Senior Project Report 2102499 Year 2017

A State-space model estimation of EEG signals using subspace identification

Satayu Chunnawong ID 5730569821

Advisor: Assist. Prof. Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering

Chulalongkorn University

Abstract

Studies of brain connectivity can bring relevant information about the activity from activated network using Autoregressive model is one of widely-used methods to analyze brain connectivity. However, this model do not completely describe EEG data components because EEG time series data may be consisted of Moving average (MA) component. In this project, we consider state-space model of EEG signals by using subspace identification and learn brain connectivity by using Granger causality. System matrices of state-space model were estimated to fit time series. Granger causality was used for learning brain connectivity from state-space model. We also used H_2-norm, H_∞-norm and pole position classified EEG signal into data in normal condition or in seizure condition. Because these features can indicate system frequencies and system output response which different in both conditions. All experiments conducted by using EEG data from the Children's Hospital of Boston. The results showed that when we estimated EEG data on AR model, Baysian Information Citerion tended to choose order 2. Moreover, fitting of time series from estimated state-space model is approximately 12.2% in average, comparing to actual EEG time series. This result caused unrealiability of EEG classification. Despite the fact that pole of seizure data have similar distribution compared to pole of normal data and system norm of seizure data had more energy than system norm of normal data. In addition, EEG model estimation caused the result of brain connectivity which we learned by using Granger causality test on state-space model and AR model.

Keywords: Brain connectivty, Subspace identification, Granger causality

Contents

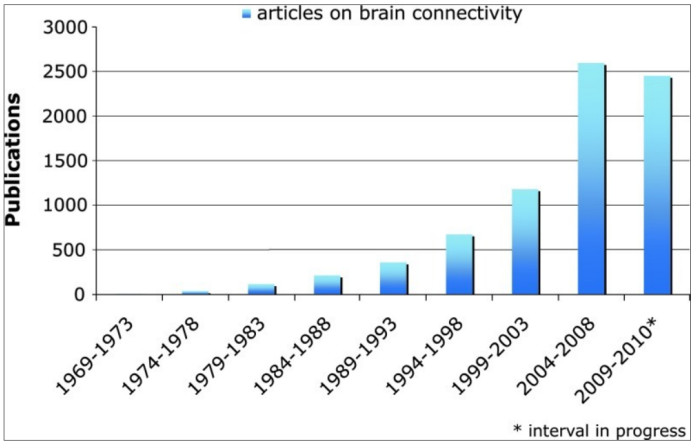


Figure 1: Number of publications in the PubMed database using the search term in "5-year increments". [1]

1 Introduction

Nowadays, there has been a growing interest in learning brain connectivity. According to number of brain connectivity publications indexed by PubMed (`https://www.ncbi.nlm.nih.gov/pmc/`), the number of publications is likely to be an exponential growth since 1969 [1] because the measure with associated signal processing are probably bring relevant information about the activity from activated network and also disrupted network that associated with tumors [2]. There are many methods to analyze how a group of neurons affects to the others, such as Dynamic causal modeling (DCM) and Granger causality (GC) [3]. One of widely-used method to analyze brain connectivity is Granger causality via Autoregressive (AR) model.

$$y(t) = A_1 y(t-1) + A_2 y(t-2) + \cdots + A_p y(t-p) + e(t) \tag{1}$$

where y is output signal which is EEG signal, e is measurement noise and A_i is matrix explains relationship between signal in the past. However, time series data for actual EEG data may have Moving average (MA) component so that pure AR modeling may not be sufficient for EEG signals [4]. According to [5] , the result of Granger causality test on model (1) explain only causality relationship between output signal (y) but indeed the objective is to find causality relations between sources generating in the brain which are cannot be directly measured cover through AR model.

This project focuses on identifying EEG sources in state space model which can be described not only AR model but also Autoregressive Moving Average with Exogenous input (ARMAX) model. A state space model have sufficient components of actual EEG data. With this advantage, we can analyze brain connectivity by using Granger causality on state space model so that we can learn causality relation between sources (x), not scalp signals that explore GC on the surface (y). A state-space model we represent in this project are shown in (2)

$$x(t+1) = \mathcal{A}x(t) + w(t) \tag{2a}$$
$$y(t) = Cx(t) + v(t) \tag{2b}$$

In fact, we do not have information about parameters in state space model. The only information that we know is EEG time series data. For this reason, our study will also focus on subspace identification to identify system matrices for state space model. Subspace identification is a tool which is used for estimating state sequence and system matrices of model. This method is very useful because we can estimate all unknown state space variables and parameters with only prior knowledge (time series data). Moreover, system parameters of state-space can be used for classifying EEG data in two conditions which are normal condition and seizure condition because system parameters indicate system characteristics, such as system frequencies, magnitude of output response and peak gain of system.

2 Project Overview

2.1 Objectives

The objective of this study are the following

1. To estimate linear EEG model described by state space model using subspace method.
2. To learn brain connectivity for EEG signal by using Granger causality test on state space model.

2.2 Scope of Work

The scope of this project are as follow :

1. We consider a brain connectivity comparison of Granger causality test
2. We consider linear time invariant model only and using estimation method based on least square.
3. Experimental results mainly consists of
   (a) Show performance of Granger causality test from Autoregressive model compare to state space model on simulated data set.
   (b) Learn brain connectivity of real EEG data sets from two conditions : normal and seizure.

2.3 Expected Outcomes

1. Schemes for estimating state space models that describe brain relationship of variables from time series.
2. MATLAB codes for the proposed scheme.
3. Comparison results of brain connectivity learned from two groups of EEG time series.

3 Methodology

In this project, we estimate of EEG time series to system matrices of state space model in terms of (2) where $x \in \mathbf{R}^n$ is brain source with n nodes at time t , $y \in \mathbf{R}^m$ is an EEG measurement contains m sources at time t which is only data we know , $\mathcal{A} \in \mathbf{R}^{n \times n}$ denotes parameters explains relationship of brain sources in the past , $C \in \mathbf{R}^{m \times n}$ is unknown parameters describe relationship between EEG measurement and brain sources , $w(t)$ and $v(t)$ are internal noise from source and noise from measurement, respectively and $w(t) = Kv(t)$ where $K \in \mathbf{R}^{n \times m}$ is state disturbance matrix. The variables and parameters describe in the following table.

Table 1: Variables and parameters in this project

Variables		Parameters	
Measured	Unknown	Known	Unknown
$y(t)$	$x(t)$		$\mathcal{A}$
	$v(t)$		C
			K

We use subspace method to estimate state space model since only EEG measurement (y) are known and brain source (x) are not measured. The assumptions for parameters are that $(\mathcal{A}, C)$ are observable, $(\mathcal{A}, K)$ are controllable and noise covariance is positive definite. After that, we examine Granger causality test from any estimated state space model. The result from GC test can refer to brain

connectivity.

Moreover, we focus on finding differences of EEG data in two conditions (normal and seizure condition). We observe estimated parameters from state space. System norm shows energy of the system. We can also observe parameters by calculating pole position of systems. Pole position gives information about system oscillation.

3.1 Stochastic subspace method

We estimate sources and system matrices (in this case : $\mathcal{A}$, C,W,V) by using stochastic subspace method. The estimation process starts by estimating sources. Since, EEG linear model have no input so that the estimation will use stochastic subspace method. In this method we focus on estimate state sequence first. The process starts by dividing data by time to obtain past data and future data. Then, project the future output (Y_f) onto the past output (Y_p) space with zero initial state ($\hat{X}_0 = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \end{bmatrix}$) [6].

$$\mathcal{O}_i \triangleq Y_{i\,|\,2i-1}/Y_{0\,|\,i-1} = Y_f/Y_p \tag{3}$$

where $\mathcal{O}_i$ is the oblique projection and $Y_{0\,|\,i-1}$ is measurement data from $t = 0$ to $t = i - 1$. After that, compute the state from single value decomposition (SVD) factorization.

$$\mathcal{O}_i = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_n V_1^T \tag{4}$$

Since $\mathcal{O}_i = \Gamma_i \hat{X}_i$ [7] and there are some non-singular matrix T that $\Gamma_i = U_1 \Sigma_n^{1/2} T$ so that we obtain

$$\hat{X}_i = \Gamma_i^{\dagger} \mathcal{O}_i \tag{5}$$

Then, estimate system matrices in least-square sense by forming the equation

$$\begin{aligned} \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i\,|\,i} \end{bmatrix} &= \begin{bmatrix} A \\ C \end{bmatrix} \hat{X}_i + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \\ \begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix} &= \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i\,|\,i} \end{bmatrix} \hat{X}_i^{\dagger} \end{aligned} \tag{6}$$

with noise covariance as

$$\begin{bmatrix} \hat{W} & \hat{S} \\ \hat{S}^T & \hat{V} \end{bmatrix} = (1/j) \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}^T \tag{7}$$

After process finished, we obtained $\hat{A}, \hat{C}, \hat{W}$ and $\hat{V}$. Finally, we calculated state disturbance matrix K by using relation between $\hat{W}$ and $\hat{V}$: $\hat{W} = K\hat{V}K^T$. After we obtained all system parameters, we used them for learning brain connectivity by using Granger causality on state space model.

3.2 Granger causality test on state space model

This process happens after structured system matrices are solved. We choose state space model to examine Granger causality. We use Granger causality test to examine brain connectivity of estimated model. In this process, Before the process starts, we assume that there is no measurement noise and $w(t)$ is also uncorrelated. In this Granger causality test on state space, we examine if y_j is a cause for y_i by removing y_j from the model. To remove y_j from the model, we force j^{th} row of C from (2) be zero so that full model become reduced model. Then, determine residual error of both models. Finally, we determine log ratio of residual error of x_i for each model [8].

$$\mathcal{F}_{y_j \to y_i \,|\, \text{All others } y} = \log \frac{|\Sigma_{ii}^R|}{|\Sigma_{ii}|} \tag{8}$$

where $|\Sigma_{ii}^R|$ and $|\Sigma_{ii}|$ are prediction error covariance of x_i for reduced model and full model, respectively. Also, both Σ_{ii}^R and Σ_{ii} are calculated from optimal mean-squared error estimation which is derived by Kalman filter.
The result of Granger causality test : when $\mathcal{F}_{y_j \to y_i \mid \text{All others } y} > 0$ because $y_i(t)$ in full model usually have more fitting than $y_i(t)$ in reduced model and $\mathcal{F} = 0$ means $\Sigma_{ii}^R = \Sigma_{ii}$. Therefore, $y_j(t)$ does not cause $y_i(t)$.

In this process, we examine Granger causality test in two models : full model and reduced model. For full model, we assume $y(t)$ has the same dimension as $x(t)$. That means we force number of measurement sources equals to number of brain sources. This means we let $C = I$. To reduce the full model, we assume each $C_j = 0$ (column j of C) which means we assume that value x_j does not cause all others x (since we assume that $y(t)_j = x(t)_j$ for all j) . Therefore, $y(t)$ is linear combination of all x except for x_j.

$$\text{Full model :} \qquad z(t+1) = \mathcal{A}z(t) + w(t) \ , \ y(t) = Cz(t) \tag{9a}$$

$$\text{Reduced model :} \qquad z(t+1) = \mathcal{A}z(t) + w(t) \ , \ y(t) = C^R z(t) \tag{9b}$$

where $C = I$ and C^R is reduced matrix that the j^{th} row of C is zero.
After that, we find estimation error covariance : $\Sigma = Cov(z - \hat{z}_{t \mid t-1})$ of both model. To obtain optimal prediction error covariance, we estimate $\hat{z}$ by using minimum mean square error because with this method, the error from noise is minimized. Therefore, we will get $\hat{x} = \mathbf{E}\{x_t \mid x_{t-1}^-\}$where y_{t-1}^- is all output data from the past up to time $t-1$ and $\hat{x} = \mathbf{E}\{x_t^R \mid x_{t-1}^{R-}\}$ for reduced model. After this, we calculate estimation error covariance by using Kalman Filter [9] because of optimal method in linear model form :

$$\begin{aligned} \hat{x}_{t+1 \mid t} &= \mathcal{A}\hat{x}_{t \mid t-1} + \mathcal{A}\Sigma_{t \mid t-1}C^T(C\Sigma_{t \mid t-1}C^T + R)^{-1}(y_t - C\hat{x}_{t \mid t-1}) \\ &= \mathcal{A}\hat{x}_{t \mid t-1} + K(y_t - \hat{y}_{t \mid t-1}) \end{aligned} \tag{10}$$

where $K = \mathcal{A}\Sigma_{t \mid t-1}C^\intercal(C\Sigma_{t \mid t-1}C^T + V)^{-1}$ is Kalman gain K from (2) and w_t can be expressed by $y_t - \hat{y}_{t \mid t-1}$ where $\hat{y}$ is estimated by MMSE ($\hat{y} = \mathbf{E}\{y_t \mid y_{t-1}^-\}$) and for reduced model we will get ε^R from $y_t^R - \mathbf{E}\{y_t^R \mid y_{t-1}^{R-}\}$. From (10), time update gives a recursive solution. Therefore, we can express measurement and time update of Σ as Riccati recursion [9].

$$\begin{aligned} \Sigma_{t+1 \mid t} &= \mathcal{A}\Sigma_{t \mid t-1}\mathcal{A}^T + W - \mathcal{A}\Sigma_{t \mid t-1}C^T(C\Sigma_{t \mid t-1}C^T + V)^{-1}C\Sigma_{t \mid t-1}\mathcal{A}^T \\ &= \mathcal{A}\Sigma_{t \mid t-1}\mathcal{A}^T + W - \mathcal{A}\Sigma_{t \mid t-1}C^T(C\Sigma_{t \mid t-1}C^T)^{-1}C\Sigma_{t \mid t-1}\mathcal{A}^T \end{aligned} \tag{11}$$

(Assume that V is zero)

From (11) , this equation is the optimal way to find state prediction error covariance [10]. However, we assume observation noise covariance is positive definite, $(\mathcal{A}, C)$ are observable and $(\mathcal{A}, W)$ are controllable so that we can solve steady state Kalman filter instead. The estimation of steady state Kalman filter satisfies Discrete Algebaric Riccati Equation (DARE) :

$$\Sigma = \mathcal{A}\Sigma\mathcal{A}^T + W - \mathcal{A}\Sigma C^T(C\Sigma C^T)^{-1}C\Sigma\mathcal{A}^T \tag{12}$$

There are two methods to examine Granger causality. The first method is to find log ratio of covarience of prediction error (Σ from solving of Riccati equation). Then, we suggest to determine the time-domain Granger causality shown as [8]:

$$\mathcal{F}_{x_j \to x_i \mid \text{All others } x} = \log \frac{|\Sigma_{ii}^R|}{|\Sigma_{ii}|} \tag{13}$$

where $|\Sigma_{ii}^R|$ and $|\Sigma_{ii}|$ is estimation error covariance of x_i for reduced model and full model, respectively. In general, Σ_{ii}^R is usually larger than Σ_{ii} because full model contains more information than reduced model so that the estimation error covariance of x_i is small when it has more information. If the result is zero, it means $|\Sigma_{ii}^R| = |\Sigma_{ii}|$ implying x_j does not affect x_i conditioning to all others x. Otherwise,

the value is always positive because reduced model is come up with more covariance magnitude.

Another method to examine Granger causality is to find the coefficient [8]

$$C_i \mathcal{A}_c^k K_j \qquad \text{where } k = 0, 1, \ldots, p-1 \tag{14}$$

Denote $\mathcal{A}_c$ as state observer closed loop observer gain. Therefore, $\mathcal{A}_c$ yields the necessary and sufficient condition by the Cayley-Hamilton Theorem. (See Appendix 7.2)

3.3 System norm calculation

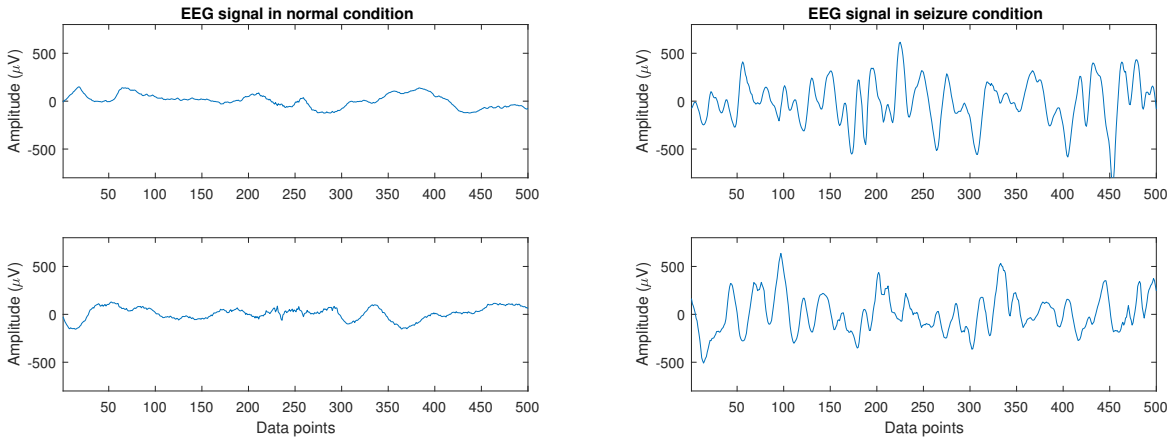
We determine systems which are time-invariant, linear and causal. Given H as transfer function of the system so that we can formulate the systems as time-domain convolution.

$$y = H * u \tag{15}$$

where $H(z) = C(zI - A)^{-1}B$
The system norm is energy of dynamical system which can be calculated either $\|H\|_2$ or $\|H\|_\infty$. $\|H\|_2\|$ gives root mean square of impulse response of system. In this case, $\|H\|_2$ are given by [11]

$$\|H\|_2 = \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Trace}[H(e^{j\omega})^G H(e^{j\omega})]\ \mathrm{d}\omega} \tag{16}$$

The other system norm is calculated by $\|H\|_\infty$ which is peak amplitude of bode plot of H.
Since system norm depends on magnitude, we use system norm for classifying EEG data into 2 conditions in Figure (2)
EEG signal at seizure condition have more magnitude than EEG signal at normal condition so in



(a) Sample of EEG signal in normal condition

(b) Sample of EEG signal in seizure condition

Figure 2: Example of EEG time series data

the following experiment, we expected that peak amplitude or system gain of seizure EEG signal are significantly more than normal EEG signal. In this project, we determined system gain by using both H_2-norm and H_∞-norm. We use MATLAB function `norm` to compute H_2-norm and use MATLAB function `hinfnorm` to compute H_∞-norm.

4 Data description

The data that we use in this project are collected at the Children's Hospital of Boston. Each case contains one hour of digitized signals. All signals were sampled at 256 samples per second with 16-bit resolution. There are 23 channels by the International 10-20 system of EEG electrode positions. Each position is labeled with a letter and a number. The letter means area that electrode lied [12]. For example, F7 means node number 7 at Frontal lobe area [13].

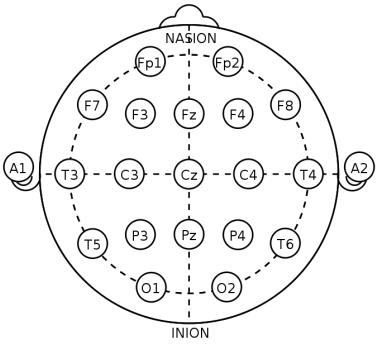


Figure 3: Electrode locations of International 10-20 system for EEG recording. The letters F,T,C,P and O stand for frontal, temporal, central, parietal, and occipital lobes, respectively.

Before the experiments started, we had to observe each case of EEG data. The case we selected for all experiments have to satisfy the following conditions.

1. The case contained both normal data with parts of data in seizure condition.
2. Normal data and seizure data in each case had obvious time series dynamic.
3. Normal data had little disturbances (the amplitude of normal data are not large as seizure data)

One of tools to observed EEG data is MATLAB program called biosig of which user interface are shown as Figure 4. Once the cases are chosen, some of project experiments required only a part of data or data

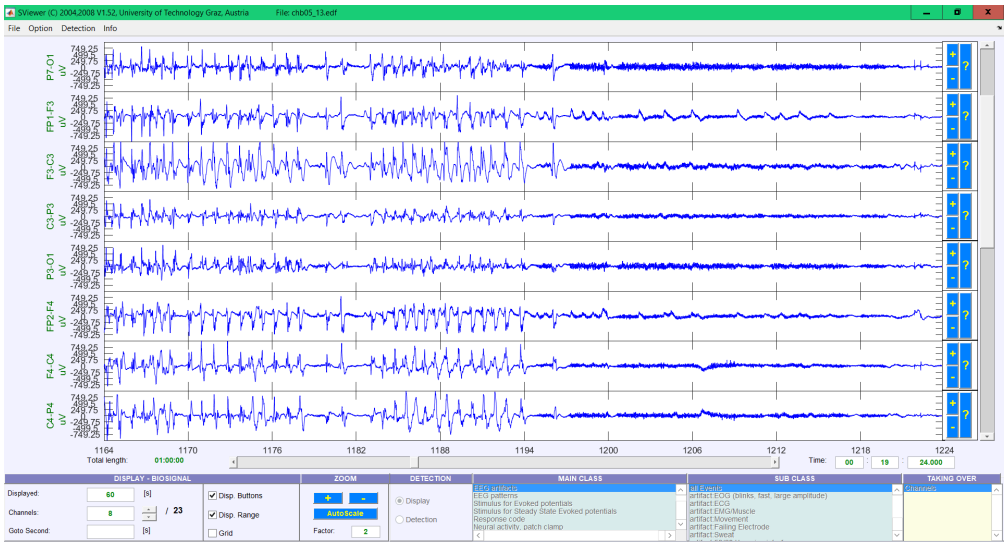


Figure 4: User interface of biosig on MATLAB

with only an interval so that we had to divded data in each case into trials. Each trial are 4 seconds long with 1024 time points and contains only one condition of signal (either normal data or seizure data) and there are 20 trials for each condition.

5 Experimental results

5.1 Model estimation of EEG signals

In this section, we assumed data from ground truth model as EEG signals. Then, we estimated the data into two forms; state-space model and linear time series model. Before the estimation starts, we divided EEG data into trial. Each trial contains 4 seconds (1024 data points) of EEG data. Also, each trial contains EEG signal only one condition (normal condition or seizure condition). We expect that the result of the estimation are likely to be nearly the same as ground truth model data in terms of model characteristic and model fitting.

State-space model In each trial, we used subspace method to estimate system parameters. We used `n4sid` function on MATLAB. This function creates a plot which suggest suitable model order. The plot contains Hankel singular values for models with various order. Moreover, we forced stability on estimated parameters so that we can calculate system gain in next experiment (otherwise, system gain is infinity). From `n4sid` function, most of trials choose order 25 so that we choose system order 25 for each trial. After we obtained system parameters, we plotted estimated time series. Then, we compared time series data with EEG data. The following figures are the comparison of EEG data estimation results. Both Figures (5) and (6) contain first 9 trials from the same channel and each trial contains

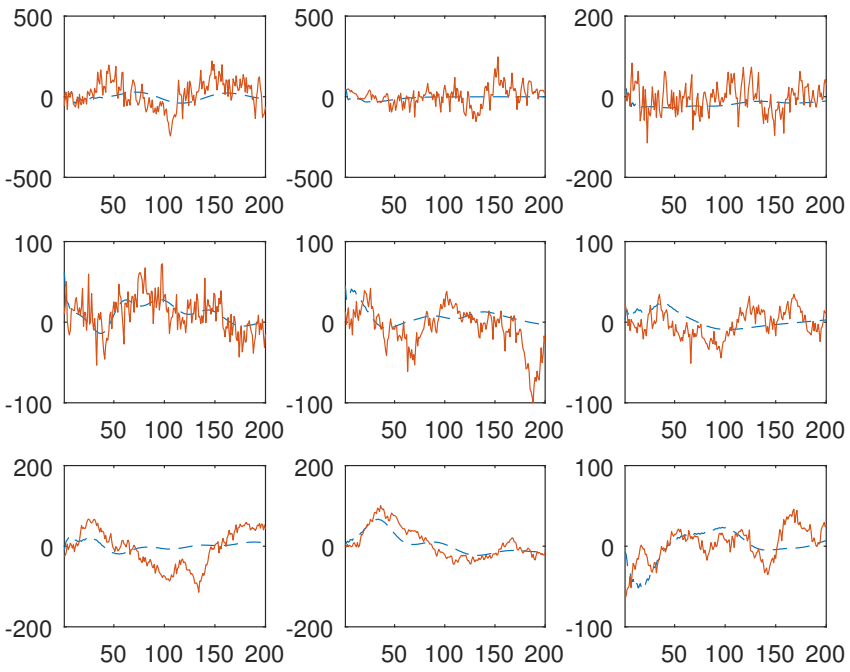


Figure 5: Fitting results of using state-space models on EEG data in *normal* condition (Line: EEG data, Dashed: estimated data)

first 200 data points.

The result showed that data from state-space with estimated parameters converged to zeros for both conditions. Estimated time series are likely to be more oscillate in seizure data results. The average fiiting for state-space model is 12.2%

Autoregressive model In each trial, we used maximum likelihood estimation to determine AR coefficients and AR order. First, we assumed that EEG data can be written as Autoregressive model with various order. After that, we selected model order by using Baysian Information Citerion (BIC).

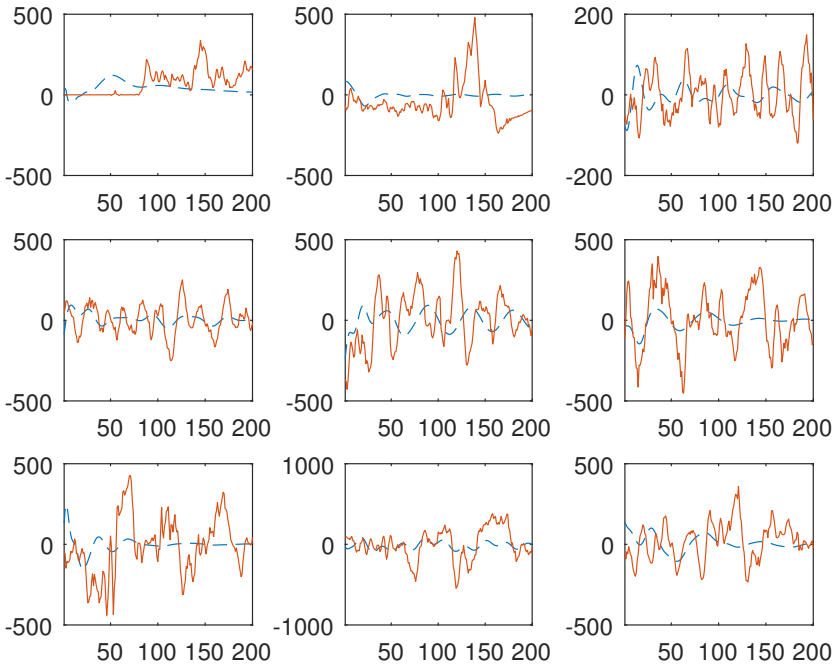


Figure 6: Fitting results of using state-space models on EEG data in *seizure* condition (Line: EEG data, Dashed: estimated data)

$$\mathrm{BIC} = -2\mathcal{L} + d\log N \tag{17}$$

where $\mathcal{L}$ is log-likelihood objective function value. For AR model order p with n sources, d is number of parameters which is n^2p and N is complexity of model which is n^2.
Finally, we compare estimation results of EEG data comparing to AR model with chosen order.
After we used maximum likelihood estimation to all trials, we plotted BIC scores shown in Figure (7) Each point in Figure (7) means BIC score for each trial when we assumed trial as $\mathrm{AR}(p)$.The resulted

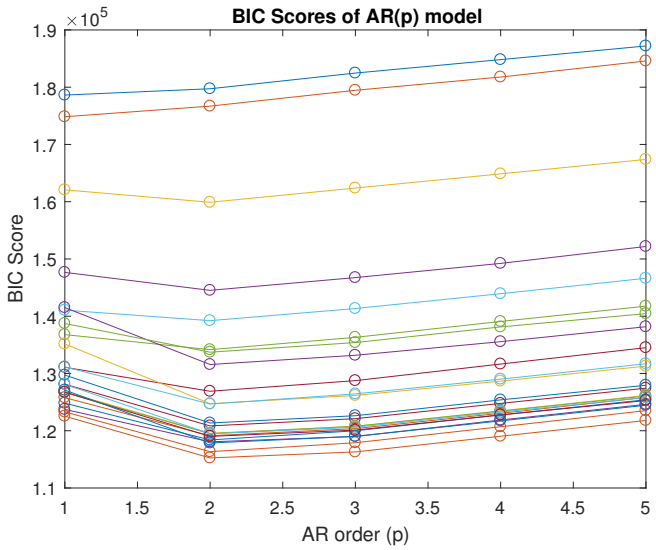


Figure 7: BIC scores of $\mathrm{AR}(p)$

show that $\mathrm{AR}(2)$ gives the best BIC scores for the most of trials. Next, we estimate all trials as $\mathrm{AR}(2)$. After that, we plotted estimated time series data and compared with actual EEG data. Both Figures

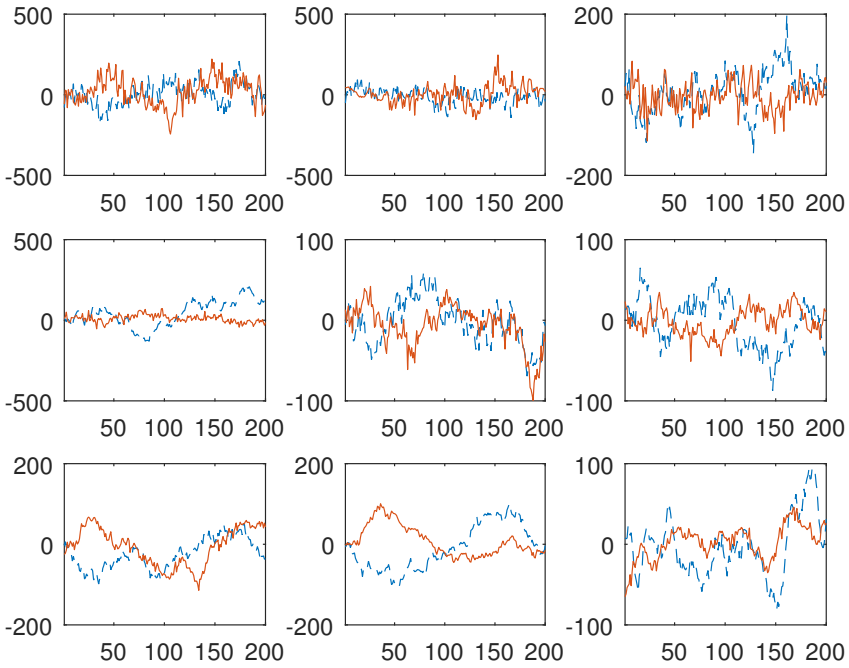


Figure 8: Fitting results of using AR models on EEG data in *normal* condition (Line: EEG data, Dashed: estimated data)

(8) and (9) contain first 9 trials from the same channel and each trial contains first 200 data points. Moreover, we compare fitting results of using AR model by using the same trials and same channel as we compare fitting on state space model.

The average fiiting for AR model is 31.8%

Discussion: The estimated EEG signals from state-space model are poor in terms of fitting. We presume that `compare` function in MATLAB that compare estimated model and EEG data is likely to not efficient. That because the result from estimated model are always deterministic and we do not know that the estimated model is accurate or not. Another reason is both functions have well fitting, if ground truth model have deterministic input. (See Appendix 7.3) The estimated EEG signals from AR model have higher fitting than one from state-space model but the estimation are still poor in terms of fitting. We presume that EEG signals are contained not only AR components but also have other component, such as MA component.

5.2 Classification of EEG data in normal and seizure condtion

The purpose of these experiments is to classify EEG data into two types : normal condition and seizure condition. For both experiments, We divided EEG data for each case into number of trials. Also, in each case, it contains EEG data with both conditions (Normal data with seizure interval). After that, we estimated each trial by using subspace identification so that we obtained state-space model. We used estimated state-space model for classifying EEG data. From EEG time series in figure (2), seizure data have more magnitude and contain only one frequency and normal data have less magnitude with higher multiple freqencies. Therefore, we used system norm to determine magnitude of the system and we used pole position to determine system frequencies.

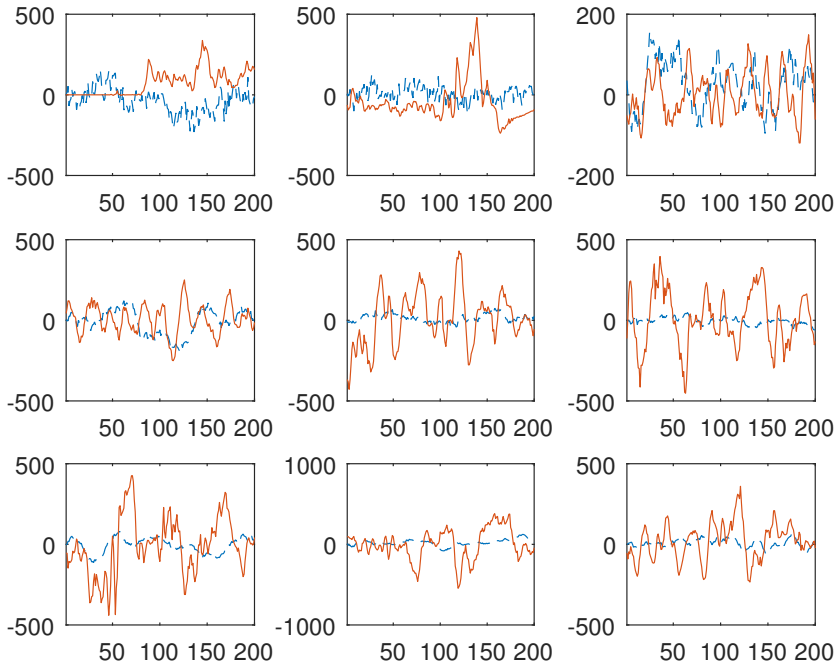


Figure 9: Fitting results of using AR models on EEG data in *seizure* condition (Line: EEG data, Dashed: estimated data)

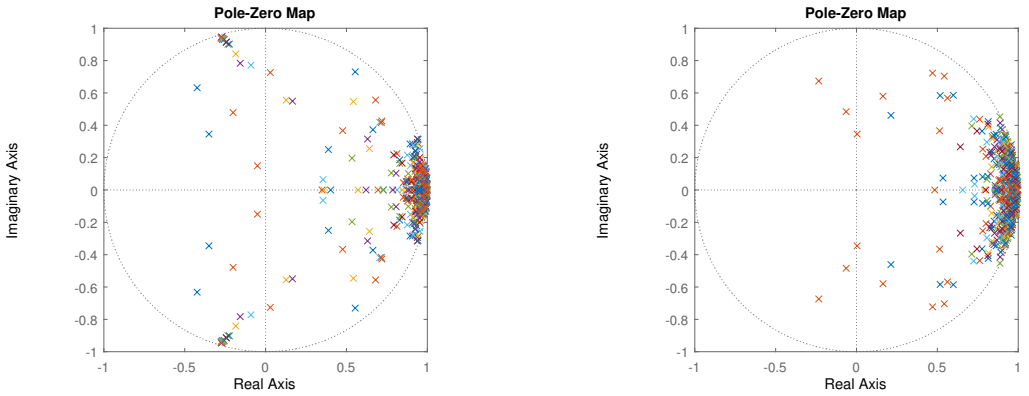
Our hypothesis is EEG data can be classified by differencing of pole position and by differencing of transfer function between noise and EEG data via H_2-norm.

Results on differences of pole position In dynamical systems, pole represents rate of decay and oscillation in time domain. In this experiment, we used estimated state-space model for determining pole position (There is no zero because of no input data). These state-space model are from divided EEG time series data. In each condition, we plotted pole position for each subtrial in the same figure so that we observed distribution of pole which tells characteristic of data. Then, we compared pole position between normal condition and seizure condition. From (2), we got relationship between EEG signal data (y) and brain sources (x) as

$$y(t) = CA^t x(0) + \sum_{n=0}^{t-1} A^{t-n-1} w(n) + v(t) \tag{18}$$

Since EEG signal data depend on A, pole position which is eigenvalue of A gives our information about decay rate and rate of oscillation. Decay rate can be observed by using magnitude of eigenvalue and oscillation rate can be observed by using angle between imaginary part and real part of eigenvalue. We expected that pole of seizure data show more oscillation compared to normal data which means pole of seizure data have more angle than pole of normal data. From figure 10, most of pole are located at unit circle. Distribution of pole location are similar in both conditions.

Results on differences of H_2-norm In dynamical system, H_2 gives a root mean square of system impulse response. This process starts after finishing state-space parameters estimation. In each condition, we calculated H_2 in each trial which means the power of output response. Then, we compared system norm between normal condition and seizure condition. We expected to see the differences of system gain frequency location since normal condition have lower output response than seizure condition.



(a) Pole of EEG data at normal condition

(b) Pole of EEG data during seizure

Figure 10: Pole position of EEG data

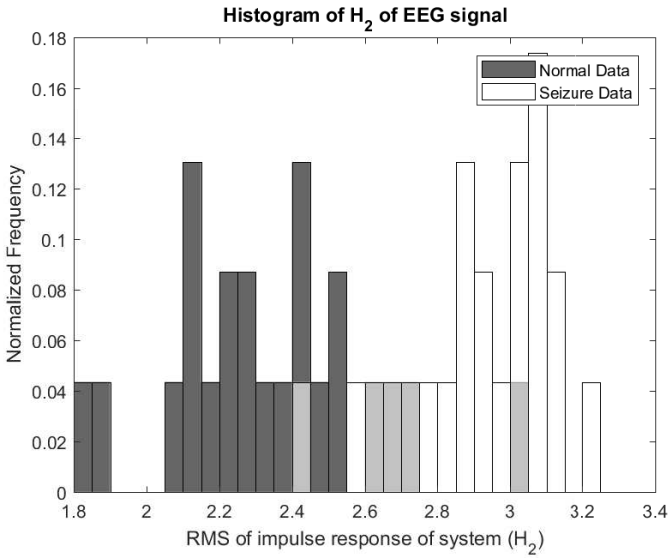


Figure 11: Histogram of H_2-norm for each trial

Results on differences of H_∞-norm Another feature to calculate system gain is H_∞-norm which means the largest value of the frequency response magnitude. We compared peak gain between system of normal condition and seizure condition. We expected to see that system of seizure condition have maximum gain at higher frequency than seizure condition maximum system gain.

Discussion: For pole position, most of poles from both conditions are located on the right bound of unit circle and pole distribution are similar for both condition. Therefore, we cannot classify EEG data by using pole position. For H_2-norm differences, Figure 11 shows that portion of trials that have higher output response are mostly from seizure data. From Figure 12, most of system of seizure data peak gain are higher than system of normal data peak gain but portion of trials that have higher peak gain is not obviously discriminate between two conditions. That because we use the system that estimated in terms of state space which are inaccurate from previous experiment. Therefore, we cannot confirm that differences of pole placement and differences of peak gain are inefficient.

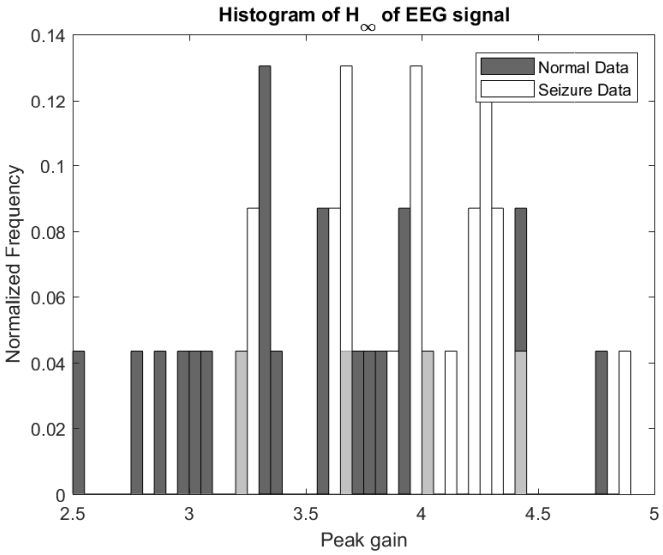


Figure 12: Histogram of H_∞-norm for each trial

5.3 Granger causality test on normal EEG data

In this section, we performed Granger causality on EEG data. We chose EEG data in normal condition. After finishing the experiments, we were able to know brain connectivity. There are three experiments in this section

1. We performed Granger causality test of $y(t)$ on state-space model. The result shows brain connectivity between EEG signal.
2. We performed Granger causality test of $x(t)$ on state-space model. The result shows brain connectivity between brain sources.
3. We performed GC test of $y(t)$ on estimated AR model. The result shows brain connectivity between EEG signal.

After we performed all three experiments, we discuss the results as the following :

1. Differences between Granger causality test of EEG sources $y(t)$ and brain sources $x(t)$.
2. Differences in terms of dynamic of Granger causality test on state-space model and the test on AR model. We expected that Granger causality test on state-space model gives result that captures more system dynamics than Granger causality test on AR model.

Granger causality test of $y(t)$ on state-space model In this experiment, we learned Granger causality test of EEG signal on state-space model. The process starts after we calculated state-space system parameters. First, we learned Granger causality on state-space model for each trial. For each trial, we observed $\mathcal{F}$. If $\mathcal{F}_{ij}$ is less than 10^{-3}, we assume that EEG signal y_j are not cause y_i. After we assumed causality of EEG signal, we repeated the process until we did every trials. For every element of $\mathcal{F}_{ij}$, we count number of trials that we assumed y_j causes y_i and plotted pairwise quantities on color map grid. The results of Granger causality test shown in Figure (13) where $y_1, y_2, \ldots, y_{21}$ refer to EEG channel.

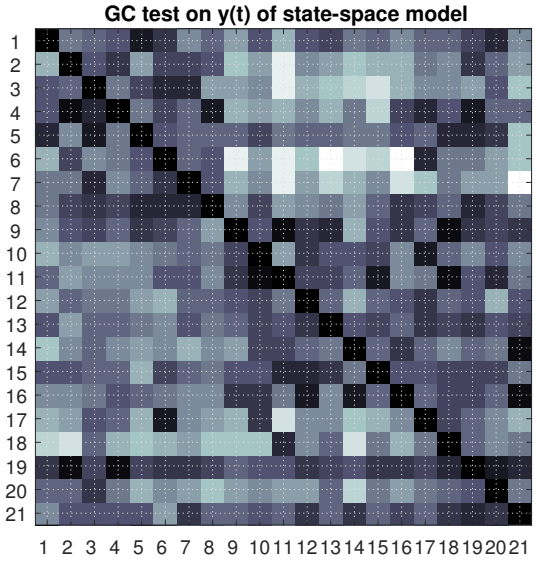


Figure 13: Granger causality test of $y(t)$ on state-space model

The cell (i, j) which is filled in black means y_j strongly causes y_i. That means the cell which is white filled means y_j slightly causes y_i.

Granger causality test of $x(t)$ on state-space model In this experiment, we learned brain connectivity of brain sources on state-space model. The process starts after we calculated state-space system parameters. From (2), the relationship between states is written as first order autoregressive model. Therefore, to performed GC test of $x(t)$ on state-space model is to performed GC test of $x(t)$ on AR model. For each trial, we observed $\mathcal{F}$ by using pairwise-conditional time-domain causality with prior information of autocovariance sequence. If $\mathcal{F}_{ij}$ is less than 10^{-3}, we assume that EEG source x_j are not cause x_i. After we assumed causality of EEG sources, we repeated the process until we did every trials. For every element of $\mathcal{F}_{ij}$, we count number of trials that we assumed x_j causes x_i and plotted pairwise quantities on color map grid.
The results of Granger causality test shown in Figure (14) where $x_1, x_2, \ldots, x_{25}$ refer to brain sources.

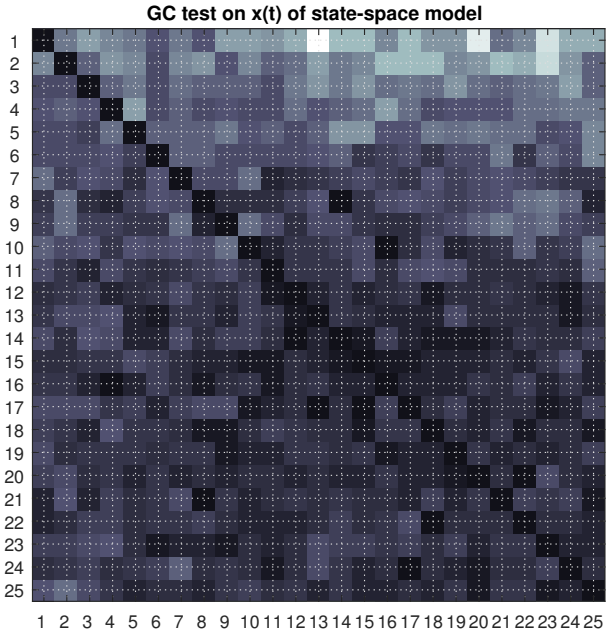


Figure 14: Granger causality test of $x(t)$ on state-space model

Granger causality test of $y(t)$ on AR model In this experiment, we learned Granger causality test of EEG sources on AR model. The process starts after we calculated coefficient of A_i for each trial. From previous experiments, we choose AR order 2 because of the best BIC scores. From GC test on AR model, if y_j is not cause y_i then [14],

$$(A_k)_{ij} = 0 \qquad \text{for all } k = 1, 2, \ldots, p \tag{19}$$

For each trial, we observed $(A_k)_{ij}$. We performed statistical test by using pairwise-conditional time-domain causality with prior information of autocovariance sequence. If $(A_k)_{ij} = 0$, we assume that EEG source y_j are not cause y_i. After we assumed causality of EEG sources, we repeated the process until we did every trials. For every element of $(A_k)_{ij}$, we count number of trials that we assumed y_j causes y_i and plotted pairwise quantities on color map grid.
The results of Granger causality test shown in Figure (14)

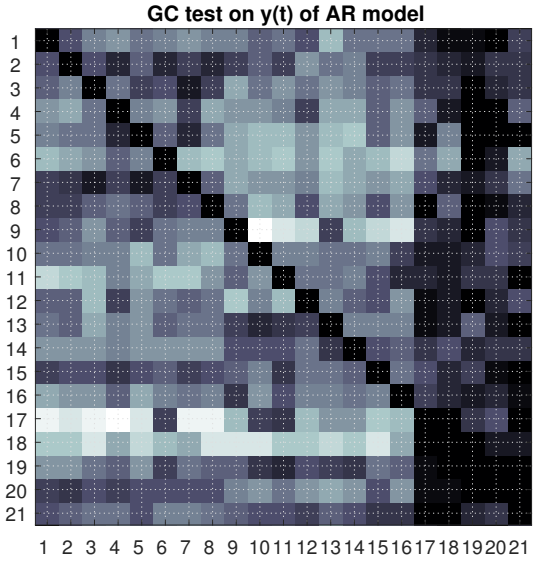


Figure 15: Granger causality test of $y(t)$ on $AR(2)$ model

Discussion: For Granger causality test on state-space model, each $\mathcal{F}_{ij}$ element of GC test of $y(t)$ shows less causality relation. That means there are some EEG channel which does not cause each others. In contrast, $\mathcal{F}_{ij}$ element of GC test of $x(t)$ shows more causality relation. That means most of brain sources causes each others. The reasons that causality relation are differences in two variables are probably that the EEG sources are measured on the surface of scalp so that signals can be corrupted by noise which may be affected distortion of causality relation. Moreover, the number of sources is not equal to number of EEG channel so that we cannot interpret the causality relation of $x(t)$. (We cannot find true position of brain sources)

6 Conclusions

This project aims to estimate EEG model described by state-space model and to learn brain connectivity for EEG signal by using Granger causality test. EEG data that we use are from the Children's Hospital Boston. We choose only the data with a little noise and obviously classify as normal data and seizure data. Moreover, we removed some channels that dependent to other channel. After we obtained satisfied EEG data, we divide the experiments into three parts. First, we estimate EEG data by using state-space model and by using AR model. The result showed that the estimated model are not accurate in fitting compare to real data. One of causes are the MATLAB function that compare simulated response are not accurate by using `compare` function in MATLAB. We use infer function instead. However, `infer` function works for univariate AR model which is not in our scheme. For AR model, the estimated model is AR order 2 according to BIC scores. The fitting of estimated AR model is slightly higher than estimated state-space model but it stills low in terms of fitting. That because EEG signals may also have other component, such as MA component. Second part is to classify EEG data into two types; normal condition and seizure condition. We divided EEG data into number of subtrials. We computed system from estimated state-space model from previous part. Then, we plot pole placement of system and also find peak gain (H_2-norm) of the system. The result showed that both pole position of normal data are similar to seizure data but system output response of normal data are obviously less than seizure data. It means seizure data has more magnitude than normal data but we cannot confirm that both algorithm are not suitable for classify EEG data since the estimated data are not accurate. The final part is to learn brain connectivity by using Granger causality test in 3 types. First, we learned brain connectivity of EEG signals ($y(t)$) on state-space model. Second, we learned brain connectivtiy of EEG sources ($x(t)$) on state-space model. Third, we learned brain connectivity of EEG signals ($y(t)$) on AR model.

References

[1] C. Pawela and B. Biswal, "Brain Connectivity : A New Journal Emerges," *Brain Connectivity*, vol. 1, no. 2, 2011.

[2] M. Hassan, O. Dufor, I. Merlet, C. Berrou, and F. Wendling, "EEG Source Connectivity Analysis: From Dense Array Recordings to Brain Networks," *PLoS ONE*, vol. 9, no. 8, 2014.

[3] K. Friston, R. Moran, and A. K. Seth, "Analysing connectivity with Granger causality and dynamic causal modelling," *Current Opinion in Neurobiology*, vol. 23, no. 2, pp. 172–178, 2013.

[4] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," *Physics Reports*, vol. 441, no. 1, pp. 1–46, 2007.

[5] A. Pongrattanakul, P. Lertkultanon, and J. Songsiri, "Sparse system identification for discovering brain connectivity from fMRI time series," in *SICE Annual Conference*, 2013.

[6] J. Songsiri, "EE531 - System Identification , Subspace Method." `http://jitkomut.eng.chula.ac.th/ee531/subspace.pdf`.

[7] P. van Overschee and B. de Moor, *Subspace Identification for Linear Systems*. Hoboken, New Jersey: Wiley, 1996.

[8] L. Barnett and A. K. Seth, "Granger causality for state space models," *Physical Review E*, vol. 91, no. 4, 2015.

[9] S. Boyd, "Kalman filter." `https://stanford.edu/class/ee363/lectures/kf.pdf`, 2009.

[10] G. Chow, *Analysis and Control of Dynamic Economic Systems*. John Wiley & Sons., 1975.

[11] K. Zhou and J. C. Doyle, *Essential Of Robust Control*. Prentice Hall, 1998.

[12] D. L. Schomer and F. H. L. da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincott Williams & Wilkins., 2004.

[13] The McGill Physiology Virtual Lab, "Biomedical signals acquisiton." `https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm`.

[14] A. Pruttiakaravanich and J. Songsiri, "A review on dependence measures in exploring brain networks from fMRI data," *Engineering Journal*, vol. 20, no. 3, pp. 207–233, 2016.

7 Appendices

7.1 Simplification of DARE applied to AR model

For AR model case, we examine GC test that $x_j(t)$ causes or does not cause $x_i(t)$ by comparing noise covariance of reduced model (Σ_{ii}^R : Noise covariance when we remove x_j from the model.) and full model by (**??**). When we determine GC test on state space model based on AR model (**??**), noise covariance can be calculated by using steady state Kalman filter that satisfies discrete Riccati equation (12). In this section, we demonstrate that the solution of discrete Riccati equation can be simplified to $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} W_1 & 0 \\ 0 & 0 \end{bmatrix}$ We can simplify Σ that we solve from into more similar form (12) by given $\left(W \in \mathbf{R}^{np \times np} \text{ and } W = \begin{bmatrix} W_1 & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & & 0 \end{bmatrix} \right)$

Given $\mathbb{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} \end{bmatrix}$ and $\Sigma \in \mathbf{R}^{np \times np}$ are in formation as : $\left[\begin{array}{c|c} U & V \\ \hline V^T & R \end{array}\right]$ where $U_{ij} \in$ $\mathbf{R}^{p \times p}$ denote the $(i,j)^{th}$ block of U
From

$$\Sigma = \mathcal{A}\Sigma\mathcal{A}^T + W - \mathcal{A}\Sigma C^T (C\Sigma C^T)^{-1} C\Sigma\mathcal{A}^T$$

we simplify in each term

$$\begin{aligned} \mathcal{A}\Sigma\mathcal{A}^T &= \left[\begin{array}{c|c} \mathbb{A} & A_p \\ \hline I & 0 \end{array}\right] \left[\begin{array}{c|c} U & V \\ \hline V^T & R \end{array}\right] \left[\begin{array}{c|c} \mathbb{A}^T & I \\ \hline A_p^T & 0 \end{array}\right] \\ &= \left[\begin{array}{c|c} \mathbb{A}U\mathbb{A}^T + A_pV^T\mathbb{A}^T + \mathbb{A}VA_p^T + A_pRA_p^T & \mathbb{A}U + A_pV^T \\ \hline U\mathbb{A}^T + VA_p^T & U \\ \vdots & \end{array}\right] \end{aligned} \tag{20}$$

$$\begin{aligned} \mathcal{A}\Sigma C^T &= \left[\begin{array}{c|c} \mathbb{A} & A_p \\ \hline I & 0 \end{array}\right] \left[\begin{array}{c|c} U & V \\ \hline V^T & U \end{array}\right] \left[\begin{array}{c} I \\ \hline 0 \\ \\ 0 \end{array}\right] \\ &= \text{First block column of } \left[\begin{array}{c} \mathbb{A}U + A_pV^T \\ \hline U \\ \\ \end{array}\right] \\ &= \left[\begin{array}{c} \sum_{i=1}^{p} A_i\Sigma_{i,1} \\ \hline \Sigma_{11} \\ \Sigma_{21} \\ \\ \Sigma_{p-1,1} \end{array}\right] \end{aligned} \tag{21}$$

$$C\Sigma\mathcal{A}^T = (\mathcal{A}\Sigma C^T)^T = \left[\begin{array}{c|ccc} \sum_{i=1}^{p} \Sigma_{1,i}A_i^T & \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1,p-1} \end{array}\right] \tag{22}$$

$$C\Sigma C^T = \Sigma_{11} \tag{23}$$

Then, combine all above terms (20), (21), (22) and (23) into DARE

$$\left[\begin{array}{c|c} U & V \\ \hline V^T & R \end{array}\right] = \left[\begin{array}{c|c} \mathbb{A}U\mathbb{A}^T + A_pV^T\mathbb{A}^T + \mathbb{A}VA_p^T + A_pRA_p^T & \mathbb{A}U + A_pV^T \\ \hline U\mathbb{A}^T + VA_p^T & U \end{array}\right]$$
$$+ \begin{bmatrix} W_1 & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & & 0 \end{bmatrix} - \left[\begin{array}{c} \sum_{i=1}^{p} A_i\Sigma_{i,1} \\ \hline \Sigma_{11} \\ \Sigma_{21} \\ \vdots \\ \Sigma_{p-1,1} \end{array}\right] (\Sigma_{11})^{-1} \left[\begin{array}{c|ccc} \sum_{i=1}^{p} \Sigma_{1,i}A_i^T & \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1,p-1} \end{array}\right] \tag{24}$$

From (20)

$$\begin{aligned} U\mathbb{A}^T + VA_p^T &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1,p-1} \\ \Sigma_{21} & \ddots & & \Sigma_{2,p-1} \\ \vdots & & \ddots & \vdots \\ \Sigma_{p-1,1} & & & \Sigma_{p-1,p-1} \end{bmatrix} \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_{p-1}^T \end{bmatrix} + \begin{bmatrix} \Sigma_{1,p} \\ \Sigma_{2,p} \\ \vdots \\ \Sigma_{p-1,p} \end{bmatrix} A_p^T \\ &= \begin{bmatrix} \sum_{i=1}^{p} \Sigma_{1,i}A_i^T \\ \sum_{i=1}^{p} \Sigma_{2,i}A_i^T \\ \vdots \\ \sum_{i=1}^{p} \Sigma_{p-1,i}A_i^T \end{bmatrix} \end{aligned} \tag{25}$$

$$\text{Determine } \Sigma_{21} \qquad \Sigma_{21} = (\text{First row blocks of } U\mathbb{A}^T + VA_p^T) - \Sigma_{11}(\Sigma_{11}^{-1})\sum_{i=1}^{p}\Sigma_{1,i}A_i^T$$
$$= \sum_{i=1}^{p}\Sigma_{1,i}A_i^T - \sum_{i=1}^{p}\Sigma_{1,i}A_i^T = 0$$

For the others i,$\Sigma_{2,i} = \Sigma_{1,i} - \Sigma_{11}(\Sigma_{1,1})^{-1}\Sigma_{1,i} = 0$. This means $\Sigma_{2,i} = 0$ for all $i = 1, 2, \dots, p$

$$\text{Determine } \Sigma_{31} \qquad \Sigma_{31} = (\text{Second row blocks of } U\mathbb{A}^T + VA_p^T) - \Sigma_{21}(\Sigma_{11}^{-1})\sum_{i=1}^{p}\Sigma_{1,i}A_i^T$$
$$= \sum_{i=1}^{p}\Sigma_{2,i}A_i^T = 0 \ (\Sigma_{2,i} = 0 \text{ for all } i)$$

For the others i,$\Sigma_{3,i} = \Sigma_{2,i} - \Sigma_{21}(\Sigma_{11})^{-1}\Sigma_{2,i} = 0$ where $i = 2, 3, \dots, p$. This means $\Sigma_{3,i} = 0$ for all $i = 1, 2, \dots, p$
Consequently, $\Sigma_{i,j} = 0$ for all $i = 1, 2, \dots, p, j = 1, 2, \dots, p$ except for Σ_{11}

$$\text{Determine } \Sigma_{11} \qquad \Sigma_{11} = A_1\Sigma_{11}A_1^T + W_1 - A_1\Sigma_{11}(\Sigma_{11})^{-1}\Sigma_{1,1}A_1^T$$
$$= W_1$$

The result of Riccati equation remains only block $\Sigma_{11} = W_1$. Therefore, it satisfies that $\Sigma = W$

7.2 $C\mathcal{A}_c^k K$ coefficients

The results of GC test on AR model (**??**) can be derived as coefficient $(A_k)_{ij} = 0, \forall k$ which means $x_j(t)$ does not cause $x_i(t)$. Meanwhile, the results of GC test on state space model (8) can also be measured by $C\mathcal{A}_c^k K$ coefficient [8]. When $C_i \mathcal{A}^k K_j = 0, \forall k,\ i = 1, \ldots, n-1$, it means $x_j(t)$ does not cause $x_i(t)$ In this section, we showed that the coefficient from GC test on AR model have same structure to the coefficient from GC test on state space based on ground truth AR model.

$$\text{Coefficient of GC test on AR model :} \qquad (A_k)_{ij} = 0\,, \forall k \tag{26a}$$

$$\text{Coefficient of GC test on state space model :} \qquad C_i \mathcal{A}^k K_j = 0\,, \forall k,\ i = 1, \ldots, n-1 \tag{26b}$$

First, we determine K from (10) based on ground truth AR model shown as :

$$\begin{aligned} K &= \mathcal{A}\Sigma C^T (C\Sigma C^T + V)^{-1} \\ &= \mathcal{A} \begin{bmatrix} \Sigma_{11}^T & \Sigma_{12}^T & \ldots & \Sigma_{1p}^T \end{bmatrix}^T \Sigma_{11}^{-1} \\ &\text{Because only } \Sigma_{11} \text{ is nonzero matrix} \\ &= \mathcal{A} \begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix}^T \\ &= \begin{bmatrix} A_1^T & I & 0 & \ldots & 0 \end{bmatrix}^T \end{aligned} \tag{27}$$

Given state observer closed loop observer gain $\mathcal{A}_c = \mathcal{A} - KC$. From (**??**) we have

$$\begin{aligned} \mathcal{A}_c &= \mathcal{A} - KC \\ &= \begin{bmatrix} 0 & A_2 & \ldots & A_p \\ 0 & 0 & \ldots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \ldots & I & 0 \end{bmatrix} \end{aligned} \tag{28}$$

Then, multiply by C on the left hand side and K on the right hand side :

$$C\mathcal{A}_c^k K = \begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} 0 & A_2 & \ldots & A_p \\ 0 & 0 & \ldots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \ldots & I & 0 \end{bmatrix}^k \begin{bmatrix} A_1 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{29}$$

when $k = 0$	CK	$= A_1$
when $k = 1$	$C\mathcal{A}_c K$	$= A_2$
when $k = 2$	$C\mathcal{A}_c^2 K$	$= A_3$
	$\vdots$	$\vdots$
when $k = p-1$	$C\mathcal{A}_c^{p-1} K$	$= A_p$

Because A_1 , A_2 , $\ldots$, A_p have the same structure so that if we assume $(A_1)_{ij} = 0$ that means $(\mathcal{A}_c)_{12} = 0$. Therefore, $\mathcal{A}_c = \mathcal{A} - KC$ yields the necessary and sufficient condition $C_i \mathcal{A}_c^k K_j = 0$ by the Cayley-Hamilton Theorem.

7.3 Estimation of time series data from state-space model

We estimated EEG signals from state-space model

$$z(t+1) = Az(t) + Bu(t) + Kv(t) \tag{30a}$$
$$y(t) = Cz(t) + Du(t) + v(t) \tag{30b}$$

From (30), we created all system matrices (A,B,C,D,K) and assume noise $v(t)$ as Gaussian distribution with zero mean and unit variance. Then, we obtained time series data from state-space. After that, we use time series data to perform subspace identification into estimated system matrices. In this experiment, we will use N4SID algorithm to calculate estimated system matrices with ground-truth model

We generated output data (y) from ground truth model with 4 conditions. First, the output data without deterministic input and noiseless. Second, the output data with deterministic input and noiseless. Third, the output data with only noise and the output data with both noise and detrministic output. As shown in Figure 16

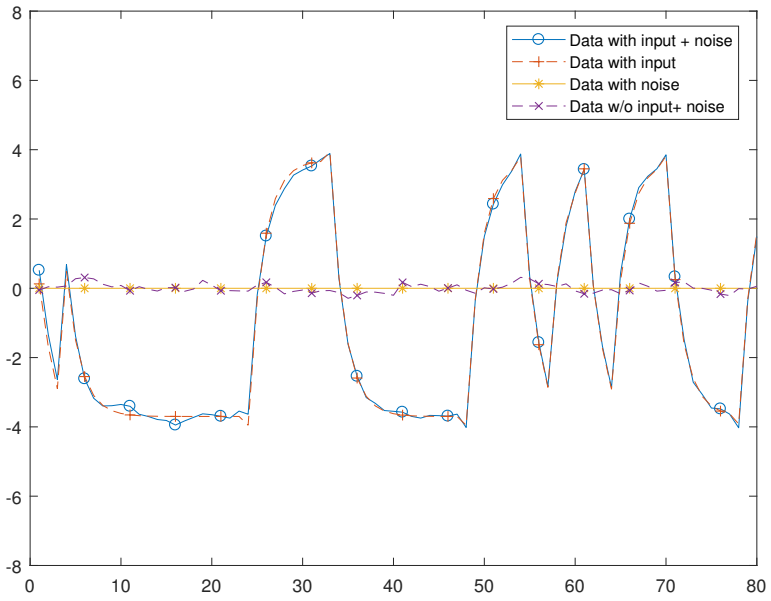


Figure 16: Time series data with all conditions

In N4SID algorithm, we divided into 2 cases. First, we performed N4SID from time series with deterministic input.

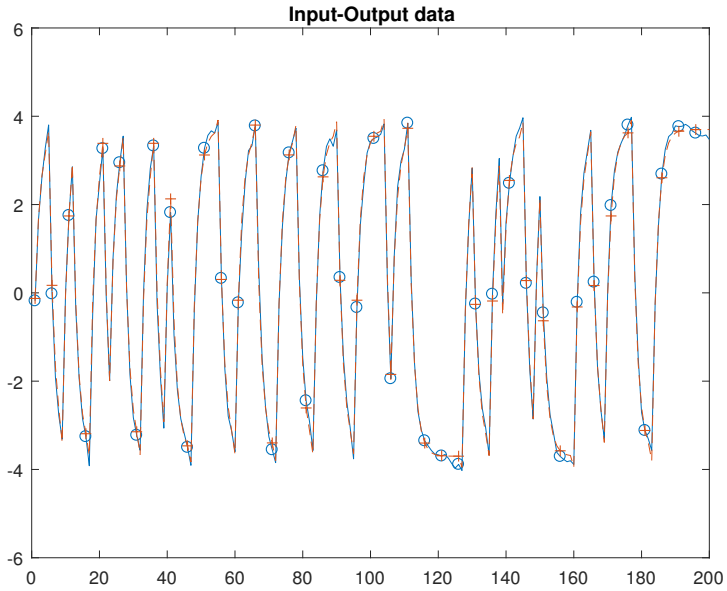


Figure 17: Time series data generated from ground truth state space model with deterministic input (Red line : with noise , Blue dash : without noise)

Then, we performed N4SID to the time series data. We choose time series data with noise and choose second order estimated model that is the same order as ground truth state space model. The result is shown in Figure 18.

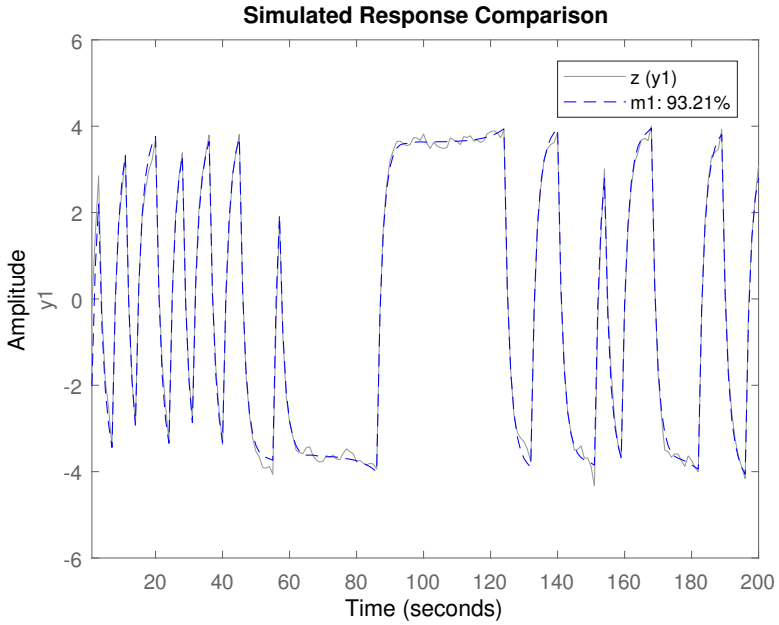


Figure 18: Comparison between time series data with noise and time series data from estimated state space model.

The result showed that N4SID can estimate time series data properly. Comparing to Figure 17, the estimated time series data is nearly the same as time series data from ground truth model without noise. Moreover, when we observed eigenvalues of ground truth model and estimated model. It shows that some modes of estimated model are near ground truth state space model.

In the second case, we performed the same method but we set B from ground truth state space model to zero ($B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$)

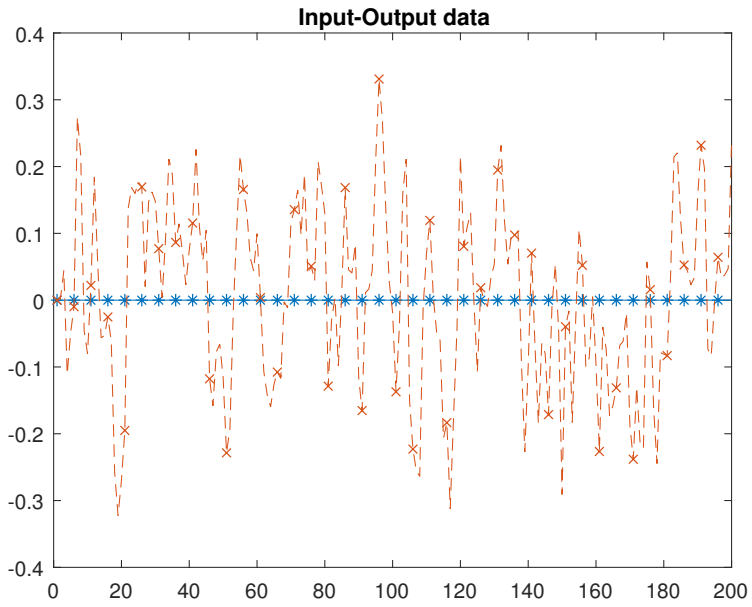


Figure 19: Time series data generated from ground truth state space model (Red line : with noise , Blue dash : without noise)

From Figure 19, time series data remains only noise despite nonzero initial state. Next, we use time series data with noise to estimate state space model by using N4SID method.

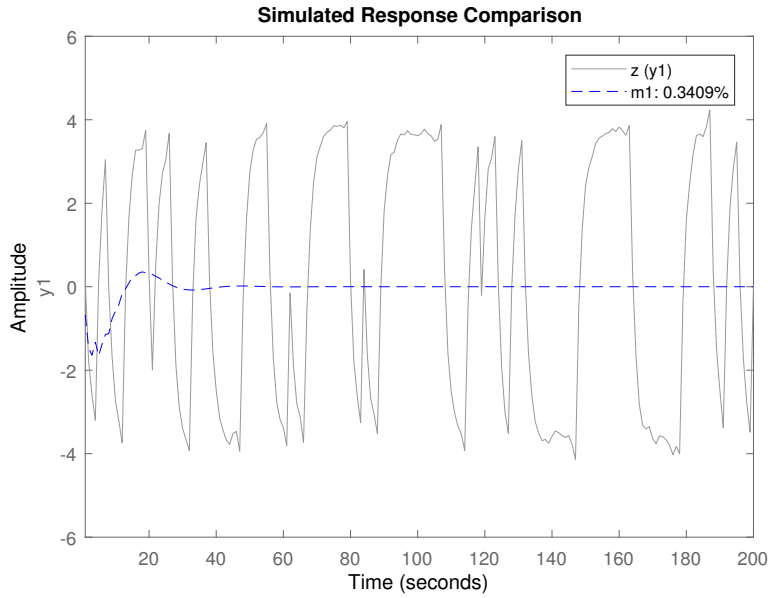


Figure 20: Comparison between time series data with noise and time series data from estimated state space model.

For eigenvalue of estimated model, it is obvious that eigenvalues of estimated model are different from eignevalues of ground truth state space model.

7.4 MATLAB function in this project

- **Data preparation:** We use this function for dividing a case of EEG data into trials, we have to defined data point that seizure condition occur and also defined duration for each trial.

```
% Developed by : Satayu Chunnawong
% Faculty of Engineering, Chulalongkorn University
%% DivideData : Divide EEG data in a case into trial.

% Input - structure variable type that contains
%       rec - actual EEG records
%       start_seizure - data point that seizure occurs
%       start_normal - normal EEG data point that user selects
%       end_seizure -  data point that seizure ends
%       window_time - time of EEG data for each trial
% Output -
%       NormalData - Data with length depends on window time.
%       SeizureData - EEG seizure data with length depends on window time.
%       NoOfTrial - Number of trial that already divided.
%       NoOfRecord - Number of data for each trial.

function [NormalData,SeizureData,NoOfTrial,NoOfRecord] = DivideData(Sample)

t = Sample.window_time;

% NoOfSubTrial depends on seizure length and window size so that we
% count subtrial for seizure case.

track = 0;


% Divide seizure data
% After data are divided, we neglect 1 second data after each trial.
for i = Sample.start_seizure:t+1:Sample.end_seizure
     track = track+1;
    SeizureData(:,:,track) = Sample.rec(:,i*256:(i+t)*256);
end

%Divide normal data by using the same number of SeizureData trials.
%After data are divided, we neglect 1 second data after each trial.
for i = 1:track
    NormalData(:,:,i) = Sample.rec(:,Sample.start_normal+(i-1)*t*256:Sample.
    start_normal+i*t*256);
end
NoOfTrial = track;
NoOfRecord = t*256+1;
```

- **State space model and AR model estimation:** After we completed preparing data, we use data from each trial to determine state space and AR model. For state-space model, we use `n4sid` function to estimate state space and use `varm.estimate` to determine $\mathrm{AR}(p)$ order of EEG model.

```
% Developed by : Satayu Chunnawong
% Faculty of Engineering, Chulalongkorn University
%-------------------------------------------------------------------
% State-space model estimation
% Output - cell which contains state space estimation for each trial
function [NormIddata,SeizureIddata,NormalModel,SeizureModel] = EEG_EstMdlSS(
    NormalData,SeizureData,NoOfTrial,NoOfRecord)

for i = 1:NoOfTrial
u = zeros(21,NoOfRecord);
%We neglect EEG channel 19 and 23 becuase depend on other channels
NormDat = NormalData(1:18,:,i);
NormDat = [NormDat; NormalData(20:22,:,i)];
SeizDat = SeizureData(1:18,:,i);
SeizDat = [SeizDat; SeizureData(20:22,:,i)];
yNormal = iddata(NormDat',u',1/256);

NormIddata{i,1} = yNormal;
ySeizure = iddata(SeizDat',u',1/256);
SeizureIddata{i,1} = ySeizure;

% We enforce stability on N4SID
opt = n4sidOptions('EnforceStability',true);
[EstMdl,x0] = n4sid(yNormal,25,opt);
[EstMdlSeiz,x0Seiz] = n4sid(ySeizure,25,opt);

NormalModel{i} = EstMdl;
SeizureModel{i} = EstMdlSeiz;
end
```

```
% Developed by : Satayu Chunnawong
% Faculty of Engineering, Chulalongkorn University
% ---------------------------------------------------------
% AR Estimation
function [NormAR,SeizAR,SIGNorm,SIGSeiz] = EEG_EstMdlAR(NormIddata,
    SeizureIddata,NoOfTrial,p)
% Determine BIC scores
logL = zeros(23,5);
for i = 1:NoOfTrial
yM = (NormIddata{i,1}.y)';
for j = 1:5
    Mdl = varm(21,j);
    [~,~,logL(i,j)] = estimate(Mdl,yM');
end
end

% Evaluate AIC and BIC scores for AR model
T = 21*1025;
for i = 1:23
numParam = zeros(5,1);
for j =1:5
    numParam(j) = 21*21*j;
end
[aic(:,i),bic(:,i)] = aicbic(logL(i,:)',numParam,T*ones(5,1));
end

% Determine AR coefficient of model
for i = 1:NoOfTrial
yM = (NormIddata{i,1}.y)';
    [NormAR{1,i},SIGNorm{1,i},ENorm{1,i}] = tsdata_to_var(yM,p);
end
for i =1:NoOfTrial
    yM = (SeizureIddata{i,1}.y)';
    [SeizAR{1,i},SIGSeiz{1,i},ESeiz{1,i}] = tsdata_to_var(yM,p);
end
```

- **State space model generation:** We use this function for generating ground truth state space model which is asymptotically stable. We have to defined number of state variables, number of output variables and number of input variables.

```
% Developed by : Satayu Chunnawong
% Faculty of Engineering, Chulalongkorn University
% ------------------------------------------------------------------
% First, we define pole of A (eigenvalue of A)
% There are complex eigenvalue when number of state is more than 3
% Otherwise, all eignvalues are real number.
% For example,
% the system that has 5 states contain one pair of complex eigenvalues.
% and the system with 6 states contains two pairs of complex
% eigenvalues

% We choose eigenvalues between -0.9 to 0.9 because we want to avoid
% oscillation poles.
function [A,B,C,D,x0] = GenStateSpace(Num_State,Num_input,Num_output)

PoleA(1:Num_State) = 0;
PoleA(1) = 1.8*rand() - 0.9;
if Num_State ~= 1
PoleA(2) = 1.8*rand() - 0.9;
end
for i = 3:Num_State
    if mod(i,2) == 0
        continue

   % Since we want coefficient of transfer function be all real number
   % So we make sure that number of state are even
   % Otherwise, the last eigenvalue is real number
    elseif mod(i,2) == 1 && i ~= Num_State
        r = 0.9*rand();
        theta = 180/pi*rand();
        PoleA(i) = complex(r*cos(theta),r*sin(theta));
        PoleA(i+1) = complex(r*cos(theta),r*sin(-theta));
    else
        PoleA(i) = 1.8*rand()-0.9;
    end
end
 P = diag(PoleA);
 Q = RandOrthMat(Num_State);
 % Since there are some eigenvalues which are complex number, we have to
 % reform matrix into equivalent matrix with all real number component
  [Q,P] = cdf2rdf(Q,P);
Q= RandOrthMat(Num_State);
% Transformation from diagonal matrix to full matrix
A = Q'*P*Q;

B = rand(Num_State,Num_input);
C = randn(Num_output,Num_State);
D = rand(Num_output,Num_input);
x0 = randn(Num_State,1);
```

- **Granger causality test:** After we completed estimating state space model, we defined system parameters. Then, calculate Granger causality function ($\mathcal{F}$) and also coefficient $C\mathcal{A}_c^k K$.

```
% This MATLAB function is developed by Nattaporn Plub-in
% (c) 2018 , Faculty of Engineering, Chulalongkorn University
%=========================================================================
%
%    This function for computing Granger causality (GC) for state-space
%    model as
%
%        x(t+1)  = Ax(t) + w(t)          w ~ N(0,W)
%        y(t)    = Cx(t) + v(t)          v ~ N(0,V)
%
%    PARAMETERS
%            A,C,W,V =   System matrices
%            FR      =   Granger causality via parameter reduction method
%            FCBK    =   Granger causality via C(A-KC)K condition [BaS:15]
%
%=========================================================================

function [FR,FCBK] = GC_SS(A,C,W,V)
    %------------------------------- Full model -------------------------------
    [P,~,K] = dare(A',C',W,V);  % solve RICCATI for full model
    K = K';
    Sigma = C*P*C';
    diagSigma = diag(Sigma);    % collect Sigma_ii of full model
    diagSigma(diagSigma==0) = 1; % imputed all zeros in Sigma to avoid inf and
    NaN
    %------------------------------ Reduced model -----------------------------
    %%
    %------------- solve RICCATI for all reduced model ----------------------
    [m,~] = size(C);
    for j=1:m                   % reduce jth parameter
        ind = 1:m;
        Creduce = C;
        Vreduce = V;
        Creduce(j,:) = [];      % force jth row of C to be zero
        Vreduce(j,:) = [];      % force jth row of v to be zero
        Vreduce(:,j) = [];      % force jth column of V to be zero
        [Preduce,~,~] = dare(A',Creduce',W,Vreduce);
        SigmaR = Creduce*Preduce*Creduce';
        diagSigmaR = diag(SigmaR);      % collect Sigma_ii of reduced model
        diagSigmaR(diagSigmaR==0) = 1;  % imputed all zeros in Sigma to avoid
    inf and NaN
        F(j,j) = 1;                     % Diagonal of GC set to be 1
        ind(j) = [];
        F(ind,j) = diagSigmaR./diagSigma(ind);
    end
    %%
    FR = log(F);                % GC from parameter reduction method
    FCBK = C*(A-K*C)*K;         % GC from C(A-KC)K condition
end
```