Quantifying Spatio-Temporal Dependencies in Epileptic ECOG

[1]Anant Hegde, [2]Deniz Erdogmus, [3]Deng S. Shiau, [1]Jose C. Principe, [3]Chris J. Sackellares
[1]Computational NeuroEngineering Laboratory, University of Florida, Gainesville, USA
[2]CSEE Department, Oregon Health & Science University, Portland, USA

Abstract—There is evidence that the mechanisms leading to epileptic seizures can be understood by continuously tracking the ongoing spatio-temporal mappings in the brain. We propose to quantify the spatio-temporal dynamics by using a SOM-based Similarity Index (SI) measure. While it is shown that this measure is statistically as accurate as the original SI measure, it is also computationally faster and therefore applicable for real-time analyses. In order to quantify the changes of SI in electrode space and along time, a spectral clustering approach is employed by interpreting the SOM-SI values as affinity matrices. Preliminary analyses on spatial mappings of multivariate epileptic ECOG data are presented using the modified spectral clustering approach. Results involving two pairs of seizures of an epileptic patient suggest that patterns associated with channels' spatio-temporal dynamics during the inter-ictal to pre-post ictal transition vary from seizure to seizure.

Keywords—**epilepsy, SOM-SI, synchronization, spectral clustering, spatio-temporal dynamics**

1. INTRODUCTION

Understanding the interactions among multiple time-series has numerous applications in signal processing and engineering. One of the key aspects of tightly coupled systems with spatial extent is their ability to interact both across space and time, which complicates the analysis greatly. In biological systems this difficulty is compounded by the fact that the systems of interest have nonlinear complicated dynamics that can dictate overall changes in the system behavior. For instance, changes in the spatio-temporal interactions among the cortical columns of the brain are known to be indicative of various pathological diseases such as Alzheimer and epileptic seizures. It is becoming clear that epilepsy is a dynamical disease [1], i.e. the macroscopic spatio-temporal dynamical across different regions of the

brain are consistent with rapid, sometimes gradual and often very subtle nonlinear dynamical interactions. It is believed that synchronization occurs due to both local and global discharges of the neurons. In quantifying this phenomenon, one of the main difficulties is that the brain is a highly complex, non-linear system. Therefore, one way to understand how physiological activities are coordinated in the brain is to understand how subsystems are coupled and how information propagates through them. From the epilepsy perspective, quantifying the changes in spatio-temporal interactions could potentially lead to the development of seizure-warning systems. This quantification would also help us identify the regions that actively participate during epileptic seizures.

Various linear and nonlinear techniques have been developed to quantify the degree of synchronization. Cross-correlation and spectral analysis are the earliest and most relied-on linear techniques in this effort [2, 3]. Efforts to improve on the correlation techniques resulted in development of asymmetric measures such as Directed Transfer Functions (DTF) and Partial Directed Coherence (PDC) [4-8]. DTF is a Multivariate Autoregressive (MVAR) model, built on the concept of Granger-causality and is widely used to analyze the flow of patterns in seizure applications [4-7]. Evidently, the MVAR models are better than most bivariate measures due to the fact that they use information simultaneously from all the channels, and thus are able to unambiguously distinguish between direct and indirect causal connectivity between nodes. The MVAR approaches have been used to determine the propagation of epileptic ECOG activity, in temporal lobe and mesial seizures [5-7]. However, these models strictly require that the measurements be made from all the nodes, else the directional relationships could be ambiguous. In addition, there is no clear evidence of causality relationships among the cortical structures and as suggested in [9], the nature of synchronization is mostly instantaneous or without any detectable delay. The applicability of MVAR measures is constrained to the linear dependencies even though there are strong evidences pointing to non-linear patterns of interactions [10].

Instantaneous phase measures using Hilbert transforms have also been used to quantify phase synchronization [11]. Hilbert transforms, however, are accurate only when the signals have a narrow-band spectrum, which is unrealistic for most real-world signals.

Empirical Mode Decomposition (EMD) is a preprocessing technique [12] in which the data is decomposed into narrow band components called the intrinsic mode functions (IMF). Contrary to wavelet techniques or conventional digital FIR filters, the EMD derives IMFs or the basis functions directly and adaptively from the data. This characteristic makes EMD particularly useful to detect local changes in the signal. However, EMD has no parameters to filter the signal into specific bands of interest. In phase synchrony analysis it is very important that the narrow band components of the time-series are in the same frequency band, otherwise inter channel relationships (such as phase) have no meaning. Due to the fact that EMD adaptively derives its basis functions based on the data, the IMFs of two different signals can easily lie in different bands. Hence, EMD in its present formulation is not a useful tool to compare phases between multiple channels with different time structures.

The applicability of phase-synchrony measures is restricted to bivariate time-series and also they are restricted to identifying only the phase locking between two signals. However, synchronization between two systems is a general phenomenon and can broadly be defined as the functional relationship governing the oscillation of the two systems. The function could be linear invertible, linear non-invertible, non-linear invertible and non-linear non-invertible. Phase synchronization, being only a subset of the general synchronization definition, fails to address the overall functional relationship between systems.

Nonlinear dependencies between multiple signals have also been studied in the last two decades, but with limited success. Generalized mutual information function (GMIF) [13] is one of the widely used techniques for quantifying non-linear correlations. One of the problems with this approach is that, it requires large data sets for probability density estimation and is, therefore, computationally not viable. Selecting an appropriate neighborhood size is an additional constraint.

Eckmann *et al.* [14] proposed the method of recurrence plots (RPs) that represents the recurrence of states in the phase-space trajectory of a chaotic signal. Since chaotic systems are non-linear in nature, this method has been fairly successful in detecting bifurcations and non-stationarities in time sequences. Recently, Arnhold *et al.* [15, 16] introduced a bivariate measure, called the similarity–index technique (SI), to measure non-linear asymmetric dependencies between time-sequences. Conceptually, this method relies on the assumption that if there is a functional dependency between two signals, the neighboring points in time will also be neighboring points in state space. In other words, then the recurrence of dynamics of one signal will mean the recurrence of dynamics in the other signal. Assuming that the connection is uni-directional, several states of the driver signal map onto a single state of the response signal. The SI does not imply any causal relationship; however, it provides a degree of similarity between the instantaneous dynamics of any two functionally coupled systems. The SI determines as to which signal is more active than the other, which in turn is closely related to the dimensional complexity of the systems. Arnhold *et al.* [15, 16] propose to search for recurrence in the signals' state-space, which poses an enormous computational burden, especially for large data sets. In this study, we propose a self-organizing map (SOM) based modification to this method [17] to reduce computational complexity, while maintaining accuracy. This is achieved by mapping the embedded data from signals onto a quantized output space through a SOM specialized on these signals, and utilizing the activation of SOM processing elements (PEs) to infer about the influence directions between the signals, in a manner similar to the original SI technique. The approach not only facilitates real-time computation but it also enables us to derive long range synchronization patterns using which we can study i) the temporal evolution of dependencies at different stages of seizure, ii) spatial-temporal distribution of channel interactions and iii) spatial-temporal clustering of channels at different stages of seizure.

2. SIMILARITY INDEX

In this section, a brief description of the original SI measure [15, 16] is provided. Assume that X and Y are two time series generated by a system, which are embedded into two vector signals in time using delays.

$N(X|Y)$ is defined as the average dependency of X on Y and it can be written as,

$$\boldsymbol{N}(\boldsymbol{X} \mid \boldsymbol{Y}) = \frac{1}{\boldsymbol{N}} \sum_{\boldsymbol{n}=0}^{\boldsymbol{N}-1} \frac{\boldsymbol{R^n}(\boldsymbol{X}) - \boldsymbol{R^n}(\boldsymbol{X} \mid \boldsymbol{Y})}{\boldsymbol{R^n}(\boldsymbol{X})} \tag{1}$$

where $R^n(X)$ is the average Euclidean distance between the state-vector of X_n and the remaining state-vectors in X. Y-conditioned Euclidean distance $R^n(X|Y)$ measures the average Euclidean distance between X_n and the vectors in X whose corresponding time-partners are the k-nearest neighbors of Y_n. This measure takes values in [0, 1], where 0 implies no coupling and 1 implies perfect synchronization. In other words, 1 suggests that recurrence of a state in Y implies a recurrence in X [15, 16]. On the same principles, $\boldsymbol{N}(\boldsymbol{X} \mid \boldsymbol{Y}) = 0$ implies complete independence between X and Y. By design, SI can quantify nonlinear dependencies. Similarly, it is possible to quantify the average dependence of Y on X by,

$$\boldsymbol{N}(\boldsymbol{Y} \mid \boldsymbol{X}) = \frac{1}{\boldsymbol{N}} \sum_{\boldsymbol{n}=0}^{\boldsymbol{N}-1} \frac{\boldsymbol{R^n}(\boldsymbol{Y}) - \boldsymbol{R^n}(\boldsymbol{Y} \mid \boldsymbol{X})}{\boldsymbol{R^n}(\boldsymbol{Y})} \tag{2}$$

Comparing $\boldsymbol{N}(\boldsymbol{X} \mid \boldsymbol{Y})$ and $\boldsymbol{N}(\boldsymbol{Y} \mid \boldsymbol{X})$, we can determine which signal is more dependent on the other. The difficulty with this approach is that at every time instant, we must search for the k nearest neighbors of the current embedded signal vectors among all N sample vectors; this process requires $O(N^2)$ operations. In addition, the measure depends heavily on the free parameters, namely, the number of nearest neighbors and the neighborhood size ε. The neighborhood size ε needs to be adjusted every time the dynamic range of the windowed data changes.

3. SELF ORGANIZING MAP (SOM)-SI

The SOM-based SI algorithm is designed to reduce the computational complexity of the SI technique. The central idea is to create a statistically quantized representation of the dynamical system using a SOM [18, 19]. For best generalization, the map needs to be trained to represent all possible states of the system (or at least with as much variation as possible). As an example, if we were to measure the dependencies between EEG signals recorded from different regions of the brain, it is necessary to create a SOM that

represents the dynamics of signals collected from all channels. The SOM can then be used as a prototype to represent any signal recorded from any spatial location on the brain, assuming that the neurons of the SOM have specialized in the dynamics from different regions.

One of the salient features of the SOM is topology preservation [18, 19]; i.e., the neighboring processing elements (PEs) in the feature space correspond to neighboring states in the input data. In the application of SOM modeling to the similarity index concept, the topology preserving quality of the SOM enables us to identify neighboring states of the signals by neighboring PEs in the SOM.

Assume that X and Y are two time series generated by a system, which are embedded into two vector signals in time using delays. Define the activation region of a PE in the SOM as the set of all input vectors (the embedded signal vectors) for which the PE is the winner based on some distance metric (Euclidean in most cases). Let X_n be the set of time indices of input vectors x_j that are in the activation region of the winner PE corresponding to the input vector x_n at time n. Similarly define the set Y_n.

The procedure to estimate the directed SOM-SI between X and Y is described in Table 1.

The normalized similarity index χ, can point out directed influences between the two signals. Specifically, positive values of χ indicate larger influence of X on Y, while negative values indicate the opposite.

Higher magnitudes of χ indicates a stronger coupling of the signals in the direction indicated by the sign. When χ is close to zero, an ambiguity occurs, since the two signals could be independent or coupled to each other equally in both directions. This ambiguity can be resolved by observing the individual values of $N(X \mid Y)$ and $N(Y \mid X)$. The computational savings of the SOM approach is an immediate consequence of the quantization of the input (signal) vector space. The search for nearest neighbors will involve $O(Nm)$ operations as opposed to the $O(N^2)$ of the original algorithm, where N is the number of samples and m is the number of neurons in the SOM ($m<<N$ by design).

4. SIMULATION

In this section, we demonstrate the viability of the SOM-based similarity index approach in determining couplings and influencing directions through synthetic simulations on signals generated from two non-identical chaotic oscillators.

Rosseler-Lorenz signals: The same Rosseler-Lorenz example used by Quiroga *et al.* [16] is used here. A synthetic nonlinear dependency between a Rosseler (X) and a Lorenz (Y) system is created by having the second state of the Rosseler system drive the Lorenz system in the following manner:

$$\begin{array}{ll} \text{Roessler} & \text{Lorenz} \\ \dot{x}_1 = -6\{x_2 + x_3\} & \dot{y}_1 = 10(-y_1 + y_2) \\ \dot{x}_2 = 6\{x_1 + 0.2x_2\} & \dot{y}_2 = 28y_1 - y_2 - y_1 y_3 + Cx_2^2 \\ \dot{x}_3 = 6\{0.2 + x_3(x_1 - 5.7)\} & \dot{y}_3 = y_1 y_2 - \frac{8}{3} y_3 \end{array} \tag{3}$$

where C is the coupling strength. To be realistic, measurement noise, up to 30dB SNR are added to both the systems. Two SOMs, one corresponding to x_2 component of the Rosseler system and the other, corresponding to the y_2 component of the Lorenz system (a new SOM for each value of C), were trained separately on embedded data, using an embedding delay of 0.3 time-units and embedding dimension of 4. Note that separate SOMs were trained because of the fact that the dynamics of the two systems were non-identical. However, in general, while working with signals generated from the same system, it suffices to train a SOM that represents all possible dynamics of that system. SOM neurons overlapped with the phase-space dynamics of the Rosseler system and the Lorenz system (for different C values) are shown in Fig. 1. Each SOM is an 8x8 rectangular grid, and is trained on a set of 4000 samples using a Gaussian neighborhood function for about 400 iterations. The neighborhood radius (standard deviation of the Gaussian neighborhood function) is exponentially annealed starting from an initial value of 4 with a time constant of 100. The step size is also annealed exponentially from 0.08 using the same time constant.

Using the SOM-SI approach, the normalized indices χ are calculated for coupling strengths of C = 0, 0.5, 1, 2, 3 and 5. The robustness to varying SNR is demonstrated in the test data, as shown in Fig 2. Firstly, we observe that the asymmetric difference in the interdependencies between the Roessler and the Lorenz system increases as the coupling (C) is increased (irrespective of the noise level). Despite the fact that the

training was done with 30dB SNR level, the absolute values of the dependencies do not change much for SNR up to around 20dB, indicating noise-robustness. However, a drastic decrease in the absolute dependency values for 10dB and 0dB SNR situations suggest that the structural relationship between the two systems is being largely overwhelmed by the stochastic noise. In summary, we see that the idea of SOM modeling of system's dynamics not only reduces computational complexity but also preserves neighborhood mappings during noisy perturbations, thus providing enhanced noise robustness.

5. ECOG DATA COLLECTION

ECOG signals were recorded from hippocampus, sub-temporal and frontal-cortex structures of epileptic patients having a history of complex, complex-partial, partial-secondary seizures of temporal-lobe focus, using bilaterally and surgically implanted electrodes (Fig 3). Using amplifiers with an input range of $\pm 0.6mv$, the recorded signals were converted to narrow-band using an anti-aliasing filter with a cut-off range between 0.1Hz and 70Hz. Using an analog-to-digital converter with 10-bit quantization precision, the narrow-band signals were sampled/digitized at 200 samples/sec. Measurements involved recording ECOGs from multiple sensors (28 to 32, with common reference channels) and the recordings spanned over 6 continuous days. A total of 55 seizures, of temporal lobe onset, were recorded from 5 patients, in the range of 6 to 18 seizures for each patient.

6. ECOG - SOM

We demonstrate the utility of SOM-SI in epileptic ECOG analysis and compare results statistically with the original SI. An 25x25 ECOG-SOM is trained using 3000 input vectors constructed by embedding (dimension=10, delay=30ms) ECOG signals collected from various regions such as temporal, sub-temporal, and orbit frontal, of an epilepsy patient. The embedding dimension was determined such that it balances to accommodate the high frequency waves as well as adequately represent the spikes. The delay was such that the components in the vector were independent of each other. The ECOG-SOM needs to represent all possible ECOG-dynamics, so the training data must include samples from the inter-ictal, ictal, and the pre-ictal states of the patient. Fig. 4 shows the phase-space trajectory of the data and the PEs

of the ECOG-SOM in two-dimensions. The normal ECOG state is represented by the smaller amplitude activity (the dominant portion of the training data), whereas the larger amplitudes correspond to the spiky, sharp and slow wave activity, mostly formed during the ictal state. We note that the distribution of the neurons is sparse in the higher amplitude region because of the density matching property of the SOM.

To ensure generalization of the SOM, a test set of ECOG signals were quantized by the trained SOM. As seen from Fig.5, the quantization results successfully approximate the dynamics of the test data set (projected in one-dimension). The correlation coefficient between the two signals was found to be 90.1%. For the most part, the correlation coefficient was between 80% and 95%. Note that the amplitude errors are higher in the larger amplitude regions corresponding to spike and slow waves. This is expected because of the sparse distribution of the neurons in the higher amplitude regions. These errors can be compensated by using a larger SOM grid (>25x25), but since the dynamics of the data are more important for the neighborhood information in the SI measure and computational complexity will be an issue, we chose not to increase the SOM grid size.

6.1. STATISTICAL COMPARISON BETWEEN SI and SOM-SI

Next, we quantify the accuracy of the SOM-SI measure relative to the original SI measure by comparing their results. SI values were calculated on nearly 39 minutes of data corresponding to a pair of signals obtained from the right temporal (RTD4) and the right sub temporal depth (RST1) electrodes, corresponding to patient P093. The entire interval of 39 minutes data was segmented into 230 non-overlapping windows of 10 seconds each. Fig. 6 shows the interdependency values of both measures. It is easy to see that the results from both the measures are to a large extent in agreement. There are also subtle differences, which need to be quantified using statistical tests.

The comparison will be two-fold: (i) identify if the number of windows in which the predicted directions of influence differ is significant or not, (ii) given time instances where both measures agree on the direction, check if significant differences exist in predicted strengths of influence. Since we found that the SOM-SI and SI values had quasi-normal distribution, we use the two-sided paired *t*-test to investigate the

extent of disagreement between the two methods. The test was performed at a significance level of α=0.05, over a size of 138 randomly selected samples out of the 230 available samples.

Null hypothesis: H_0: $\mu_d = \mu_{(\chi_{som-SI} - \chi_{SI})} = 0$

Alternate hypothesis: H_1: $\mu_d = \mu_{(\chi_{som-SI} - \chi_{SI})} \neq 0$

Paired *t*-test is chosen, because the observation in the original SI window 1 is obtained under similar conditions as the window 1 of SOM-based SI, and hence, the data may be said to occur in pairs. In this case, t_{exp} was found as -0.9441, whereas $t_{crit}=t_{(0.05),137}=1.960$. Since $t_{exp}<t_{crit}$, we do not have enough evidence to reject the null hypothesis, H_0. This was also the case in most other comparisons made using different electrode pairs from different patients. Mostly, the number of windows in which the signs of the normalized similarity index between original SI and the SOM-SI differed from each other varied in the range of 30 to 40, out of a total of 230 windows (roughly 15% to 20%). However, a detailed look of the χ revealed that the differences were mainly in windows where coupling was weak and signals were rather independent. We know that at very low activity levels between signals, quantifying the interdependence using either SI or SOM-SI could sometimes lead to ambiguous directional information. Nevertheless, there are no such statistical ambiguities in large coupling situations, either in terms of magnitude or direction. Therefore, we conclude that statistically the SOM-SI measure, computed with a 25x25 grid SOM, performs as well as the original SI measure.

7. APPLICATION OF SOM-SI TO EPILEPTIC ECOG DATA

To analyze spatio-temporal synchronization in an epileptic brain, twenty four (24) representative channels representing different regions, namely orbito-frontal, temporal and sub-temporal, from both the brain hemispheres were selected. For the preliminary analysis of inter-dependencies, we selected 4 seizures (seizure 2, 3, 9 and 10) from patient P093. Pair-wise interdependence was computed among all the C_2^{24} possible combinations of 24 channels. Seizures 2&3 were of Partial Secondary Generalized (PSG) type and seizures 9 and 10 were of Complex Partial (CP) type, with temporal lobe being the epileptogenic-foci.

Normalized similarity index was computed using $N(X|Y)$ and $N(Y|X)$, as suggested in (1,2). A number of variations of the original SI measure can be found in [15, 16], but essentially they provide the same information. We adopted (1, 2) to our improved SOM approach by computing the Euclidean distance between all the states or the processing elements (PE's) of the SOM for $R^n(X)$ and $R^n(Y)$. This will certainly not slow down the SOM-SI computation because of the fact that the distance $R^n(X)$ or $R^n(Y)$ can be computed offline after the SOM is trained and therefore can be retrieved from a lookup table.

From Fig 6, coupled with SOM-SI analyses on similar data, we make the observation that the interdependence values, though unequal, are not very different from each other. Therefore, we chose to analyze only the maximum mean driving ability of a channel. On the smoothened interdependence values, for each window and for each channel we find the maximum driving influence that the channel exerts on any other channel. Over windows (time) these maximum driving indices give the maximum driving ability of the particular channel of interest. The maximum driving abilities are evaluated for every channel under consideration and for simplicity, Fig 7 shows an example consisting of the driving abilities for just 6 channels [20]. In the inter-ictal stage, low driving ability of all the channels indicate that the channels are de-synchronized, even though they exhibit an upward trend. Synchronization goes up momentarily a few minutes pre-seizure and at the onset of the seizure, there is a sudden drop followed by a sharp increase post-seizure [20]. Higher degree of post-seizure global synchronization is followed by a gradual drop, leading to the inter-ictal state. This trend in synchronization patterns was observed in all six (6) seizures from both patients.

It is clear from fig.7 that the temporal-evolution of the interdependency values exhibit distinguishable patterns at different stages of seizure. However, we still miss information on the spatial changes during the inter-ictal activity that might possibly lead to evolution of seizures. One of the hypotheses of this study is that the neuronal activities undergo both temporal and spatial distribution changes during ictal, pre-ictal and post-ictal stages. It is absolutely essential that the spatial changes be characterized by extracting information on the undergoing channel interactions. Since the neurons are all densely

connected, it is reasonable to expect that any activity in the brain (such as a spike) can be a consequence of a number of co-ordinations between different structures of the brain. The co-ordination structure between certain regions of the electrode sites can plausibly undergo temporal changes causing neuronal activity bursts. One can imagine that these structural changes can subsequently lead to certain clinical events that manifest as spikes or slow-waves in the ECOG recordings. Characterizing those co-ordinations and understanding their dynamics can be important for many applications such as prediction and detection of seizures.

8. SPECTRAL CLUSTERING OF CHANNELS

We propose to apply a spatial-temporal clustering technique to extract information on spatio-temporal distribution of the channels, by operating on segments of data corresponding to inter-ictal, pre-ictal and post-ictal seizure stages. A 3-fold approach consisting of spatial-discretization of the data using spectral-clustering technique [21], temporal clustering using a Markov-model approach followed by K-means clustering is presented in Fig 8. The rational will become apparent during the explanation.

Spectral clustering is one of the many clustering methods that use subspace decomposition on higher dimensional features derived from the data to achieve data-clustering. Using kernel methods, the data samples are projected onto a higher dimensional space where the discriminant analysis is much easier. Projecting the data onto a feature space results in tightly formed clusters such that the between cluster entropy is maximized and within cluster entropy is minimized. Spectral Clustering is inspired by the normalized cut theory in computer science where the distance between the nodes/data samples is interpreted as an affinity matrix on which subspace decomposition yield membership labels of the nodes/data samples. In our study, we use the standard spectral clustering method by Ng et al. [21] to spatially cluster the similarity-indices obtained by the SOM-SI technique.

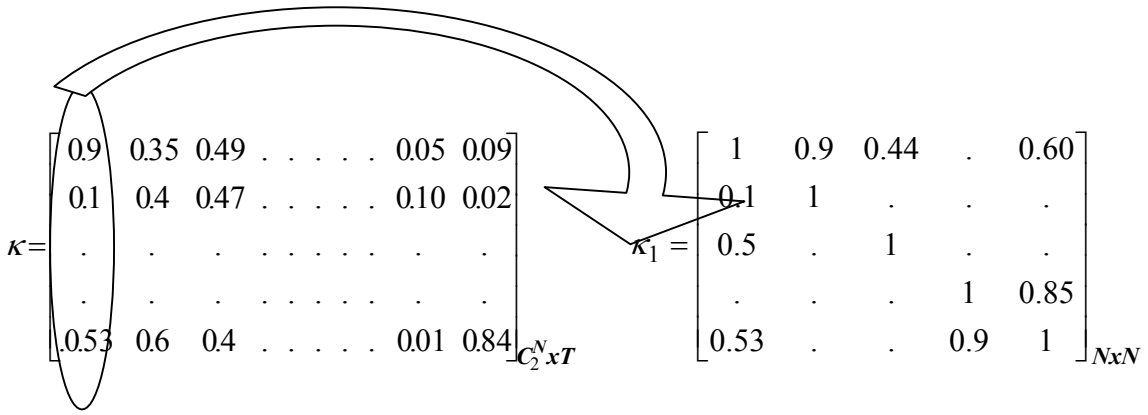
The output from computing pair-wise SOM-SI on all the possible combinations of the multivariate ECOG sources leads to a $k = 2*(C^{N}_{2})$ spatial-samples in the range of [0, 1] at every time-instance. For example:

In the (*k x t*) matrix shown in (5), '*k*' indicates the # of spatial entries at every time instance and the '*t*' columns indicate the time over which the SOM-SI is evaluated.

$$\kappa = \begin{bmatrix} 0.9 & 0.35 & 0.49 & . & . & . & . & 0.05 & 0.09 \\ 0.1 & 0.4 & 0.47 & . & . & . & . & 0.10 & 0.02 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ .0.53 & 0.6 & 0.4 & . & . & . & . & 0.01 & 0.84 \end{bmatrix} \tag{5}$$

As discussed earlier, each of these entries characterizes asymmetric inter-dependencies between the ECOG channels. In this study, κ represents the spatio-temporal correlation indices obtained by computing pair-wise similarity-index between a set of 24 channels.

One of the primary motivations of using the spectral-clustering technique is that, the κ matrix can be interpreted as a 3 dimensional spatio-temporal affinity matrix representing the pair-wise similarity between 24 nodes, indexed both in space and time. In other words, each column of the κ matrix represents pair-wise affinity between the channels/nodes at any instance in time and the columns represent the affinities in channels as a function of time. Therefore, we can represent each column of size $2C_2^N$ x 1 of (5) as a square matrix (6) of size *N* x *N*, where *N* is the # of channels. Since the weighting is normalized between 0 and 1, the diagonal elements representing the affinity of a channel with itself, are coded as 1.



(6)

As we can observe from (6), the affinity between nodes will not have a symmetric form, as required in spectral clustering. In order to perform eigendecomposition on the affinity matrix, it is essential that the affinity matrix be square and symmetric in nature. This is because the eigendecomposition yields orthogonal column vectors (also called eigenvectors) only if the projection matrix is square-symmetric. One of the ways we can transform the asymmetric matrix into symmetric is by adding it to its transpose and dividing each entry by 2. Mathematically, the transformation can be represented as,

$$\delta = \frac{(\chi + \chi^{T})}{2} \tag{7}$$

The transformed affinity matrix now represents the average information exchanged between all pairs of channels. This implies that we do not have the luxury of using the asymmetric nature of the dependencies to create a membership grouping among channels. The averaging of the dependency information will not affect channel dependencies much because of the earlier observation that there is no major difference in the driving and receiving information of the channel.

On the transformed affinity matrix δ, eigendecomposition can be done, followed by *K*-means clustering. Consequently, we have a set of labeled clusters (empirically set to 3) representing the membership of the channels. Repeating this procedure on every column in the κ matrix (6) will yield a discrete-valued matrix κ_{spect}.

$$\kappa_{spect} = \begin{bmatrix} 3 & 2 & 2 & . & . & . & . & . & 3 & 1 \\ 1 & 2 & 2 & . & . & . & . & . & 3 & 2 \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 3 & 1 & 2 & . & . & . & . & . & 1 & 2 \end{bmatrix} \tag{8}$$

Finally, to obtain an average discretization/clustering of the channels, we can use the Markov-model characterization followed by *K*-means approach, as described in the next section.

9. MARKOV-CHAIN MODELING

Spectral-clustering is a simple and a powerful technique to achieve spatial-quantization of the SOM-SI continuous data. However, the task is to be able to detect the channels that exhibit a similar behavior over a specified time-interval. In other words, we would like to find out rows of the κ_{spect} matrix that are similar with each other over the time interval *T*. We propose a Markov-chain approach as a metric to quantify the rows in the κ_{spect} matrix. This would be followed by a conventional *K*-means clustering to achieve the final grouping of channel interactions.

As we can observe from the κ_{spect} matrix, the entries across the columns indicate the temporal-transition of a channel's interaction between clusters. When there are only 3 states or clusters, one of the 9 state-transitions are possible; namely; (1-1), (1-2), (1-3), (2-1), (2-2), (2-3), (3-1), (3-2), and (3-3). The transitions can be characterized by associating a probabilistic structure that takes into consideration the amount of time a channel would be in a particular state (marginal probability) and the amount of time it makes a transition from one state to another state (transition probability). Mathematically, on each row of the κ_{spect} matrix, we describe the following:

$$\pi^{r} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}, \quad \mathbf{A}^{\mathbf{r}} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

where '*r*' is a particular row in the matrix κ_{spect} and P_{ij} corresponds to probability of transition from cluster '*i*' to cluster '*j*' and P_k corresponds to marginal probability of cluster '*k*' .

Information on marginal and transitional probability of each row will result in single-memory Markov-chain characterization of a channel's behavior. Each channel has a distinct set of states and a distinct transition between states. Finding similarity between channel interactions is now reduced to finding similarity between their Markov models. One of the ways we propose to achieve this is by finding a metric that uses the marginal and the transitional probability information and gives out a unique quantity (scalar or vector) that can be used to distinguish between two Markov characterizations. We observe that

the product of the transitional probability matrix with the corresponding marginal probability vector will result in a vector unique to each row. Mathematically,

$$\chi^{r} = \mathbf{A}^{\mathbf{r}} . \boldsymbol{P}^{r} \tag{9}$$

In this particular case, the product will result in a 3-dimensional vector for each row of the κ_{spect} matrix. *K*-means clustering, on a pre-defined number of clusters, in the 3-dimensional space will eventually result in the rows corresponding to channel-interactions to be clustered. The clustering will enable us to know the groups of channels that have similar behavioral structure in the brain.

10. SIMULATION WITH SYNTHETIC DATA

This section presents a simple simulation to demonstrate the validity of the spectral-clustering technique on spatio-temporal measurements. Ideally we could generate non-linear signals and identify interactions by using the SOM-SI measure, but the process involves training a SOM uniquely for each signal. Any kind of interaction between the signals might manifest changes in the attractor's dynamics and structure, causing the need to re-create SOMs. If there are a large number of measurements and if systems are non-identical, the whole process will soon become tedious. However, note that the demonstration is primarily aimed at testing the spectral-method's capability to cluster spatio-temporal data based on their mutual interactions. Therefore, without loss of generality, we simplify the simulation by generating a set of linear signals and applying the spectral-clustering technique to cluster the data based on their linear interactions. Consider a set of colored noise data obtained as the band-pass filtered output of an additive white-Gaussian stochastic process of zero mean and unit-variance (as shown in Fig. 10). Each of the outputs can be denoted by $s_i(t)$, $i = 1,2,.....,8$ & $t = 1,2,...,T$. We construe $x_i(t)$ as the observation measurements, sampled from eight (8) different channels of a spatio-temporal system, over an interval *T*. Observe that $x_i(t)$ is constructed by adding Gaussian white noise (of variance between 0.1 and 3) to $s_i(t)$.

$$\boldsymbol{x_i(t) = s_i(t) + n_i(t)} \tag{10}$$

Design of noise $n_i(t)$ is such that , a variance higher than 3 will destroy the structure of $x_i(t)$ to the extent that the signals become completely un-correlated with each other. Even though the two band-pass filters denoted by BPF_1 and BPF_2 respectively (as shown in fig. 11 and fig. 12) share different pass-bands, there is a significant amount of overlap as seen by the narrow separation of their pass-bands. It should also be noted from fig. 12 that some of the output measurements $x_i(t)$ can share the same band-pass filters. However the additive noise component $n_i(t)$ can introduce some stochastic differences in their linear structures. Measurements $x_i(t)$ that share the same band-pass filter can loosely be considered analogous to time-structures having reasonably strong linear-interactions. This is because information exchanged between two time-series can sometimes cause their phase dynamics to be similar.

We compute pair-wise cross-correlation indices (1) among all the channels to quantify the interactions or rather, the exact amount of information exchanged among the spatio-temporal measurements $x_i(t)$. Due to the fact that the signals are stochastic, linear and do not possess any complex or chaotic phase dynamics, we felt that the cross-correlation computation would suffice to indicate the amount of their pair-wise interactions. The cross-correlation indices as denoted by

$$\boldsymbol{R}^1 = \begin{bmatrix} \boldsymbol{R}^1_{1,1} & \boldsymbol{R}^1_{1,2} & . & . & \boldsymbol{R}^1_{1,8} \\ \boldsymbol{R}^1_{2,1} & \boldsymbol{R}^1_{2,2} & . & . & \boldsymbol{R}^1_{2,8} \\ . & . & . & . & . \\ . & . & . & . & . \\ \boldsymbol{R}^1_{8,1} & \boldsymbol{R}^1_{8,2} & . & . & \boldsymbol{R}^1_{8,8} \end{bmatrix}$$

denote the affinity matrix for the time series corresponding to say window 1. Repeating this process over a number of windows $\boldsymbol{W}$ will result in the spatio-temporal affinity matrix $\boldsymbol{R} = \left\{ \boldsymbol{R}^1 \quad \boldsymbol{R}^2 \quad . \quad . \quad \boldsymbol{R}^{\boldsymbol{W}} \right\}$, where $\boldsymbol{W}$ is the total number of windows. For each window, the spatial configuration/arrangement of the band pass filters is changed as follows:

Configuration 1 (c_1): Channels 1, 2, 3 & 4 are passed through BPF_1 and

Channels 5, 6, 7 & 8 are passed through BPF_2.

Configuration 2 (c_2): Channels 1 & 3 are passed through BPF_1 and

Channels 2, 4, 5, 6, 7 & 8 are passed through BPF_2.

Configuration 3 (c_3): Channels 1, 2, 3, 4, 6 & 7 are passed through BPF_1 and

Channels 5 & 8 are passed through BPF_2.

Configuration 4 (c_4): Channels 1, 3, 6 & 7 are passed through BPF_1 and

Channels 2, 4, 5 & 8 are passed through BPF_2.

Total number of Windows $\boldsymbol{W}$ is fixed to 100 and the probability of random occurrence of channel configurations is as follows: $\{p(c_1) = 0.5,\ p(c_2) = 0.1, p(c_3) = 0.1, p(c_4) = 0.3\}$.

Applying the spectral clustering algorithm on the spatial-correlation matrix of each window $\boldsymbol{R^i}$ resulted in

$$\lambda_{spect} = \begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 1 & 2 & . & . & . & . & 1 \\ 1 & 2 & 2 & 1 & 2 & 2 & 2 & . & . & . & . & 1 \\ 1 & 2 & 1 & 2 & 1 & 1 & 2 & . & . & . & . & 1 \\ 1 & 2 & 2 & 1 & 2 & 2 & 2 & . & . & . & . & 1 \\ 2 & 1 & 2 & 1 & 2 & 2 & 1 & . & . & . & . & 2 \\ 2 & 2 & 2 & 2 & 1 & 2 & 1 & . & . & . & . & 1 \\ 2 & 2 & 2 & 2 & 1 & 2 & 1 & . & . & . & . & 1 \\ 2 & 1 & 2 & 1 & 2 & 2 & 1 & . & . & . & . & 2 \end{bmatrix}_{(8\boldsymbol{x}100)}$$

$$c_1\ c_3\ c_2\ c_4\ c_4\ c_2\ c_1\ .\ .\ .\ .\ c_3$$

This illustration shows how the cluster labels are assigned to the channels depending on their configuration, denoted by $\mathbf{c_i}$'s. Heuristically, it is evident that the channels can either be clustered in groups of four (4) or in groups of two (2). For the case when the number of clusters equals four (4), the markov chain characterization followed by K-means clustering resulted in cluster assignment of the channels

C_1: {1, 3}, **C_2:** {2, 4}, **C_3:** {6, 7}, **C_4:** {5, 8}.

Specifying only 2 clusters resulted in

C_1: {1, 2, 3, 4}, **C_2:** {5, 6, 7, 8}.

We can easily see that the results with 4 clusters and 2 clusters are in perfect agreement with the construction of the spatio-temporal configurations. Channel 1and 3 are always together and belong to the same cluster, regardless of the configuration. Likewise, channels 5 and 8 always belong to the same cluster; although different from the cluster to which 1 and 3 belong to. This is because the source data $\boldsymbol{Z(t)}$ is always passed through BPF_2 to obtain channels 5 and 8, unlike channels 1 and 3 that are obtained by filtering $\boldsymbol{Z(t)}$ by BPF_1. The channels 2 and 4 and similarly 6 and 7 change their memberships often enough and therefore can be partitioned into separate clusters.

For comparison purposes, clustering was performed by applying the K-means algorithm directly on the time-series data. K-means treats each time-series as a multi-dimensional data vector for determining the partitions. Cluster configurations were found to be consistent with the model construction only for the 2-cluster case.

Determining the optimal number of clusters is still an open-ended problem and it is beyond the scope of this paper. For our experiments with the ECOG data, we fixed the number of clusters on an empirical basis by looking at the scatter plots of the three-dimensional vectored data, and obtained at the output of the Markov-chain modeling.

11. CLUSTERING RESULTS ON ECOG DATA

The spatio-temporal techniques described in the previous sections were applied to the SOM-SI data, computed over 2 pairs of seizures, in patient P093.

11.1. Characterizing channel groupings

The spatial-temporal clustering of the channel-dependencies were analyzed in the inter-ictal region between two seizures (seizure 2 and 3 of patient P093, in this case) and also for the pre-ictal and post-ictal region surrounding the seizure. Specifically, the time duration details of analysis are show in figure 9.

Following are the clustering results from applying the spectral-clustering procedure on different seizures. In both instances, i.e, during spectral clustering and then during K-Means step, the number of clusters were empirically chosen to as 3.

P093, Seizure 2&3, Inter-ictal activity:

C_1: 'LST1' 'LST2' 'LST3' 'LST4' 'RST1' 'LTD7' 'ROF4'

C_2: 'RTD4' 'RTD6' 'RTD8' 'RTD10' 'LOF1' 'LOF2' 'LOF3' 'LOF4' 'ROF1' 'ROF2' 'LTD9'

C_3: 'LTD3' 'LTD5' 'RST2' 'RST3' 'RST4' 'ROF3'

In cluster C_1, all the LST channels are grouped together, while it is easy to see from cluster C_2 that all the RTD and the LOF channels exhibit similar dynamical behavior. Cluster C_3 encompasses the RST channels (except RST1).

P093, Seizure 2&3, Pre-post ictal activity:

C_1: 'RST1' 'RST2' 'RST3' 'RST4' 'LTD7' 'LST2' 'LST3'

C_2: 'LTD3' 'LTD5'

C_3: 'RTD4' 'RTD6' 'RTD8' 'RTD10' 'ROF1' 'ROF2' 'ROF3' 'ROF4'
'LOF1' 'LOF2' 'LOF3' 'LOF4' 'LST1' 'LST4' 'LTD9'

The RTD and the LOF channels continue to be in the same group even during the pre-post ictal states. In fact, we also observe that the ROF channels belong to the same group as the RTD and the LOF channels. As in the inter-ictal activity, the RST channels are closely cluttered within the same group. Unlike during the Inter-ictal activity, the LST channels no longer exhibit a uniform spatio-temporal behavior (We consider an area of the brain such as RTD, LTD or LOF to be strongly belonging to a particular cluster if more than 2 channels out of the 4 channels in each area lie in the same cluster).

Comparing the cluster configuration between inter-ictal and the pre-post ictal stages, we see that the RST channels and LTD channels were similar in terms of the amount of information they exchanged with other areas of the brain, over the inter-ictal period. However, during the pre-post ictal stage, the RST channels may have been closer to the LST channels. Overall, it appears that a few areas of the brain such as the LST and ROF channels underwent drastic changes in their spatio-temporal interactions during the inter-ictal to the pre-post ictal transition.

P093, Seizure 9&10, Inter-ictal activity:

C_1: 'LST1' 'LST2' 'LST3' 'LST4' 'LOF1' 'LOF2' 'LOF3' 'LOF4' 'ROF1' 'LTD3' 'LTD9'

C_2: 'RTD4' 'RTD8' 'RTD10' 'ROF2' 'ROF3' 'ROF4' 'LTD5' 'LTD7' 'RTD6'

C_3: 'RST1' 'RST2' 'RST3' 'RST4'

Here, the left hemisphere channels, (particularly LST and the LOF channels) are closely connected within a cluster, C_1. The right hemisphere channels (namely RTD and the ROF channels) are also closely connected, in a different cluster C_2. C_3 consists of all the RST channels. Firstly, a clear separation between the left and right hemisphere channels is seen and secondly, we see that the RST channels behave differently from the RTD/ROF channels. When the number of clusters was set to 2, the RST channels merged with the other right-hemisphere channels. This implies that the RST channels have very subtle differences with the RTD and ROF channels while a clear difference exists between the RST and the left-hemisphere channels.

P093, Seizure 9&10, Ictal activity:

C_1: 'LOF1' 'LOF2' 'LOF3' 'LOF4' 'LTD3' 'LTD5' 'LTD7' 'LTD9' 'LST1' 'LST2' 'LST3' 'LST4' 'RTD4' 'ROF1'

C_2: 'RST1' 'RST2' 'RST3' 'RST4' 'ROF2'

C_3: 'RTD6' 'RTD8' 'RTD10' 'ROF3' 'ROF4'

All the Left hemisphere channels namely, the LOF, LST and the LTD channels form cluster C_1. Similar to the inter-ictal activity, the right-hemisphere channels and the left-hemisphere channels are clearly separated in terms of their overall spatio-temporal dynamics. Subtle differences exist between cluster configurations in the Inter-ictal and the Pre-post ictal activity though.

Overall comparison between inter-ictal and pre-post ictal behavior shows a very little difference in the cluster-configuration of the channels during seizure 10. This probably implies that the average spatial-distribution of the electrode sites do not have any major changes in the inter-ictal to pre-post ictal transition.

12. SUMMARY

In this study, the SOM-SI measure is proposed to detect functional dependencies among multivariate structures. The measure is simplified and faster as compared to the original SI [15, 16] measure. Simulations were performed on coupled Rossler-Lorenz chaotic systems to check the ability and accuracy of the SOM-SI measure to track spatio-temporal dependencies, both in terms of magnitude and time. It was also shown in an actual ECOG test case that the results from the SOM-SI measure were consistent with the results from the original SI measure. Paired T-tests for comparing the accuracy of the SOM-SI measure with the original SI measure showed that the SOM-SI measure is reasonably in agreement with the SI-measure, at more than 95 percent significance level. This validated the purpose of using a SOM to create a model infrastructure for ECOG and the altered SOM-SI measure to compute dependencies between multivariate systems.

A Spectral-clustering approach was adopted from the cluster literature and applied on the data samples to determine how channels are clustered at every instance in time, and how the clustering configuration changed with time. During inter-ictal to pre-post ictal transition, we observed drastic changes in the spatio-temporal configuration of a few channels in seizure 2 and 3, while on seizures 9 and 10, the analyses showed that the transitional changes were meager. This probably suggests that there may not be a regular pattern associated with channel's spatio-temporal dynamics during the inter-ictal to pre-post ictal transition. However, from the analysis on complex partial seizures (9 & 10), we find that the left hemisphere channels, for most part, may be polarized from the right hemisphere channels. A much larger observation is needed from analysis on more seizures and more patients. Statistical quantification of the changes in spatial-similarity of the ECOG recordings, based on their interaction with other sites on the brain could give us a better understanding of changes in sensory-cortical networks during inter-ictal and ictal periods. So far, we have seen that the SOM-SI coupled with the spectral clustering approach provides a reasonable description of how the brain is spatially connected at different stages of a clinical event. We believe that these results could be a value addition to the current efforts of understanding how the sensory areas are networked with each other, at different times and during different clinical events.

In many applications, it is important to identify critical-sites on the on the brain that exhibit similar behavior. Channel optimization based on data-driven criterion might enable us to do seizure prediction/detection with fewer channels. This can also be perceived as a data-mining approach in which the objective is to minimize the data measurements by extracting/identifying only the relevant data.

Future work will consist of using the spatio-temporal dependency information to identify the optimal electrode sites [22]. We believe that our data mining techniques using spectral clustering will segregate the closely connected channels more optimally because of the fact that the affinity matrix (or the spatio-temporal matrix) is derived from analyzing the spatio-temporal behavior on the actual ECOG data. One of the direct advantages of channel clustering is that it will enable us to identify the channels having similar spatio-temporal distributions. Channels within a cluster are expected to possess similar dynamical behavior. We believe that tracking the changes in the cluster configuration will lead us towards answering more difficult questions such as: How do the spatio-temporal mappings change in the brain?

Acknowledgments: This work was partially supported by a NIH grant on Brain Dynamics, NIH R01 NS 39687

REFERENCES

[1] F.H. Lopes da Silva, W. Blanes, S.N. Kalitzin, J. Parra, P. Suffczynski and D.N. Velis, " Dynamical Diseases of Brain Systems: Different Routes to Epileptic Seizures (Invited Paper)," *IEEE Transactions on BioMedical Engineering*, vlo. 50, pp. 540-549, 2003.

[2] J.S. Bendat , A.G. Piersol, *Random Data: Analysis and Measurement Procedures,* 2nd ed. New York: John Wiley, 1986.

[3] R.B. Duckrow and S.S. Spencer, "Regional coherence and the transfer of ictal activity during seizure onset in the medial temporal lobe," *Electroenceph Clin Neurophysiol*, vol. 82, pp. 415-422, 1992.

[4] M. Kaminski and K.J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern*., vol. 65, pp. 203 -210, 1991.

[5] P.J. Franaszczuk, G.K. Bergey and M. Kaminski, "Multichannel Autoregressive Analysis of Mesial Temporal Structures using the method of Directed Transfer Functions," *Epilepsia*, vol. 34 (Suppl. 6), 1993.

[6] P.J. Franaszczuk, G.K. Bergey and M. Kaminski, "Analysis of mesial temporal seizure onset and propagation using the Directed Transfer Function Method," *Electroenceph. clin. Neurophysiol.*, vol. 91(6), pp. 413-427, 1994.

[7] P.J. Franaszczuk and G.K. Bergey, "Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures," *Brain Topography*, vol. 11, pp: 13-21, 1998.

[8] L.A. Baccala and K. Sameshima, "Overcoming the Limitations of Correlation analysis for many simultaneously processed neural structures," *Brain Research*, vol. 130, pp.33-47, 2001.

[9] E. Rodriguez, N. George, J.-P. Lachaux, J. Martinerie, B. Renault and F. J. Varela, "Perception's shadow: long-distance synchronization of human brain activity," *Nature*, vol. 397, pp. 430-437, 1999.

[10] K J. Friston, "Brain Function, Nonlinear Coupling and Neuronal Transients," *The Neuroscientist*, vol. 7 (5), pp. 406 – 418, 2001.

[11] M.G. Rosenblum, A.S. Pivosky and J. Kurths, "Phase Synchronization of Chaotic Oscillators," *Physics Review Letters*, vol. 76, pp. 1804-1807, 1996.

[12] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, "The Empirical Mode Decomposition and Hilbert transform for Non-linear and Non-Stationary Time Series Analysis," *Proc. Roy. Soc. London A*, Vol. 454, pp.908-995, 1998.

[13] B. Pompe, "Measuring Statistical Dependencies in a Time Series," *Journal of Statistical Physics*, vol. 73, pp.587-610, 1993.

[14] J.P. Eckmann and K.D. Ruelle, *Europhysics Letters*, vol. 5, pp. 973, 1987.

[15] J. Arnhold, P. Grassberger, K. Lehnertz and C.E. Elger, "A Robust Method for Detecting Interdependencies: Application to Intracranially Recorded EEG," *Physica D*, vol. 134, pp. 419-430, 1999.

[16] R.Q. Quiroga, J. Arhnold and P. Grassberger, "Learning Driver-Response Relationships from Synchronization Patterns," *Physical Review E*, vol. 61, pp. 5142-5148, 2000.

[17] A.Hegde, D. Erdogmus, Y. Rao, J.C. Principe and J. B. Gao, "SOM-based Similarity Index Measure: Quantifying Interactions between Multivariate Structures," *Proceedings of NNSP'03*, pp. 819-828, Toulouse, France, Sep 2003.

[18] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice Hall, 1999.

[19] J.C. Principe, N.R. Euliano and W.C. Lefebvre, *Neural and Adaptive Systems:Fundamentals through Simulations*, John Wiley & Sons, 2000.

[20] A. Hegde, D. Erdogmus and J.C. Principe, "Synchronization Analysis of Epileptic ECoG Data Using SOM-Based SI Measure," *Proceedings of EMBS'04*, pp. 952-955, San Francisco, Sept 2004.

[21] A.Y. Ng, M.I. Jordan and Y. Weiss, "On Spectral Clustering: Analysis and an Algorithm," In T.G. Dietterich, S. Becker, & Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14,* Cambridge, MA. MIT Press.

[22] L.D. Iasemidis, P. Pardalos, J.C. Sackellares and D.S. Shiau, "Quadratic Binary Programming and Dynamical System Approach to Determine the Predictability of Epileptic Seizures," *Journal of Combinatorial Optimization*, 5, pp. 9-26, 2001.

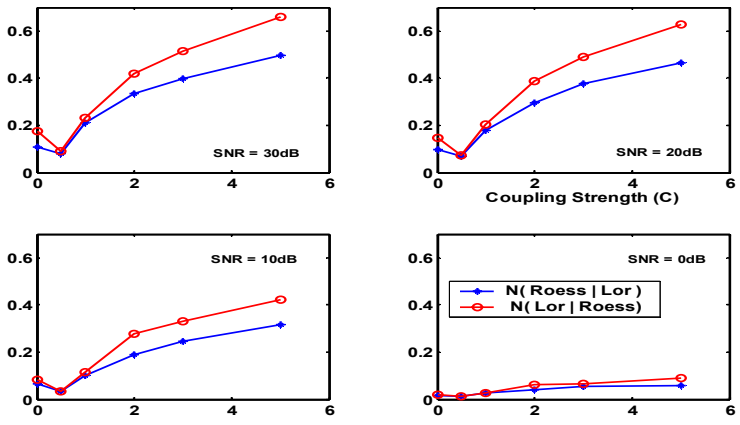


Figure 2. Illustrating the dependency relationships between the Roessler (driver) and the Lorenz system (response) as a function of coupling strength.

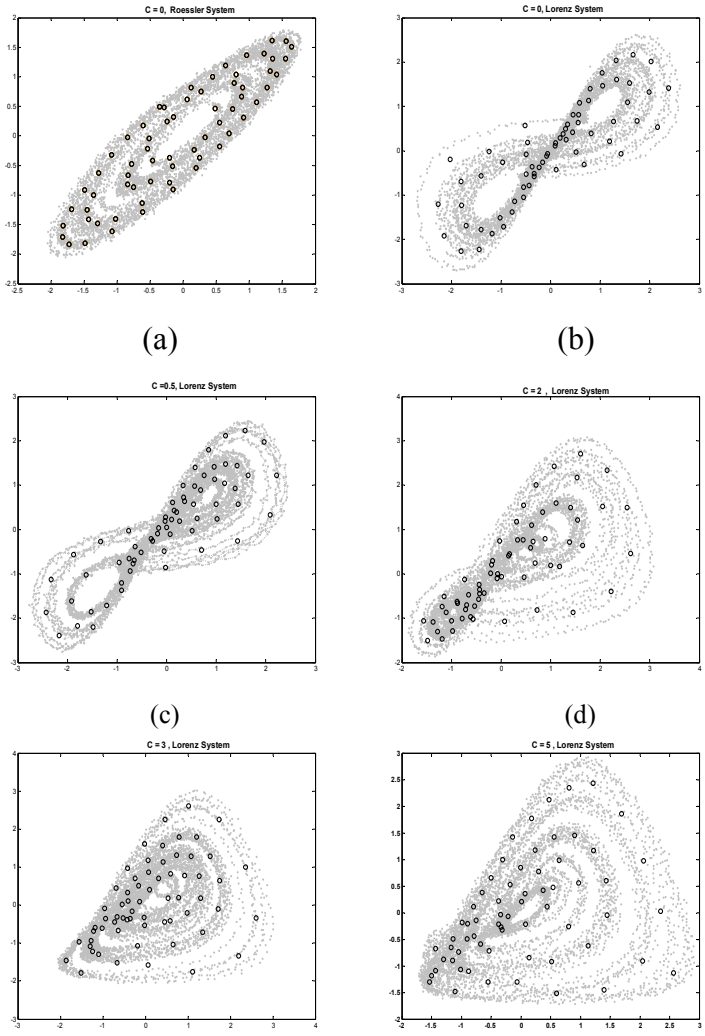


Figure 1. Phase-space trajectories of the Roessler-Lorenz system for various coupling strengths a) Rossler b) Lorenz (C=0) c) Lorenz (C=0.5) d) Lorenz (C=2) e) Lorenz (C=3) f) Lorenz (C=5). The SOM weights (circles) for each signal are superimposed on the trajectory.

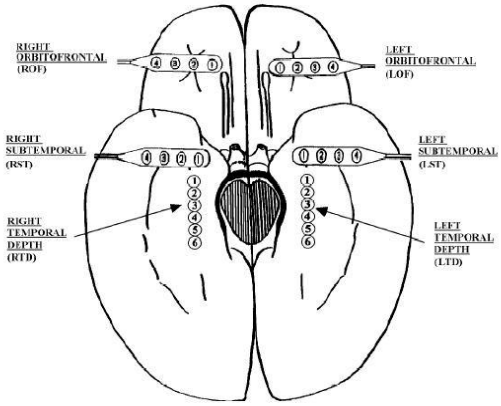


Figure 3. Diagram of the depth and subdural electrode placement. Electrode strips are placed over the left orbitofrontal (LOF), right orbitofrontal (ROF), left subtemporal (LST), right subtemporal cortex (RST). Depth electrodes are placed on the left temporal depth (LTD) and right temporal depth (RTD), to record

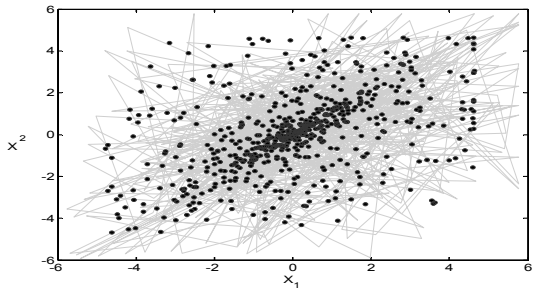


Figure 4. The phase-space trajectory of the training ECOG signal (solid lines) superimposed on the weights (dots) of the trained 25x25 ECOG-SOM grid.

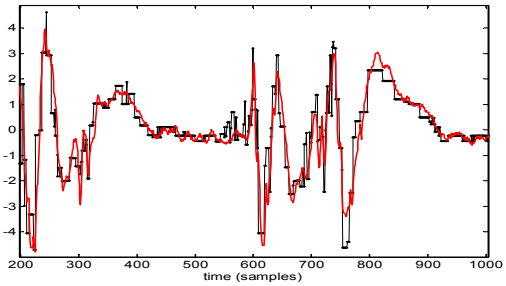


Figure 5. A qualitative illustration of the accuracy of the 25x25 ECOG-SOM. Sample test signal (solid line) overlapped on the SOM-reconstructed output (dash and dot line). In this case, the correlation coefficient = 90.1%.

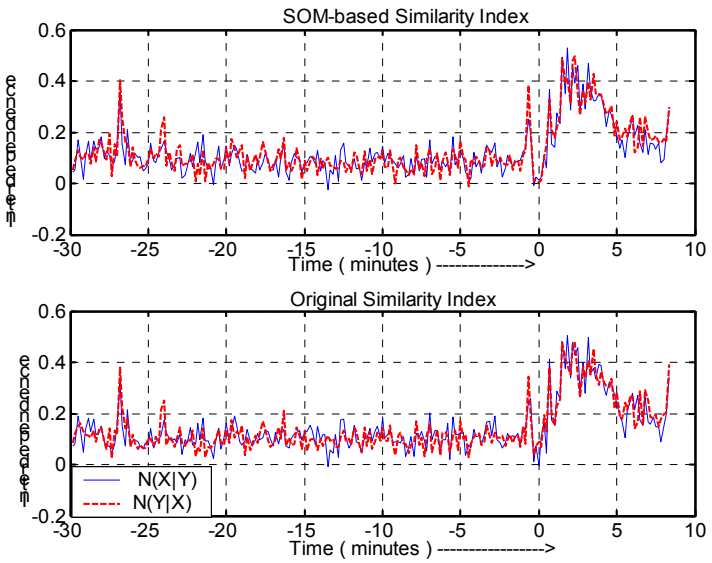


Figure 6. SOM-SI results showing the dependencies $N(X|Y)$, $N(Y|X)$. X and Y correspond to channels RTD4 and RST1, respectively. The time instant '0' corresponds to the seizure onset (top). Results produced by the original SI (bottom).

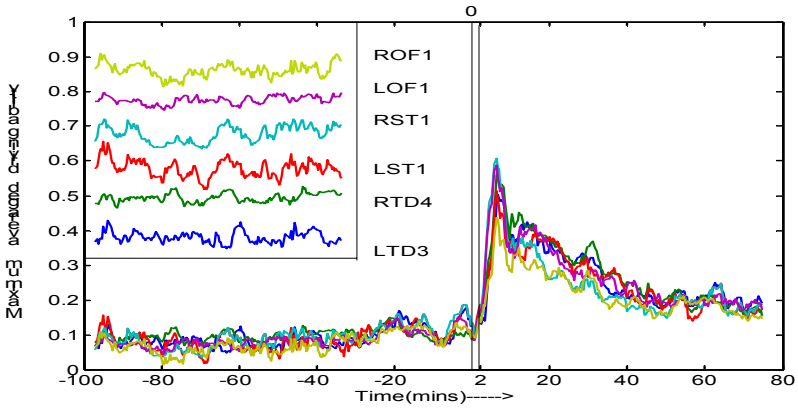


Figure 7. Maximum average driving ability of each of the six (6) channels, nearly 100 minutes before and 70 minutes after Seizure-1 in patient P092. (*The thin vertical bar corresponds to the time when seizure occurred (0 to 2 on the x-axis). For clarity, the box inside the figure shows a small portion of the maximum average driving ability of each of the 6 channels, baseline-offset by different scales.*) Drop in synchronization followed by an abrupt increase in phase synchronization at the onset of seizure is evident. Synchronization across channels during seizure is also clearly seen

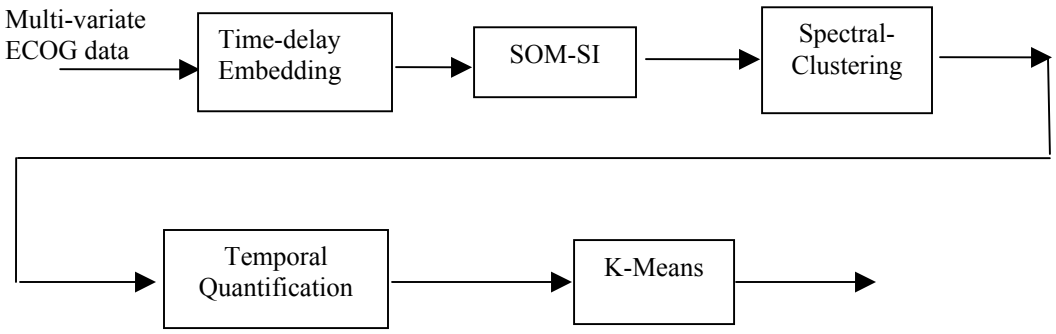


Figure 8. Block diagram to extract spatio-temporal dependency information in Multivariate ECOG structures

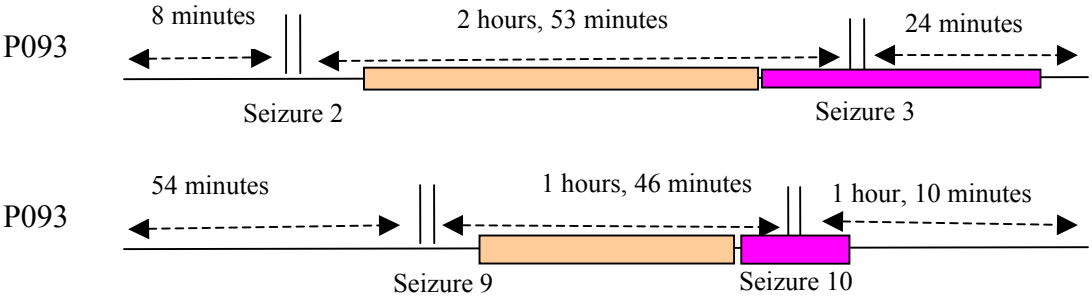


Figure 9. Graphical illustration of the data analysis on patient P093 (figures not drawn up to scale). Inter-ictal region (shaded): 20 minutes after seizure 2 (correspondingly seizure 9) upto 20 minutes before seizure 3 (correspondingly seizure 10). Pre-Post Ictal region (shaded): 20 minutes before seizure 3 (or seizure 10) and 20 minutes after seizure 3 (or seizure 10).

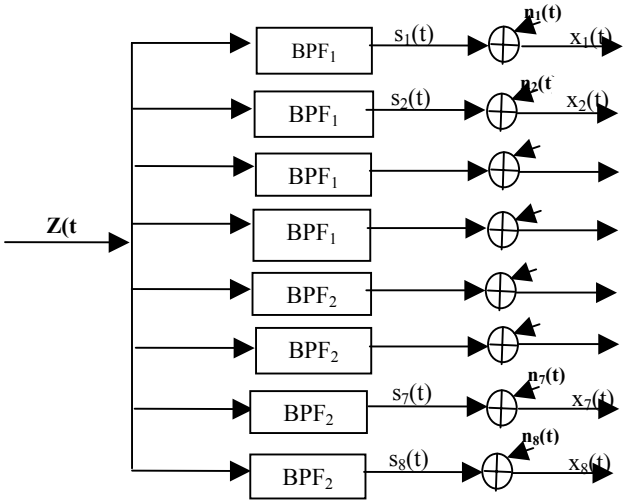


Figure 10. Model showing the generation of linearly dependent signals using configuration 1.

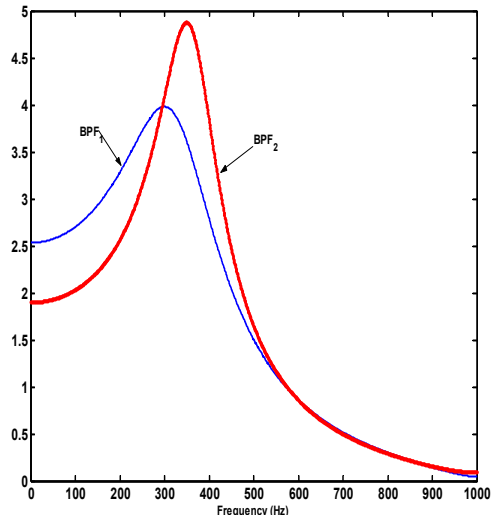


Figure 11. Plot showing the two band pass filters, separated by a narrow pass band

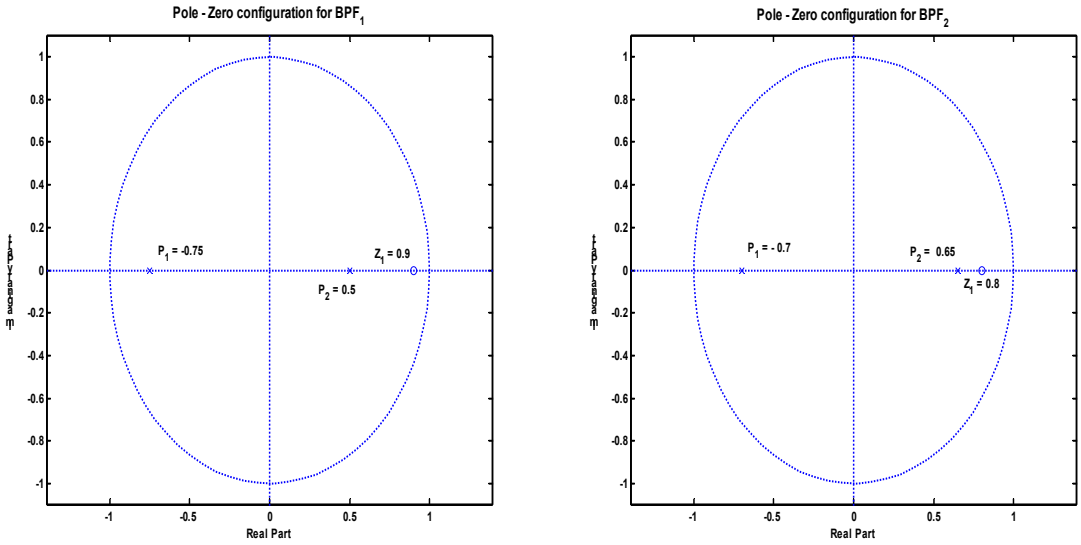


Figure 12. Z-planes showing the pole- zero configurations for the two band-pass filters in Fig 11.

Table 1. Outline of the SOM-SI algorithm

1. Train a SOM using embedded vectors from both X and Y as the input.
2. At time n, find W_n^x, the winner neuron for vector x_n, and find W_n^y, the winner neuron for vector y_n.
3. To find $R^n(X)$, compute the average Euclidean distance between W_n^x and all the other winner neurons in the SOM. Similarly, compute $R^n(Y)$.
4. Determine the sets X_n and Y_n for W_n^x and W_n^y, respectively.
5. Determine the nearest neurons $W_{n,j}^y$ corresponding to vectors y_j, where $j \in X_n$. Determine the nearest neurons $W_{n,j}^x$ corresponding to vectors y_j, where $j \in Y_n$.
6. Calculate $R^n(X \mid Y) = (1/q)\Sigma_{j=1}^q \| W_n^x - W_{n,j}^x \|$, where q is the number of elements in X_n. Calculate $R^n(Y \mid X) = (1/q)\Sigma_{j=1}^q \| W_n^y - W_{n,j}^y \|$, where q is the number of elements of Y_n.
7. Compute the ratios,
   $N^n(X \mid Y) = \left(R^n(X) - R^n(X \mid Y)\right) / R^n(X)$ (2)
   $N^n(Y \mid X) = \left(R^n(Y) - R^n(Y \mid X)\right) / R^n(Y)$ (3)
8. Find *interdependencies* $N(X \mid Y)$ and $N(Y \mid X)$ as the average of $N^n(X \mid Y)$ and $N^n(Y \mid X)$ over all n.
9. Compute the SOM-SI as the difference,
   $\chi = N(Y \mid X) - N(X \mid Y)$.