SUBSPACE ANALYSIS FOR CHARACTERIZING DYNAMIC FUNCTIONAL BRAIN NETWORKS

Ali Yener Mutlu and Selin Aviyente

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

ABSTRACT

Human brain is known to be one of the most complex biological systems and understanding the functional connectivity patterns to distinguish between normal and disrupted brain behavior still remains a challenge. Previous studies focus on analyzing functional connectivity averaged over a certain time and frequency window which is generally not sufficient to address the time-varying evolution of the connectivity patterns. In this paper, we propose a framework to describe the dynamic properties of functional connectivity in the brain. The proposed approach is based on constructing time-varying connectivity graphs from multichannel electroencephalogram (EEG) data, using subspace analysis to detect network-wide changes, identifying key event intervals and then extracting representative networks that describe the connectivity in each event interval. This framework is evaluated for EEG data, containing error-related negativity (ERN) component related to cognitive control. For each time interval, the statistically significant connectivity patterns are presented to illustrate the dynamic nature of functional connectivity.

Index Terms— Dynamic graphs, time-varying functional brain networks, dynamic network summarization

1. INTRODUCTION

In recent years, there has been a growing interest in studying the functional connectivity of the human brain, which is defined as the statistical dependencies among remote neurophysiological events [1]. Functional connectivity can be inferred from different neuroimaging data such as the electroencephalography (EEG), which offers high temporal resolution needed to quantify the time-varying relationships between neuronal populations [2]. Various measures, such as cross-correlation, Granger causality and phase synchrony, have been used for inferring the functional relationships among different brain sites [3]. However, these measures are limited to quantifying bivariate relationships and cannot reveal the collective behavior of complex brain networks.

Due to the need for understanding the underlying topologies of brain networks, a new and multidisciplinary approach to the study of complex systems based on graph theory has become popular [4]. Graph theory offers a way to quantify the multivariate relationships among neuronal activations across brain regions. For instance, Watts and Strogatz have shown that graphs with many local connections and a few random long distance connections, called 'small-world' networks, have both clustered ('cliquish') interconnectivity within groups of nodes (like regular lattices) and a short path length between any two nodes (like random graphs) [5]. Recently, there have been multiple studies which have shown small-world patterns in functional networks of healthy subjects [6]. Several studies have also shown how brain pathology, such as schizophrenia and Alzheimer's

diseases, can be related to disrupted small-world architecture [6]. Currently, graph theoretical features of functional networks, such as clustering coefficient, path length, small world parameter and modularity [4], are routinely applied to quantify the organization of these networks. However, the current studies are limited to static brain networks and thus, neglect possible time-varying properties of the topologies. A single graph is not sufficient to represent the connectivity patterns of the brain and can be considered as an unreliable snapshot of functional connectivity.

More recently, understanding the dynamic evolution of functional brain networks has been of interest. Current studies of dynamic functional brain networks are limited to extracting graph theoretic features such as modularity, small-world parameter and efficiency from each graph in the series and track the evolution of these parameters [7, 8, 9]. However, these approaches lose the spatial information provided by the graphs and cannot identify which parts of the brain contributed to the observed changes in the network. In recent years, dynamic network analysis has attracted some attention in the signal processing community. Approaches to detect anomalies or distinct subgraphs in large, noisy background [10] and to track dynamic networks [11, 12] have been proposed. For instance, in [13], direction of the principal eigenvector of a matrix based on the graph is tracked over time, and an anomaly is detected if the direction changes by more than some threshold. Tracking dynamic networks [11] using shrinkage estimation, or simple approaches such as sliding window or exponentially weighted moving averaging have also been proposed for inferring long-term information or trends [14]. However, these methods either do not offer summarization of the time-varying network patterns or do not reduce the multitude of graphs into a few representative networks.

In this paper, the goal is to reduce information from multiple connectivity graphs across time into a few meaningful time intervals over which the temporal variation of network topology is minimum and to represent these intervals with key graphs that best describe the connectivity patterns. The proposed work builds on the prior work on time-varying network analysis such that significant network-wide changes are tracked and the evolution of connectivity patterns are summarized. Subspace analysis using principal component analysis (PCA) is employed to project the high dimensional time-varying connectivity information to a lower dimensional space, which is equivalent to decorrelating the network in the spatial domain. Hence, the resulting key networks contain only the information related to the signal subspace.

2. BACKGROUND

2.1. Notation

We use uppercase bold letters, such as $\boldsymbol{X}$, to denote a matrix and lowercase bold letters, such as $\boldsymbol{x}$, to denote a vector, where $\boldsymbol{X}(t)$ and $\boldsymbol{x}(t)$ represent the matrix and vector at time point t, respectively. $X_{i,j}(t)$ represents the entry of matrix $\boldsymbol{X}(t)$ at the ith row

and j^{th} column, whereas $x_i(t)$ denotes the ith entry of vector $\boldsymbol{x}(t)$. $\|\boldsymbol{x}(t)\|_p = (\sum_{i=1}^{k} |x_i(t)|^p)^{1/p}$ is the l_p norm of a k-dimensional vector $\boldsymbol{x}(t)$.

2.2. Network Change Detection Using PCA

PCA is a dimensionality reduction technique that results in a compact representation of a multivariate data set by projecting the data onto a lower dimensional subspace defined by a set of new axes called principal components (PCs). Each PC points in the direction of maximum variance remaining in the data, given the variance already accounted for in the preceding components. Lakhina *et al.* were the first to use PCA as a tool to detect network traffic anomalies, in particular for computer network data streams, and PCA based network wide anomaly detection has been extensively investigated in [15, 16].

Let $\boldsymbol{X} = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(m)]$ be a $n \times m$ network traffic time-series data matrix, centered to have zero mean, where $\boldsymbol{x}(t)$ is an n-dimensional data vector at time point t. Then, the set of n principal components, $\{\boldsymbol{v_i}\}_{i=1}^n$, are defined as:

$$\boldsymbol{v_i} = \arg \max_{\|\boldsymbol{v}\|_2=1} \left\| \left(\boldsymbol{X} - \sum_{j=1}^{i-1} \boldsymbol{X} \boldsymbol{v_j} \boldsymbol{v_j}^T \right) \boldsymbol{v} \right\| \quad (1)$$

Solution of Eq. (1) is given by the eigenvectors of the covariance matrix, $\boldsymbol{C} = \frac{1}{m} \boldsymbol{X} \boldsymbol{X}^T$, which form an $n \times n$ matrix $\boldsymbol{V} = [\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}]$ having associated eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ arranged in decreasing order.

It has been observed that although the original data spans a high dimensional space, normal traffic patterns lie in a low dimensional signal (normal) subspace spanned by the first l PCs corresponding to the l largest eigenvalues and anomalous behavior lies in a noise (anomalous) subspace spanned by the remaining $(n - l)$ PCs [15]. Hence, the data at time point t, $\boldsymbol{x}(t)$, can be decomposed as:

$$\boldsymbol{x}(t) = \boldsymbol{x}_{\text{Sig}}(t) + \boldsymbol{x}_{\text{Noi}}(t) \quad (2)$$

where $\boldsymbol{x}_{\text{Sig}}(t)$ and $\boldsymbol{x}_{\text{Noi}}(t)$ correspond to the signal and noise components, respectively, and can be computed as:

$$\boldsymbol{x}_{\text{Sig}}(t) = \boldsymbol{P} \boldsymbol{P}^T \boldsymbol{x}(t) \text{ and } \boldsymbol{x}_{\text{Noi}}(t) = (\boldsymbol{I} - \boldsymbol{P} \boldsymbol{P}^T) \boldsymbol{x}(t) \quad (3)$$

where $\boldsymbol{P} = [\boldsymbol{v_1}, \ldots, \boldsymbol{v_l}]$ consists of the eigenvectors with the largest l eigenvalues and $\boldsymbol{I}$ is the identity matrix. When there is a large change in the noise component, $\boldsymbol{x}_{\text{Noi}}(t)$, an anomalous network wide behavior is declared [15]. By performing statistical testing using a Q-statistic on the squared prediction error, $\|\boldsymbol{x}_{\text{Noi}}(t)\|_2$, one can determine whether an anomaly is observed or not [15].

3. METHODS

3.1. Forming Time-Varying Graphs Using Phase Synchrony

In order to identify event intervals and infer the evoked network activity within each interval, we first need to obtain time-varying graphs representing the interactions across different brain sites. The nodes of the graphs correspond to different brain regions and edges correspond to the connectivity strengths. In this paper, we quantify the connectivity using a recent phase synchrony measure based on RID-Rihaczek distribution [17]. This measure has been shown to be more robust to noise and to provide better resolution as discussed in [17]. The first step in quantifying phase synchrony is to estimate the time and frequency dependent phase, $\Phi_i(t, \omega)$, of a signal, $\boldsymbol{s_i}$,

as: $\Phi_i(t, \omega) = \arg\left[\frac{C_i(t, \omega)}{|C_i(t, \omega)|} \right]$ where $C_i(t, \omega)$ is the complex RID-Rihaczek distribution[1]:

$$C_i(t, \omega) = \int \int \underbrace{\exp\left(-\frac{(\theta\tau)^2}{\sigma} \right)}_{\text{Choi-Williams kernel}} \underbrace{\exp(j\frac{\theta\tau}{2})}_{\text{Rihaczek kernel}} A_i(\theta, \tau) e^{-j(\theta t + \tau \omega)} d\tau d\theta \quad (4)$$

and $A_i(\theta, \tau) = \int s_i(u + \frac{\tau}{2}) s_i^*(u - \frac{\tau}{2}) e^{j\theta u} du$ is the ambiguity function of $\boldsymbol{s_i}$. The phase synchrony between nodes i and j at time t and frequency ω is computed using 'Phase Locking Value' (PLV):

$$PLV_{i,j}(t, \omega) = \frac{1}{L} \left| \sum_{k=1}^{L} \exp\left(j\Phi_{1,2}^k(t, \omega) \right) \right| \quad (5)$$

where L is the number of trials and $\Phi_{i,j}^k(t, \omega) = |\Phi_i(t, \omega) - \Phi_j(t, \omega)|$ is the phase difference estimate between the two nodes for the kth trial.

Let $\{\boldsymbol{G}(t)\}_{t=1,2,\ldots,T}$ be a time sequence of graphs where $\boldsymbol{G}(t)$ is an $N \times N$ weighted and undirected graph corresponding to the functional connectivity network at time t for a fixed frequency or frequency band, T is the total number of time points and N is the number of nodes within the network. The time-varying edge values are quantified by the average PLV within a frequency band and at a certain time as:

$$G_{i,j}(t) = \frac{1}{\Omega} \sum_{\omega=\omega_a}^{\omega_b} PLV_{i,j}(t, \omega) \quad (6)$$

where $G_{i,j}(t) \in [0, 1]$ represents the connectivity strength between the nodes i and j within the frequency band of interest, $[\omega_a, \omega_b]$, and Ω is the number of frequency bins in that band. Since the graphs are undirected and symmetric, we create vectors, $\{\boldsymbol{x}(t)\}_{t=1,2,\ldots,T}$, to equivalently represent these graphs where $\boldsymbol{x}(t)$ is an $\binom{N}{2}$-dimensional column vector obtained by stacking the columns of the upper triangular portion of $\boldsymbol{G}(t)$.

3.2. Event Interval Detection

Once the time-varying graphs and corresponding column vectors, $[\boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)]$, are obtained, we need to identify network-wide changes to determine time intervals which may correspond to the underlying neurophysiological events. We define an event interval as a time window during which the similarity between consecutive graphs is maximized. Since the spatial correlation in the network, such as long term trends in connectivity, may attenuate the effect of individual short-term variations or the unique features of each edge, the edge values are first projected to a lower dimensional signal subspace through an orthogonal projection operator, $\boldsymbol{P}$. $\boldsymbol{P}$ is obtained from the l eigenvectors of the covariance matrix, $\boldsymbol{C} = \frac{1}{T}[\boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)][\boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)]^T$, corresponding to the largest l eigenvalues such that:

$$\frac{\sum_{j=1}^{l} \lambda_j}{\sum_{j=1}^{\binom{N}{2}} \lambda_j} \times 100 \geq 90 \quad (7)$$

Once the edge values are projected, i.e., $\boldsymbol{P}^T \boldsymbol{x}(t)$, the similarity between consecutive time points needs to be quantified. Common approaches to quantifying similarity include correlation, l_2 or l_∞ norm or information theoretic measures [18]. These measures either favor

---

[1]The details of the RID-Rihaczek distribution and the corresponding synchrony measure are given in [17].

edges with large values, e.g., l_∞ norm favors edges with maximum difference, or require large amount of data for computing empirical histograms, e.g., information theoretic measures. For this reason, in this paper we propose tracking changes based on the normalized projected vectors, $[\boldsymbol{y}(1), \ldots, \boldsymbol{y}(T)]$, where $\boldsymbol{y}(t) = \frac{\boldsymbol{P}^T \boldsymbol{x}(t)}{\|\boldsymbol{P}^T \boldsymbol{x}(t)\|_2}$. The time-varying angle of these new set of vectors indicates the degree of alignment of the unique short-term features of the edge values with the direction of the principal eigenvectors. The angular similarity, $a_t = \boldsymbol{y}^T(t-1)\boldsymbol{y}(t)$, between these vectors is tracked to detect event intervals. If there is an abrupt decrease in the angular similarity between the subsequent direction vectors, this would indicate a significant change in the network patterns. Hence, we detect event intervals as follows:

$$E_t = \begin{cases} 1, & \text{if } a_t < \mu_t - 3\sigma_t \\ 0, & \text{if } a_t \geq \mu_t - 3\sigma_t \end{cases} \quad (8)$$

where a change detection algorithm based on adaptive thresholding is used. An event boundary is detected, $E_t = 1$, at time t depending on the deviation of a_t from the mean, $\mu_t = \frac{1}{\delta} \sum_{k=1}^{\delta} a_{t-k}$, by three σ_t, $\sigma_t = \sqrt{\frac{1}{\delta} \sum_{k=1}^{\delta} (a_{t-k} - \mu_t)^2}$. The length of the moving average window, δ, can be chosen based on the sampling frequency and total number of time samples, T.

3.3. Summarizing Signal Subspace for Key Graph Estimation

After determining the event intervals, our goal is to form key graphs which best summarize the particular intervals. An ideal key graph should describe the particular interval with minimal redundancy. In this paper, we eliminate the spatial correlation in the network, i.e., noise subspace, within that interval and consider only the features of the edges aligned with the principal eigenvectors.

Let $[\boldsymbol{x}(i), \ldots, \boldsymbol{x}(i+m-1)]$ be the set of m vectors obtained from the time-varying graphs that compose a detected event interval. The corresponding projection matrix, $\boldsymbol{P_i}$, is constructed using the eigenvectors corresponding to the largest l eigenvalues of the sample covariance matrix formed from $[\boldsymbol{x}(i), \ldots, \boldsymbol{x}(i+m-1)]$ such that 90% of the total energy is captured similar to section 3.2. For each time point within the particular event interval, signal subspace component is extracted as:

$$\boldsymbol{x}_{Sig}(t) = \boldsymbol{P_i}\boldsymbol{P_i}^T\boldsymbol{x}(t) \ \ t = i, \ldots, i+m-1 \quad (9)$$

For each event interval, we compute a weighted mean vector:

$$\overline{\boldsymbol{x}}_{Sig} = \frac{\sum_{t=i}^{i+m-1} SNRI_t \boldsymbol{x}_{Sig}(t)}{\sum_{t=i}^{i+m-1} SNRI_t} \quad (10)$$

where $SNRI_t$ is a time-dependent signal to noise ratio index:

$$SNRI_t = \frac{\|\boldsymbol{P_i}\boldsymbol{P_i}^T\boldsymbol{x}(t)\|_2}{\|(\boldsymbol{I} - \boldsymbol{P_i}\boldsymbol{P_i}(t)^T)\boldsymbol{x}(t)\|_2} = \frac{\|\boldsymbol{x}_{Sig}(t)\|_2}{\|\boldsymbol{x}_{Noi}(t)\|_2} \quad (11)$$

The $\overline{\boldsymbol{x}}_{Sig}$ in equation (9) weighs time points with higher SNR more, thus emphasizing signal subspace activity. The corresponding key graph is obtained by reshaping $\overline{\boldsymbol{x}}_{Sig}$ such that it constitutes the upper triangular part of the symmetric key graph for the particular interval.

4. RESULTS

4.1. EEG Data

The proposed framework is applied to a set of EEG data containing the error-related negativity (ERN). The ERN is a brain potential response that occurs following performance errors in a speeded

reaction time task usually 25-75 ms after the response [19]. Previous work [20] indicates that there is increased coordination between the lateral prefrontal cortex (lPFC) and medial prefrontal cortex (mPFC) within the theta frequency band (4-8 Hz) and ERN time window (25- 75 ms), supporting the idea that frontal and central electrodes are functionally integrated during error processing. EEG data from 62-channels was collected in accordance with the 10/20 system on a Neuroscan Synamps2 system (Neuroscan, Inc.). A speeded-response flanker task was employed, and response-locked averages were computed for each subject. All EEG epochs were converted to current source density (CSD) using published methods [21]. In this paper, we analyzed data from 90 subjects corresponding to the error responses.
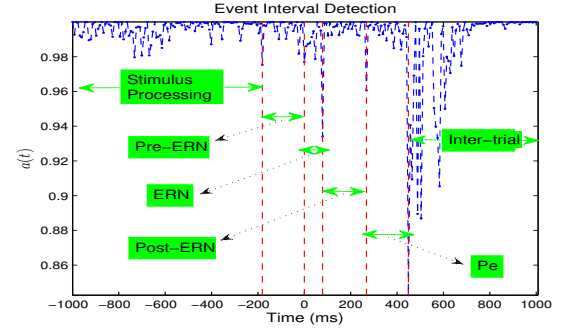

Fig. 1. Event interval detection: 6 event intervals are identified which correspond to the stimulus processing (-1000 to -179 ms), pre-ERN (-178 to 0 ms), ERN (1 to 94 ms), post-ERN (95 to 281 ms), Pe (282 to 462 ms) and inter-trial (463 to 1000 ms) intervals, respectively. The subjects respond to the stimulus at time 0 ms and the red lines indicate $E(t) = 1$.

4.2. Network-wide Change Detection

For each subject, time-varying graphs, $\boldsymbol{G}_q(t)$, $t = 1, \ldots, 256$, for the qth subject are computed using Eq. (6) where the number of nodes, N, is equal to 62, the frequency band of interest is the theta band (4-8 Hz) and the sampling frequency is 128 Hz. We compute a mean time-varying graph sequence, $\bar{\boldsymbol{G}}(t)$, over all subjects as:

$$\bar{\boldsymbol{G}}(t) = \frac{1}{90} \sum_{q=1}^{90} \boldsymbol{G}_q(t) \quad (12)$$

and the event interval detection algorithm is applied to this average sequence, $\bar{\boldsymbol{G}}(t)$, where the length of the moving average window, $\delta = 50$ milliseconds, is 5% of the sampling period and chosen such that both the abrupt changes in the connectivity patterns are detected and over-smoothing is avoided. We detected 6 event intervals using Eq. (8) which are consistent with the speeded reaction time task and correspond to the stimulus processing (-1000 to -179 ms), pre-ERN (-178 to 0 ms), ERN (1 to 94 ms), post-ERN (95 to 281 ms), Pe (282 to 462 ms) and inter-trial (463 to 1000 ms) intervals, respectively, as shown in Fig. 1.

The initial event interval corresponds to the complex processing of the stimulus before making a response. The Pre-ERN and Post-ERN intervals index activity around the incorrect motor response. The ERN and Pe intervals are detected successfully, where the Pe (error-positivity) interval corresponds to positive evoked potential similar to the well-known P300 and is observed subsequent to the incorrect response.
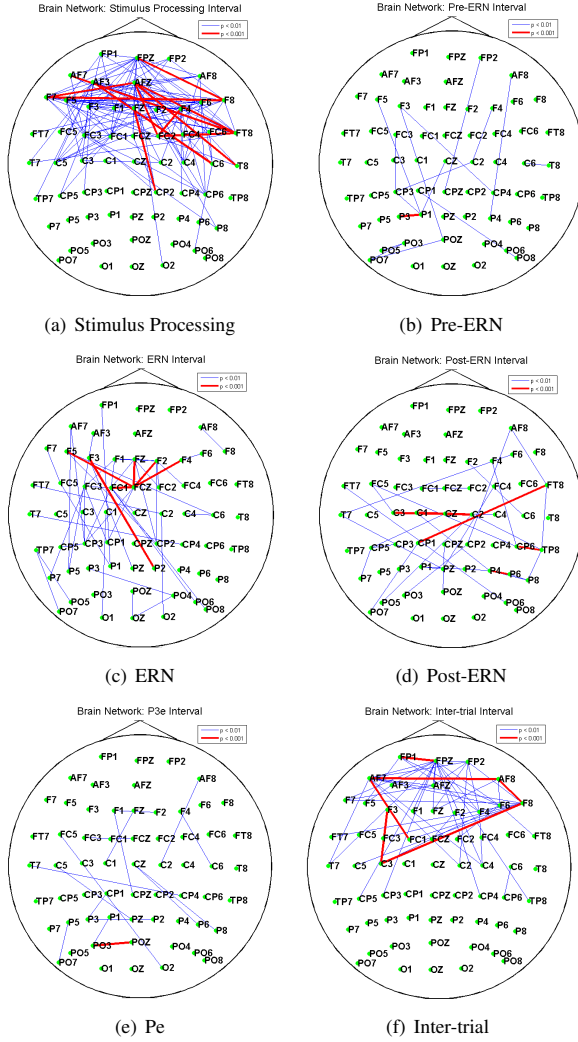
(a) Stimulus Processing   (b) Pre-ERN

(c) ERN   (d) Post-ERN

(e) Pe   (f) Inter-trial

Fig. 2. For each event interval, a key graph is obtained using the framework described in section 3.3. We compared the estimated key graphs with the ones obtained from the surrogate time-varying graphs and statistically significant interactions with $p < 0.01$ and $p < 0.001$, are shown in blue and red colors, respectively.

4.3. Key Graph Estimation

For each interval, subspace summarization approach described in section 3.3 is employed to estimate $\overline{x}_{Sig}$, which is used to construct the corresponding symmetric key graph. We compared the estimated key graphs with the ones estimated from 2000 surrogate time-varying graphs generated by randomly reshuffling the edge weights. Fig. 2 shows the statistically significant edges with $p < 0.01$ and $p < 0.001$ in blue and red colors, respectively. Due to the complex activity associated with the error commission, ERN interval has much more significant connections compared to the Pre-ERN and Post-ERN intervals as expected. Moreover, the central electrode (FCz) and the frontal electrode sites (F5, FZ, F2 and F4) have significant connections with $p < 0.001$, which is consistent with previously observed interactions in theta band between medial prefrontal cortex (mPFC) and lateral prefrontal cortex (lPFC) during

error-related cognitive control processes [20]. The remaining intervals do not include such interactions among frontal and central sites.

We also compared the connectivity values, between FCz electrode and the remaining 61 electrodes, within the key graphs for Pre-ERN, ERN and Post-ERN intervals. Welch's t-test at 5% significance level is applied to identify if FCz has stronger connectivity during the ERN interval compared to the Pre-ERN and Post-ERN intervals. For both comparisons, Pre-ERN vs ERN and Post-ERN vs ERN, the null hypothesis is rejected, where FCz has a larger mean connectivity for the ERN interval, indicating that the central electrode has significantly larger connectivity with the rest of the brain. Moreover, we compared the connectivity values for Pre-ERN and Post-ERN and found no significant difference between the connectivity values from these intervals.

5. CONCLUSIONS

In this paper, we proposed a new framework to describe the dynamic evolution of functional brain networks. The proposed approach is based on detecting the network-wide changes, identifying the corresponding event intervals and representing these intervals with key graphs. The key graphs are obtained by projecting the connectivity data in the event interval onto a lower dimensional signal subspace thus discarding the spatial correlation among the edges, such that the key graph would summarize the particular interval with minimal redundancy. Application to real EEG data containing event related potentials shows the effectiveness of the proposed framework in determining the event intervals and describing network activity with a few number of representative networks.

Future work will concentrate on exploring alternative decomposition techniques for subspace analysis and different ways to form the key graphs through weighting, which may improve the performance of dynamic network summarization. We will also extend the proposed approach to compare the dynamic nature of functional networks for error and correct responses to get a more complete understanding of cognitive control. Finally, different group analysis methods will be considered to account for the variability across individual subjects.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] K. Friston, "Functional and effective connectivity: A review," *Brain Connectivity*, vol. 1, no. 1, pp. 13–36, 2011.

[2] C. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, 2005.

[3] E. Pereda, R. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, no. 1-2, pp. 1–37, 2005.

[4] M. Valencia, J. Martinerie, S. Dupont, and M. Chavez, "Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach," *Physical Review E*, vol. 77, no. 5, p. 050905, 2008.

[5] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[6] C. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in

Alzheimer's disease," *Cerebral Cortex*, vol. 17, no. 1, pp. 92–99, 2007.

[7] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, "Time-varying graphs and dynamic networks," *Ad-hoc, Mobile, and Wireless Networks*, vol. 6811, pp. 346–359, 2011.

[8] P. Mucha, T. Richardson, K. Macon, M. Porter, and J. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *Science*, vol. 328, no. 5980, pp. 876–878, 2010.

[9] M. Chavez, M. Valencia, V. Latora, and J. Martinerie, "Complex networks: new trends for the analysis of brain connectivity," *Arxiv preprint arXiv:1002.0697*, 2010.

[10] B. Miller, M. Beard, and N. Bliss, "Matched filtering for subgraph detection in dynamic networks," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2011, pp. 509–512.

[11] K. Xu, M. Kliger, and A. Hero, "A shrinkage approach to tracking dynamic networks," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2011, pp. 517–520.

[12] K. Xu, M. Kliger, and A. Hero III, "Adaptive evolutionary clustering," *arXiv preprint arXiv:1104.1990*, 2011.

[13] T. Ide and H. Kashima, "Eigenspace-based anomaly detection in computer systems," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 440–449.

[14] H. Tong, S. Papadimitriou, P. Yu, and C. Faloutsos, "Proximity tracking on time-evolving bipartite graphs," in *Proc. of SDM*, 2008, pp. 704–715.

[15] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4. ACM, 2004, pp. 219–230.

[16] L. Huang, X. Nguyen, M. Garofalakis, M. Jordan, A. Joseph, and N. Taft, "In-network pca and anomaly detection," *Advances in Neural Information Processing Systems*, vol. 19, p. 617, 2007.

[17] S. Aviyente and A. Mutlu, "A time-frequency-based approach to phase and phase synchrony estimation," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3086–3098, 2011.

[18] D. Kifer, S. Ben-David, and J. Gehrke, "Detecting change in data streams," in *Proceedings of the Thirtieth International Conference on Very Large Data Bases*, vol. 30, 2004, pp. 180–191.

[19] J. R. Hall, E. M. Bernat, and C. J. Patrick, "Externalizing psychopathology and the error-related negativity," *Psychological Science*, vol. 18, no. 4, pp. 326–333, 2007.

[20] J. Cavanagh, M. Cohen, and J. 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.

[21] J. Kayser and C. Tenke, "Principal components analysis of laplacian waveforms as a generic method for identifying ERP generator patterns: I. evaluation with auditory oddball tasks," *Clinical Neurophysiology*, vol. 117, no. 2, pp. 348–368, 2006.