



# HHS Public Access

Author manuscript

*Conf Proc IEEE Eng Med Biol Soc.* Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:

*Conf Proc IEEE Eng Med Biol Soc.* 2019 July ; 2019: 4803–4806. doi:10.1109/EMBC.2019.8856953.

## A Novel Scanning Algorithm for MEG/EEG imaging using Covariance Partitioning and Noise Learning

Chang Cai<sup>1</sup>, Kensuke Sekihara<sup>2</sup>, Srikantan S Nagarajan<sup>1</sup>

<sup>1</sup>Biomagnetic Imaging Lab, UCSF, USA.

<sup>2</sup>Signal Analysis Inc.,

### Abstract

In this paper, we present a novel scanning algorithm, called Covariance Optimization Garnering Noise for Active Cancellation (COGNAC), for magnetoencephalography (MEG) and electroencephalography (EEG) source localization. COGNAC uses a probabilistic graphical generative model for describing sensor data. This novel generative model partitions contributions to sensor data from sources at a particular scan location and from sources outside the scan location, with corresponding multi-resolution variance parameters that are estimated from data. Maximizing a convex upper bound on the marginal likelihood of the data under this generative model results in a cost function that can be optimized efficiently. Importantly, this generative model enables learning of sensor noise without the need for additional baseline or pre-stimulus data. The resulting inference algorithm is quite robust to reconstruction of highly correlated sources and to the effect of high levels of interference and noise sources. Algorithm performance was compared to representative benchmark algorithms on both simulated and real brain activity. In simulations, performance of our novel algorithm is consistently superior to benchmarks. We also demonstrate that the new algorithm is robust to correlated brain activity present in real MEG/EEG data.

### I. INTRODUCTION

Magnetoencephalography (MEG) and electroencephalography (EEG) are two popular non-invasive techniques for detecting brain activity by measuring magnetic fields or electric potentials from on or near the scalp surface with excellent temporal resolution. Brain activity can then be reconstructed from MEG/EEG recordings. The reconstruction of brain sources is a challenging problem because it involves solving for unknown brain activity across thousands of voxels from the recordings of just a few hundred sensors. To circumvent this nonuniqueness, various estimation procedures incorporate prior knowledge and constraints about source characteristics.

Most of the source reconstruction algorithms can be viewed in a Bayesian framework [1]. This perspective is useful because at a high level, the prior distribution, implicitly or explicitly imposed, can be used to differentiate and compare the various source localization

methods. Recently, we have developed Champagne, a novel tomographic source reconstruction algorithm, that is derived in an empirical Bayesian and incorporates deep theoretical ideas about sparse-source recovery from noisy, constrained measurements. Champagne improves upon existing methods of source reconstruction in terms of reconstruction accuracy, robustness, and computational efficiency [2]. Experiments with preliminary simulated and real data, presented in [3], show that compared to other commonly-used source localization algorithms, Champagne is more robust to correlated sources and noisy data. However, baseline data is necessary for accurate noise estimation and reconstruction of source activity. In certain source configurations, such as in the resting state, such baseline or pre-stimulus information is not available for analysis. Thus, it is unclear how reconstruction methodologies such as Champagne should be implemented in these scenarios.

In order to solve the problem above, we present a novel Bayesian scanning algorithm, called Covariance Optimization Garnering Noise for Active Cancellation (COGNAC), for magnetoencephalography (MEG) and electroencephalography (EEG) source localization, which uses a probabilistic graphical generative model for describing sensor data measurements. Our novel generative model partitions contributions to the sensor data from sources at a particular scan location and from sources outside the scan location, with corresponding Gaussian variance parameters that are estimated from data. Bayesian inference of this generative model, i.e. maximizing the marginal likelihood of the data, results in a cost function that can be optimized efficiently using convex bounding techniques. The resulting inference algorithm is quite robust to reconstruction of highly correlated sources and to the effect of high levels of interference and noise sources. The performance of the algorithm is tested in simulations and real data.

## II. Theory

The generative model for the observed MEG and EEG data which use an array of sensors to take electromagnetic field (or voltage potential) measurements from on or near the scalp surface with excellent temporal resolution is as follows. In both cases, the observed sensor data  $y(t) = [y_1(t), y_2(t), \dots, y_M(t)]^T$  at time point  $t$  is used for mapping from source activity which is generated by the same synchronous, compact current sources located within the brain, where  $M$  is the total number of sensors. We express the whole sensor time courses as  $y = y(t_1), y(t_2), \dots, y(t_K)$  where  $t_\kappa$  refers to the  $\kappa$ -th time point. Typical brain activity source models apply voxel discretization over a whole brain volume, and assume a fixed source at each voxel. The sources' activities are denoted by  $s(t) = [s_1(t)^T, s_2(t)^T, \dots, s_N(t)^T]^T$ , and  $N$  is total number of voxels.  $s(t)$  is called the voxel vector and the whole sources time courses are expressed as  $s = s(t_1), s(t_2), \dots, s(t_K)$  where  $K$  is the number of time-samples. The source at  $i$ -th voxel is expressed using a  $d_c$  dimensional vector  $s_i(t) \in \mathcal{R}^{d_c \times 1}$  where  $d_c$  denotes the number of directions<sup>1</sup>. We define the lead-field matrix as  $\mathbf{L} = [L_1, \dots, L_N]$  and for  $i$ -th voxel as  $l_i \in \mathcal{R}^{M \times d_c}$ ; The  $p$ -th column of  $L_i$  represents the signal vector that would be observed at the

<sup>1</sup>It is common to assume  $d_c = 2$  (for MEG) or  $d_c = 3$  (for EEG). For simplicity, we assume  $d_c = 1$  during derivation.

scalp given a unit current source/dipole at the  $i$ -th voxel with a fixed orientation in the  $p$ -th direction.

We assume that the whole brain activity can be represented by one source at a particular scan location (grid point) plus sources at other locations in the brain. The data vector can be expressed as:

$$\mathbf{y}(t) = \mathbf{l}_i \mathbf{s}_i + \sum_{j=1, j \neq i}^N \mathbf{l}_j s_j(t) + \boldsymbol{\varepsilon}. \quad (1)$$

The model in Eq. (1) is converted into a probabilistic model when we specify prior distributions for the unknown variables, such that

$$p(\mathbf{s}_i(t)) = \mathcal{N}(\mathbf{s}_i(t) \mid \mathbf{0}, \mathbf{v}_i^{-1}). \quad (2)$$

We divide the non-scan voxel space into  $R$  a priori specified regions (or tiles). The specification of regions could be based on either anatomical or functional criteria and the algorithm is agnostic about that. We assume that regions are non-overlapping, i.e., each voxel belongs to exactly a single region, and the  $r$ -th region contains  $N_r$  voxels. Importantly, we assume that voxels within the same region have the same prior regional variance.

$$p(s_j(t)) = \mathcal{N}(s_j(t) \mid \mathbf{0}, \boldsymbol{\Omega}_r^{-1}), \quad (3)$$

where  $\mathbf{v}_i \in \mathfrak{R}^{d_c \times d_c}$  and  $\boldsymbol{\Omega}_r \in \mathfrak{R}^{d_c \times d_c}$  are the prior precisions for  $i$ -th scanning voxel activity and for the  $r$ -th region. The Gaussian noise is assumed as that

$$p(\boldsymbol{\varepsilon}) = \mathcal{N}(\boldsymbol{\varepsilon} \mid \mathbf{0}, \boldsymbol{\lambda}^{-1}), \quad (4)$$

where  $\boldsymbol{\lambda}$  is a diagonal precision matrix. The  $i$ -th scanning source vector is estimated as its posterior mean  $\bar{\mathbf{s}}_i(t_k)$  which is given by

$$\bar{\mathbf{s}}_i(t_k) = E[\mathbf{s}_i(t_k) \mid \mathbf{y}] = \mathbf{v}_i^{-1} \mathbf{l}_i^T \boldsymbol{\Sigma}_i^{-1} \mathbf{y}(t_k), \quad (5)$$

where  $\boldsymbol{\Sigma}_i$  is the data model covariance, expressed as

$$\boldsymbol{\Sigma}_i = \boldsymbol{\lambda}^{-1} + \mathbf{l}_i \mathbf{v}_i^{-1} \mathbf{l}_i^T + \sum_{r=1}^R \sum_{j \in \zeta_r, j \neq i} \mathbf{l}_j \boldsymbol{\Omega}_r^{-1} \mathbf{l}_j^T, \quad (6)$$

$\sum_{j \in \zeta_r, j \neq i}$  indicates summation of voxels belonging to the  $r$ -th region while excluding the  $i$ -th scanning voxel. Estimating the parameters of the model,  $\mathbf{v}_i$ ,  $\boldsymbol{\Omega}_r$  involve minimizing a convex upper bound of the marginal likelihood of the data [4],  $\mathcal{F}(\mathbf{v}_i, \boldsymbol{\psi}_i, \boldsymbol{\Omega}, \boldsymbol{\Lambda})$  such that

$$\begin{aligned}
\mathcal{F} = & \frac{1}{K} \sum_{k=1}^K \left[ \lambda (\mathbf{y}(t_k) - \bar{\mathbf{I}}\mathbf{s}(t_k))^T (\mathbf{y}(t_k) - \bar{\mathbf{I}}\mathbf{s}(t_k)) \right. \\
& + \bar{\mathbf{s}}_i(t_k)^T \mathbf{v}_i \bar{\mathbf{s}}_i(t_k) + \sum_{r=1}^R \sum_{j \in \zeta_r, j \neq i} \bar{\mathbf{s}}_j(t_k)^T \mathbf{\Omega}_r \bar{\mathbf{s}}_j(t_k) \left. \right] \\
& - \boldsymbol{\psi}_i^T \mathbf{v}_i^{-1} - \sum_{r=1}^R \text{tr}(\mathbf{\Lambda}_r^T \mathbf{\Omega}_r^{-1}) + \text{const},
\end{aligned} \tag{7}$$

where the posterior mean of  $\mathbf{s}(t_k)$  is denoted as  $\bar{\mathbf{s}}_i(t_k)$ . The update rules for parameters  $\hat{\mathbf{v}}_i$  and  $\hat{\mathbf{\Omega}}_r$  are given by

$$\hat{\mathbf{v}}_i^{-1} = \sqrt{\frac{\frac{1}{K} \sum_{k=1}^K \bar{\mathbf{s}}_i(t_k) \bar{\mathbf{s}}_i^T(t_k)}{\boldsymbol{\psi}_i}}, \tag{8}$$

$$\hat{\mathbf{\Omega}}_r^{-1} = \sqrt{\frac{\frac{1}{KN_r} \sum_{k=1}^K \sum_{j \in \zeta_r, j \neq i} \bar{\mathbf{s}}_j(t_k) \bar{\mathbf{s}}_j^T(t_k)}{\mathbf{\Lambda}_r}}. \tag{9}$$

The update rules for auxiliary variables  $\boldsymbol{\psi}_i$  and  $\mathbf{\Lambda}_r$  are

$$\boldsymbol{\psi}_i = \mathbf{l}_i^T \boldsymbol{\Sigma}_y^{-1} \mathbf{l}_i, \tag{10}$$

$$\mathbf{\Lambda}_r = \sum_{j \in \zeta_r, j \neq i} \mathbf{l}_j^T \boldsymbol{\Sigma}_y^{-1} \mathbf{l}_j. \tag{11}$$

Defining

$$\mathbf{n}(t_k) = \mathbf{y}(t_k) - \mathbf{l}_i \bar{\mathbf{s}}_i(t_k) - \sum_{r=1}^R \sum_{j \in \zeta_r, j \neq i} \mathbf{l}_j \bar{\mathbf{s}}_j(t_k), \tag{12}$$

the noise precision can be estimated using

$$\lambda^{-1} = \frac{1}{K} \sum_{k=1}^K \text{diag}[E[\mathbf{n}(t_k) \mathbf{n}^T(t_k)]], \tag{13}$$

where  $\text{diag}(\mathbf{A})$  indicates the operation that create a diagonal matrix whose diagonal entries are equal to the diagonal elements of a matrix  $\mathbf{A}$ . For  $j$ -th non-scan voxel which belongs to  $r$ -th region, the source vector is estimated as its posterior mean  $\bar{\mathbf{s}}_j(t_k)$ , which is given by

$$\bar{\mathbf{s}}_j(t_k) = \mathbf{\Omega}_r^{-1} \mathbf{l}_j^T \boldsymbol{\Sigma}_y^{-1} \mathbf{y}(t_k). \tag{14}$$

A summary of the proposed algorithm is shown in Table 1.

We note that the computational complexity of the proposed algorithm is on the order  $\mathcal{O}(MK(R - I)N)$ , roughly equivalent to a single dipole scan, which is on the order  $\mathcal{O}(K(M^2 + N))$ . These are much smaller than the complexity of a multi-dipole scan which is on the order  $\mathcal{O}(KN^P)$  where  $P$  is the number of dipoles. According to our tests with the right initialization described below, if we use a scalar leadfield matrix with more than 8000 voxels, COGNAC can finish a whole brain scan and localize activity within half an hour (running MATLAB on a Xeon workstation with 12 CPU cores and 16G memory).

### III. Performance evaluation

Several representative source localization algorithms were chosen to evaluate the performance of COGNAC: an adaptive spatial filtering method, linearly constrained minimum variance beamformer (referred to as Beamformer) [5]; a non-adaptive weighted minimum-norm method, standardized low-resolution tomographic analysis (referred to as sLORETA) [6]; one Bayesian algorithms, MSP [7].

We tested our novel algorithm using scalar leadfield matrix constructed within the brain volume assuming a single-shell spherical model [8] as 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 simulated 5 point sources, then set the SNR=0 dB and the correlation between all sources as 0.5. We ran this configuration for more than 50 times and then averaged results. Two parameters  $A$  prime metric and the aggregate performance (AP) were used to evaluate the performance of the novel algorithm [2], [3], [4]. The  $A$  metric estimates the area under the FROC curve for one hit rate (HR) and false rate (FR) pair, or in our case, for each simulation. If the area under the FROC curve is large, then the hit rate is higher compared to the false positive rate. The aggregate performance (AP) captures both the accuracy of the location and the time courses of algorithms.

For real data sets test, we chose the resting state MEG data analysis, subjects were instructed simply to keep their eyes closed and do not think of anything in particular, for COGNAC, we use the diagonal noise covariance estimated using a variational Bayesian factor analysis (VBFA) model [9]. The EEG data (128-channel ActiveTwo system) was downloaded from the SPM website (<http://www.fil.ion.ucl.ac.uk/spm/data/mmfaces>) and the lead field was calculated in SPM8 using coarse resolution. The EEG data paradigm involves randomized presentation of at least 86 faces and 86 scrambled faces. Here we use 86 averaged faces data to subtract the averaged scrambled-faces data to study the differential response to faces versus scrambled faces [10]. The EEG data has been reported in our prior publications using the Champagne algorithm, and details about this dataset can be found in [3].

## IV. Results

### A. Simulation Results

COGNAC shows us better source reconstruction results with scalar leadfield matrix. Figure 1 is the simulation results using scalar leadfield matrix with 5 active point sources and configurations are set as SNR=0 dB and correlation among sources are 0.5. COGNAC performs the best among all source reconstruction algorithms.

## B. Real data: Resting State Data

The localization results for resting state data analysis from three subjects are shown in Figure 2. As is seen, COGNAC can consistently localize brain activity during rest, while other benchmark algorithms do not have consistent localization of resting-state activity. For resting state analysis, even though there is no pre-stimulus data for background noise information estimation, COGNAC still recovers reasonable activity. However, Beamformer and sLORETA can not localize brain activity consistently for all subjects.

## C. Real data: Face-Processing task: EEG

In Figure 3, we present the results from using novel algorithm and benchmarks on the face-processing task EEG data set. Figure 3 shows the average power, M100 peak power, and M170 peak power at different rows separately. We see that COGNAC is able to localize brain activity with sparse peaks at visual areas and the fusiform gyrus. While the benchmark algorithms produce brain activity with incorrect and/or diffuse localization.

## V. DISCUSSION

We have described a novel scanning algorithm for MEG source localization that systematically accounts for contributions to sensors from activity outside the scan location. Importantly, the algorithm also estimates heteroscedastic sensor noise from observed data without the need for additional pre-stimulus or baseline data. We show that performance of COGNAC is superior to existing benchmarks both on simulations and real MEG/EEG data with complex source configurations. COGNAC is robust to reconstruction of highly correlated sources and to the effect of high levels of interference and noise sources. Superior performance of this algorithm arises from an explicit parameterization of the covariance partitions to sensor data from both non-scan location activity and sensor noise. No other existing source localization algorithm also simultaneously estimates sensor noise in data while retaining accuracy in localization. Our future work will include extensions to non-Gaussian source estimation and spatio-temporal smoothness in the COGNAC generative models.

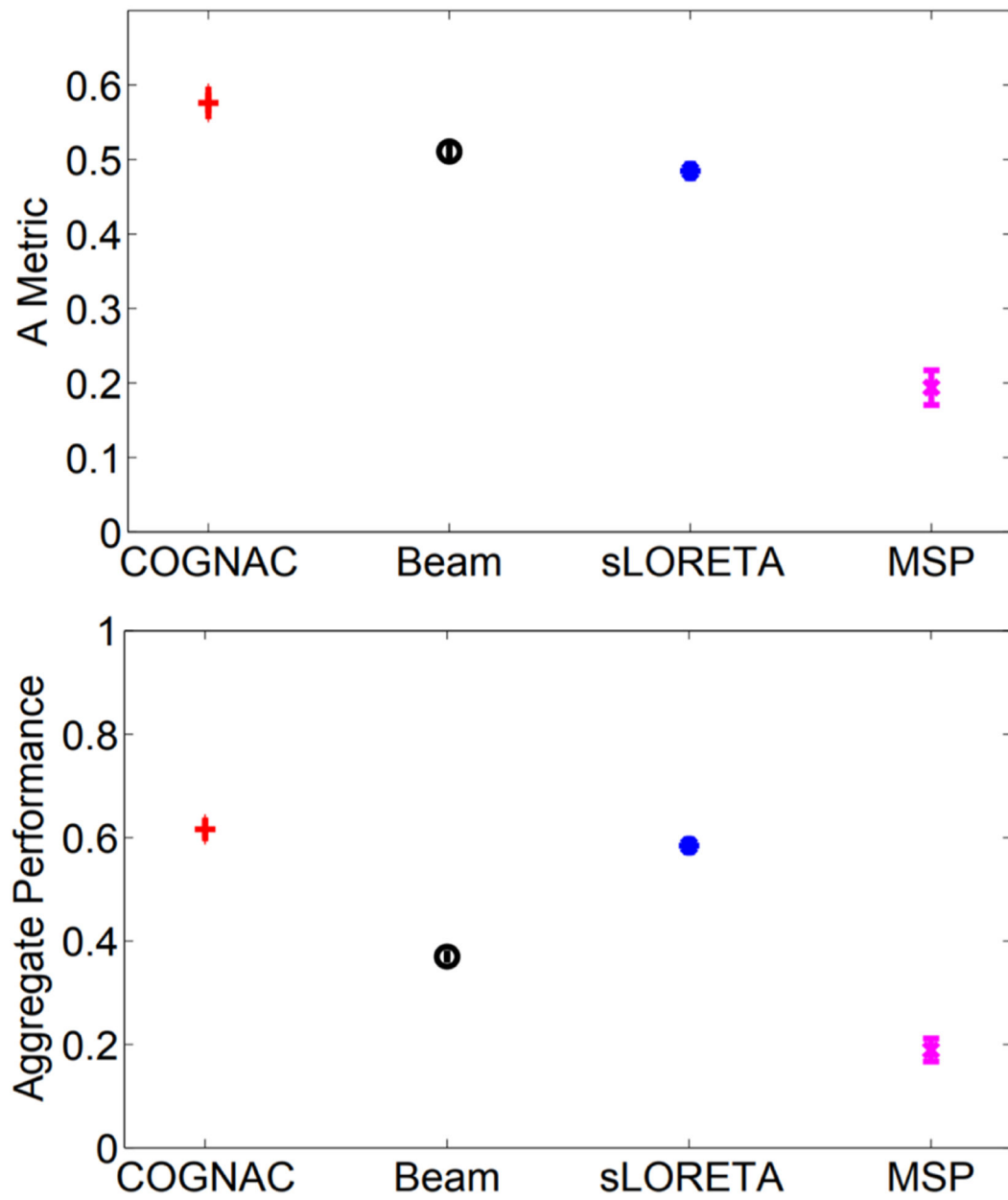
## Acknowledgments

This work was supported by NIH grants R01EB022717, R01DC013979, R01NS100440 and MRP-17-454755

## References

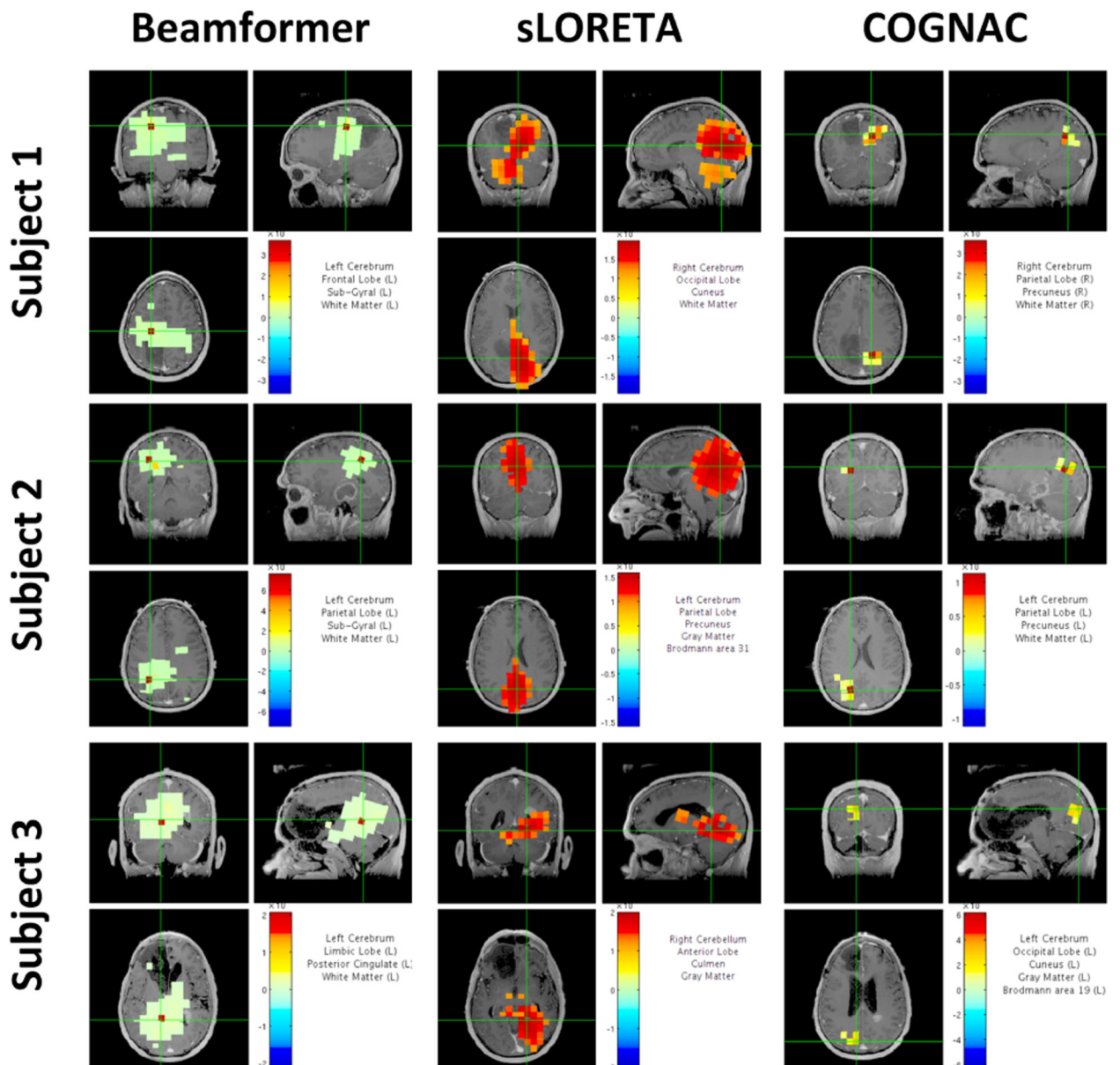
- [1]. Sekihara K and Nagarajan SS, *Electromagnetic brain imaging: A Bayesian perspective*. Berlin, Heidelberg: Springer-Verlag, 2015.
- [2]. Wipf DP, Owen JP, Attias HT, Sekihara K, and Nagarajan SS, "Robust bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG," *Neuroimage*, vol. 49, no. 1, pp. 641–655, 2010. [PubMed: 19596072]
- [3]. Owen JP, Wipf DP, Attias HT, Sekihara K, and Nagarajan SS, "Performance evaluation of the champagne source reconstruction algorithm on simulated and real M/EEG data," *Neuroimage*, vol. 60, no. 1, pp. 305–323, 2012. [PubMed: 22209808]
- [4]. Cai C, Sekihara K, and Nagarajan SS, "Hierarchical multiscale bayesian algorithm for robust meg/eeg source reconstruction," *Neuroimage*, vol. 183, pp. 698–715, 2018. [PubMed: 30059734]

- [5]. Sekihara K and Scholz B, "Generalized Wiener estimation of three-dimensional current distribution from biomagnetic measurements," in *Biomag 96: Proceedings of the Tenth International Conference on Biomagnetism*, Aine CJ et al., Eds. New York: Springer-Verlag, 1996, pp. 338–341.
- [6]. Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, and Halgren E, "Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity," *Neuron*, vol. 26, pp. 55–67, 2000. [PubMed: 10798392]
- [7]. Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, Henson R, Flandin G, and Mattout J, "Multiple sparse priors for the M/EEG inverse problem," *Neuroimage*, vol. 39, no. 3, pp. 1104–1120, 2008. [PubMed: 17997111]
- [8]. Hallez H, Vanrumste B, Grech R, Muscat J, De Clercq W, Vergult A, D'Asseler Y, Camilleri KP, Fabri SG, Van Huffel S et al., "Review on solving the forward problem in EEG source analysis," *Journal of neuroengineering and rehabilitation*, vol. 4, no. 1, p. 46, 2007. [PubMed: 18053144]
- [9]. Nagarajan SS, Attias H, H. KE 2nd, and Sekihara K, "A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data," *Stat. Med*, vol. 20, pp. 3886–3910, 2007.
- [10]. Henson RN, Flandin G, Friston KJ, and Mattout J, "A Parametric Empirical Bayesian framework for fMRI-constrained MEG/EEG source reconstruction," *Human brain mapping*, vol. 31, no. 10, pp. 1512–1531, 2010. [PubMed: 20091791]

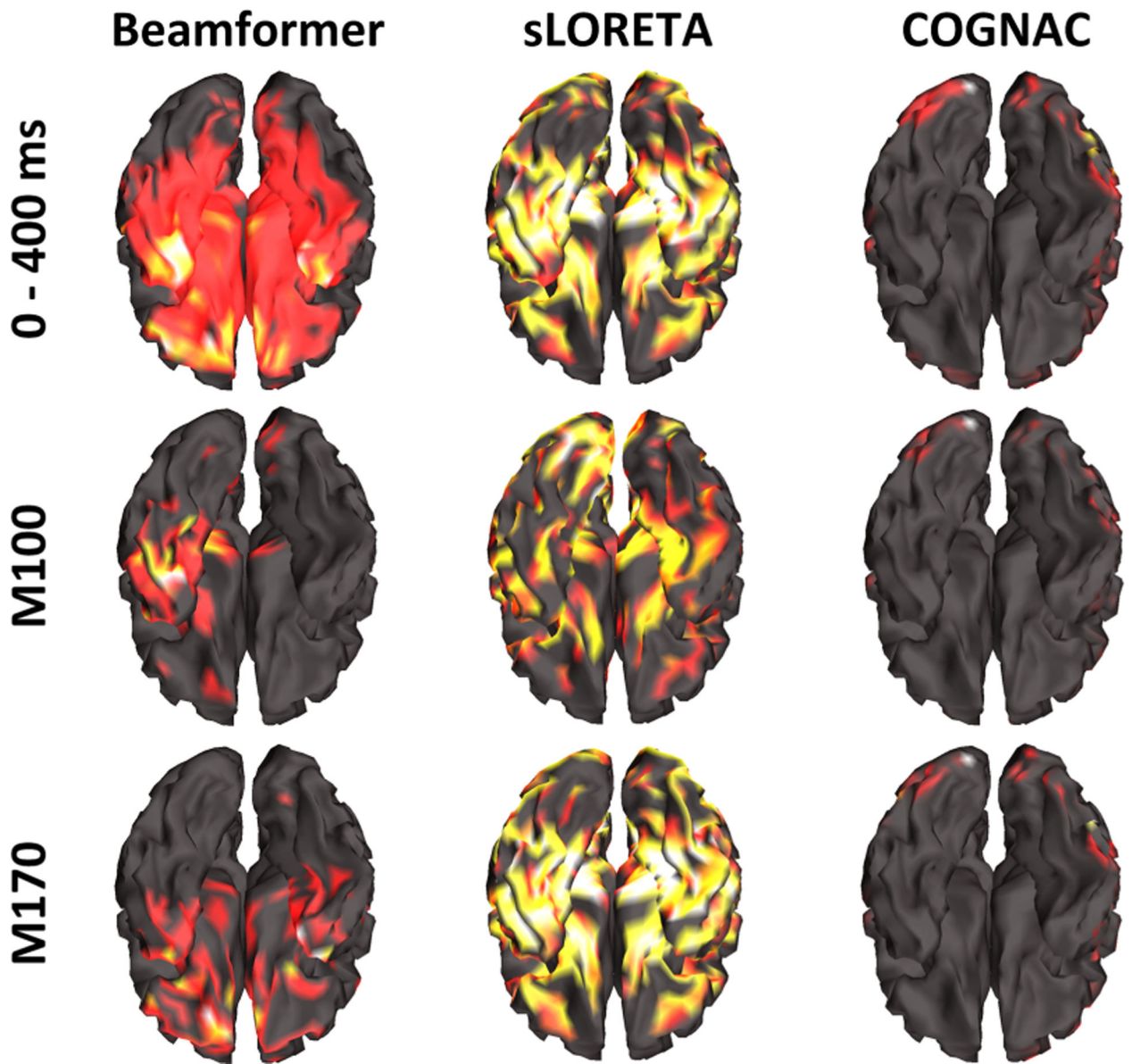


**Fig. 1:** Performance evaluation with 5 point sources. The SNR, and inter-source correlation were set as 0 dB, and 0.5. These results were averaged with 50 times simulations.





**Fig. 2:** Resting state data analysis with three subjects. Benchmark algorithms Beamformer and sLORETA are plotted at first to second columns. COGNAC is shown at the last column. Thresholding is performed at 10 percent of the maximum power.



**Fig. 3:** Results for EEG face processing EEG data from novel algorithm and benchmarks. The first row is the average power mapping from 0 ms to 400 ms. The second and third rows are for peak power activity at 100 ms and 170 ms.

**Table 1**

## The COGNAC Algorithm

---

repeat  
 1: Set the  $i$ -th target voxel as  $i = 1$ .  
 2: Set appropriate initial values to  $v_{\hat{r}}, \mathbf{Q}_{\hat{r}}, \boldsymbol{\lambda}$ , compute  $\boldsymbol{\psi}_{\hat{r}}, \mathbf{A}_{\hat{r}}$  using Eqs. (10), (11), and  $\bar{\mathbf{s}}_{\hat{r}}(t_k)$  using Eq (5) for  $k = 1, \dots, K$ . repeat  
   3: Update  $v_{\hat{r}}, \mathbf{Q}_{\hat{r}}, \boldsymbol{\lambda}$  using Eqs. (8), (9), (13), (14).  
   4: Update  $\boldsymbol{\psi}_{\hat{r}}, \mathbf{A}_{\hat{r}}, \bar{\mathbf{s}}_{\hat{r}}(t_k)$  with the updated values of  $v_{\hat{r}}, \mathbf{Q}_{\hat{r}}, \boldsymbol{\lambda}$  using Eqs. (10), (11), (5).  
 until The cost function of Eq. (7) converges.  
 5: Set the target voxel to the  $(i + 1)$ -th voxel.  
 until All voxels are scanned ( $i = N$ ).

---

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript