Title

Advancing the neurophysiological understanding of stress, a study based on recorded Electroencephalography (EEG) data in real-world classroom



Author

Wang, Siwen

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

# Advancing the neurophysiological understanding of stress, a study based on recorded Electroencephalography (EEG) data in real-world classroom

A Thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

in

Bioengineering

by

Siwen Wang

Committee in charge:

Professor Tzyy-Ping Jung, Chair
Professor Gert Cauwenberghs, Co-Chair
Professor Gabriel A. Silva

2022

The Thesis of Siwen Wang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT OF THE THESIS

Advancing the neurophysiological understanding of stress, a study based on recorded Electroencephalography (EEG) data in real-world classroom

by

Siwen Wang

Master of Science in Bioengineering

University of California San Diego, 2022

Professor Tzyy-Ping Jung, Chair
Professor Gert Cauwenberghs, Co-Chair

Stress has been a prevalent part of modern life, particularly during the time of the pandemic. While short-term stress may cause little harm to productivity, if left untreated for a long period of time, it could eventually lead to anxiety and depression, which significantly decrease the quality of life. Such a problem is even more severe among students. A recent survey conducted in 2020 with 15,346 graduate and professional students had shown that 32% of them screened positive for major depressive disorder [4]. This study, which took place in a real-world classroom, aims to uncover some of the neuronal mechanisms behind stress among young adults with their recorded EEG data. Such understanding could provide the theoretical foundation for

stress reduction and prevention techniques such as real-time stress detection and non-invasive neurostimulation. This thesis is structured into four chapters. In the first chapter, the author introduced the importance of the problem, the experiment design, and the dataset. In the next chapter, the author began the stress analysis by studying the power spectral density of the 30 EEG electrodes. Results showed that most theta (4-8Hz) and alpha (8-13Hz) frequency bands in the frontal, central, and right parietal regions showed statistical significance among the elevated and normal stress groups. While the power spectra information is helpful for understanding stress, it is important to remember that EEG signals are mixtures of source activities, which makes the underlying source activities and locations unknown. Thus, in chapter three, the author decomposed the EEG data into independent sources using Independent Component Analysis (ICA) and analyzed the effect of stress in terms of cortical source activities. The results showed that some sources responded to stress while others did not. One limitation of this study was that each source was analyzed individually. Thus, in the final chapter, the author focused on exploring the interactions between regions (i.e. effective connectivity) under stress. The results showed that the information inflow and outflow near the central region were statistically different between the elevated stress and normal stress groups.

# Chapter 1

# Introduction

## 1.1 Background

Stress has been a prevalent part of modern life, particularly in the time of a global pandemic. According to a survey conducted by American Psychological Association (APA) in 2020, nearly 80% of people reported the pandemic being a significant source of stress in their life. While stress itself can be temporary, several studies [5] [6] have shown that it is highly correlated to depression, which can significantly decrease the quality of life [7]. The problem of stress, anxiety, and depression is even more prevalent in graduate students. According to a study done by Harvard University in 2018 [8], 18% of graduate students experienced moderate or severe symptoms of depression and anxiety - more than three times the average population. Unfortunately, this number rose even higher due to the pandemic. A recent survey done in 2020 with 15,346 graduate and professional students had shown that 32% of graduate and professional students screened positive for major depressive disorder, while 39% screened positive for generalized anxiety disorder [4].

Despite stress being so prevalent and having an impact on almost everyone's life, the neuronal mechanism underlying it is still not fully understood. Past study has shown several brain regions are associated with stress regulation [9], but the direction of neuronal network interaction under stress remains largely unexplored. Several neurotechnologies with different spatial and temporal resolutions to study neural activities are available as shown in Fig. 1.1.

Electroencephalography (EEG) is often preferred in non-clinical neuroscience research and modern Brain-Computer interface (BCI) applications due to its non-invasiveness. Recent advance in EEG sensor technology has enabled dry and non-contact recording to become a reality, which makes EEG a convenient and practical technique for mobile imaging [10]. With that said, EEG is an ideal candidate for this experiment that took place in a real-world classroom during classes.

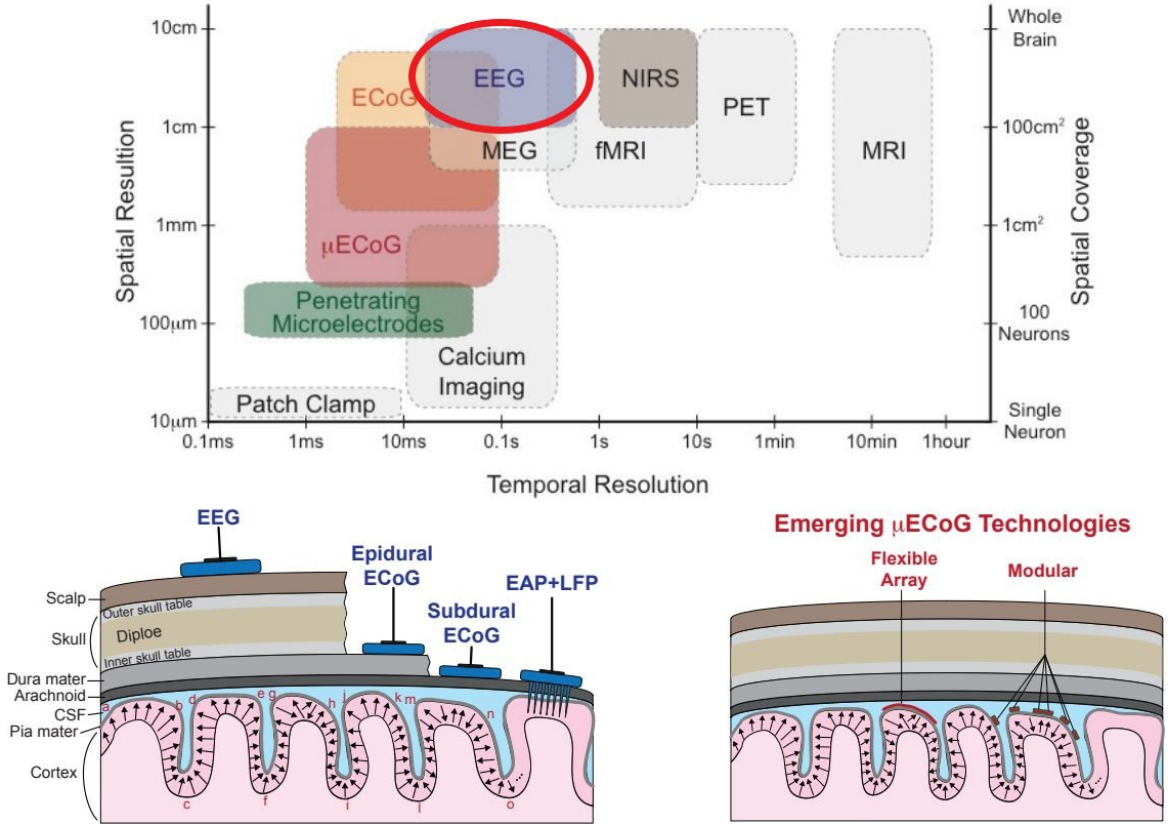


**Figure 1.1.** Minimally Invasive Neurotechologies with spatial and temporal resolution (Image from "Silicon Integrated High-Density Electrocortical Interfaces," Proceedings of the IEEE, vol.105 (1), pp. 11-33, 2017. [1])

## 1.2 Research objective and Motivation

The objective of my thesis research, broadly speaking, is to advance the current neurophysiological understanding of stress using recorded EEG data. To be more specific, this consists of three parts: in the first part, I focus on analyzing the impact of stress by studying

the frequency power information of all recording electrodes between elevated stress and normal stress groups. This gives the readers a comprehensive picture of how stress could alter regional cortical activities reflected by the increase in power in certain frequency bands. Such understanding is also informative for feature selection in traditional machine learning models; in part two, I focus on the activation of different cortical regions with versus without stress. This gives a better understanding of the cortical locations where stress could impact and alter the neuronal activation pattern; in the last chapter, I investigate the effective connectivity between cortical regions that are affected by stress. This gives a better understanding of the dependency of activation regions as well the direction of network propagation due to stress. With a better understanding of the neural mechanism under stress, not only could we build machine learning models to detect stress and offer personal health recommendations, but we could also come up with novel therapy methods based on neurostimulation.

## 1.3 Experimental Design & Dataset

### 1.3.1 Participants

Eighteen graduates ($24 \pm 1.2$ years old, 10 male and 8 female) from the National Chiao Tung University, Taiwan were recruited to participate in this study during the first semester of the 2014-2015 school year [11].

### 1.3.2 Depression, Anxiety, and Stress Scales

The Depression, Anxiety, and Stress Scales (DASS) survey was used in this study to access the depression, anxiety, and stress level of each participants. A shortened 21-question version was used in this study instead of the original 42-question version [12]. Fig. 1.2 shows a few sample questions from the survey, and score were calculated in the end to determine the severity of each condition.

**Depression, Anxiety and Stress Scale (DASS21)**

For each statement below, please circle the number in the column that best represents how you have been feeling in the last week.

| Statement | Did not apply to me at all | Applied to me to some degree or some of the time | Applied to me a considerable degree or a good part of the time | Applied to me very much or most of the time |
|---|---|---|---|---|
| 1. I found it hard to wind down | 0 | 1 | 2 | 3 |
| 2. I was aware of dryness of my mouth | 0 | 1 | 2 | 3 |
| 3. I couldn't seem to experience any positive feeling at all | 0 | 1 | 2 | 3 |
| 4. I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion) | 0 | 1 | 2 | 3 |
| 5. I found it difficult to work up the initiative to do things | 0 | 1 | 2 | 3 |
| 6. I tended to over-react to situations | 0 | 1 | 2 | 3 |

**Figure 1.2.** DASS-21 Survey sample questions

### 1.3.3 Data Collection

The experiment took place in a real-world classroom while class is in session. The signal acquisition was carried out using a Compumedics Neuroscan system, with sampling rate 1000 Hz. Thirty sensors were placed on the head using a 32-channel Quik-Cap. Prior to each data collection session, the participates were asked to fill out the DASS survey for stress, anxiety and depression assessment. After the survey had been completed, a 5-minute eye open resting session was taken as the baseline. Data was continuously being recorded when class is in session, and several stressful events such as pop quiz were purposefully presented to the participates to alter the cognitive state. The events were also time locked for later data analysis. Each class session lasted around 60 minutes, and in the end of the class, another 5-minute eye-open resting session was taken. DASS survey was also repeated to assess any possible stress level elevation due to the designed stressors. Fig. 1.3 shows a graphical representation of the events of each data recording session.

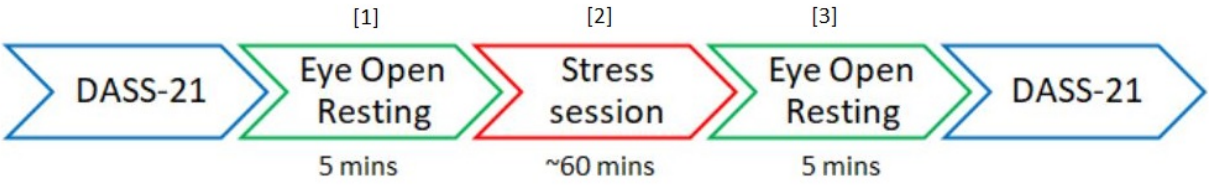


**Figure 1.3.** Data recording session with events label and corresponding length

**Figure 1.4.** Data recording in session

### 1.3.4 Dataset and preprocessing pipeline

While there are many interesting data analysis studies that can be carried around the topic of stress, anxiety, and depression, this paper focuses only on stress in the resting state. Thus, only the recordings from the first 5-minute eye-open resting session, that is the period labeled as [1] in Fig. 1.3, are used. Here on and after, when a dataset is mentioned, it refers to the recordings of the first 5-minute eye-open resting session. A total of 90 recordings (19 elevated stress, 71 normal stress) are used in subsequent analysis. Similar to any other data

collected in the real world, EEG data also require pre-processing for noise removal. First, a band-pass filter between 1 Hz and 45 Hz was implemented to reduce high-frequency artifacts and some eye activity. The data were re-referenced to the average of the left and right Mastoids channels. Next, a group of poorly recorded channels was identified and removed using EEGLab clean_rawdata function. The selection was made based on the contact quality with the scalp and the correlation between neighboring electrodes. Next, artifact subspace reconstruction (ASR) [13] was implemented to remove high amplitude artifacts, so as to further reduce the effect of the eye and muscle movement artifacts. Finally, Independent Component Analysis (ICA) was implemented, and non-cortical individual components (ICs) were automatically rejected using the EEGLab IC automatic labeling tool. The threshold for rejection was set to be 20%, that is an IC is automatically rejected if the probability of it being a brain component falls under 20%. More detailed explanations about ICA will be covered in chapter three.

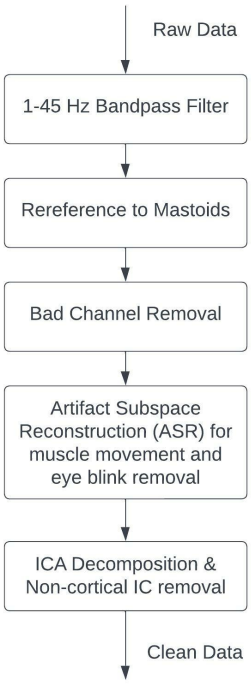


**Figure 1.5.** Pre-Processing Pipeline

# Chapter 2

# Stress analysis using frequency power information of recording electrodes

## 2.1 Introduction

In this chapter, the author analyzed the effects of stress represented by the frequency and power information of the 30 EEG electrodes. Usually, EEG activities are categorized into five bands representing different frequency ranges. They are the delta bands representing 0.5-4Hz, theta bands representing 4-8Hz, alpha bands representing 8-13Hz, beta bands representing 13-30Hz, and gamma bands representing 30Hz and above. The central goal of this chapter is to identify the channels (electrodes) and frequency bands where there is a significant difference between the elevated stress and normal stress groups. Such findings could help us to better understand the brain state under stress. Furthermore, it could also provide a basis for feature selection in machine learning algorithms and be applied to real-time BCI applications such as stress detection.

## 2.2 Method

### 2.2.1 Fourier Transform

Fourier Transform, a method to decompose any time series into a set of sinusoids that are characterized by its frequency and amplitude, is commonly used in the field of science and

engineering [14]. This concept may sound foreign to some people, but in fact, all of us are performing this “transformation” every time we listen. Our brain makes sense of sound, waves generated by the movement of air molecules, by decomposing it into distinct frequencies and amplitudes which gives rise to pitch and loudness. This provides an excellent example of how information could be better understood by looking at the same wave from a different perspective, i.e. the frequency domain. This exact same idea applies to EEG data analysis. Though it is possible to identify characteristic waveforms such as eye blink and muscle movement in EEG recordings by visual inspection, for the complex tasks such as stress detection and characterization, it is much better off to analyze the data in the frequency domain. There are two major methods to compute the frequency and power distribution (i.e. power spectrum density) of time series: the Welch method [15] and the wavelet transform [16]. Each method has some advantages and disadvantages and is often selected based on the application. Here, the author chose the Welch method for its simplicity and fit. Fig.2.1 shows the script the author used for power spectrum density (PSD) computation.

PSD.m

```
%Normal Stress Group
dataFiles = dir('N\*.set');
LengthFiles = numel(dataFiles);
%Elevated Stress group
dataFilesI=dir('I\*.set');
LengthFilesI = numel(dataFilesI);
M_f=[]; %normal
M_f_I=[];%increased

%Specify Channel names
allchanlabel={'FP1','FP2','F7','F3','FZ','F4','F8','FT7','FC3','FCZ','FC4','FT8','T3','C3','CZ','C4','T4','TP7','CP3','CPZ','CP4','TP8','T5','P3','PZ','P4','T6','O1','OZ','O2'}
for s = 1:numel(allchanlabel)
    label=allchanlabel(s);
    M_f=[]
%Load all data in elevated stress group
for i = 1:LengthFilesI
    data_name = fullfile('I', dataFilesI(i).name);
    EEG1=pop_loadset(data_name);
    x={EEG1.chanlocs.labels};

    index=find(strcmp(x,label));
    if isempty(index)==1
        continue;
    end
    %PSD computaton using the Welch method
    [spectra,freqs] = spectopo(EEG1.data(index, :), 0, EEG1.srate, 'freqrange',[1 50],'plot','off');
    idx_freqs = find(freqs>50,1)-1;
    plt_freqs = freqs(1:idx_freqs); %X values, frequency
    plt_spec =spectra(1:idx_freqs); %Y values, power
    M_f=[M_f;plt_spec];
end
ext='_I.mat';
fname=strcat(label,ext);

%save file
save(fname,'M_f')
end
```

**Figure 2.1.** Matlab script for PSD calculation

### 2.2.2 Statistical Analysis - Two-sample Independent t-test

Statistical testing is another important concept and analysis in science. It is often conducted after the result has been obtained to verify if it has any significance, as compared to the observation is just due to random chance. There is no precise definition of statistical significance. However, it is usually agreed in the scientific community that a p-value less than 0.05 is sufficient to reject the null hypothesis [17] A null hypothesis states that there is no statistical relationship or significance between the observed samples [18]. In other words, the results are just by random chance. Here, the null hypothesis would be that the elevated and normal stress groups have the same PSD distribution in all frequency bands. To decide whether to accept or reject this hypothesis, a p-value is computed for every frequency band. Each p-value tells us the probability of the null hypothesis being true for that given frequency. As mentioned before, it is generally accepted that to claim statistical significance, p-value must be less than 0.05. That is to say, there is less than 5% chance the results are from chance alone. The author follows the standard p-value cutoff in this paper.

## 2.3 Result and Discussion

Fig. 2.2 to Fig. 2.7 show the PSD profile with the statistical testing results for all 30 EEG electrodes grouped by their placement location. For those who are unfamiliar with the EEG electrodes labels, please refer to Fig. 2.8 for the approximate location where each electrode locates on the scalp surface. There are several interesting findings looking at the plots. First, one can observe that the theta (4-8Hz) and alpha (8-13Hz) frequency bands are most common among all the channels that have any frequency range that is statistically significant. Next, looking at each figure grouped by their spatial location, one can find that all channels located in the frontal, central and right temporal region have some frequency range that is statistically different between the elevated and normal stress groups. Whereas in left temporal channels, no frequency reached statistical significance. Two of the three channels (OZ, O2) in the occipital region have

reached statistical significance in a very narrow frequency band, 10Hz, and 9-10Hz respectively. Three of the five channels (PZ, P4, T6) in the parietal region reached statistical significance in some frequency bands, from 4-11Hz, 6-10Hz, and 10Hz respectively. Several pieces of evidence could be used to make sense of the results. First, we will try to explain from a neurophysiology perspective. In a typical brain anatomy graph, the cerebral hemispheres (the largest part of the brain) are often divided into four lobes: frontal, temporal, parietal and occipital as shown in Fig. 2.9. Each lobe is believed to be responsible for different functions and behaviors. For example, the frontal lobe is well-known for its involvement in language and motor function. Recent research has also shown that it is involved in various cognitive processes such as memory, task planning, attention as well as emotion and personality [19]. Stress is an emotional response, so it logically makes sense that we observe some differences in the frontal channels between the elevated and normal stress groups. It is likely that stress has altered the activation pattern of neurons in the frontal lobe, for instance, more neurons fire at a higher rate, which contributes to the difference in power we measure. A similar explanation could also explain why little statistical difference is observed in the occipital channels, that is, the occipital region is mostly responsible for vision processing. This intuitive makes sense that it is unlikely that stress will alter the processing of visual stimuli, thus little to no difference is observed in this region. Next, we will change direction and talk about a well-known bio-marker called frontal alpha asymmetry (FAA). In 1978, at the Society for Psychophysiological Research annual meeting, Richard David presented a novel finding that positive and negative experiences are associated with different asymmetrical frontal brain activity. [20]. In the experiment, participants were asked to watch portions of a TV program and indicate how much they liked the content, while EEG being recorded from the frontal and lateral regions. The results show that a positively rated scene is associated with greater relative left-hemispheric alpha activation in the frontal region, while right-hemispheric alpha activation in the frontal region associates with negatively rated scenes. Moreover, such difference is not observed in lateral channels [21]. Since this study has published, thousands of groups across the globe have started to explore this novel biomarker and its potential

applications. As of today (10/18/2022), there are more than 69,000 articles published related to FAA. Some studies show that FAA is a promising bio-marker that is significantly correlated with Major Depressive Disorder (MDD) [22]. The study by Mennella's group has also shown that by applying neurofeedback to reduce left frontal activity results in a decrease in negative affect and anxiety [23]. However, there are also studies that claim that it is still unclear if FAA is a reliable bio-marker, particularly for the diagnosis of depression [24] [25]. Combined with the results observed here, that is, all frontal channels have some frequency range that is statistically different across the elevated and normal stress group, it is certainly possible that this is due to the level of activation difference of the right and left hemispheres in frontal regions. Though this is a possible explanation, the author has also computed FAA on the classroom dataset and found no apparent association between FAA score and stress level.

## 2.4 Conclusion

In this chapter, I analyzed the effects of stress represented by the power spectrum information of all 30 recording electrodes. Fourier transform is used to convert recorded signals from the time domain to the frequency domain, then statistical testing is implemented to obtain results unique to one group. As the results show, all channels located in the frontal, central and right temporal regions have mostly theta and alpha frequency bands that are statistically different between the elevated and normal stress groups. Less difference is observed in the occipital and parietal regions, and no difference is observed in the left temporal region. The author hypothesize that the difference in activation across regions is due to the different functions of cerebral hemispheres lobes. As neuroscience research has shown that frontal regions are mainly responsible for the cognitive process such as emotion, whereas other regions may not play a major role when the brain is under stress.

This concludes chapter two, and in the next chapter, we will study stress from the perspective of independent sources.

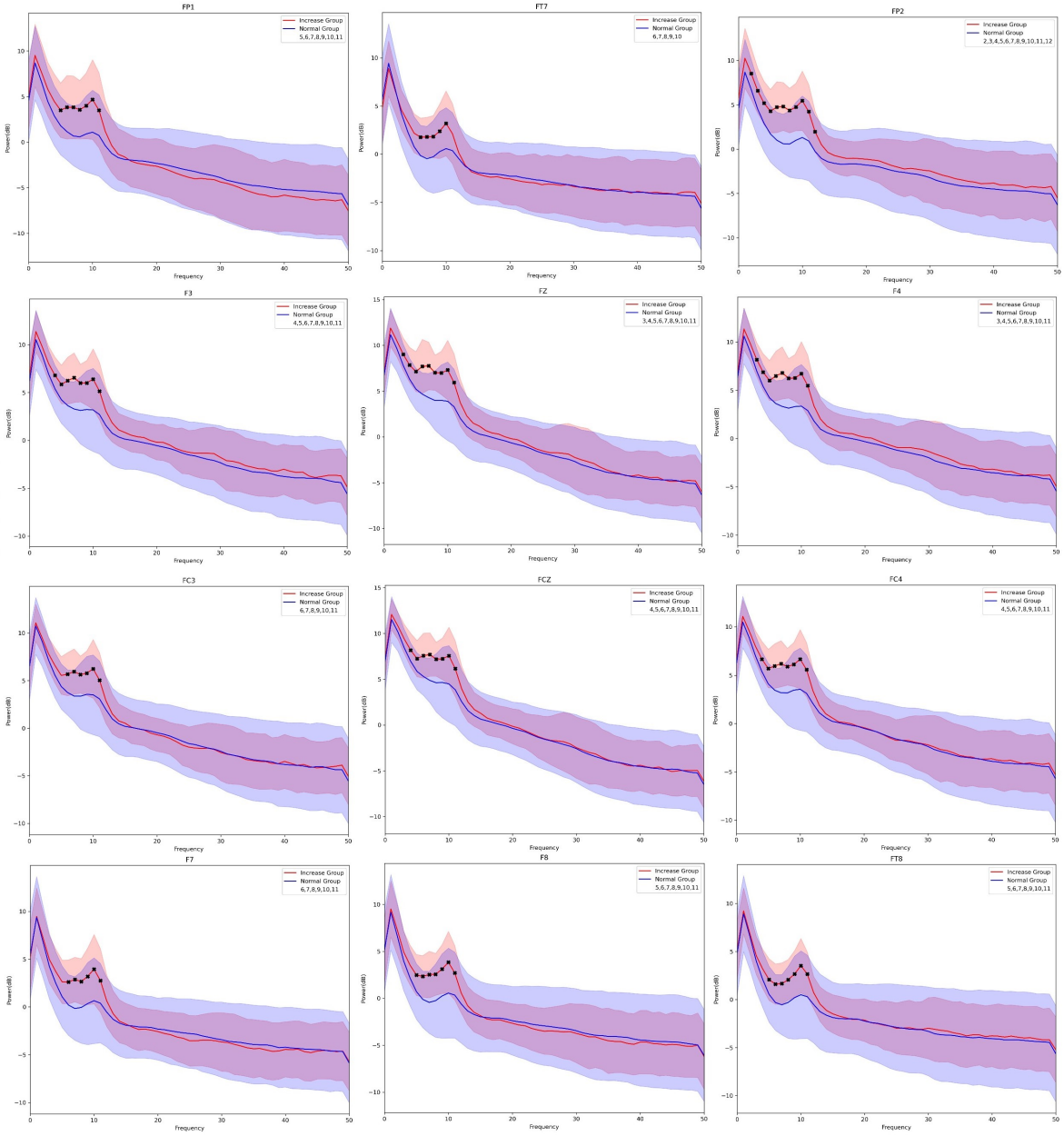


**Figure 2.2.** Frontal channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

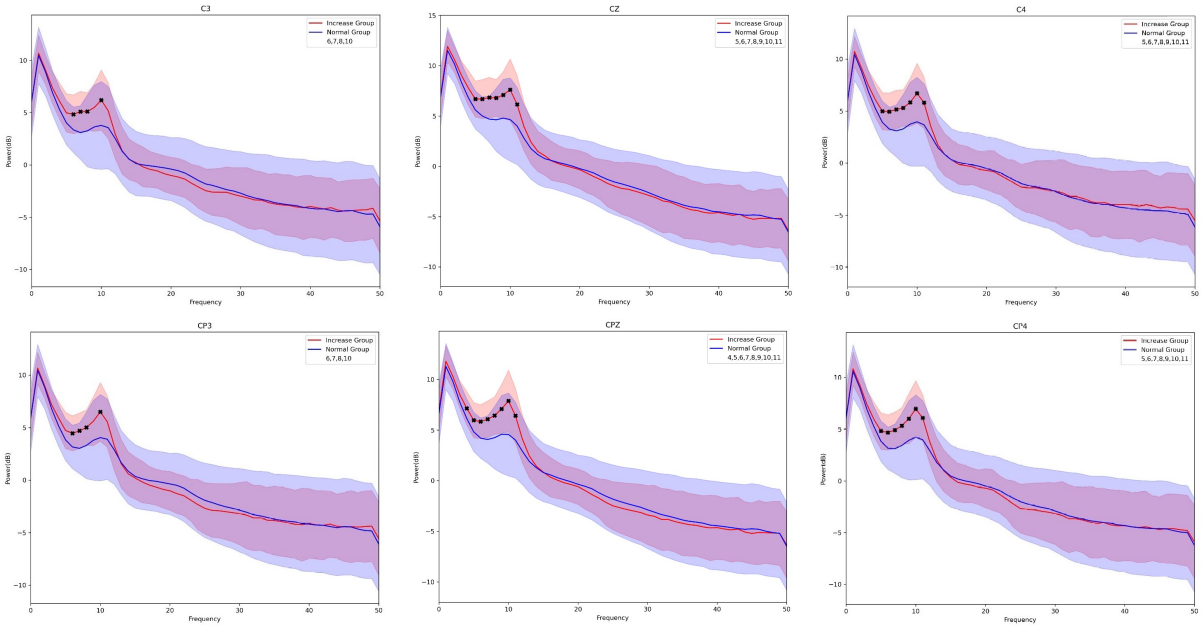


**Figure 2.3.** Central channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

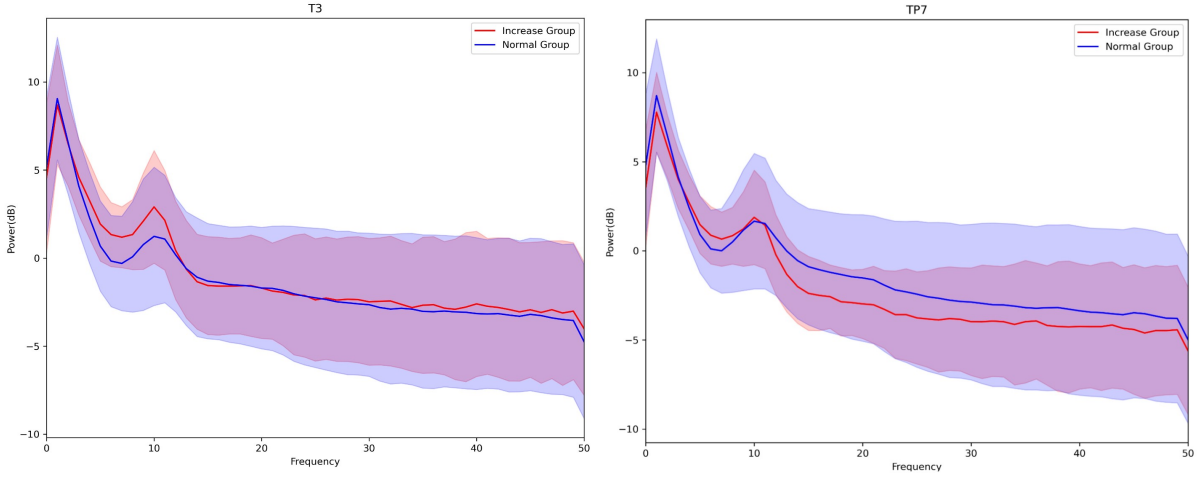


**Figure 2.4.** Left temporal channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

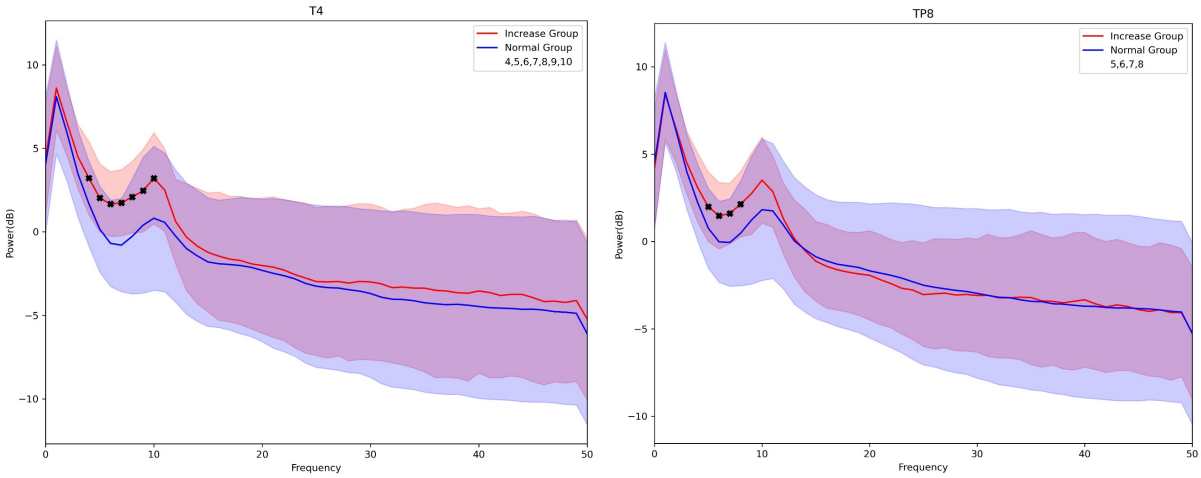


**Figure 2.5.** Right temporal channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

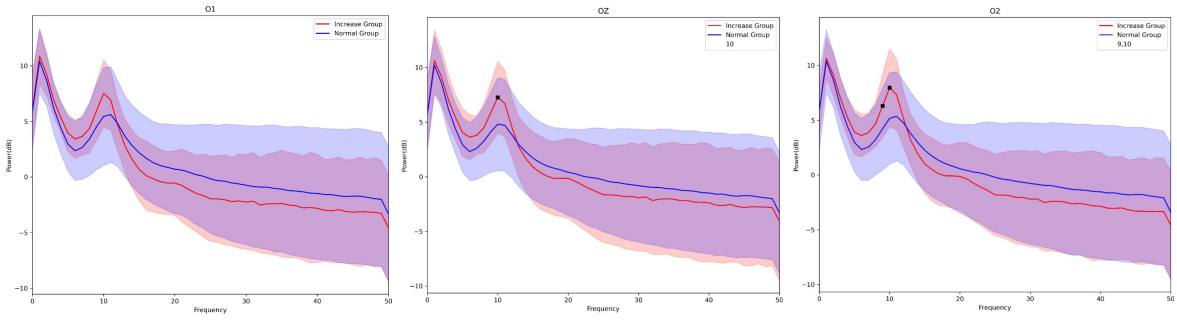


**Figure 2.6.** Occipital channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

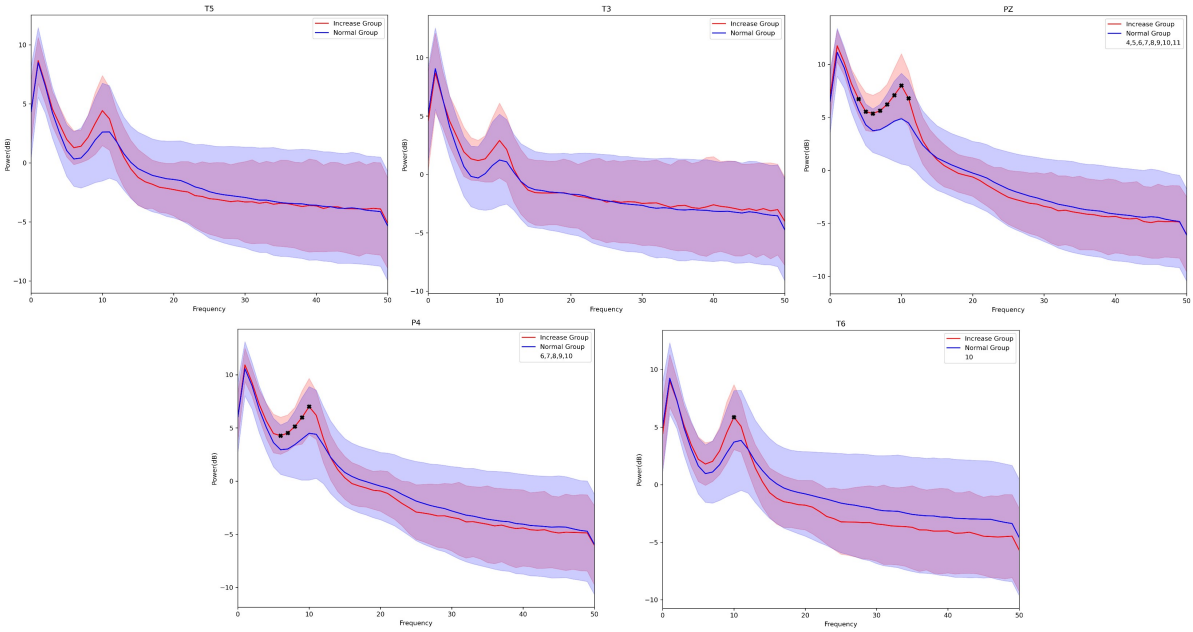


**Figure 2.7.** Parietal channels power spectrum density (PSD) shaded error bar plot with t-test result: upper and lower boundary represented +1 & -1 standard deviation respectively, black markers indicate statistical significance

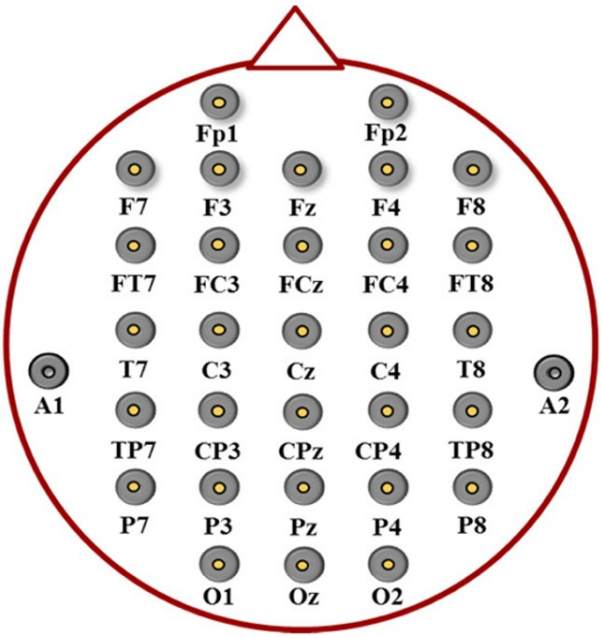


**Figure 2.8.** Electrode placements of 32 channels according to the international 10–20 system. (Figure from ”Fusion of electroencephalographic dynamics and musical contents for estimating emotional responses in music listening” [2])

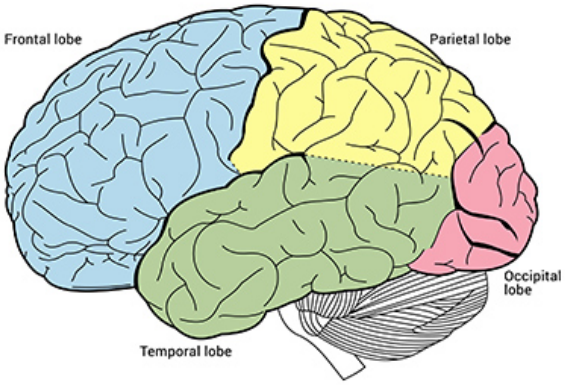


**Figure 2.9.** Brain Anatomy (Image from: https://qbi.uq.edu.au/brain/brain-anatomy/lobes-brain)


**Figure 2.10.** The effect of activity increase in left and right frontal regions (IMotions, 2017)

# Chapter 3

# Stress analysis using frequency power information of independent components

## 3.1 Introduction

In this chapter, I explored the effects of stress in terms of cortical region activation. There are two central questions I wish to answer in this chapter: Do cortical regions activate differently under elevated stress state as compare to normal stress state? If yes, where are those regions that express this difference? To identify those regions, first, independent component analysis (ICA) and dipole fitting were carried out on all subjects to identify each cortical source and its corresponding location. Next, the independent components were clustered into 10 clusters using dipole locations as criteria. After this was finished, Fourier Transform was applied to transform each component into frequency domain. As mentioned in chapter two, such technique is common in EEG data analysis to transform complex time series into organized spectra power information that is much easier to gain insight of. Then, the mean and standard deviation of the spectra power of all components in a given cluster was computed and visualized. Furthermore, statistical testing was carried out in each frequency bin on every cluster to identify the region of difference across the elevated and normal stress groups. The results showed that some sources respond to stress while others do not. We also identified the source that contributed to the channel domain observations made in the last chapter.

## 3.2 Methods

### 3.2.1 Independent Component Analysis

It is often in the researchers' best interest to identify the underlying neural source when a stimulus such as stress has taken place. While the scalp channel recordings are useful in telling the differences in amplitude and frequency across different activities, they cannot localize the brain regions that are affected or responsible for processing the stimuli. The scalp channel recordings can be seen as the superposition of many activities happening simultaneously across different parts of the brain. Thus, we cannot directly use the channel data to identify each underlying sources of activity. Instead, independent component analysis (ICA) [26] has been widely used for the purpose of blind source separation [27]. By using a statistical approach, ICA can decompose event-related potential (ERP) data as recorded using EEG into a maximum number of components equal to the number of sensors. A formal mathematical definition of ICA and psychological interpretation can be found in Bell [28] and Onton's [29] work. As mentioned in the experimental design section, ICA is applied on every dataset such that non-cortical sources are removed. In the preprocessing stage, the cutoff confidence threshold for IC removal is set up to be 20%. This is a rather relaxed cutoff boundary because one does not want to remove any cortical components at this early stage. There are roughly 2200 ICs after this soft cutoff, down from about 2700 ICs at the beginning. Despite the automatic IC labeling tool in EEGLAB being an excellent tool for providing recommendations for what each IC represents, it is not a substitution for visual inspection. To ensure that the ICs only contain cortical components, which is critical later on to obtain an accurate clustering result, visual inspection is carried out in addition to the automatic labeling algorithm. However, 2200 ICs have greatly exceeded the viable range for visual inspection, thus in the second prepossessing stage, I have decided to set up a much stricter IC removal threshold of 80%. This keeps the components only if they are labeled as a brain with a confidence measure greater than 80%. This further reduces the total ICs down to around 800, which is a much more viable range to work with for visual inspection. It is also

important to note that there is a time and accuracy trade-off for using a stricter cutoff confidence measure compare to a relaxed one. On the one hand, as the confidence measure cutoff increases, it is more likely to retain only the brain components. This helps to reduce the time needed for visual inspection in later stages. On the other hand, a strict cutoff may remove components that are coming from the brain. This could remove some useful information, decreasing the clustering result's truthfulness.

### 3.2.2 Dipole Location & Orientation

After ICA is finished, the next stage is dipole fitting. One common problem researchers encounter when comparing individual components across subjects is that there is no easy method for comparison. All ICs differ with each other in some ways such as power spectra, scalp maps and so on. Depending on the application, one could choose any reasonable metric mentioned above to compute the distance or similarity between pairs of ICs and form clusters. Here, since the goal is to identify the cortical regions that are affected by stress. This naturally makes dipole fitting a ideal method to use. Dipole fitting, in simple terms, is a method to estimate the location and direction of neuronal propagation for any given source activity (i.e. individual components). This estimation is rather challenging since one is trying to solve an "inverse" problem, that is, given the observed electrical potential data (i.e. EEG), reconstruct the physiological source [30] (Figure 3.1). "Fortunately, the problem of finding the location of a single equivalent dipole generating a given dipolar scalp map is well-posed, given a sufficiently accurate electrical "forward problem" head model that specifies the resistance between each scalp electrode location and each possible (brain) source location" [31]. Here, the boundary element model (BEM) supported by the EEGLAB Dipfit function is used. Moreover, the residual variance threshold is set up to be 15% as suggested by Delorme's work [32]. Residual variance here measures how well the dipole fitting result explains the scalp map, the higher the percentage, the worse the fitting. Examples of good and bad dipole fitting results are shown in Figure 3.2 and Figure 3.3. It is also worth to note that for simplicity, single-dipole model is used. One could attempt

using higher dipole models for fitting, but this approach would introduce ambiguity in the clustering stage as multiple dipole locations exist for one scalp map. For each component, the corresponding dipole fitting result consists of two parts representing the dipole location and orientation. The dipole location, as represented by a 3-dimensional array, represents the cortical location where the source is generated. While the dipole orientation, also a 3-D array, represents the direction of neural information flow.

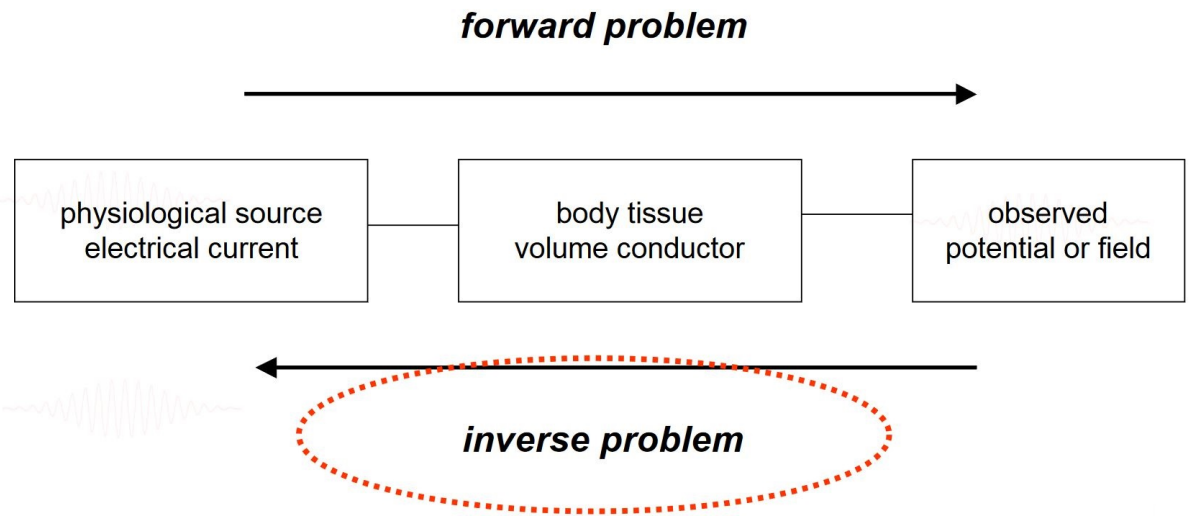


**Figure 3.1.** Dipole Fitting: an inverse problem

### 3.2.3 K-means Clustering

After dipole fitting is completed, the data is ready to be clustered. It is the researcher's choice to choose which parameters to use as clustering criteria. Power spectra, scalp map, dipole location and dipole orientation are all valid choices to use for clustering. Again, since the goal is to identify the cortical regions that are affected by stress, it is natural to use dipole locations as the criteria for clustering.

There are various types of clustering algorithms that are available to use. For example, there are distance-based methods such as K-means clustering [33], density-based methods such as DBSCAN [34] and Hierarchical-based methods. Each method can be used to handle different data, and for the purpose of dipole location clustering, distance-based method (i.e. K-means

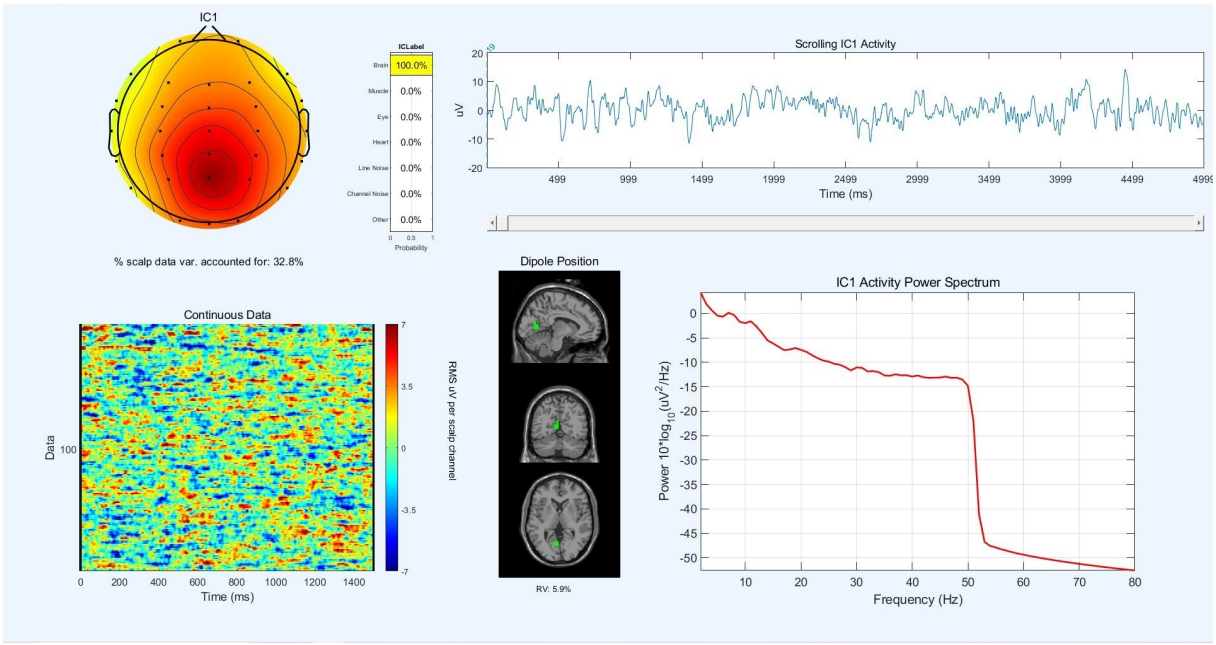


**Figure 3.2.** A ”good” dipole fitting result: residual variance <15%

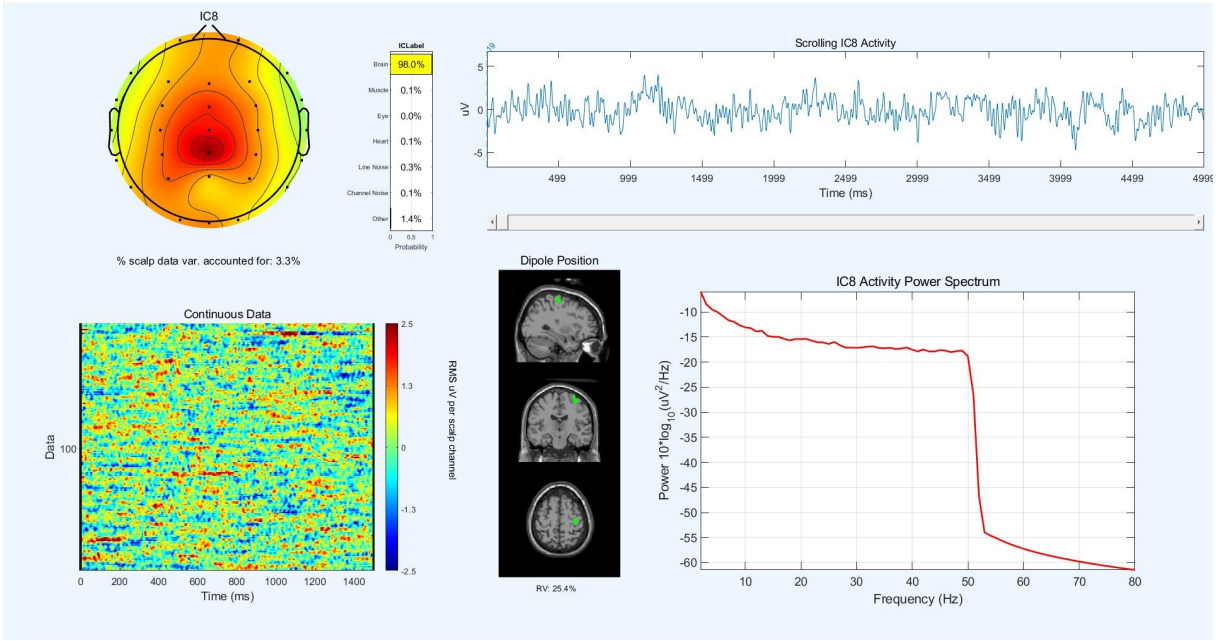


**Figure 3.3.** A ”bad” dipole fitting result: residual variance >15%

clustering) fits the best. The detailed algorithm for K-means clustering can be found in this paper [35]. For those who are familiar with this method, one of the drawbacks about K-means clustering is that it requires K, the number of clusters, to be specified before the algorithm is

applied. One of the popular ways to determine the "optimal" number of clusters is using the "elbow" method [36]. Distortion is calculated using different K values, and the value of the point where distortion starts to decrease linearly is the "optimal" number of clusters one should use (Figure 3.4). Following this approach, the "optimal" number of clusters is found to be five, and

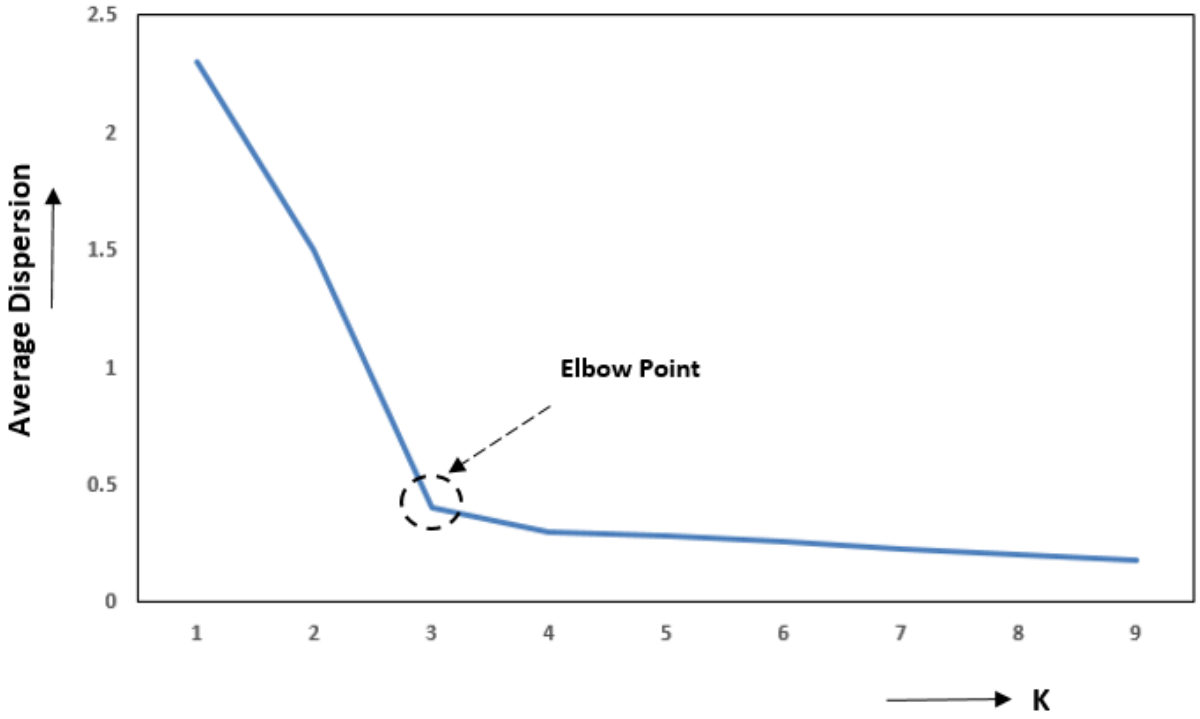


**Figure 3.4.** An example for using elbow methods in K-means clustering (Dangeti, 2022)

the average of the topoplots roughly represents the frontal, parietal, occipital, left temporal and right temporal lobe as shown in Figure 3.5. However, despite five clusters being the theoretical optimal number of clusters to use, one can clearly see that for practical reasons, there are not enough resolution to examine the exact cortical location where stress has affected. Thus, a larger K needs to be tested. After several different K implementations, I found that K equals to 10 gives a reasonable clustering results with adequate spatial resolution (Figure 3.6). In the last stage, Fourier Transform is applied on every component to transform them from time domain to frequency domain. T-test is then carried out for each cluster to identify all the frequency ranges that are statistically different across the elevated stress and normal stress groups. The results are

shown in Figure 3.7.

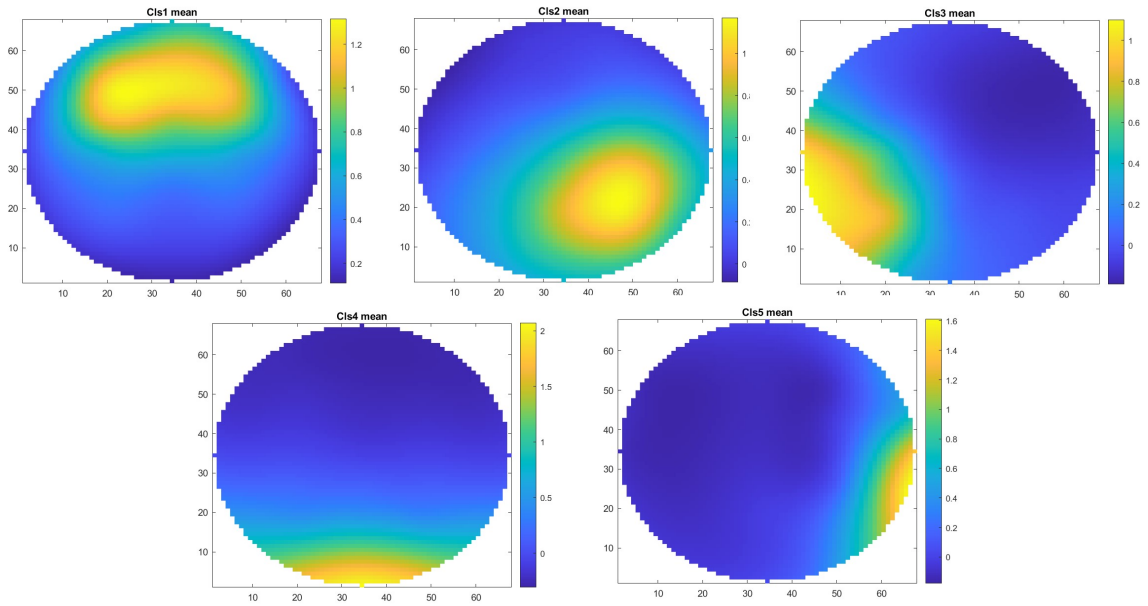


**Figure 3.5.** 5-means Clustering Result

## 3.3 Discussion

Looking at Fig. 3.7, one can observe that most sources do not respond to stress except cluster 9. Then connecting back to the results obtained in chapter one, that is, statistical significance is observed in frontal, central and right temporal channels in delta and alpha bands, we now identified the source that contributed to that observation. Cluster 9, as Fig 3.6 shows, is a source that is very deep in the brain which projects toward the central and frontal areas. Thus, the channels along the direction of source propagation will pick up the most signal, which explains why most significance is observed in frontal and central channels.

## 3.4 Conclusion

In this chapter, I explored the difference in cortical region activation between the elevated and normal stress groups. To identify those regions, first, ICA and dipole fitting are carried out on all subjects to localize each cortical source. Next, the ICs are clustered into 10 clusters using

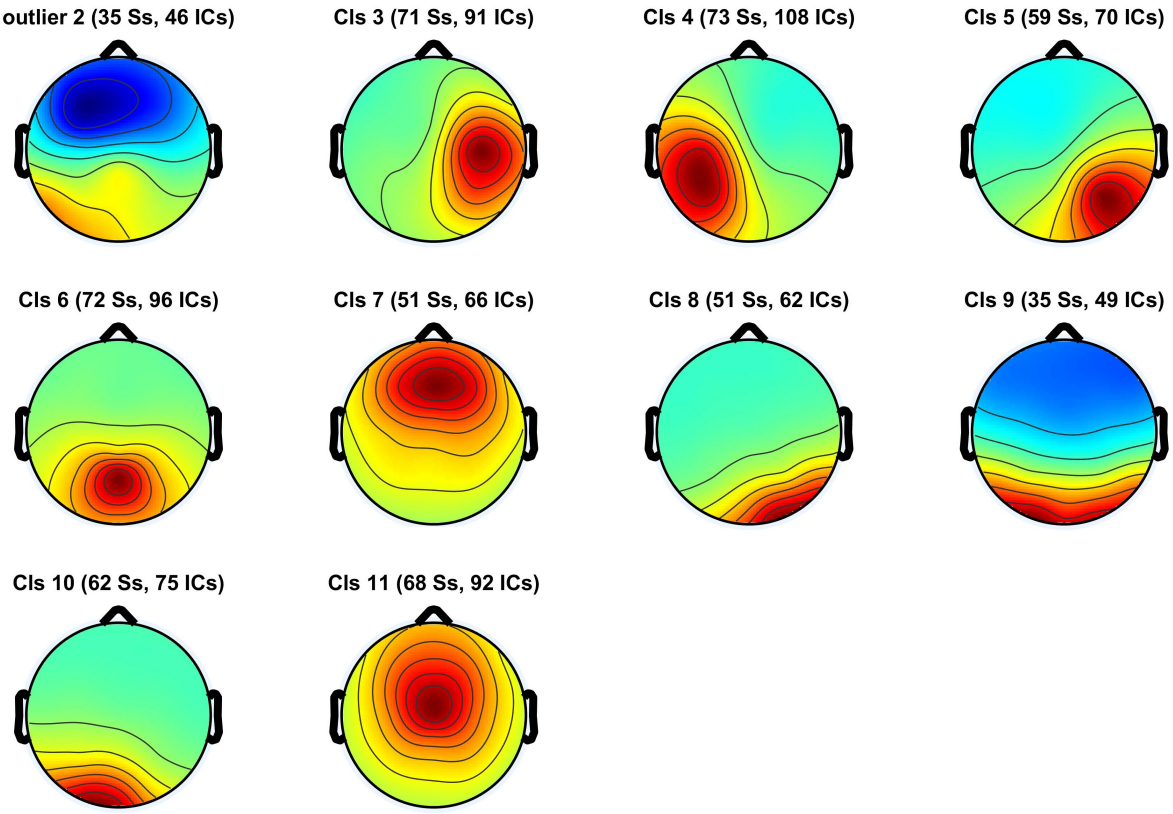


**Figure 3.6.** 10-means Clustering Result: (number of subjects, number of ICs in the cluster)

their corresponding dipole locations as criteria. Then, Fourier Transform is applied to transform each component into frequency domain. In the last step, the mean and standard deviation of the spectra power of all components in a given cluster is computed and visualized. Moreover, statistical test is carried out in each frequency bin on every cluster to identify the region of difference across the elevated and normal stress groups. The result shows that some sources respond to stress while others do not. In this case, source nine is the only source that responds to stress stimuli. Findings in this chapter helps to explain the observation made in chapter 2 from the channel domain perspective.

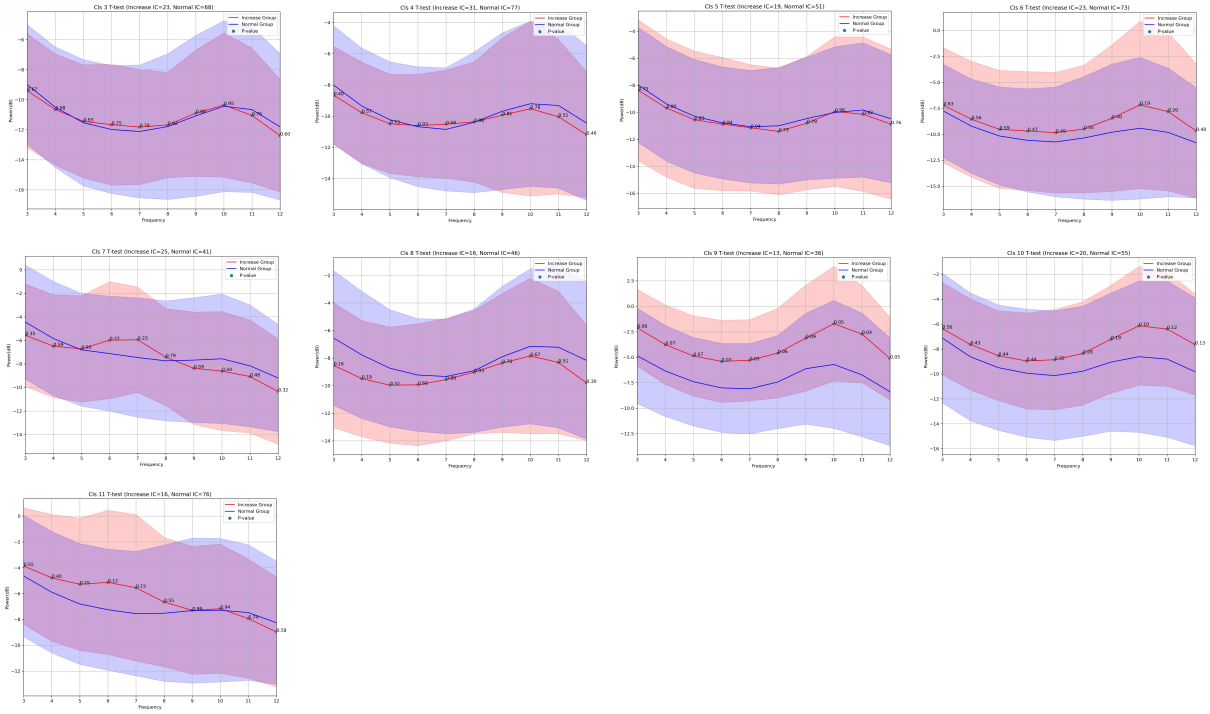

**Figure 3.7.** Shaded Error Bar plot with T-test results: upper and lower boundary represented +1 & -1 standard deviation respectively

# Chapter 4

# Effective connectivity and direction of network propagation of cortical regions under stress

## 4.1 Introduction

Understanding the functional and structural connectivity of the human brain network is an important goal of neuroscience and medicine. The U.S. National Institutes of Health launched the Human Connectome Project in 2009 with its central goal of mapping the human brain and connecting its structure to function and behavior. Since the project first launched, hundreds of papers have been published to advance the understanding of brain connectivity under various medical conditions such as Alzheimer [37], depression [38] and epilepsy [39]. However, less work is being done to understand how the brain networks react to emotional stimuli such as stress. The human brain Connectivity analysis usually falls into three categories: structural, functional, and effective [40]. Structural connectivity, as its name suggests, is a study to identify the anatomical links between different parts of the brain. This study focuses on understanding what brain structures are capable of influencing each other via direct or indirect axonal connections. Functional connectivity refers to the symmetrical correlations between brain regions under a task or stimuli. Here, the focus is to understand the functionally related regions by measuring the correlation of activity through imaging techniques such as functional magnetic resonance imaging

(fMRI). In contrast to the symmetric nature of functional connectivity, effective connectivity denotes the asymmetric or causal dependencies between brain regions. The task here is to identify the brain regions that are causally influencing other regions during information flow. This usually involves searching for the direction of information flow between structures [41].

In this chapter, I aim to explore the effective connectivity of brain structures under stress by applying Granger-causal methods (a common method for effective connectivity measure) to the recorded EEG data.

## 4.2 Methods

### 4.2.1 Granger Causality & Source Information Flow Toolbox (SIFT)

Granger Causality, a method developed by Clive Granger in 1960s, has been widely used in economics to determine if one time series could be used to predict another. The theory states that if a signal x1 Granger-causes x2, then past values of x1 should contain information that helps predict x2 above and beyond the information contained in past values of x2 alone [42]. One can intuitively see that such information could be rather useful in economic applications such as stock market forecasting. Moreover, by applying Fourier Transform, it is possible to examine Granger Causality in the frequency domain [43]. This is exactly what this chapter will use to compute effective connectivity, by examining Granger Causality in frequency domain. Several toolboxes that offer various forms of Granger Causality are available to use as shown in Table 4.1. All of the toolboxes listed in the table offer free-to-use licenses and are integrated with Matlab. Here, the author chose SIFT due to its integration with EEGLab.

The main contributor of SIFT, Tim Mullen, has published a user manual for running SIFT [41]. Furthermore, one of EEGLab's main contributors, Makoto Miyakoshi, has made a website (https://sccn.ucsd.edu/wiki/How_to_run_SIFT_simulation) that is up-to-date (last updated 8/20/2022) documenting the steps to run SIFT. Both are excellent resources for researchers who would like to use SIFT in their research, and here, the author closely followed each step listed in

the user manual and website. Note that the detailed mathematical explanation of each step will be omitted in this chapter, but it is documented in the manual for anyone who is interested. In an overview, there are three steps to compute effective connectivity using SIFT. Before running the first step in SIFT, the user must prepare the data for model fitting. The data must be properly "cleaned" and downsampled to 100-120Hz using the Matlab command pop_resample(EEG, 100, 0.8, 0.4) [44]. Non-cortical ICs should also be removed to reduce the size of the final connectivity tensor. A reasonable brain label probability cutoff threshold is 70% or above as suggested by Miyakoshi, here, the author used a stricter cutoff of 80%. Dipole fitting should also be done using the DIPFIT plugin in EEGLab. After all of the pre-processing steps have been properly carried out, SIFT can be started. First, SIFT requires the user to pre-process the data using its function in the drop-down menu. However, since in the pre-step, the data has already been properly "cleaned", there is no need to carry out any additional pre-processing. As shown in Fig. 4.1, the parameters are left as default. Next, the data undergoes model fitting and validation as shown in Fig. 4.2 and Fig. 4.7. In the model fitting stage, there are several parameters the researcher can manipulate depending on the goal. First, several algorithms are available to fit the model, the most common ones are ARFIT and Vieira-Morf. According to Mullen's SIFT user manual, ARFIT is usually faster in terms of running time while Vieira-Morf has a slightly better coefficient estimate. This is a trade-off decision the researcher can make, here, the author chose to use ARFIT. The more important parameter a researcher should consider is the window length. In this classroom resting-state recording, there is no event or stimuli designed to alter the subject's cognitive state in the middle of the experiment. Thus, time resolution can be omitted in the connectivity tensor. In this case, a window length of 300, the entire length of the recording, is used. For other applications, if the experiment has been designed to induce cognitive state transition in the middle of the recording, a smaller window length may be ideal to capture this potential change in connectivity between regions. An important and useful note the author wishes to make is that the running time for model fitting will significantly increase when window length and stepsize is small. This can easily take days of continuous fitting if the resolution is

**Table 4.1.** A list of toolboxes that offer effective connectivity analysis for neural data

| Toolbox Name | Primary Author | Release | Website |
|---|---|---|---|
| Multivariate Granger Causality (MVGC) Toolbox | Lionel Barnett | 1.3 | https://users.sussex.ac.uk/∼lionelb/MVGC/html/mvgchelp.html |
| Time Series Analysis (TSA) Toolbox | Alois Schloegl | 4.6 | https://pub.ist.ac.at/∼schloegl/matlab/tsa/ |
| eConnectome | Bin He | 2.0 | https://www.nitrc.org/projects/econnectome |
| Brain-System for Multivariate AutoRegressive Timeseries (BSMART) | Jie Cui | - | https://brain-smart.org/ |
| Source Information Flow Toolbox (SIFT) | Tim Mullen | 1.52 | https://github.com/sccn/sift |

set to be too high. Next, after clicking fitting, a warning window may appear similar to the one shown in Fig. 4.3. A detailed explanation of this warning is addressed in Miyakoshi's github (https://github.com/sccn/groupSIFT) under "SIFT tips". The researcher can read that section to decide whether the warning can be ignored. After clicking ok, two graphs should appear after model fitting has been completed as shown in Fig. 4.4 and Fig. 4.5. In this case, the author chose to use the result returned from the elbow method with optimal model order 9. After the optimal model order is determined, simply input this number into the model fit box to start model fitting as shown in Fig. 4.6. The final step in the model fitting stage is to validate the fitted model. This step makes sure that the model is a good approximation such that connectivity can be estimated within an acceptable error range (Fig. 4.7). This concludes the second step, model fitting and validation. The last step in SIFT is to perform connectivity analysis using Granger Causal method (Fig.4.8). Plot setting and connectivity result are shown in Fig. 4.9 and Fig. 4.10 respectively.

### 4.2.2 GroupSIFT

While it is certainly interesting to analyze the effective connectivity tensor of the individual subject under stress, in order to generalize this finding, group-level analysis must be sought out. However, as the author has mentioned in chapter three, it is usually challenging to compare ICs across subjects since no two subjects share the same IC. The same challenge persists here: despite having the connectivity tensor for all the subjects, there is no easy way to compare each tensor since it only captures the unique dipoles one subject has. However, one potential solution the author adopted here is to transform all the unique dipoles into their corresponding spatial regions (Fig. 4.11). By doing this, it provides a common ground for comparison since

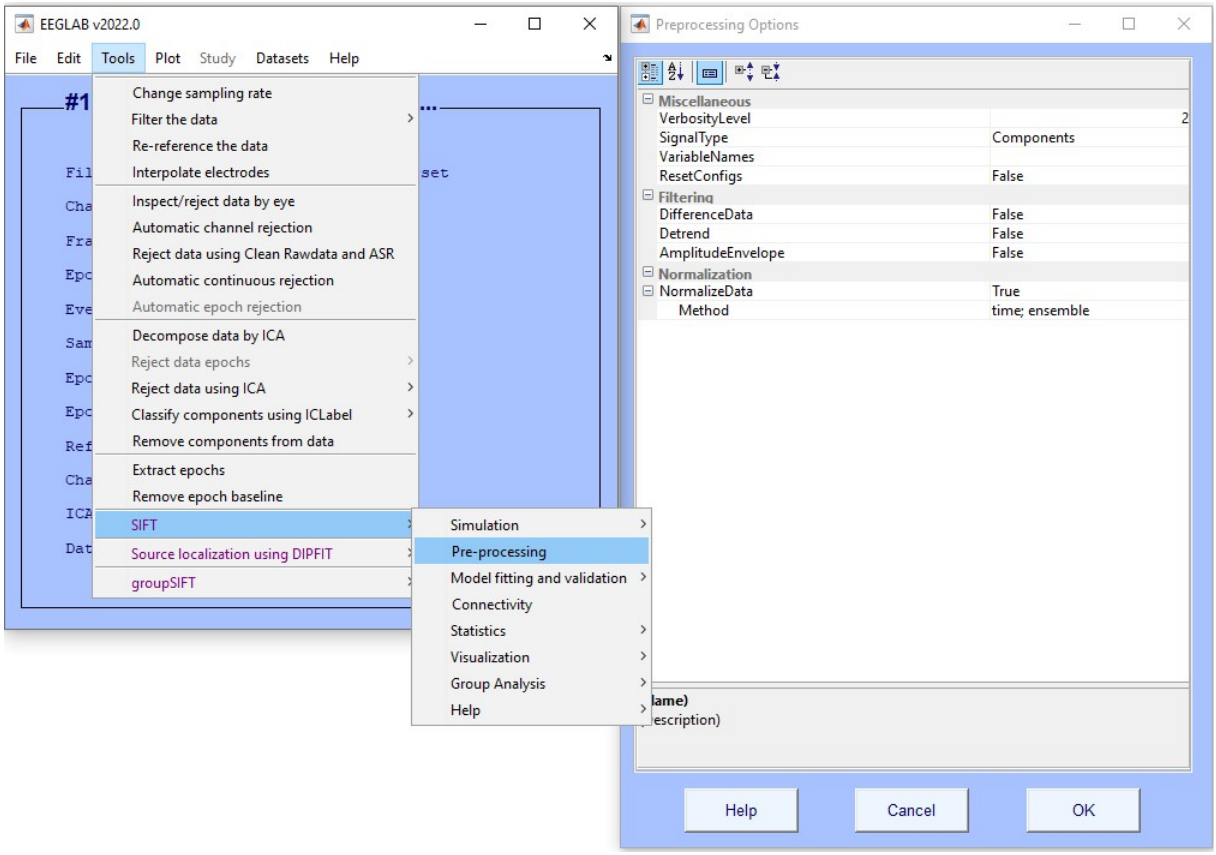


**Figure 4.1.** SIFT pre-processing parameter settings

all subjects share the same spatial region. This transformed the original problem into a much better-defined problem, and now the goal is to identify if there is such a region that is unique in terms of information outflow and inflow to the elevated stress group. GroupSIFT, designed for group-level analysis for SIFT results, conveniently provides the function to transform from dipole to dipole connectivity to cortical region connectivity. Here, the author will provide rather detailed steps and explanations for carrying out GroupSIFT connectivity analysis so that future researchers can use it as a reference for their own study. A more concise guideline for running GroupSIFT can be found here (https://github.com/sccn/groupSIFT). As a side note, GroupSIFT is currently only available to download from the link above, and one must manually place the unzipped folder inside the plugins folder under the EEGLab path. There are five steps in carrying out GroupSIFT analysis, and they are batch process, model validation, dipole connectivity to

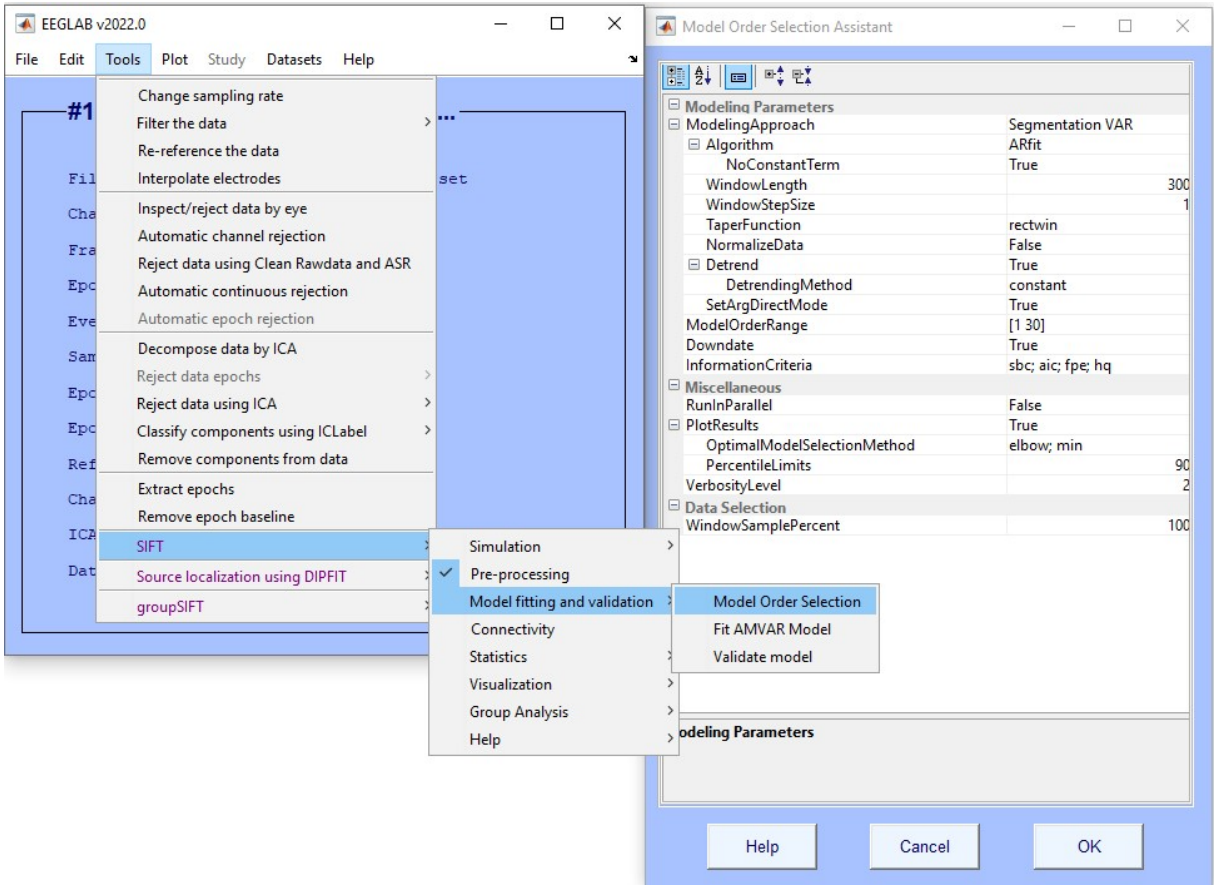


**Figure 4.2.** SIFT model order estimation parameter settings

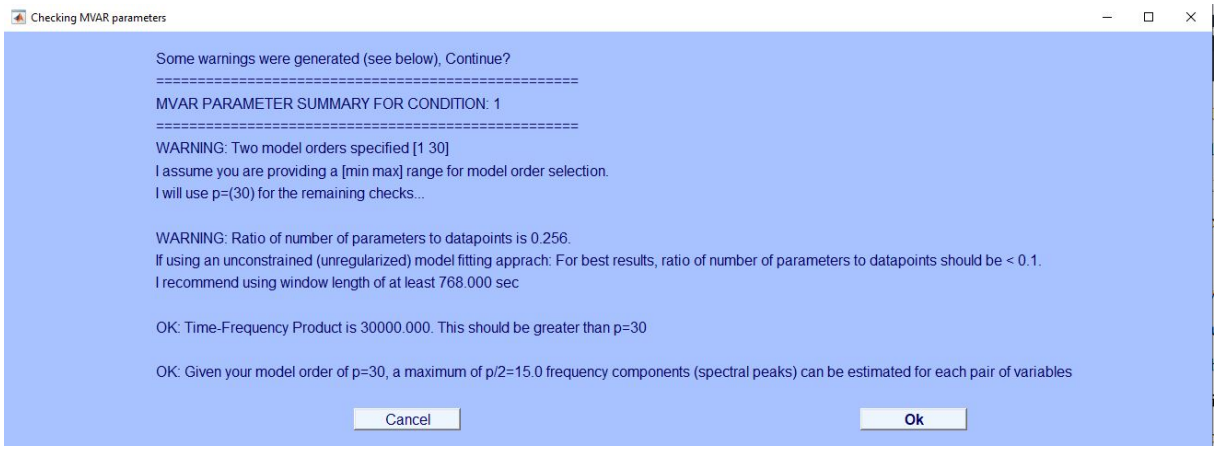


**Figure 4.3.** SIFT model order estimation warning

spatial connectivity transformation, statistical test, show result, and export for the movie. In the first step batch process, as the name suggests, it provides a convenient feature to batch process

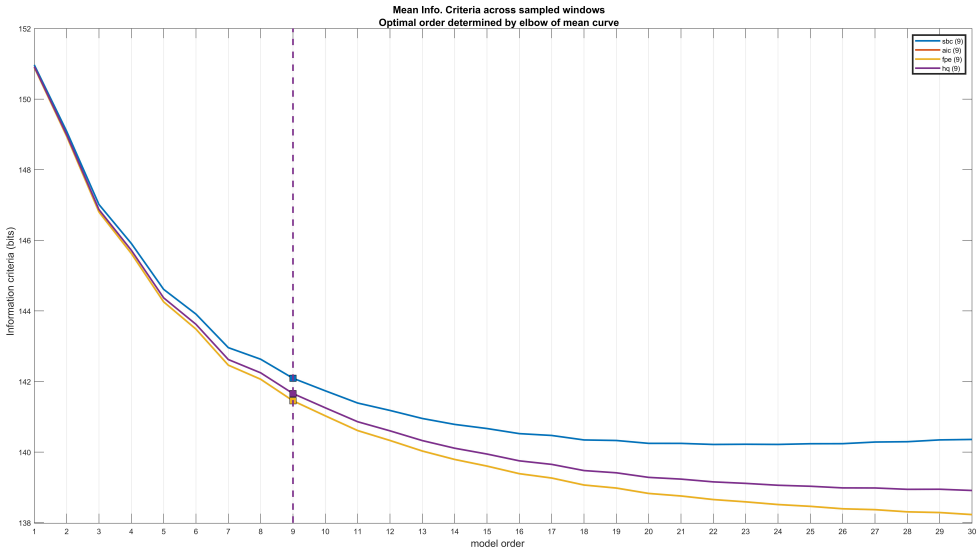


**Figure 4.4.** SIFT model fitting result elbow method

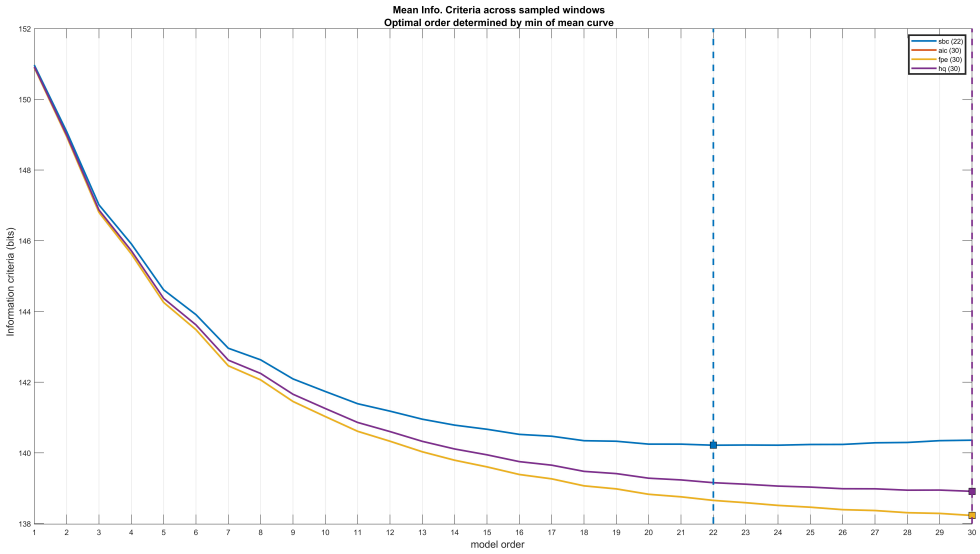


**Figure 4.5.** SIFT model fitting result min method

multiple files for running SIFT. Again, as a reminder, all files must go through SIFT first before doing the "dipole to spatial region" transformation. Fig. 4.12 shows the parameters to set for running batch process. One thing to note is that depending on the goal of the study, one may choose not to check mark the single-window analysis option, and an explanation for this has

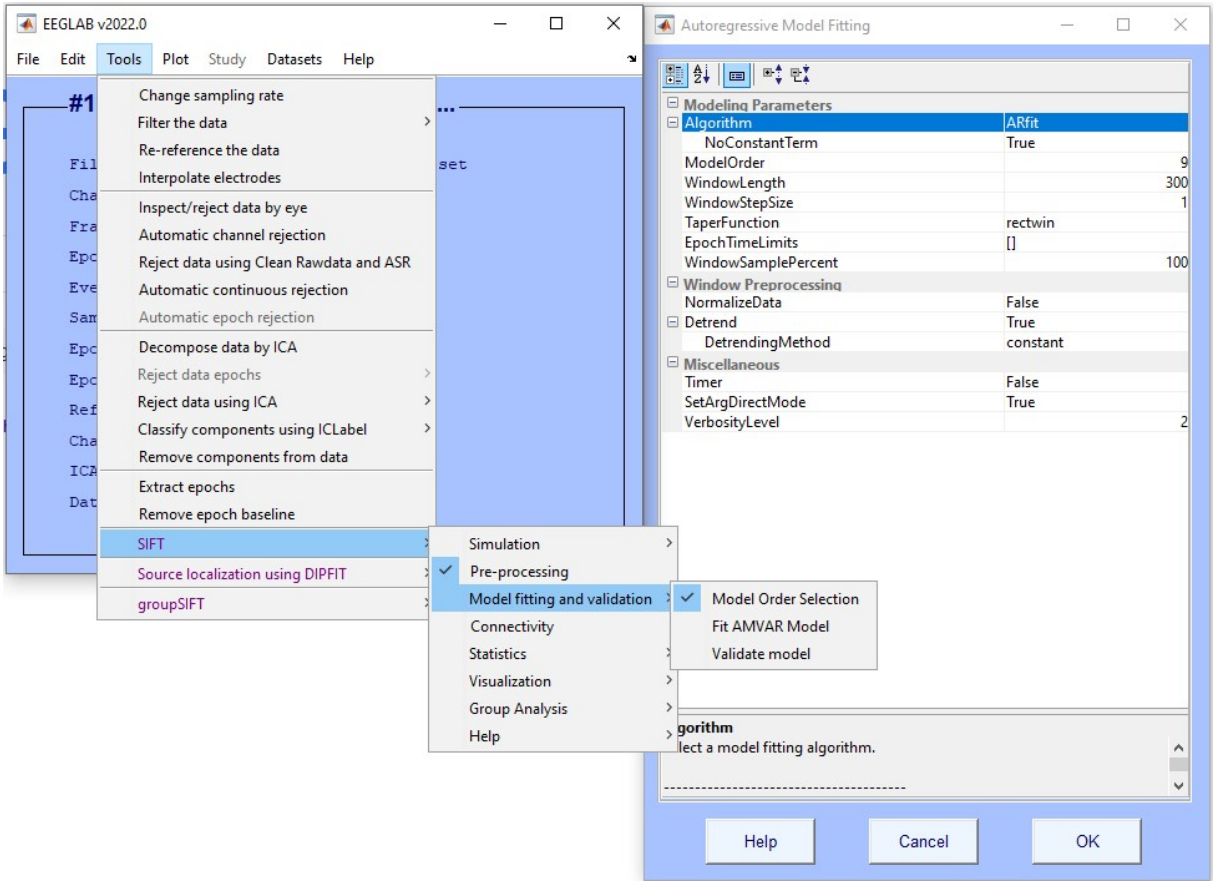


**Figure 4.6.** SIFT Model Fit parameter settings

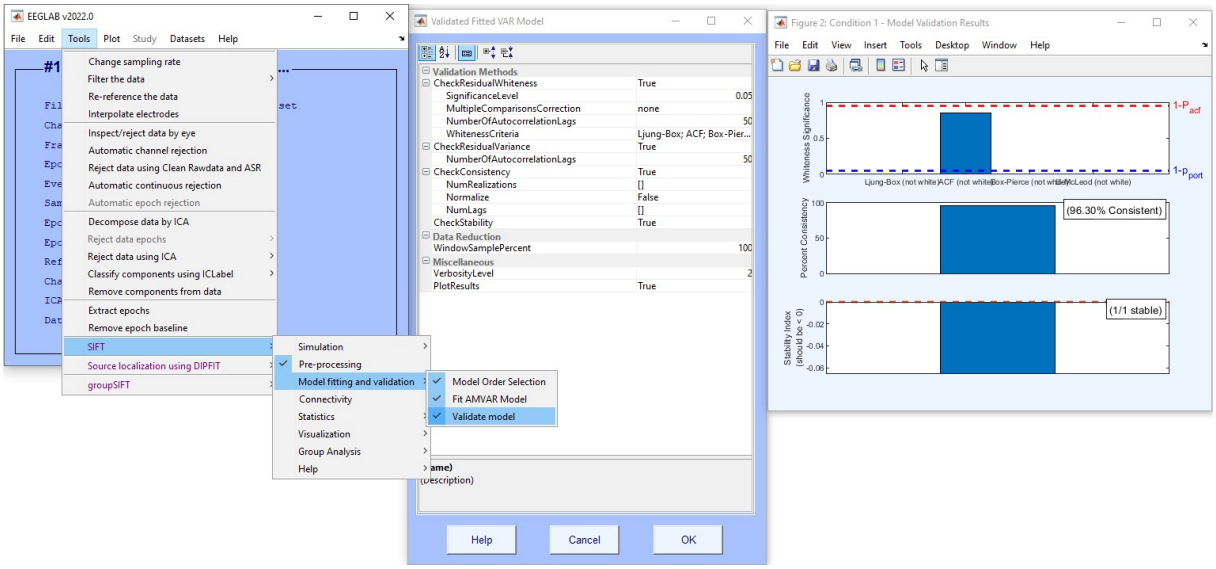


**Figure 4.7.** SIFT model validation parameter settings and result

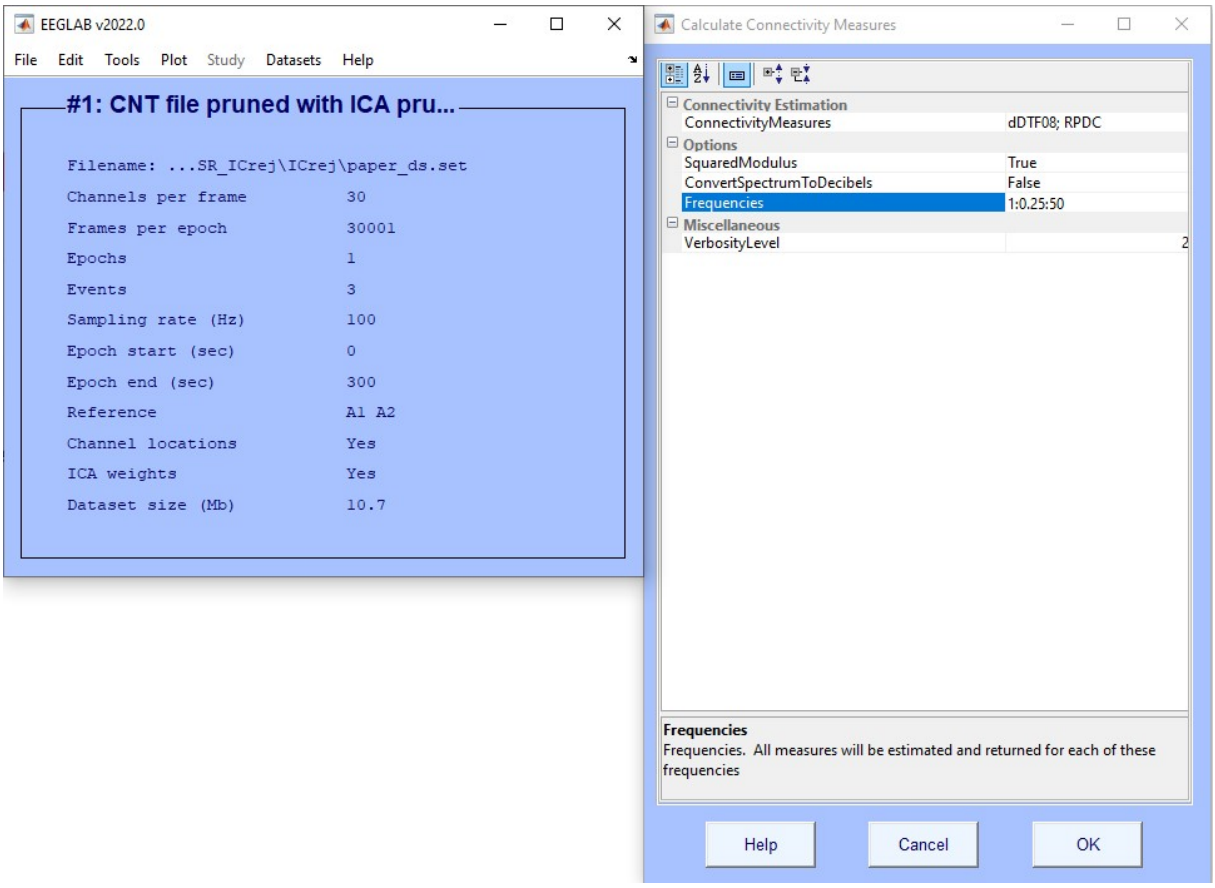


**Figure 4.8.** SIFT connectivity parameter setting

already been given in the previous section. Moreover, it is highly recommended that one follows the instruction between %, that is to comment out certain lines to prevent pop-up windows. Once the parameters are set, simply click on the button and select the files for running SIFT. Next, select the function "validate AR model" from the GroupSIFT dropdown menu. Select all the files that have finished running through SIFT. A result similar to Fig. 4.13 should show up. Now the files are ready for the third step, "transformation". As shown in Fig. 4.18, there are several parameters one can set. There are two connectivity metrics one can choose from, and below that, there are two parameters one can set to define the smoothing width and minimum percentage of subjects with non-zero dipole density respectively. For smoothing width, it is recommended to use the default value 20 mm for EEG data, and 8 mm for fMRI data. Then for the minimum percentage option, it is highly dependent on the data one is using. The idea here is that one wants

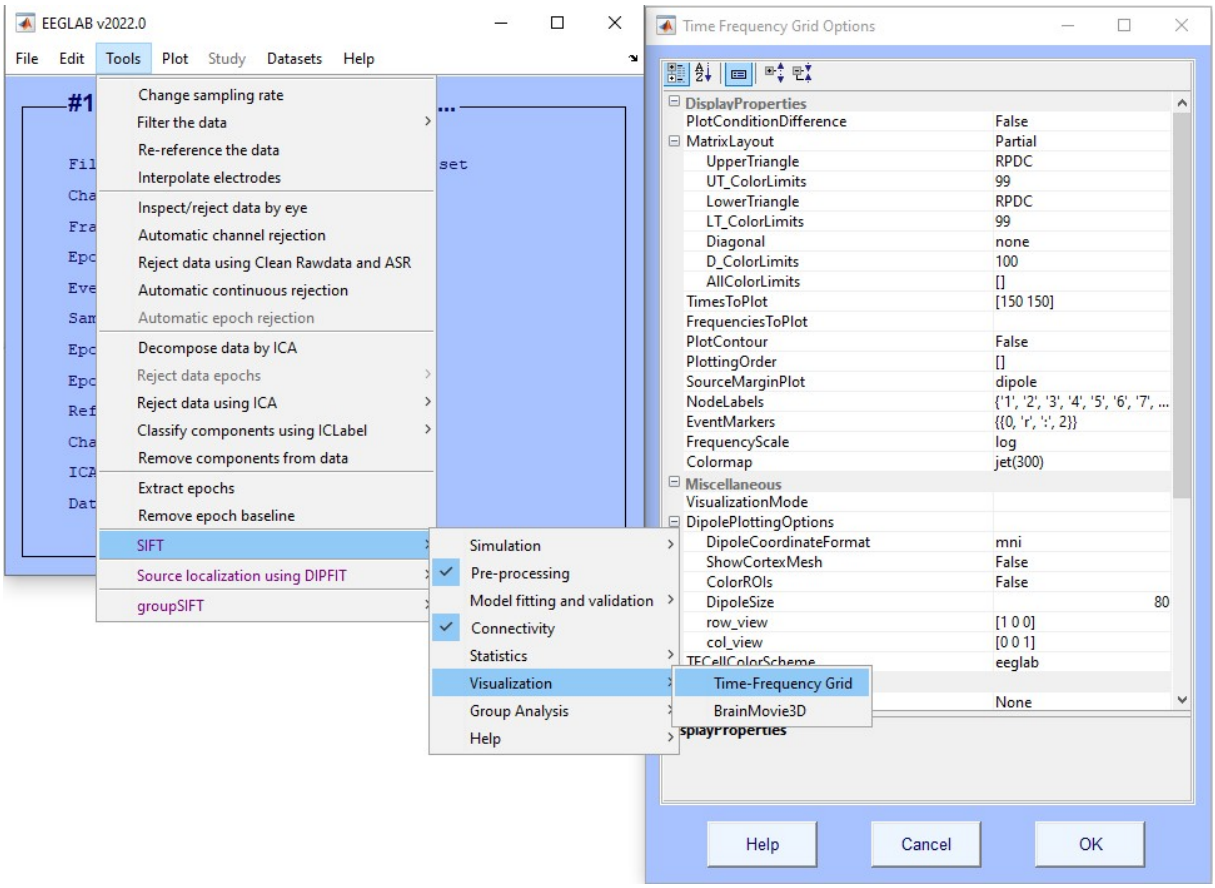


**Figure 4.9.** SIFT connectivity plot parameter setting

to focus only on the regions that have the majority of the subjects that contribute non-zero dipole density. Thus, by setting up this percentage of subjects threshold, one can filter out the regions that do not include many dipoles. For datasets with highly overlapped dipole location across subjects, the threshold could be set to a higher value without losing much information. While for datasets with low overlapped dipole locations across subjects, the threshold should be set to a lower value to retain information. One can always experiment with this parameter by clicking on the "compute upper bound for estimation" button and selecting the dataset to perform this transformation. In this case, 70% is a good threshold since 90.9% dipole density is accounted for, which means little information is lost. Fig. 4.15 shows the result of the transformation with each location and its corresponding dipole density. If one is satisfied with the result, input a filename in the box below and click the "Select All .set files and START" button. Note that this

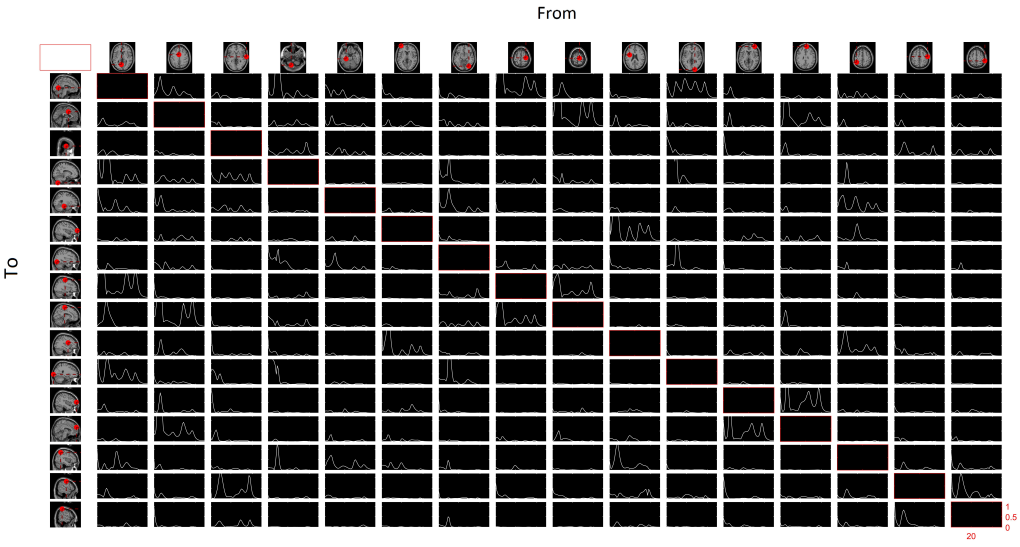


**Figure 4.10.** SIFT connectivity result (x-axis: Frequency; y-axis: information flow)

transformation must be performed separately for each condition, in this case, the elevated stress group and normal stress group are processed separately. After the transformation is finished, the next step is statistical testing, and the parameter settings are shown in Fig. 4.16. Here, the author used subtraction to compute significance, that is subtracting the normal stress condition from the elevated stress condition. Another way to view this is to treat the normal condition as the baseline period where we are trying to determine how the elevated stress differs from it. Note that by default, p value is set to 0.01, one may want to use the standard cutoff 0.05 as the author did here. After the parameters are set, simply click on the "start process" button, and two .mat files should be returned once the process finishes. The last step is to view the results, and settings are shown in Fig. 4.17.

## 4.3 Results and Discussion

The final connectivity matrix obtained from group level study is shown in Fig. 4.19. There are six pairs of regions that are unique in terms of information flow among the elevated stress

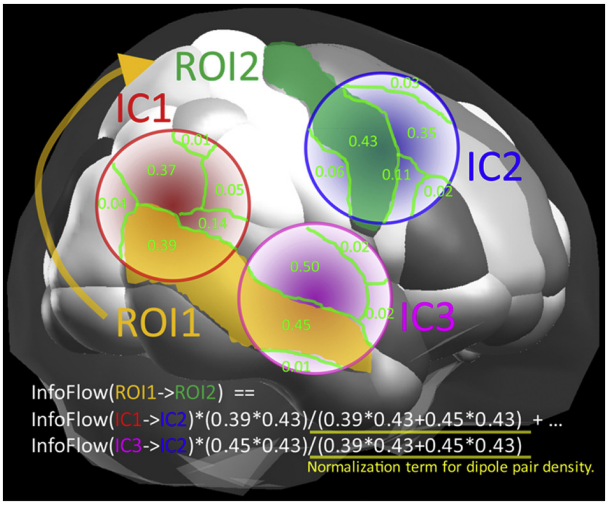


**Figure 4.11.** Computation of source-level connectivity (Figure from "Neural activation and connectivity during cued eye blinks in Chronic Tic Disorders" [3]).

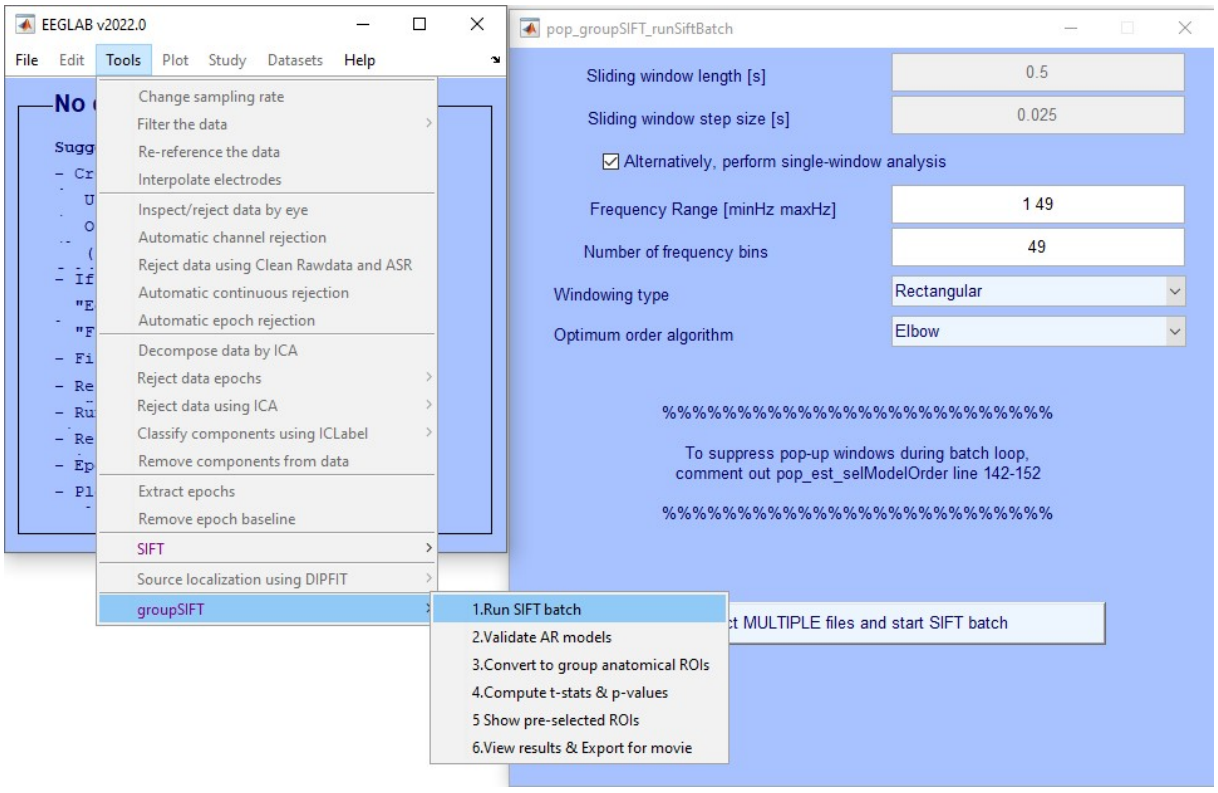


**Figure 4.12.** GroupSIFT batch process

group. Among the six pairs of unique regions, one pair (indicated by a yellow dot) represents decreased information flow in delta bands from the left lingual region to the left precuneus region

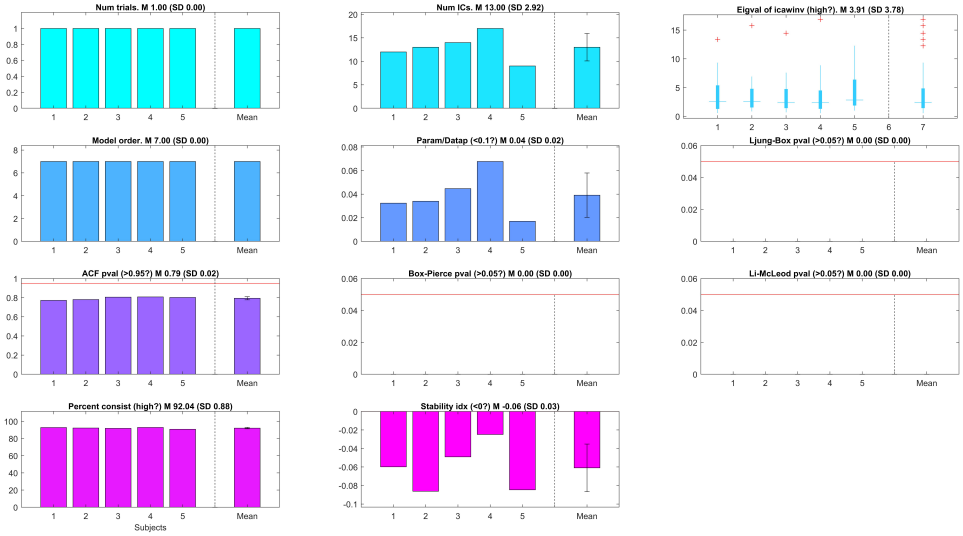


**Figure 4.13.** GroupSIFT model validation result

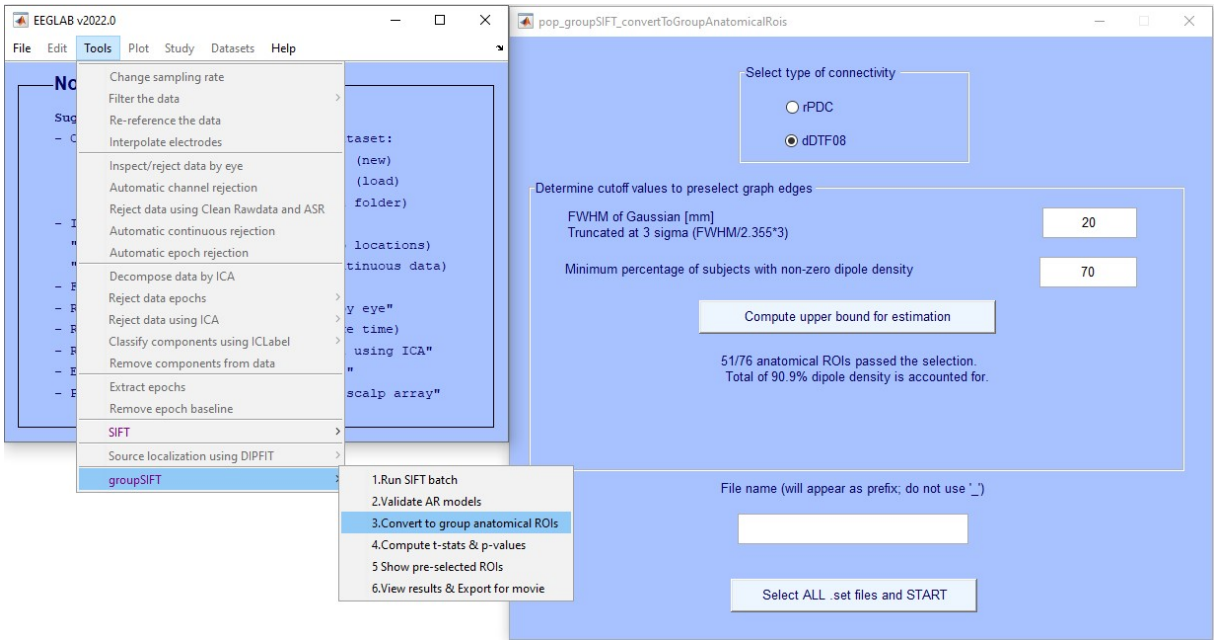


**Figure 4.14.** Convert to Group Anatomical ROI parameter setting

as shown in Fig. 4.20. While the rest five pairs of regions all indicate increased information flow in certain frequency bands as compared to the normal stress condition. To further elaborate, they are the right precentral to left calcarine region with significance found in alpha and low beta bands (Fig. 4.21); right postcentral to left lingual region with significance found in the

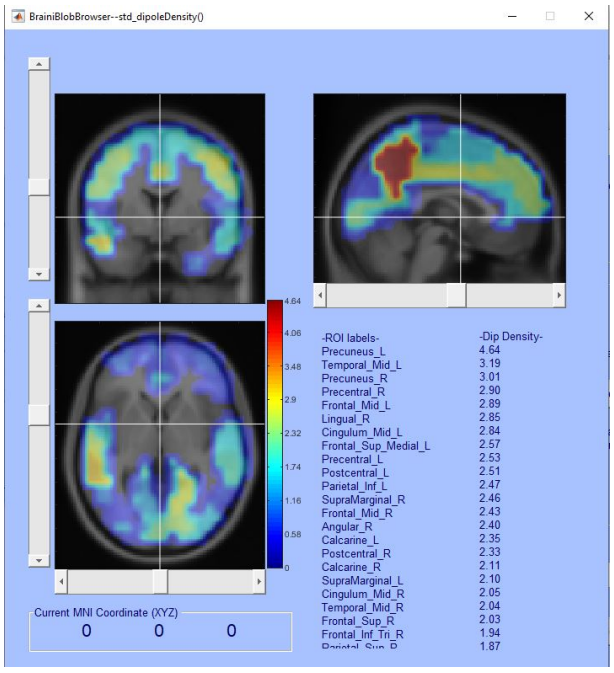


**Figure 4.15.** Convert to Group Anatomical ROI Result

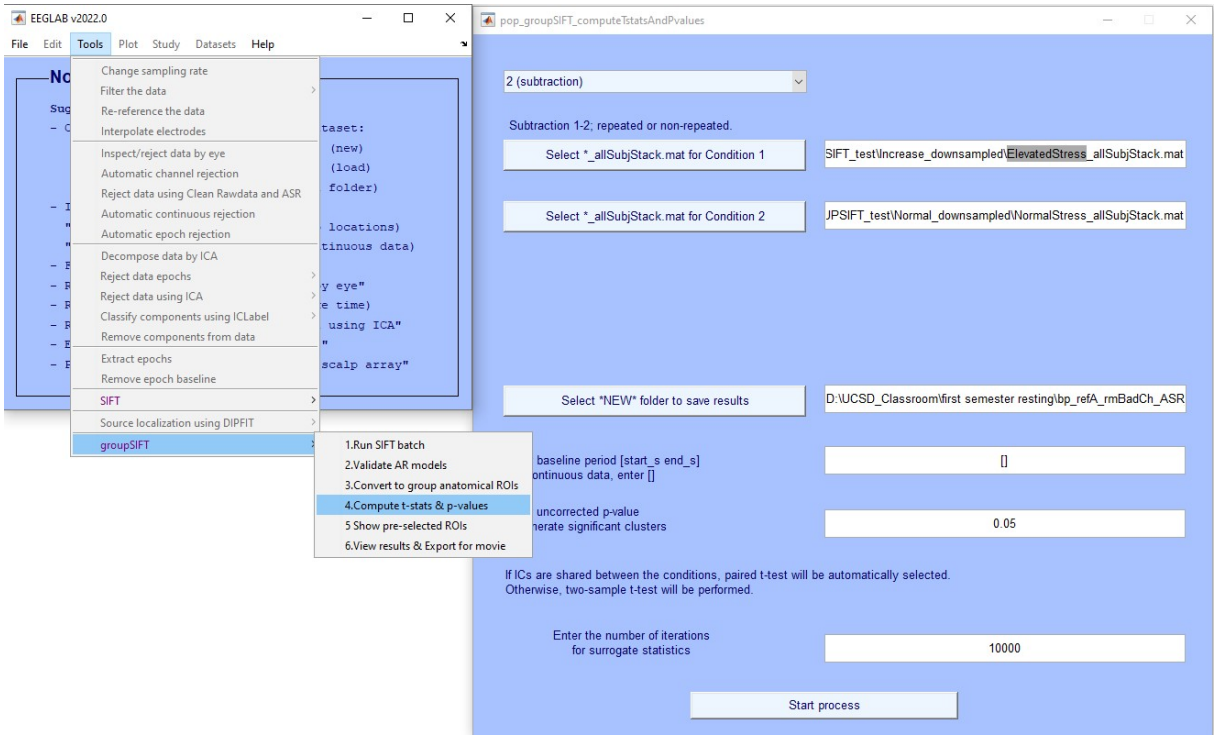


**Figure 4.16.** Compute T-stats and P-value parameter setting

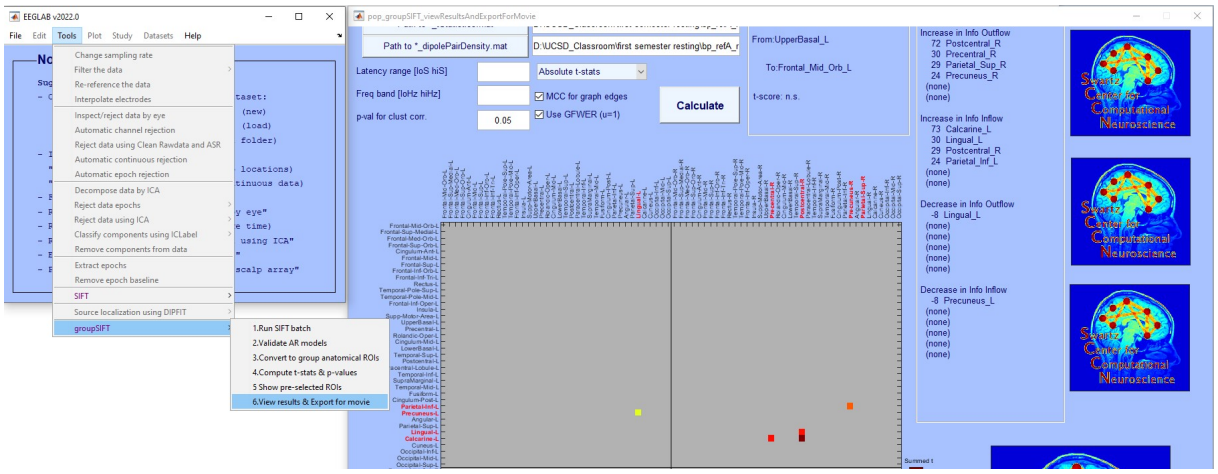


**Figure 4.17.** View Result Parameter Setting

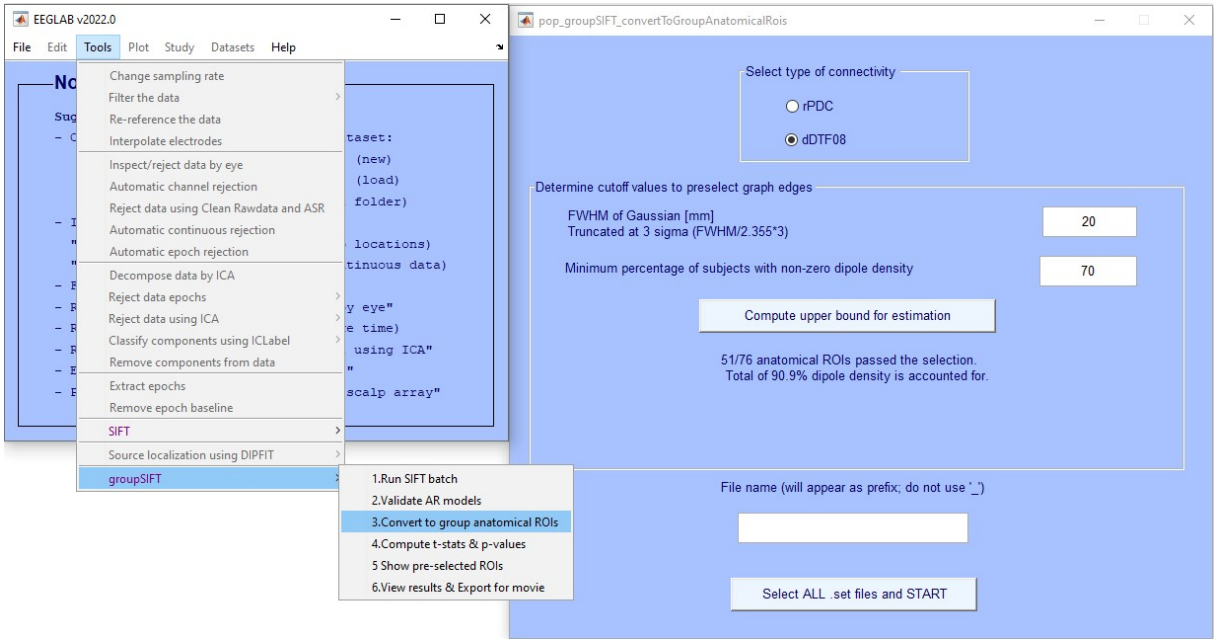


**Figure 4.18.** Convert to Group Anatomical ROI parameter setting

majority of the beta bands (Fig. 4.22); right postcentral to left calcarine region with significance found in alpha and low beta bands (Fig. 4.23); right precuneus to left inferior parietal region with significance found in theta and alpha bands (Fig. 4.24); and right supplemental parietal to the right postcentral region with significance found in gamma bands (4.25). Very little literature is available that analyzes the effective connectivity difference under stress on a whole brain scale. Most work in the field mainly focuses on functional connectivity of stress on a smaller regional

scale or task. Chang's work [45] has primarily focused on the functional connectivity difference under stress in amygdalar subregions. Their results show that the connectivity pattern is mainly altered in centromedial subregions under stress. Nair's work [46] focused on the effect of stress on task-related functional connectivity, which later showed by the experiment results that stress has no significant effect on functional connectivity during verbal problem-solving. Interestingly, Chang published another article [47] in 2019 that studied the hippocampal connectivity of the after-effect of acute social stress using fMRI data. Their results showed that stress altered the information flows in the thalamus-hippocampus-insula/midbrain circuit. Now, going back to the results we obtained in this chapter using the classroom stress data, the author wishes to point out a few observations and possible explanations. First, looking at the five pairs of regions that have increased information flow, one can observe that regions near the central brain are the most common. The right postcentral region is observed to have increased information outflow both to the left calcarine region and the left inferior parietal region as shown in Fig. 4.22 and Fig. 4.23. Moreover, the same region is also shown to have increased information inflow from the right supplemental parietal region (Fig. 4.25). The right precentral region is also shown have increased information outflow to the left calcarine region (Fig. 4.21). Connecting back to the results we observed in chapters two and three, that is, central regions of the brain are found to have distinct power spectra information compared across groups. This observation in power difference might be the effect of a difference in information flow near central regions as observed in this chapter.

## 4.4 Conclusion

In this chapter, we studied the effective connectivity by applying Granger Causality analysis using the SIFT plugin in EEGLab. Later, we discussed that it was challenging to conduct group-level connectivity analysis due to each subject's unique dipole locations. Then an alternative method for comparison is adopted, that is by transforming all the dipole locations

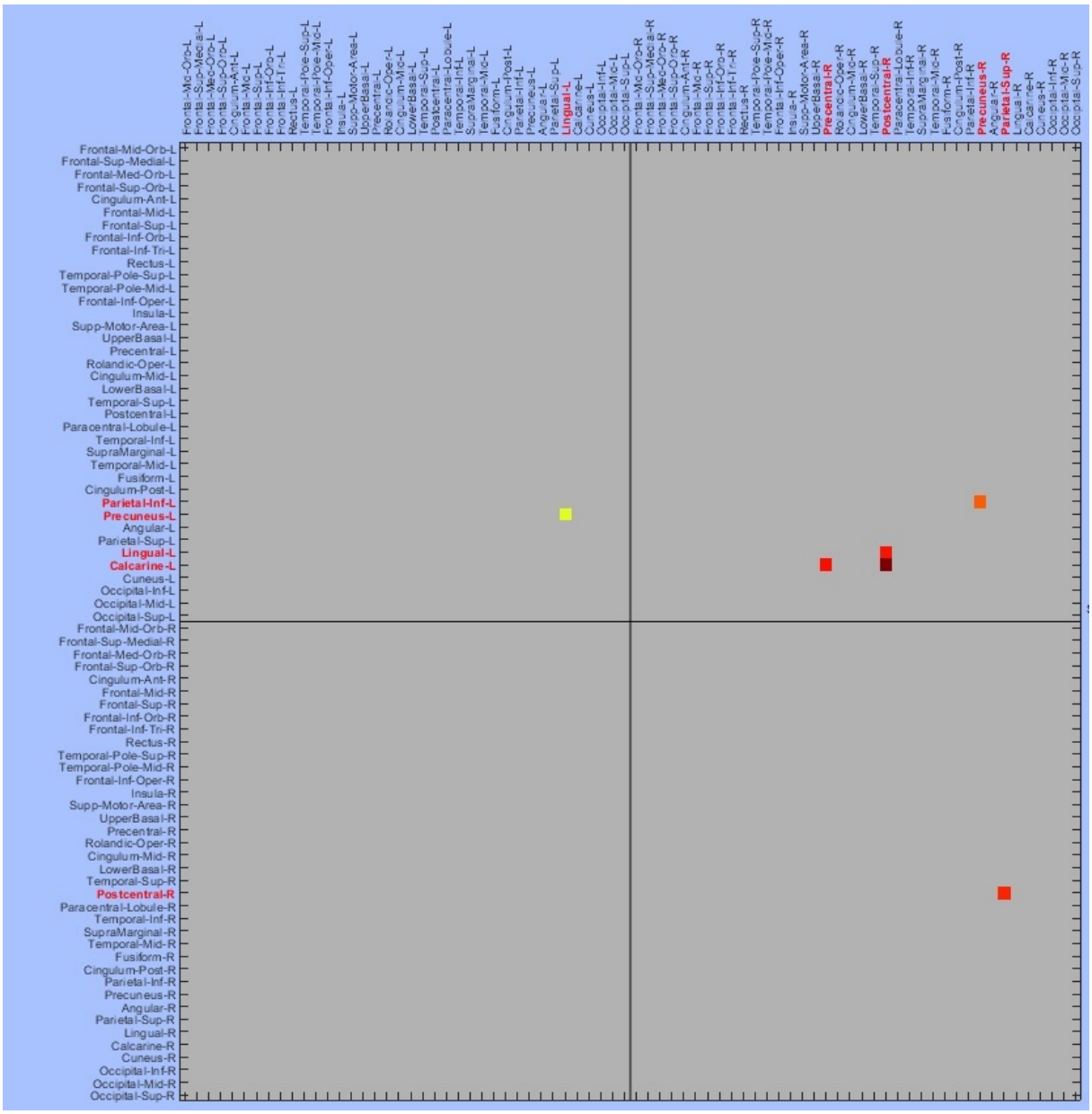


**Figure 4.19.** Overall connectivity matrix under stress, calculated using Elevated stress - Normal Stress, the yellow dot indicates decreased information flow, others indicate increased information flow

into their corresponding spatial regions. By doing this "transformation", a connectivity matrix that was based on common cortical locations was formed. Then GroupSIFT was introduced in great detail describing the steps to carry out this group-level connectivity analysis. In the end, we found that there are six pairs of regions, most near the central brain, that had unique information

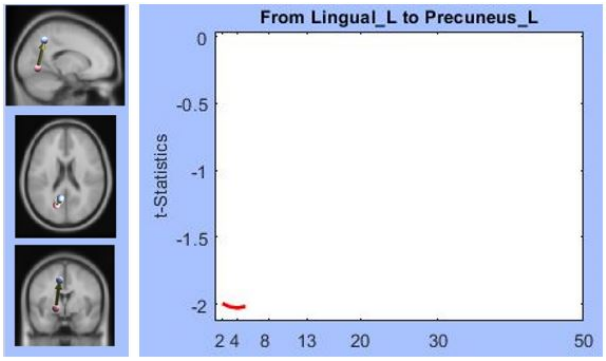


**Figure 4.20.** Decreased Information flow: from left Lingual to left precuneus. Red lines indicate the frequencies that reached statistical significance

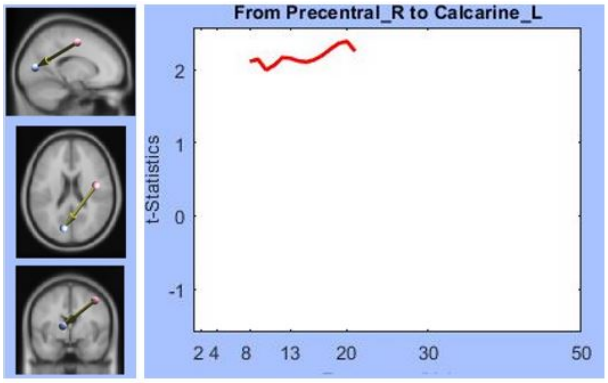


**Figure 4.21.** Increased Information flow: from right precentral to left calcarine. Red lines indicate the frequencies that reached statistical significance

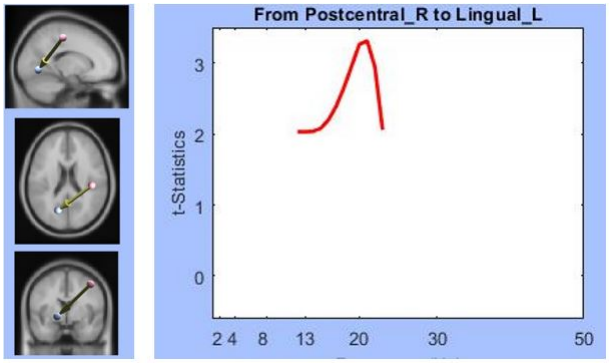


**Figure 4.22.** Increased Information flow: from right postcentral to left lingual. Red lines indicate the frequencies that reached statistical significance

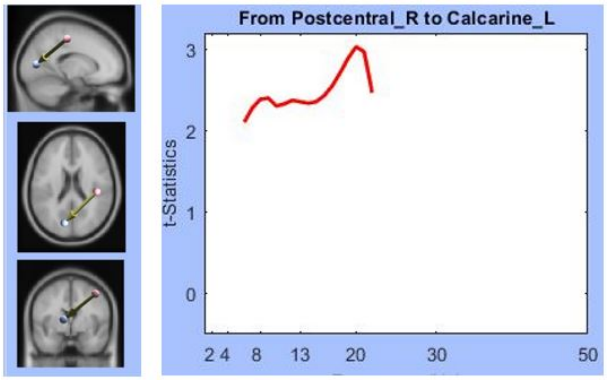


**Figure 4.23.** Increased Information flow: from right postcentral to left calcarine. Red lines indicate the frequencies that reached statistical significance

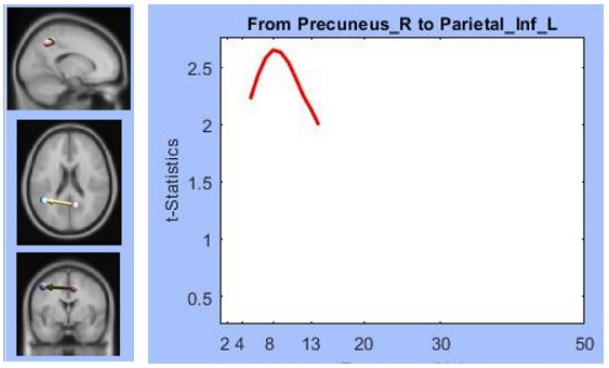


**Figure 4.24.** Increased Information flow: from right precuneus to left inferior parietal. Red lines indicate the frequencies that reached statistical significance

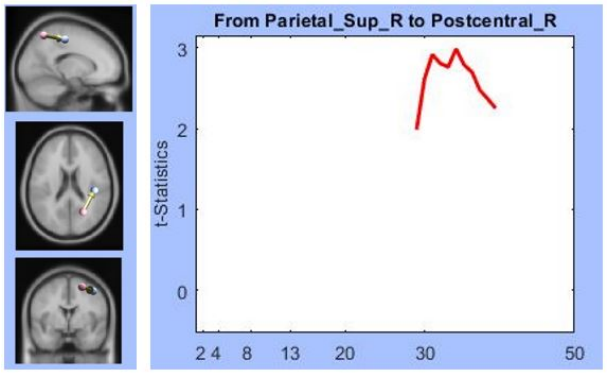


**Figure 4.25.** Increased Information flow: from right supplemental parietal to right postcentral. Red lines indicate the frequencies that reached statistical significance

flow in certain frequency bands among the elevated stress group. Finally, connecting back to the results in chapters two and three, the author makes the hypothesis that the difference in power spectra observed in the central region may be the result of the difference in information flow near the central region of the brain.

# Bibliography


[1] S. Ha, A. Akinin, J. Park, C. Kim, H. Wang, C. Maier, P. P. Mercier, and G. Cauwenberghs, "Silicon-integrated high-density electrocortical interfaces," *Proceedings of the IEEE*, vol. 105, no. 1, pp. 11–33, 2016.

[2] Y.-P. Lin, y.-h. Yang, and T.-P. Jung, "Fusion of electroencephalographic dynamics and musical contents for estimating emotional responses in music listening," *Frontiers in neuroscience*, vol. 8, p. 94, 05 2014.

[3] S. K. Loo, M. Miyakoshi, K. Tung, E. Lloyd, G. Salgari, A. Dillon, S. Chang, J. Piacentini, and S. Makeig, "Neural activation and connectivity during cued eye blinks in chronic tic disorders," *NeuroImage: Clinical*, vol. 24, p. 101956, 2019.

[4] I. Chrikov, K. M. Soria, B. Horgos, and D. Jones-White, "Undergraduate and graduate students' mental health during the covid-19 pandemic," 2020.

[5] H.-J. Park and I.-S. Jang, "Stress, depression, coping styles and satisfaction of clinical practice in nursing students," *The Journal of Korean academic society of nursing education*, vol. 16, no. 1, pp. 14–23, 2010.

[6] A. H. Farabaugh, D. Mischoulon, M. Fava, C. Green, W. Guyker, and J. Alpert, "The potential relationship between levels of perceived stress and subtypes of major depressive disorder (mdd)," *Acta Psychiatrica Scandinavica*, vol. 110, no. 6, pp. 465–470, 2004.

[7] D. Stevanovic, J. Jancic, and A. Lakic, "The impact of depression and anxiety disorder symptoms on the health-related quality of life of children and adolescents with epilepsy," *Epilepsia*, vol. 52, no. 8, pp. e75–e78, 2011.

[8] P. Barreira, M. Basilico, and V. Bolotnyy, "Graduate student mental health: Lessons from american economics departments," *Harvard Univ*, 2018.

[9] B. S. McEwen and P. J. Gianaros, "Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease," *Annals of the New York Academy of Sciences*, vol. 1186, no. 1, pp. 190–222, 2010.

[10] Y. M. Chi, Y.-T. Wang, Y. Wang, C. Maier, T.-P. Jung, and G. Cauwenberghs, "Dry and noncontact eeg sensors for mobile brain–computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 2, pp. 228–235, 2012.

[11] O. Komarov, L.-W. Ko, and T.-P. Jung, "Associations among emotional state, sleep quality, and resting-state eeg spectra: a longitudinal study in graduate students," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 795–804, 2020.

[12] M. M. Antony, P. J. Bieling, B. J. Cox, M. W. Enns, and R. P. Swinson, "Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample.," *Psychological assessment*, vol. 10, no. 2, p. 176, 1998.

[13] C.-Y. Chang, S.-H. Hsu, L. Pion-Tonachini, and T.-P. Jung, "Evaluation of artifact subspace reconstruction for automatic eeg artifact removal," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1242–1245, IEEE, 2018.

[14] R. N. Bracewell, "The fourier transform," *Scientific American*, vol. 260, no. 6, pp. 86–95, 1989.

[15] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[16] M. J. Shensa *et al.*, "The discrete wavelet transform: wedding the a trous and mallat algorithms," *IEEE Transactions on signal processing*, vol. 40, no. 10, pp. 2464–2482, 1992.

[17] E. H. Livingston, "Who was student and why do we care so much about his t-test? 1," *Journal of Surgical Research*, vol. 118, no. 1, pp. 58–65, 2004.

[18] S. K. Haldar, "Chapter 9 - statistical and geostatistical applications in geology," in *Mineral Exploration (Second Edition)* (S. K. Haldar, ed.), pp. 167–194, Elsevier, second edition ed., 2018.

[19] C. Chayer and M. Freedman, "Frontal lobe functions," *Current neurology and neuroscience reports*, vol. 1, no. 6, pp. 547–552, 2001.

[20] J. J. Allen and J. P. Kline, "Frontal eeg asymmetry, emotion, and psychopathology: the first, and the next 25 years.," 2004.

[21] R. Davidson, "Frontal versus perietal eeg asymmetry during positive and negative affect," *Psychophysiology*, vol. 16, no. 2, pp. 202–203, 1979.

[22] A. Cantisani, T. Koenig, H. Horn, T. Müller, W. Strik, and S. Walther, "Psychomotor retardation is linked to frontal alpha asymmetry in major depression," *Journal of Affective Disorders*, vol. 188, pp. 167–172, 2015.

[23] R. Mennella, E. Patron, and D. Palomba, "Frontal alpha asymmetry neurofeedback for the reduction of negative affect and anxiety," *Behaviour Research and Therapy*, vol. 92, pp. 32–40, 2017.

[24] C. Gold, J. Fachner, and J. Erkkilä, "Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression," *Scandinavian Journal of Psychology*, vol. 54, no. 2, pp. 118–126, 2013.

[25] N. van der Vinne, M. A. Vollebregt, M. J. van Putten, and M. Arns, "Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? a meta-analysis," *NeuroImage: Clinical*, vol. 16, pp. 79–87, 2017.

[26] S. Makeig, A. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems*, vol. 8, 1995.

[27] S. Makeig, T.-P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proceedings of the National Academy of Sciences*, vol. 94, no. 20, pp. 10979–10984, 1997.

[28] A. J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 11 1995.

[29] J. Onton and S. Makeig, "Information-based modeling of event-related brain dynamics," in *Event-Related Dynamics of Brain Oscillations* (C. Neuper and W. Klimesch, eds.), vol. 159 of *Progress in Brain Research*, pp. 99–120, Elsevier, 2006.

[30] R. Oostenveld, "Dipfit: localizing dipoles." https://sccn.ucsd.edu/eeglab/workshop06/handout/Robert_Oostenveld-dipfit.pdf.

[31] M. Scherg, "Fundamentals of dipole source potential analysis," *Auditory evoked magnetic fields and electric potentials. Advances in audiology*, vol. 6, no. 40-69, p. 25, 1990.

[32] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent eeg sources are dipolar," *PLOS ONE*, vol. 7, pp. 1–14, 02 2012.

[33] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[34] D. Birant and A. Kut, "St-dbscan: An algorithm for clustering spatial–temporal data," *Data & knowledge engineering*, vol. 60, no. 1, pp. 208–221, 2007.

[35] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[36] M. Syakur, B. Khotimah, E. Rochman, and B. D. Satoto, "Integration k-means clustering method and elbow method for identification of the best customer profile cluster," in *IOP conference series: materials science and engineering*, vol. 336, p. 012017, IOP Publishing, 2018.

[37] M. Yu, O. Sporns, and A. J. Saykin, "The human connectome in alzheimer disease—relationship to biomarkers and genetics," *Nature Reviews Neurology*, vol. 17, no. 9, pp. 545–563, 2021.

[38] R. P. Auerbach, D. Pagliaccio, N. A. Hubbard, I. Frosch, R. Kremens, E. Cosby, R. Jones, V. Siless, N. Lo, A. Henin, *et al.*, "Reward-related neural circuitry in depressed and anxious adolescents: a human connectome project," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 61, no. 2, pp. 308–320, 2022.

[39] C. J. Cook, G. Hwang, J. Mathis, V. A. Nair, L. L. Conant, L. Allen, D. N. Almane, R. Birn, E. A. DeYoe, E. Felton, *et al.*, "Effective connectivity within the default mode network in left temporal lobe epilepsy: findings from the epilepsy connectome project," *Brain connectivity*, vol. 9, no. 2, pp. 174–183, 2019.

[40] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature reviews neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[41] T. Mullen, "Source information flow toolbox (sift)," *Swartz Center Comput Neurosci*, pp. 1–69, 2010.

[42] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.

[43] A. Seth, "Granger causality," *Scholarpedia*, vol. 2, no. 7, p. 1667, 2007.

[44] M. Miyakoshi, A. Delorme, and D. Truong, "Groupsift," 2022.

[45] J. Chang and R. Yu, "Alternations in functional connectivity of amygdalar subregions under acute social stress," *Neurobiology of Stress*, vol. 9, pp. 264–270, 2018.

[46] N. Nair, J. P. Hegarty, B. J. Ferguson, P. M. Hecht, M. Tilley, S. E. Christ, and D. Q. Beversdorf, "Effects of stress on functional connectivity during problem solving," *NeuroImage*, vol. 208, p. 116407, 2020.

[47] J. Chang and R. Yu, "Hippocampal connectivity in the aftermath of acute social stress," *Neurobiology of Stress*, vol. 11, p. 100195, 2019.