BAYESIAN INFERENCE FOR BRAIN SOURCE IMAGING WITH JOINT ESTIMATION OF STRUCTURED LOW-RANK NOISE

Sanjay Ghosh[1], Chang Cai[2], Yijing Gao[1], Ali Hashemi[3], Stefan Haufe[3], Kensuke Sekihara[4], Fellow, IEEE, Ashish Raj[1], and Srikantan S. Nagarajan[1][*], Fellow, IEEE

[1]University of California San Francisco, USA
[2]Central China Normal University, China
[3]Technical University of Berlin, Germany
[4]Signal Analysis Inc., Japan

ABSTRACT

The inverse problem in brain source imaging is the reconstruction of brain activity from non-invasive recordings of electroencephalography (EEG) and magnetoencephalography (MEG). One key challenge is the efficient recovery of sparse brain activity when the data is corrupted by structured noise that is low-rank noise. This is often the case when there are a few active sources of environmental noise and the MEG/EEG sensor noise is highly correlated. In this paper, we propose a novel robust empirical Bayesian framework which provides us a tractable algorithm for jointly estimating a low-rank noise covariance and brain source activity. Specifically, we use a factor analysis model for the structured noise, and infer a sparse set of variance parameters for source activity, while performing Variational Bayesian inference for the noise. One key aspect of this algorithm is that it does not require any additional baseline measurements to estimate the noise covariance from the sensor data. We perform exhaustive experiments on both simulated and real datasets. Our algorithm achieves superior performance as compared to several existing benchmark algorithms.

Index Terms— EEG, MEG, brain source imaging, low-rank noise.

1. INTRODUCTION

Electromagnetic brain imaging is the reconstruction of brain activity from non-invasive recordings of magnetic fields and electric potentials. Electroencephalography (EEG) and magnetoencephalography (MEG) are widely used techniques to study the function of human brain [1, 2]. Efficient estimation of the brain activity on the cortex surface is important for neuroscience research and clinical diagnosis [3]. It is crucial to determine, to the extent possible, where and when neurophysiological activity is occurring. Thus, the inverse problem turns out to be estimating both the spatial location of the sources and the time course of brain activity from the EEG/MEG measurements. Several methods have been introduced to solve this ill-posed inverse problem of brain source imaging. The regularization based methods attempt to enforce prior information on the source signal. The minimum-norm estimation algorithm (MNE) [4] tries to minimize the L_2 norm of the solution favoring low power of the brain activity. Other variants of MNE include the weighted MNE (wMNE) [5], low resolution brain electromagnetic tomography (LORETA) [6], and standardized LORETA (sLORETA) [7], etc. Another broad class of approach is to consider Bayesian techniques [8, 9, 10, 11, 12, 13, 14] with appropriate priors assigned to the model parameters. Recently introduced Champagne algorithm [8], a novel tomographic source reconstruction algorithm derived in an empirical Bayesian fashion with incorporation of deep theoretical ideas about sparse-source recovery from noisy, constrained measurements. One of active current research focus is to improve upon the reconstruction methods under realistic noise assumption such as low-rank statistical model [11]. An enduring challenge is to jointly remove the noise of the sensor arrays and recover the brain signal. The noise statistics in the model plays a crucial role in the success of sparse source recovery. In particular, the noise covariance in sensor data is what dictates the working of Bayesian frameworks. The existing works include noise covariance matrix of both diagonal and full structure. In this work, we consider one of most realistic assumptions - low rank noise covariance. This is often the case when there are a few active sources of environmental noise or the MEG/EEG sensors are highly correlated. To best of our knowledge, no existing algorithm has addressed the brain source estimation problem under low-rank noise covariance.

In this paper, we propose a novel robust empirical Bayesian framework for brain source imaging under low-rank noise assumption. It provides us a tractable algorithm for iteratively estimating the noise covariance and the brain source activity. The proposed algorithm is found to be quite robust to initialization and computationally efficient. The proposed algorithm

---

does not require any additional baseline measurements to estimate noise covariance from sensor data. We further show that the new algorithm produces competitive performance with benchmark method on real MEG data and able to resolve distinct and functionally relevant brain areas.

This paper is outlined as follows. In Section 2, we introduce the inverse problem in brain source imaging under low-rank noise. This is followed by the proposed Bayesian reconstruction algorithm. Next, in Section 3, we show the experimental results of our approach on synthetic and real MEG data. Finally, we conclude in Section 4.

2. THEORY

2.1. Probabilistic Generative Model

In brain source imaging setup, brain activity originates from a number of electric current dipoles, where the location, orientation, and the magnitude of each dipole determine the recorded signal from the EEG electrodes. We consider the following forward model between measurements and brain sources:

$$y = Lx + \varepsilon, \qquad (1)$$

where $y \in \mathbb{R}^{M \times K}$ is the EEG/MEG sensor data measured at M sensors at K time-points, $x \in \mathbb{R}^{N \times K}$ is the underlined brain signals where N is the number of voxels. The lead-field matrix $L \in \mathbb{R}^{M \times N}$ represents the propagation of electromagnetic field from a particular source location to the EEG/MEG sensors. The additive measurement noise ε is drawn from $\mathcal{N}(0, \Lambda^{-1})$, where Λ is the precision of the noise.

In this work, the noise is modelled in terms of the covariance of the noise statistics as the follows:

$$\Lambda = \sum_{i=1}^{r} n_i n_i^T + \sigma \mathbb{I}, \qquad (2)$$

where $\mathbb{I} \in \mathbb{R}^{M \times M}$ is an identity matrix, σ is a scaling factor, and r represents the rank of the noise statistics. The above low-rank noise is simulated by placing correlated passive (correlated) source of noise at close proximity of MEG/EEG sensors. Notice that for $\sigma = 0$, the noise co-variance Λ in either (2) is purely low-rank; it's rank is r. For $\sigma \neq 0$ in (2), the noise is a mixture of low-rank and homoscedastic statistics.

We assume zero-mean Gaussian prior for the underlying source: $x_k \in \mathbb{R}^{N \times 1} \sim \mathcal{N}(0, \Phi^{-1})$, $k = 1, \ldots, K$, where $\phi^{-1} = [\phi_1^{-1}, \ldots, \phi_N^{-1}]^\top$ contains N distinct unknown variances for the brain sources and $\Phi^{-1} = \mathrm{diag}(\phi^{-1})$. We consider to solve the inverse problem by using Bayesian learning framework; by finding the maximum a-posterior probability (MAP) solution. The posterior probability $p(x_k|y_k)$ can be derived by using Bayes' rule

$$p(x_k|y_k) \propto p(y_k|x_k)p(x_k),$$

where $p(y_k|x_k) = \mathcal{N}(Lx_k, \Lambda^{-1})$ and $p(x_k) = \mathcal{N}(x_k|0, \Phi^{-1})$. It is straightforward to show that posterior probability $p(x_k|y_k)$

is also Gaussian. Suppose, the posterior probability takes the following form:

$$p(x_k|y_k) = \mathcal{N}(x_k|\overline{x}_k, \Gamma^{-1}),$$

where $\overline{x}_k$ is the posterior mean, and Γ is the posterior precision matrix. Further it can be shown that:

$$\Gamma = \Phi + L^T \Lambda L,$$

$$\overline{x}_k = \Gamma^{-1} L^T \Lambda y_k. \qquad (3)$$

2.2. Proposed Bayesian Inference Algorithm

Notice that we need both Φ and Λ to compute the Bayesian estimate of unknown x is computed using (3). However, since we have access to only y, the idea is to compute the complete data likelihood and average over it with the posterior probability as follows:

$$\Theta(\Phi, \Lambda) = \mathbb{E}\left[\log p(y, x|\Phi, \Lambda)\right]$$

where the expectation $[\cdot]$ is taken with respect to $p(x|y)$. The complete data likelihood is expressed as

$$\log p(y, x|\Phi, \Lambda) = \frac{K}{2}\log|\Phi| - \frac{1}{2}\sum_{k=1}^{K} x_k^T \Phi x_k$$
$$+ \frac{K}{2}\log|\Lambda| - \frac{1}{2}\sum_{k=1}^{K}(y_k - Lx_k)^T \Lambda (y_k - Lx_k). \qquad (4)$$

Update of Source Parameter: Here we obtain the optimal $\hat{\Phi}$ by maximizing the average likelihood $\Theta(\Phi, \Lambda)$ for a fixed Λ as follows: $\frac{\partial}{\partial \Phi}\Theta(\Phi, \Lambda) = 0$. With further derivation, we can show that it follows MacKay update equation:

$$\phi_j = \frac{[\Gamma^{-1} L^T \Lambda L]_{j,j}}{\frac{1}{K}\sum_{k=1}^{K} \overline{x}_k^2(j)}, \qquad (5)$$

where $[\cdot]$ indicates (j, j)-th element of a matrix.

Update of Noise Parameter: In this work, we propose to estimate $\widehat{\Lambda}$ via factor analysis of the following residual noise at l-th iteration:

$$\eta_k^{(l)} = y_k - Lx_k^{(l)}. \qquad (6)$$

The idea is to obtain the covariance of $\eta_k^{(l)}$ using Bayesian factor analysis. In other words, we assume a factored decomposition of η_k as follows:

$$\eta_k = A\nu_k + \delta,$$

where the residual noise at k-th time-point is $\eta_k \in \mathbb{R}^{M \times 1}$, $A \in \mathbb{R}^{M \times P}$, ν_k is a P-dimensional column vector, and δ is modeling noise. Notice that we drop the iteration symbol l for simplification of notations.

Algorithm 1: Low-rank Champagne.

Input: Sensor data y; lead-field matrix L.
Output: Brain source activity x.
1 Set random initialization: Φ and Λ.
2 Initialize $l \leftarrow 1$.
3 **repeat**
4 | Compute $\Gamma = L^T \Lambda L + \Phi$.
5 | Compute $\overline{x} = \Gamma^{-1} L^T \Lambda y$.
6 | Estimate Φ using (5).
7 | Compute $\eta = (y - L\overline{x})$.
8 | Estimate $\Lambda \leftarrow$ factor-analysis(η).
9 | $l \leftarrow l + 1$.
10 **until** *stopping condition:* $l = l_{max}$

We assume the prior probability distribution of the factor ν_k is assumed to be the zero-mean Gaussian with its precision matrix equal to the identity matrix,

$$p(\nu_k) = \mathcal{N}(\nu_k | \mathbf{0}, \mathbf{I})$$

The factor activity is assumed to be independent across time. Thus, the joint prior distribution:

$$p(\nu) = \prod_{k=1}^{K} p(\nu_k) = \prod_{k=1}^{K} \mathcal{N}(\nu_k | \mathbf{0}, \mathbf{I}),$$

where $\mathbf{I}$ is an identity matrix of size $(P \times P)$. The modeling noise δ is assumed to be Gaussian with the mean of zero:

$$p(\delta) = \mathcal{N}(\nu_k | \mathbf{0}, \Omega^{-1}),$$

where Ω is a diagonal precision matrix. Further, with few steps of mathematical derivations, we obtain the Bayesian estimate of residual noise covariance as follows:

$$\Lambda = \left(AA^T + \Omega^{-1} \right). \tag{7}$$

Finally we summarize the proposed reconstruction method in Algorithm 1. The key aspect is the novel way of estimating covariance of residual noise within each iteration of the proposed Bayesian algorithm. We note that our work is motivated by one recent empirical Bayesian source reconstruction algorithm is called Champagne [8]. We refer the proposed algorithm as Low-rank Champagne.

3. RESULTS

3.1. Simulations and Real Data

We generate source signal data by simulating dipole sources with fixed orientation. The damped sinusoidal time courses with frequencies sampled randomly between 1 and 75 Hz are created as voxel source time activity. The time-courses are
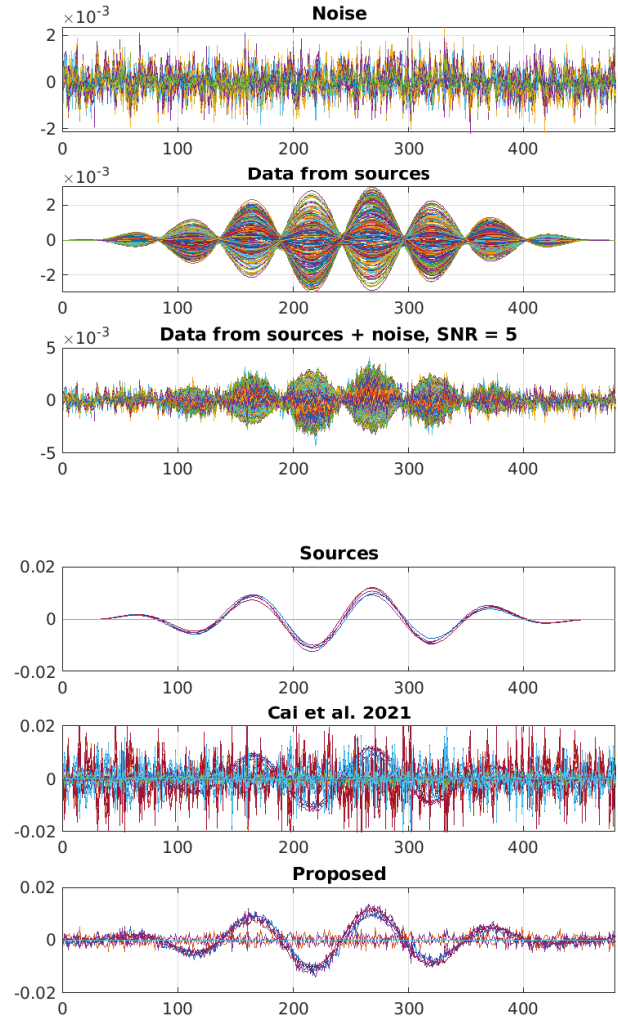

Fig. 1. Reconstruction of active brain sources under simulated brain noise of rank $r = 5$ and SNR 5 dB.

then projected to the sensors using the lead-field matrix generated by the forward model. We consider 271 MEG sensors and a single shell spherical model [15] implemented in SPM12 http://www.fil.ion.ucl.ac.uk/spm at the default spatial resolution of 8196 voxels at approximately 5 mm spacing. We set time period as 480 samples with source activities of interest and noise activity.

Real MEG data we used in our experiments was acquired in the Biomagnetic Imaging Laboratory at University of California, San Francisco (UCSF) using a CTF Omega 2000 whole-head MEG system from VSM MedTech (Coquitlam, BC, Canada) with 1200 Hz sampling rate. The lead field for each subject was calculated in NUTMEG software using a single-sphere head model (two spherical orientation lead fields) and an 8 mm voxel grid. the lead field was calculated using a three-shell spherical model at the coarse resolution.
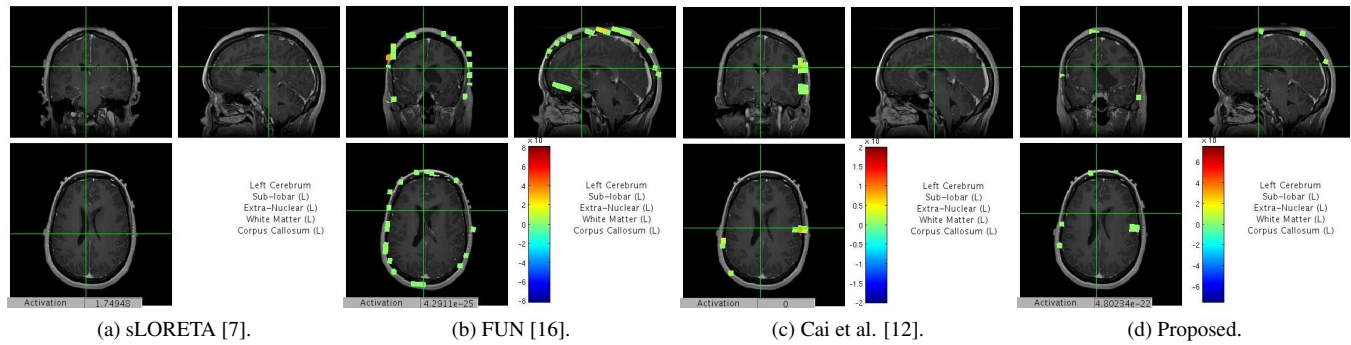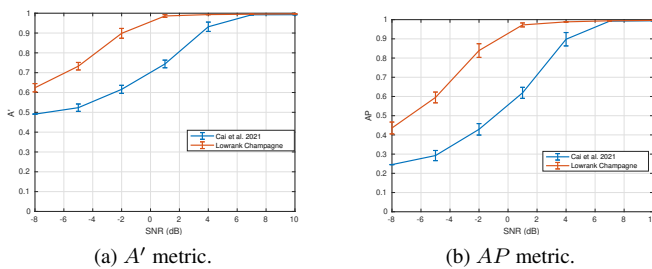
(a) sLORETA [7]. (b) FUN [16]. (c) Cai et al. [12]. (d) Proposed.

Fig. 2. Auditory evoked field (AEF) results.

(a) A' metric. (b) AP metric.

Fig. 3. Performance with varying SNR (dB) for rank $r = 10$.
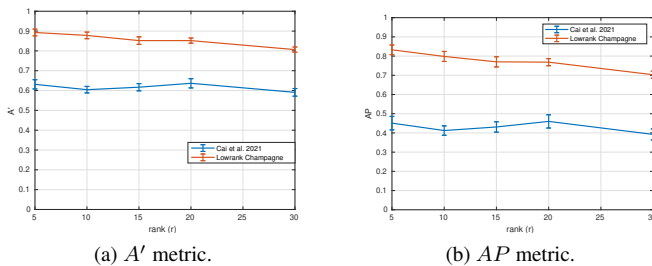

(a) A' metric. (b) AP metric.

Fig. 4. Performance with varying rank for -2 dB SNR.

3.2. Benchmarks and Performance Quantification

In this paper, we choose noise-learned Champagne [Cai et al. 2021] as the key benchmarks to compare with. Two other we compare are sLORETA [7] and full structure noise (FUN) learning method [16]. The performance of simulated brain source reconstruction is evaluated based on response receiver operator characteristics (FROC) [12]. We compute A' and aggregated performance (AP) metrics to quantify performance of source localization and reconstruction. For both, a higher value indicates better result.

3.3. Simulation Results

A brain source reconstruction example is shown in Fig. 1. We simulated a time-course with 5 active sources randomly

placed with 3D brain. We also generated low-rank noise as demonstrated in (2). The simulated brain signal is then projected to measurement space by applying the leadfield matrix and then the noise is added at SNR = 5dB. Finally, we reconstruct the underlined time-course using the proposed algorithm. We also compared the reconstruction performance with recent method by Cai et al. [12]. It is visually evident that our method performs better that the state-of-the-art method.

Performance results versus SNR (for a fixed rank) for the respective algorithms are plotted in Fig. 3. Reconstruction performance is evaluated for five randomly seeded dipolar sources with an inter-source correlation coefficient of 0.99. We also study the performance with varying the rank of the noise covariance for fixed SNR level in Fig. 4. It is clear that proposed *Low-rank Champagne* outperforms in both setting in terms of both A' and AP metrics.

3.4. Results on real datasets

The localization result for AEF (Auditory Evoked Fields) data of a single representative is shown in Fig 2. We also compare the performance with sLORETA, FUN, and Cai et al. [12]. Our method is able to localize bilateral auditory activity which is supposed to be in Heschl's gyrus, the location of primary auditory cortex. In contrast, sLORETA performance is not as good because of the presence of correlated sources.

4. DISCUSSION

In this paper, we proposed a new method for MEG source reconstruction and localization. The algorithm is able to estimate sensor noise from observed data without the need for additional pre-stimulus or baseline data. To be the best of our understanding, the improved performance of this algorithm arises from the efficient the method of estimating the noise statistics via factor analysis of residual component. In summary, this algorithm displays significant advantages over many existing benchmark algorithms for electromagnetic brain source imaging.

5. REFERENCES

[1] K. Sekihara and S. S. Nagarajan, *Adaptive spatial filters for electromagnetic brain imaging*. Springer Science & Business Media, 2008.

[2] ——, *Electromagnetic brain imaging: a bayesian perspective*. Springer, 2015.

[3] C. Phillips, M. D. Rugg, and K. J. Friston, "Anatomically informed basis functions for eeg source localization: combining functional and anatomical constraints," *NeuroImage*, vol. 16, no. 3, pp. 678–695, 2002.

[4] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical & biological engineering & computing*, vol. 32, no. 1, pp. 35–42, 1994.

[5] A. M. Dale and M. I. Sereno, "Improved localizadon of cortical activity by combining eeg and meg with mri cortical surface reconstruction: a linear approach," *Journal of cognitive neuroscience*, vol. 5, no. 2, pp. 162–176, 1993.

[6] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain," *International Journal of psychophysiology*, vol. 18, no. 1, pp. 49–65, 1994.

[7] R. D. Pascual-Marqui *et al.*, "Standardized low-resolution brain electromagnetic tomography (sloreta): technical details," *Methods Find Exp Clin Pharmacol*, vol. 24, no. Suppl D, pp. 5–12, 2002.

[8] 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*, vol. 49, no. 1, pp. 641–655, 2010.

[9] K. Liu, Z. L. Yu, W. Wu, Z. Gu, J. Zhang, L. Cen, S. Nagarajan, and Y. Li, "Bayesian electromagnetic spatio-temporal imaging of extended sources based on matrix factorization," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2457–2469, 2019.

[10] V. P. Oikonomou and I. Kompatsiaris, "A novel bayesian approach for eeg source localization," *Computational Intelligence and Neuroscience*, vol. 2020, 2020.

[11] F. Liu, L. Wang, Y. Lou, R.-C. Li, and P. L. Purdon, "Probabilistic structure learning for eeg/meg source imaging with hierarchical graph priors," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 321–334, 2020.

[12] C. Cai, A. Hashemi, M. Diwakar, S. Haufe, K. Sekihara, and S. S. Nagarajan, "Robust estimation of noise for electromagnetic brain imaging with the champagne algorithm," *NeuroImage*, vol. 225, p. 117411, 2021.

[13] A. Hashemi, Y. Gao, C. Cai, S. Ghosh, K.-R. Müller, S. Nagarajan, and S. Haufe, "Efficient hierarchical bayesian inference for spatio-temporal regression models in neuroimaging," *Advances in Neural Information Processing Systems*, vol. 34, pp. 24 855–24 870, 2021.

[14] A. Hashemi, C. Cai, G. Kutyniok, K.-R. Müller, S. S. Nagarajan, and S. Haufe, "Unification of sparse bayesian learning algorithms for electromagnetic brain imaging with the majorization minimization framework," *NeuroImage*, vol. 239, p. 118309, 2021.

[15] H. Hallez, B. Vanrumste, R. Grech, J. Muscat, W. De Clercq, A. Vergult, Y. D'Asseler, K. P. Camilleri, S. G. Fabri, S. Van Huffel *et al.*, "Review on solving the forward problem in EEG source analysis," *Journal of neuroengineering and rehabilitation*, vol. 4, no. 1, p. 46, 2007.

[16] A. Hashemi, C. Cai, Y. Gao, S. Ghosh, K.-R. Müller, S. S. Nagarajan, and S. Haufe, "Joint learning of full-structure noise in hierarchical bayesian regression models," *IEEE Transactions on Medical Imaging*, 2022.