Senior Project 2102499 Year 2019

Learning Granger causality for multivariate time series using state-space models

Anawat Nartkulpat ID 5930562521

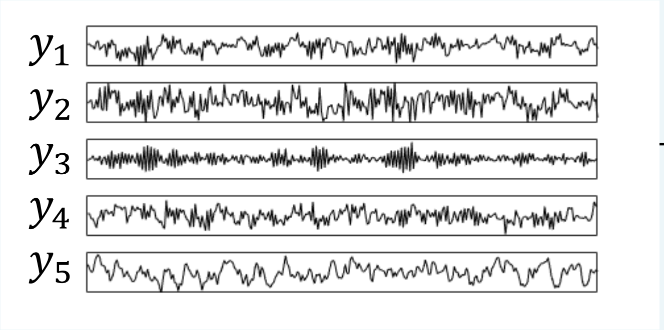
Advisor: Assist. Prof. Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering
Chulalongkorn University

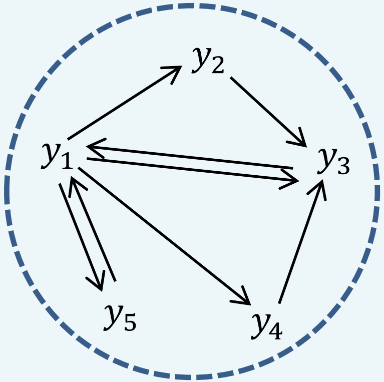
Outline

- Introduction
- Project overview
- Methodology
- Comparative method: Gaussian Mixture Model
- Result and Discussion
- Conclusion

Introduction



Time series data



Causality pattern

- **Granger causality (GC)** is a tool to measure causal connectivities between variables in time series based on model estimation. A state-space model is considered as it is more general than other linear models such as autoregressive model or moving average model.
- Learning causalities in time series data has many applications especially in neuroscience in which causal relationships between brain regions are explored.
- The statistical distribution of GC of the state-space model is unknown, so a method to classify zero and non-zero causalities is proposed in [1] by fitting averaged GC measures to a Gaussian Mixture Model (GMM).
- We consider applying permutation test, which does not required knowledge of GC distribution, to classify zero and non-zero causalities.

Project Overview

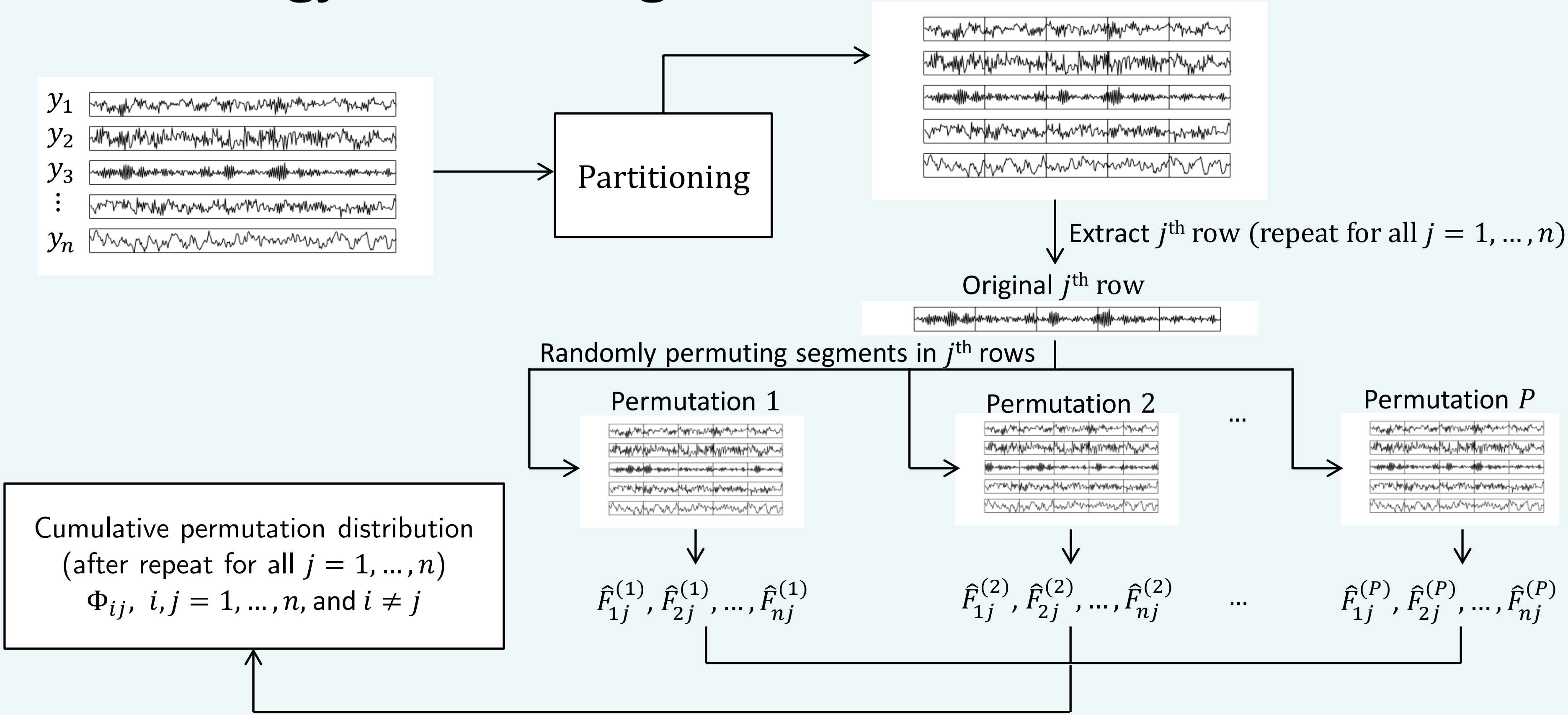
- Objectives

- To develop a scheme for classifying the zero patterns of the GC of state-space models using the permutation test.
- To compare the performance of the permutation test with the GMM method in classification of zero and non-zero entries of GC matrix obtained from state-space model

- Scope of work

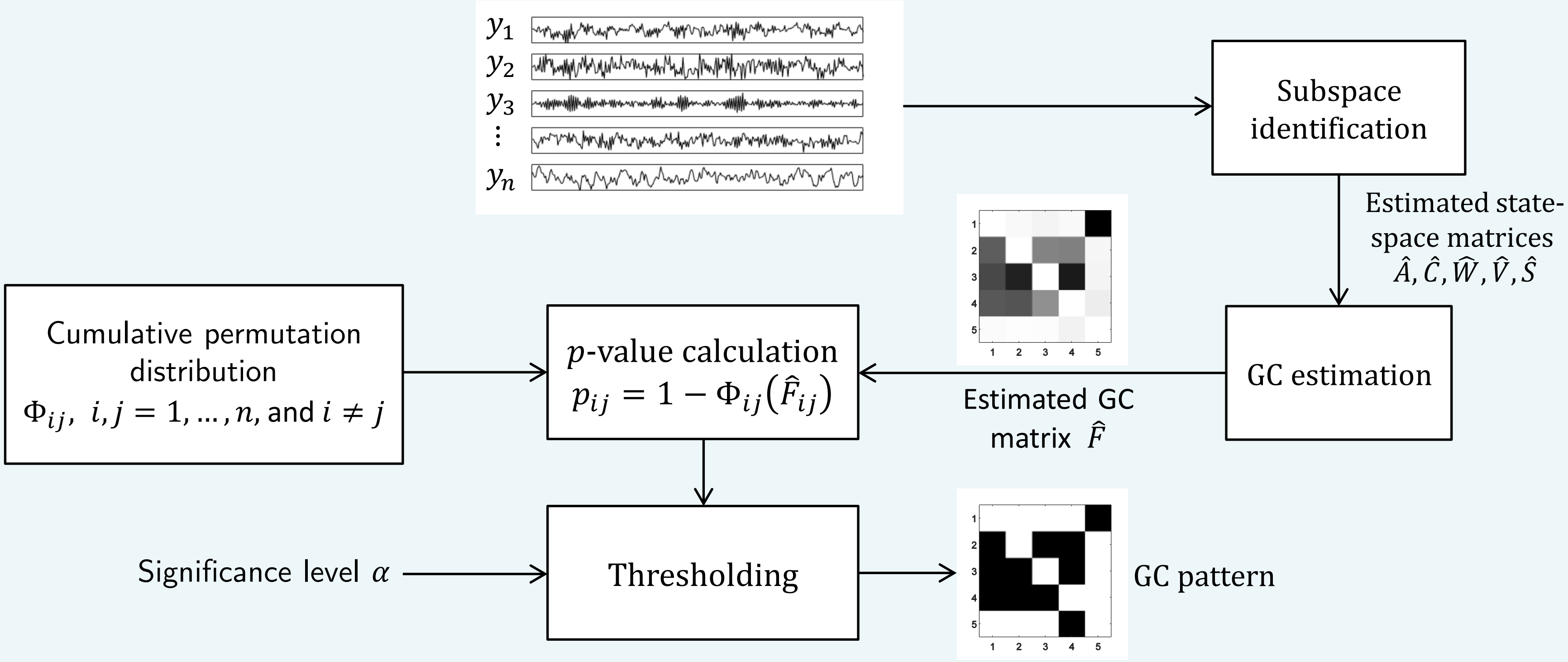
- We perform GC estimation on simulated data
- We compare performance, computational cost and assumptions required between permutation test and GMM method

Methodology: GC learning scheme



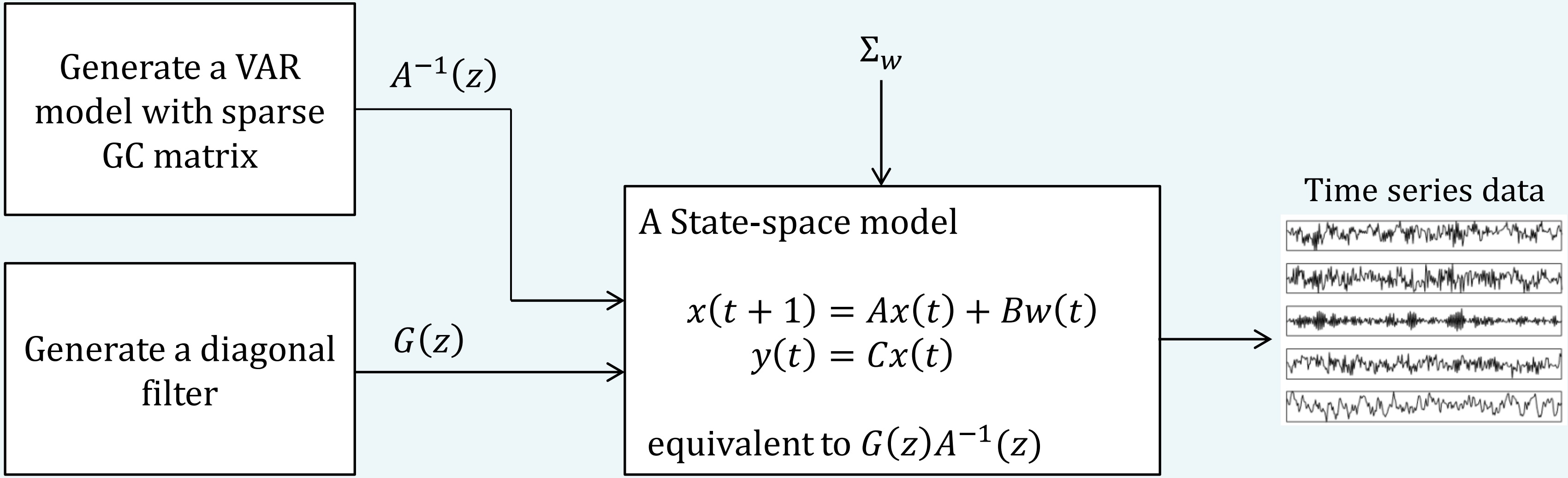
The scheme for obtaining permutation distribution.

Methodology: GC learning scheme



The scheme learning GC using permutation test.

Methodology: Ground truth model generation



The scheme for generating ground truth model and time series data.

Methodology: Subspace identification

We consider estimating parameters A, C, W, V, S of a stochastic state-space model

$$x(t+1) = Ax(t) + w(t)$$
$$y(t) = Cx(t) + v(t)$$

This method is based on orthogonal projection. Suppose that the outputs Y is known, it was shown in [2] that

$\mathcal{O}_i \equiv Y_{i|2i-1} / Y_{0|i-1} = Y_f / Y_p$ (Projecting the future outputs onto the past output space)

$\mathcal{O}_i = \Gamma_i \hat{X}_i \Longrightarrow \hat{X}_i = \Gamma_i^{\dagger} \mathcal{O}_i$ and $\hat{X}_{i+1} = \Gamma_{i-1}^{\dagger} \mathcal{O}_{i-1}$

$$\begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} \hat{X}_i + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \Longrightarrow \begin{cases} \begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix} = \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} \hat{X}_i^{\dagger} \\ \begin{bmatrix} \widehat{W} & \hat{S} \\ \hat{S}^T & \hat{V} \end{bmatrix} = \frac{1}{j} \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}^T \end{cases}$$

Methodology: Granger causality

The measure of the Granger causality from y_j to y_i is defined by

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

where Σ is the covariance of the prediction error given all other y_k and Σ^R is the covariance of the prediction error given all other y_k except y_j [3].

The calculation of Σ and Σ^R are done by solving P from the Discrete Algebraic Riccati Equation (DARE)

$$P = APA^T - (APC^T + S)(CPC^T + V)^{-1}(CPA^T + S^T) + W$$

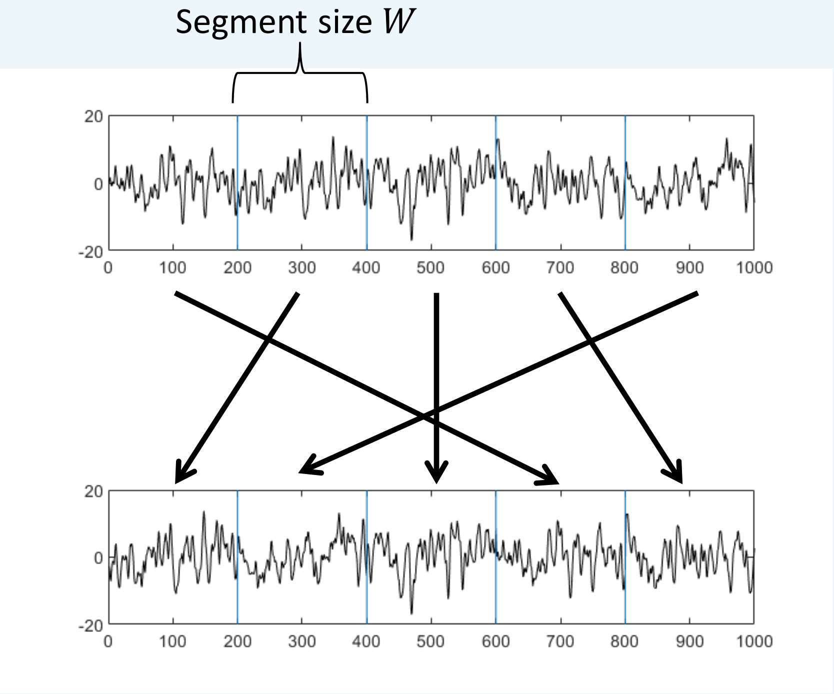
And using the fact that $\Sigma = CPC^T + V$. For Σ^R, we again solve DARE but without j^{th} row in C, and without both j^{th} row and column in V.

Methodology: Permutation test

- The distribution of F_{ij} is unknown
- The null hypothesis $H_0\colon F_{ij} = 0$ is to be tested

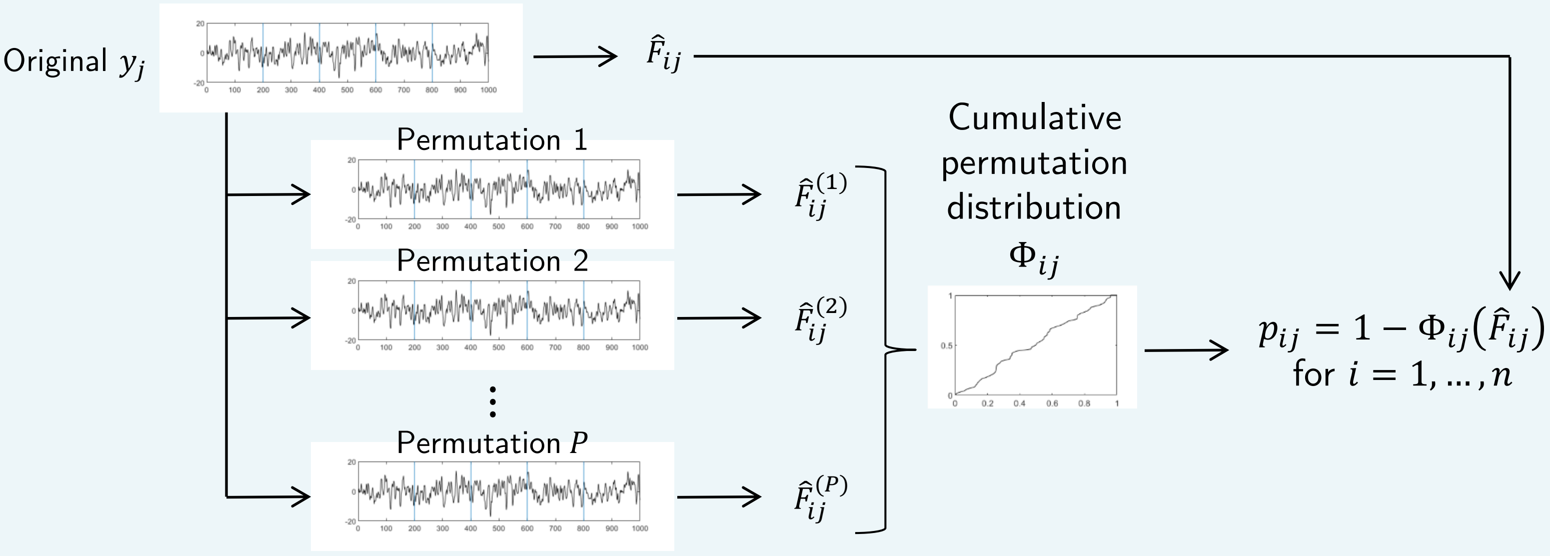
Justification: Under the true null hypothesis, y_i is not Granger-caused by y_j, so rearranging data in channel y_j does not change the outcome.

So, we may form a distribution of F_{ij} under H_0 empirically from the permutations of data [4].



One of the possible permutations.

Methodology: Permutation test



The scheme for calculating p-values.

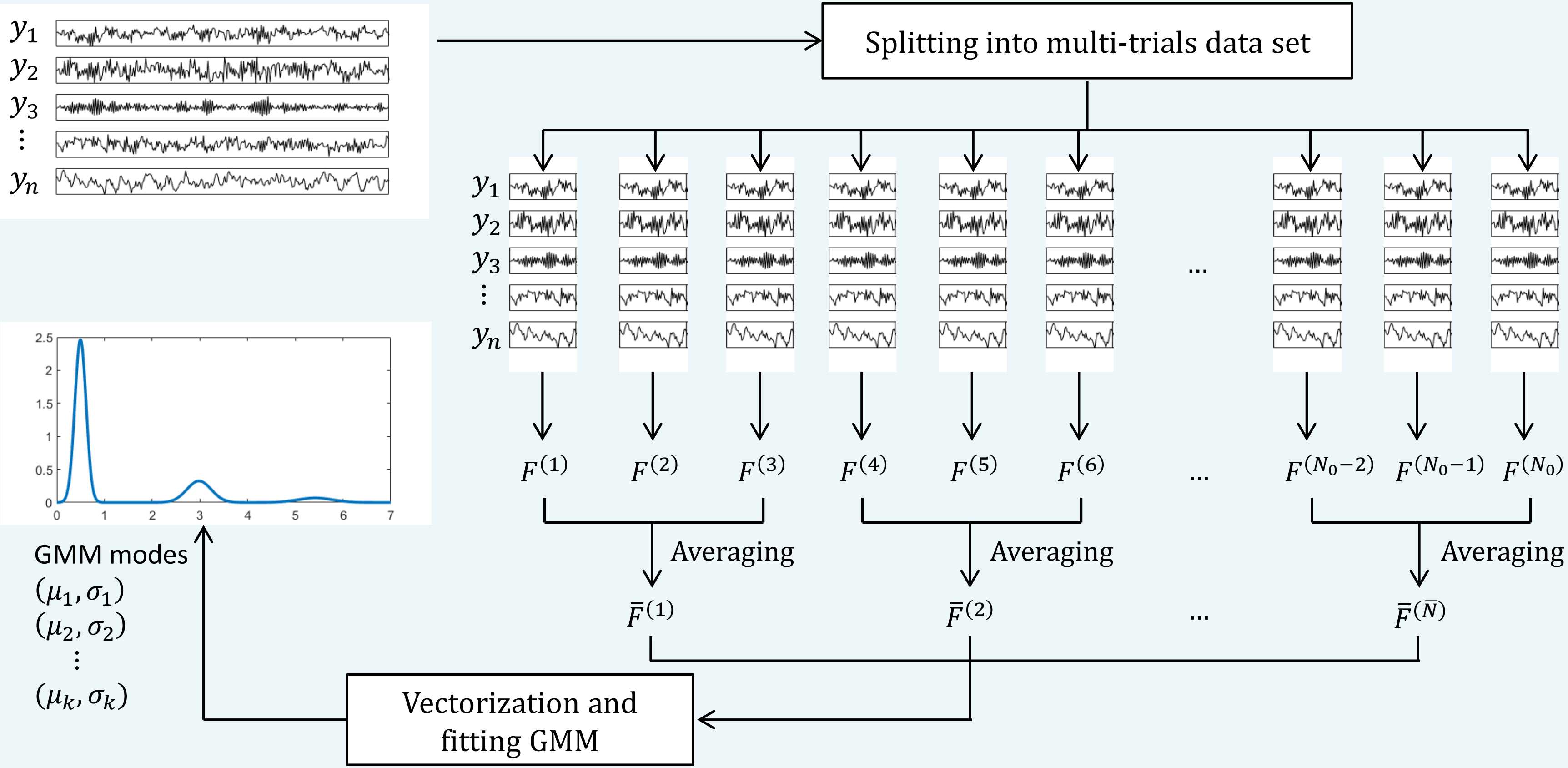
Methodology: Permutation test

For a given significance level α, F_{ij} can be tested to decide that $F_{ij} = 0$ or $F_{ij} \neq 0$ by thresholding the p-values.

- **Multiple testing issue**: Testing many hypotheses ($F_{ij} = 0$ for all i, j) at once may give overall Type I error, or a family-wise error rate (FWER), greater than α.

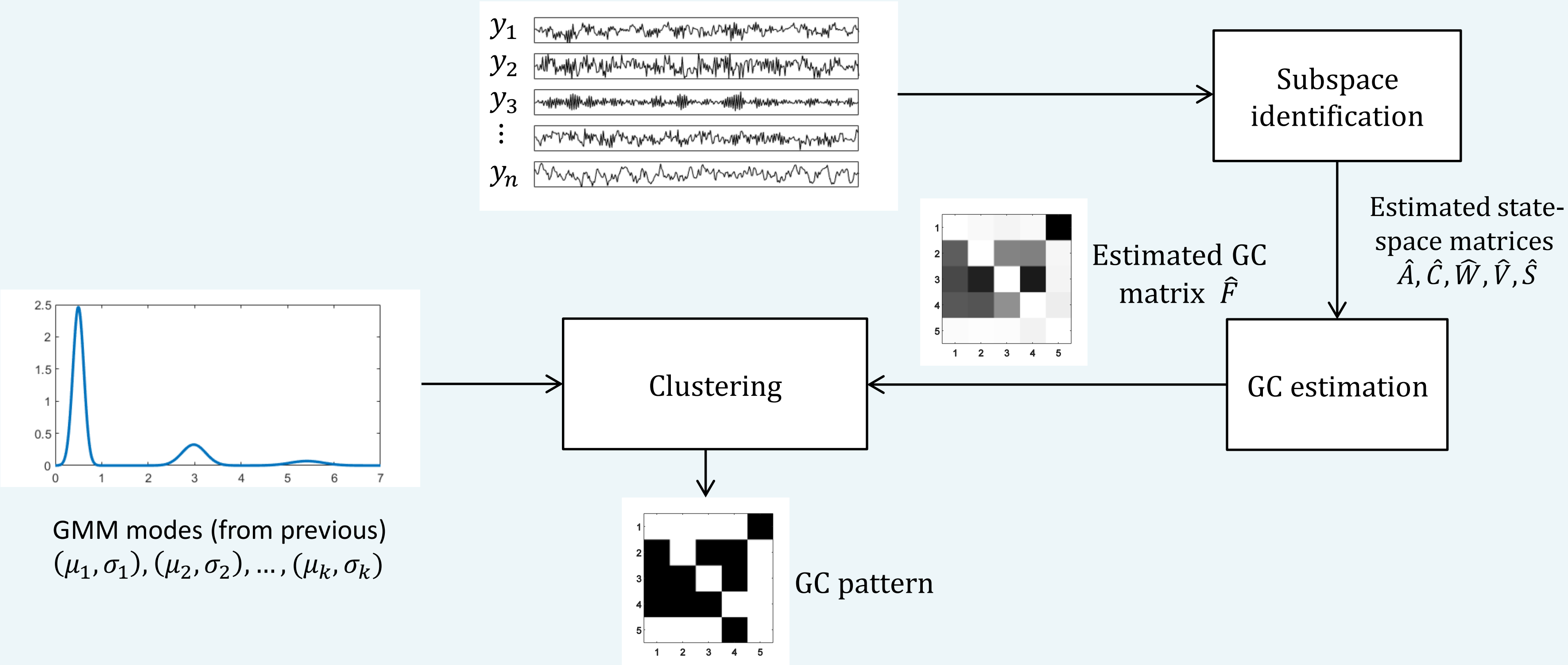
- **Remedies:** Let N be the number of hypotheses to be tested.
  - **Bonferroni Correction:** Test each hypotheses with a corrected significance level $\alpha_{Bon} = \frac{\alpha}{N}$.
  - **Benjamini-Hochberg procedure:** Sorting p-values in the ascending order
    $$p_1 \leq p_2 \leq \cdots \leq p_N$$
    Then use a corrected significance level $\alpha_{BH} = \frac{\alpha}{N} \max\{k | p_k \leq \frac{k\alpha}{N}\}$ for thresholding.

Comparative method: Gaussian Mixture Model (GMM)



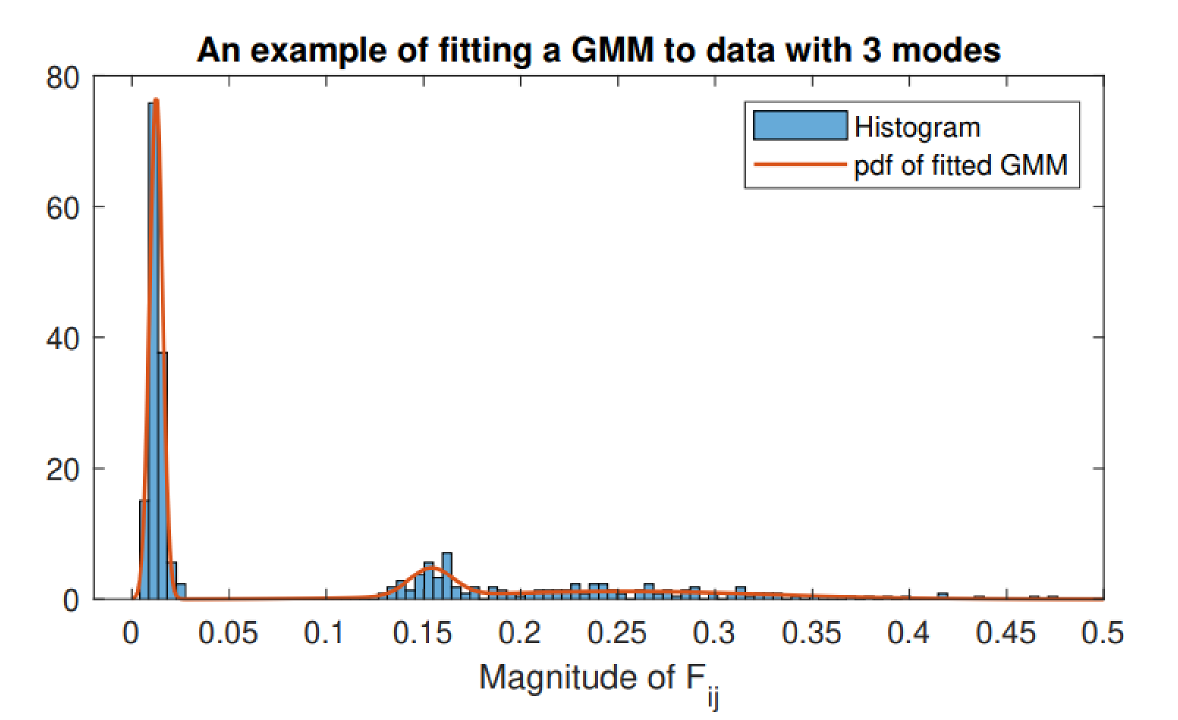
The scheme for obtaining GMM from time series data.

Comparative method: Gaussian Mixture Model (GMM)



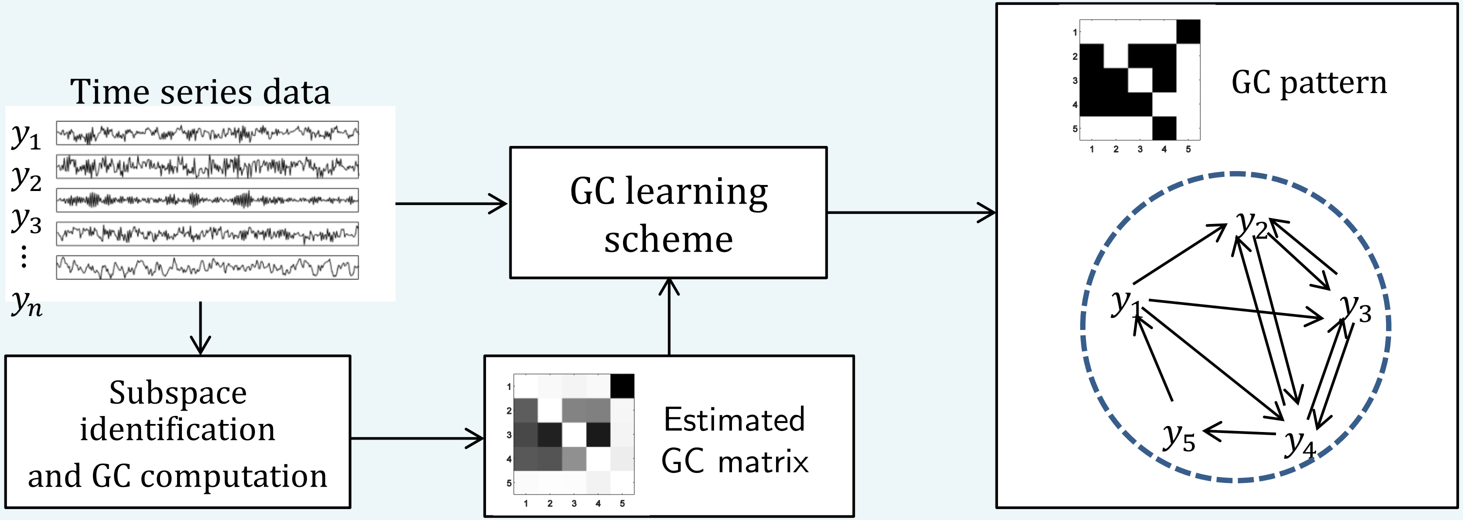
The scheme learning GC using GMM method.

Comparative method: Gaussian Mixture Model (GMM) [1]



An example of fitting a GMM to data with 3 modes.

Results & Discussion



- Complete permutations and Monte-Carlo permutation test
- Comparison of the performance between permutation test and GMM method
- Performance under different ground truth network densities
- Comparison of the computation time between permutation test and GMM method

Results & Discussion

The performance of the permutation test when choosing the number of partitioning segments to be 5 (Complete) and 10 (Monte-Carlo) segments respectively.

Performance index	5 segments			10 segments		
	Simple	Bon	B-H	Simple	Bon	B-H
ACC	0.9678	0.9047	0.9047	0.9650	0.9944	0.9944
TPR	1	0	0	1	1	1
TNR	0.9644	1	1	0.9613	0.9939	0.9939
FPR	0.0356	0	0	0.0387	0.0061	0.0061
FNR	0	1	1	0	0	0

- Complete test yielded slightly better ACC without correction.
- Monte-Carlo test allowed applying correction methods and gave much better performance.

Results & Discussion

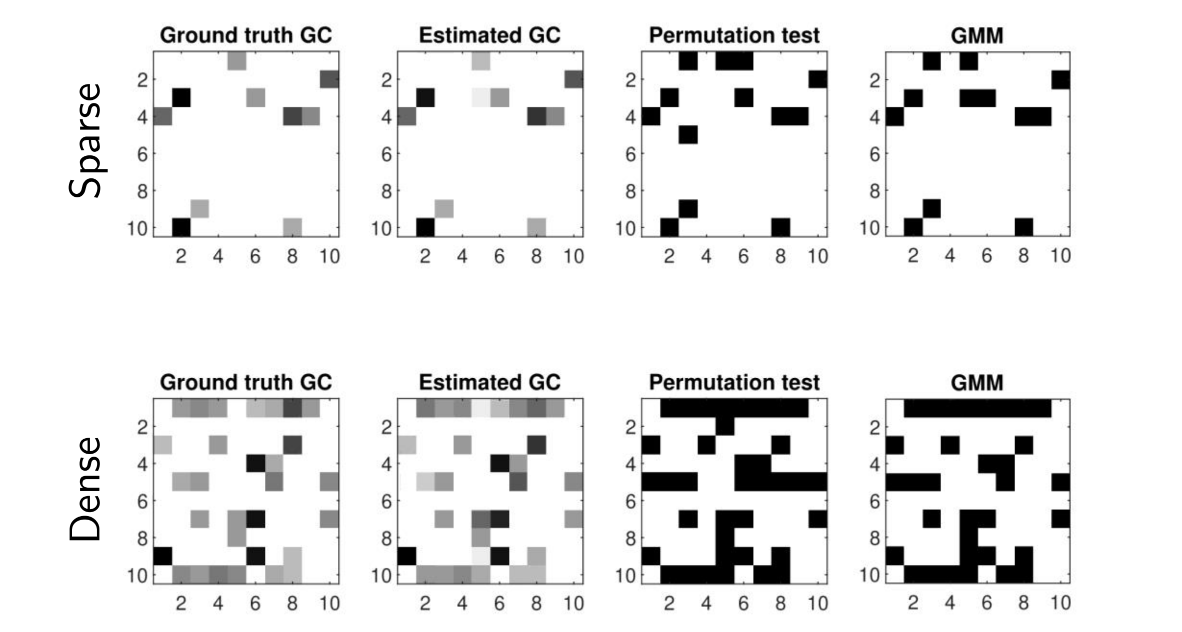
The performance of GMM method and permutation test on data from ground truth models with sparse GC.

Performance indices	20000 time points				50000 time points			
		Permutation test				Permutation test		
	GMM	Simple	Bon	B-H	GMM	Simple	Bon	B-H
ACC	0.9558	0.9297	0.9922	0.9803	0.9933	0.9447	0.9936	0.9869
TPR	1	1	1	1	1	1	1	1
TNR	0.9512	0.9223	0.9914	0.9782	0.9926	0.9389	0.9929	0.9856
FPR	0.0488	0.0777	0.0086	0.0218	0.0074	0.0611	0.0071	0.0144
FNR	0	0	0	0	0	0	0	0

- With more data, GMM method can perform as good as permutation test.

Results & Discussion

Examples of GC patterns obtained from permutation test and GMM on sparse and dense ground truths.



- Permutation test gave more false positives when the ground truth had denser GC.
- GMM method showed slightly more false positives in both sparse and dense ground truths.

Results & Discussion

The performance of GMM method and permutation test on data from ground truth models with **sparse** and **dense** GC.

Sparse

Performance indices	20000 time points				50000 time points			
		Permutation test				Permutation test		
	GMM	Simple	Bon	B-H	GMM	Simple	Bon	B-H
ACC	0.9558	0.9297	0.9922	0.9803	0.9933	0.9447	0.9936	0.9869
TPR	1	1	1	1	1	1	1	1
TNR	0.9512	0.9223	0.9914	0.9782	0.9926	0.9389	0.9929	0.9856
FPR	0.0488	0.0777	0.0086	0.0218	0.0074	0.0611	0.0071	0.0144
FNR	0	0	0	0	0	0	0	0

Dense

Performance indices	20000 time points				50000 time points			
		Permutation test				Permutation test		
	GMM	Simple	Bon	B-H	GMM	Simple	Bon	B-H
ACC	0.9472	0.9444	0.9781	0.9614	0.9869	0.9386	0.9672	0.9533
TPR	1	0.9949	0.9667	0.9898	1	0.9898	0.9411	0.9795
TNR	0.9218	0.9201	0.9835	0.9477	0.9807	0.9140	0.9798	0.9407
FPR	0.0782	0.0799	0.0165	0.0523	0.0193	0.0860	0.0202	0.0593
FNR	0	0.0051	0.0333	0.0102	0	0.0102	0.0589	0.0205

- Both methods performed worse on dense ground truths but permutation test showed significantly drop in performance.

Results & Discussion

The computation time (seconds) of GMM method and permutation

Average computation time (sec)	GMM method	Permutation test
20000 time points	3.1948	278.0139
50000 time points	4.5904	667.2358

Method	Computation time
Permutation test	$(1 + nP)T_{SSID}(N) + (1 + nP)T_{GC}$
GMM method	$(1 + N_0)T_{SSID}\left(\frac{N}{N_0}\right) + (1 + N_0)T_{GC} + T_{fitGM}$

N = length of time series data
P = number of permutations used in permutation test
n = number of dimensions in time series data
N_0 = number of GC samples used in GMM method
$T_{SSID}(n)$ = computation time of subspace identification on data of length n
T_{GC} = computation time of calculating GC matrix
T_{fitGM} = computation time of fitting GMM

- Permutation test required much more computation time than GMM method since, in general, $nP \gg N_0$ and $T_{SSID}(N) > T_{SSID}(N/N_0)$.

Conclusion

- Higher number of permutations in Monte-Carlo permutation test gives higher performance.
- Monte-Carlo permutation test is more preferable as it allows using p-value correction methods which yield higher performance.
- Overestimating order of state-space model does not hinder the performance as much as underestimating.
- On sparse ground truths, permutation test performs better than GMM method but the difference can be reduced by increasing the length of the time series data.
- Both permutation test and GMM methods perform worse on dense ground truths when compared to sparse ground truths. The decrease in performance is significant in permutation test.
- Permutation test requires much more computation time than GMM method.

Reference

[1] J Songsiri., Learning brain connectivity from EEG time series. Technical report, Chulalongkorn University, 2019. https://jitkomut.eng.chula.ac.th/pdf/eeg_bc_final_jss.pdf.

[2] P. Van Overschee, and B. De Moor. *Subspace identification for linear systems: Theory—Implementation—Applications.* Springer Science & Business Media, 2012.

[3] L. Barnett and A. K. Seth. Granger causality for state space models. *Physical Review E*, vol. 91(4), 2015.

[4] T. E. Nichols and A. P. Holmes. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1):1-25, 2002.

[5] P. Good. Permutation, parametric, and bootstrap tests of hypotheses. Springer Science & Business Media, 2005.

[6] B. Efron. *Large-scale inference: empirical Bayes methods for estimation, testing and prediction*, volume 1. Cambridge University Press, 2012.