FREQUENCY-SPECIFIC NON-LINEAR GRANGER CAUSALITY IN A NETWORK OF BRAIN SIGNALS

Archishman Biswas[1] & Hernando Ombao[1]

[1]Statistics Program, King Abdullah University of Science and Technology (KAUST)

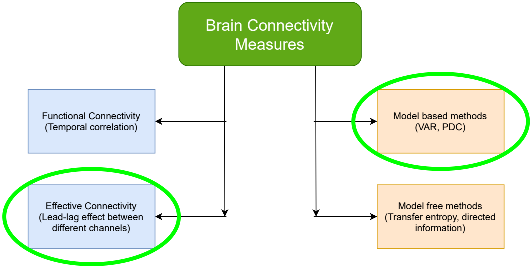
April 16, 2022

Overview

- Brain connectivity measures and non-linear connectivity analysis.
- Architecture and formulation of the Component-wise Multi-Layer Perceptrons(cMLPs) and their use.
- Modifications to include frequency band specific connectivity estimates and dealing with non-stationarity.
- Simulations to showcase the utility of proposed NLGC and Spec NLGC methods.
- Implementing proposed method to an EEG time series data recorded during an epileptic seizure.
- Conclusion and future research prospects.

Brain Connectivity Measures

- Brain connectivity network dynamics is key to understanding many complex neuronal processes.



- Model-based measures used in current studies mostly assume linear directed connections between the channels.
- Granger causality (GC) is a powerful measure that is used frequently to analyze effective connectivity in multi-channel brain signals.
- GC is often implemented in context of VAR models, which assumes that underlying connections are linear.

Non-Linear Connectivity Analysis

- **Kernel based methods:** In some studies, GC has been implemented using kernel functions to get non-linear GC estimates.
- **Our Method:** We propose NLGC and Spec NLGC models which utilizes component-wise MLPs to get non-linear GC connection estimates.
- Past studies have shown the utility of MLPs in time-series forecasting. But, due to black-box nature of MLPs, it is hard to use them for directed connectivity estimation.
- **Solution:** Use component-wise MLPs(i.e. cMLP), which is using one MLP for every channel of the data

Component-Wise Multi-Layer Perceptrons

- A generalization of the classical VAR(K) model would be to model the current values $X(t)$ using past values $X_1(t'), X_2(t'), ..., X_N(t')$ using some non-linear function $g(.)$ such that:

$$X(t) = g(X_1(t'), X_2(t'), ..., X_N(t')) + \epsilon(t)$$

- $g(.)$ is model using MLPs. We model each channel separately, i.e. using cMLPs to get a interpretable architecture:

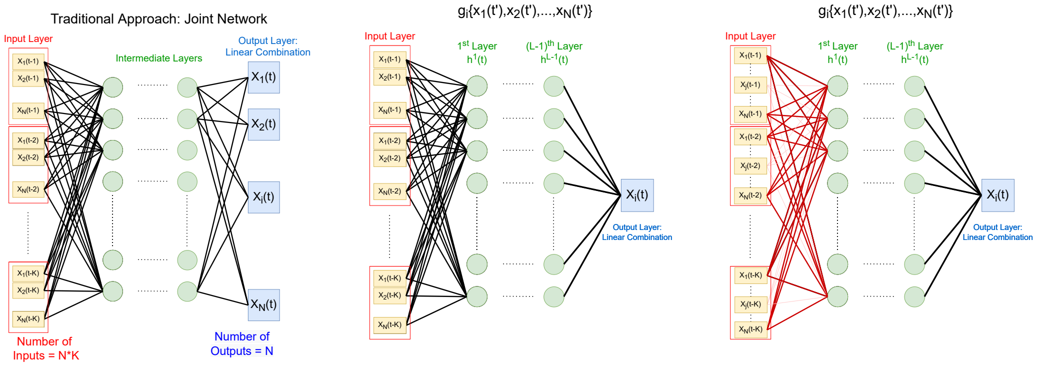
$$X_i(t) = g_i(X_1(t'), X_2(t'), ..., X_N(t')) + \epsilon_i(t)$$

- To implement each $g_i(.)$, we implement cMLPs of single hidden layer $h^1(t) \in \mathbb{R}^H$ with H neurons:

$$Hidden\ layer:\ \ h^1(t) = \sigma\left[\sum_{n=1}^{K} W^{1n} X(t-n) + b^1\right]$$

$$Output\ layer:\ \ X_i(t) = W^2 h^1(t) + b^2$$

cMLP Architecture Comparisons

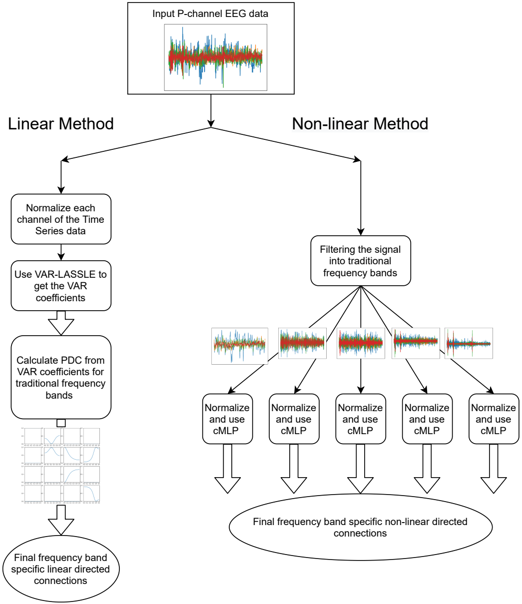


(a) Traditional use (b) Component-wise MLP (c) Condition for $X_j \not\to X_i$

$$\min_{W^1,W^2,b^1,b^2} \sum_{t=n}^{T} [X_i(t) - g_i(X(t-1),\ldots,X(t-K))]^2$$

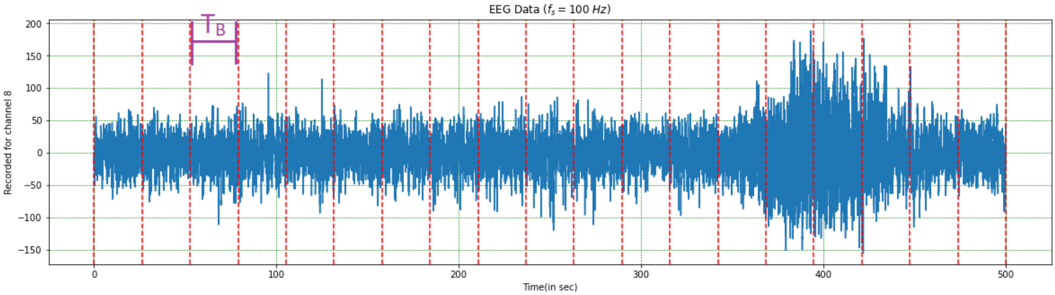
$$+\lambda \sum_{j=1}^{N} \sum_{n=1}^{K} \|(W^{1n}_{:j},\ldots,W^{1K}_{:j})\|_2$$

Frequency Band specific Connections



- Past studies have shown the existence of frequency specific connections in modalities like EEG, LFP, fMRI.
- We utilize a 3rd order Butterworth filter to decompose each channel of EEG signal into delta(0.5-4.0 Hz), theta(4.0-8.0 Hz), alpha(8.0-12.0 Hz), beta(12.0-30.0 Hz), gamma(30.0-50.0 Hz)
- Effectively this gives us a total of 5 time-series data for each of the channels.

Dealing with Non-Stationarity



- Non-stationary behaviour in neuronal time-series data can occur for many reasons, and there are many sophisticated methods to deal with it.
- For our case, we have just simply used a over-lapped time window approach.
- The time window/block size is to be selected with caution, considering the trade-off between small windows leading to better time-resolution and poorer cMLP training and vice-versa.

Experiments Overview

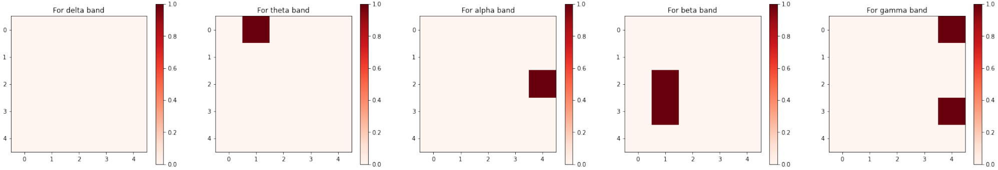
- Non-linear mixtures of AR(2) processes are used for the simulations. This gives us the ground truth to evaluate our scheme.
- To check performance under noisy conditions, the noise levels of the signals are varied using AWGN of SNR: $\{2 \text{ dB}, 5 \text{ dB}, 10 \text{ dB}, 15 \text{ dB}, 20 \text{ dB}\}$.
- For all the simulations and implementation on seizure EEG data, a single hidden layer neural-network is chosen, with number of neurons in hidden layer $= \text{H} = 100$.
- The cMLPs are trained using the hierarchical penalty and proximal gradient descent with a line search is used for training the networks.
- The mean and the median absolute deviations for the AUPR scores for 5 random realizations of each of the setting is reported.

Overall Non-Linear GC Connectivity

- The simulations for the overall NLGC are done in order to get the idea of what effect the SNR has on the proposed NLGC performance
- N $=$ 10 channels non-linear data is generate using 2 sets of AR(2) latent sources
- The non-linearity is induced using a transfer function of the form $\tau(x) = a + bx^2$
- The ground truth is set such that there are a total of 18 true connections among the 90 possible total connections

Frequency Specific NLGC Connectivity

- The ground truth is generated in a similar manner to that of the NLGC, using a different non-linear transfer function $\tau(x)$
- N $=$ 5 channels were used in the simulations, which can be decomposed into 5 bands each, leading to a total of 25 decomposed signals.
- 6 actual connections were used in the ground truth data, the figure below explains the true connectivity patterns:



Visualizing the Individual AR(2) Processes

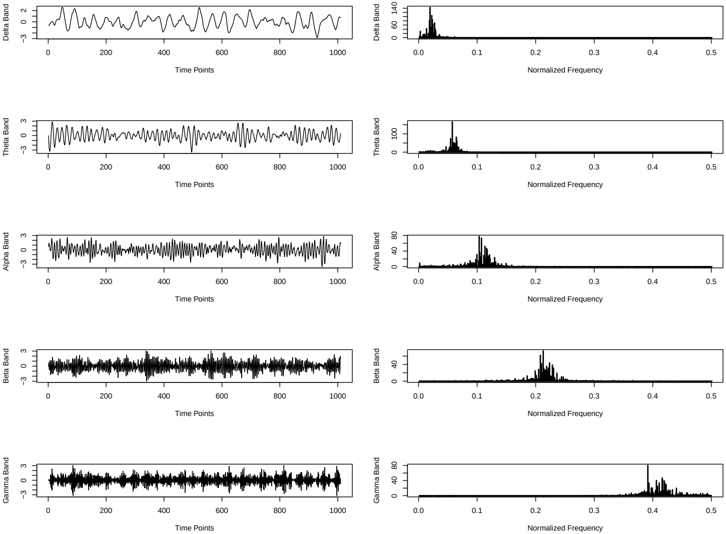


Figure: Latent sources from AR(2) processes

Table Showing the Simulation Results

Method	2 dB	5 dB	10 dB	15 dB	20 dB
VAR-LASSLE(1)	0.42± 0.01	0.42± 0.02	0.41± 0.01	0.41± 0.01	0.41± 0.01
NLGC(1)	0.62± 0.03	0.72± 0.05	0.84± 0.02	0.82± 0.02	0.8± 0.03
PDC(1)	0.26± 0.03	0.23± 0.00	0.25± 0.03	0.23± 0	0.24± 0
Spec NLGC(1)	0.79± 0.02	0.82± 0.01	0.9± 0.01	0.91± 0.02	0.81± 0.02
VAR-LASSLE(2)	0.41± 0.03	0.41± 0.02	0.41± 0.01	0.42± 0.03	0.44± 0.02
NLGC(2)	0.45± 0.03	0.45± 0.06	0.71± 0.05	0.87± 0.02	0.9± 0.05
PDC(2)	0.32± 0.03	0.26± 0.05	0.24± 0.06	0.14± 0.03	0.11± 0
Spec NLGC(2)	0.63± 0.035	0.68± 0.07	0.8± 0.02	0.9± 0.03	0.92± 0.03
VAR-LASSLE(3)	0.4± 0.02	0.4± 0.03	0.44± 0.02	0.43± 0.02	0.43± 0.02
NLGC(3)	0.36± 0.06	0.42± 0.03	0.75± 0.04	0.88± 0.03	0.93± 0.00
PDC(3)	0.26± 0.05	0.27± 0.05	0.29± 0.03	0.17± 0.04	0.28± 0.10
Spec NLGC(3)	0.65± 0.07	0.74± 0.06	0.91± 0.03	0.98± 0.01	0.99± 0.00

Analysis of Seizure EEG Data

- We apply the NLGC and Spec NLGC method on a 18-channel seizure EEG data with 50,000 time-points, having a sample rate of 100 Hz.
- We used a time-windowed approach considering the quasi-static nature of EEG signals using a 50% overlap and 2000 time samples in each window.
- This gives 500 time-points overlap on each side of the tie window, leading to a total of 33 GC matrices over the 500 second recording.
- In order to understand the network dynamics and visualize the amount of change in the GC connectivity network, we plotted the *Euclidean Distance*(ED(t)) between consecutive GC matrices:

$$ED(t) = \sqrt{\sum_{all \;\; i,j} \mid [GC(t)]_{i,j} - [GC(t-1)]_{i,j} \mid^2}$$

Comparison for Overall Connections

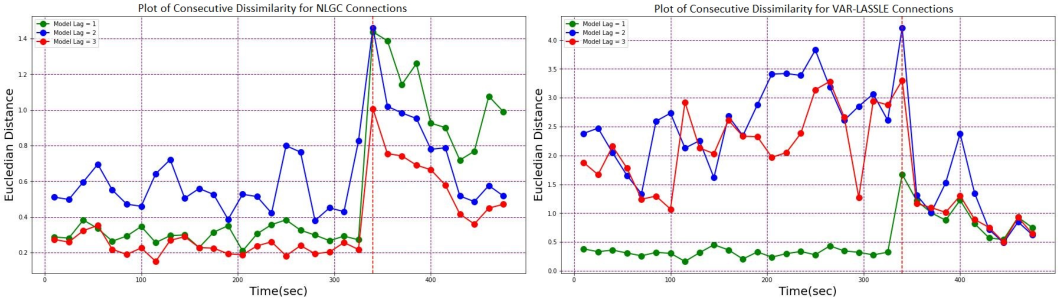


Figure: Comparison between NLGC and VAR-LASSLE

- Consecutive dissimilarity between directed connectivity is plotted using NLGC and traditional VAR-LASSLE, with model lags of $K = 1,2,3$.
- The sudden rise of the consecutive dissimilarity of NLGC method suggests that our method is able to detect the start of the seizure quite well.
- This is not true for the case of VAR method where the rise in consecutive dissimilarity not much appreciable.

Visualizing the Estimated NLGC Connections

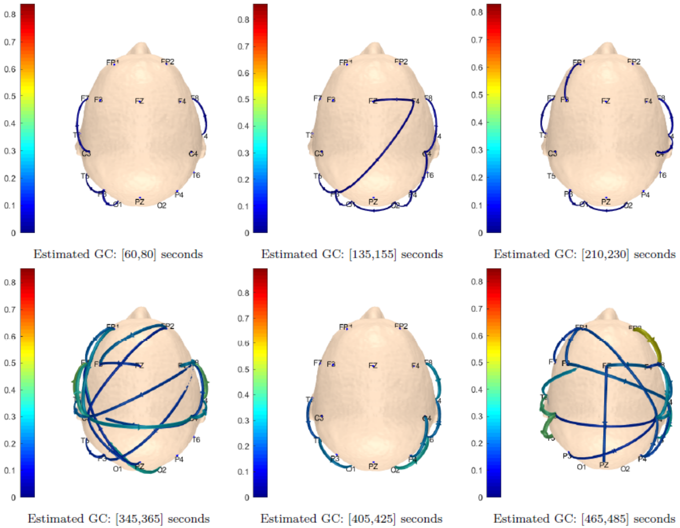


Figure: Propagation of NLGC connections from left to right hemisphere during seizure

Comparison for Frequency Specific Connections

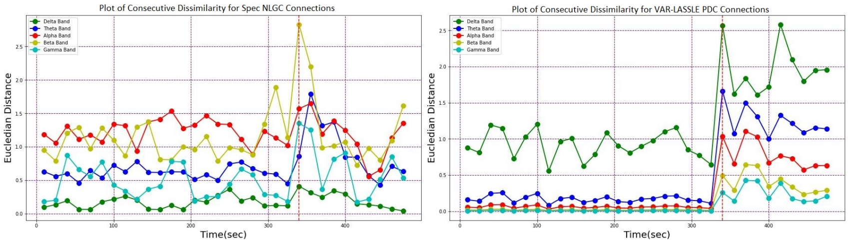


Figure: Comparison between Spec NLGC and PDC

- Consecutive dissimilarity between directed connectivity in each frequency band is plotted using Spec NLGC and traditional VAR-LASSLE based PDC, with model lag of $K = 1$.
- In Spec NLGC, sudden change occurs mostly in the theta, beta and gamma bands. This is consistent with past studies.
- In case of VAR-LASSLE based PDC, we observe sudden change in lower frequency bands which is inconsistent with past studies.

Conclusion and Future Work

- We have introduced and evaluated performance of a frequency band specific non-linear Granger causality framework combining Butterworth filters and component-wise MLP networks with hierarchical penalty.
- Simulation results on non-linear data shows the huge improvement on use of proposed methods over traditional methods.
- Implementation on epileptic EEG data provides novel findings about time evolving connectivity pattern between different EEG channels.
- Integration of Spec NLGC with sophisticated approaches to deal with non-stationarity can be explored in future studies.
- We have deployed simulations settings as per need in brain signal analysis, but implementation of the method proposed in fields like financial data analysis would also be worth exploring.

The End