ± 0.0022
	$T = 20$	0.45 ± 0.028	0.130 ± 0.0036	0.933 ± 0.0029
	$T = 50$	0.23 ± 0.016	0.086 ± 0.0024	0.947 ± 0.0018
	$T = 100$	0.14 ± 0.008	0.056 ± 0.0017	0.953 ± 0.0017
	$T = 200$	0.11 ± 0.005	0.039 ± 0.0013	0.957 ± 0.0015

Table I demonstrates that the source reconstruction performance according to all measures increases significantly with the number of samples for both the realistic lead field and a random Gaussian forward matrix. A random Gaussian forward model leads to a lower MSE than the realistic lead field, while the opposite relationship is observed for the EMD. Notably, exact recovery of the support, as quantified by the theory for RIP matrices and measured by the MSE, is rarely necessary in practice, and a reconstruction close (in terms of the Euclidean distance in 3D brain space) to the true solution (as measured by the EMD) is more relevant. The EMD performance achieved using the realistic lead field is, therefore, encouraging.

B. Comparison of algorithms for correlated samples

We investigated the ability of the various algorithms to reconstruct source activations from correlated samples. To this end, we generated time-dependent non-stationary sources in the frequency domain. For this purpose, $K = 2$ source locations were selected. Source activations at these locations were generated as lowpass-filtered white noise using a second order digital Butterworth filter with cutoff frequency chosen at random between 10 and 15 Hz, assuming a sampling rate of 200 Hz. The number of time samples was fixed to $T = 20$.

For this simulation, we used the same realistic lead field as in Section VI-A. We again added i.i.d. Gaussian noise at SNR = 0 dB to the measurement, and we evaluated the MSE, EMD and Corr performance metrics across 100 repetitions of the experiment. We compared the performance of LowSNR-MM-Champagne, T-EM-Champagne and AR-EM-Champagne presented here with commonly used baseline methods including EM- and MM-Champagne [14], [15], eLORETA [4], and S-FLEX [8]. The noise variance, σ^2, was assumed to be known (e.g., from a baseline measurement), and the regularization parameters of all methods were set accordingly. The pruning threshold γ_{thresh} was set to $\gamma_{\text{thresh}} = 10^{-3}$ for all Champagne variants in order to ensure a fair comparison. Convergence was defined for all algorithms if the relative improvement

ISBN 978-90-827970-1-5 © EURASIP 2018

TABLE II
SOURCE RECONSTRUCTION PERFORMANCE (MEAN ± SEM) OF THE PROPOSED ALGORITHMS IN COMPARISON TO BASELINE METHODS. PERFORMANCE WAS MEASURED IN TERMS OF MEAN-SQUARED ERROR (MSE), EARTH MOVER'S DISTANCE (EMD), AND TIME-COURSE CORRELATION (CORR). PERFORMANCE SCORES SIGNIFICANTLY OUTPERFORMING ALL OTHERS ARE HIGHLIGHTED IN BOLD FACE.

	MSE	EMD	Corr
eLORETA	1.08 ± 0.007	0.329 ± 0.0037	0.954 ± 0.0015
S-FLEX	0.19 ± 0.004	0.207 ± 0.0033	0.964 ± 0.0011
EM-Champagne	0.15 ± 0.003	0.162 ± 0.0027	0.965 ± 0.0012
MM-Champagne	0.14 ± 0.003	0.161 ± 0.0027	0.965 ± 0.0012
LowSNR-MM-Champagne	0.14 ± 0.003	0.141 ± 0.0025	0.967 ± 0.0011
T-EM-Champagne	0.09 ± 0.003	**0.055 ± 0.0023**	0.968 ± 0.0011
AR-EM-Champagne	0.10 ± 0.004	**0.044 ± 0.0024**	0.955 ± 0.0032

of the objective function was less than 10^{-8}. A maximum of 2000 iterations was carried out in case no convergence was reached. As can be seen from Table II, LowSNR-MM-Champagne is on par with the conventional EM- and MM-Champagne variants in terms of the MSE, but is characterized by improved localization performance in terms of the EMD. By exploiting temporal correlations in the signals, T-EM-Champagne and AR-EM-Champagne further outperform all competing methods in terms of the MSE and EMD.

VII. CONCLUSION

In this work, we have proposed sparse Bayesian learning algorithms for solving the EEG inverse problem, which are suitable in low-SNR settings and if the available samples are correlated. Using realistic simulations, we have shown that these algorithms can exploit the temporal structure in neurophysiological time series to achieve better reconstruction performance than the state-of-the-art. In the future, we aim to provide a theoretical analysis of the convergence properties of the proposed methods and to demonstrate their efficacy on real data.

ACKNOWLEDGEMENT

AH acknowledges support from Berlin International Graduate School in Model and Simulation based Research (BIMoS) as well as Berlin Mathematical School (BMS). He further wish to thank Gitta Kutyniok and Klaus-Robert Müller for valuable suggestions for improving the paper.

REFERENCES

[1] Sylvain Baillet, John C Mosher, and Richard M Leahy, "Electromagnetic brain mapping," *IEEE Signal processing magazine*, vol. 18, no. 6, pp. 14–30, 2001.

[2] Yu Huang, Lucas C Parra, and Stefan Haufe, "The new york heada precise standardized volume conductor model for EEG source localization and tes targeting," *NeuroImage*, vol. 140, pp. 150–162, 2016.

[3] Irina F Gorodnitsky, John S George, and Bhaskar D Rao, "Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm," *Electroencephalography and clinical Neurophysiology*, vol. 95, no. 4, pp. 231–251, 1995.

[4] Roberto D Pascual-Marqui, "Discrete, 3D distributed, linear imaging methods of electric neuronal activity. part 1: exact, zero error localization," *arXiv preprint arXiv:0710.3341*, 2007.

[5] Stefan Haufe, Vadim V Nikulin, Andreas Ziehe, Klaus-Robert Müller, and Guido Nolte, "Combining sparsity and rotational invariance in EEG/MEG source reconstruction," *NeuroImage*, vol. 42, no. 2, pp. 726–738, 2008.

[6] Alexandre Gramfort, Matthieu Kowalski, and Matti Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in medicine and biology*, vol. 57, no. 7, pp. 1937, 2012.

[7] Daniel Strohmeier, Yousra Bekhti, Jens Haueisen, and Alexandre Gramfort, "The iterative reweighted mixed-norm estimate for spatio-temporal MEG/EEG source reconstruction," *IEEE transactions on medical imaging*, vol. 35, no. 10, pp. 2218–2228, 2016.

[8] Stefan Haufe, Ryota Tomioka, Thorsten Dickhaus, Claudia Sannelli, Benjamin Blankertz, Guido Nolte, and Klaus-Robert Müller, "Large-scale EEG/MEG source localization with spatial flexibility," *NeuroImage*, vol. 54, no. 2, pp. 851–859, 2011.

[9] Alexandre Gramfort, Daniel Strohmeier, Jens Haueisen, Matti S Hämäläinen, and Matthieu Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, 2013.

[10] Sebastián Castaño-Candamil, Johannes Höhne, Juan-David Martínez-Vargas, Xing-Wei An, German Castellanos-Domínguez, and Stefan Haufe, "Solving the EEG inverse problem based on space–time–frequency structured sparsity constraints," *NeuroImage*, vol. 118, pp. 598–612, 2015.

[11] Rémi Gribonval, "Should penalized least squares regression be interpreted as maximum a posteriori estimation?," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2405–2410, 2011.

[12] Wei Wu, Srikantan Nagarajan, and Zhe Chen, "Bayesian machine learning: EEG\MEG signal processing measurements," *IEEE Signal Processing Magazine*, vol. 33, no. 1, pp. 14–36, 2016.

[13] Michael E Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244, 2001.

[14] Kensuke Sekihara and Srikantan S Nagarajan, *Electromagnetic brain imaging: a Bayesian perspective*, Springer, 2015.

[15] David P Wipf, Julia P Owen, Hagai T Attias, Kensuke Sekihara, and Srikantan S Nagarajan, "Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG," *NeuroImage*, vol. 49, no. 1, pp. 641–655, 2010.

[16] David P Wipf and Bhaskar D Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on signal processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[17] Ying Sun, Prabhu Babu, and Daniel P Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2017.

[18] Saurabh Khanna and Chandra R Murthy, "On the support recovery of jointly sparse Gaussian sources using sparse Bayesian learning," *arXiv preprint arXiv:1703.04930*, 2017.

[19] Saeid Haghighatshoar and Giuseppe Caire, "Massive MIMO channel subspace estimation from low-dimensional projections," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 303–318, 2017.

[20] Zhilin Zhang, *Sparse signal recovery exploiting spatiotemporal correlation*, University of California, San Diego, 2012.

[21] Zhilin Zhang and Bhaskar D Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

[22] Zhilin Zhang and Bhaskar D Rao, "Sparse signal recovery in the presence of correlated multiple measurement vectors," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 3986–3989.

[23] Benyuan Liu, Zhilin Zhang, Gary Xu, Hongqi Fan, and Qiang Fu, "Energy efficient telemonitoring of physiological signals via compressed sensing: A fast algorithm and power consumption evaluation," *Biomedical Signal Processing and Control*, vol. 11, pp. 80–88, 2014.

[24] Afonso S Bandeira, Edgar Dobriban, Dustin G Mixon, and William F Sawin, "Certifying the restricted isometry property is hard," *IEEE transactions on information theory*, vol. 59, no. 6, pp. 3448–3450, 2013.

[25] Stefan Haufe and Arne Ewald, "A simulation framework for benchmarking EEG-based brain connectivity estimation methodologies," *Brain topography*, pp. 1–18, 2016.

[26] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int J Comput Vision*, vol. 40, pp. 99–121, 2000.

ISBN 978-90-827970-1-5 © EURASIP 2018