# 2102531 System Identification

# Learning Brain Network Differences from EEG Data
# Semester 1/2018

Parinthorn Manomaisaowapak
Jitkomut Songsiri
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

November 26, 2018

## Contents

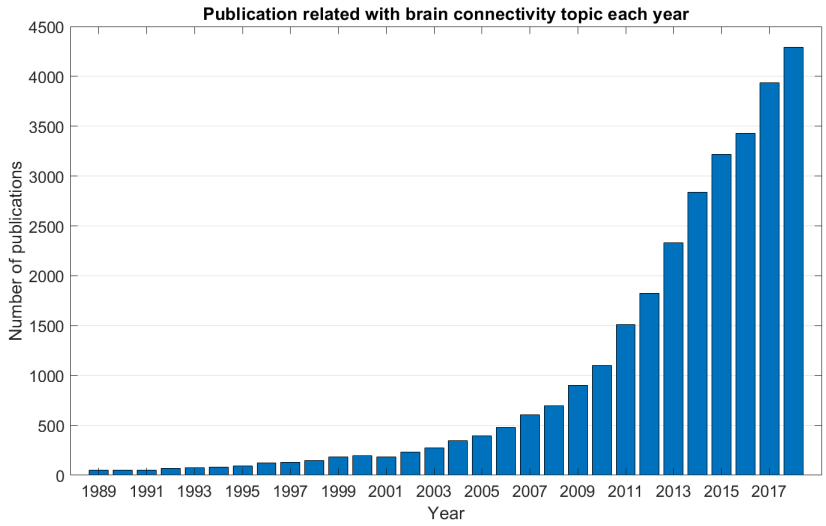


Figure 1: Publication on PubMed in topic related to brain connectivity each year, keywords: Brain connectivity OR functional brain connectivity OR effective brain connectivity.

# 1 Introduction

In past few decades, there are an enormous increment in publication per year in PubMed database on topics related with brain connectivity as in Figure 1. The PubMed database indicated that there are 4,290 publications related with brain connectivity in 2018. The rising demands of information extraction from biosignal such as fMRI (functional Magnetic Resonance Imaging), EEG (Electroencephalography), MEG(Magnetoenchephalography) to discover how brain regions interact to each other have drew us to attention in this topic.

The term brain connectivity refers to a pattern of links across brain regions that indicates causal interaction or statistical dependencies [12], [9]. There are three types of brain connectivity definition. The first is *Structural brain connectivity* which refers to the links that anatomically connected between brain regions. The second is *functional connectivity* which is the links between brain regions that can be anatomically unconnected regions defined by statistical dependencies. But statistical dependencies cannot be interpreted alone in general because statistical measures such as correlations cannot be interpreted as causality. The causal interaction is the others. Between brain regions the causal interaction are described in the last type of connectivity, the *effective brain connectivity*. The detection of brain network differences between normal groups and TBI (Traumatic Brain Injury) groups is our main interests because the classification of road accident patients, that they have undetected long-term brain injuries or not, are crucial. Our framework is based on two approaches, statistical framework and sparse estimation frame work. In statistical framework, the connectivity matrices are computed from each patient in each group individually and compute statistical measure such as average value of brain connectivity measures for representing the whole group to perform hypothesis test between group whether the statistical measure is difference or not. In sparse estimation framework in [11] exploits the sparsity patterns in brain connectivity matrix and used to regularize the estimation while controlling the difference of all subjects in the group resulting in representation of group brain connectivity. The difference can be directly identified from group brain connectivity.

In this project, we will use Multivariate Granger Causality toolbox (MVGC) [1] to compute Granger causality matrix as a measure of effective brain connectivity denoted as GC matrix. The GC matrix is based on vector autoregressive model because of model simplicity. In statistical framework, The average value of GC matrices in all trials of a group will be represent as a group GC matrix. The difference will be determined by Hotelling T-squared test. Overall process is described in 1 and the detail are in section 4

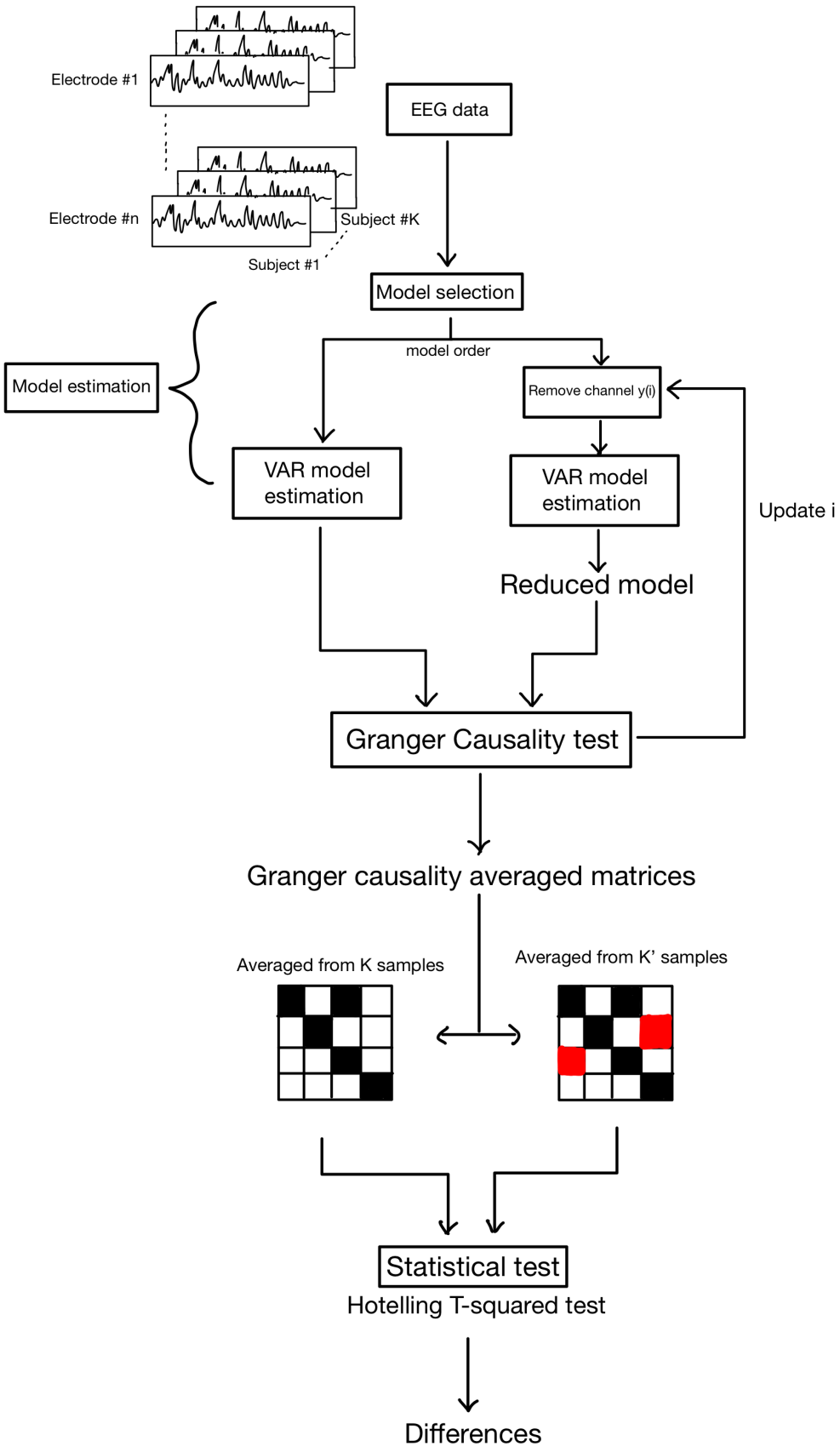


Figure 2: Group difference pathway in this project.

## 2 Problem statement

Given that brain time series data (EEG + fMRI) collected from two groups, Healthy group and TBI group. We aim to

1. Learn brain networks based on GC of two classes of brain signals.
2. Analyze brain network difference of two groups.

## 3 Background

### 3.1 Granger causality estimation

$$\mathcal{F}_{ij} = \log \frac{\det \Sigma_{ii}^R}{\det \Sigma_{ii}} \tag{1}$$

Granger causality is a concept that test if the past of one time series can help to predict another time series in sense of reducing the residual variance of the predicted time series. In this project, we will compute Granger causality on vector autoregressive model.

Vector autoregressive model order $p$ is defined as

$$y(t) = \sum_{k=1}^{p} A_k y(t-k) + e(t) \tag{2}$$

$y(t), e(t) \in \mathbb{R}^n, A_k \in \mathbb{R}^{n \times n}$ which can be expressed in state-space representation as

$$x(t+1) = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & I & 0 \end{bmatrix} x(t) + \begin{bmatrix} e(t+1) \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{3}$$

with $x(t) = \begin{bmatrix} y(t-1)^T & y(t-2)^T & \dots & y(t-p)^T \end{bmatrix}^T$

and the output equation is

$$y(t) = \begin{bmatrix} I & 0 & \dots & 0 & 0 \end{bmatrix} x(t) \tag{4}$$

VAR model parameter can be estimated by ordinary least square methods or solve via Yule-Walker equation [3].

Let's considered multivariate AR(1) process $y_i(t), y_j(t)$, note that both are vector. We want to investigate if $y_i(t)$ is depended only in its own past value, not from past of $y_j(t)$. The full fitted VAR model is

$$\hat{y}_i(t) = A_{ii} y_i(t-1) + A_{ij} y_j(t-1)$$
$$\hat{y}_j(t) = A_{ji} y_i(t-1) + A_{jj} y_j(t-1)$$

which can be expressed as

$$\hat{y}(t) = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj} \end{bmatrix} y(t-1)$$

noted that $y(t) = \begin{bmatrix} y_i(t)^T & y_j(t)^T \end{bmatrix}^T$

And the reduced model is

$$\hat{y}^R(t) = \begin{bmatrix} A_{ii}^R & 0 \\ A_{ji}^R & A_{jj}^R \end{bmatrix} y^R(t-1)$$

Granger causality can be tested by a ratio of the generalized variance [2] of reduced model and full model.

This measure is, in general, defined by equation (1) which is multivariate version of Granger causality with $\Sigma_{ii}^R$ as the residual covariance matrix of reduced model and $\Sigma_{ii}$ as residual covariance matrix of full model. In this case, both $y_i, y_j$ are vector of time series, and it has physical meaning as a test of multiple time series to another multiple time series.

Intuitively, if past of $y_j$ help to predict $y_i$, the generalized variance of full model, which included $y_j$, will less than the one that does not include, which is the reduced model, then the value of log-ratio will be nonzero and the value can be interpreted as connectivity value. Conversely, if past of $y_j$ does not help to predict $y_i$, the generalized variance of two model should be equal and leads log-ratio (1) to zero.

One of the effective brain connectivity measure is Granger causality which is the test for directional causal inference between two multivariate variables [1] which can be computed from Vector autoregressive model (VAR).

Granger causality inference can be achieved by equation (1) with generalized residual covariance $\Sigma_{ii}^R = \mathbb{E}((y_i - \hat{y_i}^R)(y_i - \hat{y_i}^R)^T)$, $\Sigma_{ii} = \mathbb{E}((y_i - \hat{y_i})(y_i - \hat{y_i})^T)$, which can be estimated by using unbiased sample residual covariance.

In this case if the log ratio (1) is zero, the generalized variance of $y_i$ is unchanged which means the past of $y_j$ have no connection with current $y_i$.

### 3.2 Hotelling's T-squared test

The effective brain connectivity matrix is usually sparse due to the sparsity of connection in structural brain [12] . In numerical computation, there will be no real zeros which emphasizes the reason to perform significant test whether the value inside connectivity matrix is actually zero. However, this project aims to detect a difference between groups, the significant test should not be performed in order to preserve distribution of GC matrices.

The difference is defined as the difference in the mean of GC matrix of trials drawn from first group and second group.

This equality will be tested by vectorizing the average value of GC matrices with dimension $n \times n$ into vector mean of a group yielding $\bar{X}_1, \bar{X}_2$ which, by central limit theorem, is $n^2 - n$-variate normal distributed, the $n$ subtraction came from causal inference in the same channel.

two samples Hotelling's T-squared is used to test whether the vector mean of two samples are equal. The test statistics is defined as

$$T^2 = (\bar{X}_1 - \bar{X}_2)^T (\frac{s_1}{n_1} + \frac{s_2}{n_2})^{-1} (\bar{X}_1 - \bar{X}_2) \tag{5}$$

where $\bar{X}_i$ is sample vector mean of $X_i$, $s_i$ is unbiased sample covariance matrix of $X_i$ respectively. The test required that both $X_1, X_2$ are drawn from Normal distribution with common covariance matrix and test statistics $T^2$ will have distribution as

$$T^2 \sim \frac{pv}{v - p + 1} F_{p, v-p+1} \tag{6}$$

where $p$ is dimension of vector, $v = n_x + n_y - 2$ is degree of freedom [6].

In practical uses, The two samples are mostly drawn from normal distribution with unequal covariance matrix, this problem is called multivariate Behrens-Fisher problem. There are many solutions, one of them is to estimate distribution of $T^2$ by modifying degree of freedom $v$ in 6. In this project, the degree of freedom is used as in [7] which proved that

$$v = \frac{p + p^2}{A_1 + A_2}$$

where

$$A_i = \frac{\mathrm{tr}[(\tilde{s}_i s_p^{-1})^2] + [\mathrm{tr}(\tilde{s}_i s_p^{-1})]^2}{n_i}; i = 1, 2$$

$$\tilde{s_i} = \frac{s_i}{n_i}$$

$$s_p = \tilde{s_1} + \tilde{s_2}$$

Hotelling T-squared test can be explained intuitively as the multivariate version of student's t statistics that used to compare mean in scalar version. But in multivariate sense, the vector mean cannot be compared element-wise because there are many components which are not necessarily independent to each other such as normal distribution with non-diagonal covariance matrix. The $T^2$ brings mean vector into scalar representation as quadratic loss function. If value of $T^2$ is low, the vector mean of two samples are more likely to be equal.

### 3.3 Group differences test

In this project, the mean of Granger causality based brain connectivity matrices of all trials in a class is used to represent brain connectivity of the class and The hypothesis is if brain connectivity of two groups are different, their population mean of brain connectivity matrix will not equal. Hence, the group differences can be determined by comparing element-wise mean of all trials between TBI and healthy group. The element-wise mean of all trials will be asymptotically normal distributed by the virtue of Central Limit Theorem. The Hotelling T-squared test is used to test hypothesis that the mean of two groups are equal under normality assumption.

## 4 Methodology

There are 4 steps to compute group difference test in statistical framework, which are

1. **data preparation**

   The data preparation method will described how data was collected.

2. **Model estimation**

   This section will involve model order selection and VAR coefficient estimation which includes 2 methods of estimation, Ordinary least square and solve via Yule-walker equation [3].

3. **GC matrix computation** This section will explain how the Granger causality matrix was estimated for all available data

4. **Group difference test** The statistical test in 6 will be performed in this step. This step will required asymptotic normality of sample mean, the testing samples must be large enough.

In step 2-3 will used MVGC toolbox to implement.

### 4.1 Data preparation

By the reason that the statistical group difference test uses central limit theorem to assume normality of the means of Granger causal inference in element-wise of all subjects, implies that the number of sample must be maximize. One way is to augment the data by splitting into multiple trials of each subject but in VAR model parameter estimation has number of parameter to be estimated as $NVAR = n^2p$, $n$ is number of EEG channel, $p$ is time lag, we will use $p = 3$. So it will be $63 \times 63 \times 3 \approx 12,000$ which need atleast $12,000$ data points but we will use rule of thumb, data points will be 10 times more than the number of parameters. Hence, it will be 2 trials per subject.

However, the real data set contained the channel that are highly correlated. After filtered out by remove the channel that have correlation with another channel more than 0.9. There are 22 remaining channel causing the number of parameters to be estimated reduced to approximately $1,000$, hence, the number of trials can increase to 29 sample per subject.

### 4.2 Model estimation

The estimation is based on assumption that the EEG time series are wide-sense stationary, the dynamic matrix in (3) must be stable.

### 4.2.1 Model order selection

Model order is selected by AIC, BIC value which, in general, described by equation (7)

$$\text{AIC} = -2\mathcal{L} + 2k \tag{7}$$

$$\text{BIC} = -2\mathcal{L} + k \log N \tag{8}$$

where $\mathcal{L}$ is log-likelihood function of VAR(p) process, $k$ is number of parameters to be estimated, in this case $k = n^2p$, and $N$ is number of all observation.

In the MVGC's source code, AIC implementation is based on [8] and BIC is the same as (8). Maximum log-likelihood function is used.

$$\text{AIC} = -2\mathcal{L} + 2k\frac{N}{N-k-1} \tag{9}$$

$$\mathcal{L} = -\frac{N}{2}\log\det\hat{\Sigma} \tag{10}$$

where $\hat{\Sigma} = \frac{1}{N-1}ee^T$ is unbiased estimator of residual covariance matrix, $e = y - \hat{y}$.

In this project, the amount of data are not sufficient to select higher order model, due to the trade-off between parameter estimation and the usage of central limit theorem in statistical test. So, the order candidates are $p = 1, 2, 3$.

### 4.2.2 VAR coefficients estimation

There are two main methods to estimate the coefficients,

1. **Ordinary least square**

   Ordinary least square is a solution of the overdetermine system. In this case, the linear system is

   $$\begin{bmatrix} y(p+1) & y(p+2) & \dots & y(N) \end{bmatrix} = \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix} \begin{bmatrix} y(p) & y(p+1) & \dots & y(N-1) \\ \vdots & \vdots & \dots & \vdots \\ y(2) & y(3) & \dots & y(N-p+1) \\ y(1) & y(2) & \dots & y(N-p) \end{bmatrix} \tag{11}$$

   This is in the form $Y = \beta X$, $Y \in \mathbb{R}^{n\times(N-p)}$, $X \in \mathbb{R}^{pn\times(N-p)}$, $A_i \in \mathbb{R}^{n\times n}$ where $n$ is the number of channel and $N$ is number of timepoints, $p$ is model order.

   Then the least square optimization formulation is

   $$\underset{\beta}{\text{minimize}} \quad ||Y - \beta X||_F^2 \tag{12}$$

   The least square solution $\hat{\beta}$ can be solved analytically by solving the normal equation.

   $$\hat{\beta}(XX^T) = YX^T \tag{13}$$

   In the MVGC toolbox, Least square method was implemented by MATLAB function **mrdivide** that simply computes least square solution via QR factorization. In this case, the regressor matrix in (13) is mostly rank deficient which caused by highly correlated EEG channel, hence, there will be infinitely many exact solution. However, the highly correlated channel must be excluded.

2. **Solve via Yule-Walker equation**

   $$\begin{bmatrix} \Gamma(1) & \Gamma(2) & \dots & \Gamma(p) \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \dots & A_p \end{bmatrix} \begin{bmatrix} \Gamma(0) & \Gamma(-1) & \dots & \Gamma(-p+1) \\ \Gamma(1) & \Gamma(0) & \dots & \Gamma(-p+2) \\ \vdots & \dots & \ddots & \vdots \\ \Gamma(p-1) & \Gamma(p-2) & \dots & \Gamma(0) \end{bmatrix} \tag{14}$$

Yule walker equation described in (14) is a system of linear equation that came from taking autocovariance of equation(2) with multiple lags. where $\Gamma(k)$ is autocovariance matrix that can be estimated by its unbiased sample autocovariance. If the datapoints are large enough, by the law of large number, the $XX^T$ in equation 13 will converge to autocovariance matrix in equation 14. The autocovariance matrix in (14) is in Toeplitz form that can be solve efficiently by LWR (Levinson Wiggins Robinson) algorithm which has been proven that this algorithm will yield stable VAR coefficients [14].

### 4.3 GC matrix computation

Granger causality is implemented by estimating full model and reduced model from the data set directly instead of compute via autocovariance sequence that recommended by MVGC toolbox [1]. The reason of this recommendation is to increase computation accuracy in frequency domain Granger causality calculation but in this project, only time domain Granger causality is used. The significant test of Granger causality measure is to perform hypothesis test. The null hypothesis is that the entry in Granger causality matrix is actually zero with asymptotic distribution as chi-squared distribution [1], [5] $(N-p)\mathcal{F}_{ij} \sim \chi^2_{p(n_i+n_j)}$ where $N, p, n_i, n_j$ denotes sample size, lags, dimension of $y_i, y_j$ respectively.

The null hypothesis will be rejected if p-value of the GC measure is below 0.05 which p-value is probability that the null hypothesis is true. Hence, the GC measure will be set to zero if the p-value more than 0.05.

But in statistical framework, the significant test will not be performed due to the hypothesis test on group difference test. Setting zero in GC matrices may cause assumed test distribution to change.

### 4.4 Group difference test

The group differences of brain connectivity is determined by comparing element-wise GC matrices mean of all trials between healthy group and TBI group. the mean of GC matrices has to be vectorized into vector mean in order to use the Hotelling T-squared test that compare vector mean of two samples, that assumed to be drawn from Gaussian distribution with common covariance matrix.

## 5 Data description

The EEG datasets are achieved from USM (University Sains Malaysia). There are two groups of data, TBI vs. Healthy. Each groups has 7 subjects. each subjects performed N-back test, emotion and real-time task. the EEG data are measured before and after those tasks with 2 different condition, eyes-open and eyes-closed. For example, data with label **after REC** is the EEG data that measured After given tasks during Resting state with Eyes-Closed. The Figure 3 is conclusion of data structure.

### 5.1 Electrode placement system

Electrode placement systems are standard methods to measure EEG signal from scalp. The example are 10-20 system and 10-10 system. The number denotes distance in % from front to back, in this case there are 10% and 20% distance between electrodes. The 10-10 system has more spatial resolution [10].

### 5.2 Measurement

The data was measured by 64 channel 10-10 EEG electrode placement system with sampling rate 1000 Hz. Only channel 32 (EOG channel) has to be removed before analysis because it is not connected. All channel's EEG signal was a voltage difference between the the EEG electrode and

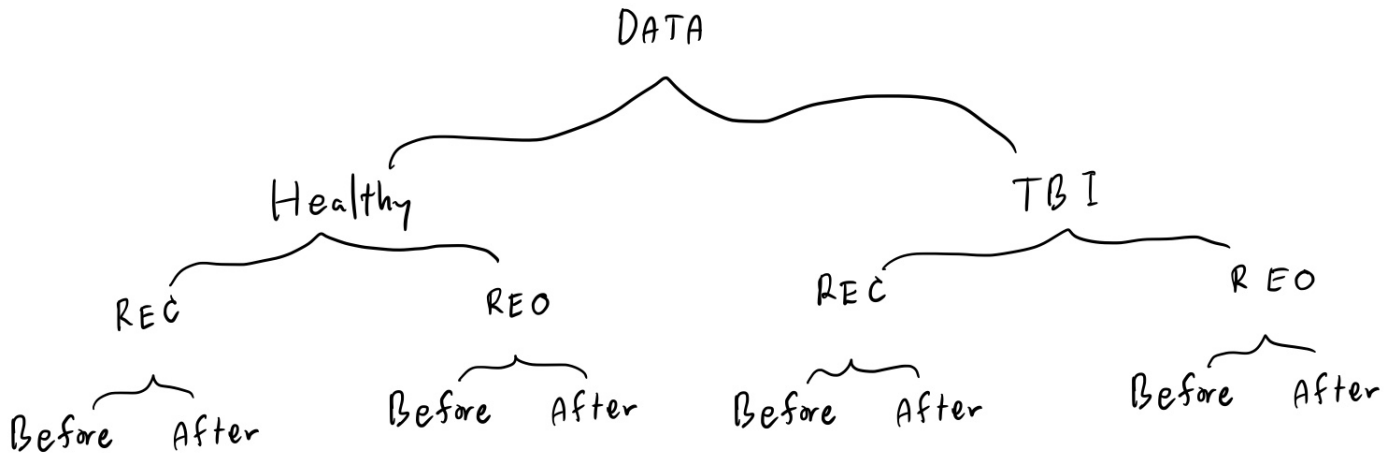


Figure 3: Data structure

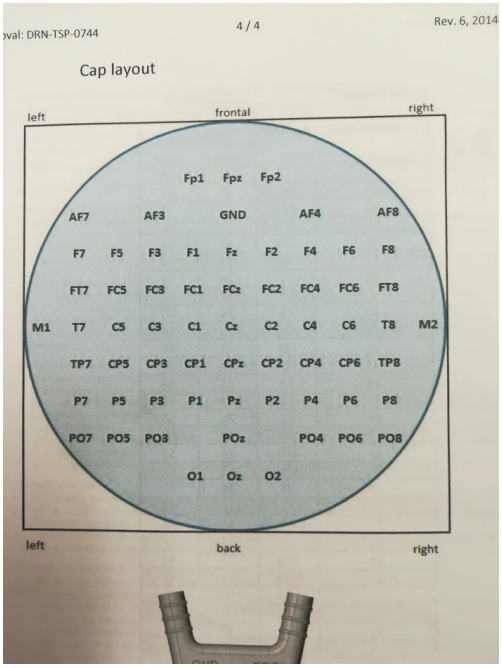


Figure 4: actual 10-10 system electrode placement system that used to measure EEG signal.

reference electrode, which is channel CPz. the Ground (GND) channel described in the datasheet in figure 5.1 is not presented in data file.

However, the real data set contained the channel that are highly correlated. After filtered out by remove the channel that have correlation with another channel more than 0.9. There are 22 remaining channel causing the number of parameters to be estimated reduced to approximately $1,000$, hence, the number of trials can increase to 29 sample per subject.

### 5.3 Data Problem

The abnormality of data are investigated by the plot comparison of fitting The hypothesis is the data that contains abnormality such as spikes in the signal should be detected when fitting the model by MSE value as described in figure 6, 7 which the MSE seems to be very high at those trials. The example of signal that contains spikes is from trial 6, which denoted as **data6** in the Figure 7, is Figure 8

Another problem on data is the high linearly correlated channel which is the cause of rank deficient least square and non unique solution of VAR coefficient which should not be acceptable.

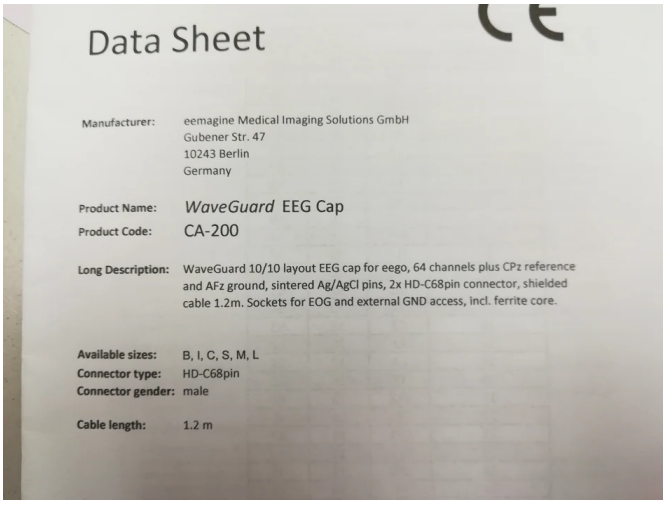


Figure 5: Actual 10-10 system electrode placement system that used to measure EEG signal.

One way to solve the problem is to remove the highly channel correlated channel which measured by sample correlation.

By removing highly correlated channels, another problem is every subjects has different highly correlated channels position and amount. The solution may be remove the channel by extracting each region of EEG electrode as one channel.

# 6 Experiment

## 6.1 Model order selection

Model order is selected by lowest BIC score, because it will generally return simpler model than AIC. In the range of $p = 1, 2, 3$ the BIC value is minimum when p=3 in all data.

## 6.2 VAR coefficient estimation

The goodness of fit is compared by mean square error. The result is in figure 10 which both methods seem to be equally fitted. The spike should be an error in estimation. But this is the result from minimum norm solution, the effect of highly correlated channel data needs further investigation.

### 6.2.1 Model estimation problem

In VAR parameters estimation, the AR coefficient matrix seem to be unstable, residual covariance matrix is singular. The stability problem is solved by using LWR algorithm from which has been proved that the VAR process from LWR algorithm always stable [14].

### 6.2.2 Solution of correlated data channel & Rank deficiency in regressor matrix

The channels of an example data were selected by correlation criteria which is to remove all channel that correlated with some another channels with correlation greater than 0.9. The number of channel is drop from 63 to 34 channels after this procedure.

The model order was reduced to 2 because the regressor matrix with higher order needs more data points to be full rank, even the channels are not highly correlated.

## 6.3 Granger causality matrix computation

The GC measure is computed by fitting full and reduce model directly but in this case, the univariate Granger causality is used because we want to indicate causal interaction between "one" channel to

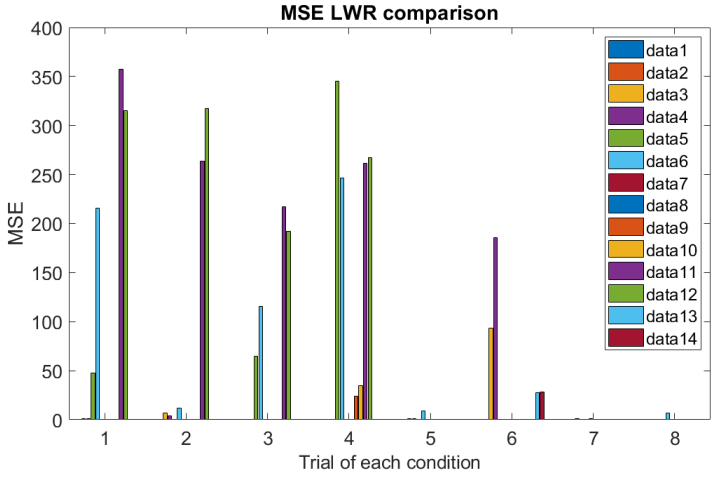


Figure 6: MSE of LWR algorithm in all trials.

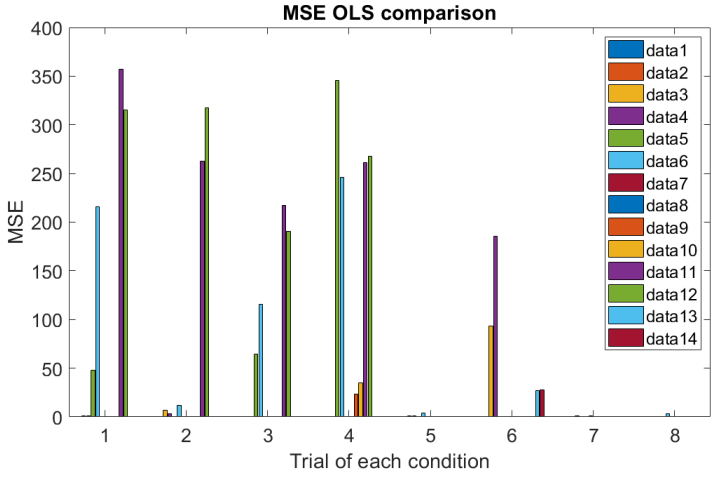


Figure 7: MSE of OLS algorithm in all trials.

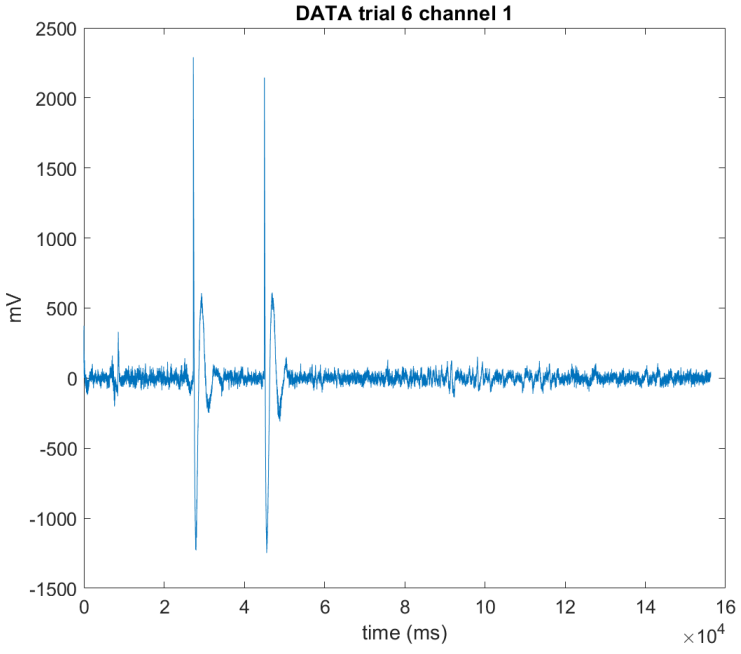


Figure 8: Example plot of abnormal data

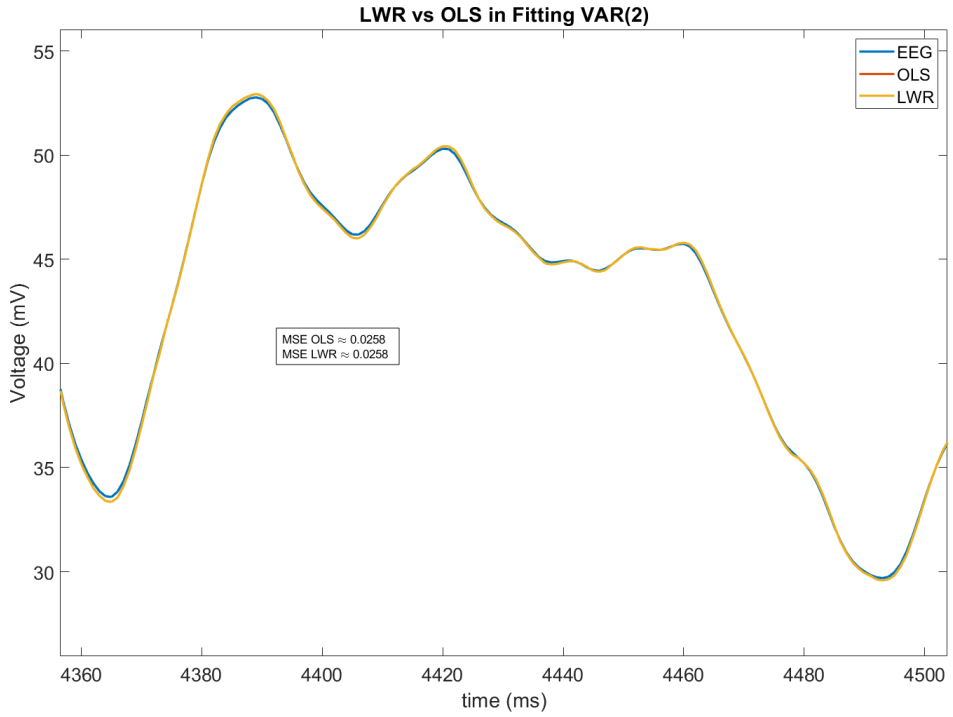


Figure 9: fitting comparison of OLS method & LWR algorithm that implemented in MVGC toolbox

another. So, the equation (1) is reduced into scalar equation. The test was done by performing Granger causality test in each pair of channels. And the significant test used chi-squared test to compute p-value for hypothesis test to test if the causal is actually zero.

Our result is in the Figure 11

## 6.4 Group difference test

The test was performed in 2 condition

1. Crossing test between TBI and Healthy with number of GC matrices sample of 29. This can measure number of true positive and false negative of the test.

2. Test among the same group, between TBI and TBI or Healthy and Healthy with number of GC matrices sample of 29. This can measure number of true negative and false positive of the test.

### 6.4.1 Experimental results

The results of group differences test are in Figure 12. The true positive rate is 81.73 %, overall true negative rate is 48.19 %. True negative rate in TBI vs TBI is 66.67 % but in Healthy vs Healthy is 0 %.

The test in the same class can detect the differences 3 from 4 class.

True negative rate is decreased mainly because of healthy vs healthy, the reason might be

1. The different class in healthy group have different brain connectivity.

2. The EEG signal in healthy group contained artifacts such as spikes.

3. The number of trials are not enough for statistical testing because it is the comparison of $n^2$ dimensional vector, where $n$ is number of channel. In this scenario, the dimension of vectorized GC matrices is $22 \times 22 - 22 = 462$, subtracted NaNs. And the samples are vary

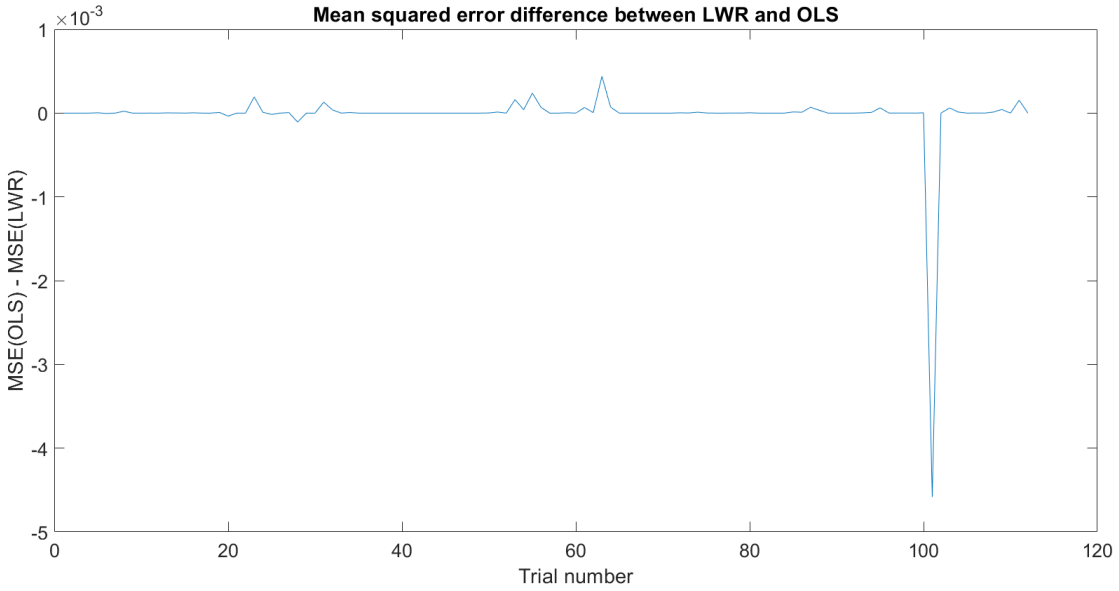


Figure 10: Difference in mean squared error of model fitted by OLS method and LWR algorithm

in ranges of 29-145 which may be not enough. In [7] denoted that the sample size should be atleast 4 times more than vector dimension.

# 7 Conclusions

The Hotelling T squared-test is able to detect the difference between TBI and healthy group but number of trials and data abnormality may cause error in hypothesis testing.

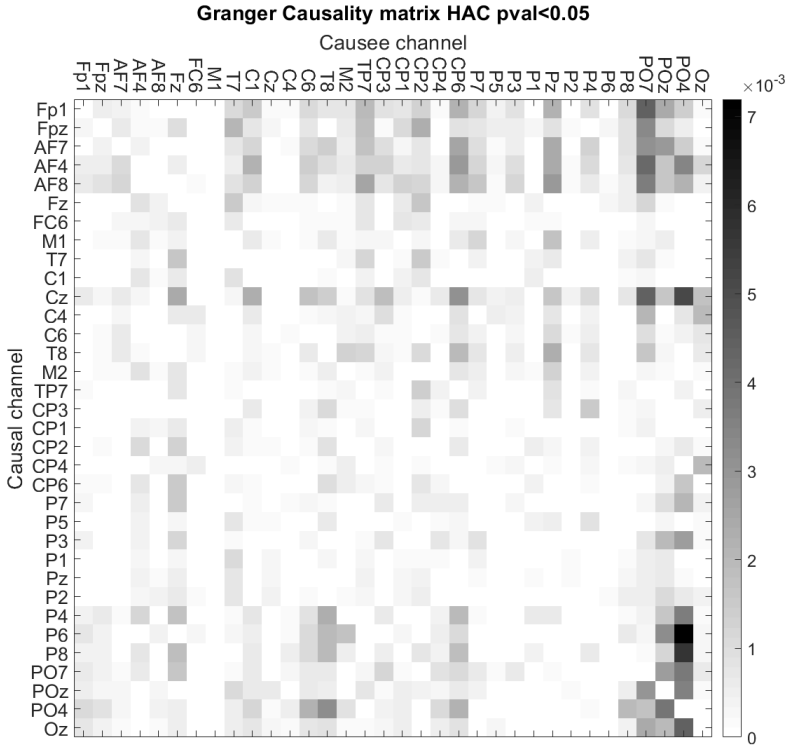


Figure 11: Example of Granger causality matrix from EEG data with highly correlated channels removed.

| | n=104 | n=23 | n=60 |
|---|---|---|---|
| | Healthy vs TBI | Healthy vs Healthy | TBI vs TBI |
| True positive | 85 | - | - |
| True negative | - | 0 | 40 |
| False positive | - | 23 | 20 |
| False negative | 19 | - | - |

Figure 12: Group differences results.

# 8 Matlab Code

## 8.1 Granger causality matrices computation

```
clear
clc
clf
close all

%% fetch data
n = input('Enter number of files to be imported\n');
fdir = dir('C:\Users\ritth\Desktop\SELECTED');
addpath('C:\Users\ritth\Desktop\SELECTED')
load('C:\Users\ritth\Dropbox\SeniorProject\Matlabfile\chan_data_reorder.mat') %newchanindex
load('C:\Users\ritth\Dropbox\SeniorProject\Matlabfile\chorder.mat')
load('C:\Users\ritth\Dropbox\SeniorProject\Matlabfile\index_equalize.mat')
%% Main program
for i=3:n+2 % accessing each class
%     i=7;
    zz =1;
    ri = i-2;
    file_name = fdir(i).name;
    disp(file_name)
    s = load(file_name);
    fname = fieldnames(s);
    fname = fname{1};
    disp(file_name(1:end-4))
    d = s.(fname); % use dynamic field, d is accessed file's structure data
    dim = length(d);
    F = zeros(22,22,29*dim); % GC matrix size 22x22 with 29*d trials
    for j =1:dim %Extracting trial from file
        tmpindex = newchanindex;
        data = d(j).data;
        data = data(order,:,:); %Reorder channel
        data(41,:,:) = []; %remove EOG
        tmpindex(41,:) = [];
        data(ind,:,:) = []; %remove high corr equally in all file
        tmpindex(ind,:) = [];
        for k=1:29 % iterate through trials in each class
            disp(k)
            data_k = data(:,:,k);
            tic
            F(:,:,zz) = GC(data_k); % Compute GC matrices
            zz=zz+1;
            toc
        end
    end

    if strcmp(fname,'HAC')
        save('F_HAC.mat','F')
```

```
47      elseif strcmp(fname,'HAO')
48          save('F_HAO.mat','F')
49      elseif strcmp(fname,'HBC')
50          save('F_HBC.mat','F')
51      elseif strcmp(fname,'HBO')
52          save('F_HBO.mat','F')
53      elseif strcmp(fname,'PAC')
54          save('F_PAC.mat','F')
55      elseif strcmp(fname,'PAO')
56          save('F_PAO.mat','F')
57      elseif strcmp(fname,'PBC')
58          save('F_PBC.mat','F')
59      elseif strcmp(fname,'PBO')
60          save('F_PBO.mat','F')
61      end
62  end
```

### 8.2 Difference test computation

```
1  clear
2  clc
3  %% sieve GC matrices from each condition and channel index
4  load('F_HAC.mat')
5  F_HAC = F;
6  load('F_HAO.mat')
7  F_HAO = F;
8  load('F_HBC.mat')
9  F_HBC = F;
10 load('F_HBO.mat')
11 F_HBO = F;
12 load('F_PAC.mat')
13 F_PAC = F;
14 load('F_PAO.mat')
15 F_PAO = F;
16 load('F_PBC.mat')
17 F_PBC = F;
18 load('F_PBO.mat')
19 F_PBO =F;
20 load('F_channel.mat')
21
22 %% Data detail
23 %HAC : 58 trials,2 subjects
24 %HAO : 29 trials,1 subject
25 %HBC : 58 trials,2 subjects
26 %HBO : 87 trials,3 subjects
27 %PAC : 58 trials,2 subjects
28 %PAO : 145 trials,5 subjects
29 %PBC : 116 trials,4 subjects
30 %PBO : 58 trials, 2 subjects
31 PERM_H(1).data = F_HAC;
32 PERM_H(2).data = F_HAO;
33 PERM_H(3).data = F_HBC;
34 PERM_H(4).data = F_HBO;
35
```

```
PERM_P(1).data = F_PAC;
PERM_P(2).data  = F_PAO;
PERM_P(3).data  = F_PBC;
PERM_P(4).data  = F_PBO;

%% Crossing Healthy vs TBI
%Stat test AC
[pval_AC,T2_AC] =Tsq(F_HAC,F_PAC);
%Stat test AO
[pval_AO,T2_AO] =Tsq(F_HAO,F_PAO);
%Stat test BC
[pval_BC,T2_BC] =Tsq(F_HBC,F_PBC);
%Stat test BO
[pval_BO,T2_BO] =Tsq(F_HBO,F_PBO);
%% Non-crossing Healthy
[pval_HACAO,T2_HACAO] =Tsq(F_HAC,F_HAO);
[pval_HACBC,T2_HACBC] =Tsq(F_HAC,F_HBC);
[pval_HACBO,T2_HACBO] =Tsq(F_HAC,F_HBO);
[pval_HAOBC,T2_HAOBC] =Tsq(F_HAO,F_HBC);
[pval_HAOBO,T2_HAOBO] =Tsq(F_HAO,F_HBO);
[pval_HBCBO,T2_HBCBO] =Tsq(F_HBC,F_HBO);
%% Non-crossing TBI
[pval_PACAO,T2_PACAO] =Tsq(F_PAC,F_PAO);
[pval_PACBC,T2_PACBC] =Tsq(F_PAC,F_PBC);
[pval_PACBO,T2_PACBO] =Tsq(F_PAC,F_PBO);
[pval_PAOBC,T2_PAOBC] =Tsq(F_PAO,F_PBC);
[pval_PAOBO,T2_PAOBO] =Tsq(F_PAO,F_PBO);
[pval_PBCBO,T2_PBCBO] =Tsq(F_PBC,F_PBO);
%% Permutation test
Tpos=0;
Fneg=0;
FposH=0;
TnegH=0;
FposP=0;
TnegP=0;
for i=1:4
    for j=1:4
        [Tpos_tmp,Tneg_tmp,Fpos_tmp,Fneg_tmp] = permute_test(PERM_H(i).
            data,PERM_P(j).data,1);
        Tpos =Tpos+Tpos_tmp; % tell that there is a difference which is
            true
        Fneg =Fneg+Fneg_tmp; %Tell that no difference which is false
    end
end
for i=1:4
    for j=i+1:4
        [Tpos_tmp,Tneg_tmp,Fpos_tmp,Fneg_tmp] = permute_test(PERM_H(i).
            data,PERM_H(j).data,0);
        TnegH =TnegH+Tneg_tmp;
        FposH =FposH+Fpos_tmp;
    end
end
for i=1:4
```

```
86      for j=i+1:4
87          [Tpos_tmp,Tneg_tmp,Fpos_tmp,Fneg_tmp] = permute_test(PERM_P(i).
                data,PERM_P(j).data,0);
88          TnegP =TnegP+Tneg_tmp;
89          FposP =FposP+Fpos_tmp;
90      end
91  end
92  Fpos =FposH+FposP;
93  Tneg = TnegH+TnegP;
94  TPR = Tpos/(Tpos+Fneg);
95  FPR = Fpos/(Fpos+Tneg);
96  PPV = Tpos/(Tpos+Fpos);
```

### 8.3 Function definition

```
1   function F = GC(X)
2   [m,~] = size(X);
3   F = zeros(m,m);
4   for i=1:m
5       for j=1:m
6           if i==j
7               F(i,j) = NaN;
8           else
9               F(i,j) = GCCA_tsdata_to_mvgc(X,i,j,2,'OLS');
10          end
11      end
12  end
13  end
```

```
1   function [pval,T2] = Tsq(X,Y)
2   [nx,ny,nz] = size(X);
3   [mx,my,mz] = size(Y);
4   Xvec = reshape(X,nx*ny,nz);
5   Yvec = reshape(Y,mx*my,mz);
6   Xvec(~any(Xvec,2),:) = [];
7   Yvec(~any(Yvec,2),:) = [];
8   tmpX = Xvec;
9   tmpY = Yvec;
10  muX = mean(tmpX,2);
11  muY = mean(tmpY,2);
12  sx = cov(tmpX');
13  sy = cov(tmpY');
14  sp = (sx/nz+sy/mz +0.00001*eye(nx*ny-nx)); %regularized covariance
        matrix 0.00001
15  p = length(muX);
16
17  %% Modified Nel and Van der Merwe test
18  A = trace(((sx/nz)*(sp\eye(nx*ny-nx)))^2);
19  B = (trace((sx/nz)*(sp\eye(nx*ny-nx))))^2;
20  C = trace(((sy/mz)*(sp\eye(nx*ny-nx)))^2);
21  D = (trace((sy/mz)*(sp\eye(nx*ny-nx))))^2;
22  v = (p+p^2)/((A+B)/nz+(C+D)/mz);
23  d1 = p;
24  d2 = v-p+1;
```

```
% disp(d2/(p*v))
T2 = (muX-muY)'*(sp\eye(nx*ny-nx))*(muX-muY);
pval = 1-cdf('F',T2*d2/(p*v),d1,d2);
end
```

```
function [Tpos,Tneg,Fpos,Fneg] = permute_test(X,Y,opt) %opt = 1 is
    different, opt=0 is not different
% n=0;
[~,~,nx] = size(X);
[~,~,ny] = size(Y);
kx = nx/29;
ky = ny/29;
Tpos = 0;
Fpos = 0;
Tneg = 0;
Fneg = 0;
itr = 0;
for i =1:kx
    for j=1:ky
        itr=itr+1;
            [pval,~] = Tsq(X(:,:,29*(i-1)+1:29*i),Y(:,:,29*(j-1)+1:29*j
                ));
            if pval <0.05 && opt== 1 % Reject Null Hypothesis (True
                positive)
                Tpos = Tpos+1;
            elseif pval<0.05 && opt== 0
                Fpos = Fpos+1;
            elseif pval>0.05 && opt ==1 %
                Fneg = Fneg+1;
            elseif pval>0.05 && opt ==0
                Tneg = Tneg+1;
            end
    end
end
```

## References


[1] Seth AK Barnett L. The MVGC multivariate Granger causality toolbox: A new approach to granger-causal inference. *Journal of Neuroscience Methods*, 2 2014.

[2] Adam B. Barrett, Lionel Barnett, and Anil K. Seth. Multivariate granger causality and generalized variance. *Phys. Rev. E*, 81:041907, Apr 2010.

[3] George.E.P. Box and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1976.

[4] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441, 2008.

[5] John Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313, 1982.

[6] Richard J. Harris. *A Primer of Multivariate Statistics*. Psychology Press, 3 edition, 2001.

[7] K. Krishnamoorthy and Jianqi Yu. Modified nel and van der merwe test for the multivariate behrensfisher problem. *Statistics and Probability Letters*, 66(2):161 – 169, 2004.

[8] A.D.R. McQuarrie and C.L. Tsai. *Regression and Time Series Model Selection*. World Scientific, 1998.

[9] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059 – 1069, 2010. Computational Models of the Brain.

[10] American Clinical Neurophysiology Society. Guideline 2: Guidelines for standard electrode position nomenclature. 2016.

[11] J. Songsiri. Learning multiple granger graphical models via group fused lasso. In *2015 10th Asian Control Conference (ASCC)*, pages 1–6, May 2015.

[12] O. Sporns. Brain connectivity. *Scholarpedia*, 2(10):4695, 2007. revision #91084.

[13] Zhuo Sun, Martijn van de Giessen, Boudewijn P. F. Lelieveldt, and Marius Staring. Detection of conversion from mild cognitive impairment to alzheimer's disease using longitudinal brain mri. *Frontiers in Neuroinformatics*, 11:16, 2017.

[14] P. Whittle. On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix. *Biometrika*, 50(1/2):129–134, 1963.