# A NOVEL METHODOLOGY TO QUANTIFY DENSE EEG IN COGNITIVE TASKS

Catia S. Silva, Jose C. Principe

Computational NeuroEngineering Laboratory Department of Electrical and Computer Engineering University of Florida Gainesville, Florida 32611, USA

# ABSTRACT

Cognition emerges from complex interaction amongst widespread brain areas. In this paper, we use a novel methodology for temporal networks quantification for EEG. We model the spatiotemporal structure of dependencies across different electrodes with respect to a single electrode as a local probability density function. This enables immediately the use of information theoretic quantities (information divergences) to quantify brain connectivity in simple two-dimensional graphs. We show that for a visual-motor-driven task, we are able to cluster subjects that performed the task with higher attention-coefficient, in an unsupervised-fashion. We test this methodology with two measures of functional connectivity: correlation coefficient and a measure of association.

*Index Terms*— Cognitive states, EEG, functional connectivity, spatiotemporal quantification, clustering.

### I. INTRODUCTION

Cognitive science is evolving from a focus on quantification of discrete brain areas towards an emphasis on distributed models of brain function, because cognition emerges from the complex interaction amongst widespread brain areas. Communication requires that there is interchange of information, which can be captured in the form of transient temporal dependence between the activities of different brain areas [1]. Functional brain networks have been constructed from functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG). In terms of measures of functional connectivity, the correlation coefficient [2][3] will only capture linear interactions between time series, whereas other measures, such as mutual information [3][4][5][6][7], time-series generalized measure of association (TGMA) [8], phase synchronization [3][9][10], copulas family [11] or Granger causality [12], are sensitive to both linear and nonlinear interactions. Since neural events occur in small windows of time, it is common to divide the neural data into windows, and, for each, compute measures of functional connectivity between pairs of electrodes. This will form a tensor or a concatenation of matrices for different time windows also known as functional temporal networks.

In the last decade, the abstract representation of these functional networks/matrices as a graph has allowed visualization of functional brain networks and description of their non-trial topological properties in a compact and objective way [13][14][15][16][17]. Despite its evident impact, temporal networks analysis are purely

\*This work was funded by Fundação Para a Ciência e Tecnologia (FCT) SFRH/BD/91329/2012.

Andreas Keil

Center for the Study of Emotion and Attention Department of Psychology University of Florida Gainesville, Florida 32611, USA

based on simple node-counting criteria (e.g. node cluster coefficient, measures of centrality, local efficiency, connected components, etc.), which is a very simple statistic of the underlying communication between brain areas (existing or not). Differentiation of cognitive states or neurological diseases are then assessed in a qualitative manner by visualizing node connections of the graphs.

In previous work, we have proposed a novel methodology for analysis of temporal functional connectivity networks [18]. This methodology quantifies the spatial structure of pairwise dependence of each cortical area (electrode with respect to a given electrode) as a probability density function (pdf). Since we are interested in quantifying communication in a spatiotemporal-fashion, we use information theoretic quantities to do so, namely Cauchy-Schwartz Quadratic Mutual Information [19]. We can now quantify how much each functional spatial structure (captured by pdf) changes from one time window to the next. Static spatial structures can also be analyzed using this methodology. For that, we simply fix the time window and quantify dependencies/similarities between pdfs from different cortical areas. Besides quantification of temporal functional networks, we can now perform clustering of the spatiotemporal patterns each subject acquires and, in this way, quantitatively assess their cognitive state or neurological disorder.

The present work illustrates the potential of this novel methodology (described in [18]) by: i) use two different measures of functional connectivity and evaluate the differences based on a visual-motor-driven task; and ii) cluster the quantification patterns, that is, agglomerate people with (potentially) similar cognitive state (attention-coefficient) in the task-solving problem in an unsupervised-fashion. We show that we can unsupervisely cluster people with the same attention state.

The rest of the paper is organized as follows. Section II introduces the neural data and its conditioning. Section III fully describes the novel methodology to quantify dense EEG as well as the additional improvements. Section IV presents results for: the use of different measures of functional connectivity; pdf estimation per cortical area; unsupersived clustering of people by cognitive state; and discussion. Finally, section V offers some concluding remarks.

#### **II. MATERIALS**

EEG data was acquired from a total of 15 participants. The data was continuously recorded from 129 sensors using an Electrical Geodesics<sup>TM</sup> HydroCel Geodesic Sensor Net, digitized at a rate of 250 Hz, using the vertex sensor (Cz) as the recording reference, with the online band-pass filter set at 50 Hz (low-pass). Sensor impedances were kept below 50 k $\Omega$ .

Subjects engaged in a task composed by avoidable (CS+) vs unavoidable (CS-) conditions. The CS+ and CS- gratings were presented centrally and flickered at 15 Hz. The avoidable vs unavoidable context was signaled by additional geometric shapes (flickering at 12 Hz). Conditions: CS+ active trials (appropriate motor response cancels UCS delivery), CS+ passive trials (no motor response is required and UCS delivery is inevitable), CSactive trials (appropriate motor response is required, but UCS is never delivered) and CS- passive trials (no motor response is required and UCS is never delivered). A salient color change of the peripheral cues, from gray to green, that occurred halfway within each trial (3 s from stimulus onset with a 2 s duration) served as the proximal impetus for motor responses. A simple button press during a circumscribed temporal window successfully canceled delivery of the white noise (UCS). Trials were 6 s in length and the inter-trial interval varied randomly between 4 and 5 s. (More details in [20]).

#### **III. METHODOLOGY**

Consider a time window t of the EEG recordings from 129 channels/electrodes. First, we compute the pairwise functional connectivity between each electrode. This results in an 129 × 129 matrix  $\mathcal{K}^t$  for time window t, where  $\mathcal{K}_{ij}^t$  is the functional connectivity between electrode i and electrode j. Now, we consider each column of  $\mathcal{K}^t$  as a random variable (r.v.) and estimate its probability density function (pdf). This results in 129 pdfs, one for each electrode. To compute these pdfs, we do not include the elements  $\mathcal{K}_{ii}^t$  (giving a total of 128 points). For the last step, we quantify the similarity between two pdfs across time using Cauchy-Schwartz Quadratic Mutual Information (CS-QMI) [19]. For example, consider electrode  $e_{75}$  (column 75), for time windows t = 4 and t = 5, we compute CS-QMI( $f(\mathcal{K}_{75}^k), g(\mathcal{K}_{75}^5)$ ), where f and g are pdfs. This result will appear in the row 75 and column 4 of the spatiotemporal quantification matrix  $\mathcal{Q}_m$  of participant m.

Finally, for each condition (CS+ACtive, CS+Passive, CS–Active, CS–PAssive), we cluster the spatiotemporal patterns  $Q_m$  where  $m = \{1, 2, ..., 15\}$ .

#### **III-A.** Correlation Coefficient

One of the simplest ways to measure the statistical dependence between r.v.s is the correlation coefficient, which only captures second order statistical interactions. Let X and Y be two r.v.s, the correlation coefficient is

$$\rho(X,Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

where  $\mu_X$  and  $\sigma_X$  are the sample mean and standard deviation of X, and  $\mu_Y$  and  $\sigma_Y$  are the sample mean and standard deviation of Y. The correlation coefficient returns values between -1 and 1. We are going to take the absolute value because we are only interested in the amount of correlation between the two times series, not the sign.

# III-B. Time Series Generalized Measure of Association

The Generalized Measure of Association (or GMA) [8] is a parameter-free spatial dependency measure of association and can be defined as:

$$GMA = \frac{1}{n-1} \sum_{i=1}^{n-1} (n-r)P(R=r)$$



**Fig. 1:** A. Experimental paradigm consists of four different conditions, which were presented to the subjects in a pseudorandom order – Avoidance conditioning: CS+ trials were always paired with the unconditioned stimulus (UCS). A loud white noise burst was used as UCS. The CS– trials were never paired. The avoidable versus unavoidable context was signaled by geometric shapes. B. Each trial as the duration of 6 s onset and 400 ms offset. The visual cue is alerted by color change at 3 s onset.

where P(R = r) is defined as  $P(R = r) = \#\{i : r_i = r\}/n$ , and represents the empirical probability of the rank variable. GMA assumes values between 0.5 and 1. Since the EEG can be modeled as a stochastic time series, GMA was modified to Time Series GMA or TGMA [8], which includes a pre-optimization step to minimize the time correlation properties of stochastic processes. For TGMA the parameters needed to estimate are the embedding dimension mand the lag value L. For this study the authors considered m = 5and L = 5. These values can be computed directly from data [8].

#### **III-C.** Spatiotemporal Quantification of Functional Activity

To quantify the pdf changes over time, we use Cauchy-Schwartz Quadratic Mutual Information (CS-QMI). This is an estimator for continuous r.v. that can be evaluated in a Reproducing Kernel Hilbert Space (RKHS) using the kernel trick if we use a positive definite function (such the Gaussian) in Parzen estimation [21]. Let  $X_1$  and  $X_2$  be two r.v.s. The CS-QMI between the two r.v.s  $X_1$  and  $X_2$  is defined as the Cauchy-Schwartz divergence ( $D_{CS}$ ) between the joint distribution of  $X_1$  and  $X_2$  and the product of the marginal distribution of  $X_1$  with the marginal distribution of  $X_2$ , that is:

$$I_{CS}(X_1, X_2) = D_{CS}(f_X(x_1, x_2), f_{X_1}(x_1)f_{X_2}(x_2)), \text{ where}$$
$$D_{CS}(f, g) = -\log \frac{\left(\int f(x)g(x)dx\right)^2}{\int f^2(x)dx \int g^2(x)dx}$$
$$= \log \int f^2(x)dx + \log \int g^2(x)dx - 2\log \int f(x)g(x)dx$$
$$= 2\hat{H}_2(fg) - \hat{H}_2(f) - \hat{H}_2(g)$$

where  $\hat{H}_2(f)$  is the Renyi's quadratic entropy estimator of the pdf f. Using a Gaussian kernel (Parzen) estimator [21], for samples  $x_i$ 

drawn from pdf f,  $\hat{H}_2(f)$  can be written as

$$\hat{H}_{2}(f) = -\log \int_{-\infty}^{+\infty} \left(\frac{1}{N} \sum_{i=1}^{N} G_{\sigma}(x - x_{i})\right)^{2} dx$$
$$= -\log \left(\frac{1}{N^{2}} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\sigma\sqrt{2}(x_{j} - x_{i})}\right)$$

Using the same nomenclature, in this step we estimate the pdfs of all  $\mathcal{K}_i$ ,  $i = \{1, 2, ..., 129\}$  as a Gaussian kernel estimator  $f(\mathcal{K}_i^t)$  and quantify the temporal dependencies as  $\mathcal{Q}_{i,t} = I_{CS}(f(\mathcal{K}_i^t), g(\mathcal{K}_i^{t+1})).$ 

### **III-D.** Clustering of Spatiotemporal Patterns

It is known that people process information somewhat similarly but different factors, such as attention, may alter their cognitive states. To quantify how similar different individuals solve a task, we are going to perform hierarchical clustering of the spatiotemporal patterns  $Q^m$  across all  $m = \{1, 2, ..., 15\}$  subjects. First, we compute the L2-norm between matrices to generate a distance matrix  $\mathcal{D}$  of size  $15 \times 15$ .

To the matrix  $\mathcal{D}$  we perform agglomerative hierarchical clustering using Ward's linkage method [22]. The Ward's linkage method minimizes the total within-cluster variance as a squared Euclidean distance. The within-cluster sum of squares is defined as the sum of the squares of the distances between all objects in the cluster and the centroid of the cluster. The sum of squares measure is equivalent to the following distance measure:

$$d(r,s) = \sqrt{\frac{2n_r n_s}{(n_r + n_s)}} \left\| \bar{X}_r - \bar{X}_s \right\|_2$$

where  $\|\bullet\|_2$  is the Euclidean distance or L2-norm,  $\bar{X}_r$  and  $\bar{X}_s$  are the centroids of clusters r and s,  $n_r$  and  $n_s$  are the number of elements in clusters r and s.

# **IV. RESULTS**

# **IV-A.** Measures of Functional Connectivity

Having computed the functional temporal networks  $\{\mathcal{K}_{corr}^t\}_m$  and  $\{\mathcal{K}_{TGMA}^t\}_m$  for all t windows of all m participants using both correlation coefficient and TGMA, respectively, we are now ready estimate the pdfs.

An example of the pdfs of  $\{\mathcal{K}_{corr}^{12}\}_8$  and  $\{\mathcal{K}_{TGMA}^{12}\}_8$  captured using a Gaussian kernel with kernel width  $\sigma = 0.05$  can be seen in Figure 2. There is a large difference between these pdfs, with the TGMA more peaky and short tails (Figure 2 first row), while the pdfs with correlation (Figure 2 second row) tend to have more samples across the range of values, perhaps because of the higher specificity of TGMA [8]. So we see that both measures are substantially different in quantifying the dependence across channels.

Because of the variability of the EEG, we averaged the spatiotemporal patterns  $Q_{corr}^m$  and  $Q_{TGMA}^m$  with  $m = \{1, 2, ..., 15\}$ by conditions (CS+Active, CS+Passive, CS–Active, CS–Passive) to visualize how the average person reacts to the experiment. Figures 3 and 4 show these averages when using correlation and TGMA respectively. The electrodes have been rearranged by brain regions as described in [18], namely frontal (Fr), left temporal (LT), right temporal (RT), parietal (Pa) and occipital (Oc).



Fig. 2: Example of pdfs from subject m = 8, condition CS+Active, time window t = 12, correspondent to [2.6, 2.8] s. First (second) row illustrates these pdfs for TGMA (correlation) in 3 electrodes of each 5 brain areas considered, namely 10, 17, 23 (frontal), 102, 114, 122 (right remporal), 6, 37, 79 (parietal), 66, 75, 84 (occipital), and 44, 51, 128 (right remporal).



Fig. 3: Average by condition of spatiotemporal quantification using correlation measures of functional connectivity.



Fig. 4: Average by condition of spatiotemporal quantification using TGMA measures of functional connectivity.

For correlation, CS-QMI has quantified major changes in the frontal cortex, mainly in the beginning of the task, during the visual cue presentation and at the end of the task. We believe these changes are related to the interpretation of the given condition trial. In addition, for the active trials (where motor response is required), we see a change across brain areas between 3-4 s which corresponds to the time the visual cue is presented and the motor action is performed (Figure 3). However, there is not much difference between the active and passive conditions nor between the conditions CS+ and CS-, which should be different on average [18]. On the other hand, TGMA captured both frontal activity and motor activity (in the RT region) aligned with the visual cue presentation (Figure 4), and there is more variability across the four conditions as expected.



Fig. 5: Agglomorative clustering with Ward's method [22] of an L2-distance matrix of spatiotemporal patterns computed from correlation.

#### **IV-B.** Clustering of Cognitive States

The previous average results assume that all subjects accomplish the task in the same way, which is highly unlikely. In fact, during the collection of this data, it was observed that certain individuals  $m^* = \{1, 5, 6, 7, 12, 13, 14\}$  were motivated to respond faster to the visual cue. However, a subject by subject analysis of the EEG is normally not pursued in cognitive science because of the noise and variability of the responses. Since our methodology is based on pairwise measures and we capture their responses over the head in terms of a single probability function per electrode, we think our methodology can be extended to subject specific quantification. Figure 5 and 6 present results of agglomerative clustering using Ward's method [22] from L2-distance matrices of the spatiotemporal quantification from correlation and TGMA measures, respectively. Correlation-based spatiotemporal quantification returns larger L2-distances amongst participants compared to TGMA, which may mean that the TGMA is less sensitive to the differences in EEG across subjects. We can also see, just by eyeballing the dendograms, that the two measures produce rather different clusterings of subjects. However, each dependence measure tends to preserve the clusters amongst the four experimental conditions. For instance, the dendrograms are similar between Active trials, as well as between Passive trials within each measure. Perhaps the dendrogram that is the most different is the CS-Passive, which is the condition where the subject is least engaged in the experiment.

Cutting the dendrograms requires some knowledge about the application. In our case, we can expect to have at least 2 clusters:

one for the faster responders and the other for the remaining people. For correlation, if we expect to see 2 clusters, the threshold cut should be at around 60 (Figure 5). In CS+Active condition (the most attention-demanding task), the cluster  $m_1^{corr} = \{1, 2, 3, 4, 5, 6, 9, 11, 12\}$  is observed. This cluster contains approximately 57.14% of faster responders. For CS+Active using TGMA (Figure 6 top left) and a threshold of around 15, the cluster  $m_1^{TGMA} = \{2, 3, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15\}$  contains approximately 85.71% faster responders. Note that the clustering is purely unsupervised.



Fig. 6: Agglomorative clustering with Ward's method [22] of an L2-distance matrix of spatiotemporal patterns computed from TGMA.

# **IV-C.** Discussion

This study relates motivated visual brain function to its behavioral consequences. Regions falling within the dorsal pathway and the frontal cortex showed stronger synchronization with the visual cortex during the active motor compared with the passive viewing condition, as showed in [20]. These areas (in the visual cue presentation time) can be extracted from the spatiotemporal matrices as supervised features so that clustering results can better agglomerate the faster responders. We show that TGMA seems a more precise measure of dependence than correlation. However, our results still lack a more direct way to establish the accuracy of the proposed methodology, because attention as a cognitive state is a somewhat subjective feature. Therefore in future work we plan to create a realistic synthetic data set for EEG where we can evaluate with more detail our methodology. The capture or quantification of different states, like neurological events vs healthy events, can also be explored using this methodology.

#### **V. CONCLUSIONS**

Several methodologies have approached neural spatial communication. In this paper, we present a novel methodology to quantify spatiotemporal communication, completely adapted from EEG data. We tested two measures of functional connectivity and unsupervisely clustered the spatiotemporal patterns of a visualmotor-driven task for each individual. We show that for a visualmotor-driven task, we are able to cluster subjects that performed the task with higher attention-coefficient, in an unsupervised-fashion.

### **VI. REFERENCES**

- [1] F. D. V. Fallani, J. Richiardi, M. Chavez, and S. Achard, "Graph analysis of functional brain networks: practical issues in translational neuroscience," *Phil. Trans. R. Soc. B*, vol. 369, no. 1653, 2014.
- [2] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clemenceau, and F. Varela, "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures," *European Journal of Neuroscience*, vol. 12, no. 6, pp. 2124–2134, 2000.
- [3] L. Li, I. M. Park, S. Seth, J. C. Sanchez, and J. C. Príncipe, "Functional connectivity dynamics among cortical neurons: a dependence analysis," *IEEE Transactions on Neural Systems* and Rehabilitation Engineering, vol. 20, no. 1, pp. 18–30, 2012.
- [4] J. Jeong, J. C. Gore, and B. S. Peterson, "Mutual information analysis of the eeg in patients with alzheimer's disease," *Clinical Neurophysiology*, vol. 112, no. 5, pp. 827–835, 2001.
- [5] S. H. Na, S.-H. Jin, S. Y. Kim, and B.-J. Ham, "Eeg in schizophrenic patients: mutual information analysis," *Clinical Neurophysiology*, vol. 113, no. 12, pp. 1954–1960, 2002.
- [6] S. Aydın, M. A. Tunga, and S. Yetkin, "Mutual information analysis of sleep eeg in detecting psycho-physiological insomnia," *Journal of medical systems*, vol. 39, no. 5, pp. 1–10, 2015.
- [7] D. S. Bassett and E. T. Bullmore, "On human brain networks in health and disease," *Current Opinion in Neurology*, vol. 22, no. 4, pp. 340–347, 2009.
- [8] B. H. Fadlallah, A. J. Brockmeier, S. Seth, L. Li, A. Keil, and J. C. Príncipe, "An association framework to analyze dependence structure in time series," in 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2012, pp. 6176–6179.
- [9] S. Dimitriadis, N. Laskaris, and A. Tzelepi, "On the quantization of time-varying phase synchrony patterns into distinct functional connectivity microstates (fcµstates) in a multi-trial visual erp paradigm," *Brain topography*, vol. 26, no. 3, pp. 397–409, 2013.
- [10] J. M. Palva, S. Monto, S. Kulashekhar, and S. Palva, "Neuronal synchrony reveals working memory networks and predicts individual memory capacity," *Proceedings of the National Academy of Sciences*, vol. 107, no. 16, pp. 7580–7585, 2010.
- [11] P. Berkes, F. Wood, and J. W. Pillow, "Characterizing neural dependencies with copula models," in *Advances in neural information processing systems*, 2009, pp. 129–136.
- [12] W. Hesse, E. Möller, M. Arnold, and B. Schack, "The use of time-variant eeg granger causality for inspecting directed interdependencies of neural assemblies," *Journal of neuroscience methods*, vol. 124, no. 1, pp. 27–44, 2003.
- [13] E. T. Bullmore and D. S. Bassett, "Brain graphs: graphical models of the human brain connectome," *Annual review of clinical psychology*, vol. 7, pp. 113–140, 2011.
- [14] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [15] M.-E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna,

M. Kitzbichler, U. Muller, and E. Bullmore, "Functional connectivity and brain networks in schizophrenia," *The Journal of Neuroscience*, vol. 30, no. 28, pp. 9477–9487, 2010.

- [16] D. S. Bassett and M. S. Gazzaniga, "Understanding complexity in the human brain," *Trends in cognitive sciences*, vol. 15, no. 5, pp. 200–209, 2011.
- [17] S. A. Weiss, D. S. Bassett, D. Rubinstein, T. Holroyd, J. Apud, D. Dickinson, and R. Coppola, "Functional brain network characterization and adaptivity during task practice in healthy volunteers and people with schizophrenia1," *Frontiers in human neuroscience*, vol. 5, p. 81, 2011.
- [18] C. S. Silva, M. K. Hazrati, A. Keil, and J. C. Principe, "Quantification of neural functional connectivity during an active avoidance task," in 2016 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2016.
- [19] J. C. Principe, Information theoretic learning: Renyi's entropy and kernel perspectives. Springer Science & Business Media, 2010.
- [20] V. Miskovic and A. Keil, "Escape from harm: linking affective vision and motor responses during active avoidance," *Social cognitive and affective neuroscience*, vol. 9, no. 12, pp. 1993– 2000, 2014.
- [21] E. Parzen, "On estimation of a probability density function and mode," *The annals of mathematical statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.
- [22] J. H. Ward Jr, "Hierarchical grouping to optimize an objective function," *Journal of the American statistical association*, vol. 58, no. 301, pp. 236–244, 1963.