Identifying Granger causal interactions in alcoholic brains using EEG data

Akshay Bhatia

November 2018

Contents

1 Background

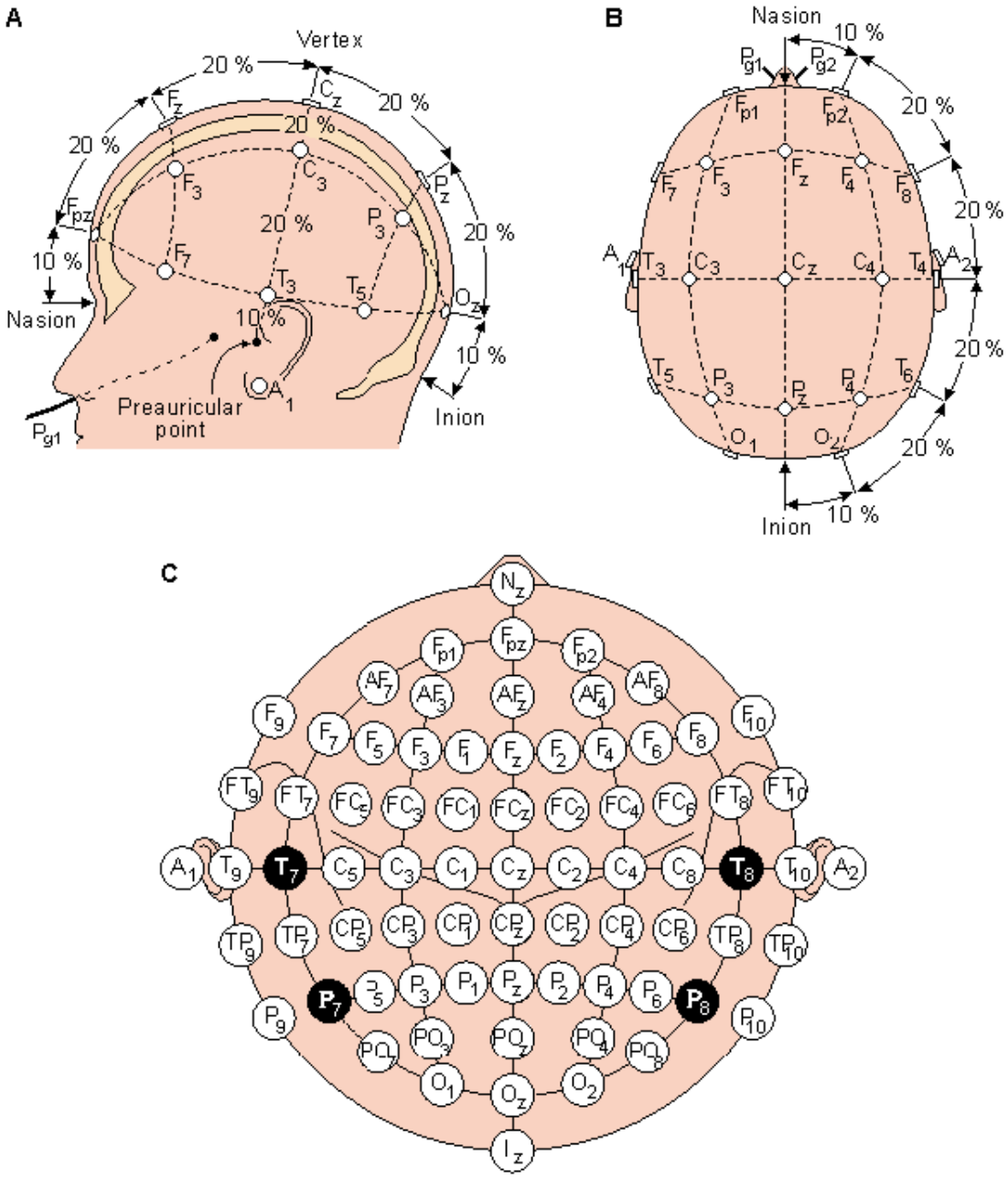
Alcoholism is a common neurological disorder caused by the mutual effect of genetic and environmental factors. It not only damages the brain system but also leads to cognitive and mobility impairments. The World Health Organization (WHO) reported that alcohol abuse is the third highest risk factor for causing diseases and results in around 2.5 million deaths each year. It would therefore be economically advantageous if we could figure out the affected regions of the brain without going into medical surgery. Electroencephalogram (EEG) is a very effective tool for studying the complex dynamics of brain activities. It can visualize complex brain activities as dynamic outputs. We will brush over the basics of EEG in the next section.

2 Introduction to enabling technology

Cognitive, perceptual, linguistic, emotional, and motor processes are fast. Most cognitive processes occur within tens to hundreds of milliseconds. Furthermore, cognitive events occur in a temporal sequence that may span hundreds of milliseconds to a few seconds.The first reason is that EEG capture cognitive dynamics in the time frame in which cognition occurs.

Another major reason for the use of EEG is that they directly measure neural activity. The voltage fluctuations that are measured by EEG (or magnetic field changes in the case of MEG) are direct reflections of biophysical phenomena at the level of populations of neurons.

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain. It is typically noninvasive, with the electrodes placed along the scalp. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain. In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time, as recorded from multiple electrodes placed on the scalp. Visual aid shown ahead illustrates The 10–20 system which is an internationally recognized method to describe and apply the location of scalp electrodes in the context of an EEG exam.


A
Vertex
20 %
Nasion
10 %
Preauricular point
Inion
B
Nasion
10 %
20 %
Inion
C

3 Objective

EEG experiment uses 64 electrodes to map ionic and electrical activity in different parts of the brain.

1. Our objective is to first isolate the brain area which is most affected among the alcoholics by appropriate comparison with their control/normal counterparts by measuring the electrode activity.

2. After isolating the affected area of the brain, we will try to identify if their exist any causal interactions taking place between two electrode readings, which could indicate the presence of an underlying component process. Causality here refers to the Granger Causality.

4 Dataset courtesy

Acknowledgments for this data goes to Henri Begleiter at the Neurodynamics Laboratory at the State University of New York Health Center at Brooklyn.

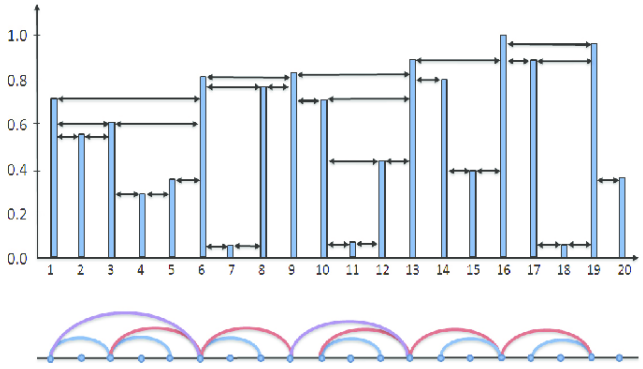
Link for the dataset:
`https://archive.ics.uci.edu/ml/datasets/eeg+database`

5 Isolation of the brain area which is most affected among the alcoholics

5.1 Horizontal Visibility Graph

The horizontal visibility graph (HVG) is a methodology that transforms a time series into a graph maintaining the inherent characteristics of the transformed time series. It basically considers each point in the time series, a node in the network, connected by the following consideration:
Let $\{x_i, i = 1, 2, ..., N\}$, be a time series of N data. Two nodes i and j in the graph are connected if it is possible to trace a horizontal line, in the time series, linking x_i and x_j not intersecting intermediate data height, fulfilling: $x_i, x_j > x_n$ for all $i < n < j$.

In the HVG, the nodes can see at least its nearest neighbors, incorporating in a natural way the time causality. One of the properties of the HVG is that it is not modified under rescaling of horizontal and vertical axes, as well as under horizontal and vertical translations. A visual example of HVG is shown next.



5.2 Kruskal-Wallis Test

This is a non-parametric method for testing whether samples originate from the same distribution.

Null hypothesis: Null hypothesis assumes that the samples (groups) are from identical populations.

Alternative hypothesis: Alternative hypothesis assumes that at least one of the samples (groups) comes from a different population than the others.

The distribution of the Kruskal-Wallis test statistic approximates a chi-square distribution, with k-1 degrees of freedom, hence hypothesis testing is done according to the significance level.

5.3 Procedure

The given dataset contains data for 10 alcoholic and 10 control subjects, with 10 runs per subject per paradigm. The frequency of sampling is 256 Hz from each of the 64 electrodes for 1 second.
However one of the electrode is used only for reference, hence not included in the analysis.
Since we have 63 electrodes, we have a hold of 630 time series each containing 256 readings in microvolts.

We have 10 HVGs for each of the 63 channels. For each channel, a significant number of features are extracted from its HVG. Some of these features include:
a.) Density
b.) K-core
c.) Entropy
d.) Degree distribution
e.) Assortativity coefficient.

Henceforth, for each channel, extracted features are evaluated with Kruskal-Wallis Test for alcoholic and control class attributed to the selected channel.

Channels with satisfy alternative hypothesis are selected (labelled as abnormal channels).
Top 3 abnormal channels came out to be O1, C3 and F5.
Support Vector Machine (SVM) algorithm is used to perform binary classification on the test set based on the features of the selected abnormal channels.

5.4 Result

The algorithm achieved an accuracy of 87.8% on the test data set based on three channels alone, which is a strong indication of abnormality.
Moreover, all three of the selected channels detect signals on the left lobe of the human brain, indicating that left part of alcoholic brains' are significantly affected.
Hence the channels O1, C3 and F5 are subjected to further analysis.

6 Granger causality analysis to detect causal relationship between abnormal channels

Since the question of *true causality* is deeply philosophical, and because of the post hoc ergo propter hoc fallacy of assuming that one thing preceding another can be used as a proof of causation, econometricians assert that the Granger test finds only "predictive causality".
For the same reason we will refer *Granger Causality* as *Granger Prediction*.

6.1 Granger Causality

Granger defined the causality relationship based on two principles:

- The cause happens prior to its effect.
- The cause has unique information about the future values of its effect.

We say that a variable X that evolves over time *Granger-causes* another evolving variable Y if predictions of the value of Y based on its own past values and on the past values of X are better than predictions of Y based only on its own past values.

6.2 Fitting an autoregressive model

To test for Granger Causality, the following two models are tested against each other:

$$X_t = \sum_{n=1}^{k} a_n X_{t-n} + e_{Xt}$$

and

$$X_t = \sum_{n=1}^{k} a_n X_{t-n} + \sum_{n=1}^{k} b_n Y_{t-n} + e_{XYt}$$

Granger prediction is the relative size of the errors from the two models — a univariate model in which current values of X are predicted only from past values of X , and a bivariate model in which current values of X are predicted both from past values of X and from past values of Y.

Consider that if the model is a good fit to the data, the errors will be small and will therefore have relatively little variability, whereas if the model is a poor fit to the data, the errors will be large and will therefore have relatively high variability. Thus, the variance of the errors from the bivariate autoregression model is compared with the variance of the errors from the univariate autoregression model.
This comparison is taken as the logarithm of the ratio of the error variances and is the mathematical definition of Granger prediction.

$$GrangerPrediction = log(\frac{var(e_x)}{var(e_{xy})})$$

where e_x comes from the first model and e_{xy} comes from the second model.

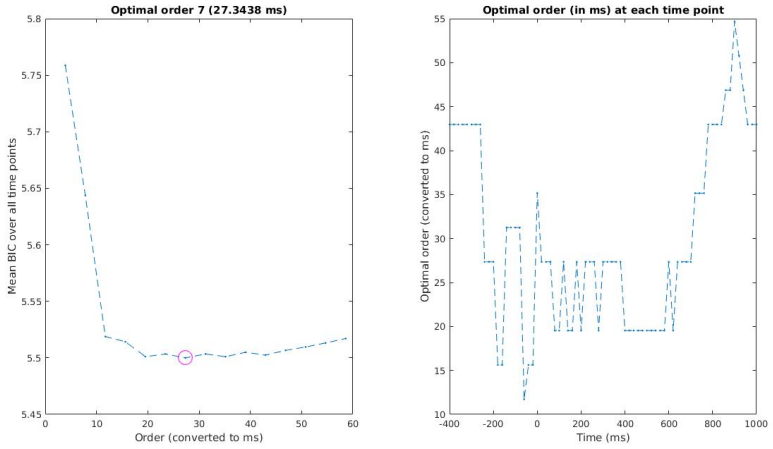
Since stationarity is an essential condition to make inferences about G-causality, Event Related Potential(ERPs) are subtracted from the daya to improve upon stationarity properties.

6.3 Model order selection by BIC

One of the most commonly used tests for appropriate model order selection is the Bayes information criterion (BIC).
To determine the optimal model order, the autoregression model is computed repeatedly with a range of orders, and the BIC is computed based on the error variances, the order, and the number of time points, according to the following equation:

$$BIC = log(det(E)) + (2^2 mlog(n))n^1$$

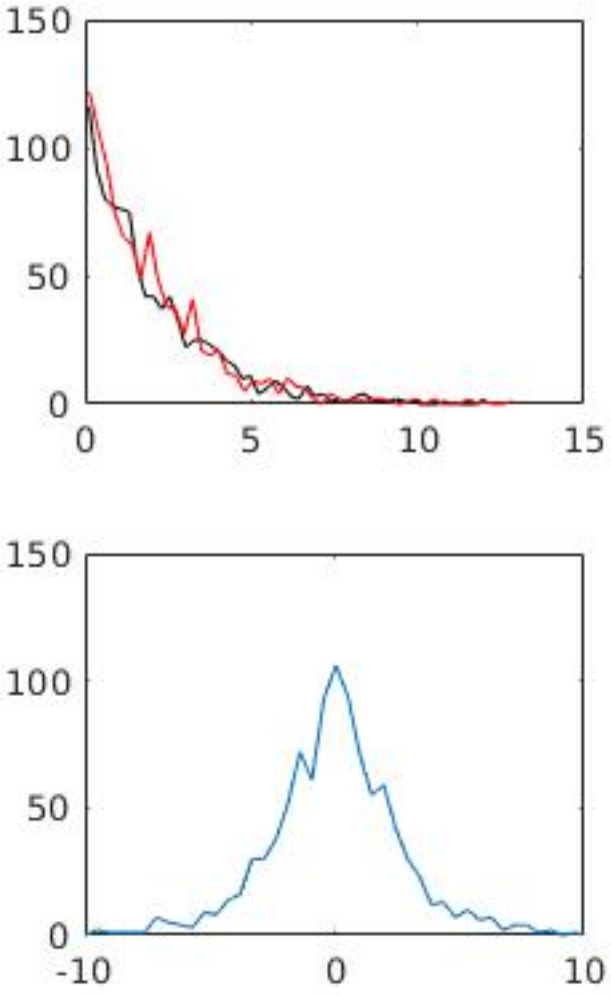
in which E is the error matrix that results from fitting the autoregression, *log* refers to the natural logarithm, *det* is the matrix determinant, m is the model order,and n is the number of time points. The 2 in the equation refers to the number of electrodes in the model.



Since the sampling rate is 256 Hz, 1 sample is recorded approximately 4 ms.
The BIC was evaluated on the EEG data which suggested that an order of seven (27 ms) is optimal for this dataset and given electrode pair because that point corresponded to the lowest BIC.
Hence, Granger Prediction is performed in successive intervals of 27 ms.

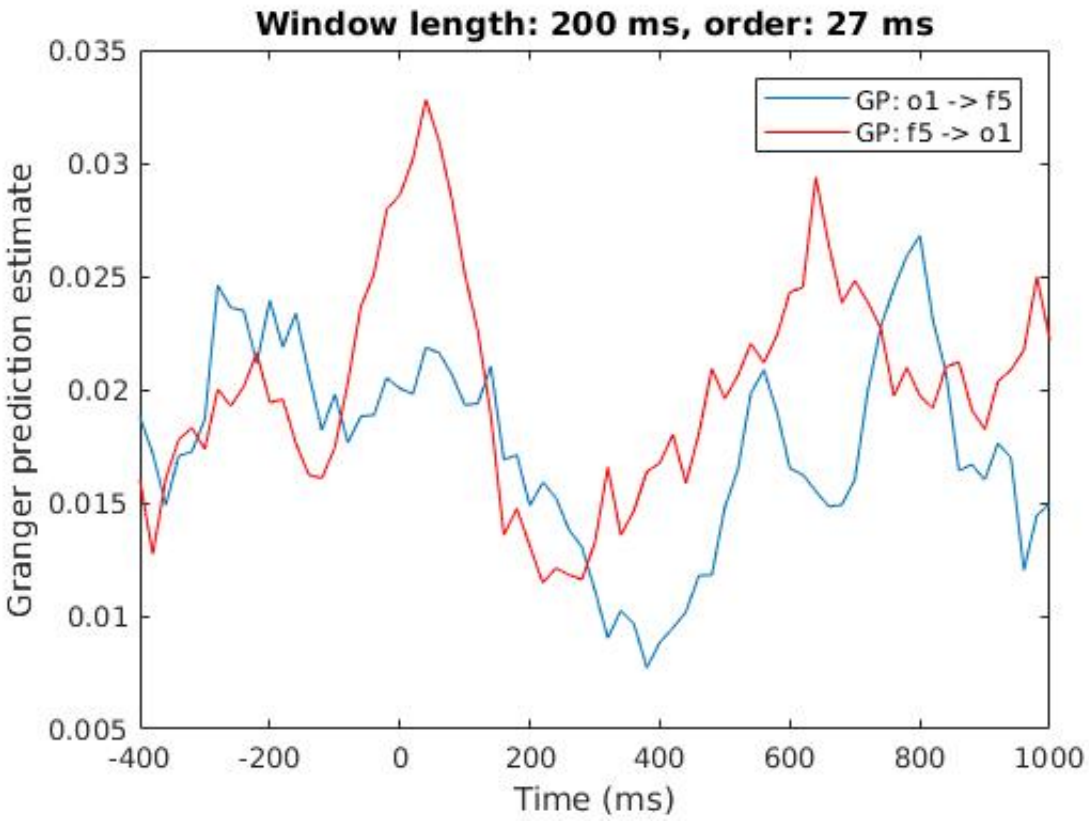
6.4 Statistical Tests

One elegant approach to statistical analysis of Granger prediction results is to focus on differences in Granger prediction results over time or across conditions or electrode pairs. Because Granger prediction results are chi-square distributed, condition difference values are approximately normally distributed.

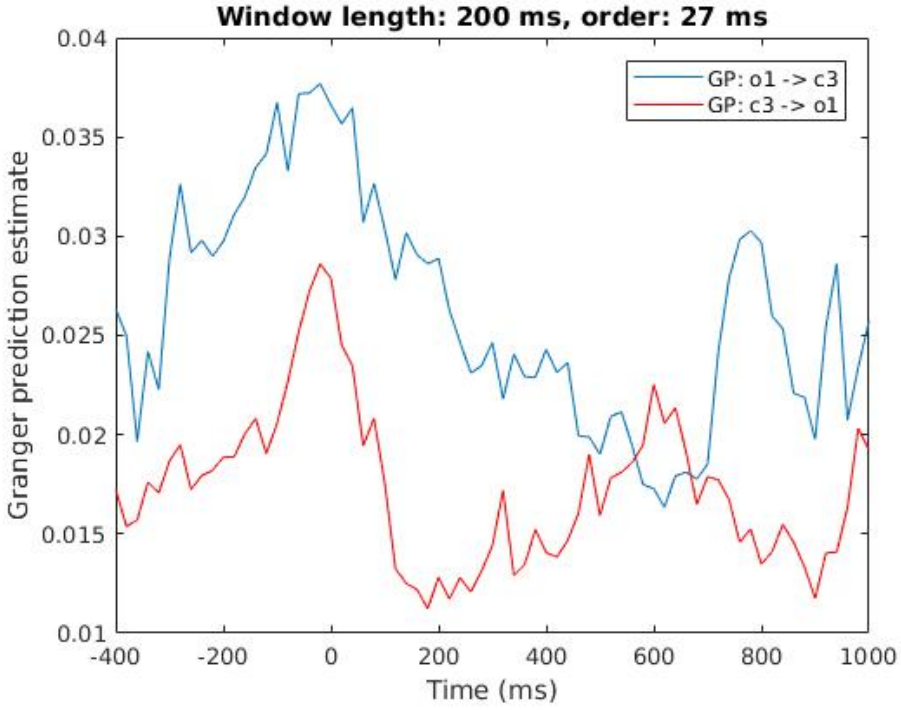


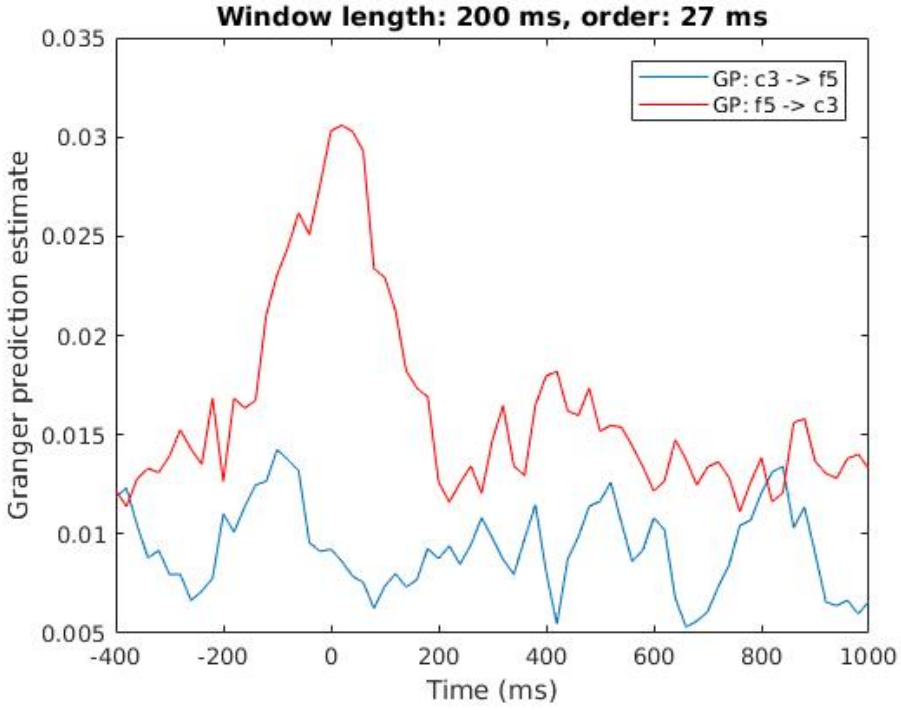
7 Observations

Following plots were observed for Granger Prediction between the three abnormal electrodes(O1, C3, F5).



Statistical tests need to be performed on the obtained plots to verify Granger Causality between the series.


Window length: 200 ms, order: 27 ms
GP: o1 -> c3
GP: c3 -> o1
Granger prediction estimate
Time (ms)
0.04
0.035
0.03
0.025
0.02
0.015
0.01
-400
-200
0
200
400
600
800
1000


Window length: 200 ms, order: 27 ms
GP: c3 -> f5
GP: f5 -> c3
Granger prediction estimate
Time (ms)
0.035
0.03
0.025
0.02
0.015
0.01
0.005
-400
-200
0
200
400
600
800
1000

8 Conclusion

The recent finding based on the study of brain tissue mitochondria found that alcohol significantly reduced amygdala reactivity to threat signals.
Conventional methods like Sample Entropy, correlation entropy were not able to give satisfactory results on EEG data for the correct identification of abnormalites in alcoholic patients. New nonlinear methods like Horizontal visibility graphs can be used to investigate the alcoholic EEG signals.

Once the correct abnormalities are diagnosed, causal analysis can then be performed for identification of the underlying component process responsible for perceptible causation.

9 References

- Analysis of alcoholic EEG signals based on horizontal visibility graph entropy
  Guohun Zhu, Yan Li, Peng (Paul) Wen, Shuaifang Wang
  Published online 2014 Sep 13. doi: 10.1007/s40708-014-0003-x

- `https://www.bci2000.org/mediawiki/index.php/User_Tutorial:EEG_Measurement_Setup`

- Analyzing Neural Time Series Data - Theory and Practice
  Mike X Cohen
  (2014, The MIT Press)

- `https://en.wikipedia.org/wiki/Granger_causality`