Published in final edited form as:

Int IEEE EMBS Conf Neural Eng. 2019 March ; 2019: 299–302. doi:10.1109/NER.2019.8717043.

Sparse Multi-task Inverse Covariance Estimation for Connectivity Analysis in EEG Source Space

Feng Liu^{1,2}, Emily P. Stephen^{1,2}, Michael J. Prerau^{1,2}, and Patrick L. Purdon^{1,2}

¹Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

²Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, 02139, USA.
fliu0@mgh.harvard.edu, emilyps14@gmail.com, prerau@nmr.mgh.harvard.edu,
patrickp@nmr.mgh.harvard.edu

Abstract

Understanding how different brain areas interact to generate complex behavior is a primary goal of neuroscience research. One approach, functional connectivity analysis, aims to characterize the connectivity patterns within brain networks. In this paper, we address the problem of *discriminative* connectivity, i.e. determining the differences in network structure under different experimental conditions. We introduce a novel model called Sparse Multi-task Inverse Covariance Estimation (SMICE) which is capable of estimating a common connectivity network as well as discriminative networks across different tasks. We apply the method to EEG signals after solving the inverse problem of source localization, yielding networks defined on the cortical surface. We propose an efficient algorithm based on the Alternating Direction Method of Multipliers (ADMM) to solve SMICE. We apply our newly developed framework to find common and discriminative connectivity patterns for α -oscillations during the Sleep Onset Process (SOP) and during Rapid Eye Movement (REM) sleep. Even though both stages exhibit a similar α -oscillations, we show that the underlying networks are distinct.

Keywords

Inverse Covariance Estimation; Sleeping α -Burst; EEG Source Imaging; Alternating Direction Method of Multipliers (ADMM)

I. INTRODUCTION

Electroencephalographic (EEG) recording of neural activity has been used widely for sleep monitoring and diagnosis of sleep disorders [1] [2] [3] [4]. EEG activity in the α band (8–12 Hz) has long been associated with wakefulness with power attenuating and becoming sporadic during the Sleep Onset Process (SOP) [2] [5], then disappearing at the onset of sleep. Bursts of α activity have also been observed during and around [1] Rapid Eye Movement (REM) sleep. It is unknown if the underlying connectivity difference for those two similar phenomenon.

Usually, we can characterize the brain functional connectivity under those distinct brain states, and compare the patterns of connectivity and pathways involved in memory

consolidation and recall. However, α -bursts in both REM and SOP only last for several seconds, which is too short in duration to study with fMRI. Instead, the high temporal resolution of EEG makes it well-suited to analyze the connectivity patterns between REM and SOP.

It is well known that the brain can be considered as a network of spatially and functionally interacting regions [7] [8]. A widely used approach to assess brain connectivity is to estimate the *sparse inverse covariance matrix*, which describes the conditional dependency among regions of interest (ROIs) by assuming that the activation pattern across ROIs follows a multivariate Gaussian distribution [9]. This approach has attracted a lot of interests since sparsity leads to good interpretability in the context of high dimensional neural imaging data.

In recent years, the sparse graphical model and its variants have been used to discover brain networks under distinct conditions. Varoquaux et al. proposed a population prior to address the variability among subjects when estimating the brain covariance matrix during spontaneous brain activity [9]. To identify the brain network alterations among HIV patients and healthy controls, Liu et al. constructed a contrasting framework and observed major differences in the occipital lobe and parietal lobe [10]. To discover the default mode alterations in schizophrenia, Lefort-Besnard et al. employed sparse inverse covariance estimation method and identified a statistically significant difference between healthy controls and patients in the default mode network (DMN), saliency network (SN), and dorsal attention network (DAN). Monti et al. proposed the Smooth Incremental Graphical Lasso Estimation (SINGLE) algorithm which uses a smoothly aggregated empirical covariance matrix and fused Lasso and suggested that the right inferior frontal gyrus and right inferior parietal lobe play key roles in attention and executive function during cognitive tasks [11].

However, to the best of our knowledge, all previous studies of estimating brain connectivity using the inverse covariance matrix method have been focused on fMRI data rather than EEG data. It is beneficial for the neuroimaging community to use EEG data to estimate the brain connectivity due to its advantages in capturing fast-changing connectivity patterns in a noninvasive way. In this paper, we proposed a general model to discover the discriminative and common connectivity patterns for the EEG recordings under distinct conditions. We used the idea of multi-task learning and come up with a novel model called Sparse Multi-task Inverse Covariance Estimation (SMICE) and an efficient algorithm under the framework of ADMM. To demonstrate its efficacy, we applied the model to find the distinct patterns for SOP and REM.

The main contributions of this paper are:

1. We incorporate the idea of multi-task learning and propose a novel model called SMICE to learn the common and discriminative networks simultaneously.
2. We present an efficient algorithm based on ADMM to solve SMICE efficiently.
3. We apply the proposed algorithm to discover distinct networks mediated by α oscillations during SOP and REM sleep.

II. PROPOSED MODEL

Assume the source signal $X \in R^{p \times n}$ contains the estimated time series from p sources with a length of n time points. The i -th column in X , denoted as x_i , is the i -th observation or realization from a p -variate Gaussian random field, following distribution $N(\mu, \Sigma)$. The empirical covariance matrix S is calculated as $S = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})(x_i - \hat{\mu})^T$. Under different cognitive tasks, we assume the source signal X_m follows different distributions $N(\mu_m, \Sigma_m)$ where m is the index for tasks. Estimating the tasks separately, however, does not utilize the combined information across different tasks, and can make the results less stable and less interpretable. We propose an objective function based on the idea of multi-task learning [12], given below:

$$\begin{aligned} \arg \min_{\{\Theta_m\}} & \sum_{m=1}^k -\log \det \Theta_m + \frac{1}{|A_m|} \sum_{m=1}^k \sum_{j \in \{A_m\}} \text{Tr}(S_{m,j} \Theta_m) \\ & + \alpha \sum_{m=1}^k \|\Theta - \Theta_m\|_1 + \beta \sum_{m=1}^k \|\Theta_m\|_1. \end{aligned}$$

For the sparse inverse covariance matrices, Σ_m^{-1} , denoted as Θ_m , the i, j -th element $\Theta_{m,ij}$ of Θ_m is zero if and only if the i th and j th variables are independent. In this paper, we use the following notation: $\text{Tr}(M)$ is denoted as trace of matrix M . $\|\cdot\|_1$ is the ℓ_1 norm, which is the sum of all of the elements of a vector or matrix. $\|\cdot\|_F$ is the Frobenius norm. In Eq.(1), k is the number of tasks, $S_{m,j}$ represents the empirical covariance matrix for the j -th trial in the m -th task, $\{A_m\}$ is the set that contains all the indices for task m

Eq.(1) describes how solutions with different features will be weighted by the estimation algorithm. The first and second terms come from the maximum likelihood estimate of the covariance matrix. The third term requires the algorithm to find a common connectivity pattern that is shared across tasks. The fourth term in Eq.(1) imposes sparsity on each task connectivity pattern. We aim to find the task specific inverse covariance matrices or *precision* matrices $\Theta_m, (m = 1 \dots k)$ and the common inverse covariance matrix (consensus variable [13]) Θ which correspond to the task specific connectivity pattern and common connectivity pattern, respectively. Letting $S_m = \frac{1}{|A_m|} \sum_{j \in \{A_m\}} S_{m,j}$ be the mean empirical covariance matrix for task m , index j can be dropped without causing any confusion.

III. OPTIMIZATION ALGORITHM

To solve Problem (1), we first introduce auxiliary variable $Z_m (m = 1 \dots k)$, and insert it into the original problem in the place of Θ_m in the last two terms:

$$\begin{aligned}
& \arg \min_{\{\Theta_m\}, \Theta} \sum_{m=1}^k -\log \det \Theta_m + \sum_{m=1}^k \text{Tr}(S_m \Theta_m) \quad (1) \\
& + \alpha \sum_{m=1}^k \|\Theta - Z_m\|_1 + \beta \sum_{m=1}^k \|Z_m\|_1 \\
& \text{s.t. } Z_m = \Theta_m, m = 1 \dots k.
\end{aligned}$$

The Lagrangian multiplier function is given below:

$$\begin{aligned}
L(\{\Theta_m\}, \Theta, \{U_m\}, \{Z_m\}) = & \sum_{m=1}^k -\log \det \Theta_m \quad (2) \\
& + \sum_{m=1}^k \text{Tr}(S_m \Theta_m) + \alpha \sum_{m=1}^k \|\Theta - Z_m\|_1 + \beta \sum_{m=1}^k \|Z_m\|_1 \\
& + \frac{\rho}{2} \sum_{m=1}^k \|\Theta_m - Z_m + U_m\|_F^2 - \frac{\rho}{2} \sum_{m=1}^k \|U_m\|_F^2,
\end{aligned}$$

where $U_m \in R^{P \times P}$, $m = 1, \dots, k$ is the Lagrangian multipliers, and ρ is the iteration step size. At each iteration, $\{\Theta_m\}$, Θ and $\{U_m\}$ are updated sequentially by fixing the other variables.

To update $\{\Theta_m\}$ at the i -th iteration, the following objective function needs to be solved:

$$\begin{aligned}
& \arg \min_{\Theta_m} L(\Theta_m, \Theta^{(i-1)}, U_m^{(i-1)}, Z_m^{(i-1)}) = \quad (3) \\
& \arg \min_{\Theta_m} \left\{ \text{Tr}(S_t \Theta_m) - \log \det \Theta_m \right. \\
& \left. + \frac{\rho}{2} \sum_{m=1}^k \|\Theta_m - Z_m^{(i-1)} + U_m^{(i-1)}\|_F^2 \right\}
\end{aligned}$$

Due to limited space, the procedure to derive the update of Θ is omitted here. Readers can refer to Page 47 of Ref. [13] for more details.

The update of Θ is done by solving the following problem:

$$\begin{aligned}\Theta^{(i)} &= \underset{\Theta}{\operatorname{argmin}} L(\Theta_m^{(i)}, \Theta, U_m^{(i-1)}, Z_m^{(i-1)}) \quad (4) \\ &= \underset{\Theta}{\operatorname{argmin}} \left\{ \sum_{m=1}^k \left\| \Theta - Z_m^{(i)} \right\|_1 \right\},\end{aligned}$$

which equals the mean of $Z_m^{(i)}$ for $i = 1, \dots, k$, calculated as $\Theta^i = \frac{1}{k} \sum_{m=1}^k Z_m^{(i)}$.

The update of Z_m involves solving the following problem:

$$\begin{aligned}Z_m^{(i)} &= \underset{Z_m}{\operatorname{argmin}} \alpha \left\| \Theta^{(i-1)} - Z_m \right\|_1 + \beta \left\| Z_m \right\|_1 \quad (5) \\ &+ \frac{\rho}{2} \left\| \Theta_m^{(i)} - Z_m + U_m^{(i-1)} \right\|_F^2, \text{ for } m = 1, \dots, k.\end{aligned}$$

Eq.(5) has an analytic solution and can be computed elementwise. Without any confusion, let z , θ , u be an element from the same location of corresponding matrices, then Eq (5) can be written as $z = \underset{z}{\operatorname{argmin}} \alpha |z - \theta| + \beta |z| + \frac{\rho}{2} (z - d)^2$ where $d = \theta + u$. The above equation is a proximal operator $\mathbf{prox}_{\rho f}(d) = \underset{z}{\operatorname{argmin}} f(z) + \frac{\rho}{2} \|z - d\|_2^2$, where $f(z) = \alpha \|z - \theta\|_1 + \beta \|z\|_1$. The analytic solution for $\theta > 0$ is:

$$\left\{ \begin{array}{ll} d \geq \theta + \frac{\alpha + \beta}{\rho}, & z = d - \frac{\alpha + \beta}{\rho} \\ \frac{-\alpha + \beta}{\rho} \leq d < \theta - \frac{\alpha - \beta}{\rho}, & z = d + \frac{\alpha - \beta}{\rho} \\ d < -\frac{\alpha + \beta}{\rho}, & z = d + \frac{\alpha + \beta}{\rho} \\ \frac{-\alpha - \beta}{\rho} \leq d < \frac{-\alpha + \beta}{\rho}, & z = 0 \\ \frac{-\alpha + \beta}{\rho} + \theta \leq d < \frac{\alpha + \beta}{\rho} + \theta, & z = \theta \end{array} \right. \quad (6)$$

And the solution for $\theta \leq 0$ is:

$$\left\{ \begin{array}{ll} d \geq \frac{\alpha + \beta}{\rho}, & z = d - \frac{\alpha + \beta}{\rho} \\ \theta + \frac{\alpha - \beta}{\rho} \leq d < \frac{\alpha - \beta}{\rho}, & z = d + \frac{-\alpha + \beta}{\rho} \\ d < \theta - \frac{\alpha + \beta}{\rho}, & z = d + \frac{\alpha + \beta}{\rho} \\ \frac{\alpha - \beta}{\rho} \leq d < \frac{\alpha + \beta}{\rho}, & z = 0 \\ \frac{-\alpha - \beta}{\rho} + \theta \leq d < \frac{\alpha - \beta}{\rho} + \theta, & z = \theta \end{array} \right. \quad (7)$$

In the end, the update of Lagrange multipliers $\{U_m\}$ is given as follows:

$$U_m^{(i)} = U_m^{(i-1)} + \Theta_m^{(i)} - Z_m^{(i)} \quad \text{for } m = 1, \dots, k. \quad (8)$$

The above procedure is summarized in Algorithm (1). The convergence of Algorithm (1) is determined by primal feasibility criteria $\|\Theta_m^{(i)} - Z_m^{(i)}\|_F^2 < \epsilon_1$ and dual feasibility

$$\|Z_m^{(i)} - Z_m^{(i-1)}\|_F^2 < \epsilon_2.$$

Algorithm 1: ADMM procedure to solve SMICE

Data: $\Theta_m^{(0)} = Z_m^{(0)} = \mathbf{0}$ for $m = 1, \dots, k$, $\Theta^{(0)} = \mathbf{0}$

Result: Θ_m, Θ

1 while not converged **do**

Update $\Theta_m^{(i)}$ by solving (3) for $m = 1, \dots, k$.

2

Update $\Theta^{(i)}$ by solving (4).

3

Update $Z_t^{(i)}$ by solving (5) for $m = 1, \dots, k$.

4

5

6

Update $U_m^{(i)}$ according to (8) $m = 1, \dots, k$.

$i = i + 1$

7 end

IV. ANALYZING EEG SOURCE CONNECTIVITY PATTERNS DURING SLEEP

A. Data Description

In this study, 10 healthy right-handed subjects (5 women and 5 men) with ages ranging from 19 to 32 years (mean: 25.8 yr; SD: 5.09) and body mass index (BMI) of 30 were asked to sleep for two consecutive nights in the Massachusetts General Hospital sleep laboratory. Experimental details for these studies have been described previously in Prerau, et al. [2]. These studies were approved by the Human Research Committee Institutional Review Board at Massachusetts General Hospital.

B. Data processing, burst segmentation and source localization

Standard EEG data preprocessing was conducted to remove bad epochs containing artifacts, and the raw data were down-sampled to 200 Hz. In order to segment α -bursts, the averaged spectrum of 8 occipital channels was used. The α -burst segmentations in SOP and REM consisted of fitting a Gaussian curve centered around the α -band (8 – 12Hz). The staged sleep data and multitaper spectrogram was given in Fig.1, where the α -burst SOP and REM are also marked in red and black box respectively. Source localization was performed using the dMAP-EM algorithm [14]. Similar to [14], we employed a reduced-dimension source space with a total number of 1284 patches. Each patch represents many densely sampled dipoles, with a total of about 300,000 dipoles across the cortical surface. In order to better interpret brain connectivity, we further bundle the 1284 patches into 68 regions of interest (ROIs) according to the parcellation of Desikan-Killiany Atlas.

C. Common and distinct connectivity patterns between SOP and REM

The proposed algorithm is applied to estimated source time series from dMAP-EM. We run SMICE for each subject independently due to high variance among subjects. We set $\alpha = 0.1$ and $\beta = 0.01$ for all subjects to establish a balance between data fidelity and sparsity. To make a fair comparison of networks across subjects and conditions, we set the maximum percentage of edges for each condition across all subjects to be 1.5%, which is about 70 edges out of 68 $67/2$ total edges in a fully connected network. The reconstructed common network and individual networks for both SOP and REM (i.e., Θ_{SOP} and Θ_{REM}) are illustrated in Fig.2. All the edges that are plotted in Fig.2 are shared at least by 5 subjects. In comparison, by random chance, $Pr(X \geq 5) < 10^{-6}$ for each edge, where X is the random variable describing the number of occurrences of an edge between two ROIs for all subjects.

As can be seen from the networks of REM and SOP, SOP has stronger connectivity between occipital and frontal areas. The convergence curves are shown in Fig.3. In Fig.3, the left subfigure is the convergence of objective function, the middle and right subfigures are convergence curves for the r -norm and s -norm, respectively, where $\|s\|_2 = \|X - Z\|_2$ and $\|r\|_2 = \|Z_t - Z_{t-1}\|_2$, which correspond to the primal feasibility criteria and dual feasibility criteria, respectively.

V. CONCLUSION

In this paper, we proposed a new framework we call SMICE to estimate common and discriminative connectivity patterns simultaneously. We describe an efficient estimation algorithm based on ADMM. We then apply this algorithm to analyze connectivity from source-localized EEG alpha oscillations observed during the sleep onset process and REM sleep. Our algorithm is able to identify both common and distinct networks during these different states. This algorithm could be useful for analyses of brain connectivity in other scenarios, such as different cognitive tasks, or in neurological or psychiatric conditions.

Acknowledgments

This work is supported by NIH grants R01NS096177 (to M.J.P.), R01AG054081 (to P.L.P.), and R01AG056015 (to P.L.P)

REFERENCES

- [1]. Prerau MJ, Brown RE, Bianchi MT, Ellenbogen JM, and Purdon PL, "Sleep neurophysiological dynamics through the lens of multitaper spectral analysis," *Physiology*, vol. 32, no. 1, pp. 60–92, 2016.
- [2]. Prerau MJ, Hartnack KE, Obregon-Henao G, Sampson A, Merlino M, Gannon K, Bianchi MT, Ellenbogen JM, and Purdon PL, "Tracking the sleep onset process: an empirical model of behavioral and physiological dynamics," *PLoS computational biology*, vol. 10, no. 10, p. e1003866, 2014. [PubMed: 25275376]
- [3]. Diych M, Li Y, and Wen P, "EEG sleep stages classification based on time domain features and structural graph similarity," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 11, pp. 1159–1168, 2016. [PubMed: 27101613]
- [4]. Lustenberger C, Mouthon A-L, Tesler N, Kurth S, Ringli M, Buchmann A, Jenni OG, and Huber R, "Developmental trajectories of EEG sleep slow wave activity as a marker for motor skill development during adolescence: a pilot study," *Developmental psychobiology*, vol. 59, no. 1, pp. 5–14, 2017. [PubMed: 27401676]
- [5]. Ogilvie RD, "The process of falling asleep," *Sleep medicine reviews*, vol. 5, no. 3, pp. 247–270, 2001. [PubMed: 12530990]
- [6]. Cantero JL, Atienza M, and Salas RM, "Human alpha oscillations in wakefulness, drowsiness period, and rem sleep: different electroencephalographic phenomena within the alpha band," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 32, no. 1, pp. 54–71, 2002. [PubMed: 11915486]
- [7]. Sporns O, Chialvo DR, Kaiser M, and Hilgetag CC, "Organization, development and function of complex brain networks," *Trends in cognitive sciences*, vol. 8, no. 9, pp. 418–425, 2004. [PubMed: 15350243]
- [8]. Liu F, Xiang W, Wang S, and Lega B, "Prediction of seizure spread network via sparse representations of overcomplete dictionaries," in *International Conference on Brain and Health Informatics*, 2016, pp. 262–273.
- [9]. Varoquaux G, Gramfort A, Poline J-B, and Thirion B, "Brain covariance selection: better individual functional connectivity models using population prior," in *Advances in neural information processing systems*, 2010, pp. 2334–2342.
- [10]. Liu X, Kong X, and Ragin AB, "Unified and contrasting graphical lasso for brain network discovery," in *Proceedings of the 2017 SIAM International Conference on Data Mining SIAM*, 2017, pp. 180–188.
- [11]. Monti RP, Hellyer P, Sharp D, Leech R, Anagnostopoulos C, and Montana G, "Estimating time-varying brain connectivity networks from functional mri time series," *NeuroImage*, vol. 103, pp. 427–443, 2014. [PubMed: 25107854]
- [12]. Zhu X, Li H, and Fan Y, "Parameter-free centralized multi-task learning for characterizing developmental sex differences in resting state functional connectivity," in *Proceedings of the... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence*, vol. 2018 NIH Public Access, 2018, p. 2660.
- [13]. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends R in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [14]. Pirondini E, Babadi B, Obregon-Henao G, Lamus C, Malik WQ, Ha MS^{ma}la^{la}inen, and Purdon PL, "Computationally efficient algorithms for sparse, dynamic solutions to the EEG source localization problem," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 6, pp. 1359–1372, 2018. [PubMed: 28920892]

Fig. 1:
Multitaper Spectrogram Analysis and SOP and REM α -Burst Segmentation

Fig. 2:

Common connectivity and task specific connectivity for SOP and REM. The middle row is the common network pattern and the top row is the network pattern of REM and the third row is the connectivity pattern for SOP.

Fig. 3:

Convergence of objective function s -norm and r -norm with $\alpha = 0.1$ and $\beta = 0.01$