

A Comprehensive Method of Evaluating the Preprocessing Stage of Functional Magnetic  
Resonance Imaging (fMRI)

By

SHENGCHAO ZHANG

Thesis

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Electrical Engineering

May 11, 2018

Nashville, Tennessee

Approved:

D.Mitchell Wilkes, Ph.D.

Richard Alan Peters, Ph.D.

*To my parents*

## ACKNOWLEDGMENTS

I am grateful to Dr. D.Mitchell Wilkes, my advisor, for patiently guiding me on writing this thesis. Also, thanks to his excellent instructions over the entire duration of my master's degree, I made a lot progresses. It is my honor to continue as his student for my doctoral career.

I extend my appreciation to two senior PhD candidates in my group, Ms. Charreau S. Bell and Ms. Mohd Khairi Nazirah, for providing me with a smart algorithm and helping me to run the preprocessing of a large quantity of data. Because of their help, my thesis has become more complete, reasonable, and can be finished on time.

During my entire master's work, Dr. Richard Alan Peters helped me a lot and always gave me precious advice when facing some tough problems.

Finally, I appreciate all of the professional editors and writing consultants from Vanderbilt Writing Studio for aiding me to improve the writing quality for this thesis.

## TABLE OF CONTENTS

	Page
DEDICATION . . . . .	ii
ACKNOWLEDGMENTS . . . . .	iii
LIST OF TABLES . . . . .	vi
LIST OF FIGURES . . . . .	vii
1 Introduction . . . . .	1
1.1 Why fMRI . . . . .	1
1.2 Functional Connectivity . . . . .	1
1.3 Preprocessing is the Cornerstone . . . . .	1
2 Background . . . . .	3
2.1 Band-Pass Filter in the Preprocessing Stage . . . . .	3
2.2 Examining the BPF-Preprocessed Signals . . . . .	3
3 Methods . . . . .	4
3.1 Chapter Introduction . . . . .	4
3.2 Infinite Impulse Response (IIR) BPF . . . . .	4
3.3 Pearson's Correlation Coefficients . . . . .	5
3.4 Posterior Cingulate Cortex (PCC) Cloud . . . . .	7
3.5 Different Seed Methods and Threshold Selection . . . . .	8
3.6 Subspace Filtering . . . . .	10
3.7 Connection Probabilities . . . . .	13
3.8 Whole-Brain Timescore (WBTc) Peaks vs. Bandwidth of BPFs . . . . .	16
4 Results . . . . .	17
4.1 Chapter Introduction . . . . .	17
4.2 Population of Subjects and the Logic of Analysis . . . . .	17

4.3 Seed Evaluations . . . . .	18
4.4 Threshold versus Variation in Connection Probabilities . . . . .	20
4.5 Subspace Filtering Evaluations and Interpretations . . . . .	32
4.6 Whole-Brain Timescore Evaluations and Interpretations . . . . .	33
4.7 Overall Evaluations and Interpretations . . . . .	34
5 Conclusions and Future Works . . . . .	35
5.1 The summary of this thesis . . . . .	35
5.2 Perspective Works . . . . .	35
BIBLIOGRAPHY . . . . .	36

## LIST OF TABLES

Table	Page
4.1 Subjects Serial Numbers Mapping Table . . . . .	17
4.2 The number of voxels in the PCC cloud based on OS and $r = 0.4$ . . . . .	19
4.3 The number of voxels in the PCC cloud based on RbS and $r = 0.4$ . . . . .	19
4.4 $CDF(0.1)$ of the <b>xpdf</b> difference values based on $r = 0.4$ and OS . . . . .	22
4.5 $CDF(0.1)$ of the <b>xpdf</b> difference values based on $r = 0.4$ and RbS . . . . .	22
4.6 $CDF(0.1)$ of the <b>xpdf</b> difference values based on AST and OS . . . . .	23
4.7 $CDF(0.1)$ of the <b>xpdf</b> difference values based on AST and RbS . . . . .	23
4.8 $CDF(0.1)$ for <b>xpdf</b> difference values based on $r = 0.5$ and OS . . . . .	24
4.9 $CDF(0.1)$ for <b>xpdf</b> difference values based on $r = 0.5$ and RbS . . . . .	24
4.10 $AUCs$ of the <b>xpdf</b> difference values based on $r = 0.4$ and OS . . . . .	25
4.11 $AUCs$ of <b>xpdf</b> difference values based on $r = 0.4$ and RbS . . . . .	26
4.12 $AUCs$ of the <b>xpdf</b> difference values based on AST and OS . . . . .	26
4.13 $AUCs$ of the <b>xpdf</b> difference values based on AST and RbS . . . . .	27
4.14 $AUCs$ of the <b>xpdf</b> difference values based on $r = 0.5$ and OS . . . . .	27
4.15 $AUCs$ of the <b>xpdf</b> difference values based on $r = 0.6$ and RbS . . . . .	28
4.16 Threshold values selected by the AST algorithm on the situation of OS . . .	29
4.17 Threshold values selected by the AST algorithm on the situation of RbS . .	30

## LIST OF FIGURES

Figure	Page
3.1 Five Different IIRBPFs . . . . .	4
3.2 One example of the PCC cloud . . . . .	8
3.3 Visualization of Pearson's Correlation Coefficients . . . . .	11
3.4 Visualization of Singular Values Matrix $\Sigma$ . . . . .	12
4.1 The Correlation Coefficients based on two Seed Methods and Different BPFs	18
4.2 One example of a CDF plot for <b>xpdf</b> difference values . . . . .	20
4.3 Plots of means for each pair of BPFs from Table 4.4 to Table 4.9 . . . . .	25
4.4 Plots of means for each pair of BPFs from Table 4.10 to Table 4.15 . . . . .	28
4.5 Number of voxels in the PCC cloud produced by AST with both OS and RbS applied . . . . .	31
4.6 Results of not-exactly-correct Subspace Filtering . . . . .	32
4.7 Results of correct, but not necessary, Subspace Filtering . . . . .	33
4.8 One example of Whole-Brain Timescore based on five BPFs-preprocessed brain data . . . . .	34

# Chapter 1

## Introduction

### 1.1 Why fMRI

Functional Magnetic Resonance Imaging (fMRI) is applied to explore the human brain and its related diseases [1] [2] [3]. Through large efforts involving mathematical computing, various relationships have been unveiled in fMRI images [1] [4] [5]. Blood oxygenation level-dependent (BOLD) signals have been considered one of the most useful and information-bearing signals on which a lot of experiments and analysis has been conducted to try to "mine" the useful information [1] [3].

### 1.2 Functional Connectivity

During the complex processing of fMRI images, one idea is to explore "connections" between every pair of voxels, the unit being artificially segmented [1] [3]. This kind of relationship is normally considered as "Functional Connectivity (FC)" [1] [5] [3] [6]. This is methodology frequently employed in fMRI processing [1]. More importantly, the Posterior Cingulate Cortex (PCC), an area believed to be highly correlated with memory and other important human properties, is a focus for FC [7] [8].

### 1.3 Preprocessing is the Cornerstone

Due to the large scale and high complexity of fMRI, a "preparing phase", usually characterized as the preprocessing stage, essentially impacts all subsequent analyses [9]. A high-quality preprocessing stage offers researchers a very solid foundation for conducting more advanced works. This thesis focuses on one step in the preprocessing stage, and to fully examine this step, a set of comprehensive analyses have been conducted.



Due to the fact that few publications have focused on preprocessing because of the simplicity of the procedures, and which therefore may be considered as "not worth exploration", in this thesis, however, some complex evaluating methods have been applied to examine the preprocessing stage. Guided by this purpose, many meaningful explorations and conclusions are obtained.

## Chapter 2

### Background

#### 2.1 Band-Pass Filter in the Preprocessing Stage

A large number of attempts for preprocessing fMRI images in a broad range of fields have been described. In the dataset upon which this thesis is based [11], a band-pass filter (BPF) has been applied to the original brain-scanned data to filter the time sequence. This step is very important because the removal of high and low frequencies largely optimizes the potential subsequent analysis and makes them more precise.

Originally, based on experimental results and academic conclusions, a BPF designed with pass band from  $0.01\text{ Hz}$  to  $0.1\text{ Hz}$  was employed to filter the time sequence, since the coherent spontaneous low-frequency fluctuations have been identified at that frequency range [13]. However, in this thesis, the upper pass band edge is increased, and different pass band widths' effects on filtering the time sequence are compared. This project developed with five different pass bands - the five different BPFs, and applies them respectively to the preprocessing stage.

#### 2.2 Examining the BPF-Preprocessed Signals

Examination of effects brought by BPFs on the preprocessing phase is a difficult task. In this thesis, several procedures are used to test the preprocessing phase. At first, a voxel addressed as a "seed" was selected from the PCC. Using this seed and an artificially determined value - the threshold value, a set of other voxels in the PCC area can be identified in terms of the degree of connectivity to the seed, and this set is named as the PCC cloud. Finally, a variable, termed as connection probabilities, was obtained via repeatedly generating PCC clouds by procedures described above. Many variations of these procedures for evaluating the preprocessing stage are performed to explore in best approach.

## Chapter 3

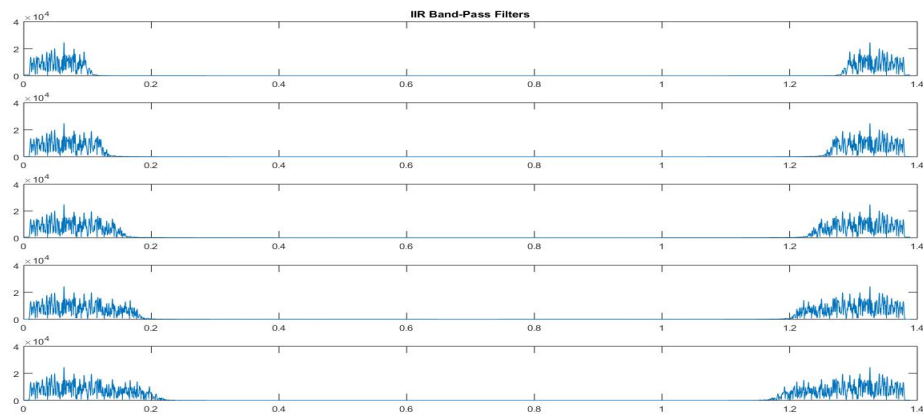
### Methods

#### 3.1 Chapter Introduction

The methods in this thesis were initiated by building five BPFs, then followed by pre-processing; after that, PCC clouds will be produced by a unique algorithm based on the preprocessing of those BPFs. Coming together with the PCC clouds will be the several significant variables addressing the signal properties of patients' brains.

#### 3.2 Infinite Impulse Response (IIR) BPF

The IIRBPF usually deals with the linear time-invariant signal systems, which is consistent with the brain signals in this thesis, and is also the necessary and core part of pre-processing fMRI. They are mainly used for filtering the temporal sequence in signals of length 1200 from four-dimension (4D) fMRI signals. There are five different IIRBPFs in the preprocessing phase in terms of five different right-most "stop-band frequencies", ranging from  $0.1 \text{ Hz}$  to  $0.2 \text{ Hz}$  with  $0.025 \text{ Hz}$  as single step, while the left-most "stop-band" frequency remains unchanged.



**Figure 3.1:** Five Different IIRBPFs

The IIRBPFs shown above have been built on the sampling period of 0.72, which is also the sampling period for processing the fMRI data. Applied with Fast Fourier Transform (FFT), **Figure 3.1** indicates that the pass band is clearly and quickly "cut off" from transition bands for all of those five IIRBPFs. Therefore, this building of IIRBPFs can be regarded as successful.

### 3.3 Pearson's Correlation Coefficients

The quantitative analysis in this thesis is mainly based upon the Pearson's Correlation Coefficients, which is widely used for signal processing in fMRI. All variables and their corresponding analyses following this chapter will be regulated by it. Some significant "preparing work" must be done before computing the Pearson's Correlation Coefficients. Firstly, a seed was selected from the Posterior Cingulate Cortex (PCC Network). Then, some numbers, used as indices, were randomly chosen from normal distribution with  $\mu = 0$  and  $\sigma = 1$ , which is denoted as

$$\mathbf{X} \sim \mathcal{N}(0, 1), \quad (3.1)$$

where the  $\mathbf{X}$  is a set comprising 200 elements. Therefore, this set is determined as a "Random Window" drawn from the original temporal sequence with length 1200.

After applied those indices from  $\mathbf{X}$ , all voxels in the brain are of length 200, and the method of Pearson's Correlation Coefficients is applied as

$$\rho = \frac{\frac{1}{N} \sum_{n=0}^{N-1} S_i[n] S_j[n]}{\sqrt{\left(\frac{1}{N} \sum_{n=0}^{N-1} S_i^2[n]\right) \left(\frac{1}{N} \sum_{n=0}^{N-1} S_j^2[n]\right)}}, \quad (3.2)$$

where the  $S_i$  and  $S_j$  are any two voxels inside the brain. In our method, equation 3.2 will be applied to our selected seed and every voxel, respectively. The entire detailed algorithm for computing Pearson's Correlation Coefficient has been summarized below:

---

**Algorithm 1** Get Pearson's Correlation Coefficient

---

```
1: procedure Getcv( $R, S, G, Seed$ )
2:    $G \leftarrow ProcessXcell(S, R)$ 
3:   for all  $G_i \in G$  do
4:      $Norm(G_i)$ 
5:   end for
6:    $Seed \leftarrow Seed - mean(Seed)$ 
7:    $Norm(Seed)$ 
8:    $cv \leftarrow GetConnections(G, Seed)$ 
9: end procedure
```

▷ The Selected Seed is from PCC Network  
▷ R is the set indicating the preprocessed brainmask  
▷ The Seed has also been truncated using the elements from **X**  
▷ Pearson's Correlation Coefficients

---

**Algorithm 1** above mainly covers a general "Pipeline" for obtaining the Pearson's Correlation Coefficients, values that are directly used for obtaining the PCC cloud. Two sub-algorithms involved are presented below.

---

**Algorithm 2** ProcessXcell and GetConnections

---

```
1: procedure ProcessXcell( $S, R$ )
2:   for all  $R_j \in R$  do
3:     if  $R_j > 0$  then
4:        $S_i \leftarrow S_i - mean(S_i)$ 
5:       if  $norm(S_i) > 0$  then
6:          $S_i \leftarrow S_i \div norm(S_i)$ 
7:          $G \leftarrow S_i$ 
8:          $Index \leftarrow i$ 
9:       end if
10:    end if
11:  end for
12: end procedure

13: procedure GetConnections( $G, Seed$ )
14:   for all  $G_k \in G$  do
15:      $xcv \leftarrow G_k \times Seed$ 
16:   end for
```

▷ Index will be used for obtaining the 3D Coefficient matrix  
▷ Applied with Pearson's Correlation Coefficients to Seed and every voxel in G, it is now vectorized

---

---

```

17:   for all  $l \in Index$  do
18:        $cv_l \leftarrow xcv_l$  ▷ Pearson's Correlation Coefficients have
been obtained in form of a 3D matrix
19:   end for
20: end procedure

```

---

With **Algorithm 1** and **Algorithm 2**, the Pearson's Correlation Coefficients can be completely obtained with truncated voxels in it - as indicated in the **cv**. In the following algorithms and descriptions, we use "*truncate()*" to represent truncating the voxels with length 1200 by using the indices from **X** in equation 3.1.

### 3.4 Posterior Cingulate Cortex (PCC) Cloud

Following the previous section, this section generates the PCC cloud in direct assistance with Pearson's Correlation Coefficients. And now a parameter, the threshold value  $r$ , will be selected for determining which voxel in **cv** could be included in the potential PCC cloud - the cloud describing the positions for all voxels in **cv**. This cloud can be reasonably denoted as a set, that is

$$p \in \mathbf{P}\{p \mid p \in \mathbf{cv}, cv_i > r\}, \quad (3.3)$$

where  $p$  represents the voxels in **cv**, and  $r$  is the threshold value. Also, the length of voxels should be truncated using the indices from **X**. A detailed algorithm is summarized as below:

---

#### **Algorithm 3** Obtain PCC Cloud

---

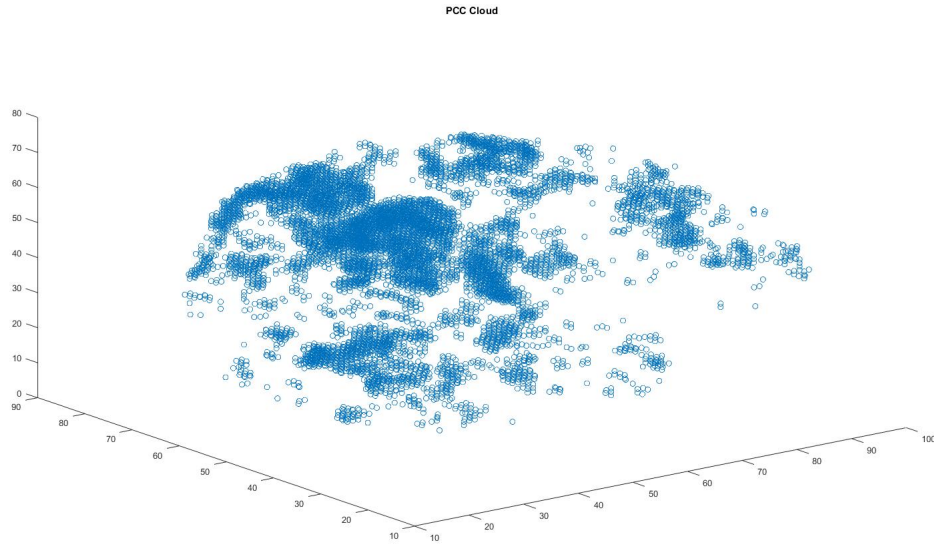
```

1: procedure  $PCCcloud(cv,r)$ 
2:   for all  $p_i \in cv$  do
3:       if  $p_i > r$  then
4:            $positions \leftarrow (i, j, k)$  ▷  $(i, j, k)$  is the position for the corresponded voxels
5:       end if
6:   end for
7: end procedure

```

---

When we plot the **positions** from the algorithm above, the figure of the PCC cloud is below



**Figure 3.2:** One example of the PCC cloud

**Figure 3.2** depicts the positions of all voxels in  $\mathbf{cv}$  with values larger than the threshold value. As described in **Algorithm 3**, this plotting of PCC clouds is yielded by 3D-coordinates of those voxels, therefore it could be regarded as the threshold value regulating the number of voxels in the PCC cloud and the shape of that cloud.

### 3.5 Different Seed Methods and Threshold Selection

All algorithms discussed above rely upon a single seed, selected from the PCC, and a threshold value determined through "engineering experience". However, experience does not necessarily lead to complete exactness. Given the fact that the seed selected previously is located in the PCC, is it reliable? Besides that, is the threshold value of 0.4 "good enough" to produce a fully reasonable PCC cloud - that is, are all voxels in the PCC cloud all "qualified"? To address these issues, other methods for selecting seed and threshold value should be taken into considerations.

The first new method introduced is a novel way for determining the seed, which is referred to as the "Region-based Seed (RbS)". Compared to that original seed selected

from the PCC, RbS is more likely to be reasonable, since it depends on the entire Cingulate Cortex (CC). Similar to the "brainmask" obtained from the preprocessing stage, there is a "cingulate cortex mask", denoted as "xccmask", also obtained from the physiological analysis of human brain. And this "xccmask" is now applied in this thesis for obtaining a reliable seed. A detailed algorithm for obtaining the RbS is described below:

---

**Algorithm 4** Obtain Region-based Seed (RbS)

---

```

1: procedure RegionSeed( $Q, S$ )
2:   for all  $Q_i \in Q$  do                                ▷  $Q$  is the set representing the "xccmask"
3:     if  $Q_i > 0$  then
4:        $RbS \leftarrow S_j$ 
5:     end if
6:   end for
7:    $Mean(RbS)$ 
8:    $Transpose(RbS)$                                     ▷ Transforming to the form of vector
9: end procedure

```

---

A new method for determining thresholds uses the characteristic of "Adaptiveness". This method is termed: Adaptively-Selected Threshold (AST). AST is the more reliable threshold value which characterizes the property of "following-up" to specific subjects. The AST is proposed by Charreau S. Bell, a PhD candidate in our group. This concept was developed due to the impact of properties in specific subjects on selection of thresholds. Specifically, Charreau S. Bell found that the original threshold value almost filters out all voxels for some subjects, but oppositely, for some other subjects, that original threshold almost does "nothing", which means that too many qualified voxels, the values of which are all larger than the original threshold value. Therefore, it is considered more reasonable to "follow up" each subject for determining the threshold value instead of arbitrarily selecting one.

**Algorithm 1** yields vectorized Pearson's Correlation Coefficients, **x<sub>cv</sub>**:



$$\mathbf{xcv} = \begin{bmatrix} xcv_1 \\ xcv_2 \\ xcv_3 \\ \vdots \\ xcv_n \end{bmatrix}, \quad (3.4)$$

where  $xcv_i = Seed \cdot S_i[n]$ . Then, ten percentiles are found for  $\mathbf{xcv}$ . Those numbers equally divide  $\mathbf{xcv}$  into ten parts, which are the  $10^{th}, 20^{th}, \dots, 100^{th}$  percentiles. After that, a set is created as  $\mathbf{H}$ , which includes those ten percentiles. Finally, if the top 20% of voxels would be selected for building the PCC cloud, then a detailed algorithm for obtaining AST is stated below:

---

**Algorithm 5** Obtain Adaptively-Selected Threshold (AST)

---

```

1: procedure AdaptiveSeed( $xcv, H$ )
2:   for all  $H_i \in H$  do
3:      $PE \leftarrow percentile(H_i^{th}) \in xcv$  ▷ 10 percentiles in PE
4:   end for
5:    $Index \leftarrow length(PE) - 20 \div 10$  ▷ Top 20% of voxels
6:    $AST = PE_{Index}$  ▷ Every percentile in PE represents a range of 10% of voxels
7: end procedure

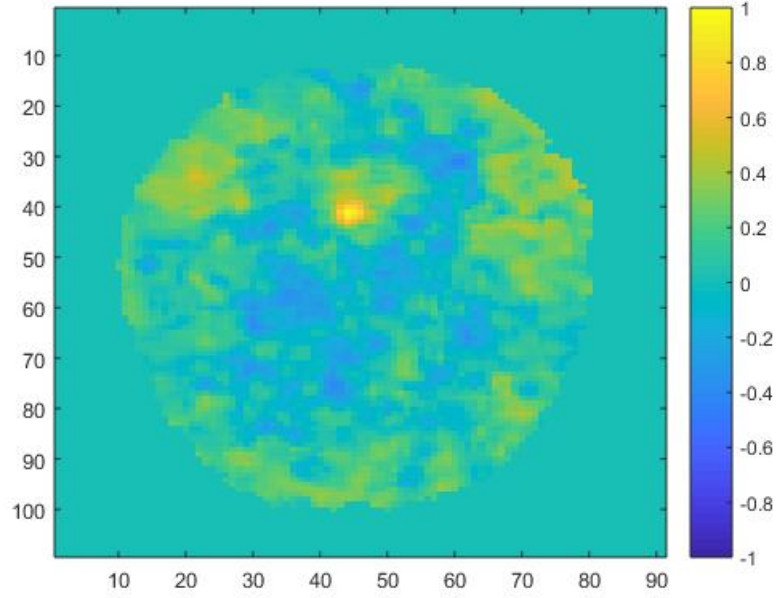
```

---

Based on those reasonable grounds, RbS and AST have both been applied in the quantitative approaches in this thesis. They are every important, and also considerably impact several critical variables and their corresponding analyses, which will be introduced soon.

### 3.6 Subspace Filtering

Yielded by **Algorithm 1**,  $\mathbf{cv}$  can be used to plot for determining the situation for correlating selected seeds with other voxels. The plane with which the selected seed shares is plotted for visualization of  $\mathbf{cv}$ .



**Figure 3.3:** Visualization of Pearson's Correlation Coefficients

Now the question is emerged as "could  $\mathbf{cv}$  be artificially increased?". It might be feasible through Principal Component Analysis (PCA) to find a lower-dimension plane onto which to project the original  $\mathbf{cv}$ . The degree of dimension depends on the number of components composed of the largest part of information among all components.

As indicated in **Algorithm 3.3**, the PCC cloud is obtained by singling out those voxels, the values of which correlation coefficients are above the threshold value. And now, for every voxel singled out, its corresponding signal sequence in  $\mathbf{xcell}$  has been collected and gathered as

$$\mathbf{Signals} = \sum_i \mathbf{S}_i \mathbf{S}_i^T, \quad (3.5)$$

where  $S_i$  has still been truncated using the indices from  $\mathbf{X}$  in equation 3.1. Then  $\mathbf{Signals}$  can be decomposed as

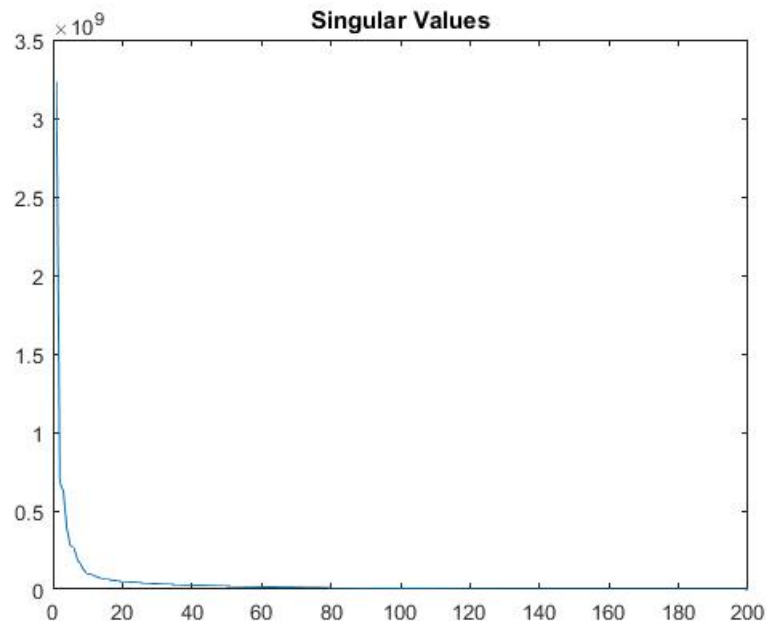
$$\mathbf{Signals} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T, \quad (3.6)$$

where  $\mathbf{U}$  is unfolded as

$$\mathbf{U} = [U_1, U_2, \dots, U_P, \dots, \dots]. \quad (3.7)$$

The value of  $P$  is the degree of plane dimension for projecting the original data. That brings us to the second question addressing the methods for choosing  $P$ . Two methods of selecting  $P$  will be provided in this thesis for intending to increase  $cv$ . The first method is based on "visual inspection", while the second one deals with this issue using a mathematically stricter approach, a methodology widely employed in the field of *Machine Learning*.

"Visually" selecting the  $P$  value is introduced first. Equation 3.6 is essentially the operation of Singular Value Decomposition, the **SVD**, which results in a matrix  $\Sigma$ , the diagonal matrix with singular values in descending order inside it. This matrix is visualized as below



**Figure 3.4:** Visualization of Singular Values Matrix  $\Sigma$

Due to the truncation, the length of  $\Sigma$  is also 200, as shown in **Figure 3.4**. Under careful observation, a "turning point" is appeared in the plot above, which is about "10" with respect to the horizontal axis. This "turning point" cuts off the ten highest singular values from the rest. Hence,  $P = 10$  could be a suitable choice.

The second operation for determining  $P$  is a mathematically stricter method of "computing" the  $P$ . The diagonal matrix  $\Sigma$  in equation 3.6 is

$$\Sigma = \begin{bmatrix} \sigma_{11} & & \\ & \ddots & \\ & & \sigma_{nn} \end{bmatrix}, \quad (3.8)$$

where elements except those along the diagonal are all equal to 0, and  $n$  denotes the length of **Signals**, which is 200. Then,  $P$ , the smallest integer, should be selected such that

$$\frac{\sum_{i=1}^P \sigma_{ii}}{\sum_{i=1}^n \sigma_{ii}} \geq 0.99. \quad (3.9)$$

Computing the  $P$  value in this method is more reasonable and logical, since the value of equation 3.9 guarantees the 99% of the variance in the original signals is retained [10]. This high degree of variance retaining is required in PCA for maintaining the necessary precision and completeness of the original data. After  $P$  has been determined,  $U$  is now truncated such as

$$\mathbf{U}_{\text{proj}} = [U_1, U_2, \dots, U_P], \quad (3.10)$$

where the  $\mathbf{U}_{\text{proj}}$  may be employed to project the original data onto the  $P$  dimensional plane.

Actually, the two methods discussed above prove that the Subspace Filtering, the PCA in this thesis, is not necessary, since the first method just "forces" the increase of the correlation coefficients at the expense of losing much of the original information, while the second method produces a very large  $P$  value in the range of 110 to 130 (through the experiments conducted in *MATLAB<sup>TM</sup>*); therefore, it is very close to the length of the original data. More complete results and analysis for this section will be provided in the next chapter.

### 3.7 Connection Probabilities

The final topic in this chapter is Connection Probabilities (CP), a matrix denoted as **xpdf**. CP measures the probabilities of every voxel's relationships to other voxels. As pre-

viously mentioned, **Algorithm 1** through **Algorithm 3** yield the PCC cloud, and now these three algorithms can be used to obtain **xpdf**. In addition, although the Subspace Filtering is not necessary in this thesis, it can be used in initiating the procedures for obtaining the **xpdf** that are contained within **Algorithm 1** to **Algorithm 3**. It should be noticed that until now, two methods of finding seeds and determining the threshold values have been presented in this thesis. Therefore, different combinations among these methods will produce different results for **xpdf**.

Although Subspace Filtering is not exactly correct for "forcing" the correlation coefficients to be increased, it can be used to generate a source for dynamically selecting the seeds to obtain the **high-probability-oriented** voxel-wise connection variable. Applied with the value of  $P$ , the original signals can be projected onto a  $P$  dimensional plane - the new PCA-processed signals are produced. Then, a new **cv**, produced again by previous algorithms, will be employed to offer a "source" for us to choose several seeds, a procedure similar to **Algorithm 3**. After that, we choose a "loop" with specific length, and in every iteration of this "loop" a seed will be dynamically selected from the "source" in the order of the sequence number of its corresponding loop. Finally, taking as input arguments those variables discussed above, procedures in **Algorithm 1** to **Algorithm 3** will be repeated, in the regulation of the length of that loop, to obtain the **xpdf**. In sum, this entire process is dynamic. However, if the RbS is applied, then the seed will remain unchanged during every iteration. Aside from that, due to different combinations of seed and threshold methods, some vital details in this dynamic process can vary greatly. In the same manner as for the previous notice, in this chapter only methods are introduced, and the further analyses, details, and results will appear in the next chapter.

The following outlines more comprehensive procedures describing how to obtain **xpdf**. These procedures "conclude" the main approaches introduced in this chapter, and includes all algorithms, all seed methods and threshold selections as previously introduced.

---

**Algorithm 6** Comprehensively Obtain the Connection Probabilities-*xpdf*

---

```
1: procedure Obtainxpdf( $R, S, G, P, OriginalSeed, r, U_{proj}, Looplength$ )
2:    $Proj = U_{proj} U_{proj}^T$ 
3:   for all  $R_i \in R$  do
4:     if  $R_i > 0$  then
5:        $Truncate(S_i)$ 
6:        $S_i - mean(S_i)$ 
7:        $S_i \leftarrow Proj \cdot S_i$  ▷ Projection is now complete
8:     end if
9:   end for
10:   $G \leftarrow ProcessXcell(S_{proj}, R)$ 
11:  SelectSeed ▷ Select RbS or Original Seed
12:  if  $Seed = OriginalSeed$  then
13:     $Seed - mean(Seed)$ 
14:     $Seed \div norm(Seed)$ 
15:  end if
16:   $cv_{new} \leftarrow GetConnections(G, Seed)$ 
17:  SelectThreshold ▷ Select AST or Original Threshold
18:  if  $r = AST$  then
19:     $r \leftarrow AST$ 
20:  end if
21:   $Logic_{index} \leftarrow find(cv_{new} > r)$  ▷  $O_s$  and  $1_s$  will be presented in  $Logic_{index}$ 
22:  for all  $L_i \in Logic_{index}$  do ▷  $L_i \in \{0, 1\}$ 
23:    if  $L_i > 0$  then
24:       $positions \leftarrow (i, j, k)$  ▷ This  $positions$  will be used for dynamically selecting the  $Seed$ 
25:    end if
26:  end for
27:  while  $Loop < LoopLength$  do
28:    if  $Seed = OriginalSeed$  then
29:       $Seed_{dyn} \leftarrow S_{positions_{Loop}}$  ▷ Dynamically selecting the  $Seed$  for each iteration in while
30:    else  $Seed_{dyn} \leftarrow RbS$  ▷ If RbS has been previously selected, the dynamic  $Seed$  will be equal for every iteration
31:    end if
32:     $Truncate(Seed_{dyn})$ 
33:     $Seed_{dyn} - mean(Seed_{dyn})$ 
34:     $Seed_{dyn} \div norm(Seed_{dyn})$ 
35:     $cv_{dyn} \leftarrow GetConnections(G, Seed_{dyn})$ 
36:     $xpdf_{dyn} \leftarrow find(cv_{dyn} > r)$  ▷ There will be  $0_s$  and  $1_s$  in  $xpdf_{dyn}$  for every iteration
37:     $\underline{xpdf} \leftarrow \underline{xpdf} + xpdf_{dyn}$  ▷  $xpdf$  has been initiated as  $0_s$ , then recursively summing up it and  $xpdf_{dyn}$  for every iteration
38:  end while
39: end procedure
```

---

### 3.8 Whole-Brain Timescore (WBTc) Peaks vs. Bandwidth of BPFs

Compared with previous discussions, the Whole-Brain Timescore (WBTc) will be much more direct. The WBTc purely measures the "time resolution" in every original voxel. It is simply computed as

$$\mathbf{WBTc}[\mathbf{n}] = \sum_i^N S_i[n] \sum_j^N S_j[n], \quad (3.11)$$

where  $i$  and  $j$  should have the same range, and  $N$  is 1200 because of the "Whole-Brain" Timescore. The equation 4.1 should be considered as summing up all the correlations between every pair of voxels in the brain.

In the next chapter, many results and analyses will be presented for illustrating the relation between WBTc peaks and bandwidth of those five different IIRBPFs which have been used for preprocessing the brain signals.

## Chapter 4

### Results

#### 4.1 Chapter Introduction

Chapter 3 comprehensively presents several computational methods for evaluating the preprocessing of fMRI images, obtaining and analyzing their significant properties. In this chapter, all the experimental results, evaluations, and interpretations will be involved for the full and quantitative illustrations of the methods previously discussed. In addition, this chapter also provides a comprehensive analysis for the selection of those methods introduced in previous chapter.

#### 4.2 Population of Subjects and the Logic of Analysis

In our data set, there are seventeen subjects in total, a data set from Massachussets General Hospital [11], the brain fMRI images of which have been processed by all the methods mentioned in the last chapter. Correspondingly, each of those subjects have five different **xcell** because those five different IIRBPFs have preprocessed the original signals. Therefore, for every method in the previous chapter, there are also five "parallel paths" with respect to every resulting variable, typically **xpdf**. Aside from that, it is even more important to explore differences and similarities for all those variables.

For the sake of simplicity of expressions in this chapter, the original serial number for each subject has been abbreviated as indicated in table below:

**Table 4.1:** Subjects Serial Numbers Mapping Table

Original#	100206	100408	100610	101006	101107	101309	101410	101915	102008	102311	102513	102816	140117	153227	170934	200109	702133
Abbreviated#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

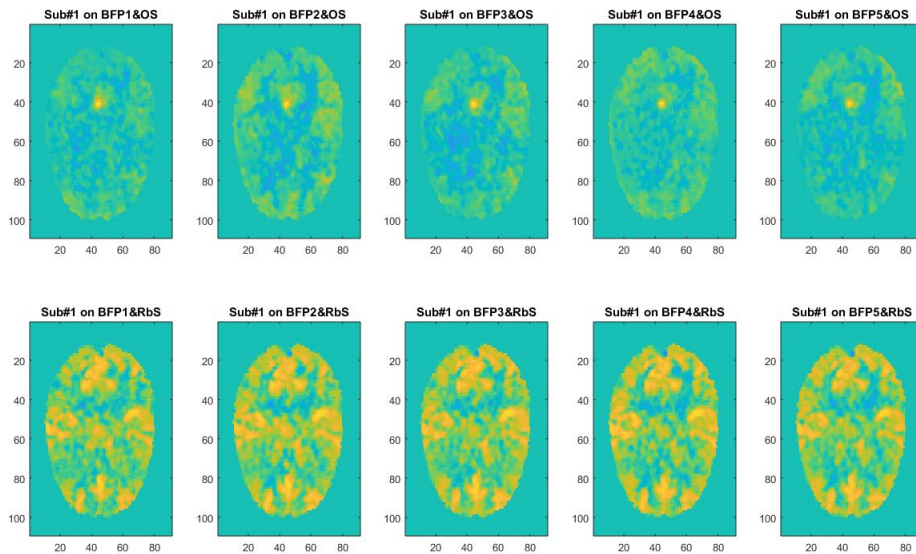
**Table 4.1** maps each subject's original serial number to a simple sequence number, which will be used for denoting every subject over this chapter.



### 4.3 Seed Evaluations

Several important sets of experiments were involved for implementing all the methods and addressing all the issues in previous chapter. Since all methods are based on seed, therefore, it is essential to evaluate its effectiveness first.

As previously indicated, two methods for seed selection, the single seed selected from the PCC and the RbS, have been presented, and the former is named as "Original Seed (OS)". Both of them are evaluated in terms of Pearson's Correlation Coefficients, the  $cv$ , at the same plane with which OS shares, and the number of voxels in the PCC cloud, while the threshold value is remained unchanged for this part. The figure below shows the  $cv$  from the two seed selection methods based on the preprocessing of five different IIRBPFs.



**Figure 4.1:** The Correlation Coefficients based on two Seed Methods and Different BPFs

Obviously, based on all those five IIRBPFs, the  $cv$  of RbS is much better than that of OS; therefore, the RbS, in this situation, makes significant improvements, which could be a sufficient reason for the favoring toward choosing RbS as the reliable seed. As mentioned before, the figure above only takes one "slice" from the whole brain, and it is necessary to know the concrete number of voxels in the PCC cloud. For the two tables presented below,

the threshold value remains as 0.4, which is the "experience-based value".

**Table 4.2:** The number of voxels in the PCC cloud based on OS and  $r = 0.4$

Subject#	BPF1	BPF2	BPF3	BPF4	BPF5
1	810	4100	1780	602	725
2	88920	81547	91800	68492	52162
3	20297	16271	6242	9039	5853
4	25	24	20	19	23
5	74543	64406	59129	53493	78976
6	32928	17871	19814	18940	8370
7	819	29	20	24	22
8	36109	36086	18917	17958	8611
9	136259	124692	126656	118551	102364
10	14101	3079	55	321	79
11	56314	69106	43631	28062	29822
12	34844	26709	30720	32912	18461
13	79	54	54	51	41
14	3801	2742	1563	156	599
15	230	323	212	91	172
16	3779	836	3737	799	34
17	51341	40180	39067	37744	32362

**Table 4.3:** The number of voxels in the PCC cloud based on RbS and  $r = 0.4$

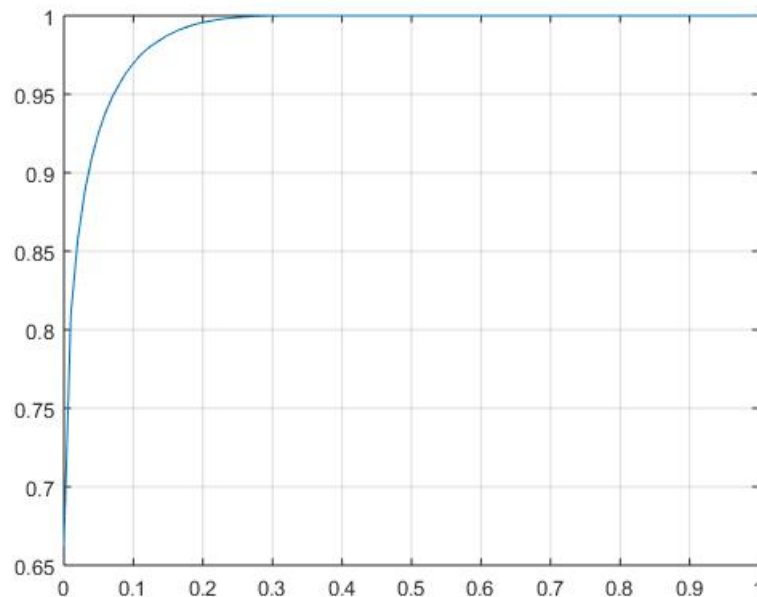
Subject#	BPF1	BPF2	BPF3	BPF4	BPF5
1	80327	93560	91214	89676	84250
2	114868	120756	110477	112459	101993
3	77106	81375	71555	72886	63714
4	34247	32457	22388	25355	21009
5	154657	159141	160173	143326	152141
6	82491	75836	53540	66346	56320
7	33550	29552	22521	31144	22728
8	139157	132824	128190	108895	117002
9	145516	149595	140373	131949	136957
10	107911	106337	97031	96079	88714
11	113597	111981	99664	106907	87848
12	104802	94718	82200	82057	88772
13	106238	108086	101370	109454	92280
14	65152	50175	56063	57413	46664
15	25602	20095	22417	17438	20827
16	43374	32398	38869	25041	26430
17	150246	142142	135346	120912	121290

As expected, the values in **Table 4.2** vary significantly much more than those in **Table 4.3**, because the RbS is more reliable compared to OS. Although this variance could also be attributed by the subjects' properties, the RbS "reconciles" those differences to some degree. In conclusion, RbS can be a better seed method.

#### 4.4 Threshold versus Variation in Connection Probabilities

The previous part evaluates seed methods with respect to the number of voxels in the PCC cloud. In this part, seed evaluations will continue on the basis of more advanced methods, which will support threshold evaluations.

In the last chapter, a complex variable has been introduced, the Connection Probabilities, denoted as **xpdfs**, which is based upon the selection of seed and threshold value ( $r$ ). For one subject, there are correspondingly five different **xpdfs**. If the absolute value of the difference in each pair of **xpdfs** is calculated, then ten difference values will be obtained. After that, for each difference value, the CDF will be calculated and a CDF plot will also be obtained. The figure below is one example of a CDF plot for difference values.



**Figure 4.2:** One example of a CDF plot for **xpdf** difference values

In **Figure 4.2**, each value of **xpdf** difference has been scaled by 1000, and the CDF values have been evaluated from 0.1 to 1. In our analysis, two significant variables regarding the difference values in each pair of **xpdfs** will be calculated for a comprehensive assessment. The first one is  $CDF(0.1)$ , the value considered as the boundary for determining the preliminary similarities in each pair, and another one is the *Area Under Curve*, the *AUC*, which will be applied to measure the overall similarities in each pair. Compared with the first variable, *AUC* is more comprehensive. However, the  $CDF(0.1)$  is stricter and "demanding". Ideally, it is expected, albeit unrealistically, to see the  $CDF(0.1) = 1$ . For instance, the situation in **Figure 4.2** is very close to ideal condition, but it should be noticed that many other situations would be worse than it.

There are many potential ways for calculating *AUCs*. In this thesis, the *Trapezoidal Rule* is employed for obtaining *AUCs*, because the calculation of *AUCs* is essentially the numerical quadrature. For every small region under the curve, the region composed of every two neighboring points along the  $x$  axis, its area can be obtained by approximating it as a trapezoid [12]. For one region, it is calculated as

$$\int_a^b f(x)dx \approx (b-a) \left[ \frac{f(a) + f(b)}{2} \right] \quad (4.1)$$

where  $a$  and  $b$  represent every two neighboring points. By summing over every small area in this way, we have

$$\int_a^b f(x)dx \approx \sum_{k=1}^N \frac{f(x_{k-1}) + f(x_k)}{2} \Delta x_k \quad (4.2)$$

where  $N$  is the number of areas calculated by equation 4.1. Due to the small number of panels in **Figure 4.2**, the error produced by equation 4.2 can be neglected.

For following pages, several plots and tables will be presented regarding the variations of **xpdf** with the changes of threshold and seed selections.

If the threshold is 0.4 with the two seed methods applied, then we have the two tables below:

**Table 4.4:** *CDF*(0.1) of the **xpdf** difference values based on  $r = 0.4$  and OS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.969757	0.816493	0.961101	0.931626	0.65673	0.758355	0.699996	0.990207	0.985193	0.882323
2	0.700817	0.764749	0.574024	0.32467	0.835071	0.736773	0.441462	0.642782	0.379745	0.49702
3	0.826001	0.68958	0.744563	0.683712	0.809316	0.892141	0.808421	0.990999	0.963223	0.907196
4	0.995995	0.985743	0.992802	0.987689	0.96848	0.997562	0.980036	0.999915	0.99971	0.95354
5	0.578512	0.426703	0.704485	0.629637	0.617555	0.728209	0.733882	0.907548	0.848394	0.63974
6	0.744263	0.732037	0.626841	0.569351	0.947533	0.820971	0.679471	0.810239	0.706219	0.840139
7	0.939747	0.904477	0.89269	0.887098	0.96548	0.971796	0.959694	0.994561	0.990479	0.996522
8	0.547945	0.460801	0.37429	0.316456	0.711829	0.676979	0.401807	0.796377	0.494377	0.515673
9	0.603331	0.653481	0.38498	0.295437	0.856447	0.62838	0.415148	0.587212	0.373772	0.638038
10	0.540965	0.491785	0.488825	0.452831	0.601012	0.598881	0.518301	0.899566	0.610233	0.61849
11	0.888512	0.658602	0.638384	0.712498	0.683087	0.670466	0.742069	0.898821	0.948459	0.962174
12	0.850322	0.716238	0.612016	0.56752	0.687593	0.599083	0.568712	0.785913	0.679177	0.826247
13	0.847084	0.858729	0.758627	0.706055	0.956656	0.859584	0.832387	0.854902	0.824968	0.948759
14	0.753798	0.717616	0.684175	0.678065	0.913801	0.837469	0.845851	0.903982	0.91388	0.973932
15	0.938876	0.973736	0.959344	0.974831	0.950257	0.933973	0.937	0.995903	0.994394	0.999961
16	0.931471	0.942348	0.878501	0.831799	0.994959	0.970703	0.885392	0.947237	0.889572	0.932306
17	0.512504	0.458109	0.446078	0.413593	0.743391	0.650807	0.680054	0.740977	0.772159	0.836323

**Table 4.5:** *CDF*(0.1) of the **xpdf** difference values based on  $r = 0.4$  and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.968229	0.958366	0.94922	0.934613	0.938863	0.934852	0.916785	0.946268	0.933545	0.941446
2	0.977274	0.947752	0.959803	0.931158	0.935695	0.939178	0.908193	0.964066	0.938859	0.935
3	0.968008	0.944147	0.943543	0.915976	0.931691	0.937837	0.909361	0.957942	0.943435	0.936907
4	0.966976	0.938224	0.944547	0.930495	0.946976	0.954254	0.939421	0.986113	0.973553	0.965408
5	0.962712	0.960263	0.885217	0.930251	0.977095	0.866962	0.942613	0.862863	0.940837	0.933656
6	0.934324	0.869171	0.903192	0.878689	0.893688	0.931864	0.897908	0.97617	0.964623	0.92345
7	0.95195	0.929368	0.95373	0.934558	0.939865	0.961282	0.944358	0.978338	0.96736	0.946148
8	0.946092	0.931705	0.861365	0.894688	0.944194	0.889684	0.92128	0.896974	0.928927	0.975794
9	0.976457	0.94979	0.925982	0.934916	0.941963	0.91632	0.930304	0.939403	0.948558	0.975919
10	0.949877	0.92558	0.923855	0.899741	0.937744	0.937717	0.911682	0.961818	0.938591	0.944516
11	0.944349	0.919309	0.931216	0.870652	0.920907	0.938554	0.872673	0.969442	0.926082	0.902044
12	0.92483	0.879226	0.880891	0.903073	0.916291	0.917553	0.939994	0.938062	0.955293	0.968078
13	0.946473	0.937097	0.947218	0.898877	0.92592	0.940272	0.893577	0.960519	0.914244	0.895566
14	0.920029	0.936226	0.940715	0.905383	0.970474	0.972329	0.953326	0.967025	0.936786	0.931442
15	0.961332	0.964985	0.950529	0.958094	0.981753	0.966477	0.972839	0.959236	0.967996	0.980796
16	0.933015	0.94789	0.913717	0.91542	0.972882	0.947436	0.952356	0.926677	0.935103	0.970253
17	0.929726	0.906218	0.84147	0.852568	0.944215	0.882792	0.893594	0.896745	0.918558	0.941814

Comparing the two tables above, the situation fits our expectations - values of *CDF*(0.1) based on Region-based Seed (RbS) are much higher and less-varying than those based on Original Seed (OS). Specifically, the values higher than 0.9 from **Table 4.5** are much more than values from **Table 4.4**. This situation is considered "good", since more similarities means more stability.

If the AST is applied, then we have another two tables below:

**Table 4.6:**  $CDF(0.1)$  of the  $xpdf$  difference values based on AST and OS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.987555	0.92841	0.916312	0.982801	0.60987	0.560432	0.873133	0.725477	0.939366	0.975348
2	0.693594	0.79935	0.744425	0.776203	0.895875	0.873579	0.931065	0.858355	0.877762	0.917272
3	0.95317	0.888408	0.86071	0.933956	0.789289	0.773187	0.877017	0.852539	0.93416	0.936277
4	0.627689	0.376227	0.868317	0.563104	0.489522	0.920131	0.683798	0.974839	0.83999	0.408734
5	0.855562	0.92263	0.947811	0.740782	0.968918	0.905682	0.709442	0.653141	0.441962	0.487362
6	0.848802	0.819177	0.736359	0.74419	0.860886	0.820465	0.776914	0.785942	0.745482	0.812141
7	0.773176	0.759757	0.910821	0.844687	0.7618	0.949876	0.860872	0.943352	0.844018	0.610041
8	0.972021	0.99884	0.788215	0.9801	0.992135	0.489762	0.895447	0.163295	0.328696	0.991516
9	0.866374	0.933224	0.799099	0.888585	0.983204	0.891349	0.962156	0.712841	0.857158	0.985985
10	0.999367	0.972035	0.968219	0.563669	0.344486	0.32913	0.067317	0.766802	0.262042	0.223226
11	0.968169	0.935643	0.988509	0.972744	0.918957	0.991164	0.974082	0.987801	0.980759	0.771989
12	0.782154	0.962851	0.734039	0.892593	0.841499	0.735568	0.776082	0.565702	0.715914	0.873315
13	0.861322	0.98461	0.910215	0.685773	0.970508	0.805658	0.533225	0.376425	0.109072	0.416454
14	0.770823	0.860211	0.395787	0.612551	0.842578	0.400315	0.586756	0.314667	0.502896	0.926412
15	0.770168	0.954281	0.913751	0.953695	0.983422	0.943931	0.97098	0.652234	0.737224	0.811842
16	0.898698	0.875609	0.942638	0.969596	0.873936	0.938864	0.968416	0.961988	0.982405	0.932941
17	0.595388	0.521785	0.585283	0.444901	0.740197	0.789524	0.620888	0.879829	0.715673	0.642144

**Table 4.7:**  $CDF(0.1)$  of the  $xpdf$  difference values based on AST and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.968364	0.964283	0.956166	0.957902	0.963081	0.957572	0.9592	0.964899	0.959321	0.956708
2	0.978404	0.974057	0.973464	0.979711	0.977097	0.975985	0.978466	0.982706	0.97591	0.97777
3	0.964779	0.963414	0.962801	0.953044	0.966578	0.969903	0.962236	0.970068	0.963914	0.96457
4	0.951129	0.957157	0.958804	0.956335	0.964051	0.954875	0.957871	0.961653	0.962716	0.961126
5	0.967849	0.972039	0.964819	0.965449	0.959966	0.972655	0.967233	0.963147	0.961306	0.966108
6	0.964369	0.964318	0.967241	0.963352	0.962519	0.966228	0.966327	0.959261	0.964361	0.962961
7	0.941606	0.924923	0.928117	0.936006	0.937305	0.943129	0.934387	0.944271	0.932838	0.938981
8	0.976933	0.977914	0.975781	0.971676	0.976095	0.978939	0.972505	0.979389	0.975035	0.975432
9	0.978362	0.97693	0.970539	0.972322	0.977798	0.97201	0.975741	0.976053	0.979876	0.975698
10	0.980876	0.976209	0.978095	0.977183	0.975275	0.97989	0.980108	0.975772	0.973524	0.979885
11	0.966957	0.967852	0.969868	0.969147	0.968808	0.969768	0.967101	0.970793	0.970124	0.971831
12	0.967316	0.966207	0.96509	0.969441	0.961962	0.954557	0.965664	0.970107	0.966295	0.964174
13	0.953249	0.955495	0.951646	0.951291	0.946018	0.939322	0.953914	0.963199	0.951046	0.946946
14	0.967047	0.967157	0.953868	0.954114	0.96155	0.962682	0.962475	0.951414	0.955057	0.959405
15	0.954855	0.955027	0.951142	0.951517	0.948196	0.946898	0.954712	0.954669	0.953703	0.950564
16	0.946416	0.954984	0.954539	0.946961	0.952537	0.953903	0.947415	0.9523	0.949784	0.947428
17	0.962478	0.954325	0.951841	0.954124	0.957789	0.955885	0.958403	0.965828	0.957353	0.956015

This is the third time the advantage of RbS has been identified based on **Table 4.6** and **Table 4.7**. More importantly, for the same reason, the AST also offers the advantage of reliability over the  $r = 0.4$ .

If the threshold value is artificially increased to  $r = 0.5$ , then we have

**Table 4.8:**  $CDF(0.1)$  for **xpdf** difference values based on  $r = 0.5$  and OS

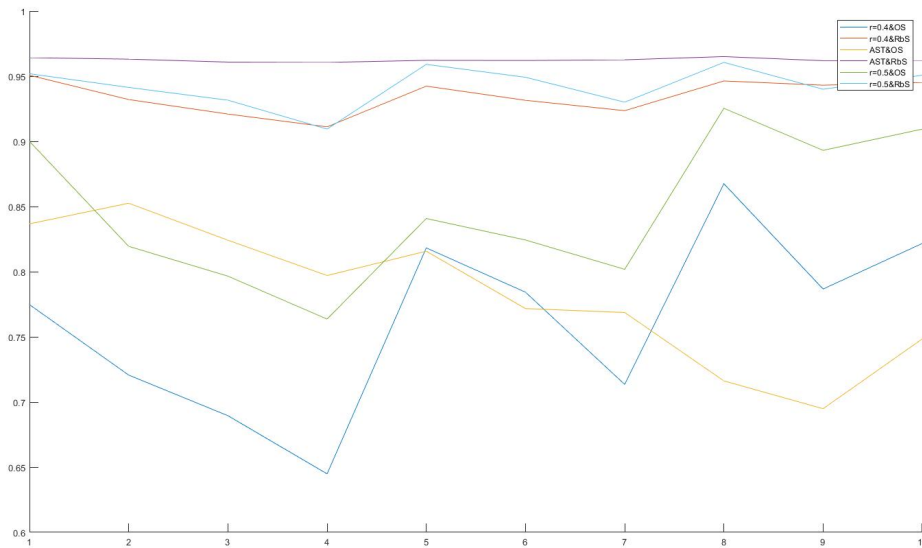
Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.999427	0.997191	0.930802	0.992078	0.901362	0.834638	0.875989	0.895037	0.991405	0.999922
2	0.643061	0.515001	0.469345	0.425786	0.713883	0.711464	0.531922	0.831495	0.593103	0.63902
3	0.895384	0.847411	0.820243	0.821551	0.98149	0.939632	0.950163	0.981677	0.982424	0.996606
4	1	1	1	1	0.988136	0.989337	0.989279	1	1	1
5	0.911333	0.639098	0.746765	0.472345	0.529323	0.670774	0.389643	0.870005	0.608068	0.544047
6	0.978334	0.868892	0.815	0.824981	0.846737	0.811025	0.813536	0.950323	0.969353	0.996377
7	0.995822	0.976745	0.985627	0.973376	0.979379	0.985999	0.973862	1	1	0.987228
8	0.976854	0.954415	0.706441	0.602957	0.695448	0.527504	0.520624	0.589338	0.567701	0.740876
9	0.835173	0.356317	0.353288	0.368418	0.390594	0.382051	0.396825	0.905147	0.955261	0.927323
10	0.945677	0.822775	0.845724	0.699634	0.748643	0.764968	0.702096	0.964626	0.794138	0.791658
11	0.662395	0.617935	0.609398	0.618678	0.801878	0.750337	0.883503	0.93858	0.992871	0.997593
12	0.747796	0.725421	0.69388	0.698205	0.920046	0.870459	0.896449	0.883546	0.893539	0.950575
13	0.9495	0.956993	0.939891	0.934281	0.999965	0.991736	0.979994	0.982294	0.971542	0.992945
14	0.967655	0.970655	0.954062	0.947657	0.999427	0.997516	0.99426	0.997903	0.994133	0.999207
15	0.971424	0.998698	0.987027	0.975918	1	1	0.999991	0.991763	0.980882	0.996093
16	0.952611	0.952153	0.962835	0.941099	0.999944	0.999974	0.99454	0.99697	0.985462	0.987632
17	0.867113	0.731007	0.721346	0.684791	0.79701	0.78575	0.738188	0.952599	0.902666	0.912449

**Table 4.9:**  $CDF(0.1)$  for **xpdf** difference values based on  $r = 0.5$  and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.968437	0.971923	0.945279	0.923062	0.973251	0.933471	0.909653	0.929078	0.897377	0.934765
2	0.986233	0.967225	0.947955	0.931876	0.945572	0.913696	0.904817	0.947756	0.935686	0.959462
3	0.955313	0.951171	0.941809	0.94436	0.959185	0.951236	0.961202	0.963292	0.964627	0.970198
4	0.980969	0.974638	0.972302	0.976442	0.978027	0.977071	0.98176	0.983787	0.987618	0.992195
5	0.96641	0.933358	0.931983	0.943496	0.939733	0.941729	0.93772	0.942932	0.927061	0.926337
6	0.94934	0.949666	0.93356	0.905437	0.970152	0.956106	0.929975	0.962584	0.937367	0.953676
7	0.985054	0.988974	0.986148	0.982481	0.993271	0.990304	0.988467	0.984284	0.979611	0.982748
8	0.943706	0.937333	0.945669	0.867581	0.963257	0.960426	0.88903	0.965604	0.899512	0.898605
9	0.943481	0.911816	0.898838	0.88015	0.942778	0.9364	0.91461	0.955682	0.934938	0.950237
10	0.925457	0.931107	0.904579	0.891769	0.972467	0.958312	0.949943	0.945751	0.933832	0.963006
11	0.882087	0.902114	0.856051	0.835185	0.97161	0.940983	0.928575	0.932876	0.915855	0.954923
12	0.951959	0.968573	0.960016	0.881631	0.973664	0.967671	0.902994	0.95957	0.884571	0.899502
13	0.872122	0.826146	0.831756	0.773872	0.925066	0.919535	0.883237	0.960935	0.929507	0.915147
14	0.956013	0.948908	0.963149	0.940547	0.960863	0.973685	0.956392	0.987948	0.976372	0.961858
15	0.982637	0.966628	0.966827	0.972882	0.975073	0.976052	0.9813	0.992431	0.994036	0.995161
16	0.99441	0.985903	0.987316	0.987619	0.9576	0.960708	0.961348	0.98921	0.987679	0.988177
17	0.938193	0.887864	0.863627	0.821808	0.902457	0.878225	0.829787	0.926044	0.894252	0.917055

From this pair of tables, the RbS is still much better than OS. Although several "1<sub>s</sub>" appear in **Table 4.9**, artificially-raised threshold filters out a lot of voxels; therefore, it is not reasonable.

The previous six tables mainly illustrate the advantages of RbS and AST in reducing the difference of **xpdfs** in terms of the values of  $CDF(0.1)$ , and there is one additional plot of drawing the lines representing the means of each subject for all previous six tables, which is the more evident comparison among them.



**Figure 4.3:** Plots of means for each pair of BPFs from Table 4.4 to Table 4.9

Obviously, the figure above shows that the advantage of RbS is evident and powerful, especially when it is associated with AST, which is an excellent combination. This combination is not only in a high similarity range, but also it no longer fluctuates, in contrast to other combinations.

The evaluations based on  $CDF(0.1)$  are now complete, and we redirect our focus to AUCs-based evaluations.

**Table 4.10:** AUCs of the *xpdf* difference values based on  $r = 0.4$  and OS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.988608	0.952863	0.984174	0.976049	0.88467	0.935856	0.917806	0.994027	0.990876	0.96767
2	0.917912	0.933952	0.886898	0.739812	0.954229	0.930602	0.798978	0.895592	0.768644	0.843655
3	0.956745	0.918076	0.935027	0.917179	0.953367	0.967532	0.951432	0.991627	0.984865	0.974042
4	0.993664	0.989817	0.992604	0.990322	0.989525	0.995136	0.991541	0.997409	0.996084	0.986689
5	0.87029	0.843815	0.910893	0.898535	0.876089	0.891666	0.920779	0.975835	0.958304	0.870136
6	0.934814	0.933107	0.897044	0.860629	0.980113	0.953443	0.920375	0.951459	0.920059	0.959938
7	0.978178	0.969723	0.967768	0.967435	0.985646	0.984827	0.983788	0.992439	0.991418	0.993785
8	0.86624	0.815807	0.807684	0.646326	0.925928	0.913604	0.768316	0.948202	0.821549	0.832873
9	0.88684	0.90771	0.819436	0.745575	0.960099	0.906052	0.831653	0.895547	0.825407	0.908953
10	0.861876	0.738129	0.755135	0.626561	0.868268	0.885587	0.756081	0.973047	0.879697	0.864733
11	0.969007	0.916914	0.906637	0.926904	0.924316	0.916352	0.934988	0.972018	0.983562	0.987429
12	0.961293	0.925576	0.894596	0.866992	0.915646	0.889028	0.863437	0.945306	0.921662	0.957025
13	0.955546	0.958048	0.931568	0.918917	0.985606	0.956458	0.942647	0.957832	0.944342	0.983121
14	0.926695	0.910473	0.889589	0.897048	0.975386	0.95622	0.958431	0.973273	0.975753	0.988946
15	0.983116	0.986323	0.982063	0.983651	0.985517	0.980226	0.980636	0.990945	0.991422	0.996771
16	0.977567	0.979381	0.965829	0.949368	0.992975	0.985036	0.967497	0.982262	0.966993	0.979603
17	0.860309	0.825508	0.774155	0.790446	0.930236	0.883123	0.913171	0.928055	0.941579	0.95399



The table above indicates the *Area Under Curve*, obtained by *Trapezoidal Rule*, for all 17 subjects each with 10 pairs of **xpdf** difference values based on ten pairs of IIRBPFs-preprocessed data. The situation in **Table 4.10** is based upon  $r = 0.4$  and RbS, and other situations will be presented below in the same order as  $CDF(0.1)$ -based evaluations.

**Table 4.11:** *AUCs* of **xpdf** difference values based on  $r = 0.4$  and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.964096	0.928996	0.917532	0.932008	0.882592	0.87424	0.893127	0.944033	0.949567	0.957806
2	0.933125	0.890536	0.88334	0.879162	0.934672	0.927046	0.925549	0.965812	0.962533	0.966645
3	0.923021	0.904376	0.903363	0.918614	0.94895	0.945269	0.960346	0.957003	0.9652	0.968677
4	0.962515	0.960836	0.955982	0.954326	0.971977	0.971436	0.969096	0.972164	0.969677	0.966395
5	0.912367	0.959518	0.973608	0.961302	0.939608	0.934345	0.924527	0.967074	0.969957	0.961772
6	0.904205	0.937157	0.923927	0.904635	0.973475	0.967465	0.952199	0.949551	0.927717	0.941167
7	0.969836	0.980779	0.961742	0.965251	0.974076	0.951346	0.958797	0.930572	0.945251	0.975012
8	0.910379	0.877038	0.895316	0.858167	0.924726	0.935016	0.897754	0.9656	0.944439	0.936888
9	0.963106	0.938986	0.943551	0.926412	0.943416	0.951513	0.930859	0.961292	0.941281	0.951348
10	0.965176	0.932714	0.913529	0.905251	0.935155	0.914472	0.903662	0.9502	0.933936	0.95354
11	0.936616	0.935209	0.930448	0.89419	0.952352	0.952269	0.915112	0.951969	0.917614	0.925008
12	0.887068	0.839286	0.859519	0.85062	0.922066	0.930548	0.927779	0.965511	0.957813	0.948305
13	0.959091	0.917389	0.876449	0.823557	0.893569	0.850474	0.795747	0.906134	0.86543	0.906212
14	0.957836	0.946864	0.928412	0.899701	0.948203	0.937279	0.90763	0.942653	0.916113	0.946362
15	0.983215	0.977897	0.970268	0.964683	0.971454	0.962268	0.957442	0.967245	0.956183	0.964645
16	0.919492	0.884726	0.88084	0.851564	0.940589	0.938181	0.912584	0.968628	0.954379	0.951353
17	0.900222	0.949309	0.830245	0.824593	0.967554	0.869279	0.883153	0.845269	0.838393	0.931578

In **Table 4.11**, the *AUC* values over the entire table are close to "1" than the condition in **Table 4.10**. Although the difference between these two tables is not as large as the corresponding situations in  $CDF(0.1)$ -based evaluations, the former conditions comprises the entire areas under curves; therefore, even a small difference can be regarded as a "large improvement".

**Table 4.12:** *AUCs* of the **xpdf** difference values based on AST and OS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.995732	0.980483	0.977473	0.994958	0.891847	0.879945	0.965706	0.926894	0.983901	0.993172
2	0.927253	0.945487	0.930764	0.943813	0.972559	0.966405	0.979424	0.961639	0.966386	0.975543
3	0.983581	0.968921	0.962412	0.981337	0.946031	0.942042	0.967527	0.962211	0.982189	0.983162
4	0.909424	0.826395	0.967484	0.882277	0.873537	0.979994	0.919414	0.993451	0.959835	0.845609
5	0.964445	0.983108	0.984132	0.927222	0.990687	0.973628	0.930445	0.890385	0.824155	0.81496
6	0.959125	0.954396	0.934043	0.925766	0.965128	0.955288	0.937678	0.947142	0.934437	0.951275
7	0.943634	0.938945	0.978596	0.959766	0.941298	0.987373	0.966265	0.985824	0.96153	0.904847
8	0.991735	0.999663	0.943069	0.993701	0.997596	0.850533	0.970787	0.615967	0.797278	0.996994
9	0.965217	0.981272	0.949816	0.97062	0.993843	0.970763	0.988253	0.929511	0.963261	0.99419
10	0.999798	0.991807	0.990797	0.878437	0.786573	0.790623	0.524511	0.939367	0.728975	0.722372
11	0.988148	0.982413	0.996033	0.992019	0.978099	0.996581	0.991884	0.995349	0.993028	0.94499
12	0.943546	0.988269	0.931598	0.972432	0.958385	0.932106	0.934521	0.879111	0.924929	0.968324
13	0.961985	0.995863	0.97551	0.908514	0.991766	0.945794	0.856279	0.826477	0.679049	0.826558
14	0.937521	0.963804	0.820144	0.885997	0.960257	0.796543	0.888918	0.770714	0.858879	0.980112
15	0.945306	0.987656	0.978099	0.987036	0.994111	0.984709	0.99081	0.915335	0.936551	0.954118
16	0.967513	0.963042	0.983829	0.99137	0.96585	0.983849	0.991244	0.98874	0.994665	0.981014
17	0.891092	0.859237	0.887354	0.826139	0.9316	0.94634	0.896324	0.968172	0.923295	0.909396

**Table 4.13:** *AUCs* of the **xpdf** difference values based on AST and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.968364	0.964283	0.956166	0.957902	0.963081	0.957572	0.9592	0.964899	0.959321	0.956708
2	0.978404	0.974057	0.973464	0.979711	0.977097	0.975985	0.978466	0.982706	0.97591	0.97777
3	0.964779	0.963414	0.962801	0.953044	0.966578	0.969903	0.962236	0.970068	0.963914	0.96457
4	0.951129	0.957157	0.958804	0.956335	0.964051	0.954875	0.957871	0.961653	0.962716	0.961126
5	0.967849	0.972039	0.964819	0.965449	0.959966	0.972655	0.967233	0.963147	0.961306	0.966108
6	0.964369	0.964318	0.967241	0.963352	0.962519	0.966228	0.966327	0.959261	0.964361	0.962961
7	0.941606	0.924923	0.928117	0.936006	0.937305	0.943129	0.934387	0.944271	0.932838	0.938981
8	0.976933	0.977914	0.975781	0.971676	0.976095	0.978939	0.972505	0.979389	0.975035	0.975432
9	0.978362	0.97693	0.970539	0.972322	0.977798	0.97201	0.975741	0.976053	0.979876	0.975698
10	0.980876	0.976209	0.978095	0.977183	0.975275	0.97989	0.980108	0.975772	0.973524	0.979885
11	0.966957	0.967852	0.969868	0.969147	0.968808	0.969768	0.967101	0.970793	0.970124	0.971831
12	0.967316	0.966207	0.96509	0.969441	0.961962	0.954557	0.965664	0.970107	0.966295	0.964174
13	0.953249	0.955495	0.951646	0.951291	0.946018	0.939322	0.953914	0.963199	0.951046	0.946946
14	0.967047	0.967157	0.953868	0.954114	0.96155	0.962682	0.962475	0.951414	0.955057	0.959405
15	0.954855	0.955027	0.951142	0.951517	0.948196	0.946898	0.954712	0.954669	0.953703	0.950564
16	0.946416	0.954984	0.954539	0.946961	0.952537	0.953903	0.947415	0.9523	0.949784	0.947428
17	0.962478	0.954325	0.951841	0.954124	0.957789	0.955885	0.958403	0.965828	0.957353	0.956015

For this pair of tables, they are both quite similar in the range of values; therefore, it is the RbS again justifying its advantage. For the reason of complete comparisons, the threshold value was once again increased from 0.4 to 0.5, and the two corresponding tables are obtained below:

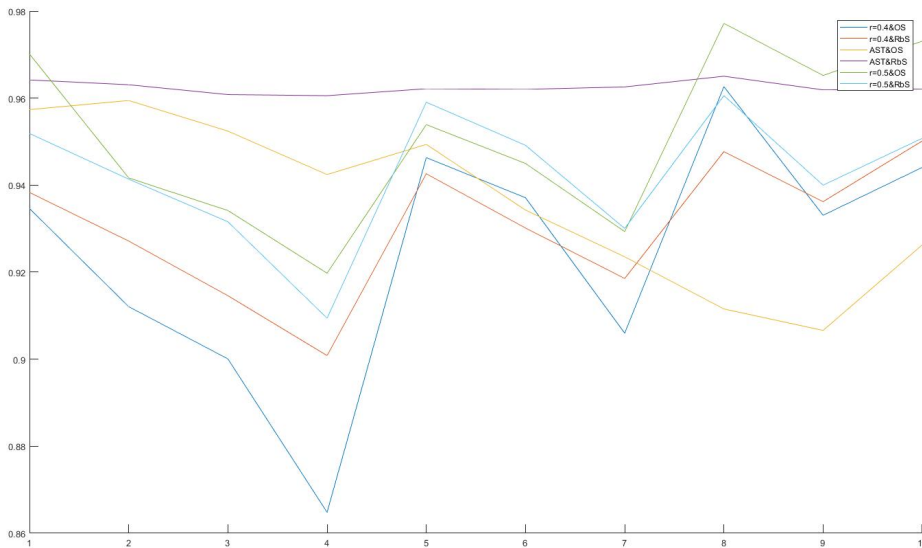
**Table 4.14:** *AUCs* of the **xpdf** difference values based on  $r = 0.5$  and OS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.996132	0.994349	0.978436	0.991295	0.975488	0.953013	0.969664	0.972826	0.989999	0.9974
2	0.908469	0.856008	0.841769	0.747097	0.93557	0.926002	0.836563	0.9552	0.886308	0.902908
3	0.972075	0.962103	0.95437	0.95571	0.987697	0.980298	0.981507	0.988801	0.989143	0.995579
4	0.999167	0.998869	0.99886	0.998957	0.994868	0.995045	0.994986	0.998983	0.998972	0.998901
5	0.975592	0.916557	0.930995	0.841324	0.891323	0.913947	0.809121	0.965334	0.908534	0.87176
6	0.98885	0.962906	0.947427	0.950953	0.954966	0.939846	0.942761	0.982684	0.985606	0.994583
7	0.995241	0.99042	0.9928	0.988839	0.991447	0.993	0.989862	0.997966	0.996885	0.993172
8	0.991985	0.985429	0.920927	0.853112	0.92684	0.812293	0.74513	0.876202	0.808999	0.926539
9	0.954056	0.744304	0.75444	0.77944	0.763841	0.77434	0.799286	0.973903	0.985106	0.979526
10	0.984034	0.955218	0.958978	0.910359	0.93253	0.942843	0.886333	0.988383	0.951079	0.939981
11	0.890407	0.850615	0.837688	0.879365	0.951792	0.940923	0.969686	0.980327	0.993413	0.99708
12	0.928486	0.925875	0.908719	0.90981	0.977268	0.966735	0.971481	0.970573	0.972495	0.985081
13	0.982181	0.984014	0.979389	0.977408	0.998539	0.995711	0.993091	0.993903	0.991072	0.996012
14	0.984667	0.985023	0.981417	0.980721	0.99616	0.993989	0.993092	0.99442	0.992912	0.99667
15	0.991172	0.996594	0.993289	0.99223	0.99838	0.997809	0.997963	0.994744	0.993785	0.997493
16	0.985159	0.985071	0.985609	0.981666	0.997192	0.997105	0.994453	0.995631	0.992859	0.994171
17	0.964728	0.914859	0.915727	0.897598	0.942521	0.941773	0.923485	0.982162	0.971375	0.975746

**Table 4.15:** AUCs of the **xpdf** difference values based on  $r = 0.6$  and RbS

Sub.#	BPFs(1&2)	BPFs(1&3)	BPFs(1&4)	BPFs(1&5)	BPFs(2&3)	BPFs(2&4)	BPFs(2&5)	BPFs(3&4)	BPFs(3&5)	BPFs(4&5)
1	0.968437	0.971923	0.945279	0.923062	0.973251	0.933471	0.909653	0.929078	0.897377	0.934765
2	0.986233	0.967225	0.947955	0.931876	0.945572	0.913696	0.904817	0.947756	0.935686	0.959462
3	0.955313	0.951171	0.941809	0.94436	0.959185	0.951236	0.961202	0.963292	0.964627	0.970198
4	0.980969	0.974638	0.972302	0.976442	0.978027	0.977071	0.98176	0.983787	0.987618	0.992195
5	0.96641	0.933358	0.931983	0.943496	0.939733	0.941729	0.93772	0.942932	0.927061	0.926337
6	0.94934	0.949666	0.93356	0.905437	0.970152	0.956106	0.929975	0.962584	0.937367	0.953676
7	0.985054	0.988974	0.986148	0.982481	0.993271	0.990304	0.988467	0.984284	0.979611	0.982748
8	0.943706	0.937333	0.945669	0.867581	0.963257	0.960426	0.88903	0.965604	0.899512	0.898605
9	0.943481	0.911816	0.898838	0.88015	0.942778	0.9364	0.91461	0.955682	0.934938	0.950237
10	0.925457	0.931107	0.904579	0.891769	0.972467	0.958312	0.949943	0.945751	0.933832	0.963006
11	0.882087	0.902114	0.856051	0.835185	0.97161	0.940983	0.928575	0.932876	0.915855	0.954923
12	0.951959	0.968573	0.960016	0.881631	0.973664	0.967671	0.902994	0.95957	0.884571	0.899502
13	0.872122	0.826146	0.831756	0.773872	0.925066	0.919535	0.883237	0.960935	0.929507	0.915147
14	0.956013	0.948908	0.963149	0.940547	0.960863	0.973685	0.956392	0.987948	0.976372	0.961858
15	0.982637	0.966628	0.966827	0.972882	0.975073	0.976052	0.9813	0.992431	0.994036	0.995161
16	0.99441	0.985903	0.987316	0.987619	0.9576	0.960708	0.961348	0.98921	0.987679	0.988177
17	0.938193	0.887864	0.863627	0.821808	0.902457	0.878225	0.829787	0.926044	0.894252	0.917055

**Table 4.14** and **Table 4.15** are also quite similar. Comparisons will be more evident among the past six tables by plotting the means along columns for every table.



**Figure 4.4:** Plots of means for each pair of BPFs from Table 4.10 to Table 4.15

Numerically, those six plotting lines in the figure above are not considerably different, but as previously discussed, for the values representing AUCs, even a relatively small difference matters for recommending selection of seed and threshold. Obviously, in **Figure 4.4**, the combination of AST and RbS is still the most appropriate.

The combination of AST and RbS yields very reliable and stable similarities between different IIRBPFs-preprocessed data due to the AST's following-up to each subject's properties and the high reliability of RbS. Different subjects have different properties, therefore its potential threshold value is also different and "subtly changed"; it is no longer reasonable for arbitrarily choosing a single value as a threshold value, regardless of whether it is based on "experience". The AST perfectly solves this problem. RbS also adaptively yields the seed on the basis of the voxels across the entire brain, therefore it always builds a seed based on "effective" voxels for different subjects. With both the reliable and safe threshold value and seed, the PCC cloud and other important variables can be obtained in a much more reasonable manner.

Before the end of this topic, it is necessary to discuss what exact threshold value selected by the AST algorithm for each set of data. Secondly, the number of voxels in the PCC cloud based on AST is also presented for further illustration of seed evaluations.

**Table 4.16:** Threshold values selected by the AST algorithm on the situation of OS

Sub.#	BPF(1)	BPF(2)	BPF(3)	BPF(4)	BPF(5)
1	0.274693	0.222655	0.231766	0.235654	0.184559
2	0.543607	0.510234	0.450738	0.463251	0.448139
3	0.307701	0.279913	0.28008	0.272949	0.247172
4	0.097494	0.090196	0.109359	0.064573	0.086821
5	0.552929	0.480519	0.356104	0.45585	0.391393
6	0.333637	0.349384	0.27091	0.27476	0.294246
7	0.136062	0.120574	0.118613	0.071118	0.079504
8	0.41043	0.333599	0.242732	0.370396	0.23511
9	0.624799	0.641597	0.604538	0.59008	0.570539
10	0.289436	0.153972	0.21934	0.206602	0.260193
11	0.458281	0.43802	0.426183	0.352677	0.33339
12	0.400423	0.350727	0.319152	0.327272	0.295247
13	0.122504	0.111195	0.064123	0.076264	0.1181
14	0.209963	0.209287	0.180953	0.220296	0.197627
15	0.172374	0.160002	0.121895	0.115844	0.114398
16	0.236377	0.24926	0.223756	0.181018	0.147648
17	0.396995	0.365285	0.406486	0.375263	0.328746

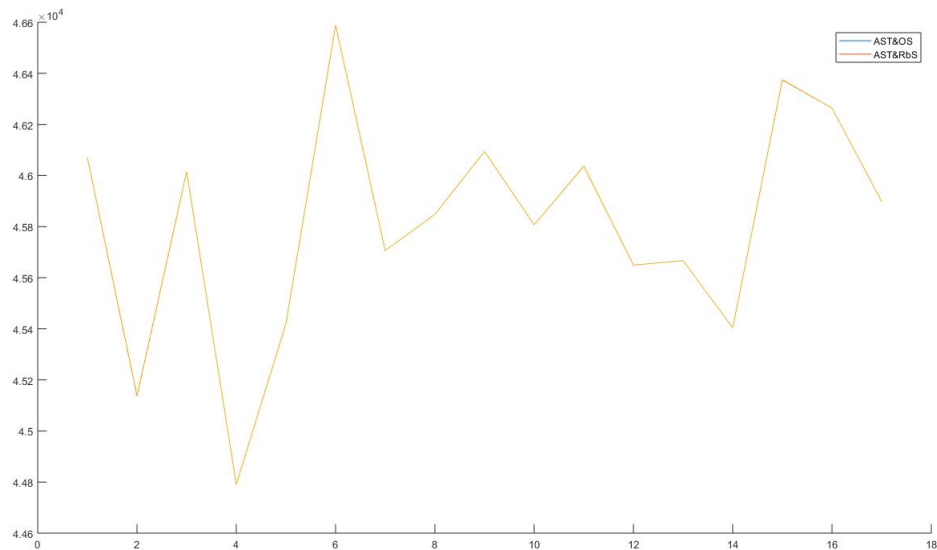
In the table above, due to the significant variations in properties of different subjects, the thresholds vary greatly along the vertical direction, while in the horizontal direction, they vary less evidently, but are still different because of some new frequency components appearing because of the effects from different IIRBPFs. The table below shows the thresholds selected by the AST algorithm with respect to the situation of RbS.

**Table 4.17:** Threshold values selected by the AST algorithm on the situation of RbS

Sub.#	BPF(1)	BPF(2)	BPF(3)	BPF(4)	BPF(5)
1	0.555431	0.532276	0.527396	0.470133	0.450542
2	0.647697	0.65928	0.649229	0.653711	0.602796
3	0.481516	0.517819	0.474489	0.509856	0.468694
4	0.369148	0.336267	0.313676	0.287636	0.3118
5	0.735484	0.70972	0.783183	0.689302	0.682313
6	0.481915	0.475878	0.439491	0.435968	0.430153
7	0.360218	0.363179	0.341174	0.353925	0.312367
8	0.683718	0.632576	0.660185	0.650961	0.602663
9	0.694333	0.702597	0.694224	0.671476	0.672109
10	0.655213	0.650721	0.574604	0.587952	0.606542
11	0.606533	0.575285	0.559257	0.51673	0.54133
12	0.541271	0.536621	0.541501	0.483249	0.527378
13	0.557319	0.516964	0.488135	0.480487	0.519238
14	0.460643	0.443943	0.390326	0.410157	0.406898
15	0.318581	0.333509	0.309448	0.29228	0.285607
16	0.39379	0.36632	0.337191	0.346456	0.328456
17	0.659295	0.635647	0.587545	0.591345	0.600621

As indicated by **Table 4.17**, the threshold values largely increase because of the effects exerted by RbS.

The AST also places a interesting effects on the number of voxels in the PCC cloud. Because AST "follows-up" every situation, therefore, the number of voxels remains the same for each subject, regardless of the effects from IIRBPFs. The following figure represents the number of voxels in the PCC cloud based on AST accompanied with both the OS and RbS.



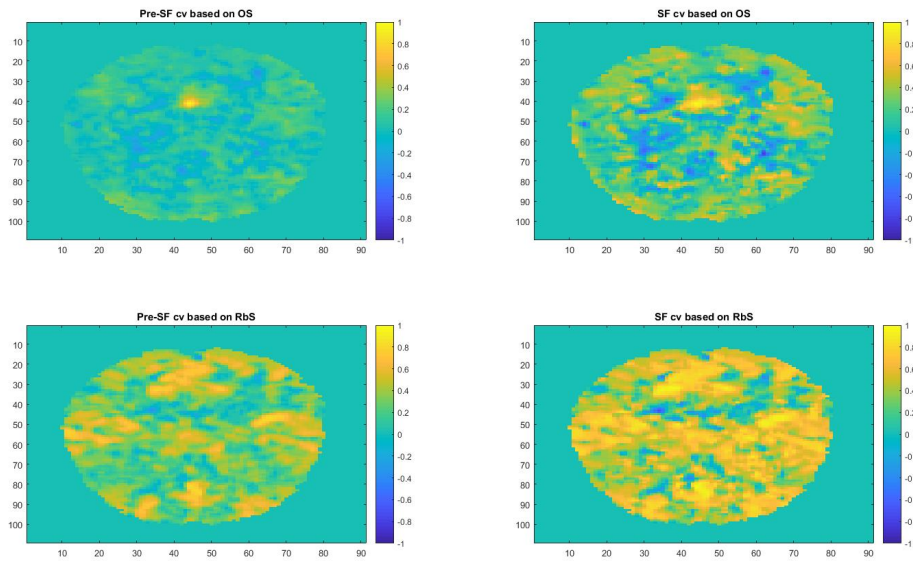
**Figure 4.5:** Number of voxels in the PCC cloud produced by AST with both OS and RbS applied

Interestingly, in the figure above, for every subject (along the horizontal axis), the number of voxels remains almost exactly same, regardless of the application of IIRBPFs and the selection of seed methods (with an extremely small difference). Again, they are stable and reliable.

Finally, a summary can be made for this thesis. For most reasoning processes in this thesis, the seed methods and threshold selections are inseparable from each other. The original seed is the one selecting from the area in the PCC network, and its selection is mainly based on "experience", and is not reliable. Furthermore, no matter which single seed is selected, there is always the possibility that it will not be sufficient to be considered as a "good" seed. However, the RbS is the most suitable seed. For the same reason, the artificially-set threshold values are also not reasonable enough, instead, the AST algorithm is the most appropriate method. In sum, the combination of AST and RbS is excellent in terms of reliability and stability. Nevertheless, there is only one concern - if it is desired to generate more voxels in the PCC cloud, more considerations should be taken before employing the RbS, since the higher threshold values triggered by RbS may filter out more voxels than does OS. In other hands, this situation may be a trade off.

## 4.5 Subspace Filtering Evaluations and Interpretations

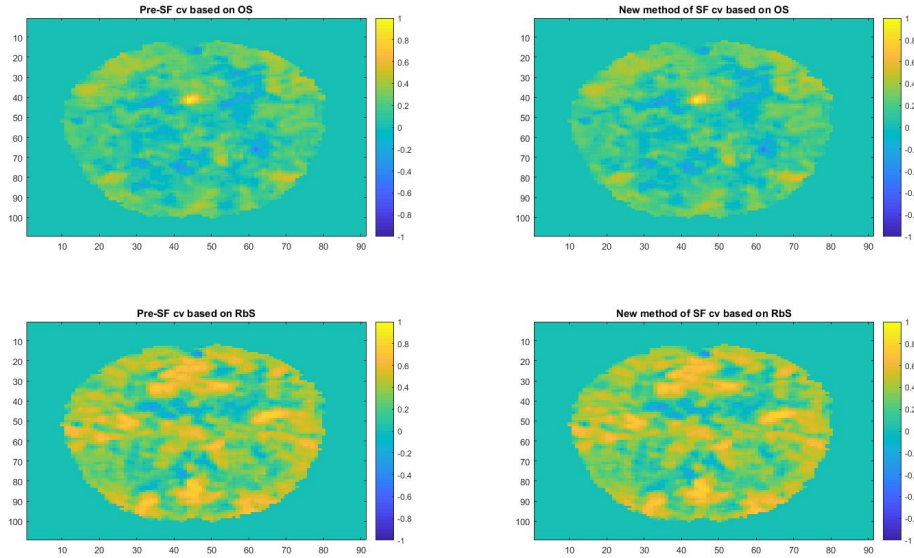
Following the previous chapter, some concrete experimental results and analysis will be presented below for supporting the assertion that Subspace Filtering (SF) is not necessary in this thesis. Previously, one of the methods for SF involves choosing the projection dimension as 10, which is also the number of the ten highest singular values by decomposing the original signal matrix. However, as discussed before, it is not the most correct and reasonable method since it forces an increase in the correlations among voxels. The following is a figure of the results in this situation.



**Figure 4.6:** Results of not-exactly-correct Subspace Filtering

As indicated by the figure above, there is a large, but not reasonable, difference between pre-SF and SF conditions in each pair. By forcing the increase in the correlation coefficients ( $cv$ ), even the stable RbS becomes unstable - that is, the  $cv$  based on RbS has been changed too much. The reason for these conditions is that too much other information has been discarded. Therefore, the ten highest singular values could not guarantee the required completeness retained compared to the original signals. This method should therefore be removed.

Another method of  $\mathbf{cv}$  is more mathematically strict, and it is yielded by maintaining the 99% of the original variance. The situation can be plotted as below.



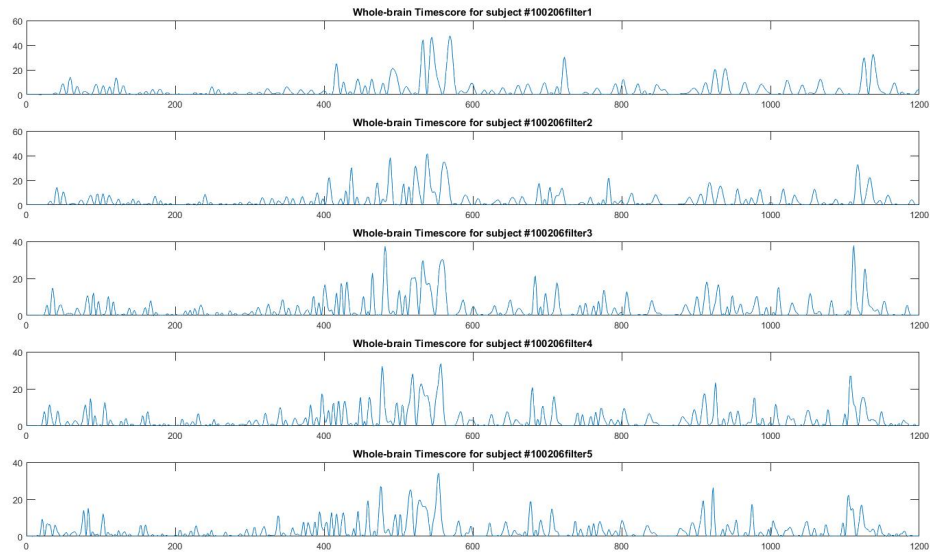
**Figure 4.7:** Results of correct, but not necessary, Subspace Filtering

As expected, the figure above does not show a detectable difference in each pair. Not inconsistent with the previous discussions, this is because of the large value of projection dimension selected for maintaining the 99% of variance in original signals. For the case above, the projection dimension has been selected as 111 out of 200. Therefore, in this thesis, the subspace filtering can be considered unnecessary.

#### 4.6 Whole-Brain Timescore Evaluations and Interpretations

Regarding the Whole-Brain Timescore (WBTc), we simply concentrate on the relation between the bandwidth of those IIRBPFs, preprocessing the brain signals, and their corresponding WBTc for each subject. Associated with the theory of *Digital Signal Processing*, it may be expected the change of spike width and increase of new frequency components in WBTc has some connections with the change of bandwidth of IIRBPFs. The following figure shows the WBTc for one of the subjects which has been preprocessed with five different IIRBPFs.





**Figure 4.8:** One example of Whole-Brain Timescore based on five BPFs-preprocessed brain data

Examining the above figure carefully, as bandwidth of IIRBPFs is increased, some new frequencies appear, the height and width of spikes change evidently, and the change in number of spikes is also detectable.

#### 4.7 Overall Evaluations and Interpretations

This chapter mainly presents many experimental results and their analyses for making comparisons among several seed and threshold methods. In the very beginning of this project, a single seed selected from the PCC network, and the threshold value is set by experience. However, these choices are not precise, and they cannot stand against the tests from all sides. Although the PCC cloud and Pearson's Correlations Coefficients are not very difficult to mathematically obtain, their reliability and reasonability are more important. Aside from that, they are just the inception for other more advanced research in this area; hence, they must be yielded by a very reasonable method for inducing high-level analyses. All factors which have been listed contribute to the excellence of the Region-based Seed and Adaptively-Selected Threshold, and they are recommended for researchers in this area.

## Chapter 5

### Conclusions and Future Works

#### 5.1 The summary of this thesis

This thesis mainly illustrates the effects of five different IIR band-pass filters in the preprocessing stage on fMRI image analysis. A procedure has been designed for evaluating these effects in terms of the corresponding changes triggered on several variables during that procedure. By the comprehensive evaluations, it has been decided that the Region-based Seed and Adaptively-Selected Threshold is the best combination for not being affected too much by the changes of IIRBPFs in the preprocessing phase. This decision is based on two reasons. First, this combination precisely selects the threshold for every IIRBPF-preprocessed signal in each subject. Second, it is very reliable and stable with respect to the number of voxels in the PCC cloud, and to the connection probabilities.

In the analysis of the Subspace Filtering, a comparison has been applied to prove it is not necessary. In the case of the Whole-Brain Timescore, this variable typically characterizes the effects of IIRBPFs in preprocessing stage.

#### 5.2 Perspective Works

A possible direction left by this thesis, which is worth exploring, is that the combination of RbS and AST is indeed unaffected too much by the IIRBPFs; nevertheless, a small but noticeable range of difference, shown in **Table 4.3** and **Table 4.4**, need further efforts for discovering whether there are any real important signals in that range because of the effects from those five different IIRBPFs.

## BIBLIOGRAPHY

- [1] R. Matthew Hutchison, Thilo Womelsdorf, Elena A. Allen, Peter A. Bandettini, Vince D. Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H. Duyn, Gary H. Glover, Javier Gonzalez-Castillo, Daniel A. Handwerker, Shella Keilholz, Vesa Kiviniemi, David A. Leopold, Francesco de Pasquale, Olaf Sporns, Martin Walter, Catie Chang Neuroimage. Dynamic functional connectivity: Promise, issues, and interpretations. Author manuscript; available in PMC 2013 Oct 25. Published in final edited form as: Neuroimage. 2013 Oct 15; 80: 10.1016/j.neuroimage.2013.05.079. Published online 2013 May 24. doi: 10.1016/j.neuroimage.2013.05.079 PMID: PMC3807588
- [2] Enzo Tagliazucchi, Frederic von Wegner, Astrid Morzelewski, Sergey Borisov, Kolja Jahnke, Helmut Laufs Automatic sleep staging using fMRI functional connectivity data NeuroImage, Volume 63, Issue 1, 15 October 2012, Pages 63-72
- [3] Tagliazucchi E, Siniatchkin M, Laufs H and Chialvo DR (2016) The Voxel-Wise Functional Connectome Can Be Efficiently Derived from Co-activations in a Sparse Spatio-Temporal Point-Process. Front. Neurosci. 10:381. doi: 10.3389/fnins.2016.00381
- [4] Xu N, Spreng RN and Doerschuk PC (2017) Initial Validation for the Estimation of Resting-State fMRI Effective Connectivity by a Generalization of the Correlation Approach. Front. Neurosci. 11:271. doi: 10.3389/fnins.2017.00271
- [5] Xu Y and Lindquist MA (2015) Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data. Front. Neurosci. 9:285. doi: 10.3389/fnins.2015.00285
- [6] Zhang H, Chen X, Zhang Y and Shen D (2017) Test-Retest Reliability of High-

- Order Functional Connectivity in Young Healthy Adults. *Front. Neurosci.* 11:439. doi: 10.3389/fnins.2017.00439
- [7] Heilbronner SR, Hayden BY and Platt ML (2011) Decision salience signals in posterior cingulate cortex. *Front. Neurosci.* 5:55. doi: 10.3389/fnins.2011.00055
- [8] Lord AR, Li M, Demenescu LR, van den Meer J, Borchardt V, Krause AL, Heinze H-J, Breakspear M and Walter M (2017) Richness in Functional Connectivity Depends on the Neuronal Integrity within the Posterior Cingulate Cortex. *Front. Neurosci.* 11:184. doi: 10.3389/fnins.2017.00184
- [9] Aurich NK, Alves Filho JO, Marques da Silva AM and Franco AR (2015) Evaluating the reliability of different preprocessing steps to estimate graph theoretical measures in resting state fMRI data. *Front. Neurosci.* 9:48. doi: 10.3389/fnins.2015.00048
- [10] Lecture notes from Coursera.org url:<https://www.coursera.org/learn/machine-learning/resources/kGWsY>
- [11] 1200 Subjects Data Release, CONNECTOME COORDINATION FACILITY url:<https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release>
- [12] Wikipedia's Explanation of Trapezoidal Rule url:[https://en.wikipedia.org/wiki/Trapezoidal\\_rule](https://en.wikipedia.org/wiki/Trapezoidal_rule)
- [13] X.-N. Zuo, A.D. Martino, C. Kelly, Z.E. Shehzad, D.G. Gee, D.F. Klein, et al., The oscillating brain: Complex and reliable, *NeuroImage*. 49 (2010) 14321445. doi:10.1016/j.neuroimage.2009.09.037.