

Synchronization analysis of Epileptic ECOG data using SOM-based SI measure

Anant Hegde, Deniz Erdogmus, Jose C. Principe

Department of Electrical & Computer Engineering, Computational NeuroEngineering Laboratory,

University of Florida, Gainesville, FL – 32611

Email: [ahegde,deniz,principe]@cnel.ufl.edu

Abstract—The exact spatio-temporal changes leading to epileptic seizures, although widely studied, are not well understood yet. We propose to investigate the mechanisms leading to epileptic seizures 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. Application of SOM-based SI measure on epileptic seizure data reveals interesting aspects of synchronization and de-synchronization at various spatio-temporal levels.

Keywords—epilepsy, seizure prediction, synchronization

I. INTRODUCTION

Epileptic studies on adults and neo-nates have been widely researched. Researchers from different disciplines have contributed immensely in understanding the various neurological aspects leading to seizures. However, we still do not have a drug that cures this disease, nor, do we have a reliable technology to prevent its occurrence. Synchronization and de-synchronization between the cortical-columns of the brain are believed to be one of the plausible reasons for most neurological disorders including epilepsy [1]. 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. The spatio-temporal changes in information across different regions of the brain are rapid and often very subtle. 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 help us develop seizure-warning systems. This quantification would also help us identify the regions that actively participate during epileptic seizures.

Many linear and nonlinear techniques that address the quantification of synchronization exist in the multivariate analysis literature. Correlation and coherence are commonly used techniques, e.g., directed coherence (DC), partial-directed coherence (PDC) [2]. The main drawback of these methods is that they only describe linear interactions, although the brain is known to be highly nonlinear. Instantaneous phase measures using Hilbert transforms and wavelet transforms have also been used to quantify

synchronization [3]. However, the applicability of these measures is restricted to identifying phase locking between two signals. Secondly, for accuracy, signals are required to be narrow-band, which is unsuitable for ECOG signals. Generalized mutual information function (GMIF) [4] is popularly used, however, this approach requires large data sets for probability density estimation, and is computationally expensive. Variants of recurrence plots [5] are used to measure recurrence of states in the phase-space between two chaotic signals. However, these plots are difficult to analyze due to the lack of quantitative measures, in addition to their computational complexity.

Earlier [6], we proposed the self-organizing-map-based similarity index (SOM-SI) measure, as a computationally simplified alternative of the SI technique, originally proposed in [7]. This technique characterizes the asymmetric synchronizations among multiple signals. Conceptually, SI is based on the assumption that if there is a functional dependency between two signals, then the recurrence of dynamics of one signal will mean the recurrence of dynamics in the other signal. Arnhold *et al.* [7] propose to search for recurrence in the signals' state-space, which poses an enormous computational burden, especially for large data sets. The SOM-SI measure reduces the computational complexity, while maintaining accuracy [6]. The central idea of the SOM-SI is to reduce complexity by quantizing the signals' state space using a topology-preserving map.

In this paper, we study the synchronization patterns in epileptic ECOG data, using the SOM-SI. First, we present a brief overview of SOM-SI. Next, the equivalence of the results obtained by SI and SOM-SI is shown by statistical tests. Section IV discusses epileptic ECOG synchronization as quantified by the SOM-SI. In Section V, we provide a brief discussion on possible future directions for research.

II. METHODOLOGY

Synchronization between two signals X and Y is understood, in principle, as a functional mapping between X and Y . In the brain, it is realistic to expect the functional mappings to be nonlinear. The attractors of functionally synchronous systems will be related. In other words, if the trajectory of Y is influenced by the trajectory of X , then Y is said to be functionally dependent on X . However, in instances where bi-directional relationships exist, the trajectory of Y can also partly influence the trajectory of X . In such a case, it is important to quantify two aspects of

synchronization: direction and strength.

A. Original SI measure

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,

$$N(X|Y) = \frac{1}{N} \sum_{n=0}^{N-1} \frac{R^n(X) - R^n(X|Y)}{R^n(X)} \quad (1)$$

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 [10,11].

By design, SI can quantify nonlinear dependencies. The difficulty with this approach is that at every time instant, we must search for the k nearest neighbors of the current embedded vector sample among all samples. Thus, the computational complexity is $O(N^2)$, N being the total number of samples. This high complexity hinders real-time implementation and analysis.

B. SOM-based SI algorithm

A SOM [8] is a neural-network in which spatial patterns from the input-space are mapped onto an ordered output space consisting of a set of neurons, called processing elements (PE). Thus each neuron in the SOM, based on its location on the map, compactly models different features/dynamics of the input. In the application of SOM modeling to the similarity index concept, the topology preservation feature of the SOM will be of added advantage, because of the fact that the neighboring neurons in the feature space will now correspond to neighboring states in the input data.

Define the activation region of a neuron in the SOM as the set of all input vectors (the embedded signal vectors) for which the neuron 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 neuron corresponding to the input vector x_n at time n . Similarly define the set Y_n . Then the procedure to estimate the directed SOM-SI between X and Y is as follows:

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 .

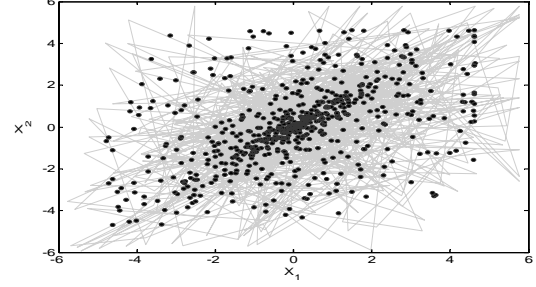


Figure 1. The phase-space trajectory of the training ECG signal (solid lines) superimposed on the weights (dots) of the trained 25x25 ECOG-SOM grid.

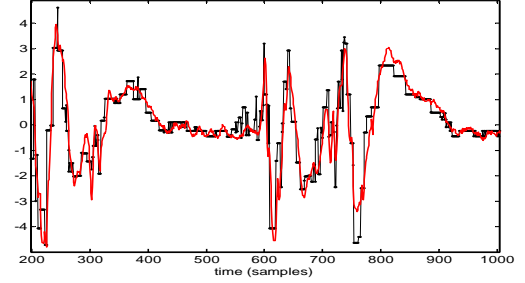


Figure 2. A qualitative illustration of the accuracy of the 25x25 ECOG-SOM. Sample test signal (solid line) overlapped on the SOM-reconstructed output (dash and dot line). In this case, the correlation coefficient = 90.1%.

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 $R^n(X|Y) = (1/q) \sum_{j=1}^q \|W_n^x - W_{n,j}^x\|$, where q is the number of elements in X_n . Calculate $R^n(Y|X) = (1/q) \sum_{j=1}^q \|W_n^y - W_{n,j}^y\|$, where q is the number of elements of Y_n .
7. Compute the ratios,

$$N^n(X|Y) = \frac{R^n(X) - R^n(X|Y)}{R^n(X)} \quad (2)$$

$$N^n(Y|X) = \frac{R^n(Y) - R^n(Y|X)}{R^n(Y)} \quad (3)$$
8. Find *interdependencies* $N(X|Y)$ and $N(Y|X)$ as the average of $N^n(X|Y)$ and $N^n(Y|X)$ over all n .
9. Compute the SOM-SI as the difference,

$$\chi = N(Y|X) - N(X|Y).$$

Positive values of χ indicate that influence of X on Y is more than the influence of Y on X , while negative values indicate the opposite. Higher magnitude of χ indicates a stronger coupling of the signals.

The nearest neighbor search involves $O(NM)$ operations as opposed to $O(N^2)$ in the original SI, where M is the number of PEs. Traditionally $M \ll N$, hence, SOM-SI offers a significant reduction in computations compared SI.

III. RESULTS

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 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. 1 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. 2, 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 ($> 25 \times 25$), 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.

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. The entire interval of 39 minutes data was segmented into 230 non-overlapping windows of 10 seconds each. Fig. 3 shows the interdependency and the SI values of both the measures. It is easy to see that the results

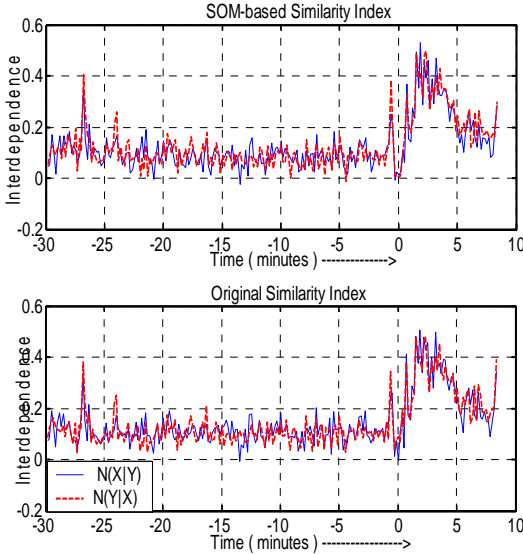


Figure 3. SOM-SI results showing the dependencies $N(X|Y)$, $N(Y|X)$. X and Y correspond to channels RTD4 and RST1, respectively. The time instant ‘0’ corresponds to the seizure onset (top). Results produced by the original SI (bottom).

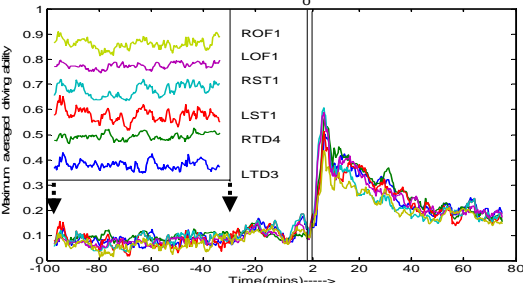


Figure 4. Maximum average driving ability of each of the six (6) channels, nearly 100 minutes before and 70 minutes after Seizure-1 in patient P092. (The thin vertical bar corresponds to the time when seizure occurred (0 to 2 on the x-axis). For clarity, the box inside the figure shows a small portion 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.

from both the measures are in agreement to a large extent. 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. Assuming that SOM-SI and SI measure values come from normal populations, 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 $\alpha=0.05$, 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 window 1 of original SI 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. Therefore, we conclude that statistically the SOM-SI measure, computed with a 25x25 grid SOM, performs as well as the original SI measure.

To analyze spatio-temporal synchronization in an epileptic brain, six representative channels representing different regions on the brain were selected. Pair-wise interdependence was computed among all the 15 possible combinations of the six channels on 2 patients involving 5 seizures (Partial secondary generalized (PSG) and Complex partial (CP) seizures) between them. The interdependence values tend to change very sharply between windows, so a smoothing is done by applying a rectangular-window moving average (length 10). Also, it can be observed from Fig. 3, the interdependence values characterizing the driving and sinking ability of the channels do not exhibit a major difference in patterns. Therefore driving and sinking abilities of the brain areas can only be addressed through statistical tests on the SI results.

On the smoothed interdependence signals, 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 are shown in Fig. 4. 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. 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 five seizures from both patients.

Interestingly, the thin lines at the seizure onset also indicate a spatial synchrony. Clinically, this behavior may indicate that the oscillators at different cortical columns in the brain interact equally with each other even though the degree of interaction is very low. As seizure subsides, it possibly resets the brain. If indeed ECOG signals are chaotic, as pointed out by Iasemidis *et al.* [9], and if we assume that ECOGs at different spatial locations are the same signals starting at different initial conditions, then it is possible that they will display higher degree of interaction in

the transient states (post-seizure) as opposed to the steady state conditions (inter-ictal state).

IV. CONCLUSIONS

In this paper, the SOM-SI measure is used to detect functional dependencies among multivariate structures. The measure is simplified and faster as compared to the original SI [7] measure. Using this measure on epileptic ECOG data, we found changes in synchronization patterns both, across time and space. In other words, the interdependency values revealed information on how spatial connectivity's change in the brain, prior to, during and after seizure. Statistical tests on SOM-SI values may reveal further insight on the exact driver/sink relationships among channels. Identifying the optimal subset of channels for prediction studies may be another important step towards real time analyses. As a possible future effort, we intend to do dependency analyses on extensive data, involving larger set of patients, electrodes and seizures. Also, quantifying characteristic changes in spatio-temporal patterns among *different types* of seizures may further aid in reliably predicting the exact type of seizure.

ACKNOWLEDGMENT

This work is supported by NIH grant R01 NS 39687

REFERENCES

- [1] W.J. Freeman, "Characteristics of the synchronization of brain activity imposed by finite conduction velocities of axons," *International Journal of Bifurcation and Chaos*, pp. 2307-2322, 2000.
- [2] 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.
- [3] M.L.V. Quyen, J. Foucher, J. Lachaux, E. Rodriguez, A. Lutz, J. Martinerieie, Francisco, J. Varela, "Comparison of Hilbert transform and wavelet methods for analysis of neuronal synchrony," *Journal of Neuroscience Methods*, vol 111, pp. 83-98, 2001.
- [4] B. Pompe, "Measuring statistical dependencies in a time series," *Journal of Statistical Physics*, vol. 73, pp.587-610, 1993.
- [5] N. Marwan and J. Kurths, "Nonlinear analysis of bivariate data with cross recurrence plots," *Physics Letters A*, vol. 302, pp. 299-307, 2002.
- [6] 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.
- [7] 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.
- [8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice Hall, 1999.
- [9] L.D.Iasemidis and J.C. Sackellares, "Chaos theory and epilepsy," *The Neuroscientist* 2, pp 118-126, 1996.