A Study of brain connectivity using statistical methods

การศึกษาการเชื่อมโยงของสมองด้วยวิธีเชิงสถิติ

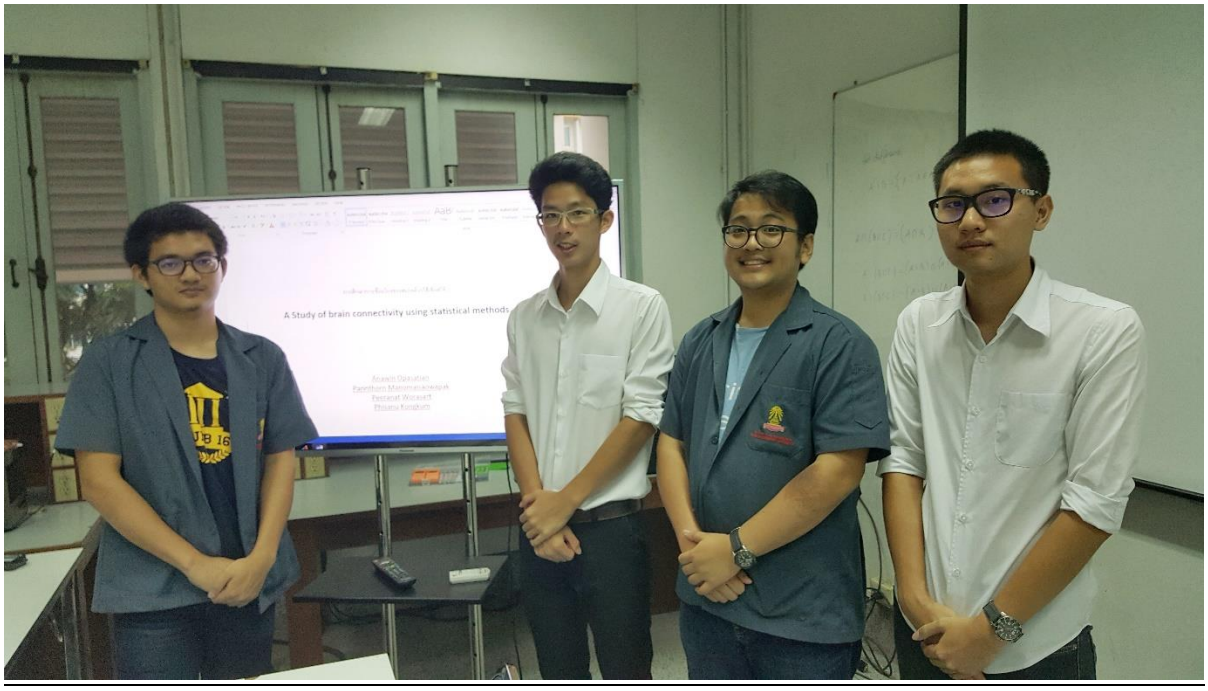
Anawin Opasatian
Parinthorn Manomaisaowapak
Peeranat Worasart
Phisanu Kongkam

นายอนาวิล โอภาสเสถียร, นายปรินทร มโนมัยเสาวภาคย์, นายพีรณัฐ วรศาสตร์ และ นายพิษณุ คงขำ

Summer project report
Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

Preface

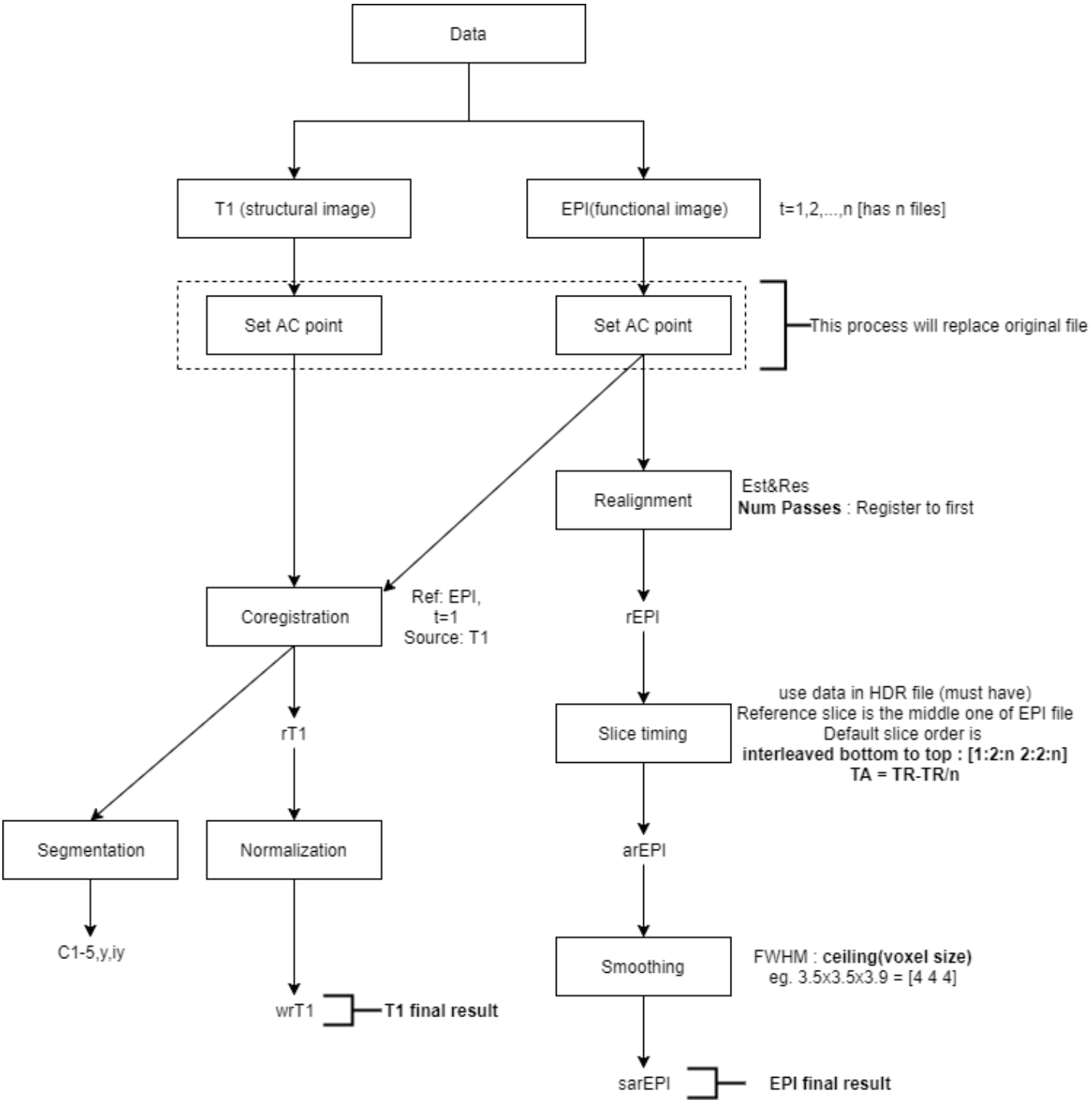
This report is a study of brain connectivity using statistical methods including hypothesis test of partial correlation matrix, sparse inverse covariance matrix estimation by Graphical Lasso and estimation by Granger Causality. The data we use to study are matrix of time-series from fMRI scan image that pre-processed by SPM12. The SPM12 software package has been designed for the analysis of brain imaging data sequences. The sequences can be a series of images from different cohorts, or time-series from the same subject.



Part I: Preprocess using SPM12

The data received from the fMRI scanner cannot be used to define the connectivity of the brain. There are many factors that can alter the result, the major ones are differences in brain size, head movements, unequal subject space, slice timing and unwanted parts of head. These problems are believed to be fixed by steps of preprocesses using the SPM software.

This is the workflow of steps in SPM12.



1. Set origin

Why?

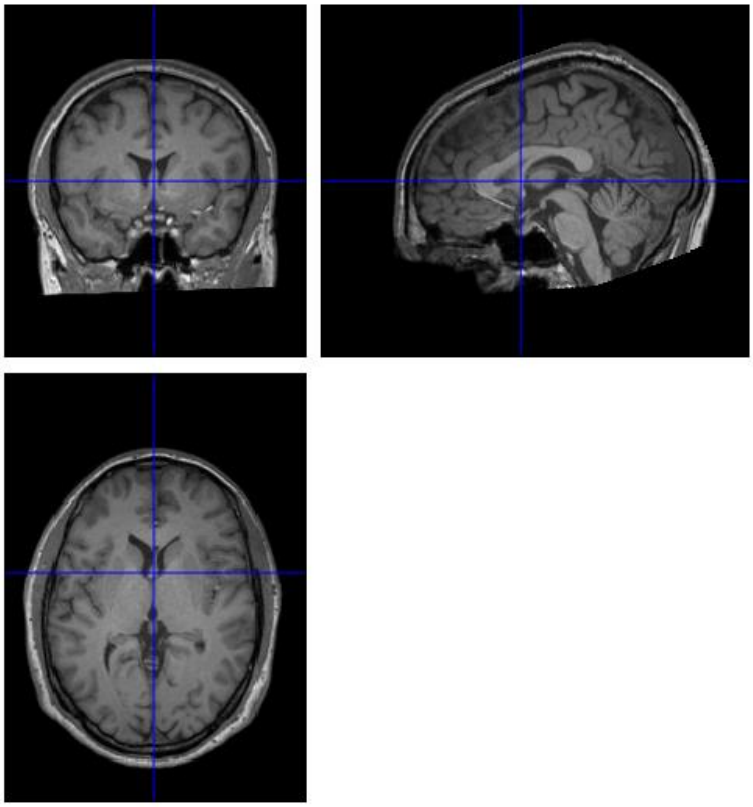
Because each subject has different brain size. When we try to pre-process brain data we have to adjust brain image using image processing to make every brain in the same position but each subject has different brain size and will have different head movement. Every brain image has different origin point. So, if we don't choose same point for every subject there will be problem in image processing. Like rotate in different axis. This may be analogous to the oscilloscope measurement must be measured from common ground.

Dr. Widhaya Sungkarat advice that we should select origin point at Anterior Commissure.

How?

In SPM12 there is no function that can set origin directly but there is function that work as the same called Reorientation. These are the steps

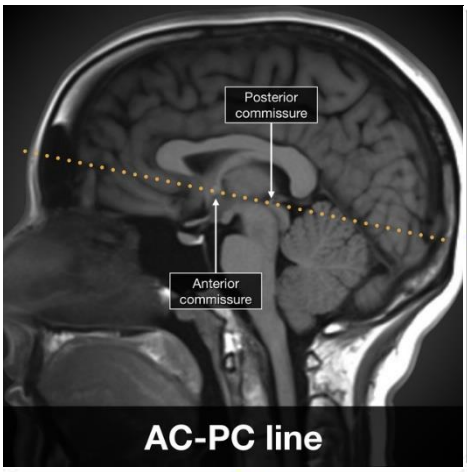
1. Display MRI scan. (Select EPI image input 1: timepoints in filter box which default is 1)
2. Choose origin point manually.

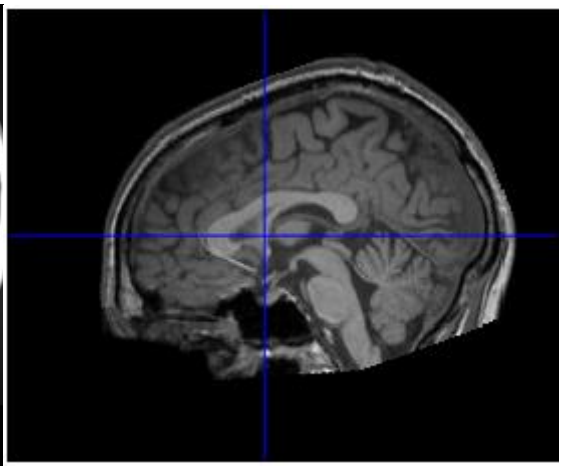


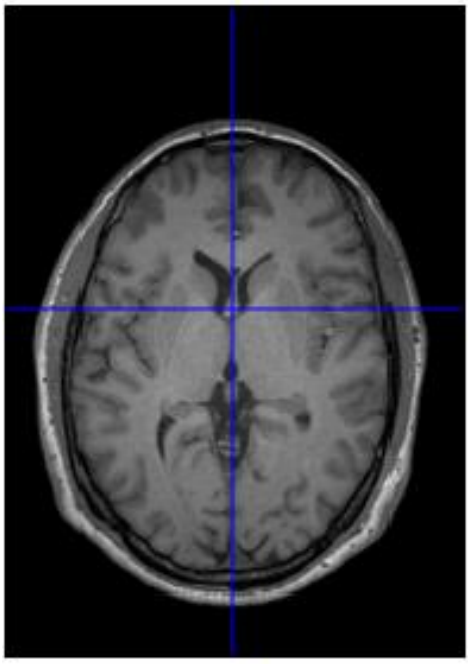
3. Calculate translation by subtracting new coordinate by default origin point and multiply result with voxel size.
4. Input value from step 3 to translation tab then check by eyes. The minus signs may be different, change until it is at the point on step 2.
5. Set origin then Reorient all EPI files

Apply to T1 image too. This process will replace files.

Tips: If you are not familiar in indicating AC point, you can rotate the image to make crosshair line cut through AC and PC point by enter the angle in radian in the box "pitch" and "yaw" first. You can see in the image below that there is a plane where AC and PC point lies together. Then rotate the image back after finding the AC point.


Posterior commissure
Anterior commissure
AC-PC line





2.Segmentation

Why?

The fMRI scans yield results that include skull and tissues which will be obstacle for our analysis. The Segment function will separate brain image to 3-5 images according to SPM version. In SPM12, the 5 images result in this process consist of white matter, grey matter, Cerebrospinal fluid, skull and tissues. We can use each image that was separated for masking or in other purpose. In SPM12 segment function also can give us the deformation matrix that can warp the subject brain into the template brain and it inverse. These deformation matrixes have y_ and iy_ prefix file which can be used in forward and inverse normalize.

How?

1. Click Segment button in SPM.

2. Select the image that you want to segment.

3. In case that you want forward or inverse deformation matrix select what you want in the bottom of batch windows.

4. In SPM12, Segmentation will result in 5 files which end with C1 to C5. C1 is white matter. C2 is grey matter. C3 is Cerebrospinal fluid. C4 is bone and skull. And C5 is other tissues.

3.Realign

Why?

No matter how we constrain the head during the scan there always still a small head movement that can cause falsify the data since our voxel is as small as $3*3*3$ mm^3. Realignment will fix this issue by apply rigid transformation (6 parameters: 3 translations & 3 rotations) to make every EPI images align with the reference image (for MRI the first (t=1) image usually picked) using the least square approach.

How?

1. Enter Realign(Est&Res)
2. Change parameter in Num Passes: Register to first

The output files will add "r" as prefix

4.Coregister

Why?

Before using the structural image as the representor, we must make sure that the T1 image match our EPI images first, since it was produce in difference time. Coregistration will corrected motion between T1 image and EPI t=1 image and warp image to same dimension. In our study, we mainly focus on EPI. Hence, we must preprocess EPI image least time as possible. So, we will coregister T1 with EPI t=1 as reference image. T1 will have lower resolution.

How?

1. Enter coregister(Est&Res)
2. Input Ref: EPI,t=1 Source: T1 and run

The output files will add "r" as prefix

5.Slice Timing

Why?

Since the 3D image file composed of many 2D slices which are not scanned instantaneously. The sequences of slices scanned are also mostly interleave (1 3 5 ... 2 4 6 ...). The problem arises because brain function with time. Time changes, the signal change. Slice timing will interpolate the correct value by interpolating in time domain.

How?

1. Read description from HDR files because some data we will know only if it described in HDR.
2. Enter Slice timing
3. Data: all EPI files, number of slices, TR, Slice order are in files description(HDR)

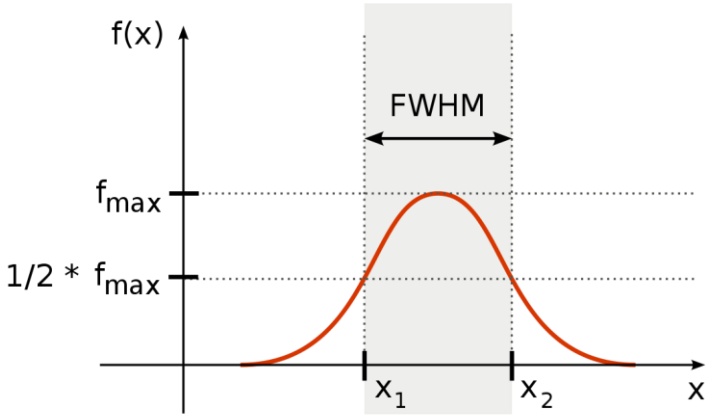
   TA = TR- [TR/number of slices]; default slice order is interleaved bottom to top (input as vector [1:2:number of slices 2:2:number of slices])

The output files will add "a" as prefix

6.Smooth

Why?

Every image file we scan have noise. This step is to suppress noise by convolving image volumes with a Gaussian kernel of a specified width, calculated using FWHM (Full width half maximum), to guarantee that it applies to all images by choose integer that is the maximum of voxel size to be FWHM. E.g. voxel size [3.5 3.5 4.5] we choose 5 to be FWHM.



How?

1. Select images to smooth (EPI t=1: ...).
2. Enter the full-width at half maximum (FWHM) of the Gaussian smoothing kernel which depend on voxel size (x y z).

The output files will add "s" as prefix

7.Normalization

Why?

Since every subject brain have difference size, they cannot be compare together in spatial space. It is necessary to compare those in standard (template) space, which in this step, there are few templates for us to choose: ICBM European Brain, ICBM East Asian Brain, Mean average Brain. However, in our study we only need the deformation matrix that transform images from ROIs space back into the subject space so we only need the inverse of deformation matrix.

How?

1. Select images to align with (the representor image should be the mean image.)
2. Select images to write (image to apply deformation matrix to).

The output files will add "w" as prefix

Part II: Determining brain connectivity

There are several statistics tools in determining brain connectivity. We were assigned to use three tools: Partial Correlation, Graphical Lasso and Granger Causality. The first two treat data as Random Variables while the latter treats data as time series.

The idea of the partial correlation and graphical lasso is the same: to estimate inverse covariance matrix which can be used to interpret value of partial correlation. Each entry of the partial correlation matrix will tell how well each element are depends on another one in case that effects from other elements are cutoff. The more the value, the more the correlation and vice versa. In the case of no correlation, the value in the entry must be 0 only. It is nearly impossible for an inverse sample covariance matrix to have entry that has an exact value of zero but it is possible to rounded to zero using hypothesis test. We set null hypothesis(H0): the entry is zero and alternative hypothesis(H1): the entry is not zero. For both methods, first calculate the sample covariance matrix from the time series by multiply the matrix of ROIs base with time points by its transpose. Note that we should set all the mean of the data to zero first. The covariance matrix will tell how well an element correlate to one another. Directly find it inverse along with the p-value of each entry. MATLAB use t-statistics to compute p-value from partial correlation matrix the output of function partialcorr(x) are partial correlation matrix and p-value matrix in same dimension. Set zero all entry that has p-value less than 0.05. to achieve inverse covariance matrix.

Graphical lasso is sparse inverse covariance matrix (precision matrix) estimator. Log-likelihood function of inverse covariance matrix (which sample is drawn from normal distribution) is given by $\log|\Theta| - trace(S\Theta) - \lambda||\Theta||_1$ [1,2] subjected to sparse inverse covariance matrix. But optimizers typically minimize function, so we should minimize the negative of log-likelihood function instead of maximizing it. In log-likelihood function. S is sample covariance matrix, $||\Theta||_1$ is L1 norm and it is weighed by regularization parameter which is the lambda. The more lambda is, the sparser precision matrix becomes. Because we are going to minimize the negative log-likelihood function: $trace(S\Theta) - \log|\Theta| + \lambda||\Theta||_1$. As we see, sparsity can be loosely seen by magnitude of L1 norm. more sparse means more zero. More zero means lower L1 norm. when we increase lambda L1 norm weight is increase. If we are going to minimize the function, L1 norm must decrease. Intuitive way to understand the idea is an example in extreme case. If we set lambda to a million, possible lowest L1 norm is just sum of absolute value in diagonal entries (other entries are zero) or inverse covariance matrix is a diagonal matrix. If we are going to estimate precision matrix. We must choose lambda value that give us best estimation. This is model selection problem. There are several methods in model selection, AIC, BIC, Cross Validation (CV) for example. In this case BIC is preferred because it has penalized term which prevents overfitting of the model. BIC score is calculated by $BIC = n\left(\log\left|\hat{\theta}\right| + trace\left(S\hat{\theta}\right)\right) + \mathrm{NZ} * \log(n)$ [3,4]. which n is number of time points, NZ is number of nonzero entries in estimated sparse inverse covariance matrix ($\hat{\theta}$). Regularization parameter with lowest BIC score is preferred.

We use SLEP toolbox [5] for graphical lasso optimization. The program is available here: https://github.com/parinthorn/summer-project

In Granger Causality models, it can be the case that x has influence on y and not vice versa. To begin, we write the Autoregressive model or AR models $y(t) = \begin{bmatrix} A_1 \, A_2 \, \dots \, A_p \end{bmatrix} \begin{bmatrix} y(t-1) \\ y(t-2) \\ \dots \\ y(t-p) \end{bmatrix} + \varepsilon(t)$ when p is the number of lags and $\varepsilon(t)$ is the noise. Consider the two models:

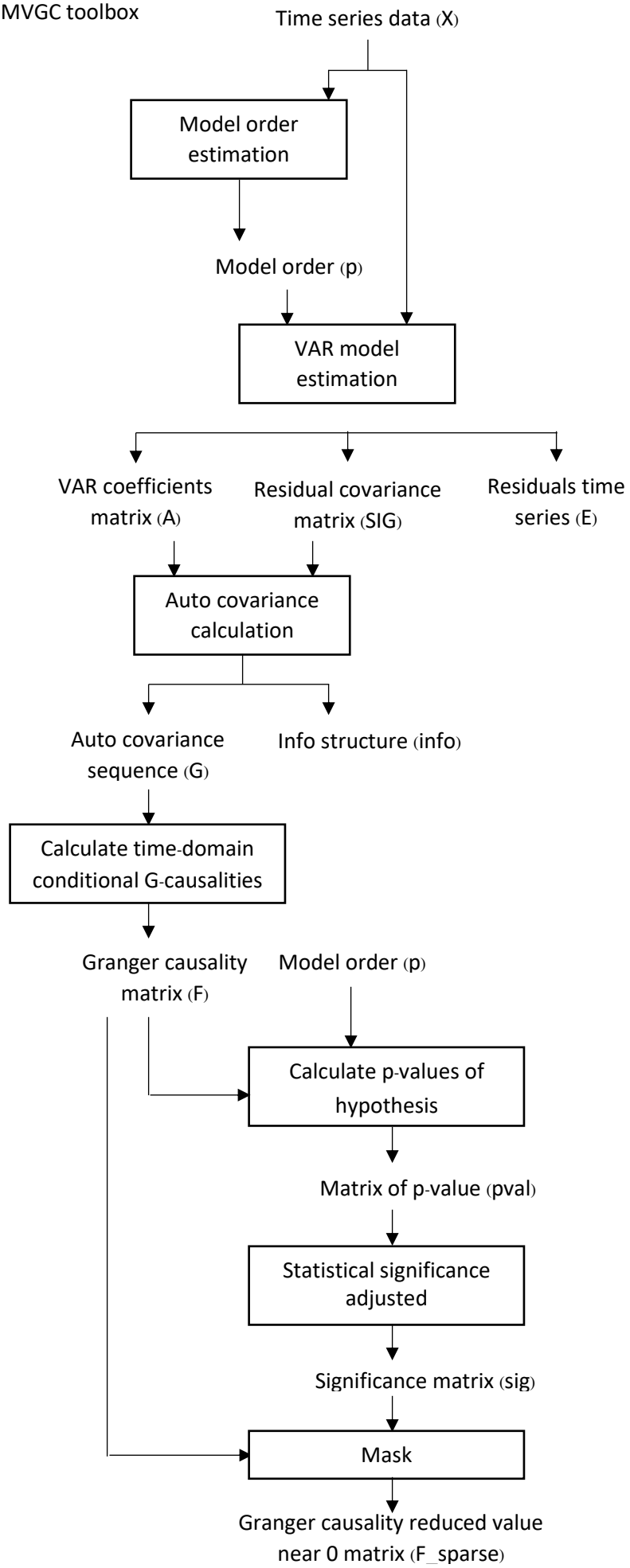
$$M_1 : x(t) = Ax^{(p)}(t-1) + \varepsilon(t)$$

$$M_2 : x(t) = \tilde{A}x^{(p)}(t-1) + \tilde{B}y^{(r)}(t-1) + \tilde{\varepsilon}(t)$$

The bivariate Granger causality from y to x is quantified by the log-likelihood ratio $F_{y \to x} = ln \frac{var(\varepsilon)}{var(\tilde{\varepsilon})}$ [6].

We use MVGC toolbox [7] to calculate Granger Causality models.

This is the flow of MVGC toolbox



1. Model order estimation: tsdata_to_infocrit

This step calculate the best model order and return as AIC (Akaike information criterion) and BIC (Bayesian information criterion)

Syntax : [aic,bic,moaic,mobic] = tsdata_to_infocrit(X,morder,regmode,verb)
Input : X is multi-trial time series data
morder is maximum VAR model order or vector of model
regmode is regression mode that can be LWR or OLS
verb is verbosity flag that can be true or false

Output : aic/bic is vector of AIC/BIC values

moaic/mobic is optimal model order according to AIC/BIC

2. VAR model estimation: tsdata_to_var

Estimate VAR model parameters for the selected model order

Syntax : [A, SIG, E] = tsdata_to_var(X,p,regmode)
Input : X is multi-trial time series data
p is model order
regmode is regression mode that can be LWR or OLS
Output : A is VAR coefficients matrix
SIG is residuals covariance matrix
E is residuals time series

3. Auto covariance calculation: var_to_autocov

Calculate the auto covariance sequence G according to the VAR model

Syntax : [G,info] = var_to_autocov(A,SIG,acmaxlags,acdectol,aitr,maxiters,maxrelerr)
Input : A is VAR coefficients matrix from VAR model estimation step
SIG is residuals covariance matrix
acmaxlags is maximum auto covariance lags to calculate
acdectol is auto covariance decay tolerance
aitr is use "accelerated iterative" Lyapunov solver algorithm
maxiters is maximum iterations
maxrelerr is maximum relative error
Output : G is auto covariance sequence
Info is info structure, with fields about error and warnings

4. Calculate the time-domain conditional G-causalities: autocov_to_pwcgc

Calculate pairwise-conditional time-domain causality

Syntax : F = autocov_to_pwcgc(G)
Input : G is auto covariance sequence
Output : F is Granger causality matrix

For our research, we use data of brain activity with 116 ROIs but the MVGC toolbox can calculate granger causality matrix for only 30 variables. Due to the limit of this toolbox, we can't calculate brain connectivity for time-series with more than 30 ROIs. By the way we have a MATLAB code that can be used to calculate brain connectivity for time-series that less than 30 ROIs.

The program is available here (require MVGC toolbox): https://github.com/PeeranatW/BrainConn

REFERENCE

[1] Yuan, M., & Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. Biometrika, 94(1), 19-35.

[2] Friedman, J., Hastie, T. and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9 432-441.

[3] Gao, X., Pu, D. Q., Wu, Y., & Xu, H. (2012). Tuning parameter selection for penalized likelihood estimation of Gaussian graphical model. Statistica Sinica, 1123-1146.

[4] Nogales, F. J., Alonso, A. M., & Avagyan, V. (2014). Improving the graphical lasso estimation for the precision matrix through roots ot the sample convariance matrix (No. ws141208). Universidad Carlos III de Madrid. Departamento de Estadística.

[5] Liu, J., Ji, S., & Ye, J. (2009). SLEP: Sparse learning with efficient projections. Arizona State University, 6(491), 7.

[6] Pruttiakaravanich, A., & Songsiri, J. (2016). A Review on Dependence Measures in Exploring Brain Networks from fMRI Data. Engineering Journal, 20(3).

[7] Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. Journal of neuroscience methods, 223, 50-68.