

Inferring Effective Connectivity in the Brain from EEG Time Series Using Dynamic Bayesian Networks

Ali Yener Mutlu and Selin Aviyente

Department of Electrical and Computer Engineering, Michigan State University
East Lansing, MI, 48824, USA

mutlualiegr@msu.edu

aviyente@egr.msu.edu

Abstract—Effective connectivity, defined as the influence of a neuronal population on another, is known to have great significance for understanding the organization of the brain. Disruptions in the effective connectivity patterns occur in the case of neurological and psychopathological diseases. Therefore, it is important to develop models of effective brain connectivity from non-invasive neuroimaging data. In this paper, we propose to use dynamic Bayesian networks (DBN) to learn effective brain connectivity from electroencephalogram (EEG) data. DBNs use first order Markov chain to model EEG time series obtained from multiple electrodes. We explore effective brain connectivity in healthy and schizophrenic subjects using this framework. Fourier bootstrapping technique is used to identify the statistically significant pairs of interactions among electrodes.

I. INTRODUCTION

The guiding principles underlying the organization of the brain are functional integration and segregation [1]. Although functional segregation is well understood, quantifying functional integration still remains a challenge. Friston and others [2] have established that measures of functional and effective connectivity are essential to understanding the integration in the brain. Functional connectivity refers to statistical relationships without causal implications, whereas effective connectivity refers to causal neural interactions [3]. Therefore, quantification of effective connectivity is important to infer the directional networks in the brain and to understand the causality of neural processes. Disruption of effective connectivity in the brain is known to be associated with neurological and psychopathological diseases such as Alzheimer's Disease and schizophrenia [4]. Therefore, assessment of brain connectivity is essential in terms of discovering the causes of these diseases.

Several mathematical methods, such as structural equation modeling (SEM) [5], dynamic causal modeling (DCM) [6], Directed Transfer Function (DTF) [7], Granger causality mapping (GCM) [2], [8], multivariate autoregressive modeling (MAR) [2], partial directed coherence (PDC) [9] and Bayesian networks (BN) [10], [11], have been used to learn the effective brain connectivity structures from neuroimaging data. However, most of these methods assume linear interactions between measured neuronal oscillations, despite the fact

that such interactions are known to be nonlinear. Moreover, they assume a known structure or model for connectivity and then try to compute the dependence relations among the regions of interest.

More recently, researchers have addressed these issues using dynamic Bayesian networks to infer the underlying connectivity structure from neuroimaging data. DBN is an extension of BN which is capable of modeling the temporal relationships among different regions of interest. Furthermore, as opposed to previous methods that quantify causality, DBNs can capture nonlinear dependencies. For example, DBN framework is employed to learn the brain effective connectivity from fMRI data in [10], [12], [13] and from electrophysiology data in [3]. However, fMRI data suffers from low temporal resolution caused by the slow scanning process. EEG recordings overcome this problem since it has better temporal resolution.

In this paper, we propose to use discrete dynamic Bayesian networks (dDBN) for modeling the effective connectivity from EEG data. To our knowledge, this paper presents the first application of dDBN to learn effective brain connectivity from EEG data. This framework allows us to quantify the nonlinear interactions between measured neural activity by modeling EEG time series as discrete random variables with multinomial distributions [12]. DBNs can assume that the structure of effective connectivity is known or unknown [10]. In our study, we prefer to start with an unknown structure and find the optimal one that represents the EEG data using Markov chain Monte Carlo (MCMC) method.

The rest of the paper is organized as follows. In Section II, we give a brief overview of DBN and its implementation as used in this paper. Section III describes the EEG data used in our study. The obtained DBN structures are validated through a Fourier bootstrapping technique which identifies the statistically significant interactions. Finally, section IV discusses the results for healthy and schizophrenic subjects and suggests extensions of the proposed method.

II. BACKGROUND ON DBNs

A Bayesian network is described by a structure S and a joint distribution, formed by a family of conditional probability distributions P and their parameters Q , over a set of random variables [14]. The graphical structure of

This work was in part supported by the National Science Foundation under Grant No. CAREER CCF-0746971.

connectivity is composed of a set of random variables and directed edges. The graphical structure of BN is in the form of directed acyclic graph (DAG) [10], which enables a joint distribution to be decomposed into conditional probabilities. For a set of random variables $x = \{x_i : i \in I\}$ (nodes) in the graph, the joint probability can be written as:

$$P(x_1, x_2, \dots, x_n | Q) = \prod_{i=1}^n P(x_i | a_i, Q_i) \quad (1)$$

where $Q = \{Q_i : i \in I\}$ represents the parameters of the conditional probabilities and a_i is the set of parents of node x_i .

DBN incorporates the time variable such that $x(t) = \{x_i(t) : i \in I\}$ is the set of random variables at n nodes at time $t = 1, 2, \dots, T$. The time index corresponds to the sampling instant and T is the total number of time samples. To account for the temporal interactions, probability distribution over the set of random variables has to be modeled for all t but this is computationally challenging. To simplify the problem, temporal interactions among random variables are assumed to be stationary and a first order Markov process:

$$P(x(t+1) | x(t), \dots, x(1)) = P(x(t+1) | x(t)) \quad (2)$$

The structure of DBN is obtained by unfolding the connectivity structure between two consecutive sampling instances for all t [10]. Hence, the transition network of DBN consists of two layers of n nodes connected in a structure S . The goal is to find the optimal structure S^* :

$$S^* = \arg \max_S P(S | x) \quad (3)$$

Using the marginal probability $P(x|S)$:

$$P(x|S) = \prod_{i=1}^n \int P(x_i | a_i, Q_i) P(Q_i) dQ_i \quad (4)$$

and the Bayes rule, $P(S|x)$ can be found as follows:

$$P(S|x) = \frac{P(x|S)P(S)}{\sum_S P(x|S)P(S)} \quad (5)$$

In order to capture the nonlinear dependencies between the nodes, we choose the conditional probabilities, $P(x_i | a_i, Q_i)$ to belong to the multinomial family, similar to [10], [12]. Modeling the time series as random variables with multinomial distributions requires the data to be discretized [14]. Consequently, we use a discrete dynamic Bayesian network (dDBN) to infer effective connectivity from EEG data. To sample networks with high posterior probability from (3), MCMC simulation is employed [14].

A. Implementation of the Algorithm

1) *Data Preprocessing*: For the EEG data used in this paper, each electrode is mapped to a node in the DBN. For each electrode, there are multiple trials of the EEG time series (mean of the number of trials is 90 and the sampling frequency is 500 Hz). In this paper, we focus on the response to the stimulus in the P300 window (200-600

ms after the stimulus) and the gamma frequency band (30-55 Hz). For this reason, each trial is filtered through a fourth order Butterworth bandpass filter and a window with 200 time samples, corresponding to the P300 range, is extracted. After the P300 time window and gamma frequency band are extracted from each trial, the time series is normalized such that it has zero mean and unit variance. The normalized EEG time series are discretized into a ternary form to implement the discrete DBN as follows:

$$d_i(t) = \begin{cases} 1 & \text{if } x_i(t) \geq \bar{x}_i + (x_{i,max} - \bar{x}_i)/3 \\ -1 & \text{if } x_i(t) \leq \bar{x}_i - (\bar{x}_i - x_{i,min})/3 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where \bar{x}_i , $x_{i,min}$ and $x_{i,max}$ are the mean, minimum and maximum values of the processed data from the node x_i , $i \in I$. The trials from each electrode are concatenated to form the training data for the DBN.

2) *Algorithmic Procedure*: The steps of the structure learning algorithm via MCMC simulation are:

- (i) *Initialization*: All connections of the network structure S are set to zero except for the self connections set to one in each interval between two consecutive sampling instances.
- (ii) *Burn-in Phase*: A new network structure is proposed by applying one of the elementary operations such as adding, deleting or reversing an edge between two nodes. Structures are accepted based on the Metropolis-Hastings criterion [14] and the ones violating the acyclicity constraint are discarded so that the only connections are from the variables at time t to variables at time $t + 1$. The iteration is repeated until the convergence to the true posterior probability is achieved.
- (iii) *Sampling Phase*: Same procedure is followed as in step ii and sampled structures are collected for every interval between two sampling instances of the data.
- (iv) *Inference*: Overall structure and parameters are obtained by averaging over the collected structures in step iii.

The software used in this paper consists of Bayes Net Toolbox [15] and MCMC [14] toolboxes for MATLAB.

III. DATA

A. EEG Data

Recently, it has been shown that large scale functional integration of the brain is mediated by neuronal groups which oscillate in the gamma band frequency [16]. Compared to healthy subjects, gamma band neural synchrony is shown to be deficient in schizophrenic subjects [4]. In this paper, we examined the gamma band EEG activity in the P300 window of four schizophrenic patients and four non-psychiatric control subjects who performed a continuous performance task (CPT), using the dDBN framework. The data used in this paper was collected using a 10-20 EEG system and cleaned from any muscle artifacts.

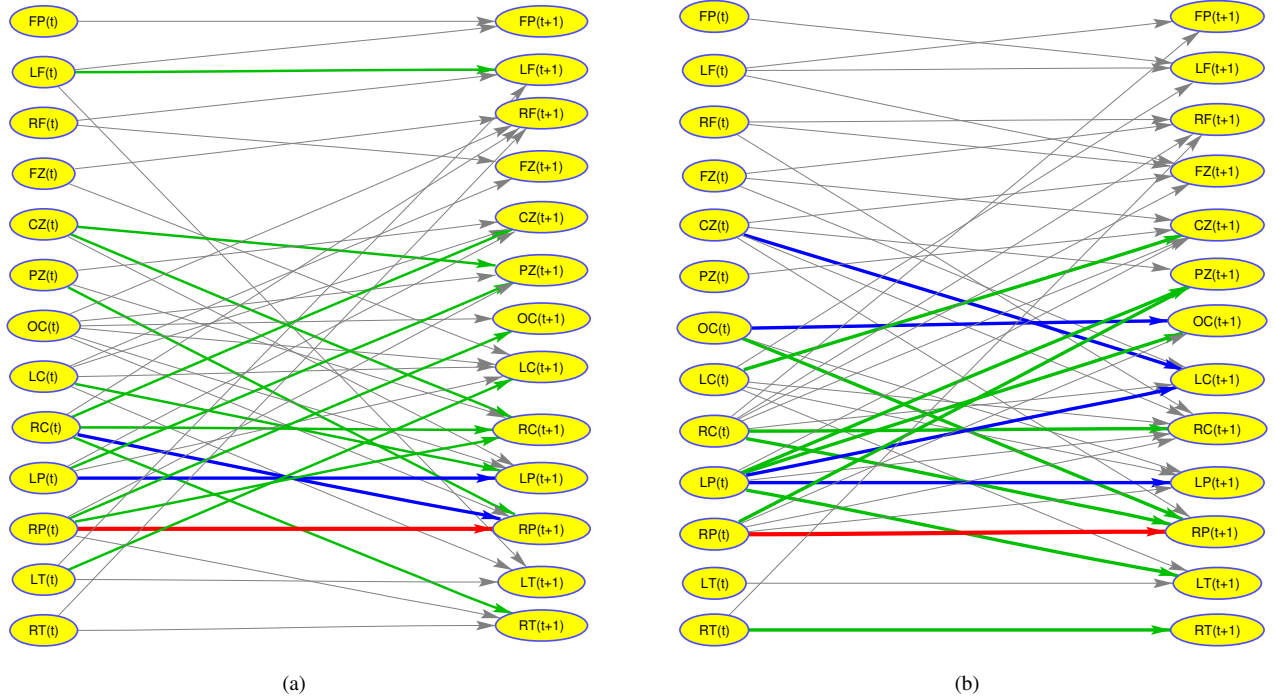


Fig. 1. Transition networks of dDBN for: (a) control and (b) schizophrenic groups. Colors of the edges (Red = 1, Blue = 0.75, Green = 0.5 and Grey = 0.25) represent the coefficients of the matrices W^c and W^s . 13 nodes are named such that they represent the brain regions as: FP (Frontal Parietal), LF (Left Frontal), RF (Right Frontal), FZ, CZ, PZ, OC (Occipital), LC (Left Central), RC (Right Central), LP (Left Parietal), RP (Right Parietal), LT (Left Temporal) and RT (Right Temporal).

B. Significance Testing

We employed Fourier bootstrapping method [17] to identify the statistically significant interactions among the electrodes. A surrogate data set is generated by first computing the Fourier transform of the data and then randomizing the phase. Finally, the inverse Fourier transform is taken to obtain the surrogate data which has the same power spectrum and autocorrelation function as the original data. For each subject, effective connectivity is inferred for 200 surrogate data sets. The number of significant interactions decreases dramatically when $p < 0.05$. For comparison purposes, we selected $p < 0.1$ and in order to determine the significant interactions at 90% significance level, a threshold value for each possible interaction is found such that 90% of the posterior probabilities obtained from the surrogate data is lower than the threshold. The same significance level is used for all possible connections.

IV. RESULTS

A structure and its parameters (posterior probabilities) are learned for each individual subject and a binary adjacency matrix of size 27×27 , which indicates the significant interactions, is obtained as follows:

$$A_l(i, j) = \begin{cases} 1 & \text{if } P_l(i, j) \geq \max(\eta_l(i, j), 0.5) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where A_l , P_l and η_l are the adjacency, posterior probability (connectivity) and threshold matrices of the l^{th} subject, respectively. Furthermore, the 27 electrodes are grouped in

13 brain regions and a new adjacency matrix, B , of size 13×13 is formed for each subject. The matrix B is formed such that if there is a significant interaction between any two electrodes that belong to two particular brain regions, then the corresponding coefficient of B is one, otherwise the coefficient is zero. After the regional adjacency matrix B is obtained for each subject, we form two summary connectivity matrices for each subject group, W^c and W^s , as follows:

$$W^c(i, j) = \frac{1}{4} \sum_{l=1}^4 B_l^c(i, j) \quad \text{and} \quad W^s(i, j) = \frac{1}{4} \sum_{l=1}^4 B_l^s(i, j) \quad (8)$$

where W^c is the connectivity matrix for the control group and W^s is the connectivity matrix for the schizophrenic group.

Fig.1 shows the effective brain connectivities, described by the weighted adjacency matrices, in the form of transition networks of dDBN for the two groups of subjects. We have identified multiple significant interactions for both groups of subjects, demonstrating that the results are not likely to be random. The total number of significant connections is similar for the two subject groups: 44 for control and 46 for schizophrenic. 70% of the connections inferred for the schizophrenic group and 66% for the control group have the value 0.25. This result demonstrates the higher variability in the connectivity patterns of schizophrenic subjects compared to the control group. Similarly, the control group has a higher number of connections with strength 0.75 indicating

consistency across the subjects. For both subject groups, there are significant connections within the parietal lobes which are directly related to visual stimulus processing and not indicative of any pathology. The most biologically significant differences between the two subject groups can be found by comparing the long-range connections. For both the control and the schizophrenic subject groups, there are significant connections between the parietal, central and temporal lobes. However, for the control group, the number of long-range connections with weights greater than or equal to 0.5, is higher with more symmetric connections compared to the schizophrenic group. Among the long-range connections, CZ-PZ and RP-RC in the control group, are the most important ones since these connections indicate the reciprocal interaction between the parietal and frontal lobes of the brain. The parietal lobe processes the visual scene and cooperates with the frontal lobe to accord attention to areas of interest and plans motor action [18]. These connections are weak or do not exist for the schizophrenic subject group. Hence, our results indicate that effective connectivity within frontal-parietal neurocognitive networks is disrupted in schizophrenia. For both subject groups, we also observe connections within a region or between closely spaced regions such as the connections between CZ and the left and right central regions which are most likely due to volume conduction in EEG recordings.

V. CONCLUSIONS

In this paper, we used dDBN to learn effective connectivity in the brain from EEG data. Effective brain connectivities are inferred for 4 control and 4 schizophrenic subjects and the significant interactions are identified through Fourier bootstrapping. The results indicate the differences between the two groups and the increased variability in the connectivity patterns for schizophrenic patients.

Future work will concentrate on identifying the linear and nonlinear portions of the effective connectivity identified through dDBN. We will also investigate the application of dDBN to neuronal sources extracted from EEG recordings to address the issue of volume conduction. Finally, we will focus on extending the analysis to a larger group of subjects, and comparing dDBN with other existing methods.

VI. ACKNOWLEDGEMENTS

We would like to thank Drs. Scott Sponheim and Edward Bernat from the University of Minnesota for sharing their EEG data.

REFERENCES

- [1] G. Tononi, O. Sporns, and G. Edelman, "A measure for brain complexity: relating functional segregation and integration in the nervous system," *Proceedings of the National Academy of Sciences*, vol. 91, no. 11, pp. 5033–5037, 1994.
- [2] L. Harrison, W. Penny, and K. Friston, "Multivariate autoregressive modeling of fMRI time series," *Neuroimage*, vol. 19, no. 4, pp. 1477–1491, 2003.
- [3] V. Smith, J. Yu, T. Smulders, A. Hartemink, and E. Jarvis, "Computational inference of neural information flow networks," *PLoS Comput Biol*, vol. 2, no. 11, p. e161, 2006.
- [4] K. Spencer, P. Nestor, R. Perlmutter, M. Niznikiewicz, M. Klump, M. Frumin, M. Shenton, and R. McCarley, "Neural synchrony indexes disordered perception and cognition in schizophrenia," *Proceedings of the National Academy of Sciences*, vol. 101, no. 49, pp. 17 288–17 293, 2004.
- [5] A. McIntosh and F. Gonzalez-Lima, "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2–22, 1994.
- [6] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, 2003.
- [7] M. Kamiński, M. Ding, W. Truccolo, and S. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biological Cybernetics*, vol. 85, no. 2, pp. 145–157, 2001.
- [8] R. Goebel, A. Roebroeck, D. Kim, and E. Formisano, "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping," *Magnetic Resonance Imaging*, vol. 21, no. 10, pp. 1251–1261, 2003.
- [9] L. Baccalá and K. Sameshima, "Partial directed coherence: a new concept in neural structure determination," *Biological cybernetics*, vol. 84, no. 6, pp. 463–474, 2001.
- [10] J. Rajapakse and J. Zhou, "Learning effective brain connectivity with dynamic Bayesian networks," *Neuroimage*, vol. 37, no. 3, pp. 749–760, 2007.
- [11] X. Zheng and J. Rajapakse, "Learning functional structure from fMR images," *Neuroimage*, vol. 31, no. 4, pp. 1601–1613, 2006.
- [12] J. Burge, T. Lane, H. Link, S. Qiu, and V. P. Clark, "Discrete dynamic bayesian network analysis of fmri data," *Hum Brain Mapp*, November 2007. [Online]. Available: <http://dx.doi.org/10.1002/hbm.20490>
- [13] J. Li, Z. Wang, S. Palmer, and M. McKeown, "Dynamic Bayesian network modeling of fMRI: A comparison of group-analysis methods," *NeuroImage*, vol. 41, no. 2, pp. 398–407, 2008.
- [14] D. Husmeier, "Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks," *Bioinformatics*, vol. 19, no. 17, pp. 2271–2282, 2003.
- [15] K. Murphy, "Dynamic Bayesian networks: representation, inference and learning," Ph.D. dissertation, University of California, 2002.
- [16] C. Tallon-Baudry and O. Bertrand, "Oscillatory gamma activity in humans and its role in object representation," *Trends in Cognitive Sciences*, vol. 3, no. 4, pp. 151–162, 1999.
- [17] D. Prichard and J. Theiler, "Generating surrogate data for time series with several simultaneously measured variables," *Physical review letters*, vol. 73, no. 7, pp. 951–954, 1994.
- [18] G. Dutton, "Cognitive vision, its disorders and differential diagnosis in adults and children: knowing where and what things are," *Eye*, vol. 17, no. 3, pp. 289–304, 2003.