LEARNING BRAIN NETWORK DIFFERENCES USING STATISTICAL METHODS

Parinthorn Manomaisaowapak

Advisor: Assist. Prof. Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering

Chulalongkorn University

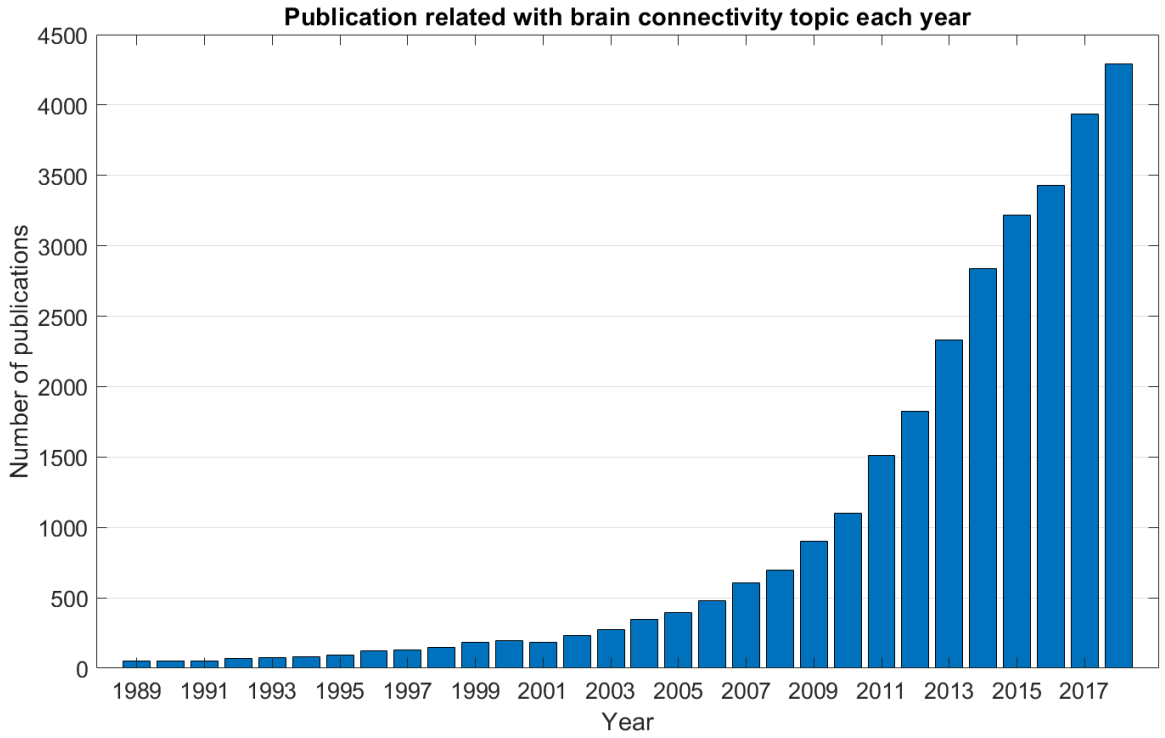
PRESENTATION OUTLINE

- Motivation
- Introduction
- Objectives
- Background
- Methodology
- Results
- Future work – In second semester

MOTIVATION

PubMed keyword:

(Brain connectivity) OR (Functional connectivity) OR (Effective connectivity)



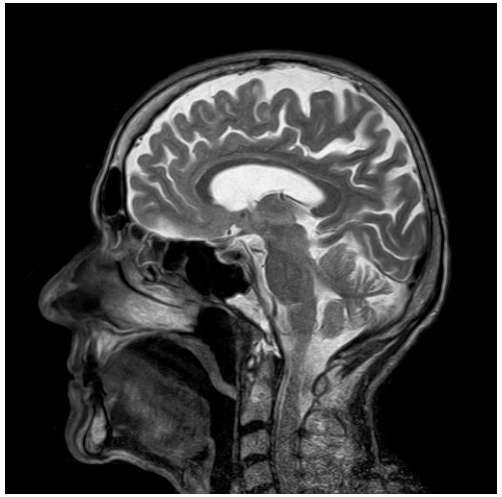
- Amount of publication per year in brain connectivity field (PubMed).
- Importance of classification between normal person and brain injured person.

INTRODUCTION

What are brain signals ?

- Electrical activities of the brain from neuronal activities.

Can be measured by

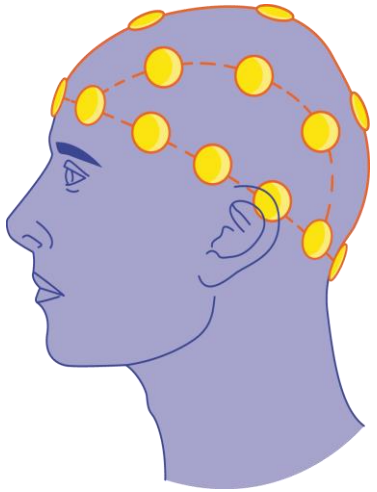


https://pixabay.com/photo-782459/

fMRI

Higher spatial resolution.

Low sampling rate.



https://smart.servier.com/smart_image/eeg-4/

EEG (Electroencephalogram)

Higher temporal resolution.

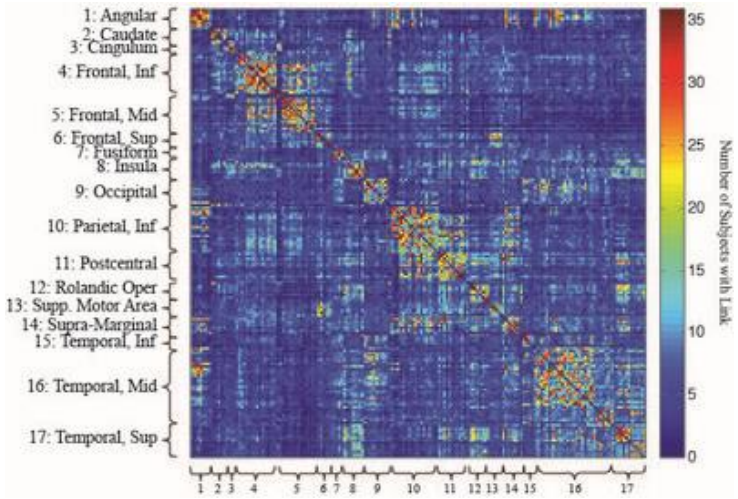
1000 Hz is used in this project.

Does not measure activity inside brain.

Interaction between brain regions can be referred as brain connectivity.

INTRODUCTION

Brain connectivity can be represented as a matrix.



Types of brain connectivity.

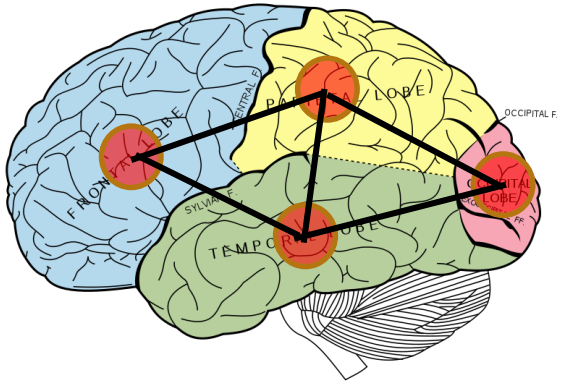
- Structural brain connectivity
  - Anatomical connection
- Functional brain connectivity
  - Statistical dependencies

Correlation does not imply causality.

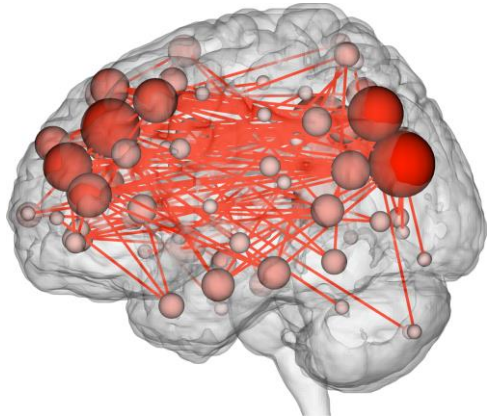
- Effective brain connectivity
  - Causal interaction

↓

Granger causality ⟶ How about the differences ?



http://disjointedthinking.jeffhughes.ca/2011/02/all-about-the-brain-part-1/



http://theconversation.com/brain-activity-is-as-unique-and-identifying-as-a-fingerprint-48723

http://www.neuroinformatics2011.org/abstracts/identification-and-classification-of-functional-modules-in-the-brain.html

OBJECTIVES

There are two objectives of this project,

- To estimate brain network using Granger causality concept from EEG or fMRI data.

  Will be explained in Background section

- To compare brain network difference between two control group and patient group.

BACKGROUND

Consider two approaches to find group differences.

- Statistical approach [semester 1]
  - Estimate Granger causality (GC) matrix individually.
  - Differences were defined by differences in average value of GC matrices.

- Sparse estimation approach [Expected to be done in semester II]
  - Estimate GC matrix two group simultaneously.
  - Differences can be determined by distance measure or statistical test.

Multivariate Granger causality (GC)

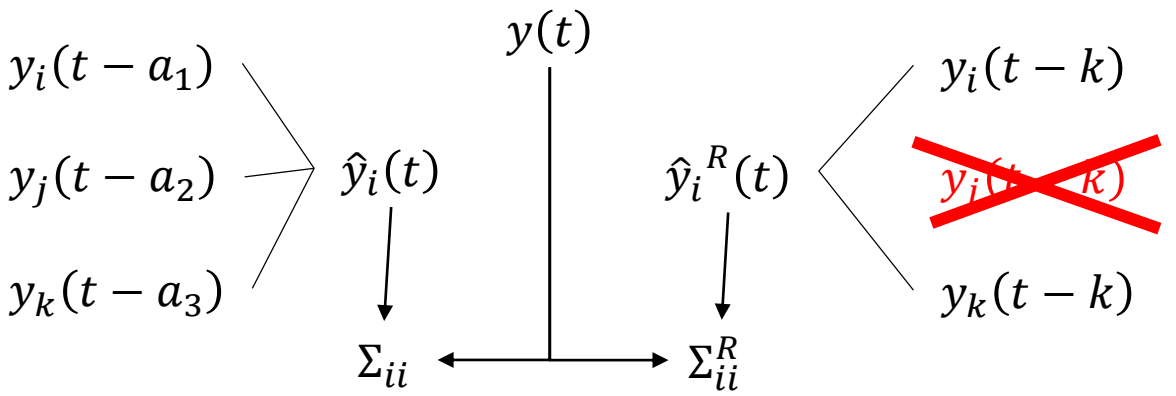
Linear vector autoregressive model (VAR) is used to compute GC.

$$y(t) = \sum_{k=1}^{p} A_k y(t-k) + e(t)$$ **VAR process order p**

Idea

- Compare model quality between full model and reduced model.

  Improve residual covariance matrix or not



GC measure

$$F_{ij} = \log \frac{\det \Sigma_{ii}^{\mathrm{R}}}{\det \Sigma_{ii}}$$

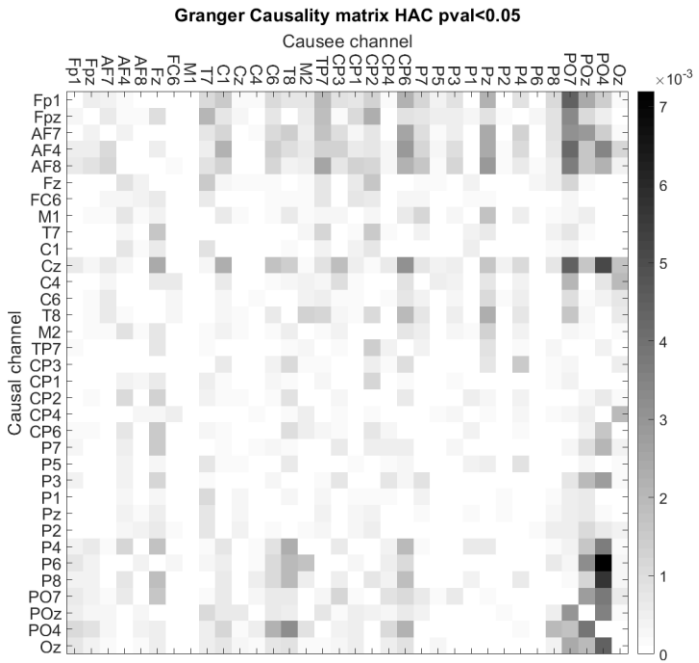
GC matrix

$$\begin{bmatrix} F_{11} & \cdots & F_{1n} \\ \vdots & \ddots & \vdots \\ F_{n1} & \cdots & F_{nn} \end{bmatrix}$$

Σ_{ii}^R is residual covariance of reduced model

Σ_{ii} is residual covariance of full model

Exact zero is nearly impossible in estimation ⟶ Significant test for finding zero



Granger causality (GC) matrix

$$H_0 : \mathcal{F}_{ij} = 0$$

$$H_1 : \mathcal{F}_{ij} \neq 0.$$

$$(N - p)\mathcal{F}_{ij} \sim \chi^2_{p(n_i+n_j)}$$

Barnett & Seth (2014)

N, p, n_i, n_j denotes sample size, lags, dimension of y_i, y_j respectively.

Significant test for group difference test.

Test vector mean difference.

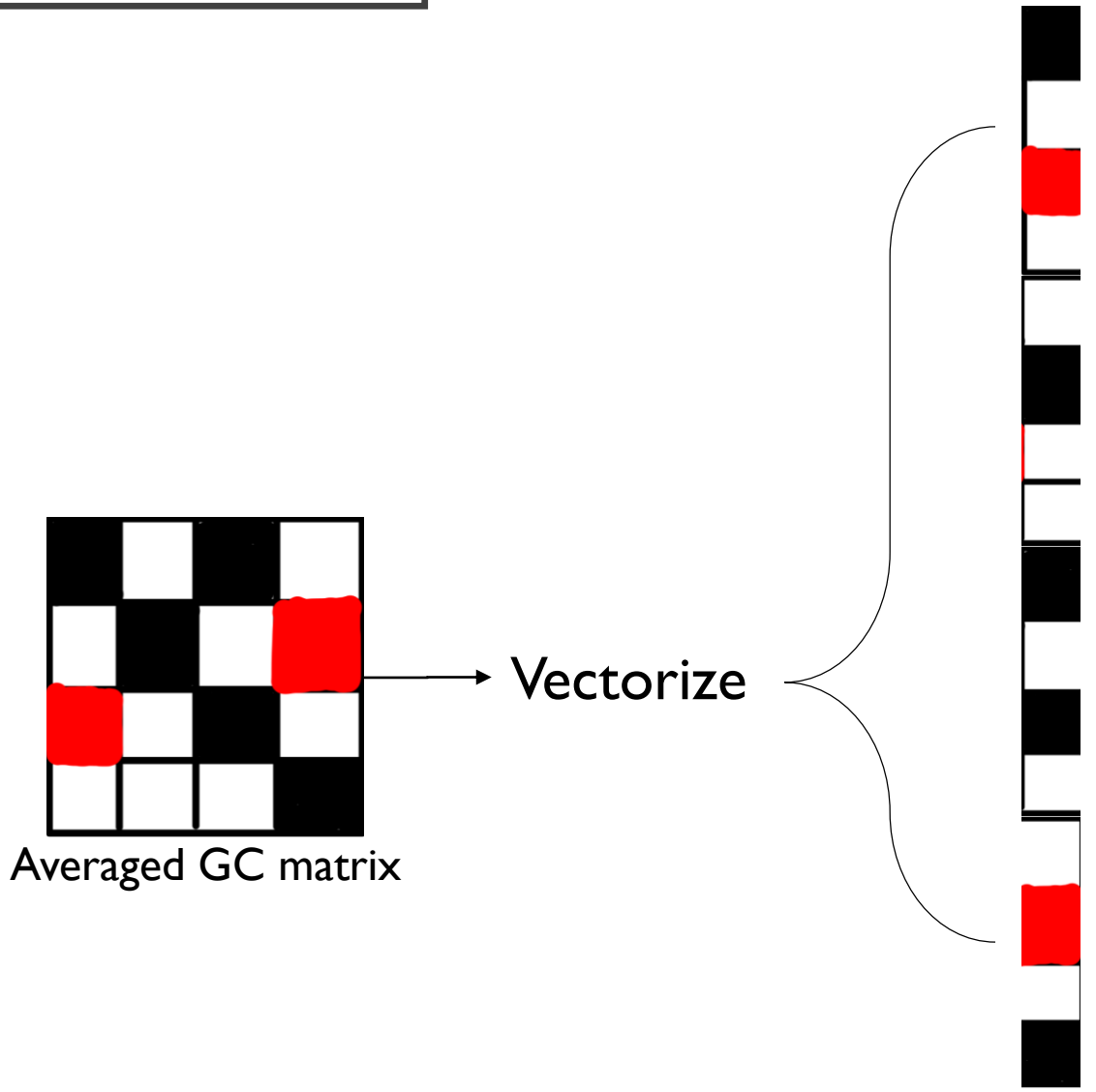
$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2.$$

μ_i denotes population mean of the ith group.

Two-sample Hotelling's T-squared test

$$T^2 = (\bar{X}_1 - \bar{X}_2)^T \left(\frac{S_1}{N_1} + \frac{S_2}{N_2}\right)^{-1} (\bar{X}_1 - \bar{X}_2)$$

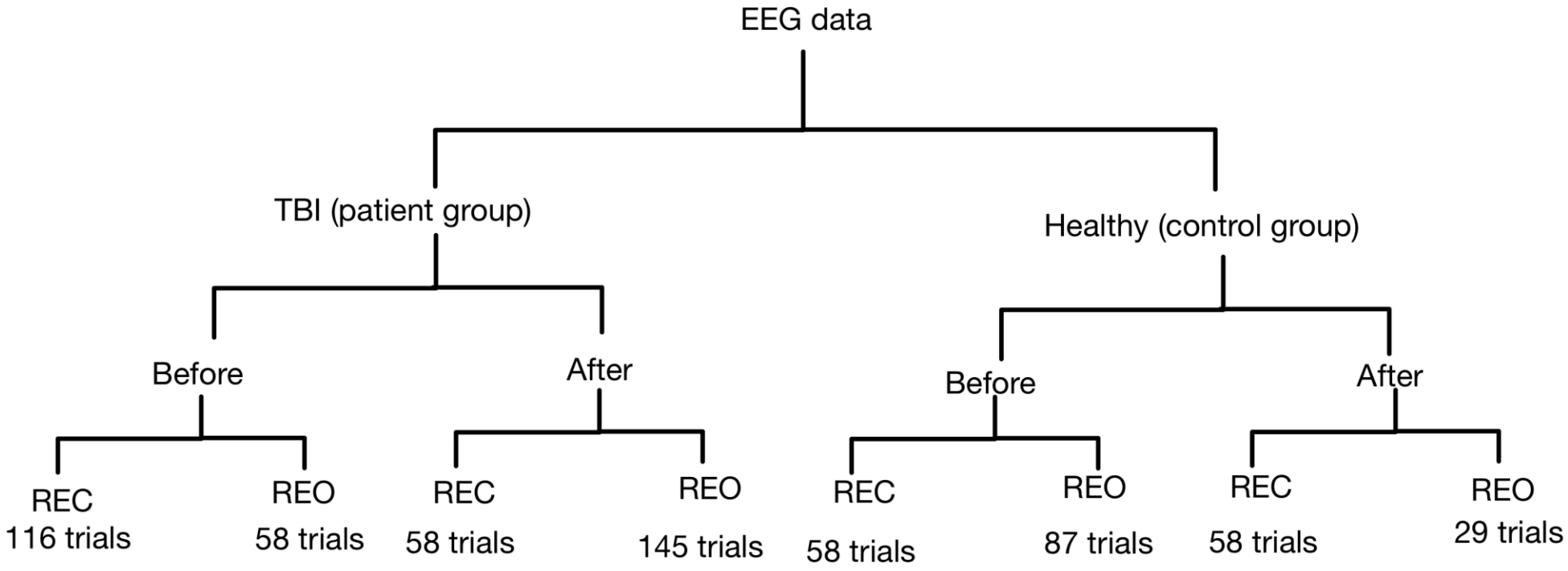
$$T^2 \sim \frac{pv}{v - p + 1} F_{p,v-p+1}$$



METHODOLOGY

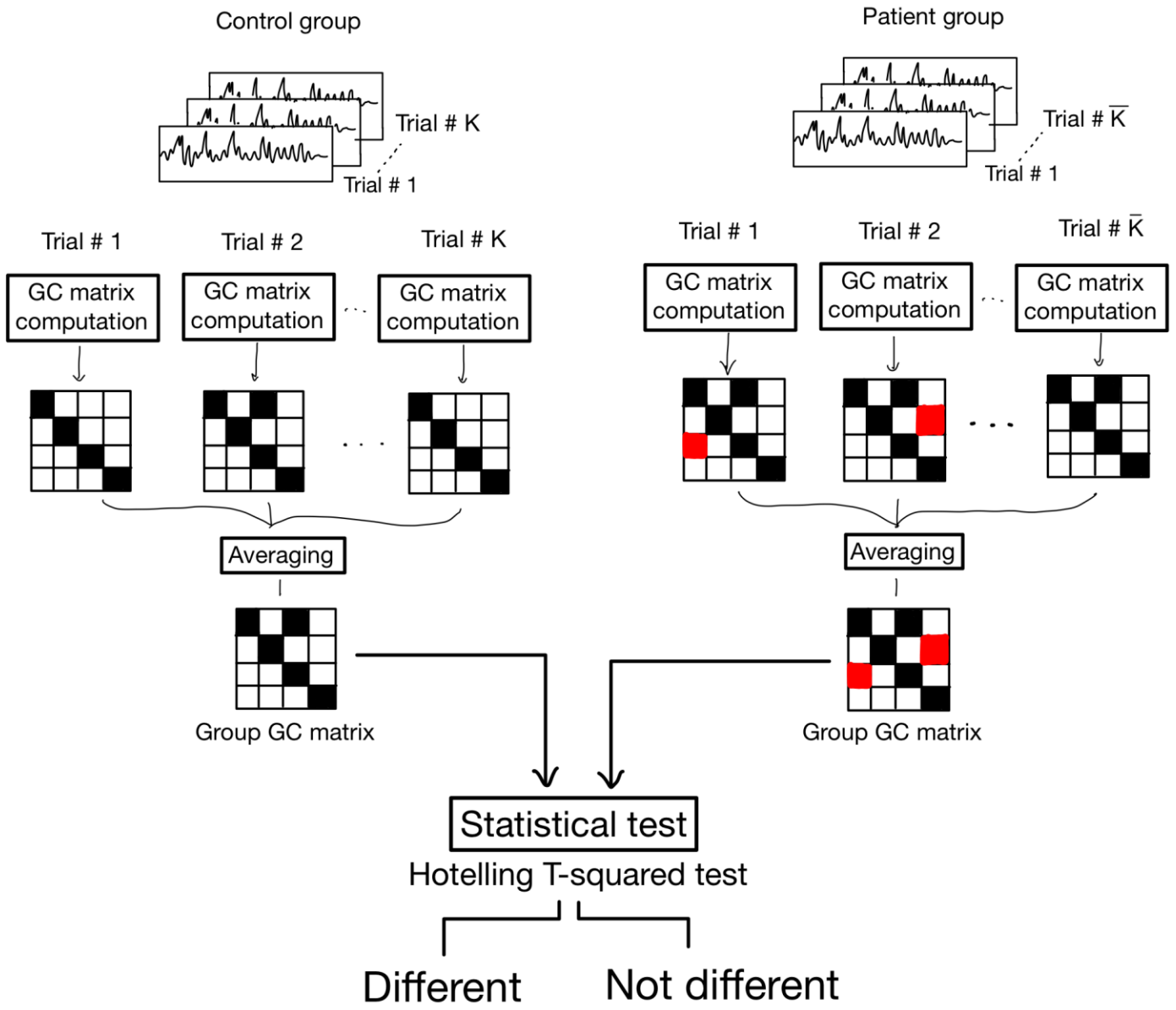
STATISTICAL APPROACH

Data from USM (Universiti Sains Malaysia)

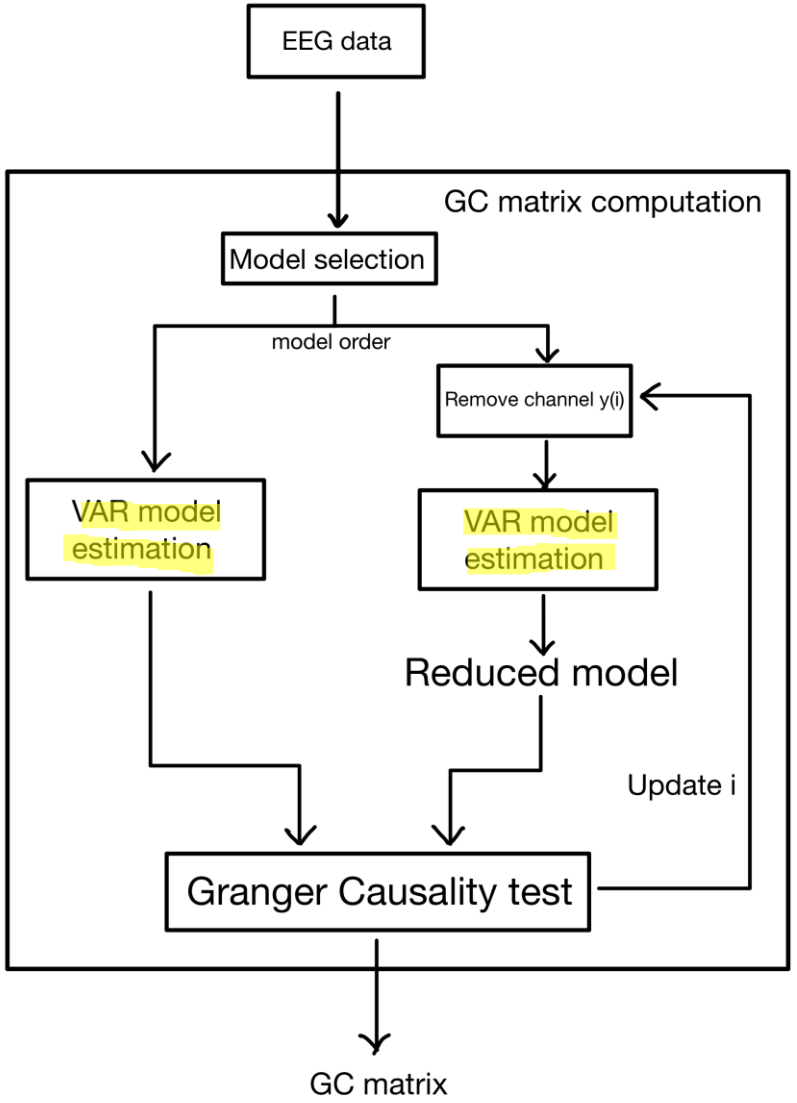


METHODOLOGY

STATISTICAL APPROACH



METHODOLOGY | STATISTICAL APPROACH



Parameter estimation

- Ordinary least square.
- Solve via Yule-Walker equation.

Both methods asymptotically yield same solution.
But Yule-Walker equation can be solve efficiently.

Hotelling's T squared test

Without equal population covariance assumption

Use Multivariate Granger Causality toolbox (MVGC),

RESULTS

STATISTICAL APPROACH

	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	-	-

Positive = groups are different

Negative = groups are not different

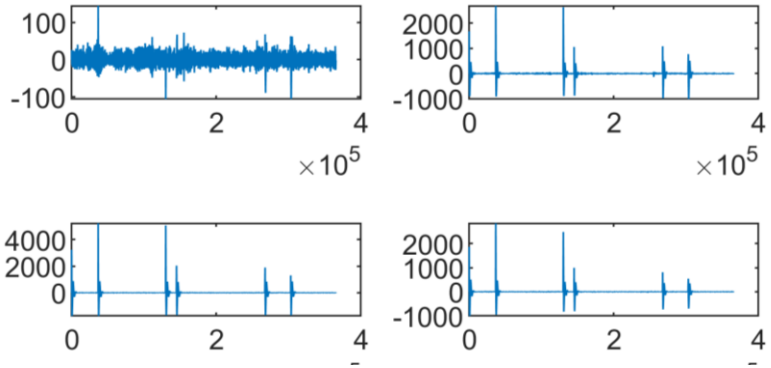
Accuracy : 66.8 %

True positive rate : 81.73%

True negative rate : 48.19%

The received data contains highly correlated EEG electrodes, abnormality such as spikes in signal.

TNR without Healthy: 66.67%



Spikes in signal

FUTURE WORK | SPARSE ESTIMATION APPROACH

Why sparse brain connectivity ?

-Can capture brain activities effectively.[2]

-Brain connections are anatomically sparse.

How to achieve ?

-Sparsity pattern of parameter such as in VAR model.

How to determine the differences ?

- Statistical testing (Hypothesis test).
- Distance measure that represents brain network differences.

General formulation

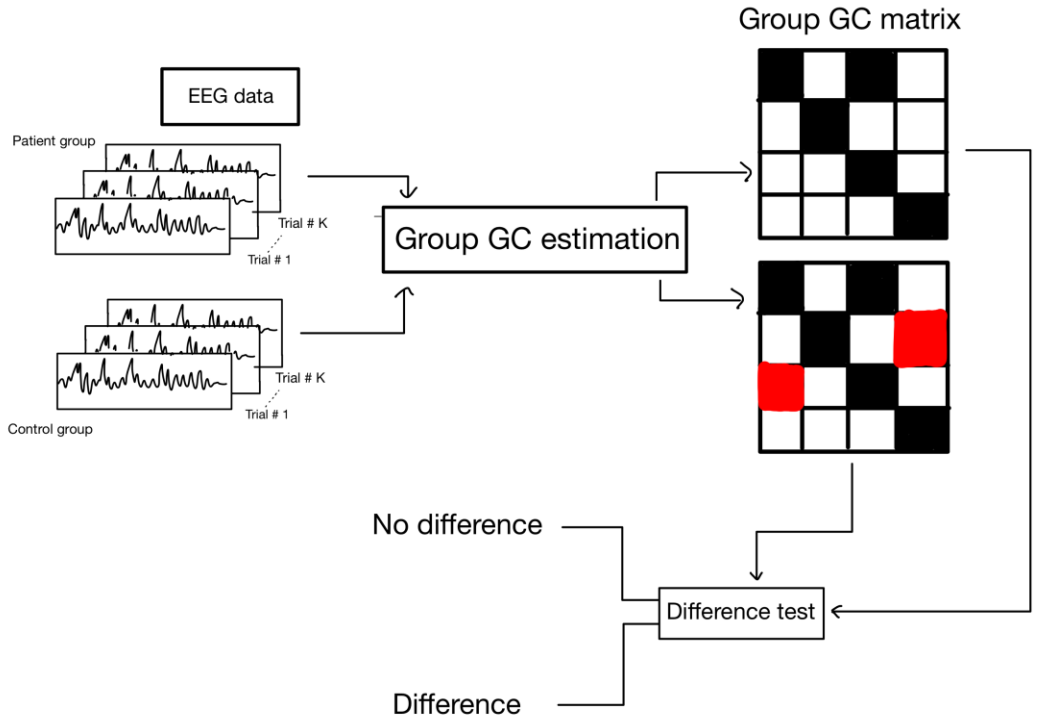
$$\underset{\theta_1,\theta_2}{\text{minimize}} \quad g_1(\theta_1) + g_2(\theta_2) + \lambda(\|\theta_1\|_1 + \|\theta_2\|_1) + \gamma(\|\theta_1 - \theta_2\|_1)$$

$g_i(\cdot)$ is defined as cost function that represent the goodness of fit

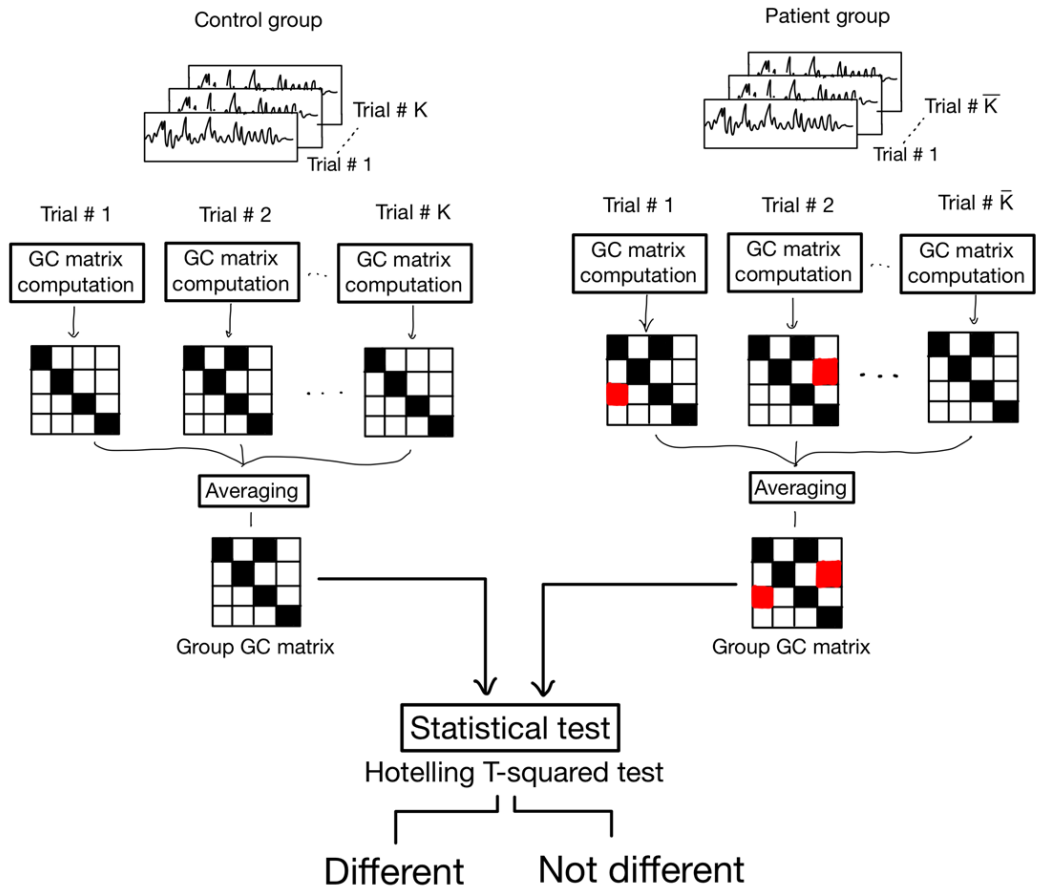
θ denotes model parameters.
λ controls sparsity pattern of model parameters.
γ controls similarity of model parameter between groups.

FUTURE WORK

SPARSE ESTIMATION APPROACH



Sparse estimation



Statistical approach

Q&A

SUPPLEMENTARY : FULL VS REDUCED MODEL EXAMPLE

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

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

FULL $\begin{bmatrix}\hat{y}_i(t)\\ \hat{y}_j(t)\end{bmatrix} = \begin{bmatrix}A_{ii} & A_{ij}\\ A_{ji} & A_{jj}\end{bmatrix}\begin{bmatrix}y_i(t-1)\\ y_j(t-1)\end{bmatrix}$ $\Sigma_{ii} = cov(y_i - \hat{y}_i)$

Past occurrence of y_j is included when estimates y_i

REDUCED $\hat{y}_i^R(t) = A_{ii}^R y_i(t-1)$ $\Sigma_{ii}^{\mathrm{R}} = cov(y_i - \hat{y}^R{}_i)$

Past occurrence of y_j does not included

SUPPLEMENTARY : PARAMETER EST.

Least square formulation

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

$$Y = \beta X$$

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

Solve via QR factorization

Derived normal equation

$$\beta(XX^T) = YX^T$$

Yule-walker equation [4]

$$[\Gamma(1) \quad \Gamma(2) \quad \cdots \quad \Gamma(p)] = [A_1 \quad A_2 \quad \cdots \quad A_p] \begin{bmatrix} \Gamma(0) & \Gamma(1) & \cdots & \Gamma(p-1) \\ \Gamma(1)^T & \Gamma(0) & \cdots & \Gamma(p-2) \\ \vdots & \cdots & \ddots & \vdots \\ \Gamma(p-1)^T & \Gamma(p-2)^T & \cdots & \Gamma(0) \end{bmatrix}$$

Solve via LWR (Levinson Wiggins Robinson) algorithm.

SUPPLEMENTARY : ASYMPTOTIC EQUAL

$$[\Gamma(1) \quad \Gamma(2) \quad \cdots \quad \Gamma(p)] = [A_1 \quad A_2 \quad \cdots \quad A_p] \begin{bmatrix} \Gamma(0) & \Gamma(1) & \cdots & \Gamma(p-1) \\ \Gamma(1)^T & \Gamma(0) & \cdots & \Gamma(p-2) \\ \vdots & \cdots & \ddots & \vdots \\ \Gamma(p-1)^T & \Gamma(p-2)^T & \cdots & \Gamma(0) \end{bmatrix}$$

$$YX^T = \beta(XX^T)$$

$$XX^T = \begin{bmatrix} \sum_p^{N-1} y(k)y(k)^T & \sum_p^{N-1} y(k)y(k-1)^T & \cdots & \sum_p^{N-1} y(k)y(k-p+1)^T \\ \sum_p^{N-1} y(k-1)y(k)^T & \sum_p^{N-1} y(k)y(k-1)^T & \cdots & \sum_p^{N-1} y(k-1)y(k-p+1)^T \\ \vdots & \cdots & \ddots & \vdots \\ \sum_p^{N-1} y(k-p+1)y(k)^T & \sum_p^{N-1} y(k-p+1)y(k-1)^T & \cdots & \sum_p^{N-1} y(k-p+1)y(k-p+1)^T \end{bmatrix}$$

$$YX^T = \left[\sum_p^{N-1} y(k+1)y(k)^T \quad \cdots \quad \sum_p^{N-1} y(k+1)y(k-p+1)^T\right]$$

SUPPLEMENTARY : HOTELLING'S T^2

Require : Both samples are drawn from normal distribution.
achieved by Central Limit Theorem.
Require : Both samples have equal population covariance matrix.
Solution of Behrens-Fisher problem can be used in case unequal covariance matrix.

Krishnamoorthy & Yu (2004)

Approximates distribution by adjusting degree of freedom

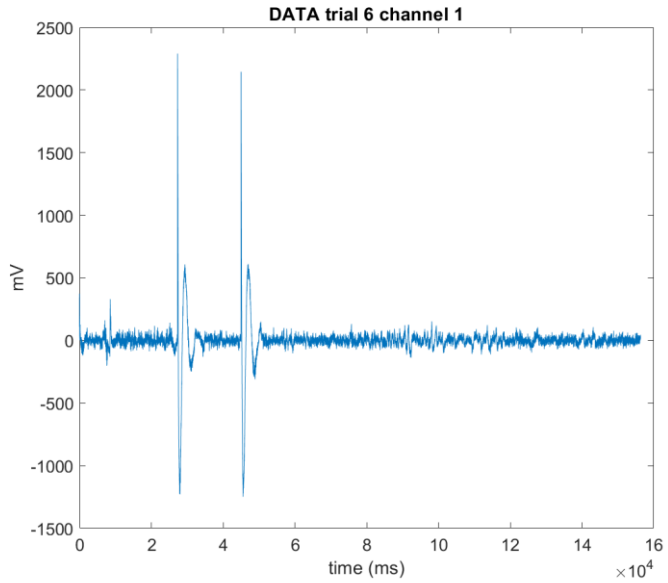
$$v = \frac{p + p^2}{A_1 + A_2}$$

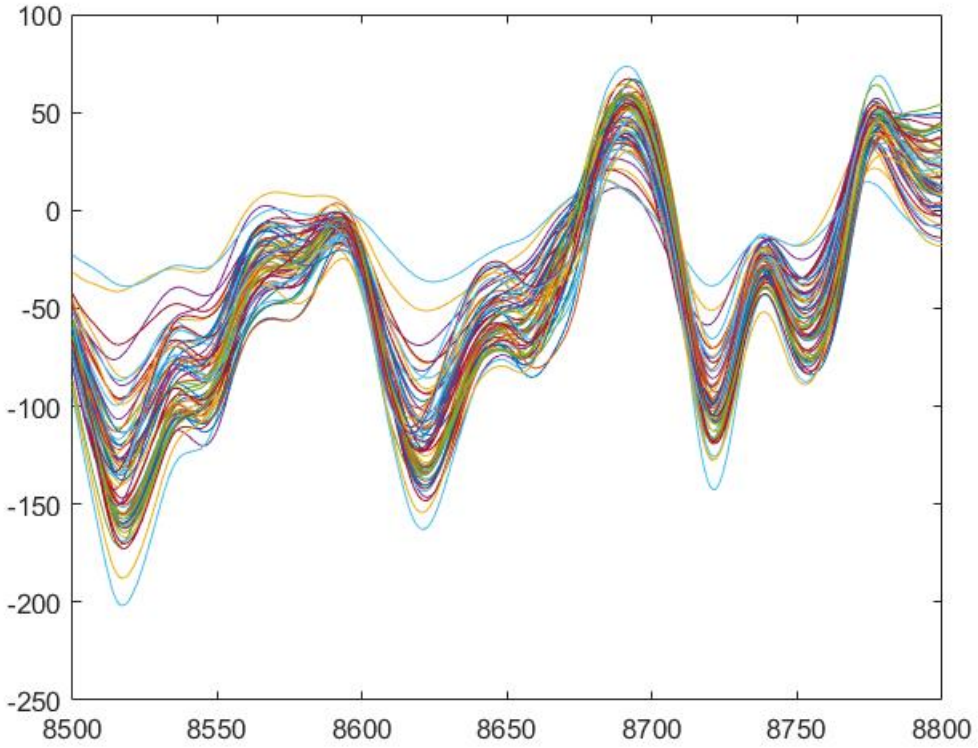
$$T^2 \sim \frac{pv}{v - p + 1} F_{p,v-p+1}$$

$$A_i = \frac{\mathrm{tr}[(\tilde{S}_i S_p^{-1})^2] + [\mathrm{tr}(\tilde{S}_i S_p^{-1})]^2}{N_i}; i = 1, 2$$

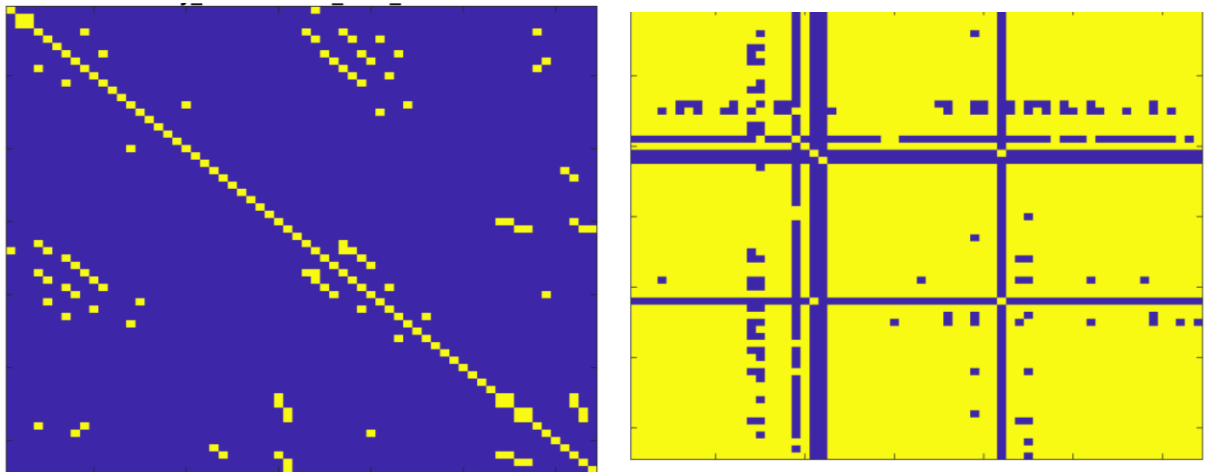
$$\tilde{S}_i = \frac{S_i}{N_i}; S_p = \tilde{S}_1 + \tilde{S}_2$$

SUPPLEMENTARY : DATA PROBLEM





SUPPLEMENTARY : DATA PROBLEM



Correlation matrix between normal(left) vs high correlated(right)

References

[1] Lionel Barnett, Anil K. Seth, The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference, *Journal of Neuroscience Methods*, Volume 223, 2014.
[2] Harini Eavani, Theodore D. Satterthwaite, Roman Filipovych, Raquel E. Gur, Ruben C. Gur, Christos Davatzikos, Identifying Sparse Connectivity Patterns in the brain using resting-state fMRI, *NeuroImage*, Volume 105, 2015, Pages 286-299.
[3] K. Krishnamoorthy, Jianqi Yu, Modified Nel and Van der Merwe test for the multivariate Behrens–Fisher problem, *Statistics & Probability Letters*, Volume 66, Issue 2, 2004, Pages 161-169.
[4] A.D.R. McQuarrie and C.L. Tsai. *Regression and Time Series Model Selection*. World Scientific, 1998.