ABSTRACT

Title of Thesis: A Study of Resting State FMRI Dynamic Functional Network Analysis of MTBI

Wenshuai Hou, 2015

Dissertation directed by: Professor Joseph JaJa
Department of ECE

Mild Traumatic Brain Injury (MTBI) is one of the most common neurological disorders. A subset of patients develop persistent cognitive deficits. A number of the brain studies have been conducted to discover the abnormalities and disruptions in the brain functional networks using similar methods to those employed in more severe brain disorders such as Alzheimer's or Schizophrenia. Static functional network analysis using resting state brain fMRI images has shown some promising results in identifying characteristics of MTBI. However, recent development in the dynamics of functional networks have been able to reveal insightful information about anomalies in brain activities that have not been observed when using traditional static analysis. Our work focuses on both static and dynamic functional analysis of the brain. Our overall analysis pipeline is data-driven using a dataset of 47 MTBI subjects and a demographically matching healthy control group size of 30. The data-driven approach proactively removes noise, focuses on the entire brain functional networks and performs advanced independent component analysis, followed by statistical tests to characterize the functional networks of MTBI patients.

A key distinction of our research is the finer labeling of MTBI subject according to their long term 6 months recovery status. Our results suggest that those MTBI subject who suffer prolonged recovery exhibit disturbed functional networks, and slowed dynamism in functional connectivity than those of the healthy control or those MTBI participants who recovered quickly. A number of useful network measurements have been found to capture the states and changes of the brain functional networks for healthy and different types of MTBI subjects in their resting state. We believe that our findings can shed more light into the impact of MTBI on the effectiveness of several functional networks and can contribute to helping clinicians make more informed decisions to aid in the recovery of MTBI patients.

A Study of Resting State FMRI Dynamic Functional Network Analysis of MTBI

by

Wenshuai Hou

Master submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of ECE
2015

Advisory Committee:
Professor Joseph JaJa, Chair/Advisor
Professor Amitabh Varshney
Professor Behtash Babadi

Acknowledgments

I would like to express my deep gratitude to those who helped me during my years of graduate study. Without them, this thesis would never have been possible.

Foremost, I would like to thank my research advisor: Professor Joseph JaJa. He inspired me to working on a set of time series problem using the combination of machine learning and data processing techniques. He helped me with generous financial support and provided me an excellent research environment to learn and grow. I also want to thank him for his guidance throughout my research work. He often come up with good idea and invaluable suggestions that guide me make breakthrough in the problems. He is an excellent advisor and mentor who have provided me with years of supports, both academically and educationally.

I also owe my thanks to our collaborators, Dr. Rao P. Gullapalli and Dr. Chandler Sours from University of Maryland - Baltimore medical school. They are the ones who recruited the participants, performed careful and detailed behavioral test and collected the resting state fMRI brain imaging data we used for our study. With their expertise , they helped me to understand the results from a neurological standpoint and keep me updated with the current related literature.

My gratitude also goes to Professor Amitabh Varshney and Professor Behtash Babadi for honor me as my thesis committee and spend their time reviewing my manuscript with valuable feedback. Their insightful questions helped me to solidate the work as well as creates future research opportunities.

I thank my fellow colleagues in lab 3332 in AVW building for making my

graduate years a comfortable and memorable experience. Jing Wu, with her expertise in high performance computing, always able to help to speed up my program and point out the potential use of software packages. With his broad knowledge on Python and Linux system, Mike Ritter helped me to customize my developing environment to boost my productivity. I thank Wang Qi and Jin Yu for many hours of intense discussion on our current researches and the cutting edge breakthrough in compute science filed in general. Whenever, I have problem with my assigned equipment, UMICAS staff is always able to aid me promptly and give advice on related problems. I thank ECE graduate office staff, Melanie Prange, Edna Walker, Arlene Schenk, Vivian Lu, Maria Hoo, Bill Churma, Stewart Heather and many others for their administrative support and their effort to help me whenever I have problem with paperwork.

I would like to thank my parents, two elder sisters, and my girlfriend Tonia Wang for their mental support and constant encourage with their best wishes. Tonia spend many hours to review this manuscript to improve the writing. She always supports me through the good times and bad.

Thanks those all who have helped me, encouraged me, inspired me, had fun with me during my graduate years.

Lastly, thank you all and wish you the best luck from the deepest of my heart.

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Brain Imaging

In the past decades, mapping anatomical brain regions with their biological properties through the use of imaging has been one of the most popular topic in neuroscience. With the developments of increasingly more sophisticated brain imaging, the higher temporal and spatial resolution have enabled researchers to study the brain in unprecedented details. Several areas in brain science have achieved major discoveries due to the advancements in brain imaging technology, areas such as cognitive neuroscience, affective neuroscience, and neuroeconomics. On the other hand, another critical topic that exploits the technology advances is neuro-pathology, which is the study of the diseases of the nervous system tissues. By studying the functional connections captured by the acquired brain imaging data, we are gaining important information about brain diseases such as mild traumatic brain injury (MTBI) and schizophrenia.

1.1.1 Structural And Functional Brain Imaging

Depending on the goals, brain imaging techniques can be mainly categorized as structural or functional, defined as follows.

1. *Structural Brain Imaging:* This type of study focuses on the study of the anatomy of the brain structure, attempting to provide detailed anatomical organization of a typical brain. Common imaging tools used to infer such information include Computed Axial Tomography (CAT), Diffusion Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET). For example, CAT measures the absorption of x-ray beams as they pass through body tissues, where the bone has the highest attenuation and appears white, whereas gas has the lowest attenuation and appears as black. Fig. 1.1 provides an illustration of a chest X-ray image. Moreover, CAT can be used to compute cross sectional images. In brain imaging applications, a ring of x-ray generators and detectors generate a serial of cross-sectional images of the subject's brain during the CT scan. The scanned images represent the direct tissue structure, and abnormalities can be detected from these images [134].

   The MRI technology involves the application of a static magnetic field which forces on the alignments of lower energy nuclei of the brain. When the magnetic field pauses, the detector coils measure the energy level that indicates the return rate of the nuclei to their original low energy state. This rate determines the type of image generated: T1 image differentiates gray and white

matters, and T2 portrays white matter hyper-intensities [138].

Diffusion MRI attempts to capture the diffusion of water molecules in body tissues. Water molecules naturally move randomly according to a Brownian motion in an isotropic medium. However, due to the special properties of body tissues, the diffusion may be an-isotropic, and the molecule moves primarily along neural fibers. The advanced development of diffusion tensor imaging (DTI) enables the measurement of diffusion in multiple directions, which allow researchers to measure brain connectivity and fiber directions of different regions of the brain [135]. Several examples of DTI imaging can be found in Fig. 1.2, [136].

2. *Functional Brain Imaging:* This type of study focuses mainly on discovering concurrent activations of different parts of the brain due to certain cognitive and affective processes. Major brain imaging technique to support this type of study includes PET, functional MRI (fMRI), Electroencephalography (EEG) and Magnetoencephalography (MEG). For example, quantitative EEG measures the coherence of brain activities between two different cortical areas. If the synchronization increases during a task, the coherence measure will increase. Advantages of quantitative EEG includes its wide availability, low cost, non-invassiveness and avoidance of radiation exposure. In particular, the change in coherence in the resting state and in performing a task may be important for diagnosing dementia [137]. fMRI measures the blood flow activity tied to neural activities. The excessive amount of oxygen brought to

Figure 1.1 CT imaging of chest. Bones and muscles have larger density and appears to be white, air in the lungs has smaller density and appear to be black.

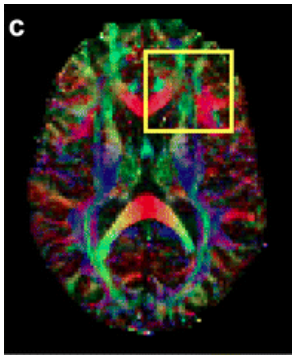


Figure 1.2 DTI technology is able to measure the water molecule diffusion in tissues which indicates neural fiber directions [136].

the brain region in use decreases the magnetic field gradient, which leads to greater magnetic homogeneity that yields greater blood flow [139]. Different functional imaging modalities provide different spatial and temporal resolutions and involve a range of invassiveness.

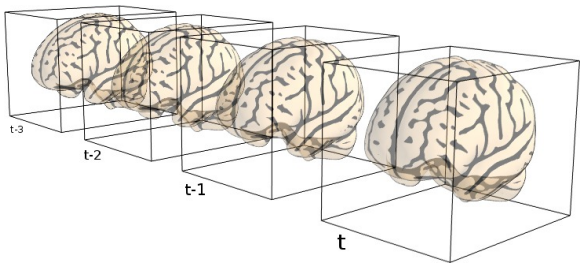


Figure 1.3 4D fMRI data is a temporal series of 3D brain scan.

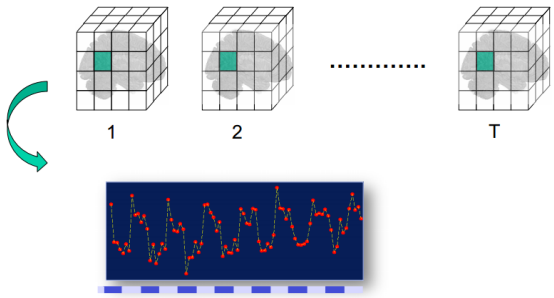


Figure 1.4 Each voxel corresponds to one spatial position in the brain, and each voxel has a time series representing its temporal activation during the fMRI session.

1.1.2 fMRI

fMRI is a functional neuroimaging technique using the MRI technology that scans brain activities by detecting associated changes in blood flow in the brain [169]. The technique is made possible because the neuronal activities demands oxygen and sugar to be brought in quickly. The oxygen level causes a difference in magnetic susceptibility in the blood. One reason for the popularity of fMRI is due to its good balance between the temporal and spatial resolutions combined with its non-invassiveness. Hence it has gained popularity in functional imaging research [140]. Subjects are not exposed to radiation as in CT technology, and therefore there are no side effect on the subjects. fMRI scans the entire brain indiscriminately, unlike EEG/MEG methods which are more effective when performed on the cortical surface. fMRI captures activities with a relatively high spatial resolution typically $2-3$mm and can be as accurate as 1 mm.

Among many useful measures of blood flows, we focus on the blood-oxygen-level dependent (BOLD) contrast, discovered by Seiji Ogawa [170], who demonstrate that the difference of magnetic susceptibility between oxyhemoglobin and deoxyhemoglobin can be detected using MRI.

An fMRI scanner places a static magnetic field through the subject's brain, which will force some low energy nuclei to align with the external field. Because the discrepancy in magnetic susceptibility in oxygen-rich and oxygen-poor blood, the two types of blood react differently to the sudden pause of magnetization. Oxygen rich blood exhibit a higher degree of magnetic homogeneity, which means aligned

nuclei return to its original lower energy state at a faster rate and emits higher energy than oxygen-poor blood. Therefore, for those brain regions that have a richer amount of oxygen, the fMRI intensity will be greater than the baseline signal. Unfortunately, the BOLD signal can be easily corrupted by noise from various sources, and hence the need to apply signal processing and multiple statistical tools to improve the signal to noise ratio (SNR) and to extract the underlying signal.

The acquired fMRI data consists of a 4D matrix, representing a time series of 3D brain scan, as shown in Fig. 1.3. The fMRI data is relatively large, there is about $100,000$ voxels in each typical brain scan. Each voxel corresponds to a volume in a spatial location. During the course of the fMRI scan, hundreds of brain scans are captured at an interval defined as TR, usually TR = 2 seconds. By tracking the intensities of a voxel over time, we obtain a time series for each voxel.

As for neuro-pathology research, fMRI is the primary tool for neuroimaging that the researchers have used to study various brain dysfunctions, such as Alzheimers [141–143], schizophrenia [144, 145], and MTBI [146–148], the latter being also the topic of this research. It is widely adopted as a diagnostic tool that aids physicians to learn how a healthy, or injured brain works, and that helps assess the potential risk of certain surgeries and treatments. Lastly, due to its vast popularity among neuro-researchers, in-depth studies have been conducted based on data-sets produced by fMRI , thereby enabling scientists to easily compare their results across studies and platforms.

1.2 Resting State fMRI and Resting State Network

1.2.1 Resting State fMRI

fMRI has been widely used to study the brain mapping related to the processes of cognition and emotion. Traditional fMRI research has been primarily task-oriented. The subjects are told to perform certain tasks to activate and increase the neural activity in related brain regions [131, 132]; for example, tapping a finger or counting, watching pictures are used respectively to understand the activations of the motor and visual networks. Often used with the task-related fMRI research is the Seed analysis, which provides a method to verify prior hypotheses by either calculating the correlation between seed areas or by modeling the relationship between the seed region and all the other regions in the brain. Many leading functional mapping techniques of the brain have been determined following this procedure [166–168]. However, forming an interesting hypothesis and the design of the corresponding task to effectively verify the hypothesis can sometimes be challenging. Data-driven approaches are often preferred for thier exploratory nature in detecting patterns that lead to valuable discoveries.

On the other hand, the rs-fMRI studies require that the subject be at rest and search for functional mappings between brain regions by examining the relationships of the corresponding fMRI time series. In the rs-fMRI setting, the subjects are told to relax and not to think of anything in particular, while their spontaneous brain BOLD signal is measured throughout the fMRI sessions. It also opens the

door for studying subjects with difficulty to perform certain tasks. rs-fMRI is a relatively new type of fMRI research which has drawn increasing attention since it was first introduced by Biswal [167] in 1995. In his work, the subjects were told not to perform any task during their fMRI sessions. The seed region was set to be the left somatosensory cortex (see figure 1.6) because it has been found to be related with bilateral finger tapping in earlier studies. The correlations between the seed region and all the other brain regions were examined and eventually a strong correlation was found between the left somatosensory cortex and homologous areas in the contralateral hemisphere. This pioneering work was confirmed and replicated later through anatomical studies which found that left somatosensory cortex is in high synchronization with primary and higher order somatosensory areas. Prior to Biswal's work [167], there have been many suspicions on rs-fMRI due to the fact that SNR in rs-fMRI is very low, which was caused by various sources of noise and other important physiological factors such as the respiratory and cardiovascular cycles measured in the BOLD signal. Biswal's [167] research dispels these doubts on rs-fMRI, since many contemporary scientists attribute the resting state signal to other sources such as heart beating and observation noise. This research has revealed that the brain network is not idle when the subject is resting, and in fact carries out a vast amount of spontaneous activities that are highly correlated among different brain regions [129].

1.2.2 Basis of rs-fMRI Analysis

rs-fMRI focuses on the low frequency fluctuations that range from 0.01 to 0.1 Hz of the rs-fMRI BOLD signal [149, 159, 167]. The underlying neuronal basis of these low frequency rs-fMRI fluctuations is still an open question, with some researchers arguing that the rs-fMRI signal is merely an artifact of various physiological processes, such as the cardiovascular and respiratory processes [160–163]. However, the support for the significance of these low frequency fluctuations originates from the underlying spontaneous neuronal activation that stems from the fact that most of these RS patterns tend to occur in brain regions that overlap in both function and neuro-anatomy [116, 129, 165, 167]. This observation suggests that multiple brain networks work together to complete functional tasks occurring during resting. The correlations are particularly high between those anatomically separated networks with similar functionality [129]. Moreover, it has also been reported that cardiovascular and respiratory components have different frequency than the rs-fMRI frequency of interest [149, 150]. Overall, an increasing number of researchers agree with the neuronal basis of rs-fMRI signal, and treat the cardiovascular and respiratory components as noise. This is the reason why it has become a standard practice to pre-process the rs-fMRI data to reduce the corruption caused by physiologic processes and other noise sources.

It has been observed that several anatomically separated regions display a high level of synchronization in their rs-fMRI time series. This implies that these regions form a network that carries out a task together when the brain is at rest. The

identification of resting state networks (RSN) has drawn increasing attention over the past decade. In the signal processing setup, this problem can be viewed as a blind source separation problem, and tools such as Independent Components Analysis (ICA) or seed analysis are adopted to support rs-fMRI applications [151–153]. rs-fMRI studies have successfully identified a number of RSNs: synchronous activation between spatially distinct brain regions without the presence of external stimulus. More details will be covered in the following section.

1.2.3 Resting State Network (RSN)

The human brain is often modeled as a set of networks that interact with each other to carry out a particular task. This type of functional mapping has become one of the most popular topics in cognitive neuroscience [121, 122]. The brain networks can be further classified as task-positive and task-negative networks [123]. The former includes brain regions that are activated when we perform tasks and cognitive processes, in contrast to task-negative regions that are more active when the subject is at rest without any external stimulus. Anatomically separated brain regions that show high synchronization behaviors during rest are defined to be RSNs [115]. A number of studies have consistently reported the existence of RSNs [115, 116, 165].

The most well known and widely studied RSN is the default mode network (DMN), which includes the Medial prefrontal cortex (MPFC), Posterior Cingulate Cortex (PCC), lateral and medial temporal lobes and posterior inferior parietal lobe [113, 114]. See figure 1.5 for an illustration [15]. On the other hand, the key

example of task-positive networks is the so called core network which is composed of a large number of sub-networks: bilateral insular regions, anterior cingulate cortex, two lateralized parietal frontal networks, primary sensorimotor network and visual networks. The DMN and core networks work in an alternative way: the activation of DMN when the subject is at rest is usually coupled with deactivation of the task-positive core network, whereas when the subject is performing a cognitive task that involves the core network, a suppression of the DMN has been observed [123, 124]. The ever-alternating activation of the core and DMN networks have been suggested to be an important factor in the brain cognitive processing [129, 130].

DMN was first identified by Raichle in 2001 [113] from PET data and by Greicius from fMRI data two years later [114]. During the PET study, the subjects were told to rest with their eyes closed. DMN regions were determined as the regions that are more active when the subjects are resting than performing cognitive tasks. This finding was confirmed by Greicius using fMRI and consistently verified by other researchers using various data and analysis tools [115–120, 165].

Although still in its development stage, DMN also has important clinical implications, Many studies of DMN have found that the direct relationship between DMN dysfunctions and major psychiatric and neurodegeneratie diseases [125–128].

Other than the DMN, seven other RSNs have been found consistently throughout many neuroimaging studies [115, 116, 129, 154–156, 165]. The locations of these eight RSNs are illustrated in Fig. 1.7 extracted from paper [129]. These networks have been reported consistently over many studies, using different subject groups, data acquisition methods, and processing procedures. These networks include

1. The primary motor network, located in the posterior portion of the frontal lobe, which works with other motor areas to perform movements.

2. The primary visual network, located in the occipital lobe, is responsible for processing visual information.

3. The extrastriate visual network, which is part of the occipital cortex next to the primary visual network. It is devoted to higher order visual processing [157].

4. The two lateralized networks consisting of superior parietal and superior frontal regions. It plays important roles in integrating sensory information from sensors of the body, knowledge of numbers and their relationships, and manipulation of objects [158].

5. The DMN is composed of precuneus, medial frontal and inferior parietal and temporal regions. DMN is active when the subject is at rest without any external stimulus. Its functionality is still an on-going debate [104].

6. The network consisting of bilateral temporal/insular and anterior cingulate cortex. It is believed to play an important role in consciousness, diverse functions such as emotions and the regulation of the body's homeostasis [102, 105, 111, 112].

7. The frontal network, which consists of neurons that are in charge of decision making, problem solving, and control of purposeful behaviors [103].

The functional synchronizations of these spatially non-overlapping RSNs coincide with our knowledge of the functional networks. These observations suggest

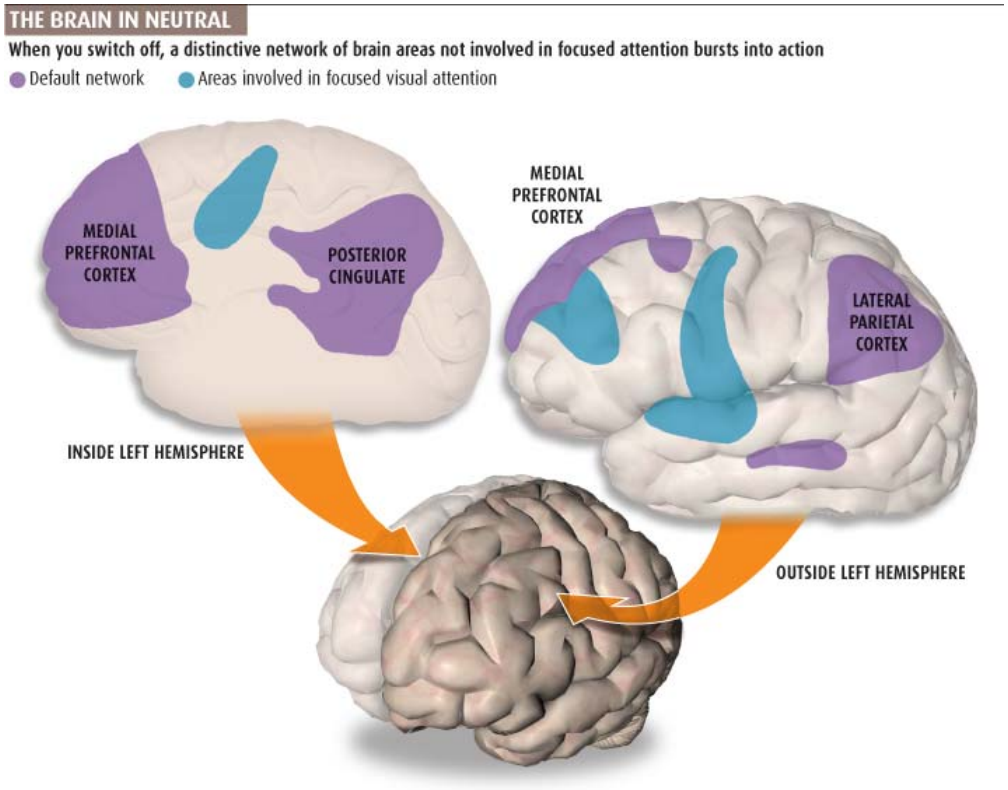


Figure 1.5 Default mode network includes the Medial prefrontal cortex (MPFC), Posterior Cingulate Cortex (PCC), lateral and medial temporal lobes and posterior inferior parietal lobe.

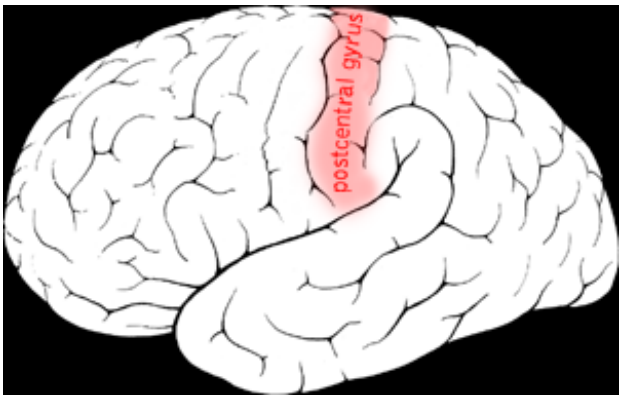


Figure 1.6 Illustration of somatosensory cortex, which locates at the lateral postcentral gyrus in the parietal lobe of the human brain.

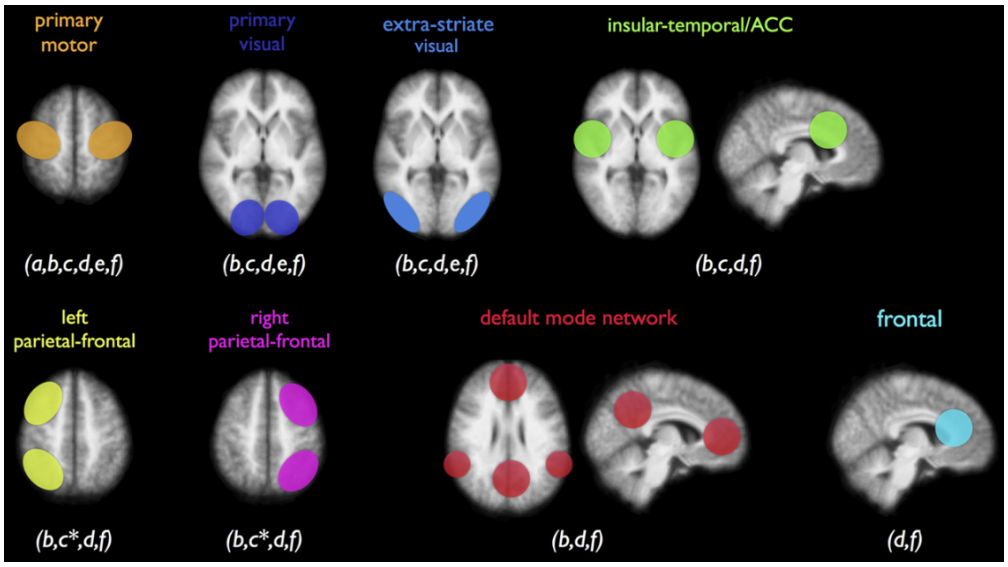


Figure 1.7 The consistently found RSN using different dataset and different methods, these networks include the primary sensorimotor network, the primary visual and extra-striate visual network, a network consisting of bilateral temporal/insular and anterior cingulate cortex regions, left and right lateralized networks consisting of superior parietal and superior frontal regions (*reported as one single network) and the so-called default mode network consisting of precuneus, medial frontal, inferior parietal cortical regions and medial temporal lobe [129].

that even when the brain is at rest, the brain is still constantly performing and coordinating tasks that require an entire network to cooperate, and there exist certain connectivity patterns in the resting brain to synchronize brain regions with functional overlapping.

1.3 Traditional Functional Analysis

The rich literature of rs-fMRI provides us with many useful tools to study the connectivity between non-overlapping brain regions, including studying subnetworks within and between RSNs. The most common tools used are: principal components analysis (PCA), singular value decomposition (SVD), clustering, independent components analysis (ICA), and seed analysis [106–108, 150, 154, 167]. Roughly, the rs-fMRI analysis can be divided into two major groups: seed analysis and multivariate decomposition [129], which will be discussed in the following sections.

1.3.1 Seed Analysis

This is the simplest way to study functional connectivity between brain regions. A brain region of interest is fixed, which is defined as a seed, and the correlations of the seed region with all the other regions in the brain are computed. The seed method produces a one to all connectivity map, and this map provides clear information to what extent other brain regions are connected to the seed, and whether this relationship is positive or negative. The choice of the seed is usually

pre-determined, either an interesting brain region in an rs-fMRI study, or in the case of task-related fMRI analysis, the seed is chosen to be the area where the task is believed to stimulate. The simplicity of seed analysis allows researchers to make conclusions with strong evidence and confidence [88]. However, the seed method is highly limited by the prior hypothesis and is not suitable for extracting interesting features for the whole brain analysis [129].

1.3.2 Multivariate Decomposition

This type of analysis is also called model-free analysis [129]. It is usually used to study the whole-brain connectivity map without making any prior assumptions about the fMRI data. Among all the techniques that belong to this category, the ones that proved the most useful are PCA, ICA, Laplacian clustering, hierarchical clustering, and normalized cut clustering [89, 90, 101, 115, 116, 140, 155, 156]. Among them, the most widely adopted and the most consistent method is the ICA [89, 115, 156].

ICA is designed to solve the blind source separation problem. In the rs-fMRI setting, it identifies a set of independent temporal or spatial patterns under the assumption that the observed fMRI data is a linear mixture of this set of independent components. ICA is a tool for the whole brain analysis; however, with an extra step of regression or back construction, it enables us to compare the time series of the independent components within a group or across groups of subjects. The extracted spatial patterns often overlap known functional sub-networks.

Other than the ICA method, clustering algorithms are also used in fMRI applications [101, 155]. They function in a similar way as the ICA method. ICA seeks to combine simultaneously activate voxels together, and to separate those voxels that are independent. A clustering algorithm, by minimizing the within-cluster distance, and maximizing between-cluster distance, achieves a similar goal, where the relationship between the voxels fMRI time series are captured by their Euclidean distance. Due to the straightforwardness of the clustering methods, the corresponding results are easier to interpret than those obtained by the ICA method. Their results overlap to a certain degree. For example, the existence of DMN is discovered by seed analysis [98, 114], clustering methods [101], and also by group level ICA method [115, 165]. Similarly, the intrinsic functional connectivity between primary visual and auditory regions have been found by all three methods [129]. A distinct advantage of the group ICA method is the convenient comparison between subjects and groups, made possible by the group level ICA Decomposition which generates a set of spatial patterns that are common for all the subjects. Moreover, advanced development of ICA can remove the noise and produce maximally reproducible and stable group level patterns [183].

1.3.3 Pre-Processing and Noise Removal

The measured rs-fMRI signal is unavoidably confounded by various random noise components that arise from both the scanner artifacts and the subjects themselves [99]. The significant and controllable sources of noises are [100]: thermal noise

due to free electrons, patient head motion, physiological processes, such as cardiovascular and respiratory process, and finally low frequency signal drift. Pre-processing attempts to increase SNR. With overwhelming noise sources, in addition to the intrinsic low SNR in the rs-fMRI, it becomes critical to pre-process the rs-fMRI data, using the following common procedures: motion correction, slice-time correction, spatial filtering, temporal filtering, registration and normalization. The rest of this subsection will be devoted to explain these noise sources and the corresponding pre-processing methods, as well as the popular software packages used.

1.3.3.1 Motion Correction

During the rs-fMRI scan, the subjects are told to stay relaxed, not to perform any task, and to stay still. However, the subjects always move during the lengthy 30-minutes fMRI scan, even a very slight head movement can cause significant signal change, and voxel misalignment. To illustrate how significant this issue is, 1% head motion near strong intensity boundaries may cause 1% voxel signals shift and this signal change maybe even bigger than the actual BOLD signal; even worse, the head movement cause voxel misalignment, which means the captured signal doesn't correspond to a consistent anatomical voxel over time [91, 93, 94], The solution of motion correction is co-registration, which is the process known to spatially align two images. For each subject, simply choose one fMRI image as the target/reference and register all the other volumes to the reference. Fig. 1.8 shows how to detect a motion. The sudden jumps in the green time series illustrates the inconsistency

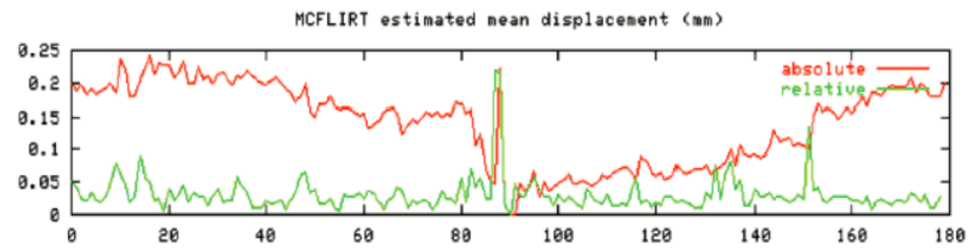


Figure 1.8 Summary of Motion. The red time series shows the difference between consecutive time points, and the green time series shows how the other volume deviates from the references. Both time series is able to show the sudden spike which corresponds to the head motion, and the red time series also displays the low frequency drift.

between the current signal strength and that of the next time point, which is likely due to a sudden head motion. Moreover, the red time series also shows that there is a low frequency drift in addition to the motion, which will be addressed later.

The registration uses 6 degree of freedom (DOF) rigid body transformation which assumes that the size of the two images are identical, and the superimposition can be controlled by a combination of three translations and three rotations. The effects of motion can be corrected by modeling each motion component as a regressor in the Generalized Linear model (GLM) which allows non-Gaussian error, and remove their effects in the observed time series [97]. Alternatively, we can model each freedom as a variable, and propose certain objective functions that will transform the problem into an optimization problem [95].

1.3.3.2 Slice Timing Correction

Within each volume, it appears that the whole brain is scanned at the same time, because most fMRI scans are only able to acquire one slice at a time, the next slice is scanned immediately after the scan of the previous, but still at a slightly later time within a certain Repetition Time (TR). In this case, the last acquired slice appears to lead the change than the earlier slice, although the actual activity happens simultaneously. Slice timing correction can be done by using temporal interpolation, which identifies the right signal amplitude from discrete nearby time points. All slices have to be synchronized to a single time point within a TR interval. The choice of this time point can be either the first time point or the middle of the TR, the former integrates easily with Gaussian linear modeled GLM, whereas the latter requires less interpolation.

Slice timing correction is often considered as an optional step in pre-processing. Because the haemodynamic response (the delivery of blood to active neuronal tissues) can be sluggish, some researchers argue that this correction is not necessary if the TR is set to be small. However, it has been shown that, under certain assumptions, this correction improves the results, especially for task related fMRI cases [96].

1.3.3.3 Spatial Filtering and Smoothing

The noise mixed with BOLD signal is usually assumed to be Gaussian noise with zero mean, and therefore, averaging the signal can even out the noise. This

is where the spatial filtering plays a role. It is simply an average of one voxel's signal with its neighbors. The smoothness is often a desirable property for analysis methods, such as the thresholding based on Gaussian field theory. However, a possible risk we encounter especially in rs-fMRI is if the actual signal is already small, it can become even harder to detect it after the smoothing process; this can be addressed by using a smaller size smoothing kernel. The smoothing kernel is similar to a window function that imposes different weights that control the smoothing process. For example, a popular choice implemented by the FSL package is Gaussian full width half maximum (FWHM) kernel [91]. With this kernel, the smoothing is implemented as a 3D convolution, where the new intensity of each voxel intensity is computed as a weighted average of neighboring intensities, the weighting and the neighborhood size is controlled by a Gaussian function that shown in Fig. 1.9. In the case of FWHM, the size is determined by the rate of decaying of the chosen kernel, the kernel includes the weights that are larger than 0.5 of its peak value.

1.3.3.4 Temporal Filtering

The spatial filtering is to remove the low frequency Gaussian noise contained in the signal and allow the signal to be sufficiently smooth for subsequent analysis. Temporal filtering can improve the SNR substantially by filtering out the high frequency noises which are likely due to scanner, physiological process and noises, and low frequency drift mentioned earlier [92, 99]. This process also serves as a mask that will be discussed later. Temporal filtering utilizes a high pass filter (HPF) that

Figure 1.9 An example of Gaussian FWHM smoothing kernel, it controls the weights, and the size of its neighborhood in consideration.

only let through high frequency components, and a low pass filter (LPF) that keeps low frequency components. Note that the cutoff frequency of the LPF is higher than the cutoff the HPF. The cutoff of the LPF is often set to be around 0.2 Hz, to remove the high frequency components due to breathing (0.3 Hz) and heartbeat (1 Hz). The general choice of the HPF is 0.01 Hz to remove the lower part of the spectrum, that is usually the noise caused by scanner drifting. There are other methods to remove this slow scanner drift. For example, the SPM package models this slow drift with cosine basis set to identify variance. FSL package uses a high pass filter but convolves the signal with a Gaussian-weighted function as a further temporal smoothing [91, 93, 94].

1.3.3.5 Image Registration and Global Intensity Normalization

The final step is to register all images to the same brain template, which is an anatomical brain template. This step is necessary because subjects have different brain shape and sizes. Registration to the same K-space (2D or 3D Fourier transform) makes the following analysis possible and easier.

Global intensity normalization makes the mean consistent among all subjects. Different subject rs-fMRI data have different means, even the same subject in two different sessions lead to a difference in the mean. This is caused by various less relevant factors such as the caffeine level [34]. Global intensity normalization is a simple procedure to scale each 4D fMRI data to obtain the same mean [91, 93, 94].

1.3.4 Functional Network

Instead of connections between a few brain regions, there has been an increasing interest in investigating whole brain networks using graph methods. Brain mapping is the study of the brain regions that are activated by certain tasks, the relationship between non-overlapping brain regions, how they interact with each other, and how to discover these interactions. Investigating the topology and properties of the brain network can provide invaluable insights on how the brain works [70–73].

Brain networks can be constructed by the seed method, ICA or clustering algorithm. The functional network is defined as a graph $G = (V, E)$, where V is the non-overlapping brain regions as nodes, and E is the set of edges. In the case of a weighted graph, the edge weight captures the quantitative relationships

between the corresponding nodes. The network can be bi-directional or undirected depending on the design of the study. The main advantage of this method is that no biological assumptions are needed to be made a priori, statistical tools and feature extractions can determine if there are interesting patterns in the dataset, or if there is a difference between the two groups of subject under study.

1.3.4.1 Connectivity Computation

The computation of the edge weight can be done in a number of ways depending on the assumption and the goal of the study. The most popular and straightforward quantitative measure that has been adopted by various studies depends on the correlation or correlation coefficient between the time series of nodes in V. For example, nodes $u,\ v \in V$ have time series $x_u,\ x_v$ respectively, each with T time points. Their correlation can be computed as,

$$r_{u,v} = E[(x_u - \overline{x_u})(x_v - \overline{x_v})]$$

this measure corresponds to the covariance, but the mean of the time series is often centered to zero, and hence correlation and covariance are the same in this case. Correlation coefficient is a popular alternative, because it leads naturally to Fisher's z-transformation, which is adopted in a number of studies [74–76].

$$\rho_{u,v} = \frac{cov(x_u,\ x_v)}{\sigma_u \times \sigma_v} = \frac{r_{u,v}}{\sigma_u \times \sigma_v}$$

where $\sigma_u,\ \sigma_v$ are the standard deviations of time series $x_u,\ x_v$ respectively. Correlation related metrics are popular for computations of connectivity because of their

low complexity, they are easy to understand, and well suited for integration with other statistical methods.

1.3.4.2 Network Properties

Given a functional network G constructed from measuring connectivity between brain regions, a number of studies have established a possible efficient topology of the graph, which is called *small world property*. This small world network property is often known for its high connectivity in its local neighborhood of major hubs measured by clustering coefficient C which will be defined later, yet still achieves short path length L at a global scale. The small worldness is quantified by comparing the network of interest G to the mean clustering coefficient and path length estimated in a random graph with the same edge degree distribution as G. For a small world network, we expect the ratios $\frac{C}{C_{random}} \approx 1$ and $\frac{L}{L_{random}} \ll 1$. This observation has been made not only by rs-fMRI studies [80, 83] but has also been supported by MEG and EEG studies [84–86]. Studies that use anatomical templates have shown that the probability that a node has k connections follows an exponential truncated power-law, which is an unique characteristic of small world network [80]. Studies have hypothesized that certain highly connected hubs in the functional brain networks are those targeted by Alzheimer's disease, which reduces the efficiency of the brain operations [87]. Overall, these studies show that the brain functional network is not a random network, but rather a highly efficient network that has several highly connected nodes that serves as hubs with strong local

connectivity.

A link between the efficiency of the functional network and the subjects' cognitive abilities has been proposed by many pioneering rs-fMRI studies, mostly focusing on the connectivity within the DMN, because DMN is a distinct network for resting state brain activities [70, 81, 82] .

The patterns of a functional brain network have also been successfully linked to indications of certain brain diseases, reported by an increasing number of studies. These studies suggest that neuro-degenerative diseases affect the interconnected cortical networks, rather than a single local region [128] Alzheimer's disease has been reported to be linked to decreased connectivity within DMN [109]. Greicius' study found reduced activities in the posterior cingulate cortex (PCC) and hippocampus in the functional network of Alzheimer's patients. Research in [125] noted a decrease in the clustering coefficient for the functional network in the Alzheimer group. These rs-fMRI studies are supported by resting state MEG studies, where decreased brain network efficiency was reported for subjects with Alzheimer's [71, 110].

Functional disconnectivities of the functional network have been reported with the schizophrenia disease [60, 61, 140], which can be potentially a disconnection disease [62]. Studies have suggested aberrant DMN connectivity in schizophrenia disease, reporting a reduction in connectivity between medial frontal cortex and precuneus [63, 98]. These studies have suggested that the DMN plays an important role in the pathophysiology of schizophrenia [98]. Anatomical studies suggest that schizophrenia is related to aberrant information integration between brain regions, which leads to an overall decreased brain network efficiency [64, 85, 142]

These studies have brought forth some interesting and useful network algorithms that can be used to explore the brain network efficiency. For a network $G = (V,\ E)$, the shortest path determines a path between the vertices such that the sum of the weights on the path is minimized. The single source SP problem can be solved efficiently using Dijkstra's algorithm which determines the shortest path using a greedy strategy [65]. The all pairs shortest paths can be computed by the Floyd Warshall algorithm [67–69]. The most straightforward evaluation of the efficiency of a network is the average shortest path, that is the average shortest path between all possible pairs of nodes,

$$a = \sum_{s,t \in V} \frac{d(s,t)}{n(n-1)}$$

where $d(\bullet, \bullet)$ is the length of shortest path value between vertices $s,\ t$. Another useful quantitative evaluation is the minimum spanning tree (MST), which serves as the backbone of a connected and undirected graph G. A spanning tree of a graph is a tree that contains all the vertices. The MST is the spanning tree with the minimal summed weights of the edges of the tree. The MST can be found efficiently with well known greedy algorithms such as Prim's and Kruskal's algorithms.

1.4 Mild Traumatic Brain Injury and Glasgow Coma Scale

Mild Traumatic Brain Injury (MTBI), often called concussion, although the latter is a sub-class of the former, is caused when a severe impact or a impulsive force is too strong to be absorbed by the cerebrosipinal fluid, which severs as a cushion around the brain [50]. There is no universally accepted definition of MTBI, but it

has been widely defined as a head injury with a temporary loss of brain function, which leads to various possible subtle or severe physical, cognitive and emotional symptoms. The common evaluations of MTBI includes some combination of loss of consciousness (LOC), post-traumatic amnesia (PTA), and the Glasgow Coma Scale (GCS).

MTBI incurs a variety of symptoms, which only appears immediately after the head/brain injury [46]. Most symptoms are resolved in a short period of $7-10$ days; however, the number and type of symptoms each patient suffers varies greatly [47]. Among them, the most common ones are headache, dizziness, vomiting, nausea, lack of motor coordination, difficulty in balancing [48]. MTBI also impairs the cognitive ability of the patients, including confusion, disorientation, difficulty to focus, and possibly LOC in some cases [49].

The physiological results of MTBI are often unpredictable and can last from hours to years [51]. For example, in one animal study, MTBI was followed by an initial growth in glucose metabolism, which was followed by a subsequent reduced metabolism that may persist up to four weeks after the incident [52]. The subsequent metabolic processes are reversible for the majority of affected brain cells; however, a small portion of the cells may die due to the event [53]. It is commonly observed that the cerebral blood flow is relatively reduced, although the causes are still unknown, this leads to an energy crisis to the cerebral cells because the reduced amount of glucose supply carried by blood [54]. MTBI causes diffusive brain injury, the dysfunction occurs over entire brain networks instead of the spot of impact [55].

Diagnosis of MTBI is a combination of physical and neurological examinations,

Eye Opening Response	Verbal Response	Motor Response
4 = Spontaneous	5 = Oriented	6 = Obeys commands
3 = To verbal stimuli	4 = Confused	5 = Localizes pain
2 = To pain	3 = Inappropriate words	4 = Withdraws from pain
1 = None	2 = Incoherent	3 = Flexion to pain or decorticate
	1 = None	2 = Extension to pain or decerebrate
		1 = None

Figure 1.10 Glasgow Coma Scale performs three tests, eye, verbal, and motor responses. The three separate values and their sum are considered for a GCS score ranging from 3 to 15, 3 indicates deep coma, and 15 is for a fully awake person.

duration of the LOC, post-traumatic amnesia (PTA), and the GCS score [56]. GCS is a neurological scale that aims to record the conscious state of a person reliably and objectively. GCS is between 3 (deep unconsciousness), and 15 (fully awake). This scale was first proposed by Graham Teasdale and Bryan Jennett [57]. The GCS is a result of three components, eyes, verbal and motor responses. The scoring table is shown in Fig. 1.10, the three separated scores and their sum are factored in to compute a final score ranging from 3 to 15 indicating the consciousness of the person.

1.5 Data Specification

In this work, we use data for the following three groups of subjects,

1. *Healthy Controls:* subjects with a healthy brain. They are used in this study

as a control group to help identify possible abnormalities with the other two groups. There are 30 subjects in this set.

2. *MTBI-A:* subjects with post-concussion symptoms after six month of the incidents, They are diagnosed with MTBI and low GCS scores in their acute phase. The post-concussion symptoms are self reported. There are 24 subjects in this set.

3. *MTBI-B:* subjects without post-concussion symptoms after six month of the incident. They have been scored low on GCS test. This decision is based on self-reported symptoms. There are 23 subjects in this category.

The subjects and dataset are carefully assembled such that their demographics (age, education, and gender) are matched between the three groups, otherwise functional Connectivity measurements are possible to be biased by these factors. Table 3.1 shows the size of each group, and their GCS scores, age, gender and education level.

Table 1.1 *Demographics Information for the Three Groups*

TYPE	SIZE	GCS	AGE	GENDER	EDUCATION
HEALTHY	30	NA	38.1 ± 19.1	NA	15.3 ± 2.3
MTBI_B	23	9	46.1 ± 16	NA	14.6 ± 3.3
MTBI_A	24	8	50.8 ± 19.2	NA	14.3 ± 3.1

Chapter 2: Methodology

In this chapter, we will discuss in detail each stage of our analysis pipeline, and justify the reasons behind our methodology.

2.1 Masking

2.1.1 Introduction

Masking is a convenient way for indexing an array. A mask M is an 3D array of the same size as the matrix X to be indexed such that each entry is either 0 or 1. Masking has the following syntax: $Y = X[M]$ where $Y_{ijk} = X_{ijk}$ if $M_{ijk} = 1$, and $Y_{ijk} = 0$ if $M_{ijk} = 0$. The masking matrix in our work is a binary {0,1} matrix with three spatial dimensions: $(54, 64, 50)$, which match the standard template (talairach template- TT_N27 template in AFNI) onto which we register each individual functional image.

Masking is commonly used as a pre-processing step prior to any statistical analysis because of the following reasons:

1. Masking serves as a regularization filter that removes noisy signals that have too high or too low magnitudes compared to the average. The removal of the

two ends of the spectrum often improves the quality of the data by reducing its skewness. The reduction in skewness assists in the stabilization and convergence in the statistical analysis which to be introduced later.

2. By focusing on signals in the middle of the spectrum which are more likely to be relevant, the amount of data that needs to be analyzed is reduced significantly. In our case, the number of voxels of the brain template is $(54 * 64 * 50)$. The remaining number of voxels after a typical masking is around $150,000$, which is a reduction of 13.2%, thereby saving computational resources and improving the quality of the results

Masking naturally converts the 4D fMRI data to a 2D matrix per subject, where the dimensions of the converted 2D matrix are $(n\_masked\_voxels, n\_timepoints)$. The conversion process is illustrated by Fig. 2.1. The 2D matrix form is more convenient for many data processing algorithms such as PCA and ICA. The mask reserves an one to one mapping between the 2D matrix and the original 4D brain imaging time series. It is saved for converting backward from 2D to 4D which is helpful when we want to draw meaningful interpretations from the results obtained using the 2D matrix.

2.1.2 Mask Computation

Our method of computing a masking is based on a heuristic proposed by T.Nichols [14]. We start by finding the least dense point of the histogram, between fractions lower cutoff and upper cutoff of the total image histogram. The least

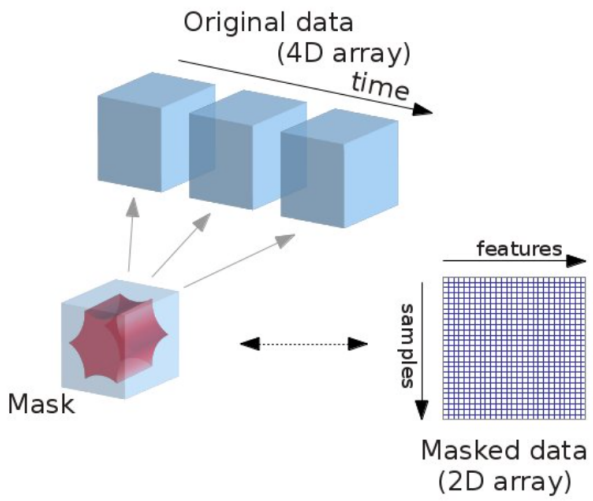


Figure 2.1 Illustration of the Masking Process. The mask is applied to every fMRI scan, and convert the 3D data into a vector. Overall, the 4D data is converted to a 2D matrix with shape $(n\_masked\_voxels, n\_timepoints)$

dense point often serves as an indicator of the gap that separates the signal and the noise of smaller magnitudes. Fig. 2.2 is a voxel activation histogram of a random subject in our dataset. The x-axis is the activation magnitude ranging from -30 to 30, centered at 0. The y-axis is the histogram count of number of voxels that fall into the corresponding bins. The masking algorithm first discards the lower and upper portions of the histogram, then finds the least dense point. Because most noisy sources introduce low-energy and low magnitude noise, the least dense point is often found on the left hand side of the peak. The x-axis value of the least dense point is marked as a threshold. The voxels that lie between the least dense point and the higher cutoff point are kept as signal for analysis later. Note this method allows the masking to be discontinuous in space, (see an example in the z direction cut of Fig. 2.3).

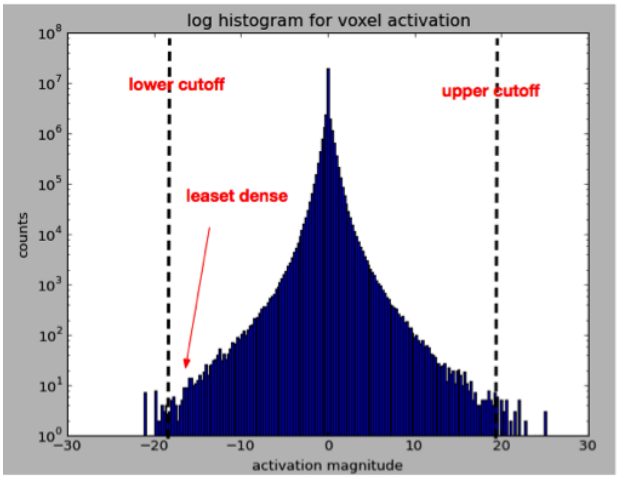


Figure 2.2 Illustration of the Masking Computation: First, the lower and upper fractions of the histogram is excluded, then the least dense part of histogram is located, which is marked as a threshold. Any voxel with magnitude larger than the threshold is saved.

In our project, the lower and upper cutoffs are set to be 20% and 85% respectively. To ensure that the masked fMRI data has the same number of voxels, each subject is processed with a common mask. This group-level mask is computed as the common intersection of each individual mask. Then common group mask is applied to each subject individually. An example of such a group mask is shown in Fig. 2.3, which illustrates the shape and areas covered by the common mask. Brain areas known without tissue are eliminated by the masking process because the signal measured in those areas are low magnitude noise.

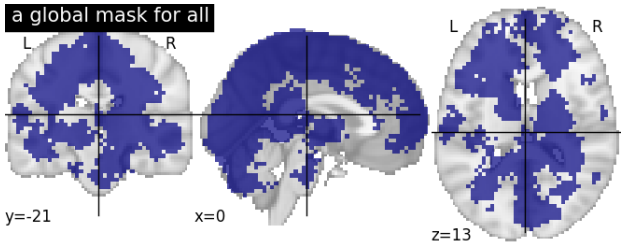


Figure 2.3 A group mask is computed as the common intersection of all individual mask. Three cross sectional x,y, and z cuts for simple illustration. The dark blue part is what will be kept, and the gray areas show what will be discarded.

2.2 Group Independent Components Analysis

2.2.1 Introduction to Independent Components Analysis

ICA decomposes a multivariate observation matrix into multiple independent non-Gaussian components. An easy example of the use of this technique is to solve the cocktail party problem which is the phenomenon of being able to focus on one conservation while filtering out all the other talking as noise during a cocktail party [13].

During a party, you constantly hear a mixture of conversations. yet it is rarely a problem for you to get confused about the speaker you are listening to. The content, location, speaker voice constitute a signature of speech, which makes each speech 'independent' from each other. The goal of ICA is similar to this blind data source separation performed automatically by our brain.

Two fundamental assumptions of ICA are:

1. The sources are independent to each other, just like each speech in a party is

independent to each other.

2. Each source follows a non-Gaussian distribution.

The justifications of these assumptions will be discussed later.

2.2.1.1 ICA's Application to fMRI

ICA has been successfully utilized in many fMRI applications. Spatial ICA parcellates brain into non-overlapping regions under the criterion of promoting statistically independence, which will be introduced later. ICA can also be used to find temporally independent components; however, most applications of ICA to fMRI use spatial ICA to parcellate brain components that are maximally independent in space. McKeown et al. claimed that the spatial ICA works well for most fMRI applications due to the fact that the spatial patterns of typical cognitive activation paradigms are distributed sparsely. Moreover, the nature of many noise sources are also sparse and localized, such as:

1. Vascular pulsation (signal localized to larger veins that are moving as a result of cardiac pulsation)

2. Breathing induced motion (signal localized to strong tissue contrast near discontinuities: tissue edges)

ICA algorithms with a sparse prior would be well suited for spatial analysis.

Each 2D masked matrix forms a natural observation matrix $X_{N\times n}$, where N is the number of time points, and n is the number of masked voxels. The process

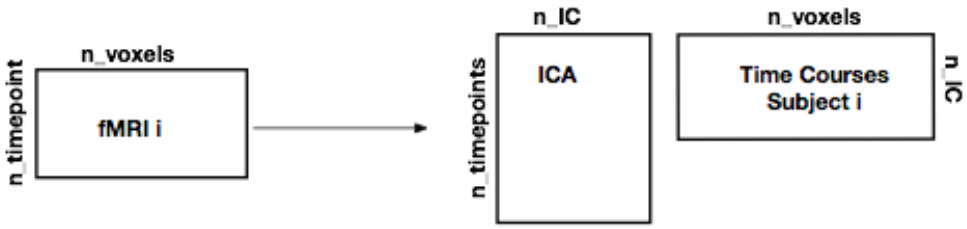


Figure 2.4 Illustration of Individual ICA. The observation matrix is decomposed into time course matrix and ICA spatial map matrix

of ICA is illustrated in Fig. 2.4. At a high level, ICA factors the left hand side matrix X into the right hand side product of time course matrix and the IC matrix, where each row of the IC matrix is an IC that corresponds to a set of regions in the brain. The corresponding column of time course matrix is the temporal activation pattern of that IC during the scan. The number of ICs needs to be carefully selected. Common approaches for estimating the number of ICs include various information theoretic approaches such as Akaike Information Criterion (AIC).

2.2.1.2 A Brief Overview of ICA Theory

In mathematical settings, the observation matrix X is formed by an array of vectors $X = (\overline{X_1}, \ldots, \overline{X_n})^T$,

where n is the number of observations, which is often assumed to be equal to the number of source patterns; the cases of under-determined and over-determined ICA are discussed in [32]. Each $\overline{X_i}$ is an realization of a random vector $\overline{X_i} \sim (x_1, \ldots, x_t)^T$, where t is the number of features (which is equal to the volume of

the mask in our case), and each random variable x_i is a mixture of n components. Overall, X is a matrix of n observations and t features. Each observation is a mixture of a set of independent components $s_k = (s_{k1}, \ldots, s_{kt}),\ k = 1, 2, \ldots, n$, each with length t:

$$x_i = \sum_{k=1}^{n} s_k a_{ik} \tag{2.1}$$

where $a_i = (a_{i1}, \ldots, a_{in})$ are the mixture weights. The vector expression above has a very concise matrix form:

$$X = AS \tag{2.2}$$

A possible justification of ICA lies in the central limit theorem, which states, the distribution of a sum of independent, identically distributed random variable converge to Gaussian distribution. In the classical version of CLT, if the x_i is independent, identically distributed random variables with expected value μ and finite variance σ^2, then $\frac{1}{N}\sum_i^N x_i \to N(\mu, \frac{\sigma^2}{N})$ as $N \to \infty$. The ICA can be understood as the reversed process of CLT. Assuming the observed brain patterns are mixtures of non-Gaussian sources, $X = AS$ where matrix A is the mixing matrix of shape $t * t$, S are the true source patterns. The goal of the ICA is to unmix observation matrix X by seeking an unmixing matrix W such that $Y = WX = WAS \approx S$ where Y is an approximation of the true source patterns, W is the unmixing matrix that converts observation back to independent sources. On the other hand, because $Y = WAS$ can be interpreted as Y is a linear transformation of S the true sources. By CLT, Y can only be more or equally Gaussian as S. By maximizing the non-Gaussianity of Y, ICA finds an independent source pattern that explains the observation X.

Note this argument is a overly simplified justification of the ICA theory, for more in-depth discussion, see [32].

2.2.1.3 Independence

The probabilistic definition of independence is the following: Two r.v.s X and Y with CDF $F_x(x)$ and $F_y(y)$ and PDF $f_x(x)$ and $f_y(y)$ are independent if and only if the combined random variable (X, Y) has a joint cumulative distribution,

$$F_{X,Y}(x, y) = F_x(x)F_y(y) \tag{2.3}$$

or equivalently,

$$f_{X,Y}(x, y) = f_x(x)f_y(y) \tag{2.4}$$

This definition is rarely used in practice, because it requires the true distributions of X, Y, and (X, Y) to be known . In reality, those distributions can only be estimated approximately if given enough observations. Widely used practical estimations of independence is the KL divergence rooted in information theory,

$$KL(X||Y) = \int f_X(x) \log \frac{f_X(x)}{f_Y(x)} dx \tag{2.5}$$

$$= H(X, Y) - H(X) \tag{2.6}$$

where $H(\bullet)$ is the entropy function of a distribution.

Another popular information theoretic measure of independence is the mutual information between X and Y:

$$I(X; Y) = \int_X \int_Y f_{X,Y}(x, y) \log \left(\frac{f_{X,Y}(x, y)}{f_X(x)f_Y(y)} \right) \tag{2.7}$$

2.2.1.4 Measure of Non-Gaussianity and Pre-processing Steps

The non-Gaussianity can be measured by the 4^{th} order cumulant - kurtosis which captures the degree of peakness of the underlying distribution, and is commonly used as a distance away from normality.

$$\kappa_4(Z) = E[Z^4] - 3E[Z^2]^2$$

where the second term is zero if the data X is centered and whitened in the pre-processing. This premise and the fact that typical ICA algorithms are iterative explains why it is a common practice to pre-process the data matrix X. Essential pre-processing steps include:

1. Center the data, by subtracting the mean from the signal. This step doesn't change the mixing matrix A or the unmixing matrix W because $Y - E[Y] = W(X - E[X]) = WA(S - E[S])$.

2. Whiten the data. Because the sources are independent from each other, $E[SS^T] = I_n$, where n is the number of independent sources and the rank of matrix SS^T. Whitening X converts the covariance matrix to an identity matrix of rank n. This is done by the SVD technique. SVD decomposes the covariance matrix into products of orthogonal matrix U and eigenvalue diagonal matrix D. $\Sigma = E[XX^T] = UDU^T$. The whitened data matrix $\hat{X}$ is $\hat{X} = D^{-\frac{1}{2}}U^T$ To obtain the desirable whitening property $E[\hat{X}\hat{X}^T] = (D^{-\frac{1}{2}}U^T)^T UDU^T(UD^{-\frac{1}{2}}) = I_n$.

2.2.1.5 ICA as an Optimization Problem

The first stage of ICA is often to center and whiten the data, which is a linear transformation such that the new zero-mean components are uncorrelated and have variance one. The whitening is not a constraint of ICA problem, but whitening step forces the mixing matrix to be orthogonal, which is often critical, as stated in section 5.2 of [32]

> Here we see that whitening reduces the number of parameters to be estimated. Instead of having to estimate the n^2 parameters that are the elements of the original matrix A , we only need to estimate the new, orthogonal mixing matrix $\tilde{A}$. An orthogonal matrix contains $\frac{n(n-1)}{2}$ degrees of freedom. For example, in two dimensions, an orthogonal transformation is determined by a single angle parameter. In larger dimensions, an orthogonal matrix contains only about half of the number of parameters of an arbitrary matrix. Thus one can say that whitening simplifies the problem of ICA significantly. Because whitening is a very simple and standard procedure, much simpler than any ICA algorithms, it is a good idea to reduce the complexity of the problem this way [32].

It is easy to verify that the recovered independent sources are ambiguous to their sign, scale, and permutation of orders. To find an optimal unmixing matrix W, the ICA can be formulated as an optimization problem: promote non-Gaussianity

as an objective function of WX.

$$\text{Observations: } X = AS$$

$$\text{Optimize: } f(WX)$$

$$\text{Subject to: } W \text{ orthogonal}$$

$$\text{and } s_i\text{'s are independent}$$

$$\text{Results: } Y = WX$$

where $f(\bullet)$ is a function that tests statistical independence empirically. It can be proven mathematically that if the $f(\bullet)$ function is chosen to be mutual information, this objective function can be derived to be,

$$\begin{aligned} f(WX) &= f(Y) \\ &= f(y_1, \ldots, y_n) \\ &= H(y_1) + \ldots + H(y_M) \end{aligned}$$

where y_i represents the random variables of matrix Y. Now the question transforms to maximizing the sum of entropy of the column variables of matrix Y. It is known that the distribution that yields the smallest entropy is Gaussian, therefore to maximize the objective function, we want each column distribution to move away from normal distribution as much as possible.

2.2.2 Group Level ICA

When ICA was first introduced by [33] for fMRI analysis in 1998, it had mainly been applied to each individual fMRI dataset. In this way, each individual receives

its own time courses that correspond to their own set of ICs. By manually choosing and comparing the set of ICs, the value of ICA was shown in its ability to generate brain spatial parcellations that correspond to known brain networks [129].

However, because each subject was characterized by its own set of time courses and ICs, it is not clear how to extract common patterns or differences between subjects in a certain population. Being able to compare across subjects is highly valuable when the main goal is to infer features that distinguish between two groups of subjects. However, ICA is by its nature a uni-subject analysis method, which cannot be easily extended to multi-variate analysis. More research needs to be conducted to understand how data from multi-subject can be considered as input of an ICA process.

To address the problem, several useful group level ICAs have been proposed and applied to real-world fMRI data with promising conclusions. Justifications were given to each type of group level ICA. Popular choices of multi-subject ICA are:

1. Temporal concatenation fMRI, where the fMRI data are concatenated along the time axis. This approach assumes that the spatial patterns that generate the fMRI data are the same across the subjects. The process is illustrated in Fig. 2.5. A single ICA is applied to the combined dataset, which decomposes the data into a set of common IC spatial patterns and time courses that explain the concatenated matrix. The latter cannot be used as inference of the subjects and it is usually discarded. The unique time courses for each subject can be derived using regression techniques.

2. Spatial concatenation fMRI, where the fMRI data are concatenated along the spatial axis. A single ICA is applied to the combined dataset, which factors the combined matrix into a commonly shared time courses and a large spatial pattern that correspond to all subjects combined. The latter doesn't have a meaningful interpretation and is often ignored. Spatial map that is specific to each subject is easily obtained by a regression step.

3. Tensor ICA, which, unlike the two methods above, estimates common time courses and common spatial maps simultaneously. The process of tensor ICA is described in Fig. 2.6 .The individual fMRI data are concatenated to a 3D dataset with (time * space * subjects). Tensor-ICA factors it to triplets of time courses, spatial maps and session modes. The session modes contain subject specific parameters that can be compared on the basis of common time courses and spatial maps.

These methods differ in the way they combine the individual data as input. Tensor-ICA and spatial concatenated ICA are suitable when there are common time courses generated consistent to the stimulus paradigm among the subjects. On the other hand, the temporal concatenated ICA assumes that the underlying spatial maps are common for all subjects, which is an reasonable for resting state fMRI study, during which the temporal variations are much larger than spatial variations. The temporally concatenated ICA has been a popular choice for the fMRI data parcellation and has been implemented by various software packages, each with its own goal and scope.

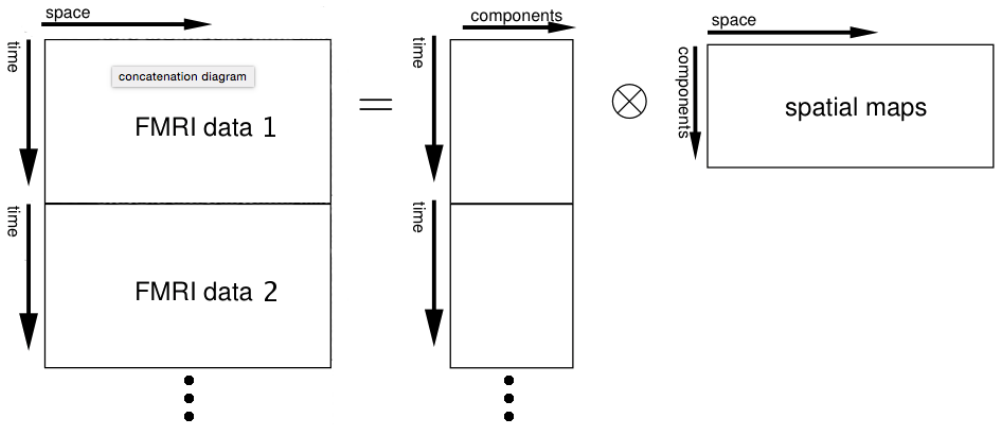


Figure 2.5 The process of temporal ICA. Individual fMRI data are concatenated along the temporal axis, and a single ICA is applied to decompose the combined data to time courses and a common spatial map.

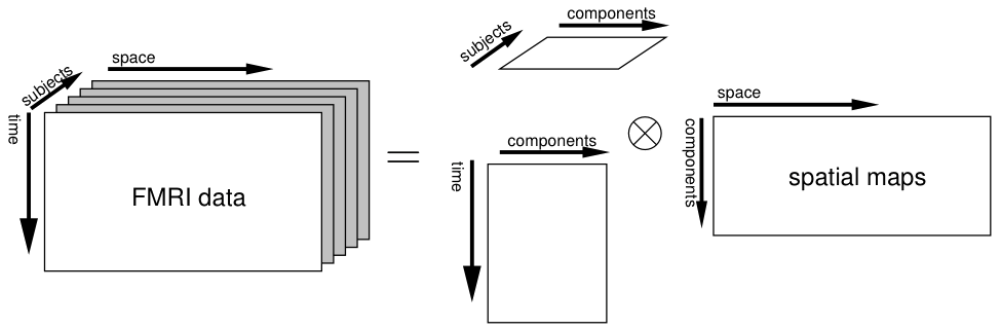


Figure 2.6 The process of tensor ICA. Individual fMRI data are combined to form a 3D dataset with dimensions (time * space * subject). Tensor ICA decompose the input to a triplets, where the session modes contains subject-specific data.

1. MELODIC Matlab software. This software is part of the FSL package, which is a comprehensive library of analysis tools for brain imaging data. The MELODIC package implements ICA by decomposing the 4D datasets into spatial and temporal components. Both temporal concatenation ICA and tensorial ICA are implemented in the package.

2. GIFT Matlab software, which implements multiple algorithms for ICA and blind source separation for fMRI and EGG brain imaging data. FIT is a module of GIFT which implements the jointICA, parallel ICA and CC with jointICA methods. Another useful tool is the FNC package which displays temporal relations amongst components, helping researchers determine functional connections within the brain. In addition to multiple implementations of ICA, such as FastICA and info-max, the GIFT package contains a duel-regression implementation which is used to derive subject-specific time courses from the common spatial maps and each individual fMRI. This back-reconstruction step allows direct and meaningful comparisons of time courses of various components across subjects or even groups. Calhoun [151] showed how spatial ICA followed by duel-regression can capture variations in subject specific spatial maps.

3. Canonical ICA (CANICA) Python package, which is another ICA method for group level fMRI functional analysis. It adopts a hierarchical model to extract common spatial patterns from fMRI dataset. The advantages of this ICA method is the stability and reproducibility of the extracted IC patterns,

and is robust to various noise sources. CANICA estimates the intermediate PCA model order by removing Gaussian like noise using statistical test, whose details will be discussed in the next section. The focus is to locate the subspace with maximal reproducibility and stability for group level principal components. Finally, the common voxel intensity thresholding is based on the value of the absolute intensity.

Each package has proven its value in developing the ICA method for fMRI analysis, Because of the superior reproducibility of CANICA and its natural integration with the rest of our Python written processing pipeline, it is adopted as the ICA decomposition method in this work. Details and theoretical framework of CANICA will be presented in the next section.

2.2.3 CANICA: Canonical ICA

Since the introduction of spatial ICA to fMRI analysis, CANICA has been one of the most successful methods for brain parcellation. In many cases, meaningful spatial patterns are derived naturally from the resting state fMRI data. However, a crucial disadvantage of most ICA methods is the vulnerability to noise and the resulting lack of reproducibility. Noise is ubiquitous in all signal processing and statistical studies. It is particularly an issue for resting state fMRI analysis, because 1) the minimal signal produced by subjects during their resting fMRI scan; 2) various noise sources rooted in the nature of the fMRI technology, ranging from the cosmic noise to subject head motion. fMRI pre-processing mitigates the noise issue to a

certain extent, but the SNR in processed data is still on the low side. Most ICA methods are vulnerable to noise, and extracted IC patterns change drastically when noise is introduced into the input dataset. The common solution to this issue is to pick the patterns that are meaningful to a neurological experts' eye. This artificial process introduces bias sourced from human, and there is no standard way to identify meaningful IC patterns. The bias due to poor reproducibility and human involved filtering process has led to various sets of IC reported in the literature and caused confusion when comparing results of different papers.

Because reproducibility has been a critical issue, Gael et al. developed a robust ICA pipeline called CANICA that is able to separate signal from noise by multiple layers of statistical tests and data processing such as Canonical Correlation Analysis (CCA). As a result, only reproducible IC spatial maps are extracted from the group level fMRI, and those patterns that are likely produced by noise are eliminated. The reproducibility of CANICA can be proven by a cross validation to generate stable group-level ICs using only a subset of participants [183].

CANICA assumes a special hierarchical model that explains the fMRI observations. At the lowest level of the model, all brain activities are expressed by a mixture of a set group level spatial patterns B of shape $(n\_patterns, n\_voxels)$. The spatial patterns P_s of subject s is generated by B with a square mixing matrix Λ_S with additional subject variability matrix R_S. The relation can be expressed mathematically:

$$P_S = \Lambda_S B + R_s$$

On the group level, the subjects' data can be concatenated vertically to obtain an even more concise relationship.

$$P = \Lambda B + R \tag{2.8}$$

where $P = P_s$, $R = R_s$, and $\Lambda = \Lambda_s$, for $s = 1, 2, \ldots, S$, and S is the number of subjects in the group.

In the resting state fMRI framework, there is no external stimulus to the subject, nor is there any task the subject is required to perform. The fMRI observation matrix Y_S for subject s can be modeled simply as a mixture of subject-specific spatial patterns P_S plus noise introduced in the scan denoted as E_S. The relationship is described as,

$$Y_S = W_S P_S + E_S$$

where W_S is another square mixing matrix, and matrices Y_S, W_S, and E_S share the same shape.

The model is illustrated in Fig. 2.7. Note that in the rightmost frame, the group level spatial maps estimated from the fMRI data is not necessarily independent from each other, the independence can only be obtained as a result of the final ICA decomposition.

The ultimate goal of CANICA is to determine the set of independent spatial map components that are common to all subjects, and this is done by multiple stages of estimation and data extractions illustrated in Fig. 2.7. The following sections will be devoted to explain each phase of CANICA.

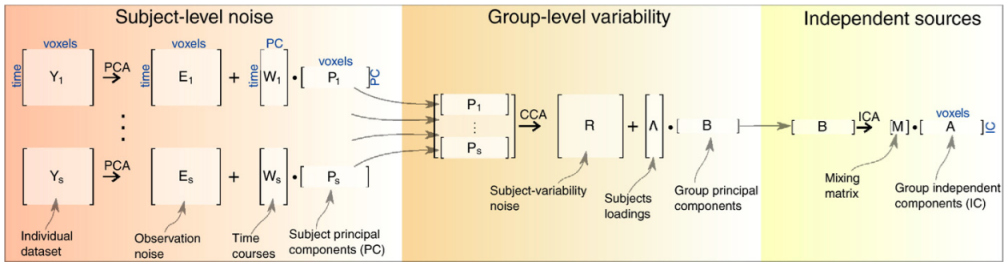


Figure 2.7 CANICA hierarchical models. Group principal components B forms subject-specific principal components, which in turn builds the fMRI BOLD observation for each subject.

2.2.3.1 Individual Level Noise Extraction

The fMRI scan is inherently a noisy process. The technique used to locate the subspace where most energy resides is the Principal Component Analysis (PCA). The residual is considered as noise or irrelevant which contribute little to the overall statistical characteristics of the dataset.

For subject s with a rectangular shape observation matrix Y_S, PCA uses Singular Value Decomposition (SVD) to decompose Y_S into the following: $Y_s = U_s \Sigma_s V_s^T$, where:

- U_S is a orthogonal matrix formed by left-singular vectors.
- V_S is also a orthogonal matrix that is formed by the set of right-singular vectors,
- The middle matrix Σ_S is a rectangular diagonal matrix formed by singular values of Y_S.

The PCA technique picks the top n dimension subspace corresponding to the

top n largest eigenvalues. The approximation of rs-fMRI Y_S in this subspace is computed as signal components $\overline{Y_S}$, and the rest is left as noise.

$$
\begin{aligned}
E_s &= Y_s - (U_S \Sigma_S V_S^T)_{1,\ldots,n} && (2.9)\\
&= Y_s - (U_S \Sigma_S)_{1,\ldots,n} (V_S^T)_{1,\ldots,n} && (2.10)\\
&= Y_s - \hat{W}_S \hat{P}_S && (2.11)\\
\hat{W}_S &= (U_S \Sigma_S)_{1,\ldots,n} && (2.12)\\
\hat{P}_S &= (V_S)_{1,\ldots,n} && (2.13)
\end{aligned}
$$

where $(U_S \Sigma_S V_S^T)_{1,\ldots,n}$ is the product of matrices limited to the top n dimension, those with largest singular values and their corresponding singular vectors.

The choice of the number of dimensions of n is a significant issue because a large n aggressively includes more signal components, which will also unavoidably include some noise components that corrupt the signal. On the other hand, a small n is a conservative approach to get rid of noise at the price of losing certain signals. CANICA uses a statistical test to choose the proper value of n. Conventional methods use a percentage of total energy/variance; however, it can be shown that n is highly dependent on the sample size under this rule, instead of the true structural variation determined by the underlying model. The dimension of the signal subspace is fixed, and hence the conventional rule includes more and more noise as the sample size increases. Research on the optimal order of PCA as a prior step of fMRI ICA model reports that $20-60$ seem to work well depending on the specific dataset [185–187]. CANICA adopts the method proposed in [184], which builds a mathematical framework for analyzing three sources of the PCA model errors together with a

stopping rule that uses the t-test for assessing the stability of the subspace spanned by the first n principal components. The method finds noise dominated modes using a t-test of bootstrap stability between real data and artificial Gaussian noise. Only those modes that are not noise dominated are retained [184]. Due to the computational expense of this order estimation, the algorithm is applied only to one subject, and the same parameters are applied to all subjects' fMRI data [183].

2.2.3.2 Group Level Variability

PCA finds the intrinsic subspace that captures most of the variance in the data. Canonical Correlation Analysis (CCA) is commonly used to find common shared subspace for two different sets of variables, whose projections have the maximum correlation. CCA is naturally adopted by CANICA as the method to identify a stable-component subspace shared by all subjects. Conventional CCA is developed for two datasets, where CANICA concatenated multiple datasets $\hat{P}_s$ as the input of an SVD decomposition,

$$\begin{aligned} \hat{P} &= \hat{P}_{s=1,\ldots,S} \\ \hat{P} &= \Upsilon Z \Theta \end{aligned}$$

where Υ and Θ are unitary matrix formed by singular vectors, and Z is the rectangular diagonal matrix formed by the singular values of $\hat{P}$. According to the CCA theory, The rows of Θ are inter-subject reproducible components, and the diagonal values of matrix Z represent the between-subject reproducibility [183]. The decision has to be made to only retain a top portion n_2 of the singular values just in our

previous PCA step. The canonical weighted forms interesting group level patterns,

$$\hat{B_S} = (\Upsilon)_{1,\ldots,n_2}$$

CANICA uses a thresholding criterion to choose the value of n_2, and the criterion is based on the reproducibility compared to observation noise $\hat{E_S}$ extracted in the previous steps. The maximum canonical correlations of noise can be computed by sampling a bootstrap distribution of $\hat{E_S}$. Only canonical variables with a probability $p < 0.05$ of being generated by noise are selected. Therefore, the retained canonical components are more reproducible than the observation noise.

Inter-subject reproducible components B are given by the vectors of Θ^T whose corresponding diagonal values in Z are each larger than the significance threshold set above. The corresponding vector of ΥZ forms the mixing matrix Λ.

The previous two sections estimate the group-level subspace of interest by first using PCA to determine the maximum stable signal subspace, and then maximizing the reproducibility of the group level components. Both steps utilize SVD and a noise model.

2.2.3.3 Group Level Independent Components

The group-level components B span the subspace of maximum correlation between the subjects. CANICA uses the FASTICA [182] on this subspace to find the group level independent components A, as shown in the last stage of Fig. 2.7. FASTICA is an iterative implementation of ICA algorithm which attempts to determine

the independent components with maximally non-Gaussian marginal distribution.

$$B = MA$$

where M is a loading matrix that controls the contributions of the ICs included in A.

The informative signal of the ICA patterns are hidden in the non-Gaussian tail of the histogram, so it is important to benchmark the Gaussian mode and only keep the rest of the distribution. This is in line with the idea of ICA which is to extract maximally non-Gaussian components. CANICA uses a simple null distribution of Gaussian of unit variance. Similar to the approach in Schwartzman's work [181], the null distribution fits the center of the empirical distribution, and CANICA keeps only the voxels with absolute intensity larger than a threshold that is computed from the null distribution.

2.2.4 IC Illustration

ICA has proven to be quite valuable because the extracted IC patterns often match with our prior knowledge of brain anatomy. See Fig. 2.8 for 42 extracted patterns of ICs by CANICA [183]. Those with a black frame are the ones that could be related to known intrinsic brain networks, such as: pre-frontal, motor network and thalamus.

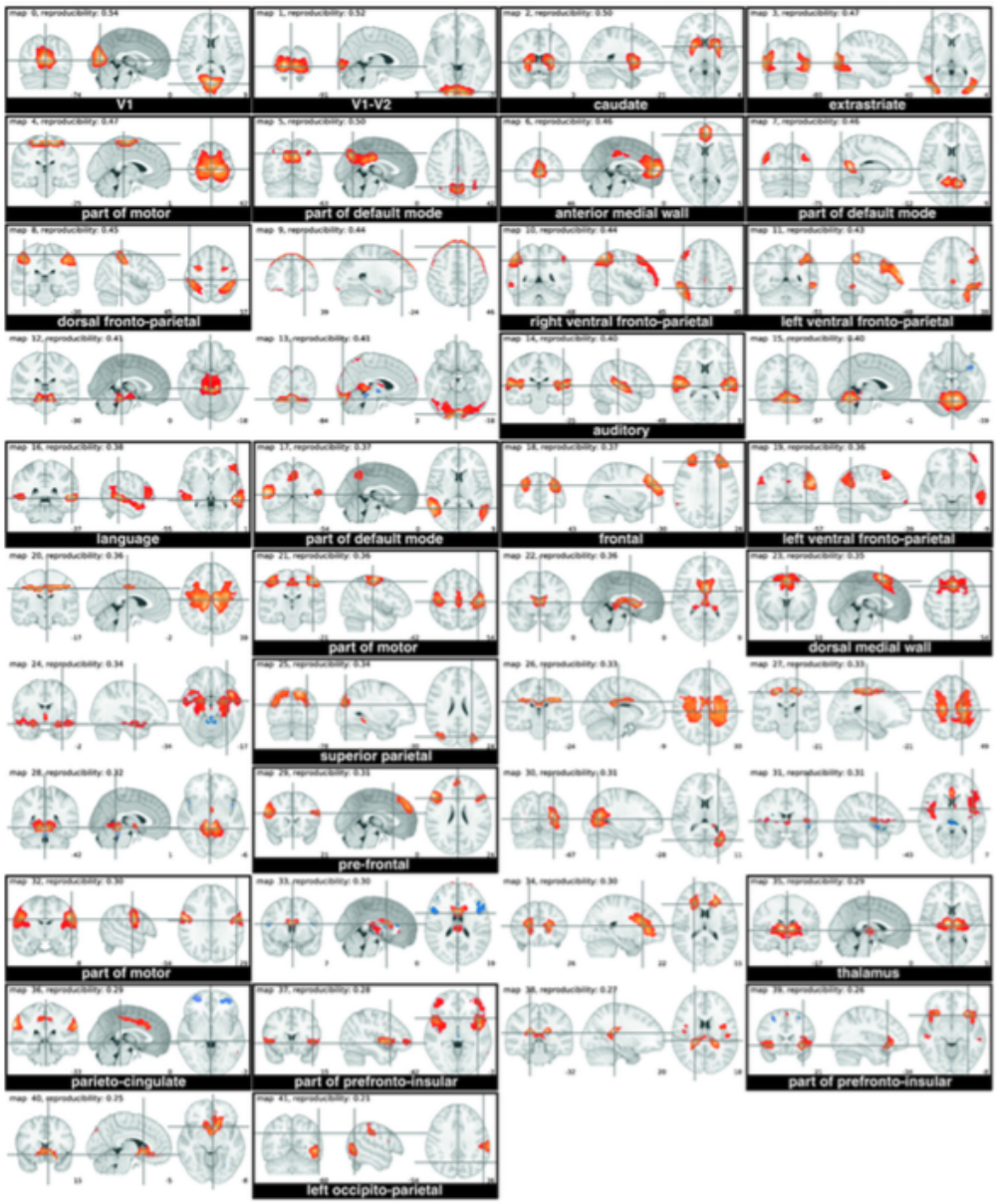


Figure 2.8 Illustration of 42 IC maps extracted by CANICA from resting state dataset. Maps corresponding to anatomically known networks are in a black frame, whereas maps likely corresponding to artifacts are not framed. Extracted brain networks are labeled with the name of the general structure they can be related to.

2.3 Regression

CANICA identifies a set of stable and reproducible IC patterns which form a common basis across the subjects. However, to compare between subject, we need to derive using their individual fMRI data and the CANICA patterns by making use of regression techniques.

Linear regression is used to find the linear relationship between target value $y(w, x)$ and input variables $\bar{x} = (x_1, x_2, \ldots x_p)$ and mixing coefficients $\bar{w} = (w_1, \ldots, w_p)^T$ with order p,

$$y(w, x) = w_0 + w_1 x_1 + \ldots + w_p x_p + \epsilon \tag{2.14}$$

$$= \bar{x} \times \bar{w} + \epsilon \tag{2.15}$$

where ϵ is Gaussian noise with distribution $N(\mu = 0, \sigma = 1)$.

In practice, there are n pairs of observations of $(\bar{x}, y)_{1,\ldots,n}$, where n is usually larger than p for the system to be determined. By stacking n observations of $\bar{x_i}$ and y_i, the system can be written in a matrix form, where

$$Y = Xw + \epsilon \tag{2.16}$$

where

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{pmatrix} \quad X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}_{n \times p} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

The simplest linear regression model is the ordinary least squares (OLS), which

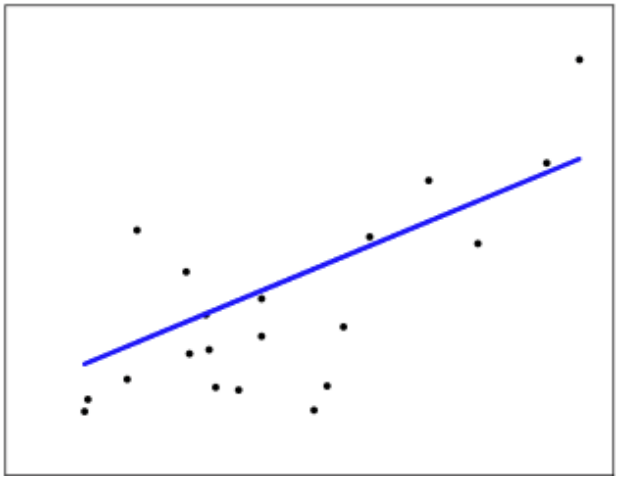


Figure 2.9 The process of OLS can be thought as trying to find a straight line that fits through the input (x-axis) and target (y-axis) with the minimum differences.

tries to fit a linear model in Eq. 2.16 with coefficients w by minimizing the squared difference between the observed target Y and the predicted approximation Xw measured in $L2$. Mathematically, the objective function can be written as

$$\min_{w} ||Xw - Y||_2^2 \tag{2.17}$$

A simple illustration of the process can be understood as fitting a straight line between the target and input variables, when there is only one independent variable x, as shown in Fig. 2.9.

Unsurprisingly, this simple OLS has an explicit solution,

$$\hat{w} = (X^T X)^{-1} X^T Y \hat{Y} = X\hat{w} = X(X^T X)^{-1} X^T Y = PY$$

where $P = X(X^T X)^{-1} X^T$ is the projection matrix onto the space spanned by the columns of X.

OLS works most effectively when the columns (regressors) of the design matrix X are independent of each other. When the regressors are linearly dependent, the values of the weights $\hat{w}$ cannot be computed because the matrix $X^T X$ is not invertible. This independence of regressors assumption is not a problem in the fMRI setting, because the design matrix is formed by the group level spatial patterns extracted by ICA whose goal is to achieve independence.

2.3.1 Ridge Regression

Ridge regression, also called Tikhonov regularization, is a powerful tool that copes well with ill-posed problems. When some assumptions of the OLS are violated and the estimation of regression coefficients w is not achievable, an $L2$ regularization term is included in the objective function,

$$\min_{w} ||Xw - y||_2^2 + \alpha ||w||_2^2$$

where $\alpha > 0$ is a positive regularization parameter that controls the contribution of the $L2$ penalty term, also called shrunkage. The larger the value of α, the greater the degree of shrinkage and the estimated weights $\hat{w}$ become more robust to colinearity of the regressors. The effect of the parameter α is shown in Fig. 2.10, the larger α is, the smaller the $L2$ norm of the coefficients.

Ridge regression provides a more robust method than OLS when the problem is ill-posed. Moreover its complexity is the same as OLS. Given a matrix X of size (n, p), both methods have cost $O(np^2)$ assuming that $n \geq p$ for an over-determined setting.

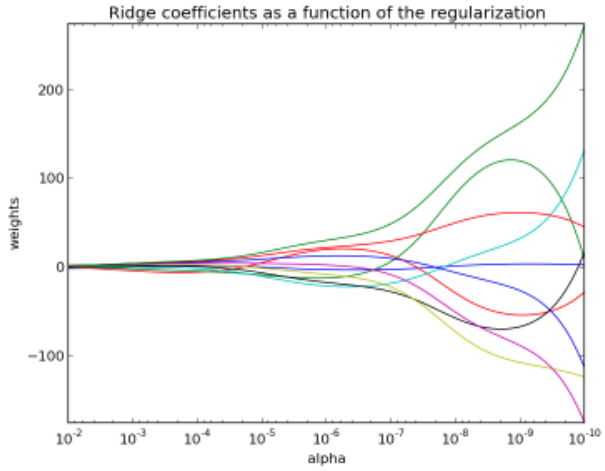


Figure 2.10 The shrunkage parameter α controls the contribution of regularization term in the objective function. The larger the α is, the smaller the output time series is.

The solution to the Ridge regression problem is given explicitly by,

$$w = (X^T X + \alpha^2 I)^{-1} X^T Y$$

which is very similar to that of the OLS, $w = (X^T X)^{-1} X^T y$, and with a ridge down term that agrees with OLS in its limiting form:

- As $\alpha \to 0$, $w^{ridge} \to w^{OLS}$.
- As $\alpha \to \infty$, $w^{ridge} \to 0$.

2.3.1.1 Choosing an Optimal α

Since the shrinkage coefficient α has a large effect on the resulting solution, choosing a suitable value for α is very important. This value is often determined by an *ad hoc* approach for practical problems. The most obvious solution is to

try a number of different values of α and pick the one that minimizes the mean square error MSE, because the goal of the regression is to approximate the target variables as close as possible. The choice of α is part of a bigger problem called model selection.

Cross validation is a useful technique to test the performance of a model. The idea is to split the labeled dataset into multiple subsets based on certain criterion. The model is trained using part of the subsets only, and the trained model is scored using the rest of the subsets. The parameter that leads to the best score is regarded as the winner.

It is common to repeat multiple runs of the training and testing cycle and take the average of the scores produced by a model for a more robust measure. Each run, the model is trained by randomly generated training sets. The most common way of splitting the data is called k-fold Cross Validation (CV). The dataset X and the label y are splits to K equal size subsets $(X_1, X_2, \ldots, X_K)$ and $(y_1, y_2, \ldots, y_K)$, where K is commonly chosen ranging from 4 to 10. Denote the approximation model that is trained with parameter θ and using $K-1$ subsets excluding X_i as $f(z, i, \theta)$. The error for the i^{th} fold is computed as the approximation error for y_i,

$$\epsilon(i, \theta) = \frac{1}{|X_i|} \sum_{j=1,\ldots,|X_i|} (y_i(j) - f(X_i(j), i, \theta))^2$$

where $X_i(j)$ and $y_i(j)$ are the j^{th} elements in training X_i and testing set y_i, respectively.

We iterate over all possible $i = 1, 2, \ldots, K$, and compute the approximation errors $\epsilon(i, \theta)$, the averaged error is the score for the model trained with parameter

θ,

$$\epsilon(\theta) = \frac{1}{K} \sum_{i=1,\ldots,k} \epsilon(i, \theta)$$

the optimal parameter θ^* is the one that minimizes the K-fold CV,

$$\theta^* = \arg\min_{\theta} \epsilon(\theta)$$

A K-fold CV can be computationally expensive in computation. If the number of possible parameters is M, K-fold CV trains MK models to find the best parameters, and the situation is severe if the complexity of the model training is high. Moreover, in many cases, the size of the dataset is not large enough to be split into K decent datasets. Under these circumstances, Leave One Out (LOO) is often preferred. LOO is a special case of K-fold CV when $K = 1$. Instead of splitting the entire dataset, and train the model multiple times, LOO provides a convenient solution for the square error incurred by parameter θ,

$$\begin{aligned} \epsilon(\theta) &= \frac{1}{|X|} \sum_{i=1,\ldots,n} (\frac{y_i - (Xw)_i}{1 - P_{ii}})^2 \\ P &= X(X^T X + \alpha^2 I)^{-1} X^T \end{aligned}$$

where P is the projection matrix which is similar to that of OLS. This result indicates that instead of training multiple models with different portions of the dataset, LOO provides a way to train the model once over the entire dataset and the error as θ is given by the equation above. Replacing the P_{ii} by the average of diagonal elements, $\frac{tr(P)}{n}$, [180],

$$GCV(\theta) = \frac{1}{n} \frac{||Y - Xw||^2}{(1 - \frac{tr(P)}{n})^2}$$

which leads to further simplification that boosts efficiency. The Generalized Cross Validation (GCV) is the CV method that we use for model selection for the optimal α shrinkage coefficient, it is conveniently implemented by Python sklearn package.

In the fMRI setting, the above Ridge regression together with the GCV method are applied to each individual fMRI dataset, with the fMRI dataset as the observation matrix Y and group level spatial maps as matrix X, the regressor coefficient matrix w of shape $(n\_timepoints, n\_voxels)$ estimated from this process is viewed as the time course specific to that subject. Each column of this time course corresponds to the temporal activation of a spatial map for that subject. Information extracted from the correlations of these time courses is used to describe the states of the brain, and the information helps us distinguish between subjects belonging to different groups.

2.4 Network Construction

Viewing each spatial pattern as a node in the brain functional network, the distance between their corresponding time courses as the edge weights, functional networks are built for each subject. It is equivalent of computing the distance matrix between each pair of rows of subject-specific time courses and convert this adjacency matrix to a network.

Conventional fMRI functional analysis implicitly assumes that the relationship between different regions in the brain are static. Therefore their relationship is extracted by measuring the distance/similarity between the entire time course.

However, recent studies have shown that the such relationships are ever changing, certain relationships that were impossible to discover under the static setting were found in the dynamic setting. Indeed, when studying the dynamics of the functional relationships of the brain, many interesting phenomena start to emerge in our research.

The most popular and well established dynamic functional analysis tool is the sliding window method. The relationship between each pair of time series are exploited only within a window that is controlled by its shape, size, and the overlap between the windows are determined by its step size. As the window moves along the time axis, each frame constitutes a network limited by that short period of time, and dynamic functional analysis focuses on the transition and variability of the series of windowed networks.

For each subject, within one sliding window, there are n_{IC} number of time courses. The relations between those time courses within one window represents the relationship of the nodes in the functional networks. Multiple distance or similarity measures can be adopted to explore this distance relations. In this work, we prefer to use authentic metrics because it makes more sense later when we compute distance based network properties such as shortest path, or MST. $L1$ metric is used for distance computation in this work.

More specifically, the time course for each subject can be expressed as a 2D matrix $T_{n\_IC*n\_features}$. This matrix is sliced along the feature axis with a rectangular window of size $window\_size = 35$ and $window\_step = 5$. Within the k^{th} slice, the truncated time course is denoted as T_k with shape $(n\_IC * window\_size)$. We

construct a distance matrix $D_{n\_IC*n\_IC}$ where $D_{i,j} = f(T_{ki}, T_{kj})$ where $f(\bullet)$ is the chosen $L1$ norm distance, $f(a,b) = \sum_i |a_i - b_i|$. Finally, for subject s has series of $L1$ distance matrix $D_{sk}, k = 1, \ldots, n\_windows$, where $n\_windows$ is the number of windows, determined by the size and step of the window parameters. We construct a network $G_{s,k}$ that corresponds for subject s and its k^{th} distance matrix, by treating the $D_{s,k}$ as the adjacent matrix.

For subject s, we map a function onto its series of networks, and get a series of scores $\bar{z} = (h(G_{s,1}), h(G_{s,2}), \ldots, h(G_{s,n\_windows}))$. One of the interesting network properties we measured is the average shortest path, which measures the average shortest path of all possible pair of nodes of the graph, namely,

$$a = \sum_{(s,t)\in V} \frac{d(s,t)}{n(n-1)}$$

where V is the set of nodes of the graph, and the function $d(s,t)$ is the shortest path between nodes s and t, computed by well know shortest path algorithm Dijkstra or Floyd-Wallshall. In fMRI setting, when a subject has a smaller average shortest path score for their functional network, it implies the brain regions are more synchronized with each other, which in turn implies a healthy and efficient brain state based on the topological meaning of shortest path.

Due to the nature of the never resting brain and various noise sources, the $L1$ distance matrix calculated is never zero. Due to the lack of reasonable benchmarking research and techniques that are available, we focus on weighted graph algorithms that can consider a fully connected graph with various weights. Fortunately, such algorithms cover a large set of graph characteristics, such as shortest

path algorithm, clustering coefficient algorithm and minimum spanning tree (MST).

In addition to the average shortest path algorithm, another network property we study is the MST, which can be viewed as the backbone of the graph. The MST can be computed very efficiently via Prim's algorithm or Kruskal's algorithm.

2.5 Statistical Tests

For each subject, we conduct a series of network properties. We'd like to compare across subjects using uni-variate statistical tests, meaning the serial scores for each subject need to be reduced to a single value, in addition to the implicit assumption that the scores for each subject obeys the same underlying distribution. It is difficult to apply multivariate statistical test in our case, because the misalignment of the scan time series.

Multiple statistical tests are carried out to evaluate group difference, especially the difference between the two MTBI types and the HC group. In terms of reducing the series of scores into one single summarizing number, we proposed two different methods that focus on different aspect of the dynamics of the functional networks.

1. The easiest method is to take the mean or median of the series, for subject s who has a series of SP scores: $(z_{s,1}, \ldots, z_{s,n_{windows}})$, we compute the mean $z_s = \frac{1}{n_{windows}} \sum_{i=1,\ldots,n_{windows}} z_{s,i}$. This method is essentially equivalent to static functional analysis, because the averaging operation along the temporal axis eliminates the dynamics captured by the sliding window method.

2. One reduction method that leads to meaningful group difference discovery is to

compute the average difference of adjacent pair of scores. For example, subject s has a series of average SP scores z_{si} $i = 1, \ldots, n\_windows$, we compute for this subject a single score: $\overline{z} = \frac{1}{n\_windows-1} \sum_{i=1}^{n\_windows-1} |z_{s,i} - z_{s,i+1}|$. This score captures the rate of change for that subject under the network measures.

These two proposed methods reduce the serial scores $(z_{s,1}, \ldots, z_{s,n_{windows}})$ for subject s into one single value z_s. Then we perform three different tests: rank-sums, t-test, and F oneway/ANOVA to test if group difference exist among the groups.

ANOVA, t-test and rank-sums are all well-known statistical tests that compare the mean between two or more groups that account for their variability. Extensive research can be found for each of them. Here we provided a brief introduction to each of them.

2.5.1 Welch's t-test

t-test estimates the probability p-value of the likelihood of a null hypothesis H_0. The null hypothesis H_0 in t-test assumes that the two groups obey distributions that have the same mean. The t-test computes a number called the t-statistics, which is related to the degree of the difference between the means of two sets of samples that are being compared. The larger the t is, the larger the difference. However, the t-statistics is not the most useful result, because this number does not tell us the degree of difference. This is why we also compute the p-value. The p-values provide an answer to the important question: once a $T = t_0$ is computed, if the null hypothesis H_0 is indeed true, what is the probability of getting a T-statistics

number at least as large as this? That is $Pr\{T \geq t_0|H_0\}$. In this sense, the smaller the p-value is, the less likely that we would find a difference similar to our dataset by chance, this means the dataset is very rare under the assumption H_0, and hence it is reasonable to argue that there is a difference between the two sets of samples when small p-values are computed.

The computation of the p-value is usually done by extrapolating pvalues from values on the discrete pvalue table once the T-statistics and the null hypothesis are determined, because the actual computation involves estimating the cumulative density function (CDF) of null distribution, which can be computationally expensive.

Multiple versions of t-tests have been developed for different purposes. The Welch's t-test is an adaptation of Student's t-test, and is intended for use for the case where the two sets have possible unequal variances. This t-test is adopted in our work because it is not clear if MTBI will have any unpredictable effects on the variance of our network properties. More specifically: Welch's t-test fits the characteristics of our datasets in the following sense:

1. Each group contains different numbers of subject, ranging from 23 to 30. A number of tests rely on the sample size being equal. An easy solution is to repeat multiple rounds of the test, each with equal sample size randomly drawn from each group, and finally take the mean of the pvalues.

2. Groups are likely to have different variances. We cannot assume that the MTBI subjects exhibit the same variability as the HC group. However, this type of test can still be performed to data that might violate this requirement.

In this case, the power of the test is mitigated. The group difference found in this case truly shows there is a group difference. If a more powerful and proper test were performed on the same dataset, an even smaller pvalue would be found.

After examining the characteristics of our dataset, Welch's t-test seems to be the most suitable t-test, because the two group variances cannot be assumed to be equal, and the group sizes may or may not vary.

The T-statistic to test whether the groups means are different is calculated as:

$$\begin{aligned} t &= \frac{\overline{X_1} - \overline{X_2}}{s_{\overline{X_1} - \overline{X_2}}} \\ s_{\overline{X_1} - \overline{X_2}} &= \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \end{aligned}$$

Here s^2 is the unbiased estimator of the variance of the each sample, n_i is the group size. And the pvalue can be computed by extrapolating from the standard discrete pvalue table.

2.5.2 One-way ANOVA

Similar to the other two tests, one-way ANOVA tests the null hypothesis that samples in two or more groups are drawn from populations with the same mean values accounting for their variances.

ANOVA is a omnibus test, meaning that a single test can tell whether there is a difference among the groups, instead of performing multiple tests for each possible pairs of groups. The advantages of this type of omnibus test is no group comparisons

need to be specified, the difference is automatically found if there is one. However, the downside is that if the null hypothesis is rejected, it is not clear which group or groups exhibit the difference.

The intuition of ANOVA lies in the CLT, for datasets $(Y_1, Y_2, \ldots, Y_K)$, where dataset Y_i consists of a number of samples, $Y_i = (Y_{i,1}, \ldots, Y_{i,n_i})$ and K is the number of datasets, and N denotes the total sample size. If the null hypothesis H_0 is true, meaning if the groups means are equal to each other, $\bar{Y_i} = \bar{Y_j} = \bar{Y},\ \forall i,\ j$, where $\bar{Y} = \frac{1}{N}\sum_{\forall i,\ j} Y_{i,j}$ is the mean of the entire datasets combined, the variances of the between group means μ_B should be smaller than the variance of the within group means μ_W,

$$\begin{aligned} \mu_B &= \frac{1}{K-1}\sum_i n_i(\bar{Y_i} - \bar{Y})^2 \\ \mu_W &= \frac{1}{N-K}\sum_{\forall i,\ j}(Y_{i,j} - \bar{Y_i})^2 \\ \mu_B &\leq \mu_W \end{aligned}$$

The F-statistic is computed as:

$$F = \frac{\mu_B}{\mu_W}$$

This F-statistic assumes $K-1$ and $N-K$ degrees of freedom under the null hypothesis. Under this definition and the implication of the CLT, a large F-statistic would imply that the samples are drawn from different distributions.

The assumptions of ANOVA are:

1. Sample variable residues are Gaussian distributed.

2. Samples are independent.

3. Variances of the groups are equal.

4. The samples for a given group are iid normal distribution random variables.

2.6 Analysis Pipeline

We now summarize our analysis pipeline. First a common mask is computed with threshold level of 0.3 and the mask regions do not need to be connected, the upper_cutoff is 85% and the lower_cutoff is 20%. with an opening parameter of 2 to ensure possible skull regions are excluded. A common mask is computed and applied to each individual RSFRMI data. CANICA is then applied to the combination of all fMRI data, with a threshold level of 3, randomly start for 10 times to ensure good results, and the FWHM Smoothing kernel of 6mm. The number of ICs is chosen to be 40. fMRI data are centered and normalized such that the standard deviation (STD) of every voxel time series is 1. We then apply the Ridge regression to each fMRI individually fitting for the best shrunkage parameter α ranging from 0.1 to 5. The dynamics functional analysis adopts a window size of 30 and a step length of 5, the window shape is rectangular. The function that is used to compute distance between the internal time series of a subject is the L1 norm. A weighted undirected functional network is constructed for each time window using the distance matrix as the adjacency matrix. Graph algorithms such as average shortest path, clustering coefficient, and MST, are applied to every network. The serial score of each subject is summarized into a single value using the two methods described before, focusing on static and dynamics aspect, respectively. Finally, one

way ANOVA tests the existence of a group differences between the three groups of subjects, and the t-test examines the detailed pairwise group difference. We also study the static and dynamic functional connectivity within the sub-networks such as DMN and Sensorimotor network to gain information about the impact of MTBI on brain functional connections and what causes the prolonged recovery of the MTBI-A type. Chapter 3 will provide the overall results, and Chapter 4 will be devoted to discuss possible explanations, the importance of our findings, and potential future research directions.

f

Chapter 3: Results

3.1 Data Acquisition

The rs-fMRI data used in this study were collected within the first 24 hours of the subjects' brain injury. All imaging was performed on a Siemens Tim-Trio 3T MRI scanner using a 12 channel receive only head coil. A high resolution T1-weighted-MPRAGE (TE = 3.44 ms, TR = 2250ms, TI = 900ms, flip angle = 9, resolution = 256 256 96, FOV = 22 cm, sl. Thick. = 1.5 mm) was acquired for anatomic reference with slices parallel to the anterior and posterior commissure points (AC-PC). For the rs-fMRI scan, T2*-weighted images were acquired using a single-shot EPI sequence (TE = 30 ms, TR =2000 ms, FOV = 230 mm, resolution = 64 64) with 36 axial slices (sl. thick. = 4 mm) over 5 min 42 s that yielded 171 volumes. During the resting state scans, extraneous auditory and visual stimuli were removed, and the participants were instructed to rest peacefully with eyes closed.

The participants consist of 30 healthy controls and 23 MTBI patients without self-reported post-concussion symptoms 6 months after the incident, and 24 patients with self-reported post-concussion symptoms 6 months after the injury. Participants are carefully assigned to match the between-group demographics, such as age, gender, and educational level. These factors are considered because they may have an

impact on the functional connectivity measurements, and thereby skew the results. Table 3.1 shows the size, basic demographic information, and the GCS scores for each group. Informed consent was obtained from each participant prior to scanning. All participants were over the age of 18. Patients were screened and excluded for history of neurological and psychiatric illness, history of stroke, history of brain tumors or seizures, and contraindications to MR. Only those patients with an admission GCS of 13-15 and mechanism of injury consistent with trauma were included in this study.

The main goal of this research is to identify the differences between two types of MTBI and the HC using both static and dynamic functional analysis. A close look at the sub-networks will reveal which brain functionality is impaired during MTBI recovery. Our work provides insights on why some MTBI patients suffer from long lasting symptoms while others do not. To achieve these goals, we found that it is necessary to have finer labeling of MTBI based on their long term recovery status. The three groups are defined as follows.

1. Healthy Control (HC): healthy subjects with a healthy brain. They are used in this study as a control group to help identify possible abnormalities with the two other groups. There are 30 subjects in this set.

2. MTBI-A: subjects with post-concussion symptoms six months after the incident, they are diagnosed with MTBI and low GCS score in their acute phase. The post-concussion symptoms are self reported. There are 24 subjects in this set.

3. MTBI-B: this group consists of the subjects without post-concussion symptoms after six months of the incident. They have scored low on the GCS test. This category is based on self-reported symptoms. There are 23 subjects in this set.

Table 3.1 *Demographics Information for the Three Groups*

TYPE	SIZE(person)	GCS	AGE(MEAN, yr)	GENDER	EDUCATION(yr)
HEALTHY	30	NA	40.17	18M/12F	15.38
MTBI_B	23	9	38.78	15M/8F	14.48
MTBI_A	24	8	44.54	19M/5F	14.04

3.2 Interpretation of pvalues and Significance Level

The p-value produced by the statistical test suggests the inconsistency between the observed data and the assumption that the null hypothesis is true. Statistical significance, which is the low probability of observing the results that are at least as extreme current results under the premise that the null hypothesis H_0 is true, is chosen to be $\alpha = 0.05$, which is a popular choice [12]. Therefore, if the pvalue derived by a statistical test is less than 0.05, then the result would be considered statistically significant and the null hypothesis would be false and the alternative hypothesis H_1 would be accepted to be true.

3.3 Data Centering and Normalization

This section is devoted to justifying certain data pre-processing steps we take before applying PCA or ICA. Our steps are:

1. Centering and normalizing each individual rs-fMRI data-set before the application of regression.

2. Test for group difference in time courses energy after the regression.

The intensity and the variance of the time series can be due to unrelated factors such as caffeine levels [34]. Each individual masked rs-fMRI data is centered and normalized before the application of Ridge regression. For each masked voxel, the mean is removed ($\mu = 0$) and STD is normalized to $\sigma = 1$. The normalization ensures that the results and conclusions are not biased by the original intensities of the rs-fMRI data.

Ridge regression allows each subject to choose its own optimal shrunkage parameter α within a pre-specified range between 0.1 to 10. The shrunkage parameter α affects the magnitude of the resulting time courses as shown in Fig 3.4. A large α leads to small magnitude of the time courses. The functional network is defined by the relationships between the time courses, as well as the magnitude of the time courses. To validate the claim that MTBI has an impact on the functional connectivity due to the relationship between brain regions, it is crucial to establish the fact that the absolute magnitude of time courses of each group are similar to each other. ANOVA one way test is carried out for this purpose, to test the potential

group difference of L1 norm of the time series. For subject s, the average absolute magnitude of time courses is computed as

$$E_s = \frac{1}{n}\frac{1}{T}\sum_{i=1,\ldots,n}\sum_{j=1,\ldots,T}|x_{s,i,j}| \tag{3.1}$$

where $n = 34$ and $T = 171$ is the number and length of meaningful time series that belongs to subject s, $x_{s,i,j}$ represents the j^{th} element of i^{th} time course of subject s.

Fig 3.1 shows the comparison between the three groups. It is clear that they have similar and comparable variance. More specifically, the three-group one-way ANOVA yields a large pvalue $p = 0.261$, which is too large to reject the null hypothesis $H_0 : \mu_{MTBI-A} = \mu_{MTBI-B} = \mu_{HC}$. Therefore it is highly unlikely that the magnitudes of the regressed time courses play role in our results and we can safely conclude that our final conclusion is determined by the inner relationships between the time courses, which is determined by the inherent characteristics of brain functional network of different types of subjects.

3.4 Group ICA with CANICA

CANICA identifies the maximal reproducible subspace from the rs-fMRI data where the group level spatial patterns resides, the set of 40 group level spatial patterns is presented in Figure 3.2. On the top left corner is their functional labels. The number of ICs was chosen to be 40 based on popular choices in the literature as well as similar results produced by multiple different values ranging from 20 to 100.

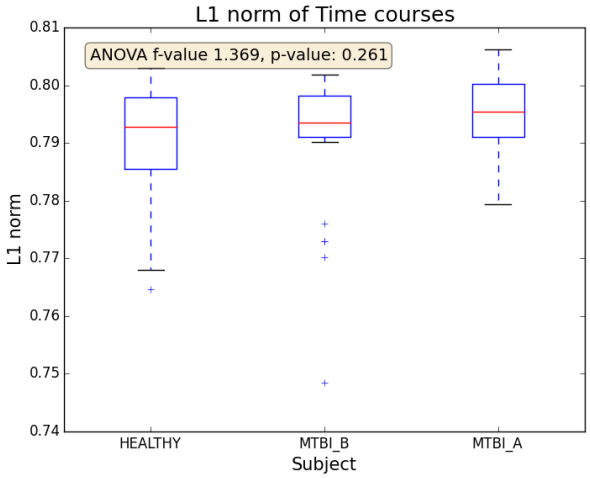


Figure 3.1 The average absolute magnitude comparison between the three groups. The value is computed based on Eq. (3.1). $p = 0.261$ for ANOVA f-oneway test between the three groups. The small pvalue argues against the existence of group difference of time courses magnitude.

3.4.1 Number of ICs

Picking the number M of ICs is important in fMRI applications, because it controls the number of nodes and the granularity of networks to be examined. A large M will create networks with excessive number of nodes, and each node corresponds to a small region of the brain, therefore the independence between different regions is compromised. It is also difficult to detect sufficiently interesting behaviors of the small regions in this case. On the other hand, having an overly small value of M will likely coalesce correlated brain regions together. There are a number of approaches to estimate the value of M, including empirical estimation and information theoretic approaches such as using the Akaike information criterion

(AIC), Kullback information criterion (KIC), and the Bayesian information criterion (BIC) [11]. The popular choice of M in fMRI ranges from 20 to 60 depending on the data-set so as to obtain the most meaningful brain parcellation according to the literature [183]. After our numerous attempts of M ranging from 30 to 100, we conclude that $M = 40$ best suits our application because of the high quality and meaningful spatial patterns it yields. In fact only 6 of the ICs out of the 40 ICs are labeled as noise by neurological experts.

Out of the 40 ICs, 34 components are labeled with meaningful known brain functionality and 6 of them are labeled as noise. These spatial patterns are computed by the CANICA hierarchical model with a pre-smoothing FWHM of 6 millimeters in size and a threshold level of 3. The entire brain functional network is split into the following list of sub-networks, their illustrations are in Fig. 3.2,

1. V - Visual network, includes: V1: primary visual, V2: visual associated area, V3: associate visual cortex, V4: associate visual cortex.

2. CC - Cognitive control network, CC1: bilateral dorsolateral prefrontal and right posterior parietal, CC2: bilateral frontal anterior frontal lobe, CC3: bilateral frontal lobe, CC4: L posterior parietal, CC5: bilateral frontal lobe.

3. D - Default mode network, DMN. D1: PCC, D2: PCC, D3: medial prefrontal, D4: medial prefrontal cortex.

4. A - Auditory network, consists of cerebral regions that controls auditory processing. A1: Left Wernickes, A2: primary/secondary auditory cortex, A3:

Left Wernickes, A4: auditory cortex, A5: left brocas/Wernickes, A6: right Wernickes.

5. C - Cerebellum, a brain region that plays an important role in motor control as well as some cognitive functions such as attention and language.

6. SC - Subcortical network. SC1: putamen, SC2: thalamus.

7. SM - Sensorimotor network. SM1: supplementary motor area, SM2: primary somatosensory, SM3: mid cingulate.

8. NOISE: 6 extracted spatial patterns are labeled as noise because they do not overlap with any known brain functionality regions.

We also study the static and dynamic functional connectivity within these sub-networks and attempt to discover the causes of the group differences between the two MTBI types and healthy controls.

CANICA produces highly reproducible ICs compared to other group ICA methods. The 40 group level spatial patterns are shown in Figure. 3.2. The set of labels can be divided into several categories based on their function, shape, and position in the template brain. The labeling scheme is similar to what has been reported in related literature [129, 183]. This set of spatial patterns serve as common basis upon which the subjects can be compared.

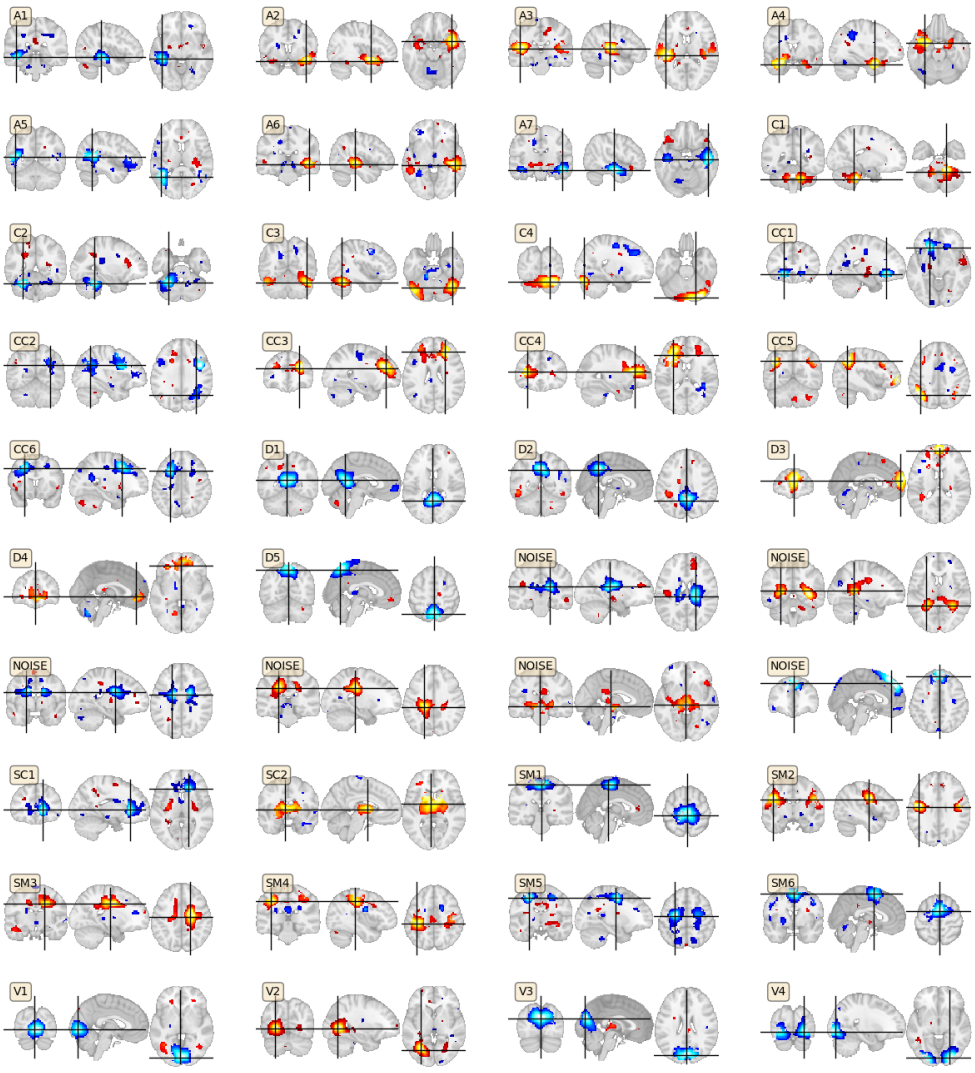


Figure 3.2 40 spatial patterns derived by the CANICA technique, the shape and position of each spatial component is illustrated with three coordinate cuts. The spatial patterns are labeled by neurological experts. 34 of them overlap with clearly known brain networks, 6 of 40 are labeled as noise components. The details of the labels can be found in a list in section 3.4.1.

3.5 Ridge Regression Results

CANICA derives 40 group level spatial components that serve as the common basis for subject and group comparison. The functional connectivity is computed between the time courses of the same subject. The time series is obtained by the regression technique, resulting in 40 time courses that correspond to the 40 CANICA extracted spatial patterns. Each time course captures the temporal dynamics of the corresponding spatial pattern. Individual masked rs-fMRI data is used as the target matrix Y and the set of group level spatial patterns as the design matrix X. The Ridge regression finds the relation between X and Y using an objective function that suppresses the magnitude of the resulting time series. Figure 3.3 shows the set of time series of a random Healthy Control (HC), each of the 40 time series has a length of 171 time points, which is equal to the duration of the fMRI sessions. The time series show similar magnitudes, their pairwise relationships define the functional network. In the static functional analysis, the whole time series will be used to compute its distance to another time series. In the dynamic functional analysis case, the time series will be sliced into a number of shorter time series using a moving window along the temporal axis, and functional analysis will be performed on every temporal slice, the series marks the dynamic changes of the functional brain of the subject.

Ridge regression suppresses the colinearity in the underlying variables by adding a shrunkage term into the objective function [14], and its contribution is controlled by a shrunkage parameter α. Figure 3.4 shows the effect of the shrunkage

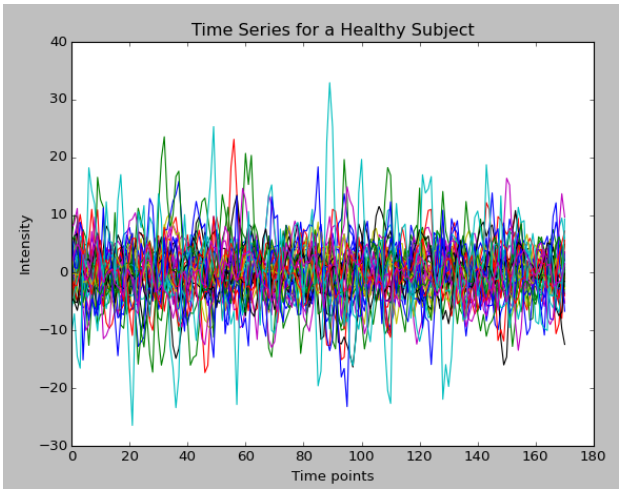
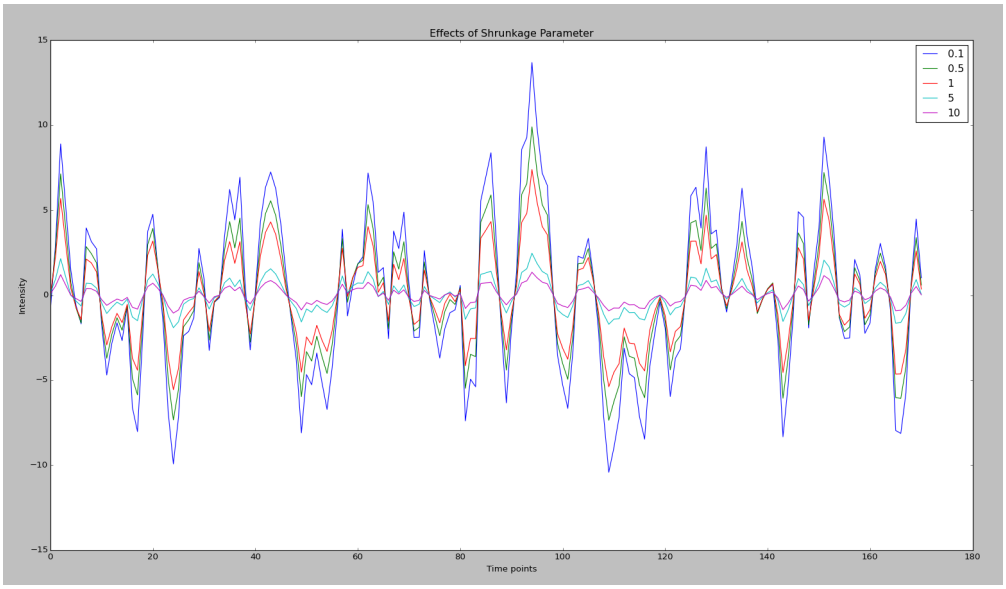


Figure 3.3 Each subject is represented by 40 time series derived from Ridge regression with the rs-fMRI data and the group level spatial patterns as input. Each time series corresponds to one spatial pattern and captures the temporal activation of those brain regions.

parameter α has on the resulting time series. The changing values of α controls the magnitudes of the resulting time series.

The subject-specific time series are derived from their rs-fMRI and the group level spatial patterns. An illustration of the time series of a randomly chosen healthy subject is plotted in Figure. 3.3. The time series represents the temporal activation of the spatial patterns.

The relationships among the time series give rise to the functional network $G = (V, E)$, where the weight of each edge is the distance between the two time series corresponding to the end points of the edge. An example of a functional network is shown in Figure. 3.5. Various distance or similarity measures have been adopted for fMRI analysis, such as the correlation coefficient, Fisher's z-score, and partial correlation computed by precision matrices. The metric we use is the $L1$



n

Figure 3.4 The effect of shrunkage parameter α. As the value of α decreases from 5 to 0.1, the regularization term in the objective function plays a smaller role, and the computed time series magnitude grows.

norm which has a more straightforward explanation to the distance based network properties such as SP and MST, and also because it produces consistent results over different window parameters during multiple tests.

3.6 Functional Brain Network Construction

A static functional brain network is constructed by computing the distance/similarity between each pair of time series for a given subject, which is then used to construct the adjacency matrix of the network. In this research, the distance metric is chosen to be the L1 norm because of its straightforward interpretation and its empirical superior results. In addition, L1 norm yields stable results under various window parameters. Figure 3.5 shows the static functional network for a healthy subject, the edge length is proportional to the distance between the two nodes. The network forms several local sub-networks, which can constitute the sub-networks with their own functionalities. The focus on this work is the dynamic functional analysis. The main difference is that instead of computing a single static functional network, the entire time series is processed by a sliding window, and a series of networks are constructed for each window.

3.7 Statistical Tests

We analyze the dynamics of the functional network by constructing a series of such networks using a moving window. This series of networks produces a series of measurements $z_{s,1}, \ldots, z_{s,n_{windows}}$, for subject s, one for each window. We use

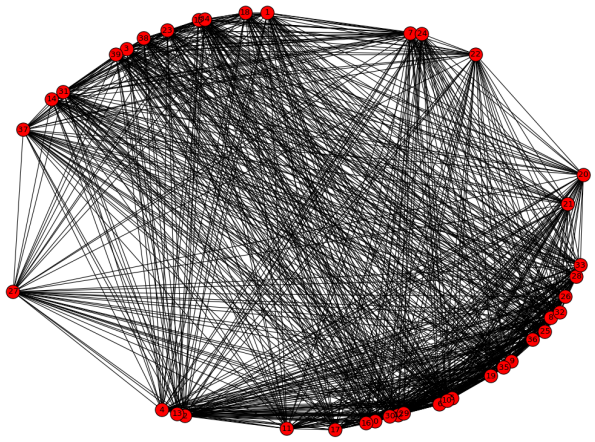


Figure 3.5 The graph of the functional network for a random healthy subject. The network reflects the relationships between the 40 spatial patterns, the lengths of the edges are proportional to the distances.

window size of 30 with a step length of 10, the shape of the window is rectangular, which yields similar results with a commonly adopted tapered sliding window.

In the resting state fMRI setting, different network properties represent different aspects of the functional network. The shortest path intuitively represents the information exchange cost between the corresponding regions of the brain. A network with a large shortest path indicates decreased efficiency. On the other hand, the large local clustering coefficient suggests that the nodes tend to create tightly knit groups characterized by a relatively high density of ties. Lastly, the total weight of MST measures the efficiency of broadcast communication among the nodes.

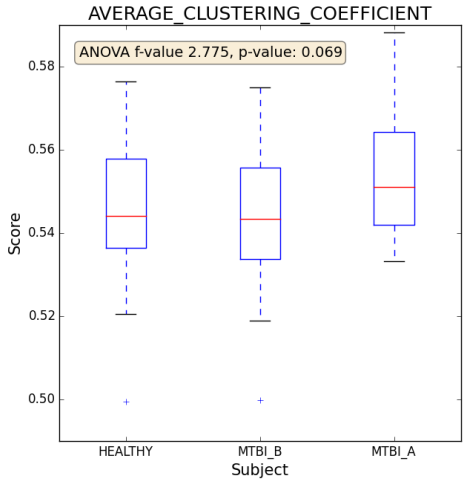
Figure 3.7 displays the summarized value for each subject, using a number of network properties including the average shortest path and average clustering coefficient. Three-group ANOVA test pvalue, group mean, and STD are shown in the same figure.

The ANOVA pvalues indicate whether there is a difference between the three groups. A detailed comparison between the three groups is shown in Table 3.2, which show the t-test pvalues between each pair of groups. The t-tests are performed in multiple runs, with equally numbers of subjects drawn from each group randomly.

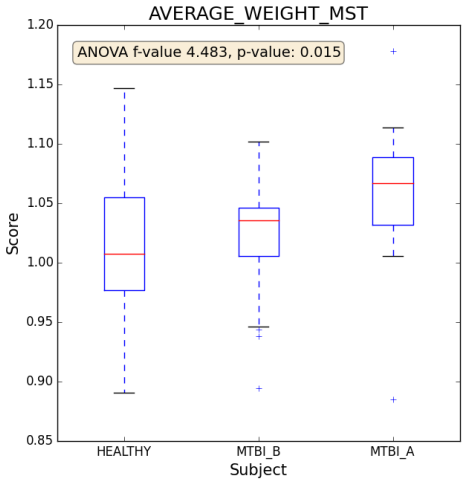
Both methods uncover differences between MTBI-A and the other two groups: MTBI-B and HC. Figures. 3.6 and 3.7 show the network properties of all subjects, divided into three groups with different marker styles and colors. Group mean and STD are also available for groups in the corresponding regions. More importantly, one way ANOVA test pvalues score for the three groups is annotated on the figure. The corresponding tables show in detail how each group is numerically different from each other. Promisingly, most of the measured properties suggest that differences exist between the groups. A number of measured network properties show group difference between MTBI-A and the other two groups. In particular, MTBI-A has larger average weights of MST ($p = 0.017$) and average clustering coefficient ($p = 0.079$) than that of HC. More interestingly, MTBI-A has larger average shortest path, average weight of MST, and average clustering coefficient than the other type of MTBI, MTBI-B. As for the changing dynamics of the network properties, MTBI-A type shows slowed dynamics displayed by average shortest path ($p = 0.09$) than that of HC, between MTBI-A and MTBI-B, MTBI-A shows reduced dynamics in average shortest path (0.019) than that of MTBI-B.

The implications are two-fold. First, the separation of the two MTBI types and the fact that no difference is found between MTBI-B and HC suggests that the MTBI incident didn't cause major impairment on the function network for MTBI-B

(a) Average Clustering Coefficient.



(b) Average Weight of MST.



(c) Average Shortest Path.

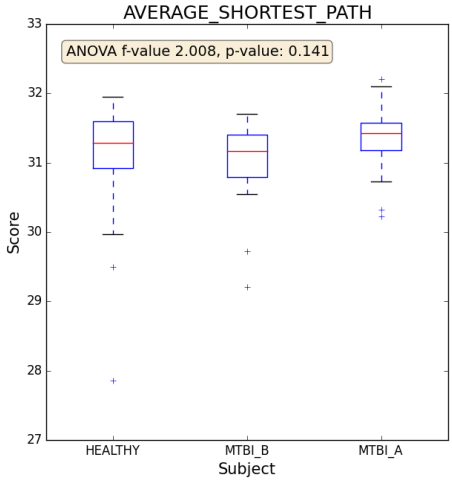
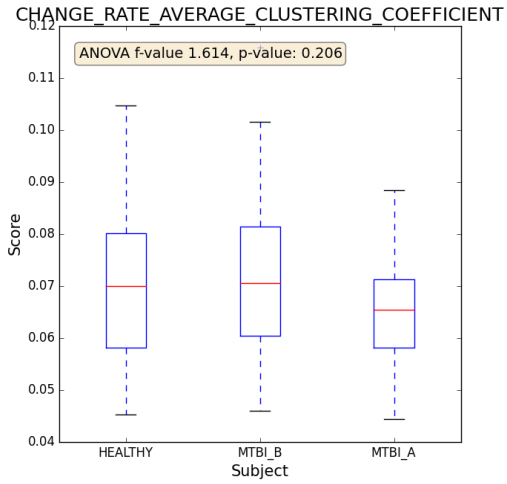
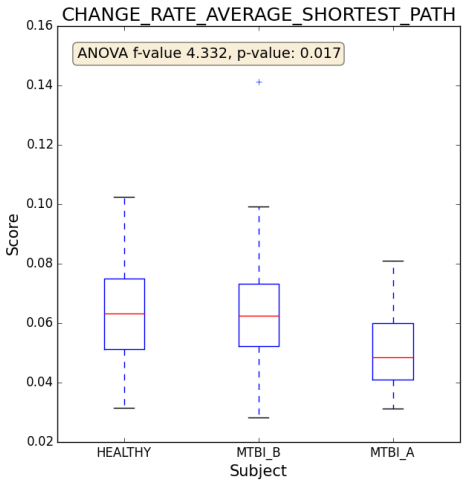


Figure 3.6 Static Network properties for the subjects. The measured properties of functional network includes average shortest path, the average weights of the MST, and the average clustering coefficient of the network. Three groups are distinguished by their line-style and marker. The ANOVA one-way pvalue indicates the possibility of the observations happen given the null hypothesis H_0 were true.

(a) Dynamics in Average Clustering Coefficient.



(b) Dynamics in Average Shortest Path.



(c) Dynamics Average Weight of MST

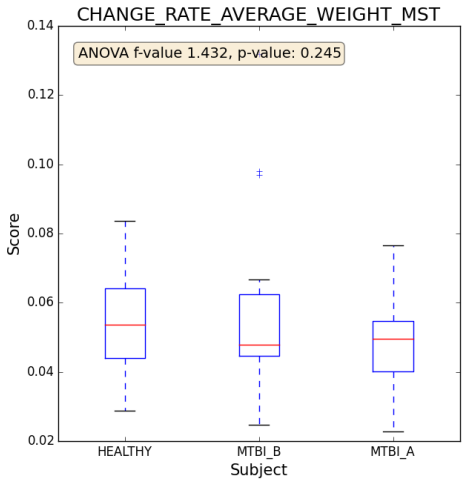


Figure 3.7 Dynamics of network properties for the subjects.

The dynamics is computed as the average absolute difference between values of adjacent time points. The dynamics is computed for these network properties: average clustering coefficient, average shortest path, and the weights of the MST. Three groups are distinguished by their line style and marker. The ANOVA one-way pvalue indicates the likelihood of the observation if the null hypothesis H_0 were true, and the mean and STD of each group is annotated within their corresponding region.

type. The mixing of the two types of MTBI is likely to prevent meaningful discovery. On the other hand, MTBI-A type is going through a distinct recovery process that separates itself from both healthy and MTBI-B type. The extended average shortest path and average clustering coefficient suggest a network that is locally clustered and lack of long range connectivity, compared to the efficient small worldness reported for a healthy brain [80]. Moreover, the larger shortest path implies distant connections between different nodes of the functional network, and hence a less coordinated brain, as is reported in other studies [16, 43]. The static results establish the topological structures of the functional network. The dynamic functional connections of MTBI have not been studied systematically before. Figure 3.7 and tables 3.2f - 3.2d show that functional dynamics of MTBI-A type are changing slower than that of the healthy control and MTBI-B. The implication of this slow down is that the MTBI-A patients are suffering from a reduced level of metabolic process in their brain.

The distinction of MTBI subjects according to their recovery six months after the incident is the foundation of our research. The different pace in recovery is commonly seen when it comes to neurological diseases, since it is usually complicated by pre-existing psychiatric or substance abuse problems, poor general health, concurrent orthopedic injuries, or comorbid problems (e.g. chronic pain, depression, substance abuse, life stress, unemployment, and protracted litigation). [36–40,48,53,54]. It is also reasonable to argue that the recovery rate is affected by the types of impact the subject has suffered. MTBI is such a broad concept that it contains various sub-categories and can be associated with broad physiological and psycho-

logical symptoms. For these reasons, the long term outcome is strongly related to recovery pace, and the various types of MTBI. As an example, A recent publication found no group difference between MTBI and HC [179], the authors focused on the default-mode network (DMN) and sub-cortical structures with dynamic sliding window method, followed by group comparison using the ANOVA and the MANCOVA methods. Their final results, which show no significant group differences, were based on measures of static or dynamic connectivity within a prior ICN. Reduced (HC > MTBI patients) static connectivity was observed in the DMN at uncorrected ($p < 0.005$) thresholds [179]. Reduced cerebral blood flow was observed in the study reported [54], which can play an important role in why the reduced functional connectivity is consistently found in MTBI fMRI studies [41–43].

3.8 Measurements Within the Sub-networks

Studying the possible existence of group differences within the sub-networks can provide valuable insights on which sub-networks are responsible for the difference in the whole-brain functional network and also which sub-networks are most affected by MTBI disorder in a resting brain. We study the set of time series that belong only to each of these sub-networks, and examine with the same graph-theoretic methodology to determine if group differences exist within the sub-networks. Note that by limiting our attention on the sub-networks, we are intentionally ignoring the connections between these sub-networks, which can be a topic for future studies. The sub-networks are listed below, and each sub-network is shown in Fig 3.8. Instead

of showing the ICs of Fig 3.2, this figure shows combined spatial regions for better illustration.

1. Visual network.
2. Cognitive control network.
3. Default mode network, DMN.
4. Auditory network.
5. Cerebellum.
6. Subcortical network.
7. Sensorimotor network.

As we delve deeper and study the internal gateways of each sub-network, we can learn which networks are responsible for the group differences between MTBI-A and the other two groups. It also provides us with an possible explanation on the effects of MTBI, and may help to answer questions such as "which damaged functional sub-network is responsible for prolonged MTBI recovery", "which network is not affected by MTBI at all". In essence, the integrity of the functional network does not imply the well being of the actual brain regions; more concrete anatomical studies have to be conducted to draw more definitive conclusions. As shown by the pvalue tables, none of the sub-networks reveal differences between MTBI-B and HC, which is consistent with the global brain result we obtained in the previous sections. Difference between MTBI-A and other two groups are found in the auditory network,

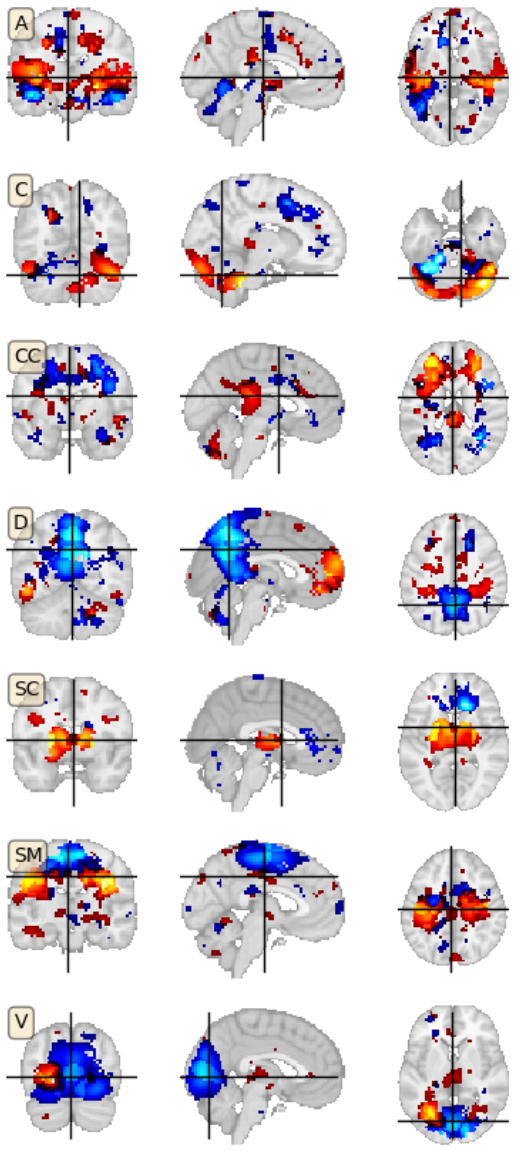


Figure 3.8 Illustrations of sub-networks: visual, auditory, DMN, sensorimotor, subcortical, and cerebellum. These sub-network are composed of the ICs derived from CANICA and labeled by their functionalities: V, C, A, CC, D, SC, SM.

visual network, sensorimotor functional network, and DMN network, whereas the rest of the subnetworks seems to be unaffected by MTBI, and they are: cerebellum, subcortical and cognitive control network.

The Ridge regression derives 40 time series matching the group level spatial pattern computed by CANICA, for each subject. The subgraph of each sub-network is defined by the relationship between the time courses of the endpoint ICs belonging to the subnetwork. We limit our attention to the sub-networks and attempt to find similar group differences in the sub-networks.

The graph theoretic algorithms applied to each sub-network are: average shortest path, average clustering coefficient, and average weight of the MST. Then we summarize the network measurements along the sliding windows to a single measure, using either the static method or the dynamic method. Finally, independent t-test is performed to each sub-network for each possible pair of groups.

The results of the auditory sub-network are summarized in Table 3.3. MTBI-A type shows functional abnormalities from HC and MTBI-B. Compared to HC, MTBI-A has a larger average shortest path ($p = 0.074$), and slowed dynamics of average shortest path ($p = 0.053$). The difference between MTBI-A and MTBI-B is more obvious. MTBI-A exhibits larger weights of MST ($p = 0.016$), larger average shortest path ($p = 0.013$) and slowed dynamics of average shortest path ($p = 0.061$). The auditory sub-network is affected by A-type MTBI, the sub-network efficiency is reduced compared to HC and b-type MTBI. There is no statistical difference between MTBI-B and HC using one graph-theoretic measures.

Results of the DMN network are included in Table 3.4. DMN is a resting state network that is commonly related to brain disorders. Numerous studies have reported to observe either increased or reduced functional connectivities in DMN sub-network for patients with brain pathologies such as: Schizophrenia, autism, and Alzheimer's diseases. In our sub-network study, we find substantial difference between the two type of MTBI. Indicated by the static and dynamic measurements of average shortest path, MTBI-B has efficient and more active functional brain, its static average shortest path is smaller than that of MTBI-A ($p = 0.036$) and the changing frequency of shortest path is faster ($p = 0.03$).

Results for Sensorimotor network are shown in Table 3.5. Sensorimotor network is commonly impaired by MTBI disorder, the lack of orientation and balancing are commonly observed symptoms of MTBI patients. The difference between the two types of MTBI is again revealed by the analysis of this sub-network, $p = 0.053$ for average weight of MST, $p = 0.001$ for average clustering coefficient. MTBI-A also differs from HC group on average clustering coefficient ($p = 0.036$) and dynamics of shortest path ($p = 0.026$).

The results for visual networks can be found in Table 3.6. The difference between MTBI-A and HC is prominent in this sub-network. HC group has a smaller weight of MST ($p = 0.047$) and smaller average shortest path ($p = 0.056$). These differences indicate the superior efficiency of the functional network of HC. The information exchange is more efficient and cheaper for HC group. MTBI-A group is likely to experience difficulty in visual processing.

Results for cerebellum regions are shown in Table 3.7. Cerebellum plays an important role in motor control and other movement-related functionalities. Although it does not initiate movement, it is in charge of coordination, precision, and accurate timing by processing the input from the sensory systems of the brain. No difference among the three groups are found by analysis of this sub-network, possible explanations will be discussed later in this section.

Results for cognitive control network are shown in Table 3.8. Cognitive control involves brain regions related to cognitive processes, including working memory, reasoning, task flexibility, and problem solving as well as planning and execution [10]. No group difference is found by cognitive control sub-network analysis.

Finally, the results for subcortical network are shown in Table 3.9. Similar to cerebellum and cognitive control network, no group difference is revealed by our analysis on this functional sub-network.

Four sub-networks our of the sense show group differences similar to the results of the whole-brain analysis. These are the auditory, visual, sensorimotor, and DMN networks. They are highly related to the most common MTBI symptoms, such as blurry vision, difficulty in balancing and orientation, and ringing in the ear [18–22, 24]. Research has shown that the MTBI groups perform lower on the full and secondary scales for attention and response accuracy on the integrated visual and auditory (IVA), continuous performance test (CPT) and Neuro-psychological Impairment Scale (NIS) [23]. More specifically, the MTBI groups perform worse than HC on reaction time, lack of attention, impulsivity, and variability on reaction time. On the NIS, the MTBI group reports more neuro-psychological symptoms

than the healthy control group. Early research [25, 26] has shown that disrupted white matter activities have been detected across certain networks in DMN, sensorimotor network, and attentional and frontal network. An animal study [27] on rats has also reported that the sensorimotor functionality is impaired when the subject rat underwent MTBI. Among all the sub-networks that showed differences between the three groups, DMN has been reported most frequently to be related to MTBI in resting state fMRI research. Another study [42] found that there are substantial reductions of connectivity in the PCC and parietal regions and increased connectivity around the MPFC in patients with MTBI ($p < .01$). The reduced posterior connectivity correlates positively with neuro-cognitive dysfunction, while the increased frontal connectivity correlates negatively with post-concussion symptoms. Another study focused on the sub-acute phase of MTBI [43], the authors found an association between MTBI and the reduction of connection strengths within DMN in the posterior cingulate and lateral parietal cortices. Their work suggests the existence of brain alterations in the resting state default mode network in the sub-acute phase of MTBI.

It is notable is that the other three sub-networks: cognitive control, cerebellum, and subcortical networks appear to be irrelevant to MTBI, because these networks seem to be similar in the three groups. This result conflicts with the obvious cognitive behaviors exhibited by MTBI subjects. Studies suggest that individuals who have experienced a concussion or mild traumatic brain injury (MTBI) show deficits in cognitive control [17]. The absence of group difference is possibly due to the nature of resting state study and the high order functionalities of these

sub-networks. For example, the cerebellum is related to motion in a way that it does not initiate motion, but it fine tunes the balancing and timing. These sub-networks, especially cognitive control network are expected to be turned off during the resting state study. Most cognitive studies require the subject to perform certain cognitive tasks or processes in order to evaluate their cognitive performance. In the resting state study, any cognitive process is undesirable and prohibited. Although it is highly likely that the patients cannot entirely empty their mind during the 5 minutes fMRI scan, the thinking should happen rarely and intermittently, and therefore should not be a dominant component of the rs-fMRI signal.

3.9 State Analysis

In research [4, 5], the k-means clustering technique is applied to the distance matrices computed in each window to identify a small number of functional connectivity (FC) states, which is analogous to the EEG micro-states: short periods during. which scalp topography remains quasi-stable [3].

The states are identified as the centroids of the k-means clustering applied on the distance matrix calculated within each window; those distance matrices reflect the most typical and common states the brain goes through to carry out various cognitive processes. By studying the characteristics of each brain state, and the distribution of the time each group spends in each state, we may find possible explanations of the static and dynamic characteristics of the MTBI-A type.

Compared to the study using state analysis on Schizophrenia [5], in which the

decision on the number of clusters or typical states is made using the elbow criterion on the ratio of within and between cluster distances, our study uses Bayesian Information Criterion (BIC) to determine the value of k, the number of clusters. BIC assumes a likelihood model, which can be constructed by viewing the k-means clustering as a Gaussian mixture model.

The formula of BIC is defined in Eq. (3.2):

$$BIC = -2 \ln p(x|\hat{\theta}, M) + k(\ln n - \ln 2\pi) \tag{3.2}$$

where $\ln p(x|\hat{\theta}, M)$ is the maximum value of the marginal likelihood function given the model M and the maximizing parameter $\hat{\theta}$. The parameter k is the number of free parameters to be estimated, n is the number of data points included in the dataset x.

BIC can also be rewritten as an increasing function of the error variance σ_e^2 and of k [2]. Thereby it is preferable to pick the model with the lowest BIC value, which is $k = 5$ in Fig. 3.9.

The states are defined by the distance matrix calculated within each sliding window because it uniquely determines the connection and topology of the functional network during that short period of the window time. According to the BIC measurements in Fig 3.9, five typical states are extracted from all the distance matrices from every subject and every sliding window. These centroid states can be used to label all the other states based on their memberships in the k-means clustering. The individual functional activities can be described as a temporal transition

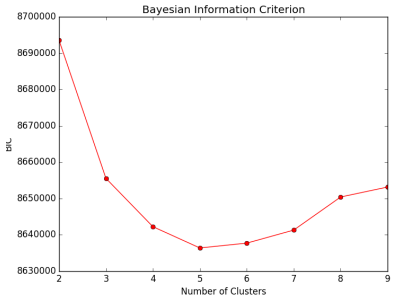


Figure 3.9 BIC scores with various numbers of clusters k. The k with the lowest BIC is most preferable.

between the centroid states. The three groups do not spend an equal amount of time in each state. Each one of them has its own 'favorite' state and 'least favorite' state, and the characteristics of those states explain the behavior of the type. The characteristics of a state is defined as follows,

1. A: Average connectivity strength in auditory sub-network.
2. V: Average connectivity strength in visual sub-network.
3. D: Average connectivity strength in DMN sub-network.
4. SM: Average connectivity strength in sensorimotor sub-network.
5. SC: Average connectivity strength in subcortical sub-network.
6. CC: Average connectivity strength in cognitive control sub-network.
7. C: Average connectivity strength in cerebellum sub-network.
8. ASP: Average shortest path length of the whole functional network.

9. ACC: Average clustering coefficient of the whole functional network.

For ease of comparison of the states, we compute the rank from the raw values, with 1 being the smallest and 5 being the largest for the five states. The ranks and the amount of time each type stays in is summarized in Table 3.10. State2 is characterized by average connectivity strength in all sub-networks, as well as the values of network measurements. All three groups spend similar amount of time in state2, around 34%.

State5 and state3 show the strongest and weakest connections within all sub-networks, respectively.

State5 also has the smallest average shortest path and a low clustering coefficient, which reflects high efficiency transmission between the regions. MTBI-A type spend significantly less time in this state compared to the two other groups (10.0% compared to 18.0% and 17.4%).

On the contrary, state3 displays the decreased link efficiency and a high tendency of local cliqueness. MTBI-A visits state3 only half as frequently as MTBI-B and HC group, thereby it spends more time in states 1 and 2, and explains why MTBI-A has reduced transition between the states..

Finally state1 and state4 have comparable network measurements somewhere between those of state3 and state5. Compared to state4, state1 has reduced connectivity in sub-network: auditory, DMN, visual, sensorimotor, and cognitive control, the same results as in the sub-network analysis. Those are the sub-networks that are found affected by the MTBI disorder. In addition, state1 is characterized by

reduced clustering coefficient and large shortest path. Together they indicate disruption of the small world topology. It is also interesting that MTBI-A stays in this state for the majority of its time, 35.0%, twice as much as HC and MTBI-B group. The fact that MTBI-A has two major states: 1 and 2, to a certain degree, explains why MTBI-A has reduced functional connectivity and slowed dynamics.

Chapter 4: Conclusions and Future Directions

This study uses an analysis pipeline that includes stages of pre-processing steps, noise removal, advanced ICA, cross validating Ridge regression, graph-theoretic analysis, and statistical tests. By identifying independent brain regions and studying their functional relationships to model the brain as a functional network, this work is able to detect both static and dynamic differences between the group pairs: MTBI-A and MTBI-B, MTBI-A and HC, using their brain scans within the first 24 hours of the brain injury. This is a new finding that has not been previously explored, to the best of our knowledge.

A major contribution of our work is to identify a functional brain basis to distinguish MTBI-A and MTBI-B groups based on the patient's symptoms after 6 months. This new understanding of MTBI is supported by a combination of facts: 1) there is no universally accepted definition of MTBI, 2) MTBI can be caused by various reasons and has demonstrated a broad spectrum of symptoms, 3) differences

in recovery pace are commonly reported but they are still controversial [31, 52, 53]. It is possible that prior research did not reveal dynamic differences between HC and MTBI, due to absence of the long term tracking of the subjects. According to our results, the two types of MTBI showed significant differences in their brain functions. From what we can tell, there are no significant functional difference between MTBI-B and HC, using either static or dynamic functional analysis. However, group differences between A-type MTBI and the other two groups are detected by several network properties and tests. The uniqueness of MTBI-A suggests that subjects go through a different recovery process than those of MTBI-B, and the divergence can be detected within the first 24 hours of the brain injury, because we use acute stage MTBI brain scans. Persisting post-concussional symptoms which last $3-6$ months after the injury may have different acute symptoms (albeit the same initiating event). Determining the prevalence of lasting symptoms is challenging. Around 75% MTBI patients' symptoms are resolved within the first 3 months, but $5-15\%$ of patients suffer post-concussional symptoms 12 months after the initial injury [29, 53]. The etiology that prolongs the recovery of MTBI-A may not be present in the cases of MTBI-B. Not all MTBI subjects are created equal, various brain regions can be impaired during the accident and lead to comparable symptoms. The key differences between MTBI-A and MTBI-B types were not observable from their behavioral studies, nor are they predictable by the GCS tests [30]. Nevertheless, the underlying impaired brain functionality seems to alter their rs-fMRI functional network [31, 43].

Static analysis discovers the relationship between the prolonged recovery process of MTBI-A type and the disruption of the small world topology of functional

network. The functional network of a healthy brain assumes small world topology [72, 80], which is characterized by a dense local clustering yet a short path length between any distant pair of nodes because of the existence of a few long-range connections. This type of graph topology has proven to be an attractive model for many empirical graphs, such as social networks, internet, and gene networks, because of its architectural advantage [84]. It is commonly used to model the organization of brain functional network because this topology can facilitate both segregated/specialized and distributed/integrated information processing [84].

The whole-brain functional network of MTBI-A is characterized by locally clustered nodes, while lack of long range functional connections of their functional networks measured within 24 hours of the initial injury, ($p < 0.03$). The enlarged "traveling distance" of the functional network is in line with what has been reported in related literature. Johnson [43] reported that patients with a concussion showed a reduced number of connections and strength of connections in the posterior cingulate and lateral parietal cortices within the resting state DMN. Moreover, connections between the left dorso-lateral prefrontal cortex and left lateral parietal cortex showed a significant reduction in magnitude as the number of concussions increased. Regression analysis also indicated an overall loss of connectivity as the number of MTBI episodes increased.

A recent publication of Mayer [179] reported reduced (HC > MTBI patients) static connectivity was observed in the DMN at uncorrected ($p < 0.005$) thresholds. Moreover, a trend ($p = 0.07$) for decreased dynamic connectivity in patients across all ICN was observed using spatially invariant analyses (DisCo-Z).

The reduced connection in functional networks is associated with more expensive information exchange, delayed response to external stimuli. Because complex cognitive processes are possible the product of information transmission between different functional regions of the brain [130] . The disruption of the normal network topology can undermine the high order functional processes and can be exhibited as the difficulty in organizing thoughts for MTBI patients. The reduced functional connectivity reflects the obstructions in the communication paths [1], due to the impact of MTBI disorder.

Using dynamic functional analysis, MTBI-A type is characterized by its slowed change rate from a time window to the next window. Specifically, the value of average shortest path is found be change slower than that of MTBI-B and HC groups along the window. Shortest path may indicate the efficiency of the functional networks, and the dynamics of shortest path may reflect the flexibility of the network. The brain functional network is ever-changing and ever-updating, the changing relationship between the regions facilities and supports the higher order cognitive processes [130]. The inability to change fast between different functional states may relate to the mitigated cognitive and emotional impairments of MTBI-A type of patients.

The dynamic functional analysis suggests that MTBI-A group may suffer from slowed metabolic processes compared to HC and MTBI-B groups. This claim is in line with the frequently reported MTBI symptoms [53, 171]. Studies [172–175, 178] have shown that age plays as an important factor in the MTBI recovery process, its implication is even stronger than that of the CT results, and aging in turn has

a strong effect on the brain metabolism [176, 177]. These links between slowed metabolism and persisting post-concussional symptoms are potential reasons for a prolonged recovery process of MTBI-A type. More explicit connection between the prolonged MTBI recovery and slowed brain metabolism is an interesting topic for future study.

Different results have been reported by Mayer [179] that there is no association discovered between abnormalities in functional network within DMN and MTBI during the semi-acute phase of recovery- our results of the well-recovered group MTBI-B matches their results. Neither static or dynamic functional analysis found the differences between MTBI-B and HC groups. The discrepancy can be caused by the following reasons, 1) our data-set contains 77 subjects in total, which produces more convincing statistical test results than a subset of research with relatively small data-sets (less than 20). 2) Our research is the first to adopt the concept of separation of MTBI patients based on their long term recovery outcome. This can make a difference because the combination of MTBI-A and MTBI-B groups is more similar to HC than MTBI-A alone. 3) Our methods remove noise and improves SNR aggressively, in addition to a number of noise removal procedures in the pre-processing stage, masking before ICA and the adopted CANICA method can account for our results. 4) Finally, our method is data driven that models the whole brain, whereas the majority of previous work mainly focused on the DMN and other well known seed regions.

In addition to whole brain analysis, subgraph formed by sub-networks with different brain functionalities are investigated. Out of the seven sub-networks, MTBI-

A shows functional abnormalities in auditory, visual, sensorimotor and DMN, when compared to HC and MTBI-B. Again, the well recovered B-type MTBI group displays no difference to HC group. These sub-networks are found to be related to the most common symptoms of MTBI, such as ringing in the ear, difficulty in balancing, and blurred vision [18–22, 24]- they are also the most explicit symptoms of MTBI. These sub-networks are also the ones that are most consistently studied during resting state study. It has been suggested that these sub-networks maintain certain activations when the brain is at rest. Our results suggest that these sub-networks are most susceptible to MTBI disorder, which confirms the results of many non-functional studies of the MTBI disorder.

A study [7] that compares the visual sensitivity of children with MTBI to that of matched non-injured children and to assess the evolution of visuo-perceptual performance over time. The processing threshold for complex visual stimuli were significantly affected for the MTBI group, and this affect last as long as 12 weeks after the injury. Their finding suggests that MTBI children have selective visual deficits for high-order information processing, and this deficits persists over a long period (12 weeks).

Similarly, another study [6] investigates the relationship between the visuospatial orientation of attention and obstacle avoidance during gait in MTBI patients individuals over a month post-injury. They observe that MTBI patients present impairment in orient attention in space immediately after the injury, they also have statistically more obstacle contacts than the healthy controls. They conclude with the discovery of the association between MTBI with spatial attention deficits and

the ability to avoid obstacles during movement.

The fMRI study conducted by [9] examines the association between semi-acute phase of MTBI and certain cognitive attentional, memory deficit. They conclude that MTBI patients who have difficulty in disengaging and reorienting auditory attention also show hypo-activation abnormalities in various regions including sensorimotor, auditory and DMN sub-networks. Functional connectivity abnormalities related to multiple sensorimotor regions are observed in a study [8] with children participants. They suggest that the altered connectivity may be related to the participants' impaired sensorimotor performance tests. On the other hand, the conclusion on the other three sub-networks which do not show group differences is hard to interpret, as these networks are related to high-order functional processes, and also because they are generally seems to be turned off during the resting state study.

Finally, the state analysis identifies five states using k-means clustering technique. MTBI-A type spends significantly less time in state5 with the strongest functional connectivity (10.0%). Compared to HC and MTBI-B, MTBI-A stays in state1 for 35.0% of its total time, and it is a state with high-degree disruption of small world topology. The distribution of how MTBI-A spends its time explains why it has reduced strength in functional connectivity, as well as reduced dynamics in multiple network properties.

Our findings implies the MTBI disorder reveals itself in sub-networks such as auditory, visual, sensorimotor and DMN when the brain is at rest. Without conducting behavior tests, these biomarkers are able to provide information on the functional disorder of the patients due to MTBI.

In conclusion, by studying the static and dynamic functional network of the whole brain as well as the region with known functionality, we are able to find a number possible bio-markers that tell apart MTBI patients who will have prolonged concussional symptoms from those who will recover quickly within 6 months. To the best of our knowledge, our study is the first functional study to find differences between different types of MTBI, and suggest their recovery pace. It is also the first to identify the dynamic functional difference between MTBI and healthy control. Our finding is important to the separation of MTBI subjects according to their self-reported 6 months recovery status. We advocate that brain recoverable disorder functional study should pay more attention to the recovery rate, which indicates the possible different pathologies of the common symptoms on the surface. Our results support the static functional hypo-connection reported in the literature. Furthermore, the slow dynamics in multiple network measurements suggests that the prolonged recovery may be related to slowed brain metabolism. Finally, the value of the bio-markers extracted within 24 hours after the injury are related to the subject's 6 month recovery. These bio-markers can be valuable to help clinicians to create the best recovery course of action for MTBI patients.

Table 3.2 *pvalue of Group Comparisons for Network Properties*

(a) Average Weight of MST

	HC	M-B	M-A
HC	1	0.718	0.018
M-B	0.64	1	0.021
M-A	0.017	0.02	1

(b) Average Shortest Path

	HC	M-B	M-A
HC	1	0.747	0.139
M-B	0.721	1	0.031
M-A	0.142	0.031	1

(c) Average clustering Coefficient

	HC	M-B	M-A
HC	1	0.73	0.067
M-B	0.763	1	0.04
M-A	0.07	0.039	1

(d) Dynamics in Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.738	0.197
M-B	0.714	1	0.097
M-A	0.148	0.094	1

(e) Dynamics in Average Shortest Path

	HC	M-B	M-A
HC	1	0.786	0.01
M-B	0.785	1	0.02
M-A	0.008	0.019	1

(f) Dynamics in Weights of MST

	HC	M-B	M-A
HC	1	0.568	0.228
M-B	0.58	1	0.135
M-A	0.206	0.136	1

Table 3.3 *pvalues of Auditory Network*

(a) Average Weight of MST

	HC	M-B	M-A
HC	1	0.385	0.12
M-B	0.393	1	0.016
M-A	0.108	0.016	1

(b) Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.616	0.585
M-B	0.645	1	0.31
M-A	0.519	0.31	1

(c) Average Shortest Path

	HC	M-B	M-A
HC	1	0.352	0.074
M-B	0.318	1	0.013
M-A	0.008	0.014	1

(d) Dynamics of Average shortest path

	HC	M-B	M-A
HC	1	0.566	0.053
M-B	0.608	1	0.061
M-A	0.051	0.059	1

(e) Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.769	0.358
M-B	0.811	1	0.387
M-A	0.291	0.385	1

Table 3.4 *pvalues for DMN Sub-Network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.725	0.101
M-B	0.801	1	0.116
M-A	0.129	0.123	1

(b) pvalues of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.519	0.373
M-B	0.552	1	0.795
M-A	0.343	0.806	1

(c) pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.745	0.134
M-B	0.752	1	0.036
M-A	0.108	0.034	1

(d) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.441	0.176
M-B	0.456	1	0.03
M-A	0.147	0.031	1

(e) pvalues of Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.341	0.11
M-B	0.35	1	0.388
M-A	0.109	0.381	1

Table 3.5 *pvalues for Sensorimotor Sub-Network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.376	0.212
M-B	0.407	1	0.053
M-A	0.229	0.056	1

(b) pvalues of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.316	0.036
M-B	0.266	1	0.001
M-A	0.044	0.001	1

(c) pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.745	0.509
M-B	0.752	1	0.249
M-A	0.536	0.241	1

(d) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.217	0.026
M-B	0.195	1	0.333
M-A	0.027	0.36	1

(e) pvalues of Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.691	0.13
M-B	0.652	1	0.286
M-A	0.123	0.3	1

Table 3.6 *pvalues for Visual Sub-network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.507	0.047
M-B	0.494	1	0.117
M-A	0.044	0.123	1

(b) pvalues of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.442	0.081
M-B	0.462	1	0.238
M-A	0.094	0.221	1

(c) pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.569	0.056
M-B	0.51	1	0.172
M-A	0.059	0.171	1

(d) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.424	0.8
M-B	0.448	1	0.332
M-A	0.773	0.318	1

(e) pvalues of Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.175	0.654
M-B	0.199	1	0.374
M-A	0.666	0.35	1

Table 3.7 *pvalues for Cerebellum Sub-network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.647	0.464
M-B	0.624	1	0.805
M-A	0.475	0.844	1

(b) pvalues of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.151	0.222
M-B	0.177	1	0.873
M-A	0.178	0.879	1

(c) cerebellum Network, pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.611	0.767
M-B	0.667	1	0.745
M-A	0.776	0.712	1

(d) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.742	0.539
M-B	0.822	1	0.484
M-A	0.567	0.474	1

(e) pvalues of Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.399	0.604
M-B	0.378	1	0.658
M-A	0.613	0.643	1

Table 3.8 *pvalues for Cognitive Control Sub-network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.652	0.786
M-B	0.608	1	0.51
M-A	0.774	0.514	1

(b) pvalues of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.541	0.7
M-B	0.503	1	0.307
M-A	0.721	0.301	1

(c) pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.706	0.735
M-B	0.695	1	0.861
M-A	0.752	0.864	1

(d) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.71	0.389
M-B	0.713	1	0.272
M-A	0.416	0.271	1

(e) pvalues of Dynamics of Average Clustering Coefficient

	HC	M-B	M-A
HC	1	0.59	0.358
M-B	0.637	1	0.66
M-A	0.32	0.688	1

Table 3.9 *pvalues for Subcortical Sub-network*

(a) pvalues of Average Weight of MST

	HC	M-B	M-A
HC	1	0.667	0.541
M-B	0.692	1	0.338
M-A	0.554	0.343	1

(b) pvalues of Average Shortest Path

	HC	M-B	M-A
HC	1	0.613	0.512
M-B	0.726	1	0.345
M-A	0.558	0.362	1

(c) pvalues of Dynamics of Average Shortest Path

	HC	M-B	M-A
HC	1	0.716	0.714
M-B	0.696	1	0.865
M-A	0.766	0.875	1

Table 3.10 *State Analysis: Rank and Time Consumption of Each State*

	A	C	CC	D	SC	SM	V	ACC	ASP	HC	M-A	M-B
state1	4	3	4	4	3	4	4	5	4	16.4%	35.0%	13.6%
state2	2	2	2	2	2	2	2	3	2	32.0%	34.4%	35.7 %
state3	5	5	5	5	5	5	5	2	5	11.5%	5.3 %	11.3 %
state4	3	4	3	3	4	3	3	1	3	22.0%	15.3%	22.0%
state5	1	1	1	1	1	1	1	4	1	18.0%	10.0%	17.4%

Bibliography

[1] de la Iglesia-Vaya, Maria, et al. *Brain Connections-Resting State FMRI Functional Connectivity.* INTECH Open Access Publisher, 2013.

[2] Kass, Robert E.; Raftery, Adrian E. (1995), *Bayes Factors*, Journal of the American Statistical Association 90 (430): 773795, doi:10.2307/2291091, ISSN 0162-1459.

[3] Lehmann, Dietrich. *Brain electric microstates and cognition: the atoms of thought.*Machinery of the Mind. Birkhuser Boston, 1990. 209-224.

[4] Allen, Elena A., et al. *Tracking whole-brain connectivity dynamics in the resting state.* Cerebral cortex (2012): bhs352.

[5] Damaraju, E., et al. *Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia.* NeuroImage: Clinical 5 (2014): 298-308.

[6] Catena, Robert D., et al. *Spatial orientation of attention and obstacle avoidance following concussion.* Experimental brain research 194.1 (2009): 67-77.

[7] Brosseau-Lachaine, Odile, et al. *Mild traumatic brain injury induces prolonged visual processing deficits in children.* Brain Injury 22.9 (2008): 657-668.

[8] Risen, Sarah, et al. *Increased Connectivity between the Sensorimotor Cortex and Dorsal Attention Network in Children after Mild to Moderate Traumatic Brain Injury (S11. 006).* Neurology 82.10 Supplement (2014): S11-006.

[9] Mayer, Andrew R., et al. *Auditory orienting and inhibition of return in mild traumatic brain injury: a FMRI study.* Human brain mapping 30.12 (2009): 4152-4166.

[10] Chan, R. C. K., Shum, D., Toulopoulou, T. & Chen, E. Y. H., R; Shum, D; Toulopoulou, T; Chen, E (2008). *Assessment of executive functions: Review of instruments and identification of critical issues.* Archives of Clinical Neuropsychology. 2 23 (2): 201216. doi:10.1016/j.acn.2007.08.010. PMID 18096360.

[11] Li Yo, Adali T, Calhoun VD (2007): *Estimating the number of in- dependent components for functional magnetic resonance imaging data.* Human Brain Mapping 28:12511266.

[12] Schlotzhauer, Sandra (2007). *Elementary Statistics Using JMP* (SAS Press) (PAP/CDR ed.). Cary, NC: SAS Institute. pp. 166169. ISBN 1-599-94375-1.

[13] Bronkhorst, Adelbert W. (2000). *The Cocktail Party Phenomenon: A Review on Speech Intelligibility in Multiple-Talker Conditions* Acta Acustica united with Acustica 86: 117128. Retrieved 2010-04-18.

[14] *Scikit-learn: Machine Learning in Python*, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[15] Moran, JM, Kelley, WM, and Heatherton, TF. (2013, Jul 17). *What can the organization of the brains default mode network tell us about self-knowledge?* Frontiers in Human Neuroscience; 7:391. doi: 10.3389/fnhum.2013.00391

[16] Tang, Cheuk Ying, et al. *Diffuse disconnectivity in traumatic brain injury: a resting state fMRI and DTI study.* Translational neuroscience 3.1 (2012): 9-14.

[17] Larson, Michael J., Thomas J. Farrer, and Peter E. Clayson. *Cognitive control in mild traumatic brain injury: Conflict monitoring and conflict adaptation.* International Journal of Psychophysiology 82.1 (2011): 69-78.

[18] Tinius, Timothy P. *The integrated visual and auditory continuous performance test as a neuropsychological measure.* Archives of clinical Neuropsychology 18.5 (2003): 439-454.

[19] Gentilini, M., Nichelli, P., & Schoenhuber, R. (1989). *Assessment of attention in mild head injury.* In Mild head injury (pp. 163175). New York: Oxford University Press.

[20] R. Parasuraman, S.A. Mutter, R. Molloy *Sustained attention following mild closed head injury.* Journal of Clinical and Experimental Neuropsychology, 13 (5) (1991), pp. 789811

[21] L.F. Collins, C.J. Long *Visual reaction time and its relationship to neuropsychological test performance.* Archives of Clinical Neuropsychology, 11 (7) (1996), pp. 613623

[22] D.T. Stuss, L.L. Stethem, H. Hugenholtz, T. Picton, J. Pivik, M.T. Richard *Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance* Journal of Neurology, Neurosurgery and Psychiatry, 52 (6) (1989), pp. 742748

[23] ODonnell, W. E., DeSoto, C. B., DeSota, J. L., & Reynolds, D. McQ. (1994). *Neuropsychological Impairment Scales (NIS) Manual.* Los Angeles, CA: Western Psychological Services.

[24] E. Arcia, C.T. Gualtieri *Neurobehavioural performance of adults with closed-head injury, adults with attention deficit, and controls* Brain Injury, 8 (5) (1994), pp. 395404

[25] Andrei A. Vakhtin, Vince D. Calhoun, Rex E. Jung, Jillian L. Prestopnik, Paul A. Taylor, Corey C. Ford *Changes in intrinsic functional brain networks following blast-induced mild traumatic brain injury.* Brain Inj. 2013; 27(11): 13041310. doi: 10.3109/02699052.2013.823561

[26] Dobkin, Bruce H. *Motor rehabilitation after stroke, traumatic brain, and spinal cord injury: common denominators within recent clinical trials.* Current opinion in neurology 22.6 (2009): 563-569.

[27] Henninger, Nils, et al. *Differential recovery of behavioral status and brain function assessed with functional magnetic resonance imaging after mild traumatic brain injury in the rat.* Critical care medicine 35.11 (2007): 2607-2614.

[28] McAllister, Thomas W. *Mild brain injury and the postconcussion syndrome.* Textbook of traumatic brain injury. Washington, DC, US: American Psychiatric Publishing, Inc (2005): 279-308.

[29] King, Nigel S. *Post-concussion syndrome: clarity amid the controversy?.* The British Journal of Psychiatry 183.4 (2003): 276-278.

[30] McNett, Molly. *A review of the predictive ability of Glasgow Coma Scale scores in head-injured patients.* Journal of neuroscience nursing 39.2 (2007): 68-75.

[31] McAllister, Thomas W., et al. *Mechanisms of working memory dysfunction after mild and moderate TBI: evidence from functional MRI and neurogenetics.* Journal of neurotrauma 23.10 (2006): 1450-1467.

[32] Hyvrinen, Aapo, Juha Karhunen, and Erkki Oja. *Independent component analysis.* Vol. 46. John Wiley & Sons, 2004.

[33] McKeown, Martin J., and Terrence J. Sejnowski. *Independent component analysis of fMRI data: examining the assumptions.* Human brain mapping 6.5-6 (1998): 368-372.

[34] Laurienti, Paul J., et al. *Relationship between caffeine-induced changes in resting cerebral perfusion and blood oxygenation level-dependent signal.* American journal of neuroradiology 24.8 (2003): 1607-1611.

[35] Bowen AP (2003). *Second impact syndrome: A rare, catastrophic, preventable complication of concussion in young athletes.* Journal of Emergency Nursing 29 (3): 2879.

[36] Cantu RC (2001). *Posttraumatic Retrograde and Anterograde Amnesia: Pathophysiology and Implications in Grading and Safe Return to Play.* Journal of Athletic Training 36 (3): 2448.

[37] Rees PM (2003). *Contemporary issues in mild traumatic brain injury.* Archives of Physical Medicine and Rehabilitation 84 (12): 188594.

[38] Erlanger DM, Kutner KC, Barth JT, Barnes R (1999). *Neuropsychology of sports-related head enjury: Dementia pugilistica to post concussion syndrome.* The Clinical Neuropsychologist 13 (2): 193209.

[39] McCrory PR, Berkovic SF (1998). *Concussive convulsions. Incidence in sport and treatment recommendations.* Sports Medicine 25 (2): 1316.

[40] Perron AD, Brady WJ, Huff JS (2001). *Concussive convulsions: Emergency department assessment and management of a frequently misunderstood entity.* Academic Emergency Medicine 8 (3): 2968.

[41] Mayer, Andrew R., et al. *Functional connectivity in mild traumatic brain injury.* Human brain mapping 32.11 (2011): 1825-1835.

[42] Zhou, Yongxia, et al. *Default-mode network disruption in mild traumatic brain injury.* Radiology 265.3 (2012): 882-892.

[43] Johnson, Brian, et al. *Alteration of brain default network in subacute phase of injury in concussed individuals: resting-state fMRI study.* Neuroimage 59.1 (2012): 511-518.

[44] Yoshino, Atsuo, et al. *Dynamic changes in local cerebral glucose utilization following cerebral concussion in rats: evidence of a hyper-and subsequent hypometabolic state.* Brain research 561.1 (1991): 106-119.

[45] Giza, Christopher C., and David A. Hovda. *The neurometabolic cascade of concussion.* Journal of athletic training 36.3 (2001): 228.

[46] Aubry, Mark, et al. *Summary and agreement statement of the first International Conference on Concussion in Sport, Vienna 2001.* British journal of sports medicine 36.1 (2002): 6-7.

[47] McCrory, Paul, et al. *Consensus statement on Concussion in Sportthe 3rd International Conference on Concussion in Sport held in Zurich, November 2008.* South African Journal of Sports Medicine 21.2 (2009).

[48] Kushner D (1998). *Mild Traumatic brain injury: Toward understanding manifestations and treatment.* Archives of Internal Medicine 158 (15): 16171624.

[49] Parkinson D (1999). *Concussion confusion.* Critical Reviews in Neurosurgery 9 (6): 335339.

[50] Shaw NA (2002). *The neurophysiology of concussion.* Progress in Neurobiology 67 (4): 281344.

[51] Mouzon, B; Bachmeier, C (February 2014). *Chronic neuropathological and neurobehavioral changes in a repetitive mild traumatic brain injury model.* Ann Neurol. 75 (2): 241254.

[52] From American Academy of Pediatrics, Clinical Report, *Sport-Related Concussion in Children and Adolescents* Pediatrics, Mark E. Halstead, MD, Kevin D. Walter, MD, The Council on Sports Medicine and Fitness, Vol. 126 No. 3, September 1, 2010.

[53] Iverson GL (2005). *Outcome from mild traumatic brain injury* Current Opinion in Psychiatry 18 (3): 30117.

[54] Bowen AP (2003). *Second impact syndrome: A rare, catastrophic, preventable complication of concussion in young athletes.* Journal of Emergency Nursing 29 (3): 2879.

[55] Hardman JM, Manoukian A (2002). *Pathology of head trauma.* Neuroimaging Clinics of North America 12 (2): 17587, vii.

[56] Borg J, Holm L, Cassidy JD, et al. (2004). *Diagnostic procedures in mild traumatic brain injury: Results of the WHO collaborating centre task force on mild traumatic brain injury* Journal of Rehabilitation Medicine 36 (Supplement 43): 6175.

[57] Teasdale G, Jennett B. (1974). *Assessment of coma and impaired consciousness. A practical scale.* Lancet 13 (2): 814.

[58] Song, Ming, et al. *Brain spontaneous functional connectivity and intelligence.* Neuroimage 41.3 (2008): 1168-1176.

[59] van den Heuvel, Martijn P., et al. *Efficiency of functional brain networks and intellectual performance.* The Journal of Neuroscience 29.23 (2009): 7619-7624.

[60] Andreasen, Nancy C., Sergio Paradiso, and Daniel S. OLeary. *'Cognitive dysmetria' as an integrative theory of schizophrenia.* Schizophrenia Bulletin 24.2 (1998): 203-218.

[61] Friston, K. J., et al. *Psychophysiological and modulatory interactions in neuroimaging.* Neuroimage 6.3 (1997): 218-229.

[62] Bleuler, E. (1911). *Dementia Praecox, or the Group of Schizophrenias* (translated 1950 by J. Zinkin). International Universities Press: New York

[63] Bluhm, Robyn L., et al. *Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: anomalies in the default network.* Schizophrenia bulletin 33.4 (2007): 1004-1012.

[64] Liu, Yong, et al. *Regional homogeneity, functional connectivity and imaging markers of Alzheimer's disease: a review of resting-state fMRI studies.* Neuropsychologia 46.6 (2008): 1648-1656.

[65] Dijkstra, Edsger W. *A note on two problems in connexion with graphs.* Numerische mathematik 1.1 (1959): 269-271.

[66] Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). *A Formal Basis for the Heuristic Determination of Minimum Cost Paths.* IEEE Transactions on Systems Science and Cybernetics SSC4 4 (2): 100107.

[67] Roy, Bernard (1959). *Transitivit et connexit.* C. R. Acad. Sci. Paris 249: 216218.

[68] Floyd, Robert W. (June 1962). *Algorithm 97: Shortest Path.* Communications of the ACM 5 (6): 345.

[69] Warshall, Stephen (January 1962). *A theorem on Boolean matrices.* Journal of the ACM 9 (1): 1112.

[70] Bullmore, Ed, and Olaf Sporns. *Complex brain networks: graph theoretical analysis of structural and functional systems.* Nature Reviews Neuroscience 10.3 (2009): 186-198.

[71] Stam, C. J., et al. *Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease.* Brain 132.1 (2009): 213-224.

[72] Sporns, Olaf, and Jonathan D. Zwi. *The small world of the cerebral cortex.* Neuroinformatics 2.2 (2004): 145-162.

[73] Reijneveld, Jaap C., et al. *The application of graph theoretical analysis to complex networks in the brain.* Clinical Neurophysiology 118.11 (2007): 2317-2331.

[74] Zhou, Yuan, et al. *Functional disintegration in paranoid schizophrenia using resting-state fMRI.* Schizophrenia research 97.1 (2007): 194-205.

[75] Vincent, J. L., et al. *Intrinsic functional architecture in the anaesthetized monkey brain.* Nature 447.7140 (2007): 83-86.

[76] Canli, Turhan, et al. *fMRI identifies a network of structures correlated with retention of positive and negative emotional memory.* Psychobiology 27.4 (1999): 441-452.

[77] Fransson, Peter, and Guillaume Marrelec. *The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis.* Neuroimage 42.3 (2008): 1178-1184.

[78] Smith, Stephen M., et al. *Network modelling methods for FMRI.* Neuroimage 54.2 (2011): 875-891.

[79] Liu, Peng, et al. *Partial correlation investigation on the default mode network involved in acupuncture: an fMRI study.* Neuroscience letters 462.3 (2009): 183-187.

[80] Achard, Sophie, et al. *A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs.* The Journal of Neuroscience 26.1 (2006): 63-72.

[81] Achard, Sophie, and Ed Bullmore. *Efficiency and cost of economical brain functional networks.* PLoS computational biology 3.2 (2007):e17.

[82] Grigorov, Martin G. *Global properties of biological networks.* Drug discovery today 10.5 (2005): 365-372.

[83] Watts, Duncan J., and Steven H. Strogatz. *Collective dynamics of small-worldnetworks.* nature 393.6684 (1998): 440-442.

[84] Bassett, Danielle Smith, and E. D. Bullmore. *Small-world brain networks.* The neuroscientist 12.6 (2006): 512-523.

[85] Micheloyannis, Sifis, et al. *Small-world networks and disturbed functional connectivity in schizophrenia.* Schizophrenia research 87.1 (2006): 60-66.

[86] Stam, C. J., et al. *Small-world networks and functional connectivity in Alzheimer's disease.* Cerebral cortex 17.1 (2007): 92-99.

[87] Buckner, Randy L., et al. *Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease.* The Journal of Neuroscience 29.6 (2009): 1860-1873.

[88] Buckner, Randy L., and Justin L. Vincent. *Unrest at rest: default activity and spontaneous network correlations.* Neuroimage 37.4 (2007): 1091-1096.

[89] Calhoun, V. D., et al. *A method for making group inferences from functional MRI data using independent component analysis.* Human brain mapping 14.3 (2001): 140-151.

[90] Thirion, Bertrand, et al. *Dealing with the shortcomings of spatial normalization: Multisubject parcellation of fMRI datasets.* Human brain mapping 27.8 (2006): 678-693.

[91] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith. *FSL.* NeuroImage, 62:782-90, 2012

[92] Friston, K. J., et al. *To smooth or not to smooth?: Bias and efficiency in fmri time-series analysis.* NeuroImage 12.2 (2000): 196-208.

[93] M.W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, S.M. Smith. *Bayesian analysis of neuroimaging data in FSL.* NeuroImage, 45:S173-86, 2009

[94] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and

P.M. Matthews. *Advances in functional and structural MR image analysis and implementation as FSL.* NeuroImage, 23(S1):208-19, 2004

[95] Churchill, Nathan W., Oder, Anita, Abdi, Herv, Tam, Fred, Lee, Wayne, Thomas, Christopher, Ween, Jon E., Graham, Simon J., Strother, Stephen C., *Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods* Human Brain Mapping 2012, Volume 33 Issue 3

[96] Sladky, Ronald, et al. *Slice-timing effects and their correction in functional MRI.* Neuroimage 58.2 (2011): 588-594.

[97] Torben E. Lund, Minna D. Nrgaard, Egill Rostrup, James B. Rowe, Olaf B. Paulson, *Motion or activity: their role in intra- and inter-subject variation in fMRI,* NeuroImage, Volume 26, Issue 3, 1 July 2005, Pages 960-964, ISSN 1053-8119

[98] Whitfield-Gabrieli, S., Thermenos, H. W., Milanovic, S., Tsuang, M. T., Faraone, S. V., McCarley, R. W., ... & Seidman, L. J. (2009). *Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia.* Proceedings of the National Academy of Sciences, 106(4),1279-1284.

[99] Tanabe, Jody, et al. *Comparison of detrending methods for optimal fMRI preprocessing.* NeuroImage 15.4 (2002): 902-907.

[100] Lindquist, Martin A. *The statistical analysis of fMRI data.* Statistical Science 23.4 (2008): 439-464.

[101] van den Heuvel, Martijn P., et al. *Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain.* Neuroimage 43.3 (2008): 528-539.

[102] Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ (February 2004). *Neural systems supporting interoceptive awareness* Nat. Neurosci. 7 (2): 18995.

[103] Kimberg DY, Farah MJ (December 1993). *A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex, organized behavior* Journal of Experimental Psychology. General 122 (4): 41128. doi:10.1037/0096-3445.122.4.411. PMID 8263463.

[104] Broyd, Samantha J.; Demanuele, Charmaine; Debener, Stefan; Helps, Suzannah K.; James, Christopher J.; Sonuga-Barke, Edmund J.S. (2009). *Default-mode brain dysfunction in mental disorders: A systematic review* Neuroscience

& Biobehavioral Reviews 33 (3): 27996. doi:10.1016/j.neubiorev.2008.09.002. PMID 18824195.

[105] Lamb K, Gallagher K, McColl R, Mathews D, Querry R, Williamson JW (April 2007). *Exercise-induced decrease in insular cortex rCBF during postexercise hypotension* Med Sci Sports Exerc 39 (4): 6729.

[106] Andrews-Hanna, Jessica R., et al. *Disruption of large-scale brain systems in advanced aging.* Neuron 56.5 (2007): 924-935.

[107] Larson-Prior, Linda J., et al. *Cortical network functional connectivity in the descent to sleep.* Proceedings of the National Academy of Sciences 106.11 (2009): 4489-4494.

[108] Lawes, I. Nigel C., et al. *Atlas-based segmentation of white matter tracts of the human brain using diffusion tensor tractography and comparison with classical dissection.* Neuroimage 39.1 (2008): 62-79.

[109] Greicius, Michael D., et al. *Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI.* Proceedings of the National Academy of Sciences of the United States of America 101.13 (2004): 4637-4642.

[110] Stefansson, Hreinn, et al. *Large recurrent microdeletions associated with schizophrenia.* nature 455.7210 (2008): 232-236.

[111] Williamson JW, McColl R, Mathews D, Mitchell JH, Raven PB, Morgan WP (April 2001). *Hypnotic manipulation of effort sense during dynamic exercise: cardiovascular responses and brain activation* J. Appl. Physiol. 90 (4): 13929. PMID 11247939.

[112] Williamson JW, McColl R, Mathews D, Ginsburg M, Mitchell JH (September 1999). *Activation of the insular cortex is affected by the intensity of exercise* J. Appl. Physiol. 87 (3): 12139. PMID 10484598.

[113] Raichle, Marcus E., et al. *A default mode of brain function.* Proceedings of the National Academy of Sciences 98.2 (2001): 676-682.

[114] Greicius, Michael D., et al. *Functional connectivity in the resting brain: a network analysis of the default mode hypothesis.* Proceedings of the National Academy of Sciences 100.1 (2003): 253-258.

[115] Beckmann, Christian F., et al. *Investigations into resting-state connectivity using independent component analysis.* Philosophical Transactions of the Royal Society B: Biological Sciences 360.1457 (2005): 1001-1013.

[116] De Luca, M., et al. *fMRI resting state networks define distinct modes of long-distance interactions in the human brain.* Neuroimage 29.4 (2006): 1359-1367.

[117] Smith, Stephen M., et al. *Correspondence of the brain's functional architecture during activation and rest.* Proceedings of the National Academy of Sciences 106.31 (2009): 13040-13045.

[118] Yeo, BT Thomas, et al. *The organization of the human cerebral cortex estimated by intrinsic functional connectivity.* Journal of neurophysiology 106.3 (2011): 1125-1165.

[119] Power, Jonathan D., et al. *Functional network organization of the human brain.* Neuron 72.4 (2011): 665-678.

[120] Van Den Heuvel, Martijn, Rene Mandl, and Hilleke Hulshoff Pol. *Normalized cut group clustering of resting-state FMRI data.* PloS one 3.4 (2008): e2001.

[121] Rykhlevskaia, Elena, Gabriele Gratton, and Monica Fabiani. *Combining structural and functional neuroimaging data for studying brain connectivity: a review.* Psychophysiology 45.2 (2008): 173-187.

[122] Bassett, Danielle S., and Edward T. Bullmore. *Human brain networks in health and disease.* Current opinion in neurology 22.4 (2009): 340.

[123] Fox, Michael D., et al. *The human brain is intrinsically organized into dynamic, anticorrelated functional networks.* Proceedings of the National Academy of Sciences of the United States of America 102.27 (2005): 9673-9678.

[124] Buckner, Randy L., Jessica R. Andrews-Hanna, and Daniel L. Schacter. *The brain's default network.* Annals of the New York Academy of Sciences 1124.1 (2008): 1-38.

[125] Supekar, Kaustubh, et al. *Network analysis of intrinsic functional brain connectivity in Alzheimer's disease.* PLoS computational biology 4.6 (2008): e1000100.

[126] Greicius, Michael. *Resting-state functional connectivity in neuropsychiatric disorders.* Current opinion in neurology 21.4 (2008): 424-430.

[127] Helps, Suzannah K., et al. *Altered spontaneous low frequency brain activity in attention deficit/hyperactivity disorder.* Brain research 1322 (2010): 134-143.

[128] Seeley, William W., et al. *Neurodegenerative diseases target large-scale human brain networks.* Neuron 62.1 (2009): 42-52.

[129] van den Heuvel, Martijn P., et al. *Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain.* Human brain mapping 30.10 (2009): 3127-3141.

[130] Fox, Michael D., and Marcus E. Raichle. *Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging.* Nature Reviews Neuroscience 8.9 (2007): 700-711.

[131] Kimberg, Daniel Y., Geoffrey K. Aguirre, and Mark DEsposito. *Modulation of task-related neural activity in task-switching: an fMRI study.* Cognitive Brain Research 10.1 (2000): 189-196.

[132] Dove, Anja, et al. *Prefrontal cortex activation in task switching: an event-related fMRI study.* Cognitive brain research 9.1 (2000): 103-109.

[133] Tomasi, Dardo, and Nora D. Volkow. *Resting functional connectivity of language networks: characterization and reproducibility.* Molecular psychiatry 17.8 (2012): 841-854.

[134] Rauch, Scott L., and Perry F. Renshaw. *Clinical neuroimaging in psychiatry.* Harvard review of psychiatry 2.6 (1995): 297-312.

[135] Filler, Aaron G. *The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI, and DTI.* Nature Precedings 7.1 (2009): 1-69.

[136] VanCott, Rachel. DTI Images. Digital image. Http://www.pbs.org. N.p., n.d. Web.

[137] Leuchter, Andrew F., et al. *Electroencephalographic spectra and coherence in the diagnosis of Alzheimer's-type and multi-infarct dementia: a pilot study.* Archives of General Psychiatry 44.11 (1987): 993-998.

[138] Shellock, F. G., S. Morisoli, and E. Kanal. *MR procedures and biomedical implants, materials, and devices: 1993 update.* Radiology 189.2 (1993): 587-599.

[139] Connelly, Aet al, et al. *Functional mapping of activated human primary cortex with a clinical MR imaging system.* Radiology 188.1 (1993): 125-130.

[140] Friston, Karl J., et al. *Event-related fMRI: characterizing differential responses.* Neuroimage 7.1 (1998): 30-40.

[141] Rombouts, Serge ARB, et al. *Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: an fMRI study.* Human brain mapping 26.4 (2005): 231-239.

[142] Wang, Liang, et al. *Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI.* Neuroimage 31.2 (2006): 496-504.

[143] Wang, Kun, et al. *Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study.* Human brain mapping 28.10 (2007): 967-978.

[144] Gur, Raquel E., et al. *An fMRI study of facial emotion processing in patients with schizophrenia.* American Journal of Psychiatry 159.12 (2002): 1992-1999.

[145] Demirci, Oguz, et al. *A review of challenges in the use of fMRI for disease classification/characterization and a projection pursuit application from a multi-site fMRI schizophrenia study.* Brain imaging and behavior 2.3 (2008): 207-226.

[146] Sours, Chandler, et al. *Default mode network interference in mild traumatic brain injuryA pilot resting state study.* Brain research 1537 (2013): 201-215.

[147] McAllister, Thomas W., et al. *Differential working memory load effects after mild traumatic brain injury.* Neuroimage 14.5 (2001): 1004-1012.

[148] McAllister, Thomas W., et al. *Brain activation during working memory 1 month after mild traumatic brain injury A functional MRI study.* Neurology 53.6 (1999): 1300-1300.

[149] Cordes, Dietmar, et al. *Frequencies contributing to functional connectivity in the cerebral cortex in resting-state data.* American Journal of Neuroradiology 22.7 (2001): 1326-1333.

[150] Cordes, Dietmar, et al. *Mapping functionally related regions of brain with functional connectivity MR imaging.* American Journal of Neuroradiology 21.9 (2000): 1636-1644.

[151] Calhoun, Vince D., Jingyu Liu, and Tlay Adal. *A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data.* Neuroimage 45.1 (2009): S163-S172.

[152] Kiviniemi, Vesa, et al. *Independent component analysis of nondeterministic fMRI signal sources.* Neuroimage 19.2 (2003): 253-260.

[153] Cherkassky, Vladimir L., et al. *Functional connectivity in a baseline resting-state network in autism.* Neuroreport 17.16 (2006): 1687-1690.

[154] Fransson, Peter. *Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis.* Human brain mapping 26.1 (2005): 15-29.

[155] Salvador, Raymond, et al. *Neurophysiological architecture of functional magnetic resonance images of human brain.* Cerebral cortex 15.9 (2005): 1332-1342.

[156] van de Ven, Vincent G., et al. *Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest.* Human brain mapping 22.3 (2004): 165-178.

[157] Orban, Guy A. *Higher order visual processing in macaque extrastriate cortex.* Physiological Reviews 88.1 (2008): 59-89.

[158] Blakemore, Sarah-Jayne, and Uta Frith. *The learning brain: Lessons for education.* Blackwell publishing, 2005.

[159] Lowe, M. J., B. J. Mock, and J. A. Sorenson. *Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations.* Neuroimage 7.2 (1998): 119-132.

[160] Birn, Rasmus M., et al. *Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI.* Neuroimage 31.4 (2006): 1536-1548.

[161] Birn, Rasmus M., Kevin Murphy, and Peter A. Bandettini. *The effect of respiration variations on independent component analysis results of resting state functional connectivity.* Human brain mapping 29.7 (2008): 740-750.

[162] Shmueli, Karin, et al. *Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal.* Neuroimage 38.2 (2007): 306-320.

[163] Chang, Catie, and Gary H. Glover. *Relationship between respiration, end-tidal CO2, and BOLD signals in resting-state fMRI.* Neuroimage 47.4 (2009): 1381-1393.

[164] Lee, M. H., C. D. Smyser, and J. S. Shimony. *Resting-state fMRI: a review of methods and clinical applications.* American Journal of Neuroradiology 34.10 (2013): 1866-1872.

[165] Damoiseaux, J. S., et al. *Consistent resting-state networks across healthy subjects.* Proceedings of the national academy of sciences 103.37 (2006): 13848-13853.

[166] Friston, Karl J. *Functional and effective connectivity: a review.* Brain connectivity 1.1 (2011): 13-36.

[167] Biswal, Bharat, et al. *Functional connectivity in the motor cortex of resting human brain using echo-planar mri.* Magnetic resonance in medicine 34.4 (1995): 537-541.

[168] Hutchison, R. Matthew, et al. *Dynamic functional connectivity: promise, issues, and interpretations.* Neuroimage 80 (2013): 360-378.

[169] Matthews, P. M., and P. Jezzard. *Functional magnetic resonance imaging.* Journal of Neurology, Neurosurgery & Psychiatry 75.1 (2004): 6-12.

[170] Ogawa, Seiji, et al. *Brain magnetic resonance imaging with contrast dependent on blood oxygenation.* Proceedings of the National Academy of Sciences 87.24 (1990): 9868-9872.

[171] Henry, Luke C., et al. *Metabolic changes in concussed American football players during the acute and chronic post-injury phases.* BMC neurology 11.1 (2011): 105.

[172] Jacobs, Bram, et al. *Outcome prediction in mild traumatic brain injury: age and clinical variables are stronger predictors than CT abnormalities.* Journal of neurotrauma 27.4 (2010): 655-668.

[173] Signorini, David F., et al. *Predicting survival using simple clinical variables: a case study in traumatic brain injury.* Journal of Neurology, Neurosurgery & Psychiatry 66.1 (1999): 20-25.

[174] Stulemeijer, Maja, et al. *Recovery from mild traumatic brain injury.* Journal of Neurology 253.8 (2006): 1041-1047.

[175] Thornhill, Sharon, et al. *Disability in young people and adults one year after head injury: prospective cohort study.* Bmj 320.7250 (2000): 1631-1635.

[176] Hoyer, Siegfried, and Claude Krier. *Ischemia and the aging brain. Studies on glucose and energy metabolism in rat cerebral cortex.* Neurobiology of aging 7.1 (1986): 23-29.

[177] Chtelat, Gal, et al. *Relationships between brain metabolism decrease in normal aging and changes in structural and functional connectivity.* Neuroimage 76 (2013): 167-177.

[178] Mosenthal, Anne C., et al. *The effect of age on functional outcome in mild traumatic brain injury: 6-month report of a prospective multicenter trial.* Journal of Trauma-Injury, Infection, and Critical Care 56.5 (2004): 1042-1048.

[179] Mayer, Andrew, et al. *Static and dynamic intrinsic connectivity following mild traumatic brain injury.* Journal of neurotrauma ja (2014).

[180] Craven, Peter, and Grace Wahba. *Smoothing noisy data with spline functions.* Numerische Mathematik 31.4 (1978): 377-403.

[181] Schwartzman, Armin, et al. *Empirical null and false discovery rate analysis in neuroimaging.* Neuroimage 44.1 (2009): 71-82.

[182] Hyvarinen, Aapo. *Fast ICA for noisy data using Gaussian moments.* Circuits and Systems, 1999. ISCAS'99. Proceedings of the 1999 IEEE International Symposium on. Vol. 5. IEEE, 1999.

[183] Varoquaux, Gal, et al. *A group model for stable multi-subject ICA on fMRI datasets.* Neuroimage 51.1 (2010): 288-299.

[184] Mei, Lin, et al. *Statistical shape modelling: How many modes should be retained?* Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on. IEEE, 2008.

[185] Himberg, Johan, Aapo Hyvrinen, and Fabrizio Esposito. *Validating the independent components of neuroimaging time series via clustering and visualization.* Neuroimage 22.3 (2004): 1214-1222.

[186] Kiviniemi, Vesa, et al. *Functional segmentation of the brain cortex using high model order group PICA.* Human brain mapping 30.12 (2009): 3865-3886.

[187] Li, Yi-Ou, Tlay Adal, and Vince D. Calhoun. *Estimating the number of independent components for functional magnetic resonance imaging data.* Human brain mapping 28.11 (2007): 1251-1266.