33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Multivariate Synchrony Modules Identified Through Multiple Subject Community Detection in Functional Brain Networks

Marcos E. Bolaños, Edward M. Bernat, and Selin Aviyente

Abstract— **The functional connectivity of the human brain may be described by modeling interactions among its neural assemblies as a graph composed of vertices and edges. It has recently been shown that functional brain networks belong to a class of scale-free complex networks for which graphs have helped define an association between function and topology. These networks have been shown to possess a heterogenous structure composed of clusters, dense regions of strongly associated nodes, which represent multivariate relationships among nodes. Network clustering algorithms classify the nodes based on a similarity measure representing the bivariate relationships and similar to unsupervised learning is performed without a priori information. In this paper, we propose a method for partitioning a set of networks representing different subjects and reveal a community structure common to multiple subjects. We apply this community identifying algorithm to functional brain networks during a cognitive control task, in particular the error-related negativity (ERN), to evaluate how the brain organizes itself during error-monitoring.**

I. INTRODUCTION

Functional connectivity is defined as the temporal correlations between spatially remote neurophysiological events [1] and describes the neural processes required for cognitive and motor tasks. Functional connectivity has been quantified by applying coherence or nonlinear synchronization measures to various neuroimaging data. In previous work, the bivariate relationships between neuronal populations have been represented as graphs composed of vertices and edges. These graphs were then analyzed using various measures from graph theory including small-world measures and centrality measures for hub classification. One remaining issue is to identify the functional modules in these neural networks through community detection. A community structure is defined as the natural tendency of a network's vertices to divide into modules containing a dense number of intra-connecting edges within each module and a sparse number of inter-connecting edges between modules. Community detection methods can be categorized as divisive, agglomerative, or optimal where a particular objective function is maximized. These categories can include methods pertaining to spectral analysis [2], random walks [3], and min-cut problems [4].

Despite the availability of numerous clustering approaches, some fundamental problems still endure. These include the uncertainty of the number of clusters in the community, determining the optimality of a particular community structure, and representing community structure across multiple subjects. The literature often presents community detection on a single network but rarely on multiple networks. Evaluating the community structure across subjects is important for describing both the common structures and identifying individual variations. Various studies [5] [6] have searched for commonality within multiple subjects by introducing random effect analysis and Bayesian classifiers. We propose a new method for identifying this common community structure across multiple subjects by representing community membership as a probability distribution. The proposed method is based on the hierarchical application of the spectral graph clustering based on the Fiedler vector.

We apply our method to a set of graphs representing the pair-wise synchrony among neural assemblies quantified by a recently introduced time-varying phase synchrony measure. This approach is applied to EEG data collected during a study of the cognitive control in the brain based on error-related negativity. In particular, we are interested in determining the neural networks responsible for error processing. It has been hypothesized that the medial prefrontal cortex (mPFC) interacts with the lateral prefrontal cortex (lPFC) in a dynamic loop during conditions of conflict or after an error [7]. We apply the proposed community detection algorithm to determine whether such a change in the organization of the brain occurs after an error response compared to a correct response.

II. METHODS

A. Time-Varying Measure of Phase Synchrony

Recently [8], we have introduced a new time-varying measure of phase synchrony based on the complex time-frequency distribution known as the Rihaczek distribution [9]. For a signal, $x(t)$, Rihaczek distribution is expressed as

$$C(t,\omega) = \frac{1}{\sqrt{2\pi}} x(t) X^*(\omega) e^{-j\omega t} \tag{1}$$

and measures the complex energy of a signal at time t and frequency ω.

One of the disadvantages of Rihaczek distribution is the existence of cross-terms for multicomponent signals. In order to get rid of these cross-terms, we introduced a reduced interference version of Rihaczek distribution by applying a kernel function such as the Choi-Williams (CW) kernel with $\phi(\theta,\tau) = \exp(\frac{-(\theta\tau)^2}{\sigma})$ to filter the cross-terms to obtain

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

M. E. Bolaños (bolanosm@msu.edu) and S. Aviyente (aviyente@egr.msu.edu) are with the Department of Electrical Engineering, Michigan State University, 2120 Engineering Building, East Lansing, MI, 48824, USA.

E. M. Bernat is with the Department of Psychology, Florida State University, 600 W. College Avenue, Tallahassee, FL 32306, USA.

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

$$C(t,\omega) = \iint \exp\left(\frac{-(\theta\tau)^2}{\sigma}\right) \exp\left(j\frac{\theta\tau}{2}\right) A(\theta,\tau) e^{-j(\theta t+\tau\omega)} d\tau d\theta \quad (2)$$

where $A(\theta,\tau) = \int x(u+\frac{\tau}{2})x^*(u-\frac{\tau}{2})e^{j\theta u}du$ is the ambiguity function of the signal and $exp(j\theta\tau/2)$ is the kernel corresponding to the Rihaczek distribution [9]. The phase difference between two signals based on this complex distribution is computed as

$$\Phi_{12}(t,\omega) = \arg\left[\frac{C_1(t,\omega)C_2^*(t,\omega)}{|C_1(t,\omega)||C_2(t,\omega)|}\right] \quad (3)$$

and a synchrony measure quantifying the intertrial variability of the phase differences, phase locking value (PLV), is defined as

$$PLV(t,\omega) = \frac{1}{N}\left|\sum_{k=1}^{N} \exp(j\Phi_{12}^k(t,\omega))\right| \quad (4)$$

where N is the number of trials and $\Phi_{12}^k(t,\omega)$ is the time-varying phase estimate between two electrodes for the kth trial. If the phase difference varies little across the trials, PLV is close to 1.

B. Graph Measures

A graph is defined as $G = (V,E)$ where V is the set of N vertices and E is the set of edges assigned to a node pair, v_i and v_j. For weighted graphs, G is represented by the weighted connectivity matrix W such that w_{ij} is the similarity between nodes v_i and v_j. Our proposed method for community detection utilizes the Laplacian matrix of a graph which has been shown to represent a more accurate representation of the network than the connectivity matrix alone [10]. We define the Laplacian for a weighted connectivity matrix as

$$L_{ij} = \begin{cases} \sum_j W(i,j) & \text{if } i = j \\ -W(i,j) & \text{if } i \neq j \end{cases} \quad (5)$$

The spectral decomposition of the Laplacian matrix has been shown to carry important information about the community structure [2], particularly the eigenvector associated with the second smallest non-zero eigenvalue referred to as the Fiedler vector, u_F. The Fiedler vector is a solution to the minimum cut problem of a graph such that most inter-edges (links between clusters) are removed and intra-edges (links within a cluster) are retained. This method allows a network's vertices to be categorized into two clusters, C_1 and C_2, based on the signs of their corresponding elements within u_F such that

$$v_i \in \begin{cases} C_1 & \text{if } u_F(i) \geq 0 \\ C_2 & \text{if } u_F(i) < 0 \end{cases} \quad (6)$$

This minimum cut splits the network into only two clusters but may be repeated iteratively to each successive cluster to find multiple communities. The divisions may continue until a preselected number of clusters is attained or until an optimization criteria such as modularity [11] is satisfied.

C. Optimal Community Structure

The primary problem with iteratively dividing a network is determining the optimal number of clusters. The most commonly used optimization measure, modularity [11], compares a community structure to the expected community structure of a random graph such that there exists a high number of edges within clusters and low number of edges between clusters. Modularity for weighted graphs is defined as

$$Q = \frac{1}{2m}\sum_{ij}\left[W_{ij} - \frac{s_i s_j}{2m}\right]\sigma(i,j) \quad (7)$$

such that m is the number of edges, s_i (indegree) is the sum of the ith row in W, and $\sigma(i,j) = 1$ if v_i and v_j are in the same cluster and 0 otherwise. Modularity ranges in $[-1,1]$ with the highest modularity corresponding to a 'good' community structure. Modularity does not always result in the highest value for the true community structure [12] and can reveal a suboptimal structure which may be due to the simplicity of the random model computed through $\frac{s_i s_j}{2m}$ in Equation 7. Therefore, we propose to compare the detected weighted community structure to numerous randomly weighted community structures. It is expected that vertex pairs connected with large weights are more likely to be in the same cluster as opposed to pairs connected with a low weight. We define a community matrix as

$$C^k(i,j) = \begin{cases} W(i,j) & \text{if } v_i \text{ and } v_j \text{ are in the same cluster} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where k is the number of clusters. The random community matrix is obtained by randomly assigning each edge to one of k clusters and computing equation 8 for the random assignments to obtain $C_{rand}^{(k,p)}$, such that $p = \{1,2,...,r\}$ is the number of permutations. Membership size of each cluster may differ between the random model and the detected community structure for each p, but the number of clusters must be the same. A large p is recommended such that numerous random clustering possibilities are evaluated. The average inter-cluster weight corresponding to each clustering matrix is computed as

$$\hat{C}^k = \frac{1}{N(N-1)}\sum_{ij} C^k(i,j) \quad (9)$$

and

$$\hat{C}_{rand}^{(k,p)} = \frac{1}{N(N-1)}\sum_{ij} C_{rand}^{(k,p)}(i,j) \quad (10)$$

Averaging over all p such that $\hat{C}_{rand}^k = \frac{1}{|p|}\sum_p \hat{C}_{rand}^{(k,p)}$ provides the random inter-cluster value. We then find k such that $Q = \hat{C}^k - \hat{C}_{rand}^k$ is maximum. A maximum value indicates that the vertices within a cluster share significantly larger weights than if they were clustered with any other set of vertices. This method relies heavily on weights and demonstrates how weight distribution is a significant contributor to community structure.

D. Multiple Subject Clustering Algorithm

When multiple graphs representing similar information are available for clustering analysis it would be advantageous to develop a probabilistic clustering technique by determining which nodes are more likely to cluster across all graphs, or all subjects. Therefore, we introduce a probability of clustering matrix, P, where the entries of $P(i,j)$ are values between 0 and 1, indicating the likeliness that a node pair is in the same cluster across multiple graphs. Given m weighted connectivity matrices, the Fiedler partitioning method can be applied to each matrix thus separating the nodes into two distinct cluster sets. The probability matrix, P, is introduced to keep track of how many times, out of m similarity matrices, nodes v_i and v_j are placed in the same cluster set. We then compute the Laplacian of P and identify the Fiedler vector and form a bi-partition of P into a community structure composed of clusters c_1 and c_{-1}. Since P represents the clustering relationships among nodes throughout all subjects, the Fiedler partition of P represents the common community structure to all subjects. The initial partition set, $C = \{c_1, c_{-1}\}$, contains $k = 2$ clusters but if $k > 2$ is desired, the process can be repeated by selecting a cluster in C to partition. In this case, c_1 or c_{-1} may be selected based on the cardinality of their nodal memberships. Next, sub-matrices are extracted from the original weighted connectivity matrices such that they only contain the nodes of the chosen cluster. These sub-matrices are used to derive the new probability of clustering matrix, P^y, where $y = 1$ if cluster c_1 was selected and $y = -1$ if cluster c_{-1} was selected. The Fiedler partition will result in two new clusters, c_1^y and c_{-1}^y. The final cluster set is $C = \{c_{-y}, c_1^y, c_{-1}^y\}$ which is a concatenation of the two new clusters with the original cluster which was not chosen for bi-partitioning. Algorithm 1 describes this process for obtaining community structures for a given number of clusters.

Algorithm 1 Probabilistic Fiedler Clustering Algorithm

Input: m $N \times N$ dimensional graphs, $G = \{G^1, G^2, ..., G^m\}$ with vertices $V = \{v_1, v_2, .., v_N\}$ and edges $E^r = \{w_{ij}^r : v_i,\ v_j \in V\}$ such that $G^r = (V, E^r)$ and $r = \{1, 2, ..., m\}$.
Input: Number of clusters, k.
Output: k clusters $C = \{c_1, c_2, ..., c_k\}$ where $c_j \subset V$.
$C = \emptyset$
for $t = 2$ to k **do**
  $M = \mathbf{0}_{|\mathbf{V}| \times |\mathbf{V}|}$
  for $s = 1$ to $|G|$ **do**
    submatrix $\hat{G}^s \subset G^s | \hat{G}^s = (V, E^s)$
    $(V_1, V_2) =$ SubRoutine(Fiedler Partition($\hat{G}^s$))
    $M(i,j) = M(i,j) + \sigma(v_i, v_j)$ where
$$\sigma(v_i, v_j) = \begin{cases} 1 & \text{if nodes } v_i, v_j \in V_1 \text{ or } v_i, v_j \in V_2 \\ 0 & \text{otherwise} \end{cases}$$
  end for
  $P = \frac{M}{|G|}$
  $(V_1, V_2) =$ SubRoutine(Fiedler Partition(P))
  $C = C \cup \{V_1, V_2\}$.
  if $t \neq k$ **then**
    $V = c_i | i = \max_q\{|c_q|\}$ and $c_q \in C$
    $E^r = \{w_{ij}^r : v_i,\ v_j \in V\}$
    $C = C \backslash \{c_i\}$
  end if
end for

Algorithm 2 Fiedler Partition

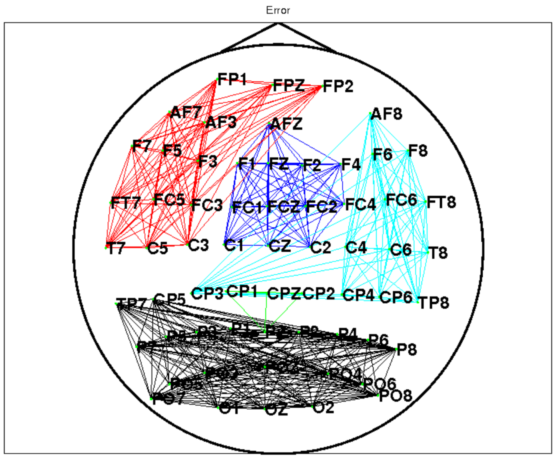
Input: graph $G = (V, E)$.
Output: Vertex sets V_1 and V_2.
Compute Laplacian Matrix, L, of G.
Compute $|V|$ eigenvectors, $\mathbf{u}$, and eigenvalues, λ, of L.
Order eigenvalues in ascending order: $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{|V|}$.
$u_F = u_i$ where $i = \min_q\{\lambda_q\} | \lambda_q \neq 0$.
for $j = 1$ to $|V|$ **do**
$$v_j \in \begin{cases} V_1 & \text{if } u_F(j) \geq 0 \\ V_2 & \text{if } u_F(j) < 0 \end{cases}$$
end for

III. Results

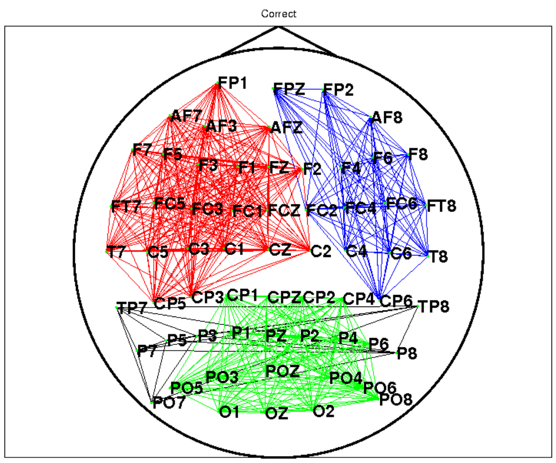
In this paper, we are interested in understanding the neural networks involved during error-related negativity which is a brain potential response that occurs following performance errors in a speeded reaction time task. Ninety undergraduate students (55 male) were recruited from introductory psychology courses at the University of Minnesota whose EEG was recorded during a speeded-response flanker task (as previously reported with amplitude-only effects [13]). The proposed phase synchrony measure was applied to this set of EEG data containing the error-related negativity (ERN). Both error (ERN) and correct (CRN) response-locked averages were computed for each subject , with the number of trials matched between error and correct for each subject. Our previous work indicates that there is increased phase synchrony associated with ERN for the theta frequency band (4-7 Hz) and ERN time window (25-75 ms) for Error responses compared to Correct responses [8].

The multiple subject clustering method was then applied to the set of Error and Correct data in order to divide the nodes into $2 \leq k \leq 20$ clusters. We applied the proposed optimization measure for identifying the best community structure illustrated in Figures 1a and 1b.

Our community structure quality measure revealed Error data were best represented by a community structure composed of 5 clusters; whereas Correct data by 4 clusters. The high number of clusters is indicative of increased inhomogeneity throughout the network during error responses versus correct responses. The community structure in Figure 1a is composed of three groups containing highly synchronized node pairs in the frontal region of the brain with a cluster located in the mPFC and two on the lPFC. The existence of the central cluster in this community structure may be indicative of the ongoing action-monitoring processes oc-



(a)



(b)

Fig. 1: Optimal community structures representing the brain during error responses (top) and correct responses (bottom).

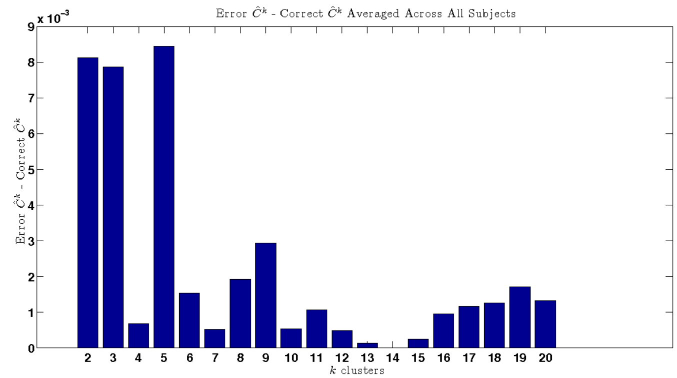


Fig. 2: Optimal community structures representing the brain during error responses (top) and correct responses (bottom).

curring in the structure, as there is no distinct cluster in the mPFC for the Correct data. As expected, the parietal regions did not show distinct clusters in either case since most activity during the task-response test is observed in the the frontal region of the brain. Finally, we compared the node pair strengths within the clusters across response types. We evaluated $\hat{C}^k$ differences between Error and Correct across all subjects with respect to the number of clusters, k (Figure 2). The plot indicates all community structures derived for Error responses showed increased synchronization within the clusters as opposed to the Correct responses, particularly for $k = 2, 3$, and 5. These results show that if the discrimination between the two response types is used as an optimization criterion, $k = 5$ would be the best community structure to use.

IV. Conclusions

In this paper, we have introduced a cluster detection algorithm for discovering a common community structure for multiple networks. We also introduced a new community quality measure in lieu of modularity. The community detection algorithm and quality measure were used to identify an optimal community structure representing the multivariate relationships among neural assemblies of the brain. Our algorithm may be applied to any set of multiple subjects represented by weighted or unweighted graphs. The future work will focus using this clustering method for classifying between the two response types.

References

[1] L. Lee, L. M. Harrison, and A. Mechelli, "A report of the functional connectivity workshop," *NeuroImage*, 2003.

[2] U. Luxburg, "A tutorial on spectral clustering," *CoRR*, 2007.

[3] Y. Deng, V. Chokalingam, and C. Zhang, "Markov chain correlation based clustering of gene expression data," in *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II - Volume 02*, ser. ITCC '05. Washington, DC, USA: IEEE Computer Society, 2005, pp. 750–755.

[4] E. L. Johnson, A. Mehrotra, and G. L. Nemhauser, "Min-cut clustering," *Mathematical Programming*, vol. 62, pp. 133–151, 1993.

[5] I. Rustandi, "Hierarchical gaussian naive bayes classifier for multiple-subject fmri data," in *In NIPS Workshop: New Directions on Decoding Mental States from fMRI Data*, 2006.

[6] W. D. Penny and A. J. Holmes, "Random effects analysis," 2004.

[7] J. F. Cavanagh, M. X. Cohen, and J. J. B. Allen, "Prelude to and resolution of an error: Eeg phase synchrony reveals cognitive control dynamics during action monitoring," *The Journal of Neuroscience*, vol. 29, no. 1, pp. 98–105, 2009.

[8] S. Aviyente, E. M. Bernat, W. S. Evans, and S. R. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human Brain Mapping*, vol. 32, no. 1, pp. 80–93, 2011.

[9] A. Rihaczek, "Signal energy distribution in time and frequency," in *IEEE Transactions: Information Theory*, vol. 14, 1968, pp. 369–274.

[10] J. Forman, P. Clemons, S. Schreiber, and S. Haggarty, "Spectralnet - an application for spectral graph analysis and visualization," *BMC Bioinformatics*, vol. 6, no. 1, p. 260, 2005.

[11] M. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, p. 8577, 2006.

[12] K. Steinhaeuser and N. V. Chawla, "Identifying and evaluating community structure in complex networks," *Pattern Recognition Letters*, vol. 31, no. 5, pp. 413 – 421, 2010.

[13] J. R. Hall, E. M. Bernat, and C. J. Patrick, "Externalizing psychopathology and the error-related negativity," in *Psychological Science*, vol. 18(4), 2007, pp. 326–333.