SEED-BASED CORRELATION ANALYSIS AND INSTANTANEOUS GLOBAL CORRELATION ANALYSIS FOR RESTING STATE FMRI

By

Charreau S. Bell

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

May 11, 2018

Nashville, Tennessee

Approved:

D. Mitchell Wilkes, Ph.D.

Zhaohua Ding, Ph.D.

Alan Peters, Ph.D.

Baxter Rogers, Ph.D.

Benoit Dawant, Ph.D.

To my parents,

and

And my siblings,

and

my most loyal and devoted friend, Chickpea.

ACKNOWLEDGEMENTS

I would first like to acknowledge and thank my advisor, Dr. Mitch Wilkes. I would also like to thank my committee members, Drs. Zhaohua Ding, Baxter Rogers, Alan Peters, and Benoit Dawant for their efforts in providing me direction, advice, and suggestions to complete this work.

I would also like to thank my labmate Nazirah Mohd Khairi for all of her assistance and support. I'd lastly like to acknowledge and thank my best friends Arnetta Bell and Kepra McBrayer, who have provided unending support as I worked through my dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION AND SIGNIFICANCE

Brain disorders have an increasingly poignant socioeconomic impact, and persons with mental illness and mental disorders, Alzheimer's disease, dementia, Parkinson's disease, and epilepsy are intensely affected by these disorders. The effect on those with mental illness is especially profound, with more than 50% of illnesses beginning by age 14, and with 75% manifesting by the age of 24. More than 43 million adults are affected by mental illness each year, which corresponds to about one in every five people, including anxiety disorders, major depression, bipolar disorder, and schizophrenia. In fact, depression is the leading cause of disability worldwide. The 10th leading cause of death in America is suicide, and 90% of those who commit suicide have underlying mental illness. 24% of state prisoners have had recent episodes with a mental health condition, and 26% of adults staying in homeless shelters are affected by mental illness. Mental illnesses cost the United States more than 193 billion dollars in lost earnings each year [6].

With the penetrating effect of mental illness and brain disorders on society and the increasing number of people that are being affected, it is essential to be able to effectively diagnose and treat people with these types of disorders. For people with mental illness, diagnostic challenges are complicated by the absense of a diagnostic test based on physiological signals to identify or differentiate mental disorders. Magnetic resonance imaging (MRI) presents a feasible method for providing neuroimaging data of the brain which may assist in correctly diagnosing patients. Structural MRIs are able to provide a primary view of the anatomy of the brain, whereas functional MRIs (fMRIs) are able to provide insight into the metabolic consumption of oxygen in the brain, which is linked to regions of effort exertion in the brain. The signal measured is the blood-oxygen-level dependent (BOLD)

signal. Because of the phenotypical differences in the presentation of a normal healthy brain versus one with an underlying disease pathology, it stands to reason that through processing neuroimages, differences in anatomical and behavioral characteristics can be identified.

Using schizophrenia as an example of the work that has been done in neuroscience and psychiatry for studying mental illness and brain disorders, there has been extensive use of sMRI for assessment of anatomical differences in the hippocampus and other brain cortical and subcortical regions including the parahippocampus and amygdala [7, 8, 9, 10, 11]. This analysis is generally performed by an expert, and requires manually identifying, tracing, and calculating the volume of ROIs to compare against a control population. Analyses of fMRI data have led to discoveries of functional connectivity networks, linking disparate anatomical regions through functional similarity. In schizophrenia, this has led to the observation of altered dynamic connectivity of brain networks including the default mode network, dorsal attention networks, and executive control networks [12].

An essential facet of research and science that cannot be under-emphasized is the usage of *tools* to process and analyze the data. The introduction of the general linear model to perform statistical parametric mapping (SPM) [13, 14], and the voxel-based morphometry (VBM) framework [15, 16] first introduced the ability to examine statistical differences across subjects, which dramatically improved efficiency, and allowed for rapid voxel-wise univariate statistical testing to identify differences, including t-tests, F-tests, and correlation analyses. Tools such as seed-based correlation analyses, graph theory, and independent components analysis (ICA) [17, 18, 19] have been used extensively to investigate the synchronizing behaviors in the brain and understand its functionality.

The development of novel tools is thus of overwhelming importance to science, and particular to the field of neuroscience, investigating the activity of the brain. The purpose of this work is to introduce and characterize two new methods for understanding the activity of the brain. The first method investigates the networks in the brain by scrutinizing the

behavior of the entirety of the brain relative to one voxel. This allows the disparate regions in the brain which perform similar activities or have similar functions to be identified. Additionally, the ability to perform Bayesian analyses due to the probabilistic nature of the approach is introduced. The second method investigates the brain during regions of high synchrony, or when many regions of the brain appear to be performing similar activities. These periods are investigated both temporally and spatially to delve into the movement of information in the brain. In this way, this work performs the function of introducing new tools to explain the behavior of the brain, which can be used to study and understand brain disorders to provide better diagnoses and treatments for those affected.

CHAPTER 2

BACKGROUND

Magnetic resonance imaging (MRI) is a powerful tool in neuroimaging for investigating the structure and functional behavior of the brain. Two related sub-modalities of MRI used to investigate brain activity are structural MRI (sMRI), which provides a single detailed anatomical 3D image of the brain, and functional MRI (fMRI) which shows the physiological behavior of the brain by measuring the blood oxygen-level dependent (BOLD) signal. In this imaging modality, a sequence of 3D images is obtained in time at a regular sampling interval (TR) for a given spatial resolution. Technological advances in MRI are enabling higher sampling frequencies combined with higher spatial resolution. With these advances, a clear interpretation of the BOLD signal is expected to demonstrate the functionality of brain regions under situations of rest as well as during tasks. The following sections will describe how fMRI signals are acquired and processed such that meaningful information regarding brain functionality may be obtained.

2.1 Basic overview of fMRI signal acquisition and image formation

Figure 2.1 shows the basic components of an fMRI scanner. At its most basic level, the scanner emits a static magnetic field generated by superconducting electromagnets which are cooled to temperatures near absolute zero. This allows large currents to be passed through the wire coils which compose the electromagnet with low resistance, generating high, stable magnetic field strengths. Common strengths of magnetic fields produced by MRI scanners are 1.5 Tesla (T), 3, 4, and 7T, although scanners up to 11T for humans and 24T for animals are currently in use. By comparison, the Earth's magnetic field strength is 0.00005T and large electromagnets used to lift cars are on the order of 1T [20].

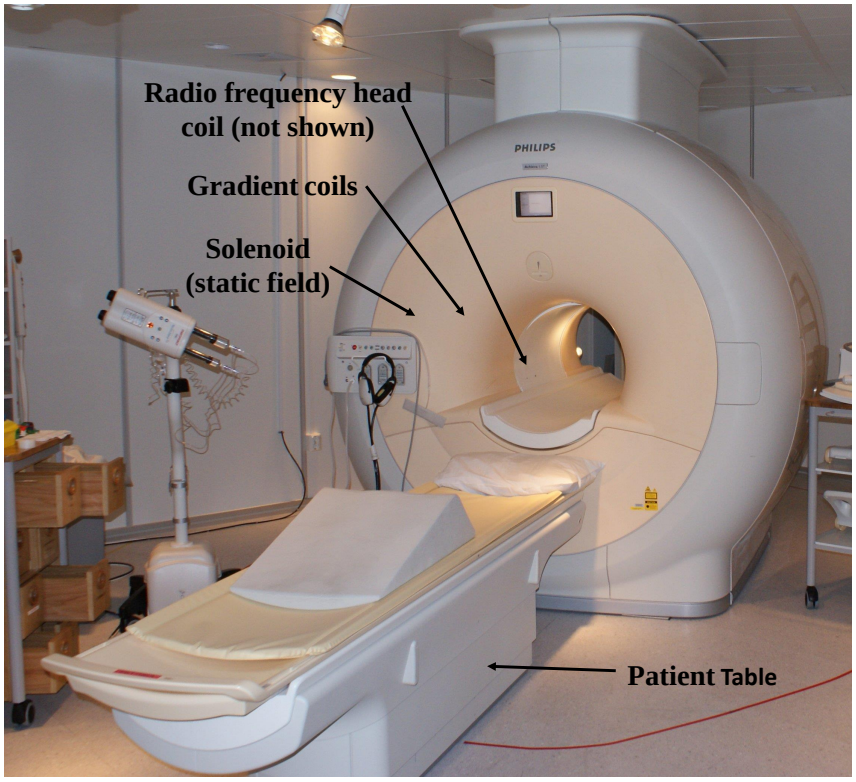


Figure 2.1: Basic components of fMRI scanner (head coil not shown) [1]

The scanner additionally includes a set of gradient coils, usually arranged to cause slight perturbations in the static magnetic field along the x, y, and z directions. The shimming coils are used to correct for small inhomogeneities in the static magnetic field and is set per-subject at the commencement of scanning. The radio frequency (RF) head coils are used to excite the atomic nuclei within the tissue given the static magnetic field. Since most atomic nuclei of interest for fMRI studies have resonant frequencies in the RF spectrum, RF signals can be used to perturb the nuclei from its equilibrium state to an excited states. Once the RF signal is removed, these fMRI coils located directly around the head can measure the desired relaxation signal.

MRI is based on detecting characteristics of nuclei possessing the nuclear magnetic resonance property (NMR) property, which means that the nucleus of the particle has both a magnetic moment and angular momentum. The human body contains many particles with the NMR property, particularly hydrogen due to the high concentration of water molecules (H_2O) in the body. When placed in a strong external magnetic field usually denoted $\mathbf{B_0}$, for example an MRI scanner, although all the particles spin axes will not be in exactly the same orientation, the net magnetization ($\mathbf{M}$) will align along $\mathbf{B_0}$. The protons (i.e., nucleus of hydrogen atom) will initiate precession about the magnetic axis at a frequency known as the Larmor frequency which is based on the strength of the external magnetic field.

Figures 2.2 A, B, and C show the excitation and relaxation steps necessary to generate a T1 signal. Figure 2.2A shows an initial spin field of protons spinning about $\mathbf{B_0}$ with a $\mathbf{M}$ along the same longitudinal axis. Excitation via a 90° excitation radio frequency pulse at the nuclei's Larmor frequency causes $\mathbf{M}$ to tip into the transverse plane as more spins accumulate at a higher energy state as shown in Figure 2.2B. If the T1 measurement time is desired (i.e., the time required for the longitudinal component of the $\mathbf{M}$ to return to its low energy state shown in A), this can be measured by receiver coils.

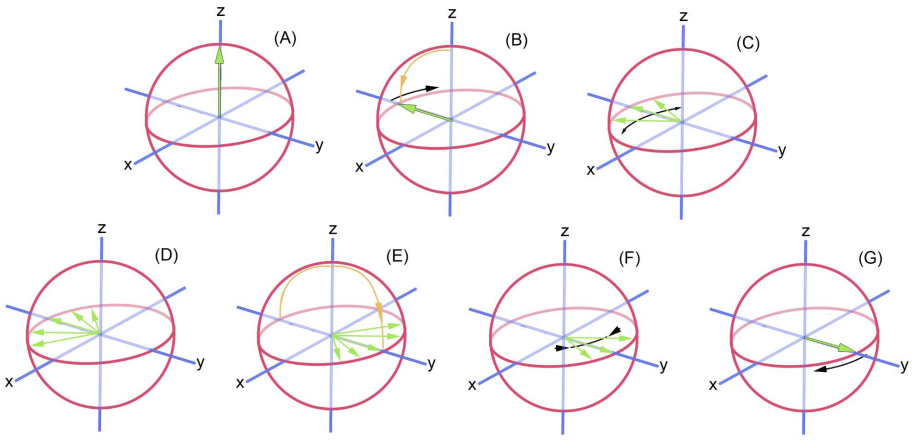


Figure 2.2: Excitation and relaxation of spin echo pulse sequence [2]

Figures 2.2 C and D show transverse relaxation of **M** due to spin-spin interactions and loss of coherence of the particles. This is referred to as the T2 decay or the T2 relaxation time. Additionally, inhomogeneities in the magnetic field lead to spatial differences in the spins of the protons which leads to a loss of coherence. The signal due to the combined effects of the spin-spin interactions and the incoherence in spins due to the magnetic field is known as T2* decay, which is always faster than T2 decay alone. The T2* signal is associated with the BOLD signal, and reflects how the blood oxygenation of various regions of the brain vary depending on usage.

Figures 2.2 E, F, and G show the effects of a specialized 180° refocusing pulse which reverses the incoherence due to lack of field homogeneity. Together, Figures 2.2 A-G show an entire spin-echo pulse sequence. A pulse sequence is a pattern of changing magnetic field gradients and electromagnetic fields that allow an image to be constructed from MR signals. Constructing the 3D image occurs in 2 major steps: slice selection to isolate a 2D slice from the brain volume, and frequency and phase encoding to identify MR signals from unique voxels in the slice. Other types of pulse sequences used in MR imaging include gradient-echo, magnetization-prepared rapid gradient echo (MPRAGE), and echo planar (EPI) imaging.

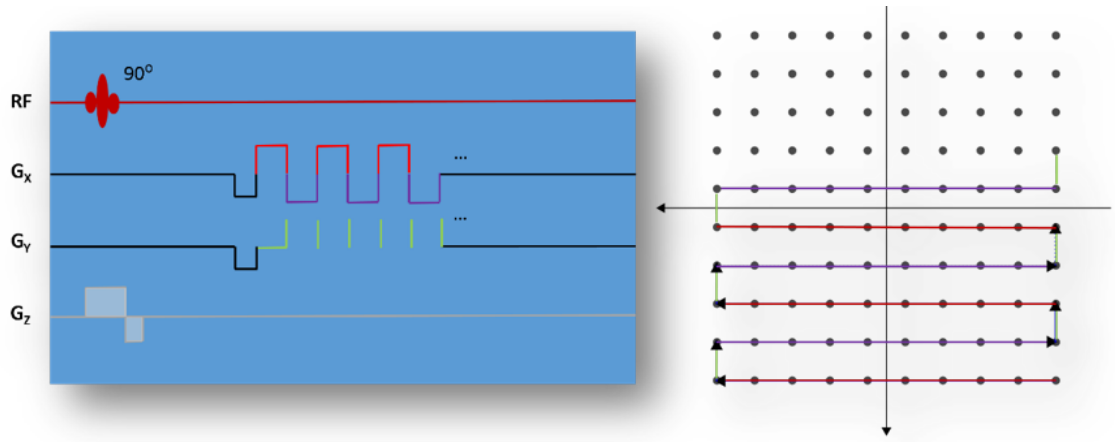


Figure 2.3: EPI pulse sequence timing diagram and sampling of k-space [3]

Due to the increased speed of sampling, EPI pulse sequences are a popular method of acquiring fMRI images. Figure 2.3 shows the timing diagram corresponding to k-space sampling using the EPI pulse sequence. As shown, a 90° excitation pulse is applied to the sample to tip the magnetization into the transverse plane. This is physically executed by manipulating the scanner's G_z gradient coils (i.e., gradient along the z-axis). Following this step, k-space is then sampled by transversing its perpendicular axes. As shown, the G_x gradient can be applied for a time to sample in the horizontal direction, and then the gradient G_y can be applied to move vertically to the next horizontal line of sampling. The G_x gradient can then be applied in the opposite direction in order to sample along the reverse direction of the original orientation. This mechanism of sampling leads to fast acquisition time, as only one 90° excitation pulse is necessary to collect an entire slice.

Unfortunately, the speed comes at the cost of post-processing to realign the data due to the way in which EPI images are acquired, EPI artifacts from imperfections in the magnetic field, and geometric distortions. Boundaries between air-filled cavities and tissue types are particularly responsible for these susceptibility artifacts. Magnetic field maps can aid in compensating for these inhomogeneities.

2.2 Physiological basis of fMRI

The BOLD signal has been shown to vary with relation to the metabolic consumption of oxygen due to brain activity [21]. The mechanism for this lies in the differing magnetic qualities of oxygenated and deoxygenated hemoglobin which are diamagnetic and paramagnetic, respectively. A number of experiments were required to uncover the relation between the BOLD signal and de/oxygenated hemoglobin. A first complication in the measurement of brain activity from fMRI is that the measurement must be indirect; there must be a relation between neuronal activity and oxygenated/deoxygenated hemoglobin. This pathway is through blood flow due to increased metabolic activity in the brain.

The idea that differences in blood oxygenation could be measured by the T2*-weighted contrast was demonstrated by a study on rats that breathed air with varying levels of oxygenation and through an ex-vivo experiment with blood vials in a saline casing. Both experiments showed that deoxygenated blood decreases the MR signal in T2*-weighted images.

Figure 2.4 shows an example of a T2*-weighted image, resulting in the BOLD contrast imaging of the brain slice shown. As can be seen in the image, there are regions of lower signal which correspond to the darker regions of the image. These correspond to where there are significant inhomogeneities due to deoxygenation of the blood in these regions. There are additionally the lighter areas of the image, which correspond to where there is more oxygenated blood and less iron to distort the sampling of the magnetic field in this region.

A reasonable hypothesis would be that since deoxygenated blood decreases the T2*-weighted MR signal, one would expect for period of high brain activity (high oxygen consumption and hypothesized rapid deoxygenation) to correspond to decreased MR signal, whereas low brain activity (low oxygen consumption and hypothesized low deoxygenation) to correspond to an increased MR signal. On the contrary, an influential experiment showed that increasing brain activity (e.g., selectively activating a region in an animal brain) in-

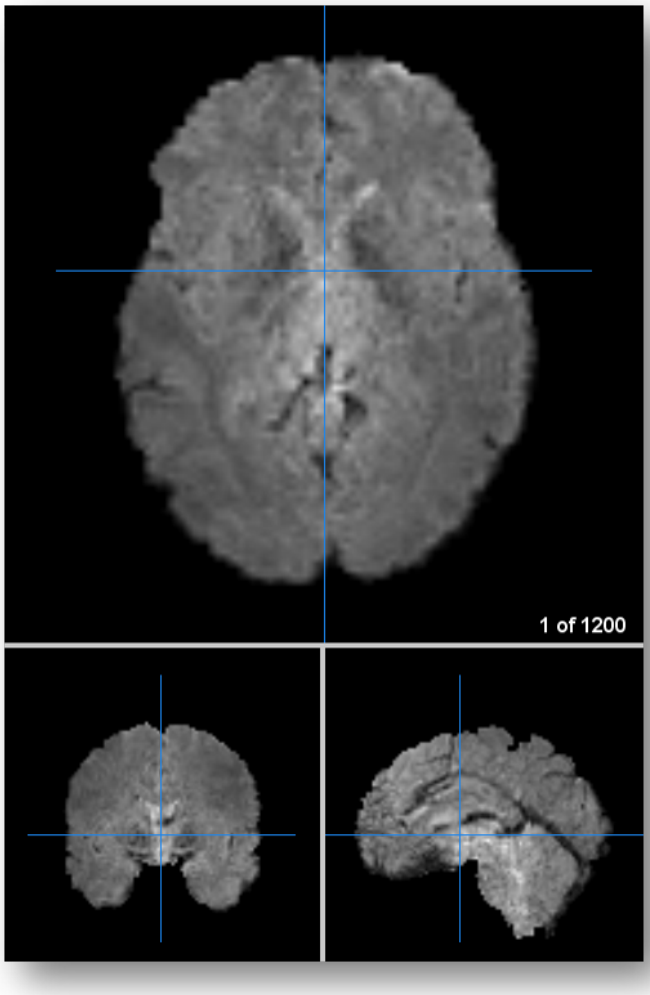


Figure 2.4: An example T2*-weighted BOLD contrast image

creased the MR signal (corresponding to decreased deoxygenated hemoglobin) [22]. The conclusions drawn from this experiment are numerous, the most applicable being that the activation of brain activity results in hyperperfusion of oxygenated blood to brain to the area; this exceeds the amount that is necessary to carry out the task. Thus, an increase in brain activity can be extracted from the T2*-weighted fMRI since the MR signal will increase due to increased oxygenation. During periods of decreased activity, there will not be a hyperperfusion of oxygenated hemoglobin and more deoxygenated hemoglobin will persist, leading to a decreased MR signal for that region.

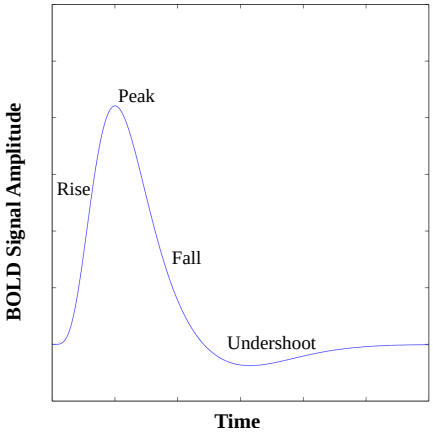


Figure 2.5: An archetypal hemodynamic response function [4]

Figure 2.5 shows a characteristic response curve for the hemodynamic response function (HDR). The HDR is the resultant BOLD signal response due to an input impulse; for an experiment, this is usually due to a task-based activation. The HDR as shown in Figure 2.5 lacks a region known as the *initial dip*, which is the subject of some controversy. The initial dip is marked by a pronounced reduction in the MR signal approximately 1-2 seconds after a stimulus; it has been attributed to increased localized uptake of oxygen due to neuronal activity, although it is usually only detected with sufficient averaging in lower magnetic fields and under high magnetic field strengths [23, 20]. Regardless, the rest of the HDR signal shows a distinct behavior of reaching a peak approximately 5 seconds after the onset

of stimulus. If the stimulus is maintained, the HDR will reach a plateau at a value slightly lower than the peak. Finally, after the stimulus is removed, the HDR will experience what is known as undershoot, for reasons not fully agreed upon including differing amounts of blood flow versus blood volume [24] or changes in oxygen metabolism [25], and then will again reach the initial baseline. The HDR imposes limits on spatial and temporal resolution due to the physiology behind signal acquisition [26].

2.3 Common preprocessing steps

A number of preprocessing steps are commonly applied to ameliorate inconsistencies introduced by the pulse sequence and from various sources of noise. A popular pulse sequence method is through interleaved slice acquisition, where non-sequential slices are imaged sequentially. For example, instead of slices 1-5 being acquired in that order, slices 1,3,and 5 might be acquired first, and then 2 and 4. This reduces cross-slice excitation. However, this means that the BOLD time courses in adjacent slices are misaligned. A temporal interpolation is usually performed to realign the time points.

An additional source of signal noise is through head motion. Small rotations or translations in the alignment of a participant's head can cause misalignments in voxels in the brain. This reduces the accuracy of results and if great enough in magnitude, can render a participant's scan wholly unuseable. Many different methods are employed to prevent the occurrence of head motion, including bite bars, special dental molds, or individualized masks, but specialized algorithms for estimating rotational and translational components of head motion are often applied in practice.

Distortions in the magnetic field and field inhomogeneities can also cause signal loss or intensity variations within the image. Distortions in the magnetic field are usually avoided by using shims, which locally change problems in the magnetic field. However, a common

preprocessing step is bias field estimation, where a map of intensity variations over space are determined and these can then be corrected in the data.

fMRI scans are additionally usually co-registered to a structural scan through the estimation of translational and rotational components. To facilitate the comparison of multiple subjects, scans are usually normalized to a common space such as Talaraich space or Montreal Neurological Institute (MNI) space. There are numerous methods to achieving this normalization through surface mapping, estimation of rigid body and affine transformation components, and warping, and many software packages available that provide this step of preprocessing.

The data are furthermore filtered, usually both temporally and spatially. Temporal filtering is used mostly to remove physiological noise; heart rates usually occur at 1.0 to 1.5 Hz and the rate of respiration at approximately 0.2 to 0.3 Hz [20]. The maximum frequency that can be accurately sampled is half the sampling rate, which is dependent on the scanner. However, depending on the frequency range of the data that is desired, a high, low, or bandpass filter can be used to capture the desired frequency region. Spatial filtering is also utilized in order to raise the signal to noise ratio (SNR) and improve the power of statistical tests usually applied to interpret the data. Gaussian filters of size 6-10mm full width at half maximum (FWHM) are usually employed for spatially fitering the data [20].

2.4 Resting-state fMRI

Analysis of the brain in the resting state fMRI (rs-fMRI) has been shown to be very effective in revealing mechanisms of brain activity [27]. A resting state scan of the brain is performed while the subject is lying in an MRI scanner, either with their eyes closed or open, and not engaged in any particular task [28]. Several studies have demonstrated that the brain is very active while a subject is at rest, and disparate regions of the brain exhibit high correlations and coactivations, particularly in the frequency spectrum at <0.1

Hz [29, 30]. Some research still indicates that a significant portion of the signal in this region comes from head motion and aliased respiratory and heart rate artifacts [31].

Additionally, a number of networks have been consistently identified with high inter-network connectivity. These networks are known as the resting state networks (RSNs), and have been identified by different groups [32, 29, 33, 34, 35, 36], although sometimes the networks somewhat vary [28, 37]. RSNs are also referred to as intrinsic connectivity networks (ICNs) to correct the notion that the brain is ever at rest in a living being [38]. RSNs persist during all times in the human brain [39], and task-based paradigms serve only to highlight specific arrangements of these networks [40]. These networks include the primary motor, primary visual, extra-striate visual, insular-temporal/anterior cingulate cortex (ACC), left and right parietal frontal, frontal, and default mode networks [28]. A commonly used parcellation of the RSNs into 7 coarse regions and 17 fine regions by clustering can be found in Yeo, 2011 [5]. The 7 RSNs into which the brain is parcellated are visual, somatomotor, dorsal attention, ventral attention, limbic, frontoparietal, and default mode network, and shown in Figure 2.6. Additionally, several disease pathologies have been linked to abnormalities in the resting state behavior of the brain [40, 41, 42, 43, 44].

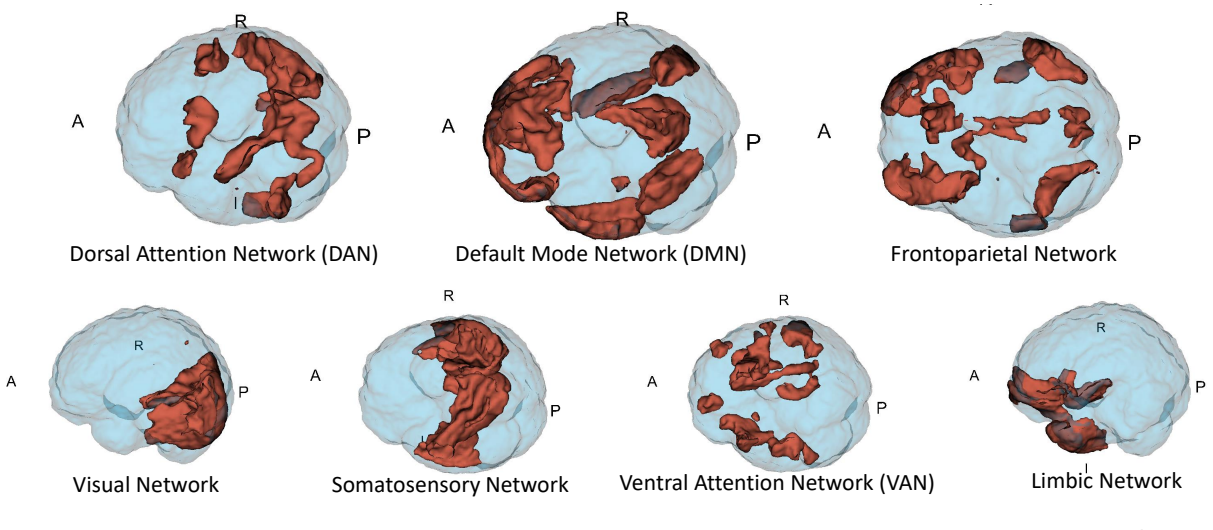


Figure 2.6: 7 resting state networks as identified by Yeo, 2011 [5]

2.5 Functional connectivity computation

Functional connectivity (FC) has emerged as a popular tool in investigating the spatially disparate, temporally related dynamics of brain functionality. FC is the temporal co-activation behavior of spatially disparate regions of the brain [13] and can be observed in the BOLD signal behavior during the resting state [45]. Popular methods include performing seed-based and whole-brain correlation analyses [29]. These reveal the similarities in brain signals between a seed region and all other brain regions, and the correlation of all brain regions with all other brain regions. Alternative methods include data-driven methods such as principal components analysis (PCA) [46], clustering [47, 48, 49, 50] and other graph-based connectivity methods [51, 28, 52, 53] have also been used to explore FC in the brain. Other voxel-wise measures of connectivity have been proposed via identification by point processes [54]. Probabilistic approaches employing Bayesian methods have also been proposed to offer further analysis into the mechanisms of FC.

2.5.1 Dynamic functional connectivity

Recent studies have demonstrated that spontaneous brain activity appears to violate assumptions of stationarity [55, 56, 57, 58], and that FC between brain regions fluctuates meaningfully over several scales in time. Observed scales of this time-varying behavior include from milliseconds (performed by other imaging modalities such as electroencephalogram (EEG), magnetoencephalogram (MEG) [59], or ultrafast MR encephalography (MREG) [60], to seconds and minutes [55, 61], to even over months and years [62, 54]. This observation has led to a subfield of FC coined *dynamic functional connectivity* (dFC). dFC obtained from resting state scans has been popular due to the unconstrained wandering of the mind during fMRI acquisition [63].

Several methods are commonly used to study dFC. The most popular method is through sliding window analysis, where a temporal window size is selected, and any of the many methods for evaluating FC is performed on this identified window. Following analysis on the initial window span, the window is shifted by a desired step length, and the analysis is performed again. This is repeated to the end of the timecourse. Sliding window analysis is advantageous due to its simplicity in implementation, its approach allowing arbitrary selection of the length of time that is desired to be studied (i.e., window length), and its efficacy in revealing dFC features [61]. However, the sliding window analysis technique is limited due to the difficulty of determining window sizes although some analytical guidelines have been derived [64], and inability to resolve very complex multi-frequency interactions [61]. Additionally, depending on the activity of the brain, sliding window analysis can still cause unwanted averaging over periods with very different levels of co-activations.

CHAPTER 3

BAYESIAN METHODS FOR SEED-BASED CORRELATION ANALYSIS

Abstract

The first aim of my work is to introduce a method to address the limitation encountered by depending on a single seed point for seed-based correlation analysis (SCA), and instead submit a probabilistic formulation of SCA which is robust to variations in the initial seed point. The method will firstly produce the strength of the correlations for voxels strongly correlated to the posterior cingulate cortex. Additionally, this approach provides a probabilistic interpretation of functional connectivity network behavior in the brain, and a maximum *a posteriori* (MAP) estimation of regions belonging to the DMN and DAN will be demonstrated as an example application of the procedure. In order to establish the ability of proposed method, seed cloud SCA (SC-SCA), to provide comparably reliable results to those of SCA, the group maps resulting from the method will be compared against those calculated from performing SCA and region-based SCA. The SC-SCA method will also be compared against its region-based counterpart, region-based SC-SCA in order to determine any extra advantages gained by performing a region-based correlation. The statistical significance of the difference between methods will be evaluated by comparing the difference in the median of the standard deviation of voxels across subjects. The components resulting from the decomposition of the FC network from SCA and SC-SCA via ICA will also be compared. This method represents a fully automated approach with probabilistic interpretations which confers augmented understanding of the DMN and its relation to underlying brain functionality.

3.1 Introduction

3.1.1 Correlation analyses

SCA [29] is useful for addressing the fundamental question of FC - which regions of the brain exhibit similar temporal patterns to those of a particular region of interest, regardless of spatial distribution or anatomical connectivity? The temporal behavior of a particular seed region of interest (ROI) is defined by its average BOLD signal time course, and is then correlated with the time course of every other ROI in the brain. ROIs with coefficients exceeding some threshold or level of significance using hypothesis testing are then considered to be functionally connected to the seed region.

There are several types of ROIs conventionally used, depending on the granularity of the desired FC network. Typical ROI selections include a single expertly-selected voxel, geometric sets of voxels (usually a sphere), or an entire brain region as defined by a brain atlas. Voxels or regions can also be identified based on statistical significance in task-based fMRI scans [65].

Although single-seed ROI SCA methods are popular due to their simplicity and the ease of interpreting the results, the FC networks generated depend notably on the initial selection of the seed location, particularly if the seed point is a single voxel. The resultant FC networks can vary widely if the seed location is varied even slightly [37, 66]. There is no general consensus on locations for ROI seed centers, even for investigating a single brain network or specific application. Thus, seed choice causes variety in the FC networks generated for identical topics, obfuscating the meaning of the individual networks generated. Although whole brain correlation can be performed (i.e., correlation between all voxels pairs) to overcome the limitation to a single seed, this clouds the interpretation of connectivity with a single region or seed.

Despite this drawback, SCA has been used extensively to study resting state functional connectivity and investigate the default mode network (DMN) and other resting state networks (RSNs) [67, 68, 69, 70]. Additionally, SCA has been used for many different applications including surgical planning [71], investigating FC of different brain pathologies including epilepsy [72, 73], Alzheimer's disease and aging [74, 75], autism spectrum disorders [76, 77], mood disorders and schizophrenia [78, 79], the effects of chronic back pain [80], and other clinical applications [52].

3.1.2 Independent components analysis (ICA)

Another popular method for investigating resting state connectivity and FC is through independent components analysis (ICA) [81, 82, 83, 34, 84], which has been used to reliably extract the resting state networks (RSNs) from several datasets [32, 85]. ICA is also known as an approach to solving the *cocktail party problem*, a problem which is described as trying to determine the underlying speech patterns from a room in which multiple speakers are speaking and recording devices are spatially distributed throughout the room of interest. In order to solve this question, the mixing problem can be formulated as:

$$\mathbf{X} = \mathbf{AS} \tag{3.1.1}$$

where $\mathbf{X}$ is the matrix of fMRI data, or the recordings of the speakers through the recording devices, $\mathbf{S}$ are the signals in a matrix optimized to have statistically independent spatial maps or speech signals in its rows, and $\mathbf{A}$ is the mixing matrix, where time courses are contained in the columns. For fMRI data, $\mathbf{X}$ has size p by n, where p is the number of time points, and n is the number of voxels in an fMRI image. $\mathbf{S}$ has size q by n, where q is the number of components, and thus $\mathbf{A}$ has size p by q. Analogous to the mixing formulation is

its inverse, the unmixing formulation, which is given as:

$$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} \tag{3.1.2}$$

In fMRI, ICA tends to be separated into two different flavors: spatial ICA (SICA), where the independent components are spatially independent, and temporal ICA (TICA), where the independent components are temporally independent. These are shown in Figure 3.1. Independence can be formulated in a number of ways, including through mutual information, infomax, and projection pursuit. An additional method of ICA frequently employed in fMRI is probabilistic ICA (PICA), which employs probabilistic methods in order to determine the best number of independent components to extract, and allows non-square mixing in the presence of Gaussian noise.

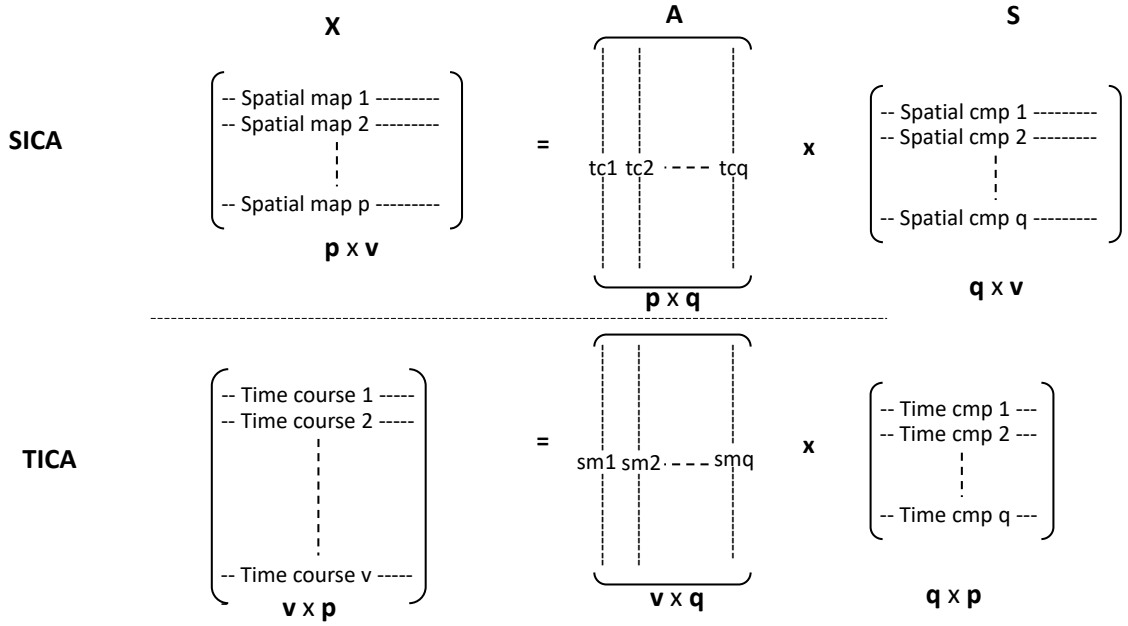


Figure 3.1: Spatial and Temporal formulations of ICA for fMRI

3.1.3 Probabilistic methods in FC

Probabilistic interpretations of FC via Bayesian methods have also been employed. One of the desirable features of probabilistic approaches for analysis of fMRI data is that they are flexible in both the spatial and temporal domains [86]. However, Bayesian methods have been predominantly limited to studying effective connectivity via data-driven methods without using a priori models based on expert knowledge of brain structure and function [86, 87, 88, 89]. Effective connectivity differs from FC in that it identifies causality (i.e., directed influence) between brain regions. Bayesian approaches have also been employed clinically in studying FC differences in Alzheimer's patients [90] and Parkinson's disease [91] for both task-based and resting state activity conditions [92].

3.1.4 Evaluating statistical significance between two populations

One possible method of evaluating whether a specific treatment (or method) is better than another treatment is through hypothesis testing. In this approach, a hypothesis is made about whether a parameter has a certain value or is in a particular range. Generally, the default hypothesis is the *null hypothesis* (H_0), or the hypothesis of no difference; that is, the parameter for both treatments has the same value and there is no statistically significant difference in applying the two treatments. The converse hypothesis, known usually as the *alternate hypothesis* (H_A) states that there is indeed a statistically significant difference between the two treatments, and the parameter has a substantially different value or lies outside the null range. The decision to accept or reject the null hypothesis and accept or reject the alternate hypothesis is then based on test statistics calculated from the data [93].

There are two major types of errors that can result from falsely rejecting or accepting the null hypothesis: type I and type II errors. A type I error is frequently referred to as a *false positive*, and is an error that results from falsely rejecting the null hypothesis. This

means that in the case that there is truly no difference between two treatments, a difference has instead been incorrectly identified. On the contrary, type II errors, *false negatives*, are errors that result from incorrectly accepting the null hypothesis. In this case, one has falsely identified that there is no difference between two treatments when there is indeed a difference. The acceptable rates of type I and type II errors is highly dependent on the application [94].

Hypothesis tests can be either one-sided or two-sided. The efforts of a one-sided test are usually focused towards determining whether there is a positive difference or a negative difference between two treatments. On the other hand, in a two-sided test, the thrust is determining whether there is any difference, regardless of whether the difference is positive or negative [93].

Bootstrapping

The other critical parameters in performing the hypothesis test are the null and sampling distributions. Although the sampling distribution is usually unknown for a particular sample, it can be estimated. One popular method for estimating the sampling distribution of a sample is through bootstrapping the distribution. The purpose of the bootstrap method is to determine a parameter of the data such as a mean, median, or standard deviation, and also the parameter's distribution. The bootstrap method utilizes random samples drawn from the data with replacement in order to generate a new distribution of a parameter of interest. In this way, the sampling distribution can be formed for a particular sample, and depending on the formulation, it can be used determine whether to accept or reject the null hypothesis [94].

3.2 Methodology

This section describes the steps necessary to execute the method and the procedures taken to evaluate the results. Processing was carried out using a combination of Matlab

R2016a (The Mathworks Company; Natick, MA, USA), Statistical Parameteric Mapping, Version 12 (SPM12) (The Wellcome Department of Neuroscience; Oxford, UK), and the Brain Extraction Tool (BET) [95, 96] and Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) packages of the FMRIB Software Library (FSL) (FMRIB Analysis Group; Oxford, UK) [97].

3.2.1 Data acquisition

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

The subjects were drawn from a population of 1200 healthy adult volunteers in the age range of 22-35 years. The dataset contains high-resolution T1-weighted sMRI and fMRI brain images. A total of 16 sets of fMRI data were randomly selected from the HCP database.

All HCP subjects were scanned using a customized Siemens Skyra 3T scanner with identical imaging parameters. The T1w image was acquired using the 3D MPRAGE sequence with 0.7mm isotropic resolution ((FOV = 224 mm, matrix = 320, 256 sagittal slices in a single slab), repetition time (TR) = 2400 ms, echo time (TE) = 2.14 ms, inversion time (TI) = 1000 ms, flip angle (FA) = 8°, bandwidth (BW) = 210 Hz per pixel, echo spacing (ES) = 7.6 ms). The fMRI scans were obtained using the following parameters: TR=720 ms, TE=33 ms, multiband factor=8, image matrix=104x90, 72 slices, 1200 volumes (time points), and isotropic slice size=2x2x2 mm^3. Full details regarding the acquisition of the data can be found in [98].

3.2.2 Data preprocessing

The HCP dataset minimizes the amount of preprocessing applied to the data; relevant preprocessing steps include removal of spatial artifacts and distortions and registration of surfaces and volumes to standard volume and surface spaces. Further preprocessing steps were applied subsequent to retrieving the data. First, each voxel time series was temporally filtered using a finite impulse response (FIR) bandpass filter (0.01Hz-0.1Hz) and linearly detrended. Since convolution of an N^{th}-order FIR filter with the data produces a finite impulse response of length (N+1) at the beginning of the data and due to other computational limitations, a total of 200 time points were removed from the beginning and end of the data. The data were a resultant 800 volumes in length.

The temporally filtered and trimmed data was then spatially smoothed in FSL using a 6mm full width half maximum (FWHM) Gaussian filter. The data was then normalized by de-meaning and dividing by the standard deviation of each voxel time series.

3.2.3 Probabilistic fMRI interpretations

Probabilistic network generation

In order to generate a network map demonstrating the probability of being highly connected to the PCC, an initializing fixed seed point was first selected from the DMN. For this work, we identified the initializing seed point at the MNI coordinates [0, -53, 26] [99, 100] in the posterior cingulate cortex (PCC). SCA was performed using the selected PCC seed, and the voxels with the top 20% of the calculated Pearson's correlation coefficients ($r > r_t$) were selected to be included in a *PCC seed cloud*. A typical threshold value for r_t is approximately 0.4. This seed cloud represents a set of voxels that are highly correlated with the initializing PCC seed. Two thousand voxels were then randomly selected from the PCC seed cloud as new initial seed points. SCA was performed for each of these seed points. Two types of

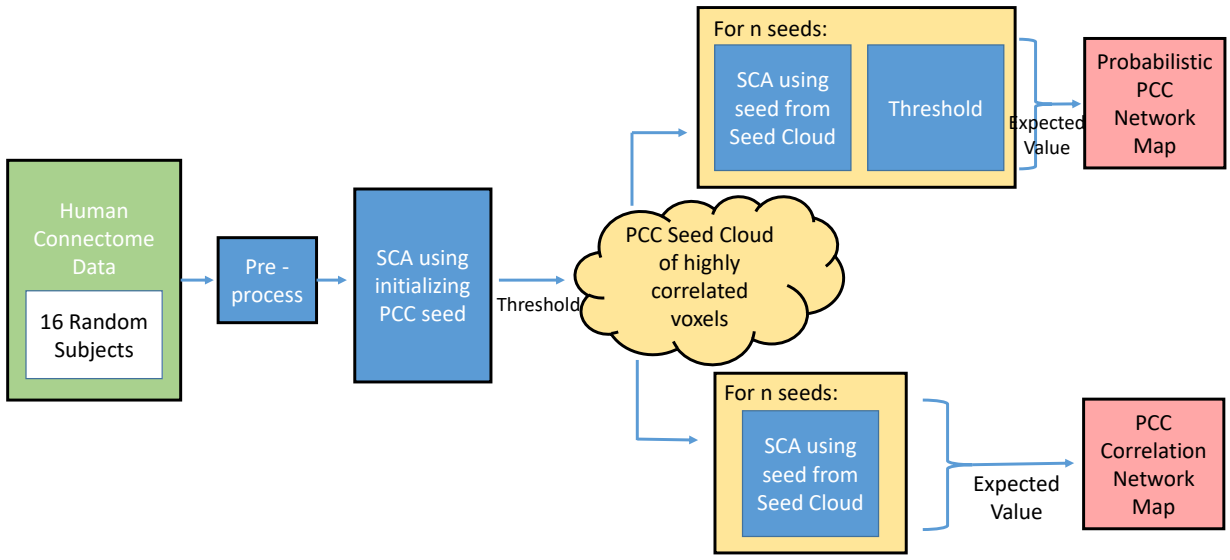


Figure 3.2: Processing pipeline for generation of PCC network maps

maps were generated to understand the relations with the PCC as shown in Figure 3.2: a probabilistic map and an average correlation coefficient map. To generate the probabilistic maps, the random variable X was defined to be the correlation value at a particular voxel, and takes on values in the range of r (i.e., [-1,1]). The event A_r is defined as an event on X which occurs when $X > r_t$, where r_t is a threshold value for Pearson's r. The indicator function is then defined as:

$$\mathbb{1}(X) = \begin{cases} 1 & X \in A_r \\ 0 & X \notin A_r \end{cases}$$

This allows the probability of event A_r to be defined as:

$$P(A_r) = \mathbb{E}[\mathbb{1}(X)]$$

This was implemented by thresholding each of the 2,000 SCA maps at $r_t + 0.1$. This slight increase was included to reduce artifacts around the initializing seed. These 2,000 maps were averaged to generate a probabilistic PCC network map for each subject.

The average correlation coefficient map was generated in a similar way, except after the 2,000 SCA maps were generated, they were simply averaged to generate the PCC network map reflecting the expected value of the correlation with the highly connected PCC seeds.

To perform a group analysis of all 16 subjects, the PCC network maps for each subject were averaged to generate a group map. To determine binary inclusion of a voxel in the PCC network for the average correlation map, the group PCC network correlation map was thresholded at r=0.4, and voxels with r>0.4 were considered to be a part of the PCC network.

Maximum *a posteriori* network estimation: an example application

In order to demonstrate the efficacy of this method for a Bayesian interpretation, an example application of maximum *a posteriori* (MAP) estimation for the DMN and DAN networks is presented. To use this method, a binary *a priori* estimate of the DMN and DAN networks was obtained from the liberal 7 network Yeo cortical parcellation map [5]. This was used as an initializing seed cloud in order to generate the expected value of the correlation coefficient map ($\mathbb{E}[correlation|seed \in Network]$), denoted θ for each of the 16 subjects. For each voxel v_i in the map, a probability density function (PDF) for the correlation was generated using the 16 subjects, and was modeled as a Gaussian PDF by computing the group mean μ and standard deviation σ as parameters. This computation is the *likelihood function* for a voxel being in the network, and denoted as $f(\theta|v_i \in Network)$.

The binary estimate of the network of choice was then used to generate a smoothed estimate of the network by convolving each of the voxels with a Gaussian, producing an

a priori estimate denoted $P(v_i \in Network)$ which reduces strict cutoffs in the probability estimation.

According to Bayes' Rule, using these elements, a maximum *a posteriori* estimation of the probability of $v_i \in Network$ given the average correlation estimate θ can be expressed according to Equation 3.2.1.

$$P(v_i \in Network|\theta) = \frac{f(\theta|v_i \in Network)P(v_i \in Network)}{f(\theta)} \tag{3.2.1}$$

The term $f(\theta)$ can be computed using the smoothed *a priori* network estimates and the likelihood function as shown in Equation 3.2.2.

$$\begin{aligned} f(\theta) = f(\theta|v_i \in Network)P(v_i \in Network)+ \\ f(\theta|v_i \notin Network)(1 - P(v_i \in Network)) \end{aligned} \tag{3.2.2}$$

3.2.4 Sensitivity analyses

Sensitivity analyses are essential in evaluating the dependence of the proposed algorithm on the parameters used to generate the results. Two sensitivity analyses were performed in order to determine the robustness of the PCC network map algorithm: (1) the number of seeds used to produce the cloud of highly correlated points, and (2) the location of the initial PCC seed point.

Number of cloud seeds

In order to assess the effects of varying the number of seeds to generate the point clouds on the generated PCC network map, a subject whose PCC map was representative of the PCC network was selected. An identical procedure to the one described in the above section was performed; however, the number of seeds was varied from 1,000 to 10,000 in increments

of 1,000, generating 10 maps. In order to assess the similarity between the generated networks, each map was thresholded at an average correlation value of r = 0.4. Then, Dice's coefficient [101] was calculated between each of these binary maps and the base map (i.e., the map generated by choosing a 2,000 seed point cloud). Dice's coefficient for two sets (e.g., maps) X and Y are given by the following equation, and is a measure of the similarity between the two maps.

$$D = \frac{2(|X \cap Y|)}{|X| + |Y|}$$

The coefficient varies between 0 and 1 with 1 indicating that the maps are identical, and the $|*|$ notation indicates the cardinality of the set.

Initial PCC seed location

A similar technique was employed to assess the robustness of the PCC network map against variations in the location of the initial seed point. In order to assess this, the initial seed point was varied from the original PCC seed point ($[0, -53, 26]_{MNI}$) by $\pm$10mm in 2mm increments along the vertical, transverse, and sagittal axes. This generated a total of 31 maps, including the map corresponding to the original seed point. The maps were compared in the same way as in Section 3.2.4 by thresholding the maps at r = 0.4, and calculating Dice's coefficient between the base map (i.e., the map generated with the PCC seed point of $[0, -53, 26]_{MNI}$) and each of the variational maps.

3.2.5 Comparison with SCA and region-based correlation

In order to compare the efficacy of this method against existing methods, the seed cloud-based SCA (SC-SCA) was compared against the commonly used method, SCA. Additionally, to further assess the significance of the initial seed point and associated time series, a region-based method was also investigated.

Generation of region time course

The region identified for use in the region-based SCA and SC-SCA was the PCC region. In order to generate a binary image of the region, Yeo's parcellation [5] of the DMN was first attained and then transformed into MNI space using SPM12. Following this computation, Multi-image Analysis Graphical User Interface (Mango) [102] was then used in order to calculate the ROI corresponding to the PCC. This was done by drawing a 22mm sphere around the estimated center of the PCC region in the DMN, and then performing a *Shrink Wrap* on the enclosed region. This was then saved and exported as the new region.

Following the binary identification of this region into a mask, the mask was then applied to each subject's time course, allowing all the time courses for all the voxels in the included region to be determined. These were then averaged in order to form the average time course for the region. Thus, instead of a single point being used for the seed point, the whole averaged time course for the region was utilized.

SCA

In order to assess the performance of SC-SCA, it was compared against the traditional method of SCA. In fact, SCA simply requires a subset of the steps required for SC-SCA. In order to perform SCA, first, a time course was identified based on either the PCC seed described in Section 3.2.3 or the region-based time course described in Section 3.2.5. After this time course was identified, it was then correlated with each other voxel in the brain using the following equation:

$$r_j = \frac{1}{n} \frac{(\mathbf{x} - \bar{x}) \cdot (\mathbf{y_j} - \bar{y_j})}{\sigma_x \sigma_{y,j}} \tag{3.2.3}$$

where r_j is the correlation value for the j^{th} voxel in the brain, n is the number of volumes in the time course, $\mathbf{x}$ is the seed time course, $\mathbf{y_j}$ is the time course for the j^{th} voxel, and σ_x and $\sigma_{y,j}$ are the standard deviations of each of the time courses.

The final generated product of SCA is a correlation map describing the correlation of the time course with each other voxel of the brain.

Modality of method comparison

To determine if there is a statistically significant difference in the different methods, the standard deviation of the correlation value in each voxel for each of the different methods was determined. This was done by calculating the standard deviation per voxel across the 16 subjects for each method of generating correlation maps (SC-SCA vs SCA and seed-based vs region-based). This generated single volume of standard deviations was considered as a distribution of the standard deviations of the correlation values over all the voxels.

The following methods were then compared on the basis of the KS statistic:

- Seed-based SC-SCA vs seed-based SCA
- Region-based SC-SCA vs region-based SCA
- Seed-based SC-SCA vs region-based SC-SCA
- Seed-based SCA vs region-based SCA
- Seed-based SC-SCA vs region-based SCA
- Region-based SC-SCA vs seed-based SCA

In this way, the statistical significance between all of the different methods can be adequately evaluated.

Bootstrap of sampling distributions

In order to visually inspect the differences between the standard deviation in voxels of all four methods, first, bootstrap sampling was performed on each of the four methods. To perform this for a single method, a sample was generated by randomly choosing 16 subjects (with replacement), and the standard deviation for each of the voxels was calculated. Then, the median of the standard deviation of all the voxels was calculated for the sample. This was repeated n times (1,000 in this work) in order to form a distribution of the median values of the standard deviation of the voxels across the subjects. The 95% confidence interval was then calculated for each of the distributions. This was repeated for each of the four methods.

Following this qualitative assessment of the difference between methods, a calculation of the statistical significance between the difference between the medians of the standard deviations across voxels was calculated. For a comparison between two methods, method A and method B, a single sample was generated of 16 randomly chosen subjects with replacement. Using this sample, the standard deviation was calculated for each voxel, and the medians, μ_A and μ_B, were determined for each method. Following this calculation, the difference between the medians was calculated as $\mu_{diff} = \mu_A$ - μ_B. This allows for a signed difference to be estimated between the two methods. This was repeated 1,000 times in order to form a distribution of μ_{diff}. In the case that there is no difference between the two methods, the data should nearly be centered around a standard deviation difference of 0. A statistically significant difference between the two methods, on the other hand, should be indicated by where the y-axis (y_0) lies with respect to the distribution. If $P(y_0 \leq \alpha/2)$ or $P(y_0 \geq (1-\alpha/2))$, then the null hypothesis is rejected and there is a statistically significant difference between the two methods. In this work, the level of significance was set to α=0.05. This was repeated for each of the of the combinations of the different methods.

3.2.6 Decomposition of networks via ICA

In order to determine the constituent components of the networks produced by SCA and SC-SCA, the networks were decomposed into independent components (ICs) using ICA. This was performed by first thresholding the seed-based correlation maps produced by SC-SCA and SCA at 0.5 to produce an estimate of the voxels which are most highly correlated to the seed of interest. These networks were then decomposed using MELODIC in the FSL [97] software package. The number of constituent ICs for the software to generate was set to 10.

Following ICA on the SC-SCA and SCA networks, the ICs were visually inspected for similarities between the two groups. Similar ICs were grouped together and dissimilar ICs were noted.

3.3 Results

A probability value or a correlation value at a particular voxel for a single subject in a generated PCC network map can confer several interpretations and expand current understanding of the DMN and connections to the PCC. Robust definition of the DMN is critical to understanding structure-function relations within this network, and how it remodels during the disease process or evolves with development or aging [103, 104, 105, 106]. These interpretations are extended for the group of 16 subjects and their dependence on particular parameters is reported in the following sections.

3.3.1 Probabilistic network map

Figure 3.3 shows the expected value of the probability of being highly correlated to the PCC. The average probabilities range from 0 to 0.6. As shown, the regions with the highest probabilities (in white, >0.55) are the precuneus (pC) and the PCC, and neither of these two regions contains the location of the initializing PCC seed. Other regions with high

probabilities of connection are closely related to the commonly accepted regions of the DMN, which include the medial prefrontal cortex (mPFC), the left and right inferior parietal lobule (R/L IPL) with probabilities >0.45, and also the parahippocampus (PH) and hippocampal formation (HF) regions with probabilities >0.35.

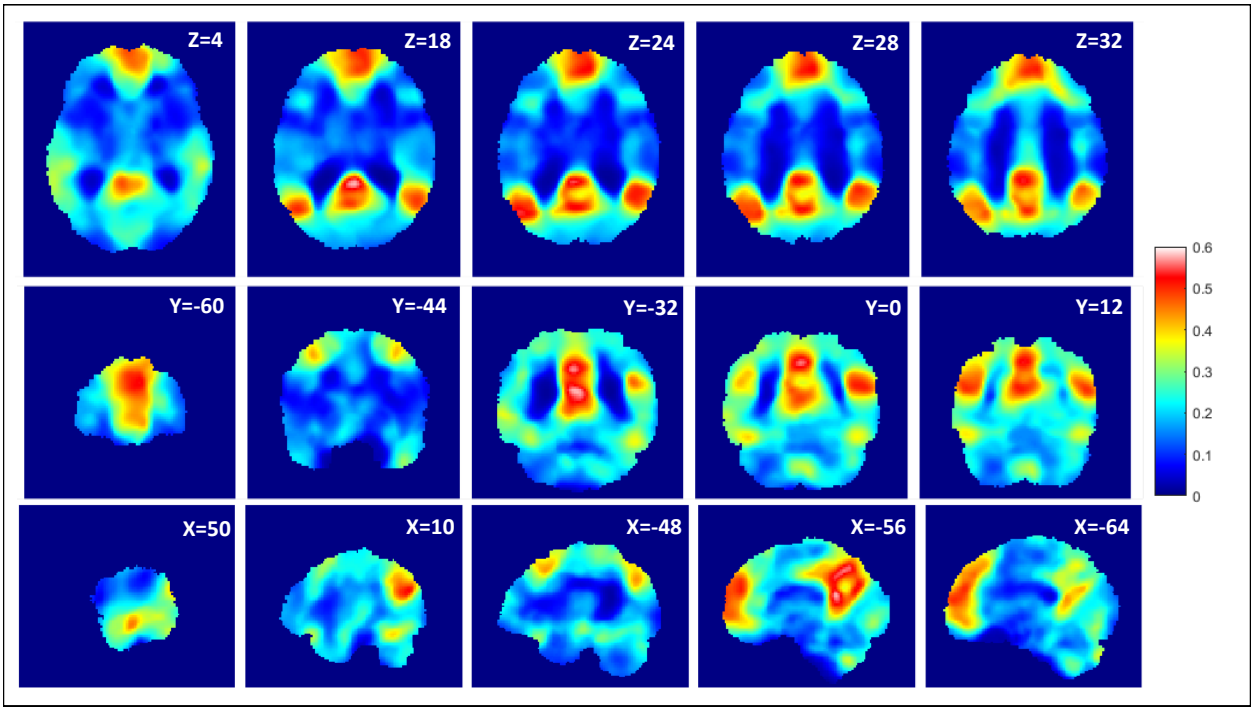


Figure 3.3: Probabilistic PCC network for set of 16 subjects

The most straightforward interpretation of these probabilities is a probability of connection; that is, if a random seed was identified in the brain, what is the probability that it would be highly connected to the PCC? An alternate interpretation also allows the relation of the percentage of the PCC network with high correlation to the voxel of interest; that is, what percentage of the PCC network has a high correlation to a particular voxel? This method allows a deeper probabilistic understanding of connections in the DMN, while also being in accordance with the current understanding of DMN FC.

3.3.2 Correlation network map

Figure 3.4 displays the mean PCC correlation network map for the group of 16 subjects. As shown, there is high correlation (r>0.65) between the main parts of the PCC network, mPFC, the PCC, pC, and the right and left IPL. The expected values of the correlations range from 0 to 0.7; although correlation values with the DMN can often dip into negative values, these values are likely averaged out due to the computational averaging over an entire time series and different seeds for a voxel to compute the correlation coefficient.

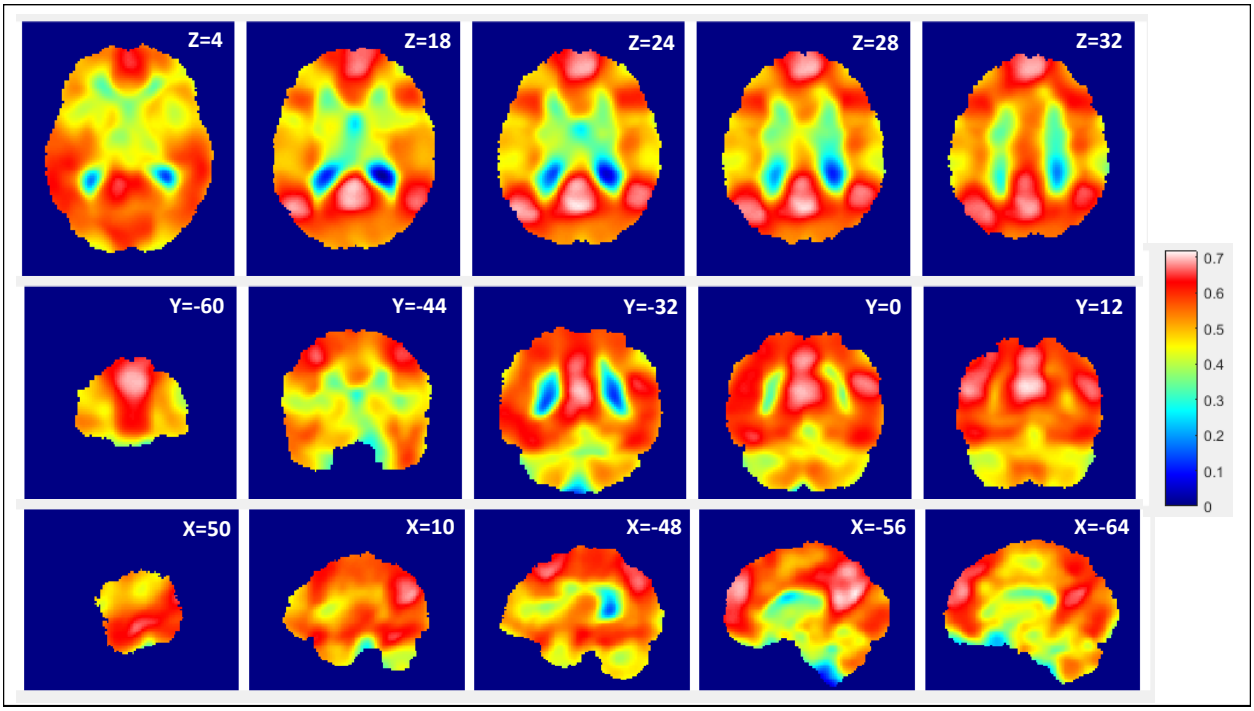


Figure 3.4: Mean Pearson's correlation coefficient for set of 16 subjects

This method also reveals other sets of sub-networks associated with the DMN, while also showing the *extent* of the connectivity between the sub-networks and the PCC network. This is clearly demonstrated when evaluating the correlation values (i.e., r) on intervals, e.g., $(-\infty,0.3)$, [0.3, 0.4), [0.4,0.5), [0.5,0.6), [0.6,0.7), and $[0.7,\infty)$. Although these ranges

were defined arbitrarily, the delineations via correlation coefficient tend to segment the PCC network into regions referred to as the DMN and others sometimes associated with the DMN.

As Figure 3.4 shows, the r≥0.7 (in white) region identifies the most highly correlated regions in the PCC network; as shown, this includes the pC and PCC, neither of which are within a 4mm sphere of the seed region. However, these are essential components of the DMN.

The second most correlated segmentation (interval [0.6,0.7)) shown in red includes a larger region surrounding the highly correlated regions in white, but additionally include all of the regions which are known to be in the DMN (mPFC, right and left IPL, PH). However, there are an additional set of regions including the right and left mid frontal gyrus, the right and left fusiform gyrus, and the middle and superior right and left temporal gyrus, the left superior parietal lobule, the left cingulate gyrus, and the left culmen. An unexpected region of correlation of between 0.6 and 0.7 with the PCC network is the lingual gyrus, which is not usually associated with the DMN.

The third most correlated segmentation are the areas with r between 0.5 and 0.6 (orange), and correspond to the largest volume of the brain. Although this is the third level of connection identified by the arbitrary thresholds, the correlation values are still relatively high, meaning that the PCC network spans and touches many regions and networks of the brain with high levels of connectivity. These regions include the right and left mid frontal gyrus, which are areas speculated to be associated with the DMN, as well as the insula, cuneus, and several other regions.

Correlation values of r<0.5 (yellow, green, and blue), appear to correspond very strongly to white matter pathways along which information can be passed along the transverse, vertical, and sagittal axes. Partial volume effects and the effects of smoothing are likely responsible for the clear interface values (e.g. outlining of ventricles) between white matter

and gray matter values. The ventricles correspond to the lowest areas of correlation to the PCC network.

It is additionally interesting to note the interpretation of these results based on the hypothesis considered about brain functionality. One hypothesis about the functionality of the brain is that there is a certain time-multiplexing of different subsystems of the brain to the PCC; in that case, the PCC network generated may reflect the amount of time that the subnetworks or regions are connected to the PCC network. On the other hand as represented here, if the assumed hypothesis is that the brain is organized into constant networks of varying connectivity, then the PCC network generated represents the average strength of the connectivity.

3.3.3 MAP DMN and DAN estimation

Figures 3.5, 3.6, and 3.7 show a set of MAP estimates for the DMN based on the *a priori* probability maps used for their calculation. Figure 3.5 shows the probability estimates for voxel-wise inclusion in the DMN given a 4 mm Gaussian smoothing of the binary *a priori* DMN map. The probabilities range from 0 to 0.99, and as shown, the mPFC, the right and left IPL, the PCC, and PC all have probabilities >0.9 of being in the DMN. The amount of smoothing used can be interpreted as reflecting the amount of confidence in the *a priori* estimate; more smoothing corresponds to less confidence and vice versa.

In contrast, Figure 3.6 shows the DMN for a 10 mm Gaussian smoothing of the binary *a priori* DMN map. Due to the larger area of smoothing, the DMN inclusion estimates have a slightly lower range, from 0 to 0.87, but all of the same regions with the highest probabilities are included in the DMN MAP estimate. As shown, there is a significant difference in the detail obtained from the 4 mm vs the 10 mm smoothing; for example, fine "u" and "n" shapes can be seen in the axial and coronal slices at Z=-4 and Y=60, respectively. However, at the 10 mm smoothing, these levels of fine detail are lost. This can also be seen in the

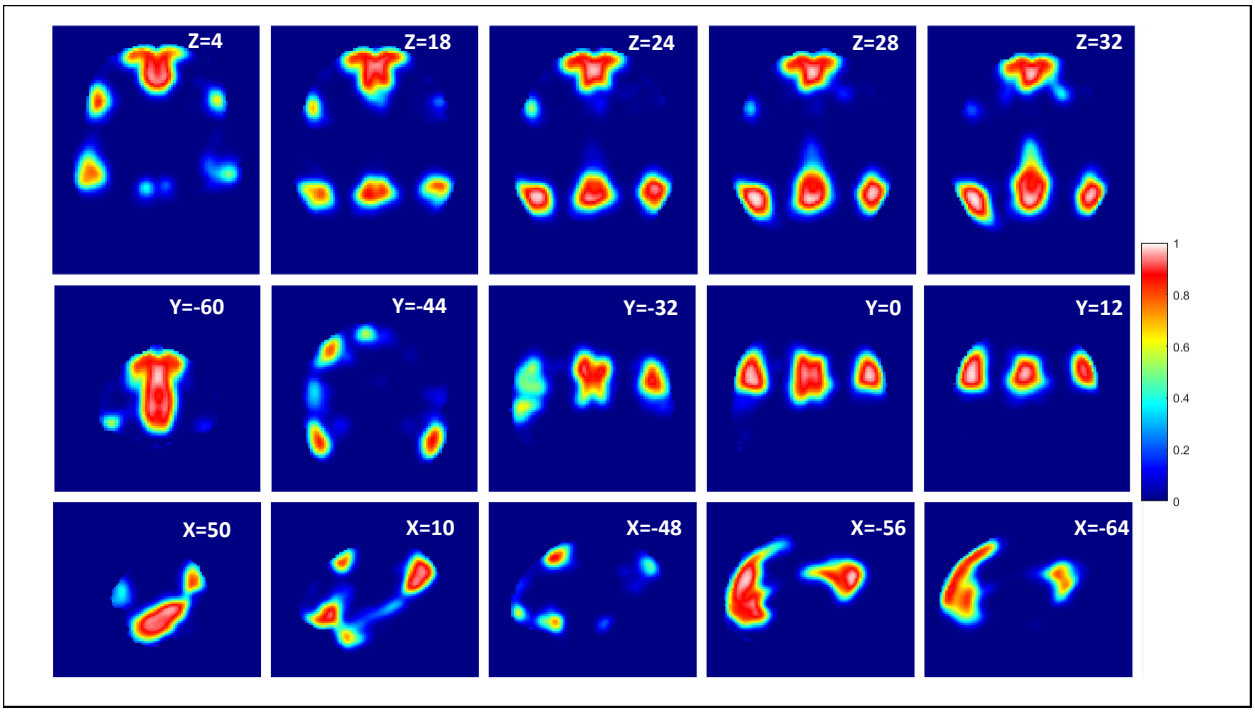


Figure 3.5: 4mm Gaussian kernel *a priori* smoothing for DMN

X=-56 sagittal slices of the brain, where two regions of very high probability can be discerned in the mPFC; however, this is again lost in the 10 mm MAP estimate.

Figure 3.7 shows the effect of complete ignorance (i.e. uniform *a priori* probability of 0.5 for all brain voxels) of the *a priori* probabilities on the MAP estimate. This is equivalent to a maximum likelihood estimate. As shown, most regions have a probability of 0.5; however, the regions that are known to be in the DMN still have higher probabilities (e.g., >0.7) of inclusion in the DMN. This map has the smallest maximum MAP probability value of 0.77.

Figures 3.8, 3.9, and 3.10 show a similar set of MAP estimates for the DAN network. Figure 3.8 shows the MAP estimate using a relatively "certain" *a priori* DAN network with a 4mm smoothing applied. As shown, the voxels with high *a posteriori* estimates are those belonging to the DAN network. This can be seen in the component brain regions, including the right and left intraparietal sulcus (IPS) and the right and left frontal eye fields (FEF).

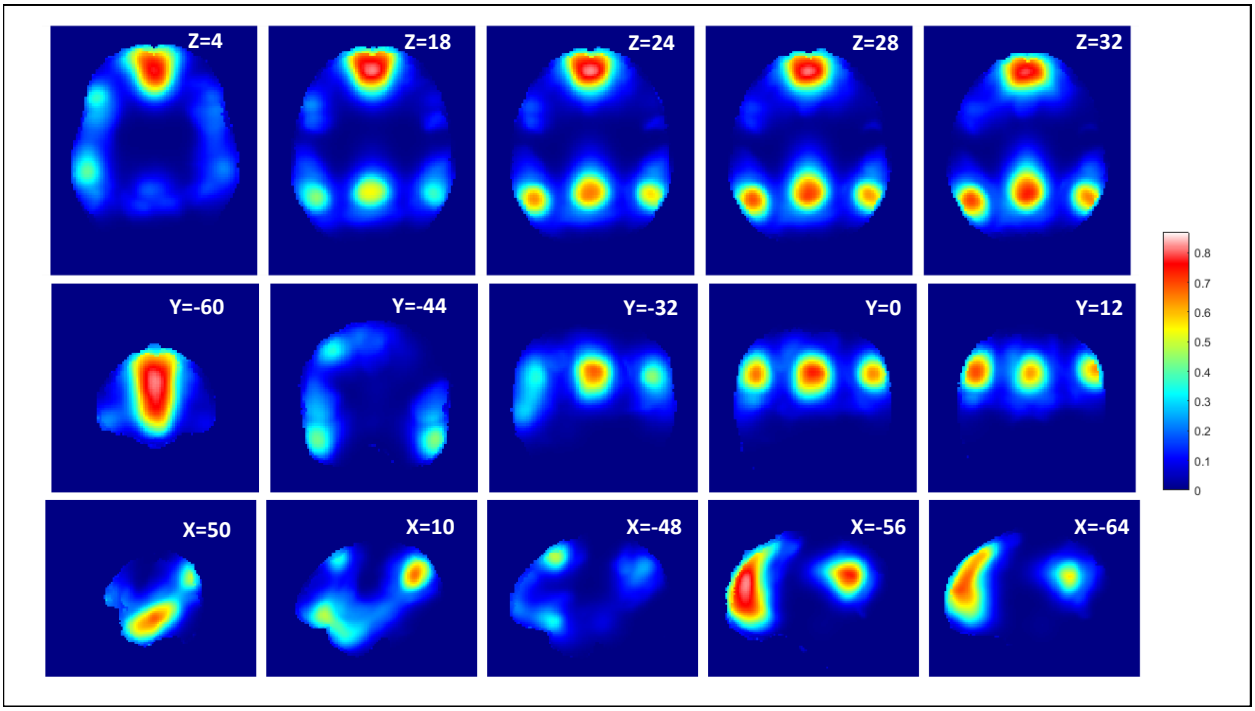


Figure 3.6: 10mm Gaussian kernel *a priori* smoothing for DMN

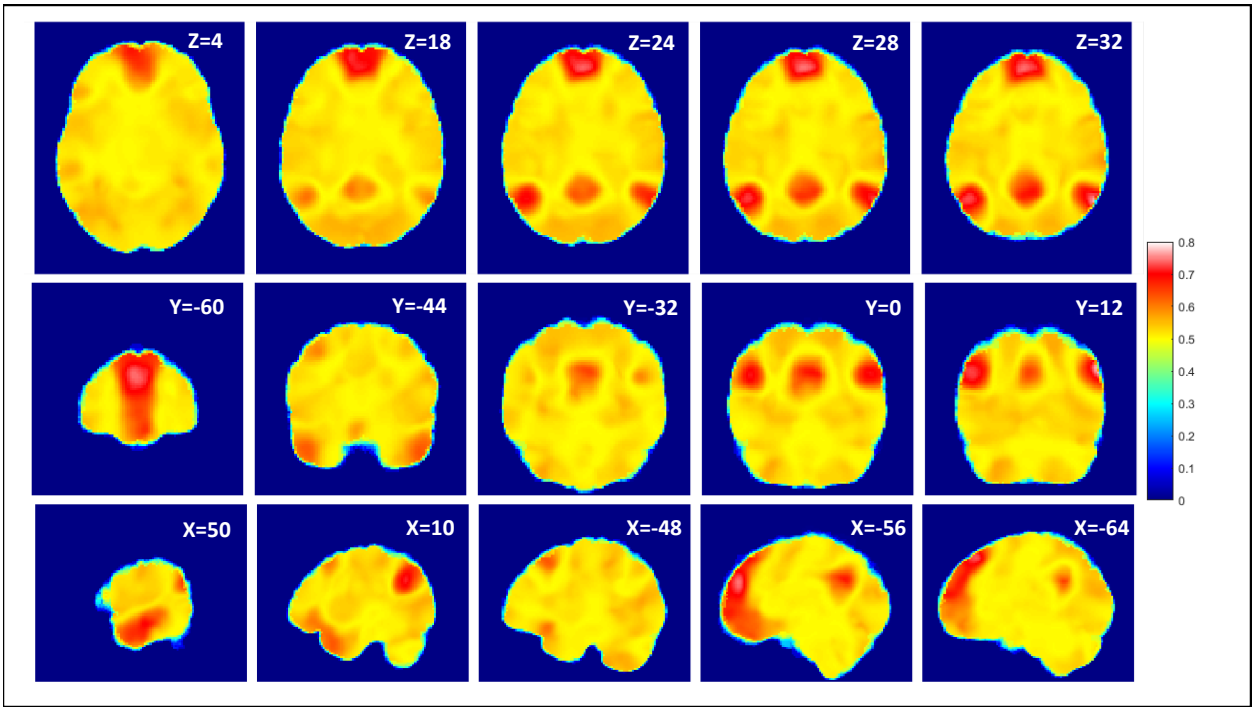


Figure 3.7: Maximally ignorant (assuming a constant) *a priori* distribution (p=0.5), equivalent to maximum likelihood estimation for DMN

There are also regions of high MAP value in the right and left fusiform gyrus. The MAP estimates range from 0 to 0.96, which the voxels with the highest MAP estimates more likely to be in the DAN.

Figure 3.9 shows a MAP estimate for a less certain *a priori* distribution for the DAN network. Similarly to Figure 3.6, the produced estimates appear a little more diffuse than the 4mm networks, with a lower range of values for the MAP estimates. The range for the MAP estimates for this level of smoothing is from 0 to 0.78 and the regions with the highest certainty around 0.7 are the IPS, FEF, and fusiform gyrus regions.

The maximum likelihood estimation for the DAN network is shown in Figure 3.10. As shown, most areas have an value close to the *a priori* estimate of p=0.5, but regions associated with the DAN have slightly elevated values. These values range from about 0.6 to the maximum value of 0.77. In this case, these regions for the IPS and the FEF are in the most superior regions of the brain; in contrast to Figure 3.8, the FEF and IPS regions in the axial slices are not nearly as pronounced.

3.3.4 Sensitivity analyses

The effect of the number of seeds parameter and the location of the initializing PCC seed were specifically assessed in this study. There are additionally other parameters which have an impact on the generated PCC network maps; however, their contributions have predictable results. For example, a correlation threshold corresponding to the 80th percentile and higher (i.e., top 20% of the correlation values) was used to identify the seed cloud of voxels highly correlated with the PCC. This threshold can be raised or lowered; raising the threshold results in PCC network maps that have higher overall correlation values, whereas lowering the threshold results in lower overall correlation values. Similar behavior exists with the r = 0.4 threshold used to generate the sensitivity maps; lowering the threshold simply creates larger binary maps and raising the threshold creates binary smaller maps.

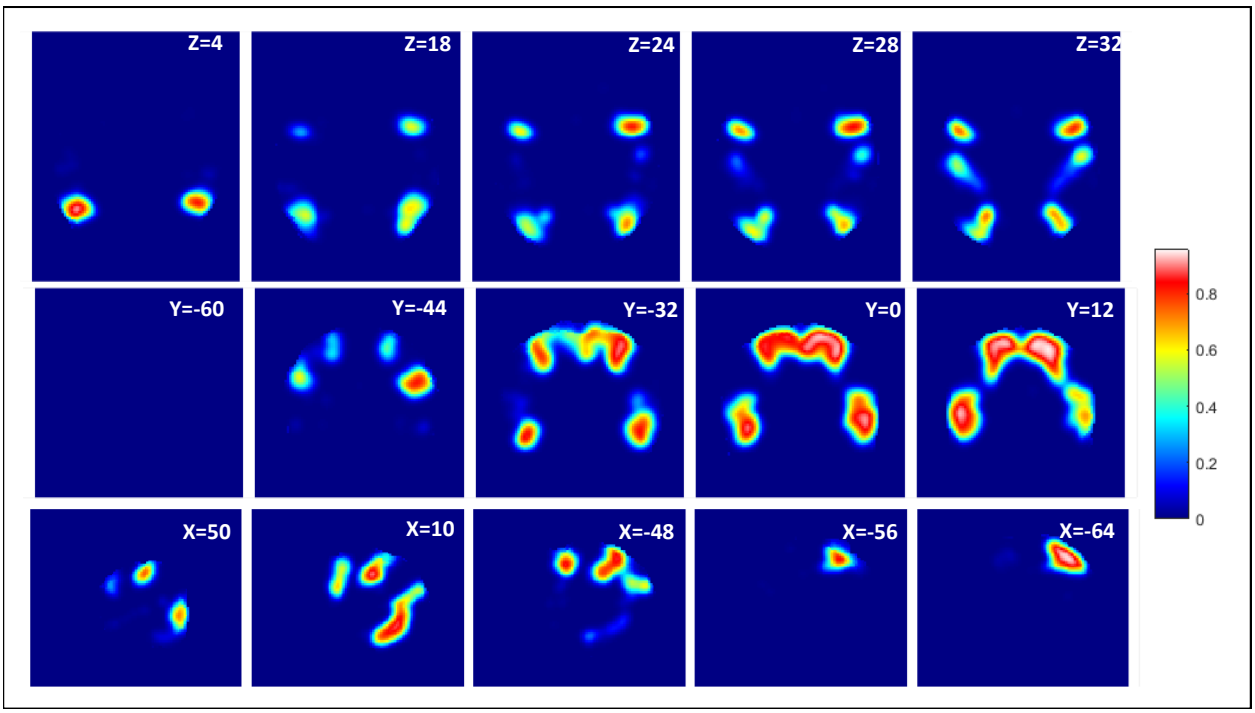


Figure 3.8: 4mm Gaussian kernel *a priori* smoothing for DAN

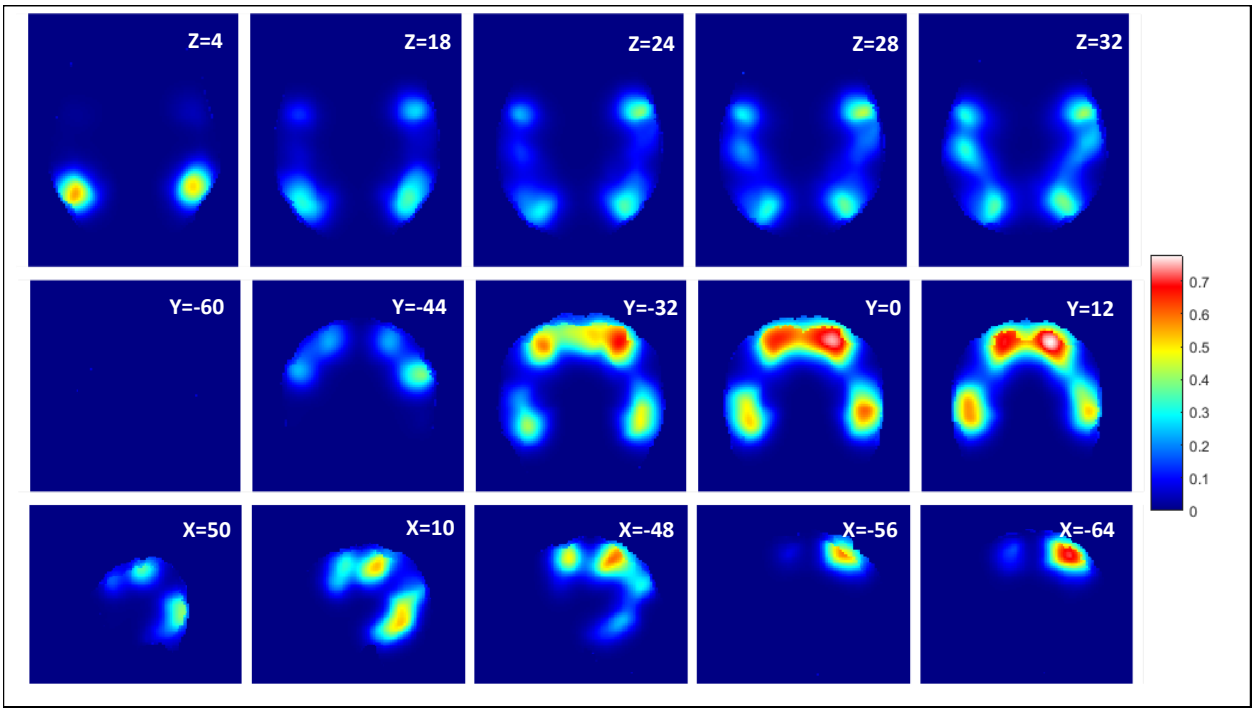


Figure 3.9: 10mm Gaussian kernel *a priori* smoothing for DAN

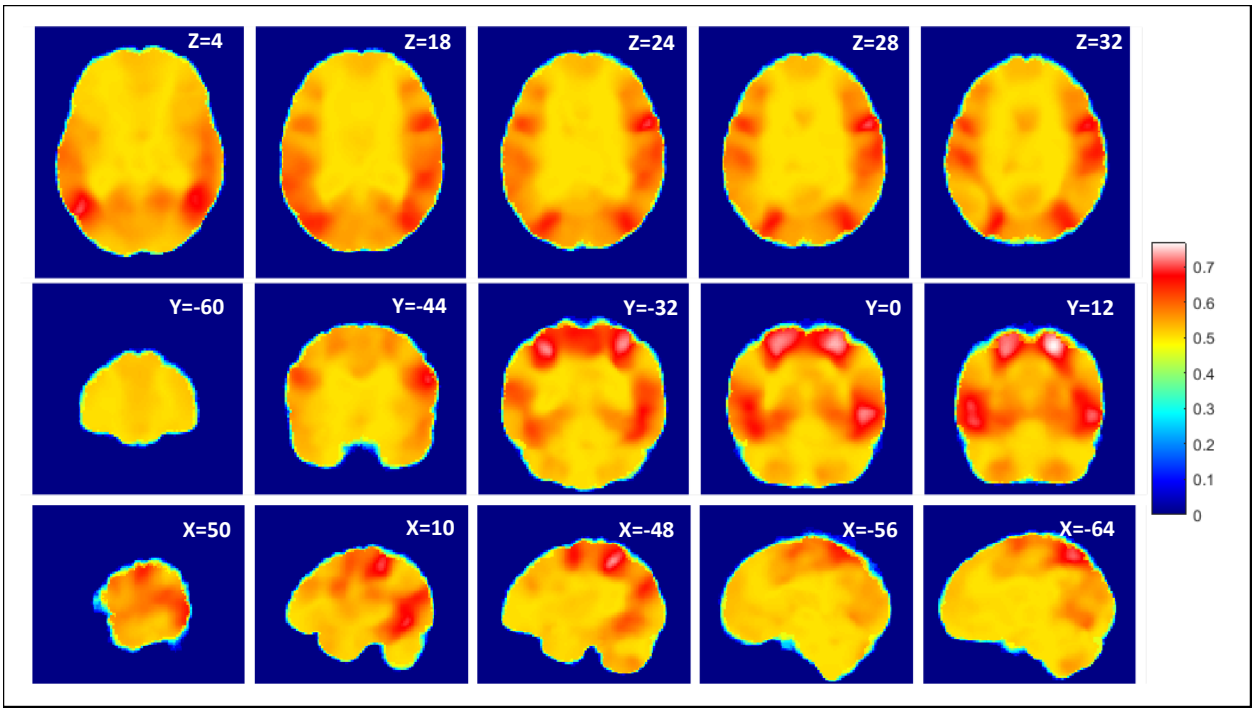


Figure 3.10: Maximally ignorant (assuming a constant) *a priori* distribution (p=0.5), equivalent to maximum likelihood estimation for DAN

Number of cloud seeds

Figure 3.11 displays the effect of varying the number of seeds selected from the PCC seed cloud on the PCC network generated. As shown, there is a similarity of at least 99% for the maps generated via any of the numbers of seeds that were investigated. Utilizing 1,000 seeds results in a slightly smaller similarity than the larger number of seeds (2,000 seeds has an overlap of 100% since it is compared against itself), and the most effective number of seeds as compared to 2,000 seeds are 3,000 and 4,000.

As the number of seeds increases, the similarity between the networks generated and the 2,000 seed map decreases slightly. This >99% similarity thus validates the usage of 2,000 seeds to generate the networks; the usage of this number of seeds thus optimizes the speed of the algorithm (fewer seeds results in faster processing time) with the fidelity of the results.

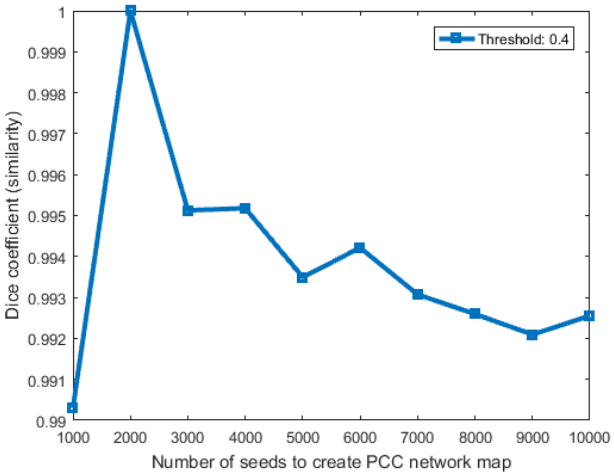


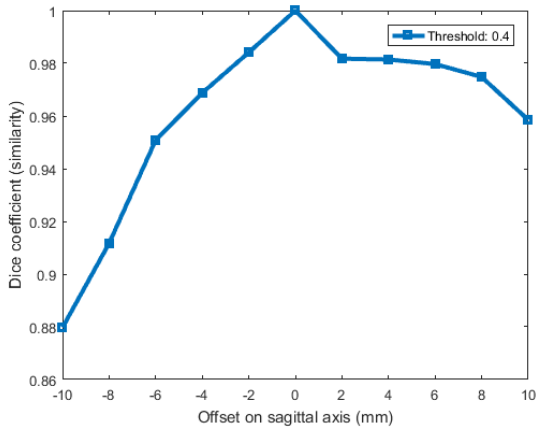
Figure 3.11: Dice's coefficient for sensitivity analysis for effect of number of seeds utilized on the PCC network generated

Initializing PCC seed location

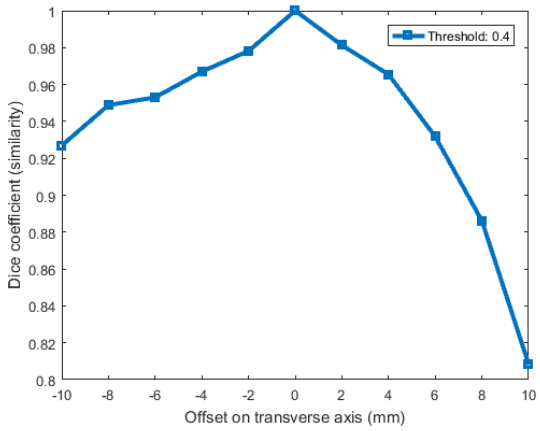
The effect of varying the initializing seed of the PCC network along the sagittal, transverse, and axial axes is shown in Figure 3.12. All of the results are compared initializing the algorithm with the $[0,-53,26]_{MNI}$ seed, which corresponds to the 0 mm offsets in Figure 3.12. Since this is the location which is being compared against, Dice's coefficient of similarity with this seed location is 100%.

As shown in Figure 3.12(a), varying the initializing seed location along the sagittal axis does towards the right represents only a minimal 4% change in the similarity of the networks generated. Varying the seed location along the sagittal axis towards the left however, has a slightly increased change of $\sim$12%, and thus, varying along one direction rather than another has a small, asymmetric impact on the network generated.

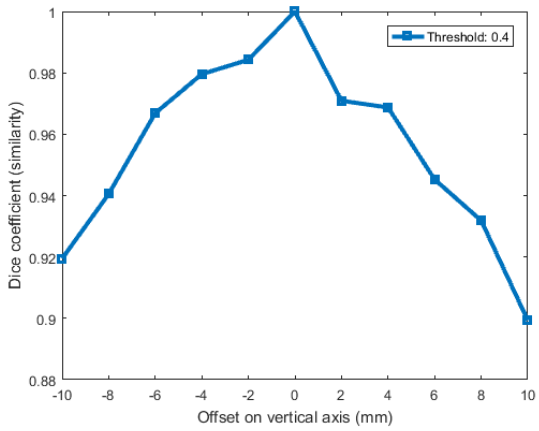
Varying along the transverse axis in an anterior direction represents the largest deviation of all from the network generated at a difference of 20% in the network, as shown in Figure 3.12(b). Due to the importance of sub-organs represented in the vicinity of the PCC (e.g., pC) along the transverse axis, varying in the anterior direction can correspond to



(a) Sensitivity of varying ±10mm on sagittal axis



(b) Sensitivity of varying ±10mm on transverse axis



(c) Sensitivity of varying ±10mm on vertical axis

Figure 3.12: Sensitivity analysis varying the location of the initializing PCC seed point

becoming increasingly closer to an organ that has strong connections to other subnetworks, thus leading to a slightly different network produced. This 10 mm anterior venture results in a network with slightly stronger correlation values to the pC, PCC, mPFC, and other areas associated with the DMN, but also has strong correlation (>0.61) to regions such as the cerebellar tonsil, declive, and culmen. Varying the location of the initializing PCC seed towards the back of the head only results in a small difference (∼8%) in the compared networks.

Figure 3.12(c) shows that varying the location of the initializing PCC seed along the axial direction in either the superior or inferior directions produces a rather symmetric effect on the generated network, and does not particularly affect the network generated. All perturbations of the location in the axial direction resulted in Dice similarities exceeding ∼90%.

3.3.5 Comparison of region-based and seed-based SCA and SC-SCA

The following results outline the generated correlation maps generated from SCA and SC-SCA. The statistical differences between the methods will additionally be described.

Region identification

Figure 3.13 shows selected slices of the the Yeo, 2011 [5] parcellation of the DMN (in white) in the vertical, transverse, and sagittal directions. The area identified to be the PCC is shown in red. The other brain region near the PCC and in some slices appear to be adjoint to the PCC is the pC.

Correlation maps of different methods

Figure 3.4 shows the generated correlation map for the seed-based correlation analysis for SC-SCA, and discussion of this map can be found in the preceding sections. Figures 3.14, 3.15, and 3.16 show the remaining 3 correlation maps for the combined treatments of SCA vs SC-SCA and seed-based vs region-based correlation analysis.

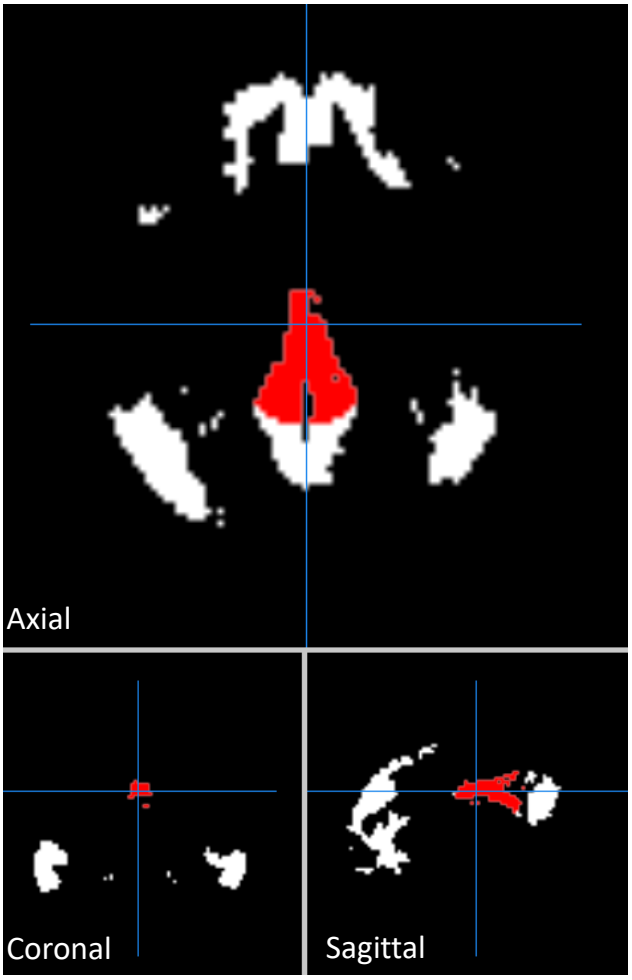


Figure 3.13: PCC region (in red) identified using Yeo, 2011 parcellation of the DMN (in white)

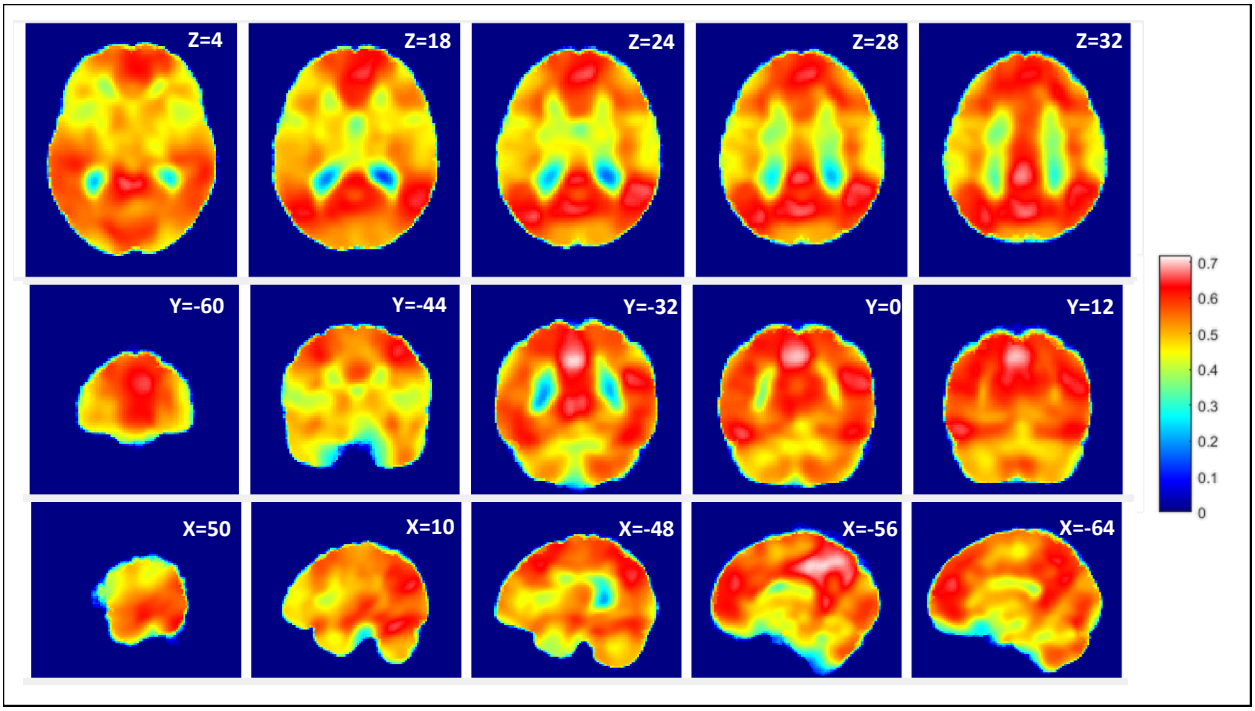


Figure 3.14: Group correlation map for region-based SC-SCA

Figure 3.14 shows the correlation map for the region-based SC-SCA method. The range of this correlation map is a minimum of -0.03 to 0.71. In this case, the correlation map does dip into negative values, although it is only very slightly less than zero. As shown, in this case, compared to the seed-based method, the strength of the correlation values in the classical DMN regions including the mPFC and the left and right IPL are all much weaker and less pronounced. The only place that retains the high level of correlation is in the PCC region, particularly near the region from which the comparison time course was calculated. All other correlations in DMN regions are less obvious than in its seed-based analogue.

Seed-based SCA is shown in Figure 3.15. This correlation map differs greatly from the seed-based and region-based methods of SC-SCA. As shown, the area of highest correlation is directly around the seed point, which has a correlation value of 1. Other DMN regions including the mPFC and right and left IPL have substantially lower correlation values around 0.6-0.7. Looking at this from a network perspective, it would not be unreasonable to consider

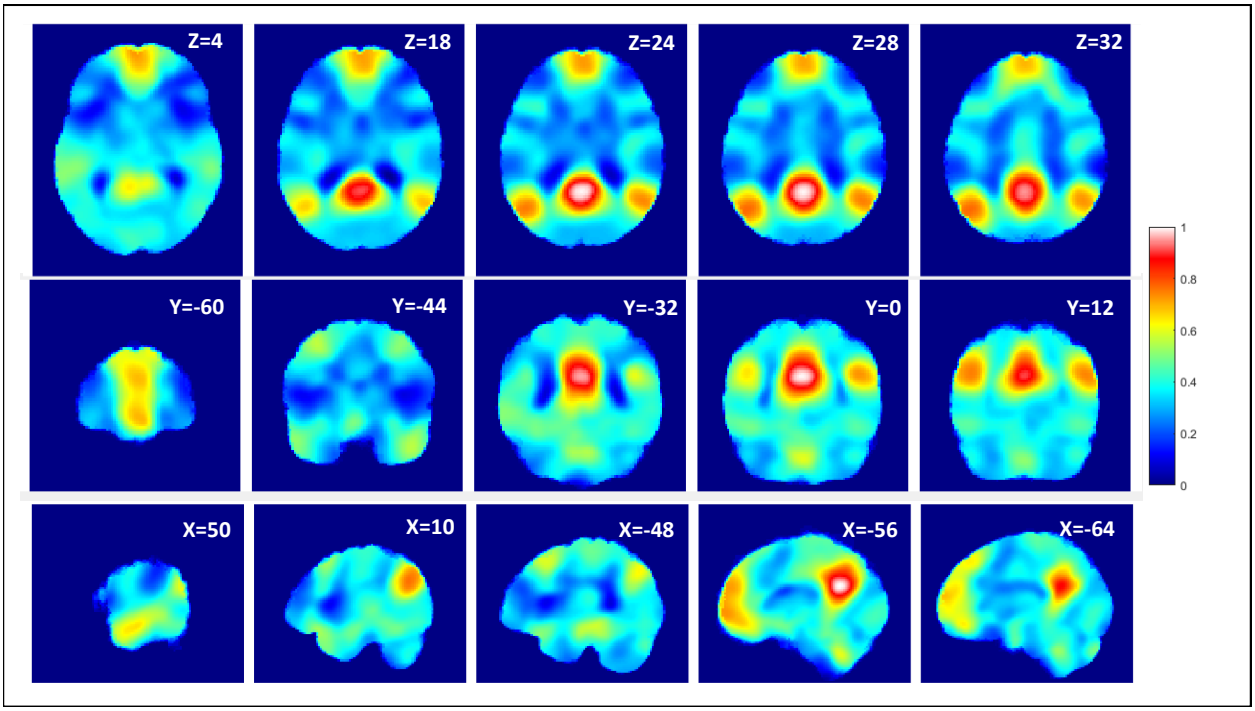


Figure 3.15: Group correlation map for seed-based SCA

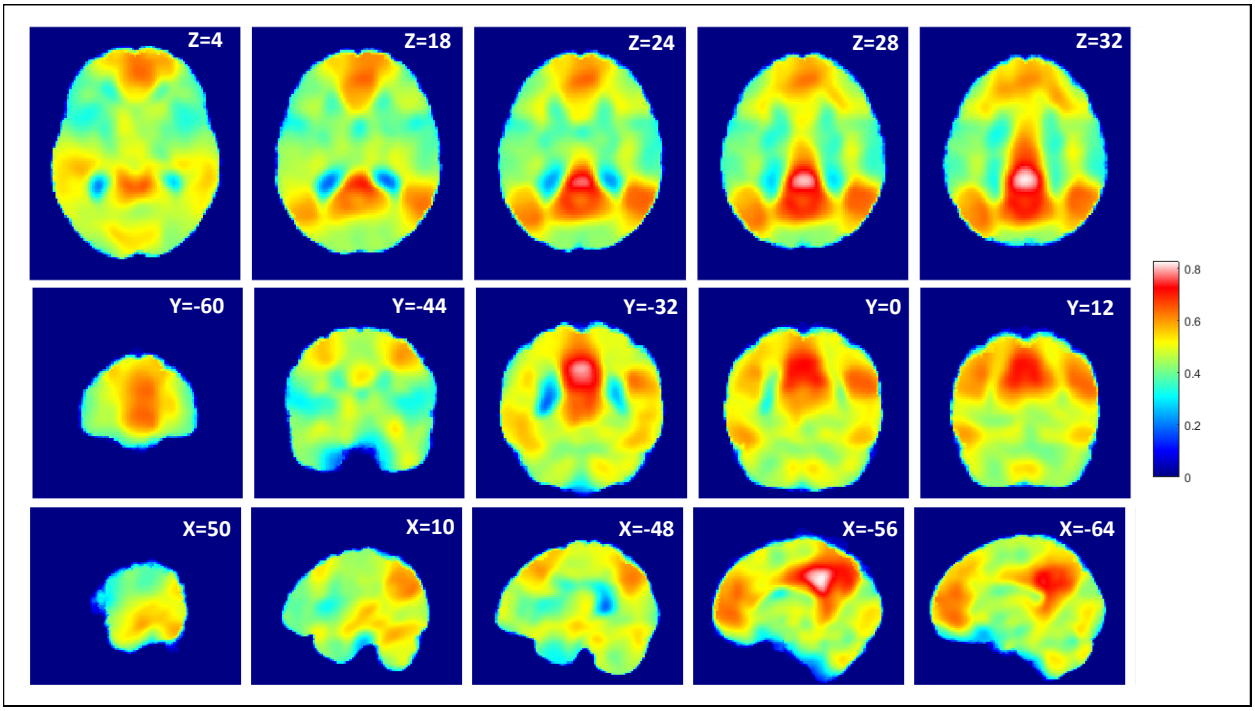


Figure 3.16: Group correlation map for region-based SCA

the PCC and pC as their own private network, and the left and right IPL and the mPFC as a totally different network. This is because their correlation values are so disparate. The minimum value of the correlation map generated by this method is -0.05, which is only a very slight negative value, and the maximum is 1.

The correlation map resulting from region-based SCA is shown in Figure 3.16. As shown, the map appears to be similar to the region-based SC-SCA method in that it has a somewhat ambient correlation value of about 0.5 with nearly the whole brain. There is a maximum in the correlation in the region where the time course was calculated from in the PCC. However, again and similarly to seed-based SCA, there is relatively little correlation to the right and left IPL and the mPFC. The majority of the maxima in correlation appears to be within the PCC region. The maximum value of correlation in this map is 0.82, and the minimum is slightly negative at -0.03.

Assessment of statistical significance

Figure 3.17 shows a comparison between the distributions of the median of the standard deviations of voxels across 1,000 bootstrap samples. As shown, seed-based SCA has the highest median value of 0.23, and it visually differs in location from the other three methods. The rest of the methods have similar medians of 0.1811, 0.1712, and 0.1615 for seed-based SC-SCA, region-based SC-SCA, and region-based SCA.

Figure 3.18 shows the sampling distribution of μ_{diff} for the comparison between seed-based SC-SCA and seed-based SCA. The 95% confidence interval for this comparison is (-0.0876, -0.0178). As shown, the entirety of the distribution is to the left of 0. This means that there is definitively a statistically significant difference between seed-based SC-SCA and seed-based SCA. Since the differences are negative, this shows that seed-based SC-SCA has a statistically significant lower median than that of seed-based SCA, which presents an

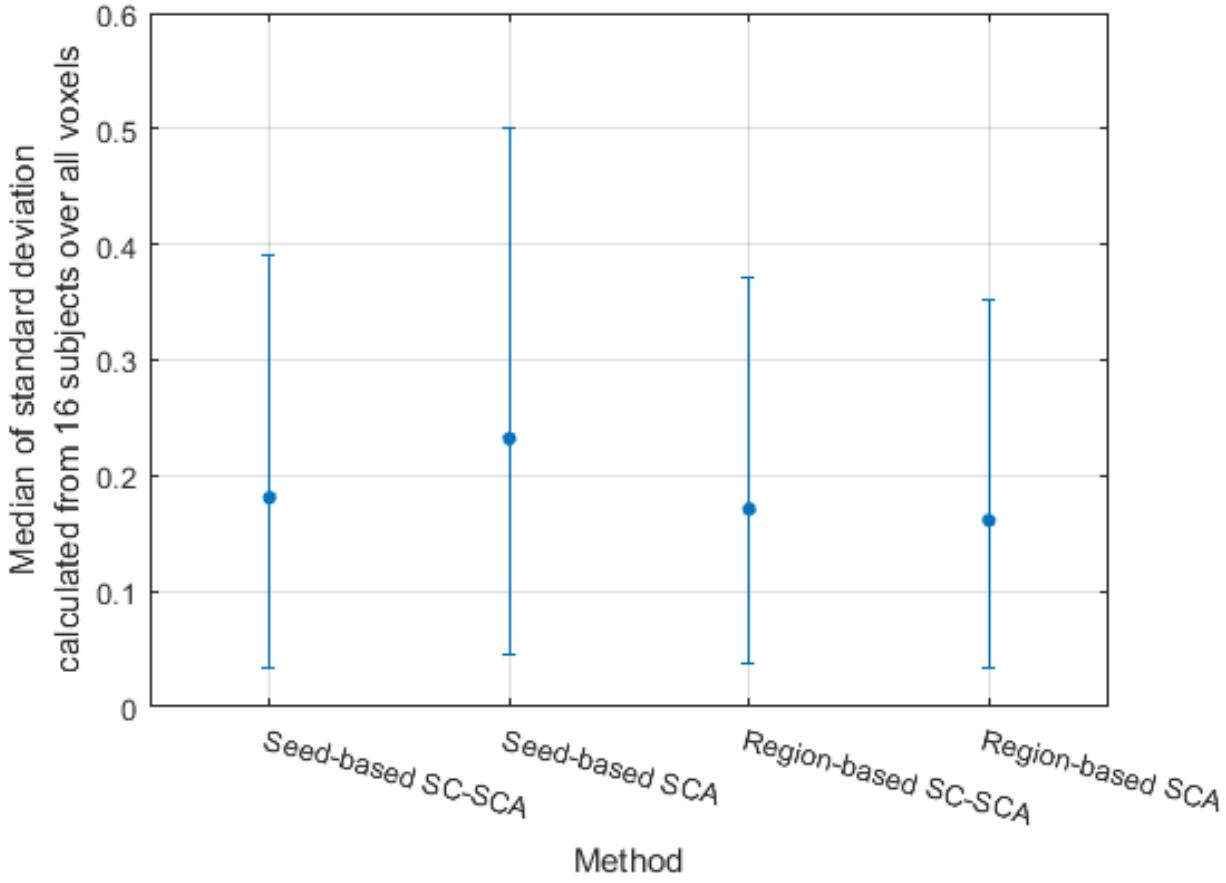


Figure 3.17: Distribution of medians bootstrapped from four methods of correlation analysis

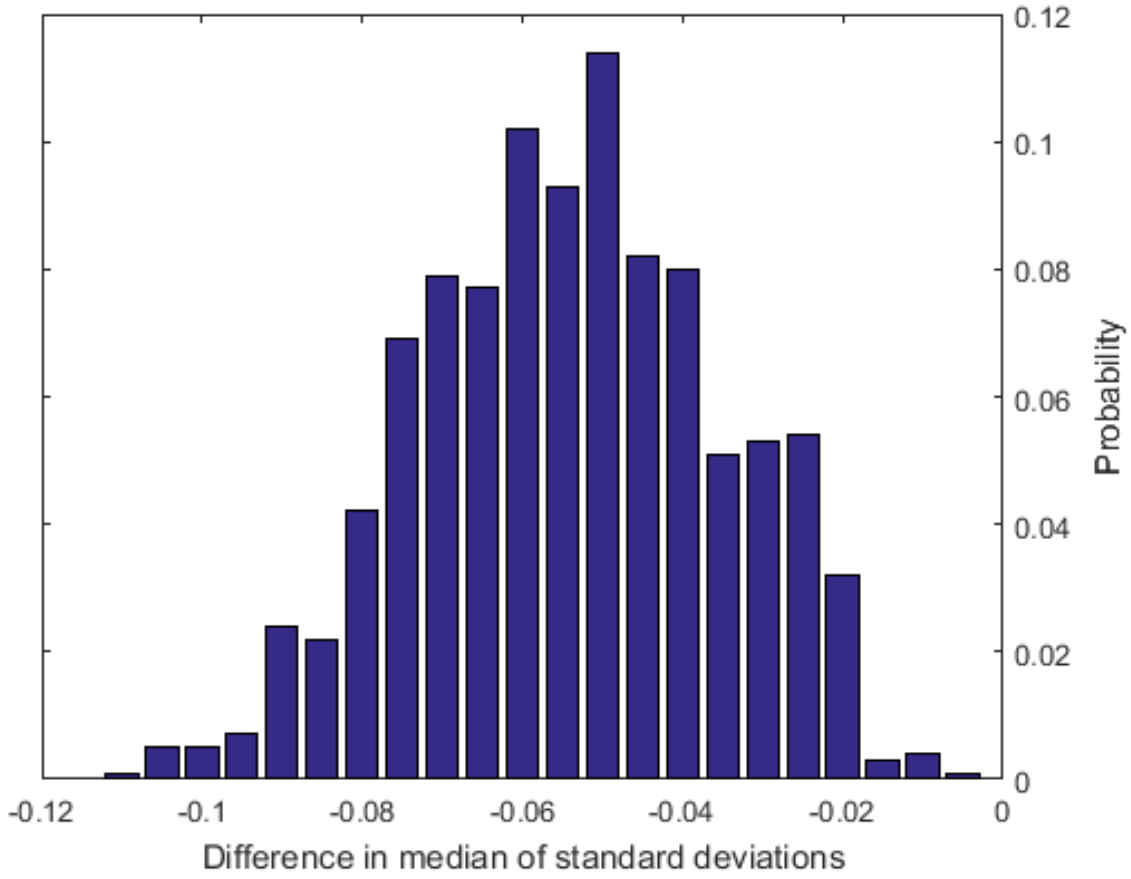


Figure 3.18: Comparison of difference in medians of standard deviations of voxels for seed-based SC-SCA and seed-based SCA

advantage of SC-SCA over SCA: there is statistically significant lower inter-subject variance in the correlation maps created by SC-SCA.

The sampling distribution of μ_{diff} for the two region-based methods for SC-SCA and SCA are shown in Figure 3.19. The figure shows that the distribution overlaps 0 and has values on both the positive and negative side. For this comparison, the confidence interval of the data was found to be (-0.0087, 0.0299). A value near the lower bound for 2.5% of the data (-0.0087) is shown in the figure, and as shown, it is not on the right side of the y-axis; in other words, this means that the comparison of these methods do not meet the criteria for statistical significance of a difference between them, and quantitatively, there is no particular

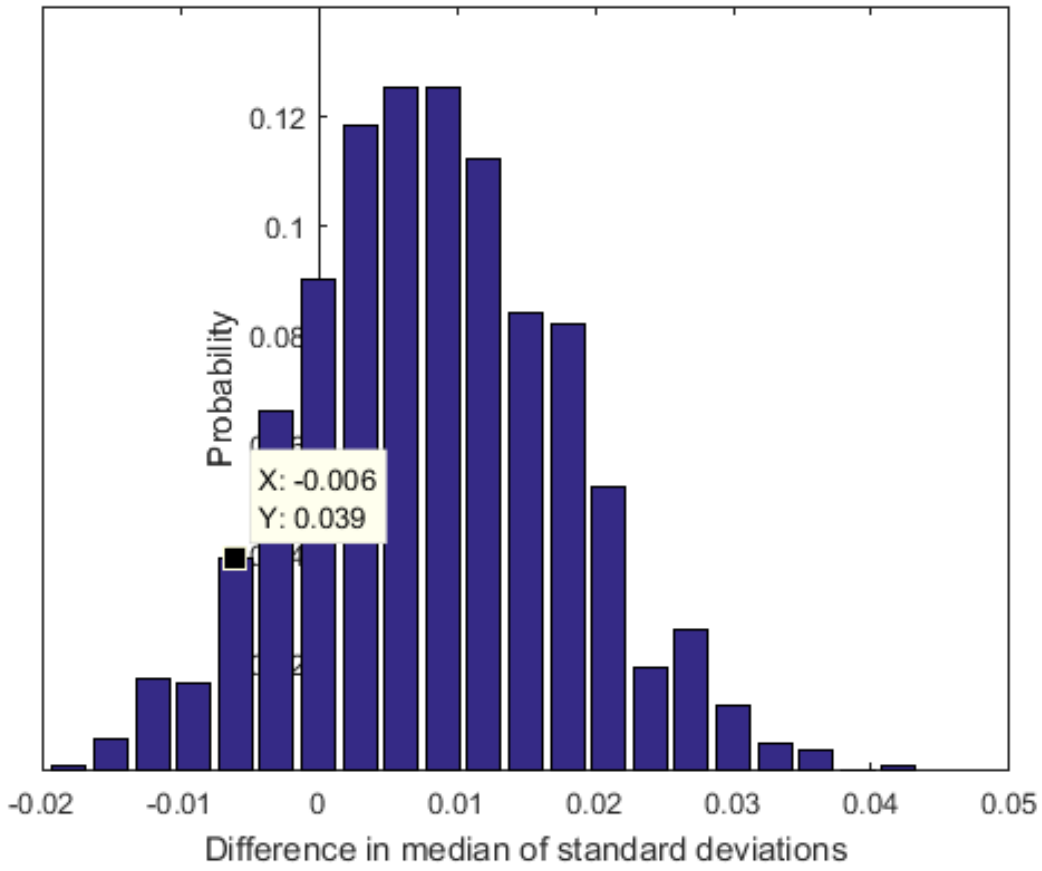


Figure 3.19: Comparison of difference in medians of standard deviations of voxels for region-based SC-SCA and region-based SCA

difference in using one method rather than the other. This means that it is equally effective to use region-based SC-SCA or region-based SCA; they produce similar results in terms of inter-subject variability for the correlation maps created.

Figure 3.20 shows the sampling distribution of μ_{diff} for seed-based SC-SCA and region-based SC-SCA. As shown, the distribution overlaps zero and results on data on both sides of the y-axis. The 95% confidence interval of the data can be found in the range (-0.0086, 0.0265), and as shown in the figure, the lower bound falls on the left side of the y-axis. Thus, again, the difference between these two methods is not statistically significant, meaning that

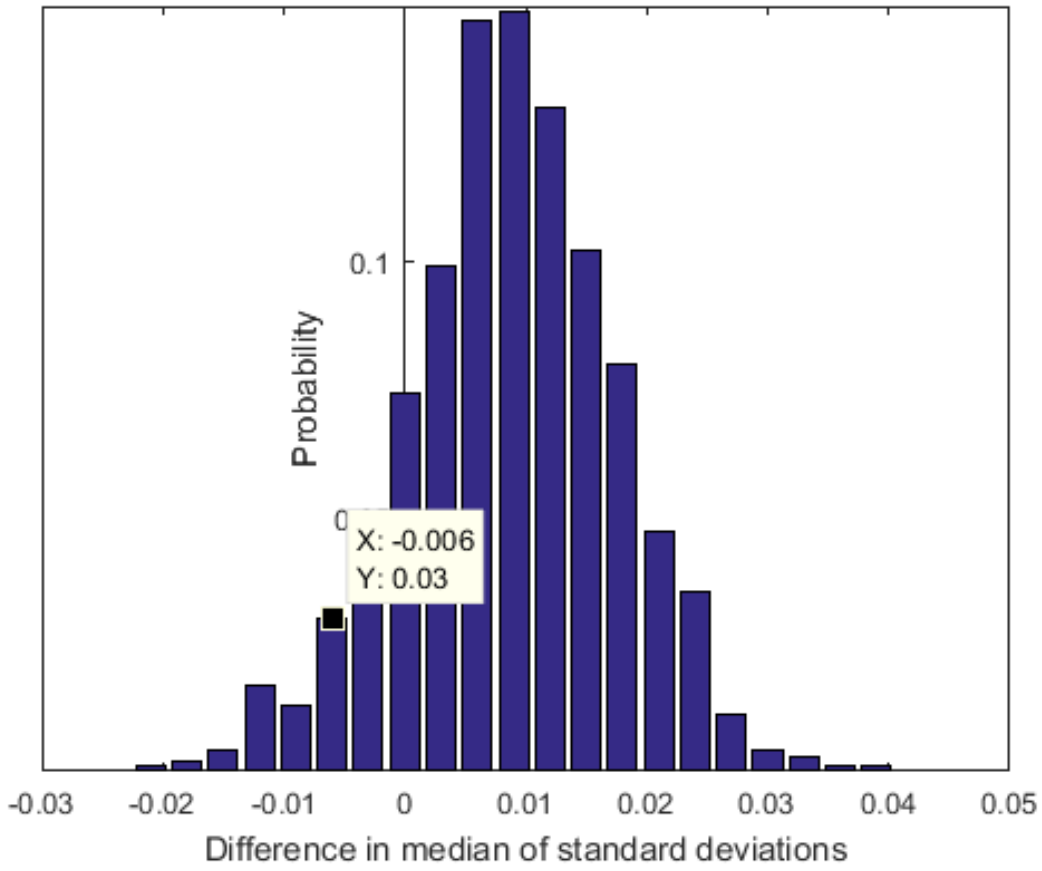


Figure 3.20: Comparison of difference in medians of standard deviations of voxels for seed-based SC-SCA and region-based SC-SCA

utilizing a region-based method vs. a single seed point does not make a substantial difference when using SC-SCA. SC-SCA performs equally well with both methods.

The comparison between using a seed-based approach versus a region-based approach for SCA is shown in Figure 3.21. In this case, the entirety of the distribution is on the either on the y-axis or to the right side of the y-axis, which means a statistically significant difference is found between these two methods. The 95% confidence interval here is (0.0250, 0.1171), and statistical significance can again be seen by observing the y-axis with respect to the confidence interval lower bound, 0.0250; since the y-axis is less than this lower bound, the results are statistically significant. Considering the side of the graph on which the distribution falls, this

means that seed-based SCA has higher inter-subject variance in the correlation maps than using its region-based analogue. This means that for SCA, the choice of a seed-based method versus a region-based method again does, in fact, matter. This confirms the point of view often expressed in the seed-based correlation literature for resting state fMRI; a region-based method is often preferred to a seed-based method for SCA. Compared to the previous finding about seed-based SC-SCA and region-based SC-SCA, this presents a particular advantage for SC-SCA; since region-based methods are used in SCA to overcome the problem of producing different networks for different close seed points, SC-SCA alone is able to overcome this problem as shown by its ability to produce similar maps, regardless of whether seed-based SC-SCA or region-based SC-SCA is utilized.

Figure 3.22 shows the comparison between seed-based SC-SCA versus the region-based version of SCA. As shown, the distribution overlaps the y-axis and so it is distributed about y=0. In order to determine statistical significance, the bounds of the 95% confidence interval of the distribution were calculated to be (-0.0062, 0.0479). As shown, the lower bound falls to the left hand side of the y axis (i.e, the y-axis is greater than the lower bound), so that the difference between these two methods is not statistically significant. This is also an advantage conferred to seed-based SC-SCA, since this comparison shows that it produces similar results to region-based SCA, which has shown to be superior to seed-based SCA.

Figure 3.23 shows the distribution of μ_{diff} for the comparison between region-based SC-SCA and seed-based SCA. The 95% confidence interval for the comparison between these two methods is (-0.1003, -0.0240). As shown, the distribution lies entirely on the left-hand side of the y-axis, showing that the difference between these two methods is statistically significant. In interpreting the directionality of the difference, the negative signifies that region-based SC-SCA has a lower variance between the maps produced by this method than those produced by seed-based SCA. This shows that using the region-based method with SC-SCA is statistically and quantitatively different from that of seed-based SCA, and confers

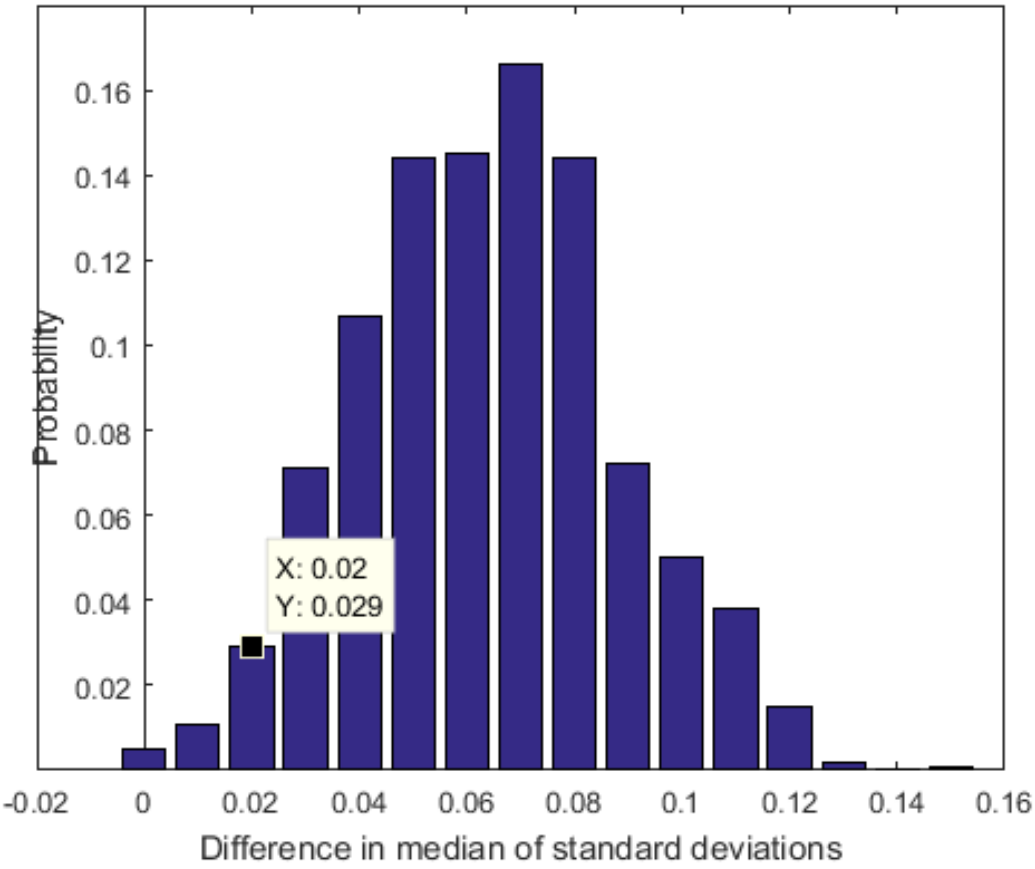


Figure 3.21: Comparison of difference in medians of standard deviations of voxels for seed-based SCA and region-based SCA

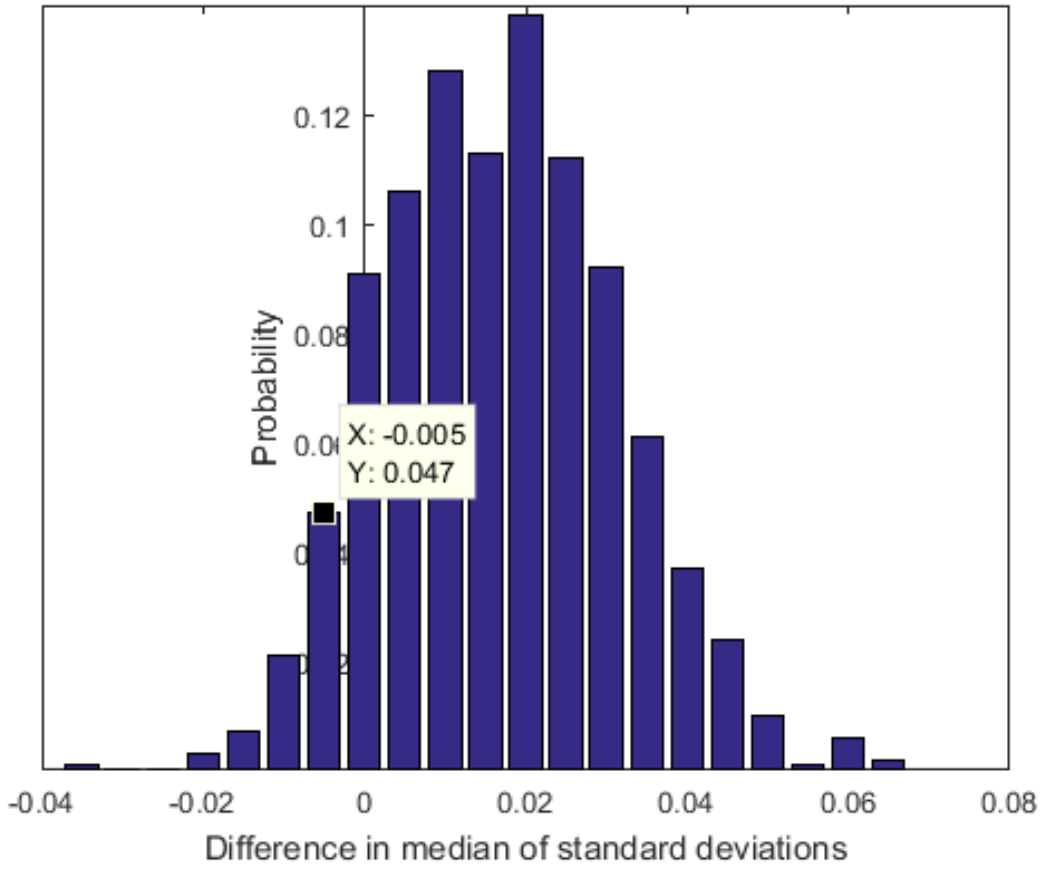


Figure 3.22: Comparison of difference in medians of standard deviations of voxels for seed-based SC-SCA and region-based SCA

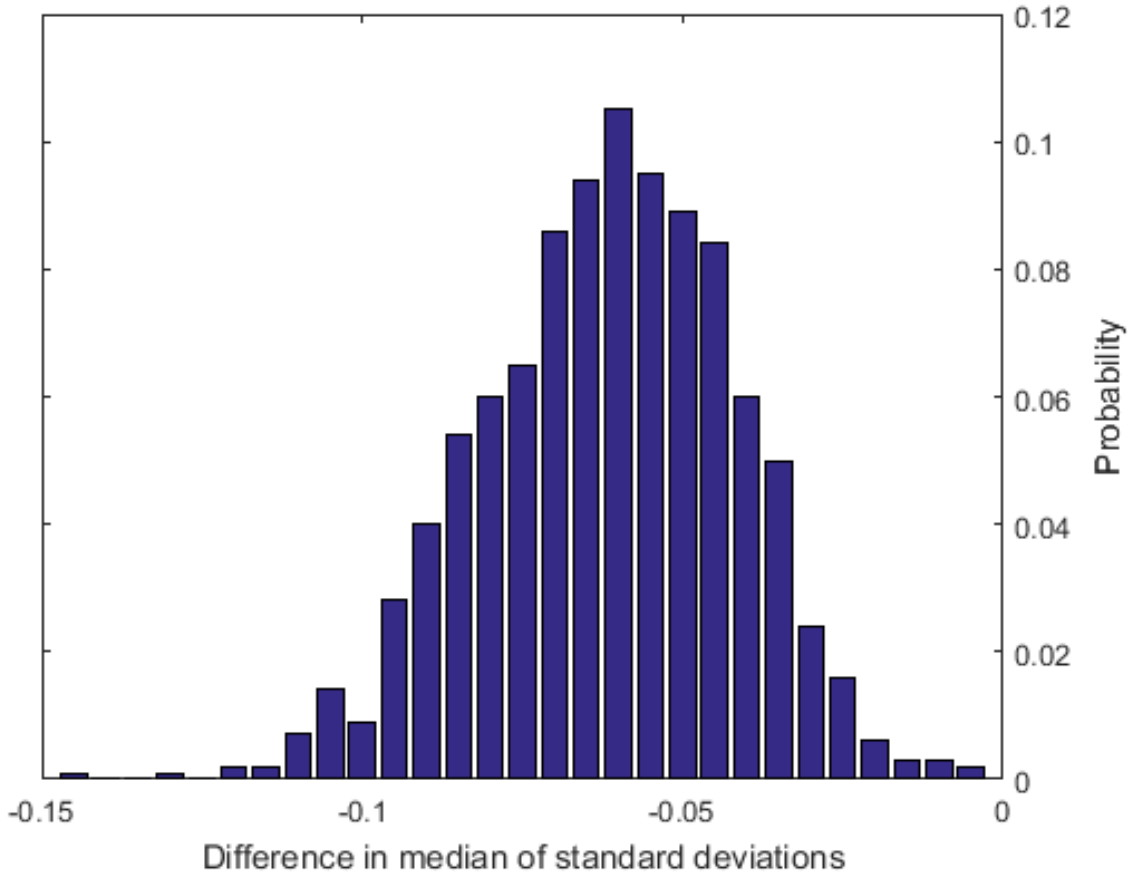


Figure 3.23: Comparison of difference in medians of standard deviations of voxels for region-based SC-SCA and seed-based SCA

yet another advantage to SC-SCA. Additionally, previous findings are consistent with these findings; it was found that there was no significant difference between region-based and seed-based SC-SCA, and seed-based SC-SCA is superior to seed-based SCA in terms of inter-subject variance.

The results of the following 3 comparisons suggest that it is immaterial which method is used in the following cases:

1. The two region-based methods: region-based SC-SCA and region-based SCA
2. Seed-based SC-SCA or region-based SCA

3. Seed-based SC-SCA or region-based SC-SCA

This shows that the two region-based methods are quantitatively similar, and it does not matter which is used because they produce similar results. Using a region-based method may introduce a disadvantage in general due to the difficulty in identifying a region as well as possible inhomogeneity in the time courses gathered from the region. Additionally, it is irrelevant whether seed-based SC-SCA and region-based SCA is used; this presents an advantage for SC-SCA because the seed-based method can possibly be easier to use by identifying a single seed voxel and due to the probabilistic advantages conferred by the method. Additionally, there are no significant differences between using seed-based vs region-based SC-SCA; using the region does not improve the consistency among the maps created with the subjects. This is additionally another advantage of SC-SCA.

On the other hand, the results of the other half of the comparisons suggest that there are distinct differences in the standard deviation of the maps that are produced in the following cases:

1. Seed-based SC-SCA vs seed-based SCA
2. Region-based SC-SCA vs seed-based SCA
3. Seed-based SCA vs. region-based SCA

Firstly, there are disparities in the stability of voxels as measured by the standard deviation in the maps created by seed-based SC-SCA vs. seed-based SCA. It is not clear that seed-based SC-SCA is more stable in terms of inter-subject variability of voxels between generated maps. Also, there are differences between using region-based SC-SCA vs seed-based SCA. This can also be easily seen in the maps produced in the figures comparison of the correlation maps of the 4 different methods. Region-based SC-SCA has lower variance in the maps between subjects than seed-based SCA. Also, there is a statistical difference in using seed-based SCA vs region-based SCA, where region-based SCA has been shown to be superior as measured

by inter-subject map variability. This makes sense since region-based SCA methods are often used in the literature instead of seed-based methods. Region-based methods, particularly with SCA, increase the stability of the generated map in terms of the location of the seed region. This also makes sense in this context; choosing the same seed in each subject may correspond to slightly different regions or voxels due to the acquisition and preprocessing of the data, thus slightly different maps per subject are created from using the seed-based method. SCA is particularly sensitive to these discrepancies, whereas the results show that SC-SCA does not improve in stability due to using a region as the seeding time course.

3.3.6 Decomposition of networks into constituent ICs

Figures 3.24 and 3.25 show the decomposition of the SCA and SC-SCA thresholded networks into their constituent 10 ICs. These figures are divided for clarity and convenience. Figure 3.24 and the first 3 frames of Figure 3.25 show a set of ICs in which an SC-SCA match was found for each SCA IC. As shown, although each of the ICs generally matches its SCA analogue, the SC-SCA ICs are much more diffuse and expanded.

Additionally, the ICs shown resemble several of the RSNs and cover several anatomical areas of the brain. Similar IC (SIC) 1 shows a part of the DMN, with the PCC and right and left IPL well-represented. There is a small area identified which is located within the mPFC. As shown, the PCC region is divided into two parts in the SCA network, but the two regions are connected into one contiguous region in the SC-SCA network. SIC 6 displays the mPFC region of the DMN, and again the IC in the SC-SCA method is more widespread than in the SCA method.

SICs 2 and 3 resemble the DAN, divided by the more prominent side that is represented (i.e., SIC 2 has a more prominent right part of the DAN and SIC 3 has a more prominent left part of the DAN). This is particularly pronounced in the SC-SCA network IC. SIC 4 has

regions in common for SCA and SC-SCA, but there are several more regions of activation in the SC-SCA formulation.

SIC 5 most strongly shows the difference in the FC formulation of SCA vs the SC-SCA formulation. As shown, the ICs contain regions in the frontal lobe and in the PCC area; however, the SC-SCA region additionally shows the path of activation between them. It additionally demonstrates connectivity with the cerebellum and lateral sides of the brain. SIC 7 shows a similar level of connectivity between regions. Again, as shown in the SC-SCA method, there is enhanced connectivity between a PCC region and an mPFC region, whereas in the SCA method, these regions are completely separate.

SIC 8 and Dissimilar ICs (DSICs) 1 and 2 show regions and networks of the brain that are highly recognizable. SIC 8 shows a small piece of the cerebellum for the SCA method, but shows a larger region of connectivity in the cerebellum, on the brain stem, and in the occipital lobe of the brain. DSICs 1 and 2 for SCA appear to both be regions of the cingulate cortex. On the other hand, DSICs 1 and 2 for SC-SCA appear to be the somatosensory network and the occipital lobe.

SC-SCA presents several advantages over SCA, the most important being the Bayesian inferences and calculations that can be performed due to the probabilistic nature of the method. Additionally, the method relies less on the identification of the initial seed - firstly, as shown by the sensitivity analyses, but also because the initial seed is only used to identify points that are highly correlated to the initial seed point. This corresponds in this work to approximately 2,000 points which are averaged to determine the network of correlation. SC-SCA also appears to have fewer artifacts around the seed region as compared to SCA, and is able to identify the finer structures of connection despite being near the seed region. Additionally, because the network generated by SC-SCA is larger and shows secondary connectivity (i.e., connectivity to voxels highly connected to the PCC), the decomposition of the network into ICs is highly illuminating. Instead of simply disparate regions of the brain,

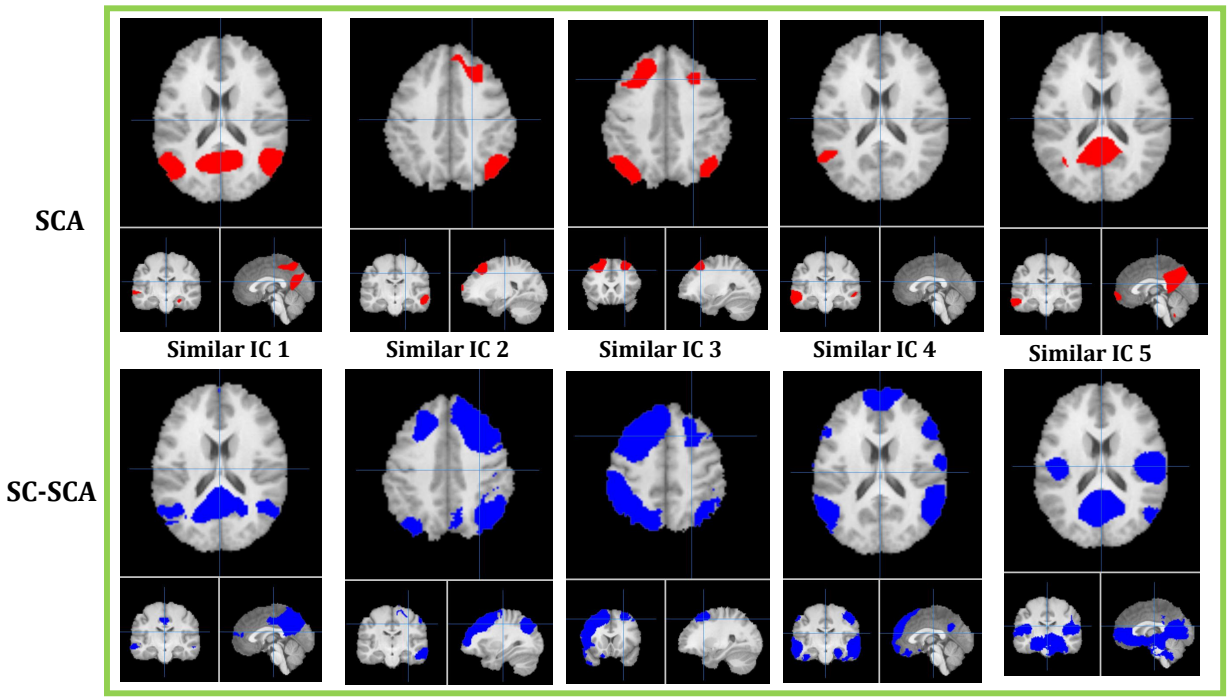


Figure 3.24: Decomposition of generated SCA (top) and SC-SCA (bottom) networks into constituent networks. The ICs for each method which are similar are shown with a green outline.

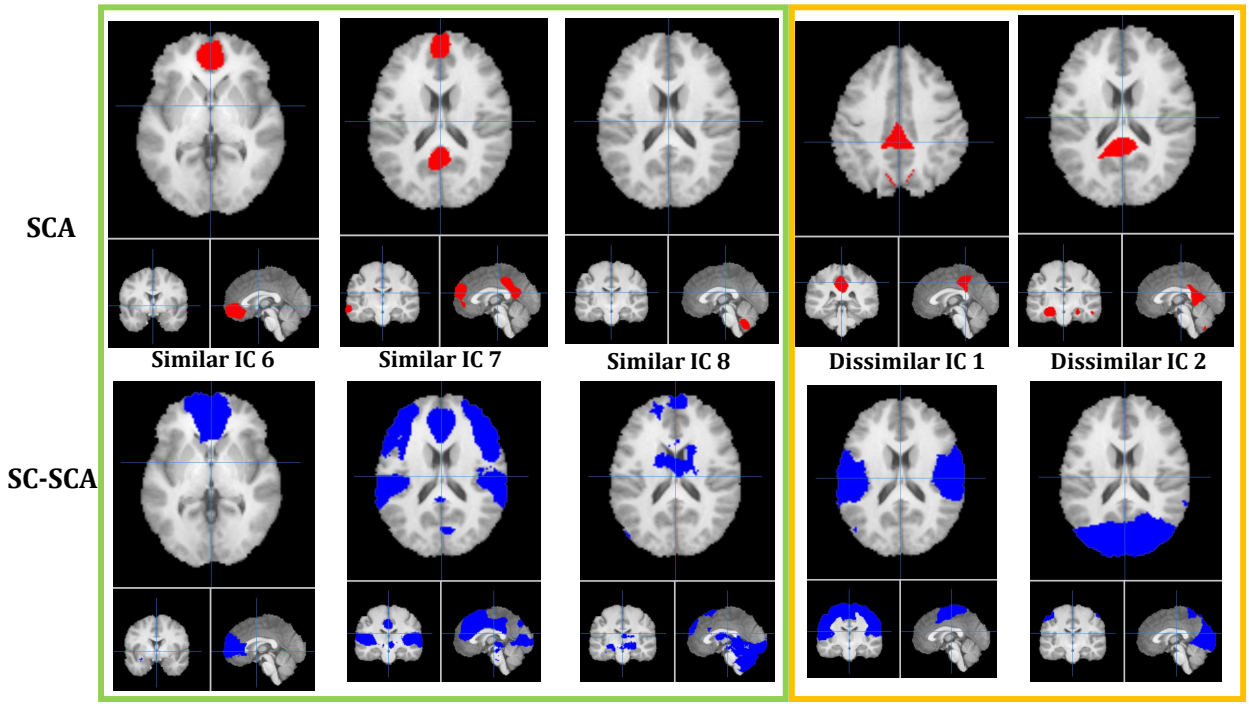


Figure 3.25: Decomposition of generated SCA (top) and SC-SCA (bottom) networks into constituent networks. The ICs for each method which are similar are shown with a green outline. Groups of ICs with no matches are shown with an orange outline.

more networks are shown, and the pathways and areas of activation between regions are highlighted. Thus, in addition to the quantitative advantages shown over SCA, the qualitative advantages of SC-SCA are also numerous.

CHAPTER 4

INSTANTANEOUS GLOBAL CORRELATION ANALYSIS (IGCA) FOR AVALANCHE ANALYSIS

Abstract

The purpose of this work is three-fold in studying dynamic functional connectivity (dFC) at the time scale of seconds. My first contribution is to use data-driven techniques with minimal extrinsic expert knowledge to distinguish between the FC of the brain during periods of large co-activations (avalanches), and during intervals when the brain has low co-activations. The method is inspired by the point process technique [107, 54], but instead defines a time *interval* which identifies avalanching periods based on the magnitude of whole-brain correlation. This allows the data to be reduced into two classes - avalanching periods and non-avalanching periods. Secondly, a set of dFC co-activation patterns (CAPs) will then be characterized, and are the networks of connectivity that are formed during avalanching periods and non-avalanching periods. Lastly, the spatial propagation of the avalanche will be demonstrated by using the regions of highest activation to track the epicenters of activity.

4.1 Introduction

4.1.1 Avalanches

Recent studies have suggested that brain activity remains in a critical state throughout processing, and localized neural activity may cause cascades in neural activity throughout the brain [108, 109, 110]. Thus, brain activity may be dominated by brief peaks or intervals of brain activity [107, 111, 112]. Appropriately termed brain *avalanches*, these periods of

high organized activity throughout a significant portion of brain regions may be responsible for the majority of the brain's behavior during a scan. Thus, important network information may be overlooked when studying functional connectivity using the entire period of the scan rather than brief instances [57]. This likely extends to the concept of the sliding window analysis, which may indiscriminately average over regions with dissimilar temporal behavior, leading to convoluted results.

A number of methods have additionally been introduced to study brain behaviors at brief periods of time including via point processes and peaks in the BOLD signal [107, 54, 111, 112]. Tagliazucchi, 2012 [107] approaches the detection of avalanches through using a point-process approach, in which a seed region is identified, and suprathrehold crossings, where the threshold was considered to be one standard deviation above the mean, were considered to be points in the point process of interest. The thresholded time frames were used to extract clusters of voxels which were activated during these frames through a nearest neighbor clustering scheme. These clusters were then organized into avalanches by investigating the spatial overlap between clusters and transition from inactivation to activation of these clusters. This work concluded that the brain persists in a critical state and avalanches correspond to a system undergoing order-disorder phase transitions, and that resting state activation maps can be created by a few subset of the points in the entire time series for a set of voxels.

A second of the Tagliazucchi publications from 2016 [54] came to similar conclusions about the condensation of the entire time series into a few different points, but followed a slightly different methodology. In this method, the binarized suprathrehold crossings of a seed point were utilized to calculate a correlation-like calculation of co-activation matrices between voxels. The functional connectivity between all voxels in the brain was then calculated by summing all of the co-activation matrices over time. This resultant matrix reflects the coupling between signals. The group determined that less than 1% of the data was necessary to reproduce findings obtained from a full time series.

Two works by Liu, 2013 [111, 112] follow a similar methodology of using seed points to extract suprathreshold crossings as in [107, 54]. However, they deviate from the methodology by then clustering the maps derived from the point process thresholding into networks with similar spatial extent. Through this method using the posterior cingulate cortex (PCC) as a seed, they are able to divide the default mode network (DMN) into co-activation maps. This group produced a similar finding that very little of the data (15%) is necessary to almost perfectly reconstruct the PCC-seeded correlation map generated from all of the data.

In addition to demonstrating that the majority of brain activity occurs in a few frames, they also suggest that all of the data can be condensed into a few frames which are able to accurately represent the entire time series. This is a desirable quality with the increasing data sizes from fMRI scans due to advances in fMRI technology which have created higher spatial and temporal resolution in fMRI scans. This reduction in data size while simultaneously preserving most of the information in the data produces accurate results with substantially decreased data size, faster computational times, and decreased computational hardware demands.

4.2 Methodology

This section describes the steps necessary to execute the method and the procedures taken to evaluate the results. Processing was carried out using a combination of Matlab R2016a (The Mathworks Company; Natick, MA, USA), Statistical Parameteric Mapping, Version 12 (SPM12) (The Wellcome Department of Neuroscience; Oxford, UK), and the Brain Extraction Tool (BET) [95, 96] and Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) packages of the FMRIB Software Library (FSL) (FMRIB Analysis Group; Oxford, UK) [97].

4.2.1 Data acquisition

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

The subjects were drawn from a population of 1200 healthy adult volunteers in the age range of 22-35 years. The dataset contains high-resolution T1-weighted sMRI and fMRI brain images. A total of 16 sets of fMRI data were randomly selected from the HCP database.

All HCP subjects were scanned using a customized Siemens Skyra 3T scanner with identical imaging parameters. The T1w image was acquired using the 3D MPRAGE sequence with 0.7mm isotropic resolution ((FOV = 224 mm, matrix = 320, 256 sagittal slices in a single slab), repetition time (TR) = 2400 ms, echo time (TE) = 2.14 ms, inversion time (TI) = 1000 ms, flip angle (FA) = 8°, bandwidth (BW) = 210 Hz per pixel, echo spacing (ES) = 7.6 ms). The fMRI scans were obtained using the following parameters: TR=720 ms, TE=33 ms, multiband factor=8, image matrix=104x90, 72 slices, 1200 volumes (time points), and isotropic slice size=2x2x2 mm^3. Full details regarding the acquisition of the data can be found in [98].

4.2.2 Data preprocessing

The HCP dataset minimizes the amount of preprocessing applied to the data; relevant preprocessing steps include removal of spatial artifacts and distortions and registration of surfaces and volumes to standard volume and surface spaces. Further preprocessing steps were applied subsequent to retrieving the data. First, each voxel time series was temporally filtered using a finite impulse response (FIR) bandpass filter (0.01Hz-0.1Hz) and linearly detrended. Since convolution of an N^{th}-order FIR filter with the data produces a finite im-

pulse response of length (N+1) at the beginning of the data and due to other computational limitations, a total of 200 time points were removed from the beginning and end of the data. The data were a resultant 800 volumes in length.

The temporally filtered and trimmed data was then spatially smoothed in FSL using a 6mm full width half maximum (FWHM) Gaussian filter. The data was then normalized by demeaning and dividing by the standard deviation of each voxel time series.

4.2.3 Temporal avalanche detection and CAP extraction

Following the additional preprocessing steps applied to the data, an instantaneous whole brain correlation (WBC) analysis was then performed. The full processing pipeline is shown in Figure 4.1. At each of the remaining 800 time points, a WBC for a single time point was performed. The instantaneous WBCs for a single time point t were calculated according to following equation:

$$WBC_t = \frac{1}{2}\sum^{i}\sum^{j}\frac{1}{n}\frac{(\mathbf{x_i} - \bar{x_i})(\mathbf{y_j} - \bar{y_j})}{\sigma_{x,i}\sigma_{y,j}} \tag{4.2.1}$$

where x and y are voxels and i and j iterate over v voxels for $i \neq j$. An important difference between the pairwise correlation calculation in Equation 4.2.1 which reflects its *instantaneous* nature is that the length of the time series is $n = 1$, but $\bar{x}$ and $\bar{y}$ are calculated over the entire time series for voxels i and j. Following this correlation, the mean value of each WBC was computed, leading to an 800-point plot of average WBC for each subject. High and low correlation thresholds γ_H and γ_L were then identified for each subject, such that values in excess of γ_H value were considered in the *avalanching state*, while time points lower than γ_L were considered in the *non-avalanching state*. In order to ensure consistency among subjects, γ_H was set at the value of one standard deviation greater than the mean of the entire WBC series, and γ_L was set to be the mean of the entire WBC series.

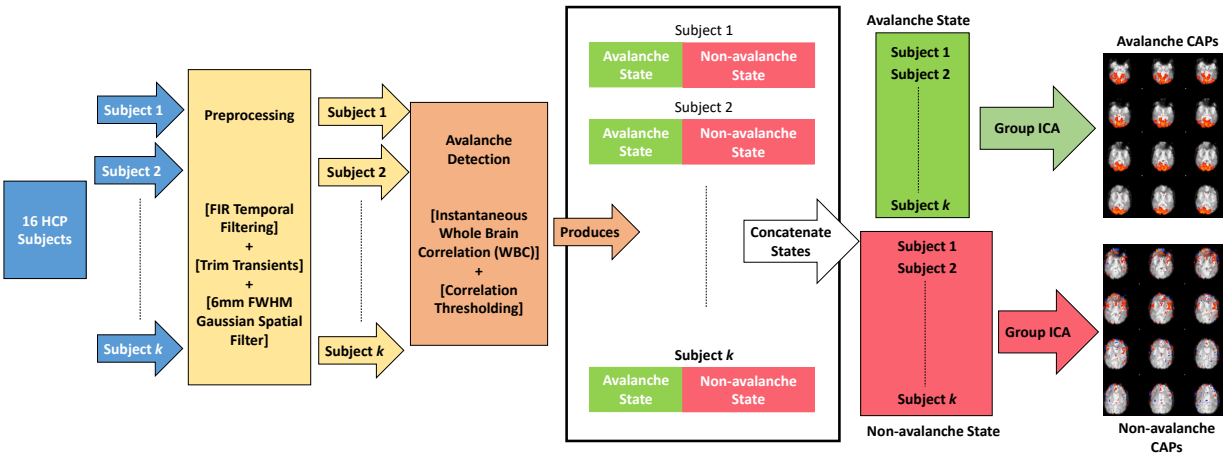


Figure 4.1: Processing pipeline for avalanche analysis

The dataset for each subject was then divided as shown in Figure 4.1 using these time points. An avalanches state dataset was created by including only volumes identified as being in the avalanche state; a complementary dataset was also formed using only volumes classified as the non-avalanching state. This divided data for all of the subjects were then temporally concatenated, and group ICA was performed using MELODIC in FSL. The independent components (ICs) generated were defined to be the CAPs for the avalanche state versus the non-avalanching state.

Additionally, a comparison between CAPs in the avalanching versus non-avalanching state was performed. Assuming that CAPs are similar although not identical between states, Dice's coefficient [101] D (a measurement of similarity) was calculated between each of the active CAPs and all of the inactive CAPs to identify which inactive CAPs correspond to those identified during the avalanche activation period. Dice's coefficient was calculated as shown in Equation 4.2.2, and ranges from 0 to 1, where 0 is completely different and 1 is identical.

$$D = \frac{2(|X \cap Y|)}{|X| + |Y|} \tag{4.2.2}$$

4.2.4 Spatial avalanche detection and identification

The avalanche detection algorithm will proceed similarly to the method described by Tagliazucchi et. al. in [107] with an important differentiation in the originating data to determine the clusters. In order to determine clusters of high activity, the WBC will be first normalized at each time point; that is, each instantaneous WBC will be de-meaned and divided by the standard deviation, producing a z-score map for each time point. These will then be thresholded at a significance level of α=0.05. This produces voxels of significant activation in the data for each of the time points. This differs from [107] in that the threshold is applied to the WBC, not based on suprathreshold crossings of the BOLD signal of a seed region.

The voxels were then clustered into groups by identifying if another activated voxel was in the vicinity of its 3D nearest neighbors (i.e., 26 points surrounding the voxel). For the set of voxels with nearest neighbors activated, this was grouped into a cluster. The location of the 3D centroid of the cluster was first determined, and then the maximum value of the WBC within the cluster was identified and the location recorded. The locations of both the centroid and the local maxima were matched to the Harvard-Oxford regions of interest [113] in order to determine their physiological brain region. In performing this labeling, both the lateralized cortical and subcortical atlases were used; this means that for some regions, there are potentially no labels, and for others, there are potentially two labels. The regions that contained no labels were eliminated, and the region corresponding to the lateralized cortical atlas was utilized in the case that there were multiple labels.

4.2.5 Investigation of avalanche properties

In order to investigate properties of avalanches such as the frequency of a particular type of avalanche's occurence or the frequency of a particular type of avalanche common between

all subjects, the avalanche was then converted to a directed graph. The nodes of the graph are the locations of the WBC correlation maxima of the cluster, and the information was considered to flow from one node to another if the two corresponding clusters were spatially overlapping at a time t and $t+1$. The information was considered to flow from the cluster at time t to time $t+1$. Additionally, an adjacency matrix was generated for each of the graphs for simplicity of calculations.

In order to determine the frequency of each type of avalanche, all the graphs for all 16 subjects were investigated simultaneously. In order to determine the similarity between graphs, the element-wise multiplication of two graph adjacency matrices A_{Gi} and A_{Gj} (i.e., $A_{Gi,Gj} = A_{Gi} \cdot^* A_{Gj}$) was calculated, and the sum of all the entries in the matrix was calculated; this quantity is representative of the number of nodes and edges which are common between both the graphs. In order to compare all the graphs of different sizes, this value was scaled by the total number of edges in both graphs. This calculation is similar to a Dice coefficient calculation, and is shown in the following equation:

$$D_{Gi,Gj} = \frac{2 * \sum(A_{Gi} \cdot^* A_{Gj})}{\sum(A_{Gi}) + \sum(A_{Gj})} \tag{4.2.3}$$

This pairwise calculation of similarity was entered into a matrix S, a $n_G \times n_G$ (where n_G is the total number of graphs for all of the subjects) matrix, where the i, j^{th} entry corresponds to $D_{Gi,Gj}$. In the case that two graphs are completely identical, the coefficient will be 1, and in the case they are completely different, the coefficient will be zero.

In order to calculate the graphs with the highest rate of occurrence, S was binarized to S_b to identify only that graphs that had 100% similarity with another graph. The sum of each row of S_b was calculated to identify the most popular graphs. This list was then sorted to remove redundant graphs and calculate the total number of times a particular graph occurred. In this way, the frequency of each graph was calculated.

4.2.6 Evaluation of avalanche trajectory

Another interesting aspect of a spatial characterization of avalanches is to determine how local maxima in WBC values appear to traverse the brain. A question of interest is whether these maxima follow known trajectories in the brain. In order to investigate this question, the Johns Hopkins University (JHU) white matter tractography atlas [114] was employed, which identifies certain well-known white matter fiber tracts.

In order to determine whether the maxima of the avalanches traverse a fiber tract, the locations of the maxima in the avalanche were first identified. Then, the percent intersection as defined in the following equation was calculated:

$$\%\, intersection = \frac{|tract\, points \cap avalanche\, maxima\, points|}{|avalanche\, maxima\, points|} * 100 \tag{4.2.4}$$

Avalanches that had more than 50% intersection were considered to be moving along a JHU white matter tract.

4.3 Results

4.3.1 Avalanche detection and CAP extraction

Figure 4.2 shows a representative example of the instantaneous WBC calculated for a subject. The sampling period is marked by strongly varying levels of instantaneous correlation rather than a constant or slowly varying value. The γ_H and γ_L cutoffs are shown in yellow and orange, respectively. Thus, the avalanching period is shown by the high and peak WBC values at particular time points.

The left side of Figure 4.3 shows the CAPs identified by group ICA during the avalanching period. The right side of this figure are the CAPs which correspond to the same networks, but are determined during the non-avalanching period. As shown, there are major similarities

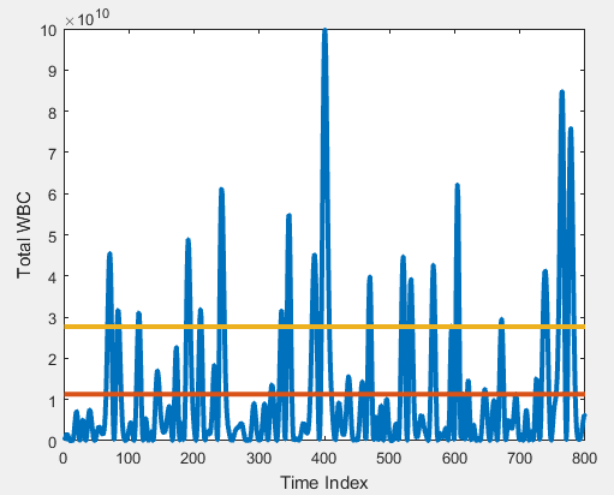


Figure 4.2: Representative subject example of total instantaneous WBC

between the two types of networks; however, during the non-avalanching periods, the active regions of the brain appear to be smaller and more segmented. On the other hand, activation during avalanching confers greater connectivity between different disparate regions in the brain. This is most evident in the appearance of an entire anterior/posterior connectivity region that appears in the brain in CAP 8, whereas the non-avalanching CAP has a limited, local region of activity.

As shown, several commonly known networks and brain structures can be readily identified upon inspection of Figure 4.3. CAP 1 strongly resembles the visual network. CAP 2 contains the corpus callosum, including the body, the genu, and the splenium. This CAP is likely to be an actual functional component rather than a consequence of the difference in the BOLD properties of white matter and gray matter because it does not include all of the white matter and is very specific to the corpus callosum. Other CAPs also include regions of white matter. CAP 7 includes many of the regions of the default mode network (DMN), and CAP 13* includes most of the cerebellum with an additional connection to the midbrain.

Table 4.1 shows the decomposition of each of the CAPs into their constituent resting state networks (RSNs) as defined by Yeo in 2011 [5]. All values reflect the percent of the

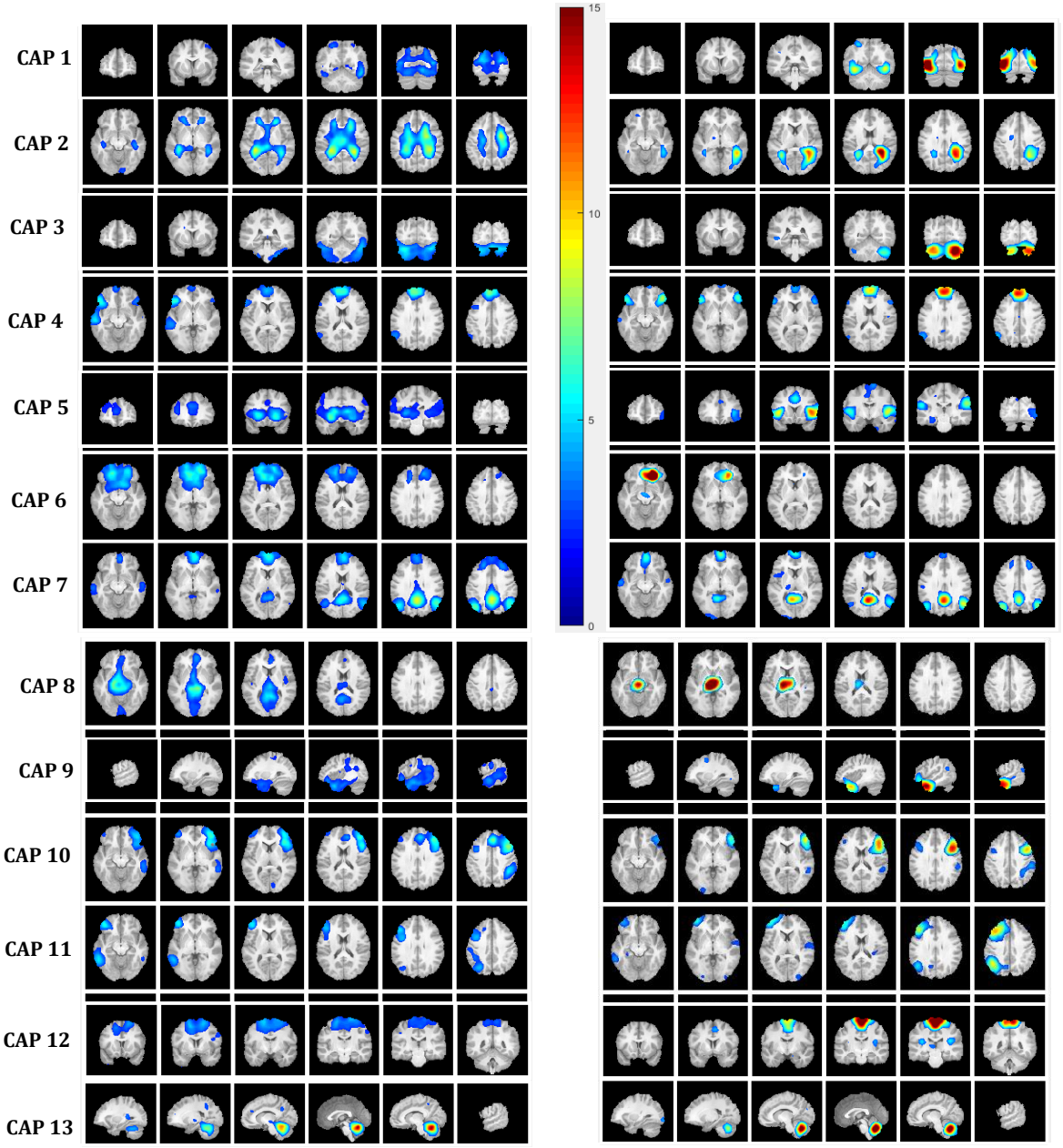


Figure 4.3: Positive activations of 13 CAPs. left: active state, right: inactive state

RSN that intersects with the CAP; values with parentheses indicate the non-avalanching CAPs whereas values without parentheses are the avalanching CAP intersection values. As indicated by the *, CAPs 1, 2, 3, 4, 6 and 8 have significant negative activations (regions not shown in Figure 4.3.

CAP 4* demonstrates a known relation between the DMN and the dorsal attention network (DAN), part of the task-positive network, where the significant activation of the DMN also corresponds to the deactivation of the DAN. More than 10% of the DAN is accounted for by 10 of the 13 CAPs, showing the interconnectedness of this particular network within most of the avalanching CAPs; however, it is only significantly present in 2 of the non-avalanching networks. CAPs 4* and 7 have significant intersection with the DMN, and it is particularly prominent in CAP 7.

The total number of times a network is activated in the avalanching and non-avalanching states is shown at the bottom of Figure 4.1. As shown, during the activation state, there are more networks represented in the CAPs than in the inactive state. A similar metric can be seen for the CAPs on the right of the table, reflecting how many networks are active in each CAP. Together, these metrics suggest that more information is being passed between networks or among networks in the avalanching states than in the non-avalanching states.

Table 4.1: Composition of RSNs by Active CAPSs. All values are percentages of the network of interest; only percentages greater than 10% are shown. CAPs marked by * correspond to CAPs which have significant regions of negative activation in addition to positive activations. Values in red indicate percentages of networks that are negatively activating. Values in parentheses correspond to the non-avalanching regions, whereas those without correspond to the avalanching periods.

CAP# \ RSN#	Visual	Somatomotor	DAN	VAN	Limbic	Frontoparietal	DMN	Active	Inactive
1*	44.4 (41.2)		18.3 (13.9)					2	2
2*	(10.7)							0	1
3*	12.4	11.3	11.2					3	0
4*			22.3		15.9	(10.1)	43.4 (28.0)	3	2
5		23.8 (25.8)		45.9 (55.8)		10.5		3	2
6*				(10.5)	44.6 (27.3)	16.8 (12.4)	17.3	3	2
7			16.7			21.4	52.4 (34.3)	3	1
8*	19.6		12.2		17.1			3	0
9			12.4		31.7 (19.8)		13.5 (12.7)	3	2
10			14.1	19.5 (11.6)		52.0 (22.4)	13.9	4	2
11			14.7 (12.6)			22.1 (28.6)	12.4 (17.0)	3	3
12		30.3 (33.5)	12.7	14.6				3	1
13*	14.0		13.9					2	0
Active	4	3	10	3	4	5	6		
Inactive	2	2	2	3	2	4	4		

4.3.2 Spatial avalanche identification

Figures 4.4, 4.5, and 4.6 show the evolution of an avalanche through 11 0.72s sampling periods. The figures are divided in this manner for display purposes. As shown, the avalanche begins with two small regions in the first frame, and then grows first into 4 larger and distributed areas which grow towards one another to form a large contiguous region. The area of activation appears to be in the cortical region of the brain. The region grows until about halfway through the entire avalanche, where it begins to dissipate. The avalanche appears to terminate in a region somewhat lower than the originating cluster areas.

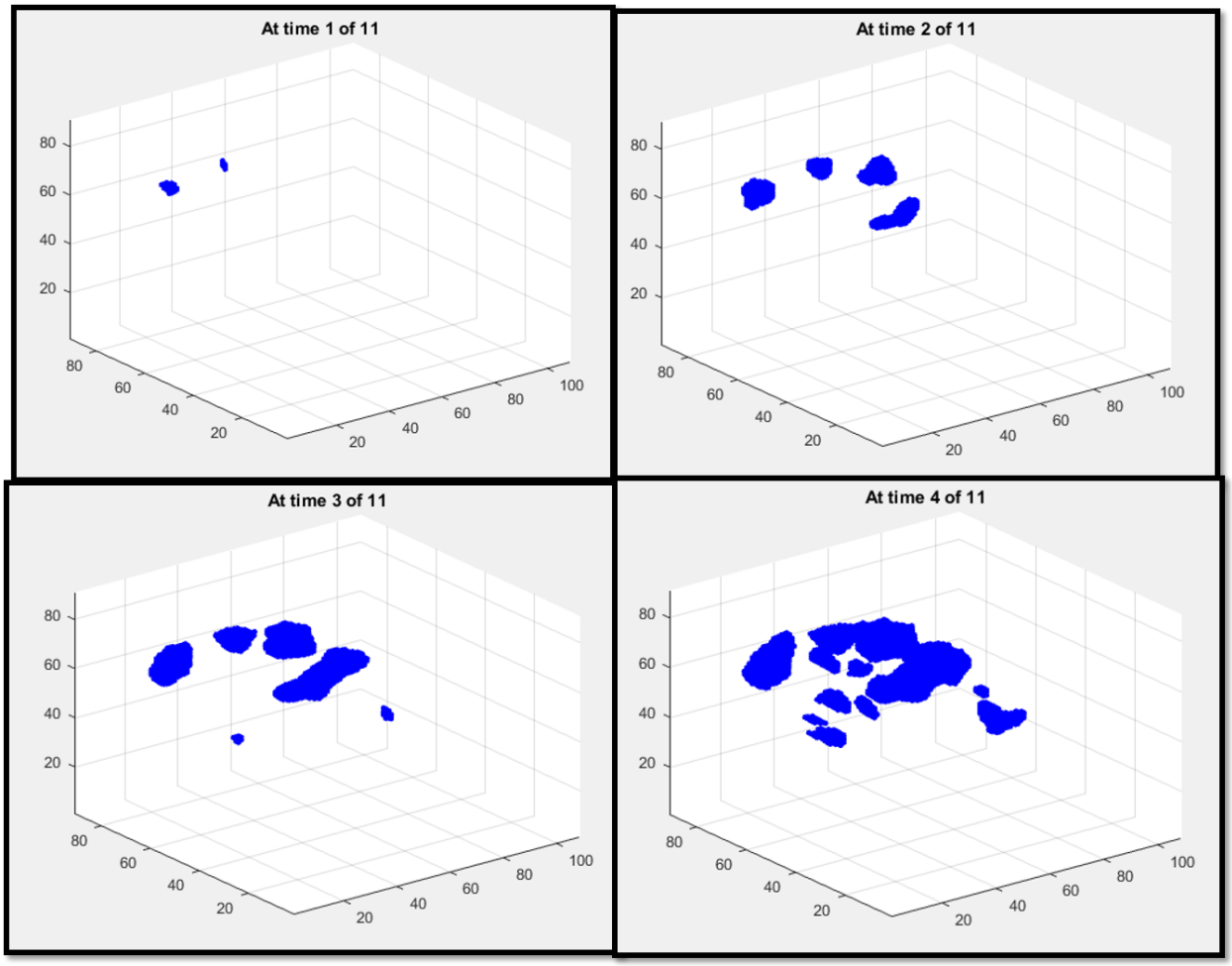


Figure 4.4: Frames 1-4 out of 11 for a sample avalanche

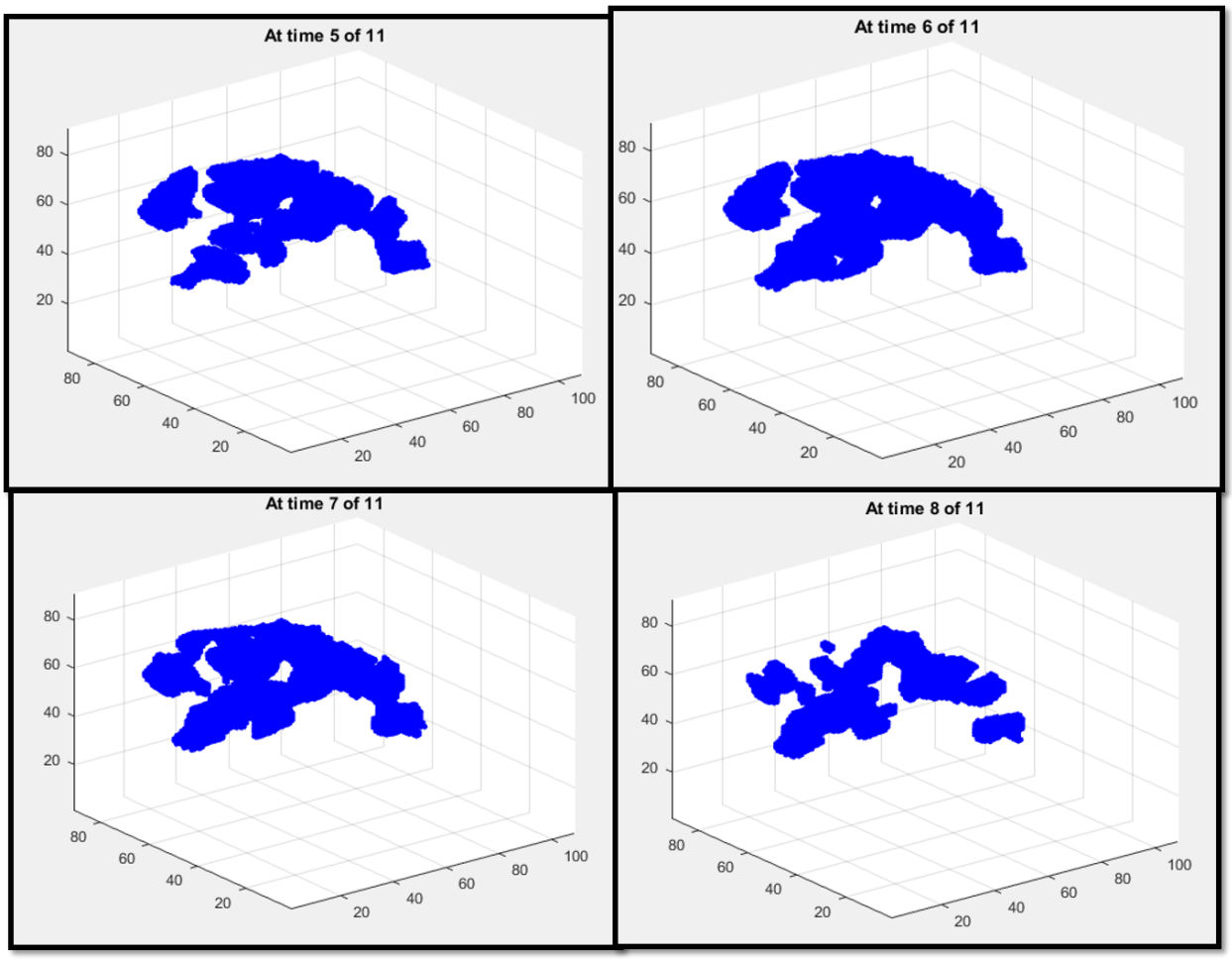


Figure 4.5: Frames 5-8 out of 11 for a sample avalanche

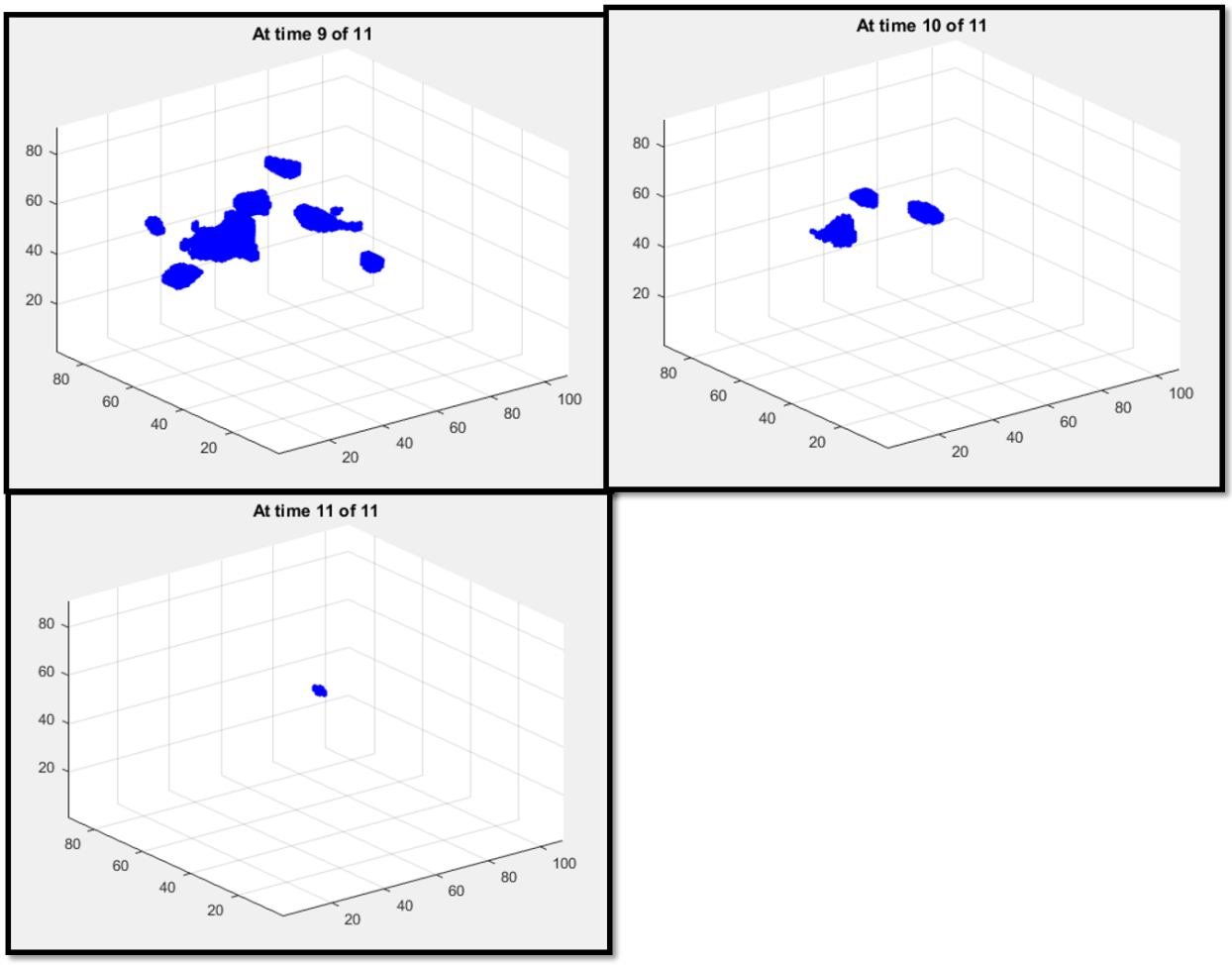


Figure 4.6: Frames 9-11 out of 11 for a sample avalanche

4.3.3 High frequency of occurrence spatial avalanches

Figure 4.7 shows an example of one of the graphs generated by the method described. As shown, the nodes are labeled according to the Harvard-Oxford ROIs which occur at the maximum correlation value in the local cluster. As shown, the information in the graph moves between two regions of the brain - nodes 107 and 108, which correspond to the right cerebral white matter and the right cerebral cortex, respectively. The graph additionally has two edges which are labelled with arrows in order to describe the direction of connection (in our case, information motion).

This particular graph illustrates one of the disadvantages of this particular method: in some cases, the temporal behavior of the information flow is lost. For example, in this example, it is unclear whether the exchange was initiated in the white matter or in the cerebral cortex. Additionally, it is unknown how many times this exchange occurred between these two regions. Regardless, the graph shows that information is passed between these two regions.

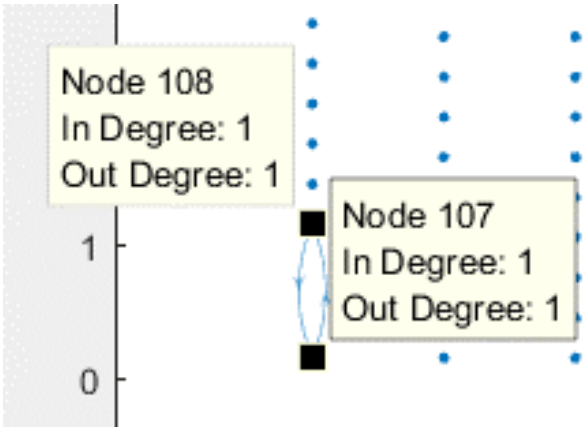


Figure 4.7: Sample avalanche graph. Graph shows the avalanche with the 19th highest frequency.

Tables 4.2 - 4.6 show the top 50 occurring avalanche types in addition to their frequency of occurrence in order of descending frequency. Indices are given for ease of reference.

These tables are divided in this manner for clarity and display purposes. As shown, the most frequently occurring avalanche is information moving within the left cerebral white matter. This accounts for more than 17% of the avalanches occurring in the brain for all the subjects. The second most occurring is information being passed within the right cerebral white matter, which accounts for approximately 16% of all of the avalanches. Together, these two regions of avalanche - white matter - define more than 30% of all the avalanches that occur in the brain. Again, due to the nature of the graphs, it is impossible to determine how many times information was passed through the white matter; it is only known that information is passed from an area of white matter to a different region of white matter at least once.

Information passing within the right and left cerebral cortex accounts for the approximately 18% of all of the avalanches, although there appears to be slightly more information flow in the left part of the cerebral cortex rather than the right. It is interesting to note that this is gray matter, and although most of the processing in the brain is considered to occur in the gray matter, the maxima of the avalanches indicate that during the high activation times in the brain, it is mostly white matter that is activated. Following the cerebral cortex, information passed within the brain stem accounts for 4% of all of the avalanches.

The first 4 avalanche types account for more than 50% of the avalanches. The other 50% is accounted for through many less frequent avalanches. The most frequent type of avalanches following the top 5 is passing information between white matter and gray matter. Avalanche 12, at a frequency of 1.1%, shows information passing within the frontal pole becoming of interest, and although the frequency is decreasing, information begins being passed back and forth from the frontal pole to white matter. Additionally, at a frequency of around 0.84%, the occipital cortex begins to become prevalent, and again, at lower frequencies, information begins being passed between the white matter and the occipital poles and occipital cortex.

At approximately avalanche 20, information circuits begin to become obvious, with information passing through the same nodes multiple times. This can be seen at avalanches 20-22, in which information is passed back and forth for an undetermined number of times within the avalanche. It is interesting to note that in these three cases, information does not appear to be passed between hemispheres, and all communication is done within a single hemisphere. This seems to be the trend with the majority of the top 50 types of avalanches that occur.

At a frequency of around 0.65%, information begins being passed within the temporal pole, and at a frequency of about 0.52%, information begins being passed from the temporal pole to other regions of the brain, particularly the gray matter of the left cerebral cortex. At a slightly lower frequency of 0.39%, this also occurs except the communication is between the right frontal pole and the right cerebral cortex.

Although usually a region in which communication occurs is first shown in the left hemisphere (e.g., the left cerebral white matter and left cerebral cortex have higher frequencies than the right cerebral white matter and cortex), this is violated at avalanche 37 (frequency of 0.32%) in which the communication between the right putamen appears before the left putamen (frequency of 0.19%). This occurs again with the right precuneus cortex (avalanche 39), and its left analogue does not occur within the top 50. This suggests that perhaps a few very distinct regions have stronger operation in the right hemisphere, whereas most of the rest of the general operations are carried on in the left hemisphere of the brain.

At frequencies of 0.26%, information flows in both hemispheres from the cerebral white matter to the occipital cortex, suggesting that perhaps some information is being passed through the white matter whose terminal location is the occipital cortex, and this usage of the white matter as a medium occurs at high correlation activity in the brain. Additionally, at these lower frequencies, information begins to flow in regions including the left temporal fusiform gyrus, right parahippocampal gyrus, left putamen, and left thalamus.

Table 4.2: Type and rate of occurrence for avalanches 1-10 out of 50 high frequency occurring avalanches

Index	**Rate (%)**	**Pathway of Information Motion**
1	17.4	Within left cerebral white matter
2	15.8	Within right cerebral white matter
3	11.2	Within left cerebral cortex
4	7	Within right cerebral cortex
5	4.1	Within brain stem
6	2.3	Left cerebral cortex to left cerebral white matter
7	2.1	Right cerebral white matter to right cerebral cortex
8	1.9	Right cerebral cortex to right cerebral white matter
9	1.8	Left cerebral cortex to within left cerebral white matter
10	1.4	Left cerebral white matter to left cerebral cortex

Table 4.3: Type and rate of occurrence for avalanches 11-20 out of 50 high frequency occurring avalanches

Index	**Rate (%)**	**Pathway of Information Motion**
11	1.1	Left cerebral white matter to within left cerebral cortex
12	1.1	Within right frontal pole
13	0.98	Within left cerebral white matter to left cerebral cortex
14	0.84	Within left lateral superior occipital cortex
15	0.84	Right frontal pole to within right cerebral white matter
16	0.78	Right frontal pole to right cerebral white matter
17	0.78	Right cerebral cortex to within right cerebral white matter
18	0.78	Within right cerebral white matter to right cerebral cortex
19	0.78	Within right cerebral white matter to right frontal pole
20	0.78	Back and forth between right cerebral white matter and right cerebral cortex

Table 4.4: Type and rate of occurrence for avalanches 21-30 out of 50 high frequency occurring avalanches

Index	**Rate (%)**	**Pathway of Information Motion**
21	0.71	Within right cerebral white matter to back and forth with right cerebral cortex
22	0.71	Within left cerebral white matter to back and forth with left cerebral cortex
23	0.65	Within left temporal pole
24	0.65	Back and forth between left cerebral white matter and left cerebral cortex
25	0.58	Right cerebral white matter to right frontal pole
26	0.52	Right cerebral white matter to within right cerebral cortex
27	0.52	Within left cerebral cortex to left cerebral white matter
28	0.52	Within left occipital pole
29	0.52	Left temporal pole to left cerebral cortex
30	0.45	Within right lateral superior occipital cortex

Table 4.5: Type and rate of occurrence for avalanches 31-40 out of 50 high frequency occurring avalanches

Index	**Rate (%)**	**Pathway of Information Motion**
31	0.39	Within right cerebral white matter to within right cerebral cortex
32	0.39	Right cerebral cortex to right frontal pole
33	0.39	Left lateral superior occipital cortex to left cerebral white matter
34	0.32	Left lateral inferior occipital cortex to left cerebral cortex
35	0.32	Left middle temporal gyrus, temporooccipital part to left cerebral cortex
36	0.32	Within right occipital pole
37	0.32	Right putamen to right cerebral white matter
38	0.32	Within right cerebral cortex to right cerebral white matter
39	0.32	Within right precuneus cortex
40	0.26	Within left cerebral white matter to left lateral superior occipital cortex

Table 4.6: Type and rate of occurrence for avalanches 41-50 out of 50 high frequency occurring avalanches

Index	**Rate (%)**	**Pathway of Information Motion**
41	0.26	Right cerebral white matter to right lateral superior occipital cortex
42	0.26	Back and forth between right cerebral white matter and right lateral superior occipital cortex
43	0.26	Within left temporal posterior fusiform cortex
44	0.26	Within right cerebral white matter and back and forth with right frontal pole
45	0.26	Within right cerebral white matter to back and forth with within right cerebral cortex
46	0.26	Within right anterior parahippocampal gyrus
47	0.26	Within left superior frontal gyrus
48	0.19	Right lateral superior occipital cortex to within right cerebral white matter
49	0.19	Left putamen to within left cerebral white matter
50	0.19	Left thalamus to within left cerebral white matter

4.3.4 Avalanche motion about white matter fiber tracts

Figure 4.8 shows the trajectory of the maxima (in red) of an avalanche through a fiber tract (in green). As shown, the traversal is not temporally extensive, and lasts only 3 frames. The motion of the cluster maxima within the tract is laterally and then in a axially downward direction.

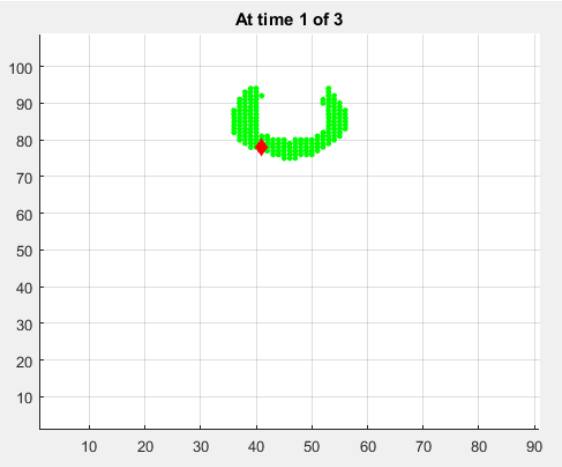
In total, the number of avalanches moving about a fiber tract corresponds to approximately 0.05% of all the avalanches in total. Of this percentage, 51% of these avalanches are moving about the forceps minor. The forceps minor is a fiber bundle that connects surfaces of the frontal lobes through the genu of the corpus callosum. This bundle connects the frontal poles.

The following fiber tracts occurred more than once and their frequency is indicated as follows:

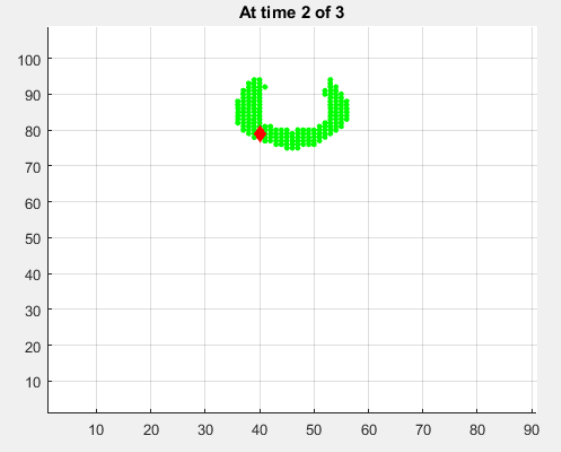
- Left superior longitudinal fasciculus, 22%
- Left anterior thalamic radiation, 10%
- Right anterior thalamic radiation, 7%

The left and right anterior thalamic radiations are located laterally and also posterior to the forceps minor, which suggests that there is information in this entire region that is being passed through the fiber tracts. However, these avalanches which correspond to the fiber tracts appear to be very short, usually only 2 or 3 frames of 0.72 seconds.

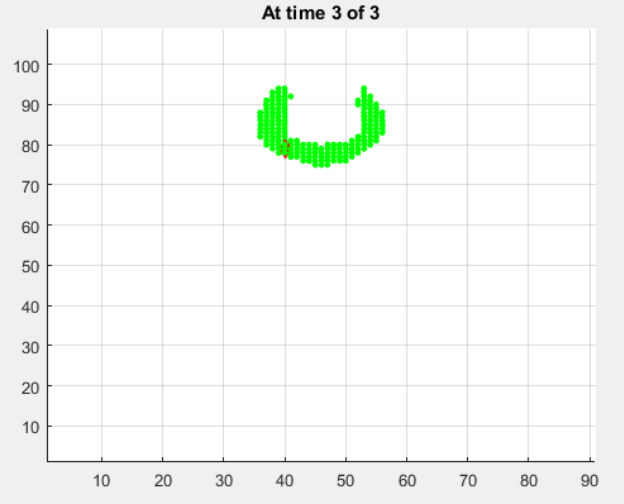
Avalanches can be formulated as both temporal and spatial activities, as shown by the preceding work. The times in which these avalanches occur reflect changes in the decomposition of the brain into independent components, or CAPs, and reflects spatial motion of information moving throughout different maxima in the brain. This work assumes a correspondence between the location of maxima in instantaneous WBC and the motion of information through the brain, and this assumption is bolstered by the movement of the



(a) Frame 1 of 3 of avalanche following forceps minor fiber tract



(b) Frame 2 of 3 of avalanche following forceps minor fiber tract



(c) Frame 3 of 3 of avalanche following forceps minor fiber tract

Figure 4.8: Trajectory of cluster maxima along forceps minor

cluster maxima through space in reasonable trajectories. It is interesting to note that most of the avalanches are related to information flow within white matter and can be postulated that avalanches represent times where information is being sent through the white matter to reach other locations in the brain. Although sometimes the initiating location is lost, the avalanches appear to capture the transfer to and through the white matter and the delivery of the information to its terminal location.

CHAPTER 5

CONCLUSIONS

5.1 Probabilistic fMRI Interpretations

MRI is a non-invasive and increasingly accurate method for studying the structure and function of the brain. Although the structures of the brain are statically and structurally connected, through the use of fMRI which measures the BOLD signal, a reflection of metabolic activity of the brain, the FC of the brain can be investigated. A popular method for investigating FC is through the use of seed-based correlation analysis, although varying the location of the desired seed even slightly in the same region can produce significantly different networks. In this work, a method is presented which produces a stable network with probabilistic interpretations, that is robust to both variance in parameters used to generate the network as well as the choice of the initial seed location. To produce this network, an initial seed is selected, and Pearson's correlation coefficient is calculated between this seed and each voxel's time series in the brain. A seed cloud of highly correlated voxels are then determined, and an SCA executed for each of the randomly selected seeds from the cloud. Depending on whether a probabilistic map or a correlation map is desired, the SCA maps are thresholded, and these are then averaged to create a PCC network map.

This method was specifically applied to the DMN, using the PCC as the initializing seed location. 16 subjects were used in our study, and a group map was generated as the average of the 16 PCC network maps generated from the subjects. The group probabilistic PCC network map showed the highest probability of connections with the pC and the PCC ($p>0.55$), followed by the mPFC, left and right IPL ($p>0.45$), and the PH and the HF ($p>0.35$). This result demonstrates that these regions have a high probability of high correlation with the PCC and that a high percentage of the PCC network is highly correlated to

these regions. The group correlation PCC network map had the highest correlation with the pC and the PCC (r>0.7), and a slightly lower correlation with the mPFC and right and left IPL (r>0.65), a set of regions known as the DMN. Other regions were shown to be slightly less correlated that are not always associated with the DMN. Depending on the hypothesis of brain connectivity employed, this can mean that the overall strength of connectivity between these regions is lower although the regions are still functionally connected, or the temporal FC of the regions varies through time.

This method was then employed to demonstrate that common forms of Bayesian estimation, including MAP estimation, can be performed to produce unique probabilistic interpretations of the data. The detail of the MAP estimation and range of the probabilities was strongly dependent on the *a priori* probability map used; however, all maps confer the same conclusions: the mPFC, the right IPL, the PC, and the PCC are all regions with a high probability of being in the DMN. The different *a priori* probabilities represented different confidences in a binary map of the DMN, and as the *a priori* maps approach the binary probability map, the MAP maps become more detailed in the finer details of regions in the DMN. This was also shown in the DAN network, where the IPS and the FEF were shown to be regions of high probability of being in the DAN.

The generated PCC network was shown to be robust to changes in the number of seeds selected from the seed cloud and variation of the initializing seed location up to ±10 mm. Variation in the number of seeds as compared to the number of seeds that was used in this study produced almost no effect; the Dice's coefficient of similarity was >99% for all of the investigated numbers of seeds (from 1,000 to 10,000). Additionally, varying the location of the initializing PCC seed differed along each axis, often asymmetrically based on moving in a positive direction or a negative one; however, in general, the similarities between the maps were >88%. The only deviant from this high similarity was variance in the anterior direction along the transverse axis; this resulted in a similarity closer to 80% at a 10 mm

variation; this is likely due to approaching other brain regions which may have very different connectivity maps.

The dependence of the method on the initial time course location was also investigated using the difference between the median value of standard deviations of voxels across subjects to assess the difference between the standard deviations of 16 subjects using each of the variations in the method. The SC-SCA method presented was additionally compared against traditional SCA to determine any advantages gained by using SC-SCA. The results show that qualitatively, SC-SCA generates correlation maps that are more detailed than SCA, and have less artifacts near the seed point. When compared against region-based SC-SCA and seed- and region-based SCA, there is a statistically significant advantage in using seed-based SC-SCA vs seed-based SCA, region-based SC-SCA vs seed-based SCA, and region-based SCA vs. seed-based SCA. There is no statistical difference in using either of the region-based methods, seed-based SC-SCA vs region-based SCA, and seed-based SC-SCA vs region-based SC-SCA.

Most notably, this shows that the SC-SCA stability in the standard deviation of voxels among subjects is not improved by using a region-based seeding time point; on the other hand, there is an improvement of the stability of the maps produced by region-based SCA rather than seed-based SCA. In the light of current literature, this suggests that region-based SCA is superior to that of seed-based SCA due to its resistance to change when different close voxels are chosen as the seed point, but this characteristic is not present in SC-SCA. Seed-based SC-SCA and region-based SC-SCA work equally as well as each other, and both better than seed-based SCA. They are on par with region-based SCA, but additionally have the qualitative advantage that probabilistic interpretations are conferred by SC-SCA.

Additionally, the decomposition of the thresholded SCA and SC-SCA networks was performed using ICA. Although 8 out of the 10 generated ICs were similar, the SC-SCA ICs were much more diffuse and showed higher connectivity between disparate regions in the

brain. This reflects potential pathways through which information flows in the brain. Several networks of similarity were found in the matching ICs, including the posterior region of the DMN, the mPFC, and the right and left parts of the DAN. There were 2 dissimilar ICs found; in SCA, these were two different regions of the cingulate gyrus, and in SC-SCA, these were the somatosensory network and the occipital lobe.

The method proposed in this work offers the potential of engineering fully automated definition of the DMN network or other resting state networks. Particularly in light of increasingly large fMRI database that is available to the community [115] complete automation as exemplified in this work will greatly accelerate the productivity of brain research. Furthermore, this work represents an introduction to a method that could have far-reaching applications due to its stability and ability to show a gradated correlation measure of how sets of regions and networks may be connected to a single seed. The introduction of this more stable method of performing seed-based correlation analysis can have very far-reaching applications in understanding the FC of the brain.

5.2 Avalanche detection and CAP extraction

Recent studies have shown that assumptions of stationarity in brain functionality may be too strong, and that brain functionality is dominated by periods of high activity in the brain. This work addresses this non-stationarity by investigating the behavior of the brain during *avalanches*, which are periods of unification or similar temporal behavior in the whole brain.

To investigate brain behavior during avalanches, a WBC was performed for each of 16 subjects at each of 800 time points at which the BOLD fMRI data was sampled. A threshold was introduced to distinguish periods of avalanching versus period of non-avalanching behavior. All the periods of avalanching behavior for all 16 subjects were concatenated to create an entire time series of avalanching, and the same was done for the non-avalanching

regions. Group ICA was performed on both time series separately, and the DICE coefficient was calculated between the independent components generated from the avalanching time series and the non-avalanching time series to determine regional correspondence between the two types of components.

CAPs generated from the avalanching region included some readily identifiable networks including the visual system, the DMN, and a connection between the cerebellum and the midbrain. The anatomical structure of the corpus callosum was also distinguished in the avalanching CAPs. In comparison to the non-avalanching CAPs, the avalanching CAPs were more diffuse, and showed more connection between brain regions. Additionally, when the CAPs were segmented into their RSNs, there were more RSNs present in the avalanching time periods than the non-avalanching time periods. This work suggests that brain activity is dominated by avalanches of activity where there is more communication between RSNs and disparate regions of the brain during the avalanching period as compared to the non-avalanching period.

The approach presented in this work represents a novel perspective in viewing activations in the brain. The approach uses a whole brain correlation produced at each time point, which describes how each voxel in the brain is correlated to each other voxel in the brain. When summed, this gives an description of the overall levels of correlation in the brain. This activity fluctuates throughout the duration of the sampling, reaching large peaks at some points with periods of low activity as well. This approach represents the first time this has been reported, and additionally, the spatial motions of the activated clusters in the avalanches throughout the brain have been evaluated for their frequency.

5.3 Future Work

There are many directions in which the two thrusts of this work may be extended. For the seed cloud SCA method, the network generated appears to be larger than just the DMN,

particularly when the region-based method is used. Performing principal or independent components analysis could show how the network is able to be subdivided into its subcomponent networks. Additionally, this method can be easily extended to become a Bayesian analysis to potentially further understand the causality in activation of these brain regions. The method may additionally be employed to investigate other networks within the brain, including the language and control networks.

The avalanche approach is particularly interesting, in that many different analyses may be performed once the instantaneous WBC is generated. The spatial interpretation of the avalanche demonstrates this. One of the hindrances of spatial method in particular was the necessity to label each of the maxima of the clusters using an atlas. If an expert in neuroscience and brain anatomy was brought in, many of the data points with no label could be labelled, expanding the clusters that are represented in the brain. Additionally, an expert could view the motion of the clusters within the avalanches to identify if there are any known tracts or movement of information which coincide with the avalanches.

In addition, studies of the internal movement of the most popular motions (i.e., motions within white matter) can also help to interpret how data moves along the white matter. If the thresholds for significance are lowered, perhaps even the gray matter initiating the signal transfer can also be captured and a full thread of information transmission can be assessed. Lastly, instead of generating the thresholds at the global level (i.e., using the entire time series for each voxel) to calculate the z-scores for thresholding, the mean and standard deviation calculations may instead be made at the volume level, so that the most significant voxels at each time point may be identified. This may assist in identifying the flow of information.

Finally, the most interesting future directions of this work can be in the assessment of brain disease pathologies and task-related experiments. It would be interesting to assess how the generated DMN using SC-SCA varies throughout a particular task, and how it varies

among subjects with a particular disorder of the brain. It would also be intriguing to assess if there is any difference in the motion of clusters in avalanches between control patients and those with a disorder. In this way, the potential causes and effects of disease pathologies on the brain could be investigated in a new way.

In conclusion, in-depth understanding of the structure and functionality of the brain is essential, from both a scientific and clinical standpoint. Brain disorders and diseases have a profound impact socioeconomically and in the personal lives of those affected and their families. Crucial to the treatment of this population, particularly those with mental illness, is diagnostic imaging of the brain which can help identify the type of illness as well as the best way to treat the illness. One of the prerequisites of this type of diagnosis is the ability to understand the normal way the brain functions versus the altered mechanism due to the illness. The purpose of this work is to that end - to provide tools that produce an unique view into the complex interplay of signals sent between regions of the brain. This work presented two innovative new tools that allow novel probabilistic interpretations of the brain and investigate behavior during the critically times of high brain voxel signal coherence. These tools will aid in the development of science relative to the function of the brain to improve the diagnostic outlook of those with brain disorders in the future.

REFERENCES

[1] J. Ainali. Philips MRI in Sahlgrenska Universitetsjukhuset, Gothenburg, Sweden. [Online]. Available: https://commons.wikimedia.org/wiki/File:MRI-Philips.JPG

[2] A. Filler. Spin-echo diagram. [Online]. Available: https://en.wikipedia.org/wiki/File:Spin_Echo_Diagram.jpg

[3] IngFrancesco. Transversing k-space with EPI sequence. [Online]. Available: https://it.wikipedia.org/wiki/Echo-planar_imaging#/media/File:MRI_k-space_epi.svg

[4] SPM-Wiki. SPM hemodynamic response function. [Online]. Available: https://commons.wikimedia.org/wiki/File:SPM_hemodynamic_response_function.png

[5] B. T. Yeo, F. Krienen, J. Sepulcre, M. Sabuncu, D. Lashkari, M. Hollinshead, J. Roffman, J. Smoller, L. Zöllei, J. Polimeni, B. Fischl, H. Liu, and R. Buckner, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of Neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011.

[6] N. Alliance on Mental Illness (NAMI). Mental Health Facts in America. [Online]. Available: https://www.nami.org/NAMI/media/NAMI-Media/Infographics/GeneralMHFacts.pdf

[7] F. Adriano, C. Caltagirone, and G. Spalletta, "Hippocampal volume reduction in first-episode and chronic schizophrenia: A review and meta-analysis," *Progress in Neuroscience*, vol. 18, no. 2, pp. 180–200, 2012.

[8] D. Velakoulis, M. Wong, A. Yung, and et. al., "Hippocampal and amygdala volumes according to psychosis stage and diagnosis," *Archives of General Psychiatry*, vol. 63, pp. 139–149, 2006.

[9] P. Yushkevich, R. Amaral, J. Augustinack, and et. al, "Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in

vivo mri: Towards a harmonized segmentation protocol," *Neuroimage*, vol. 111, pp. 526–541, 2015.

[10] S. Heckers and C. Konradi, *Hippocampal Pathology in Schizophrenia*. Springer Berlin Heidelberg, 2010, pp. 529–553.

[11] N. Woodward and S. Heckers, "Brain structure in neurophysiolocially defined subgroups of schizophrenia and psychotic bipolar disorder," *Schizophrenia Bulletin*, vol. 41, no. 6, pp. 1349–1359, 2015.

[12] N. Woodward, B. Rogers, and S. Heckers, "Functional resting state networks are differentially affected in schizophrenia," *Schizophrenia Research*, vol. 130, pp. 86–93, 2011.

[13] K. Friston, "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, pp. 56–78, 1994.

[14] S. Kiebel and A. Holmes, "The general linear model," *Human Brain Function*, vol. 2, pp. 725–760, 2003.

[15] J. Ashburner and K. Friston, "Voxel-based morphometry - the methods," *NeuroImage*, vol. 11, pp. 805–821, 2000.

[16] A. Mechelli, C. Price, K. Friston, and et. al., "Voxel–based morphometry of the human brain: Methods and applications," *Current Medical Imaging Reviews*, vol. 1, no. 1, pp. 1–9, 2005.

[17] N. Woodward, B. Rogers, and S. Heckers, "Functional resting-state networks are differentially affected in schizophrenia," *Schizophrenia Research*, vol. 130, pp. 86–93, 2011.

[18] N. Woodward and S. Heckers, "Mapping thalamocortical functional connectiity in chronic and early stages of psychotic disorders," *Biological Psychiatry*, vol. 79, no. 12, pp. 1016–1025, 2016.

[19] S. Costafreda, C. Fu, M. Picchioni, and et. al., "Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder," *BMC*

Psychiatry, vol. 11, no. 18, 2011.

[20] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, 2nd ed. Sinauer Associates, Inc.

[21] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent blood on oxygenation," *Proceedings of the National Academy of Science*, vol. 87, no. 24, pp. 9868–9872, December 1990.

[22] D. Malonek and A. Grinvald, "Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: Implications for functional brain mapping," *Science*, vol. 272, no. 5261, pp. 551–554.

[23] P. Jezzard, P. M. Matthews, and S. M. Smith, *Functional MRI an introduction to methods*.

[24] R. Buxton, E. Wong, and L. Frank, "Dynamics of blood flow and oxygen metabolism during brain activation: balloon model," *Magnetic Resonance in Medicine*, vol. 39, pp. 855–864, 1998.

[25] H. Lu, X. Golay, J. Pekar, and P. V. Zijl, "Sustained poststimulus elevation in cerebral oxygen utilization after vascular recovery," *Journal of Cerebral Blood Flow Metabolism*, vol. 24, pp. 764–770, 2004.

[26] G. Glover, "Overview of functional magnetic resonance imaging," *Neurosurgery Clinics of North America*, vol. 22, no. 2, pp. 133–139, 2011.

[27] M. Lowe, "The emergence of doing "nothing" as a viable paradigm design," *NeuroImage*, vol. 62, no. 2, pp. 1146–1151, 2012.

[28] M. van den Huevel and H. Pol, "Exploring the brain network: A review on resting-state fmri functional connectivity," *European Neuropsychopharmacology*, vol. 20, pp. 519–534, 2010.

[29] B. Biswal, F. Yetkin, V. Haughton, and J. Hyde, "Functional connectivity in the

motor cortex of resting human brain using echo-planar mri," *Magnetic Resonance in Medicine*, vol. 34, pp. 537–541, 1995.

[30] D. Cordes, V. Haughton, K. Arfanakis, J. Carew, P. Turski, C. Horitz, M. Quigley, and M. Meyerand, "Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data," *American Journal of Neuroradiology*, vol. 22, no. 7, pp. 1326–1333, 2001.

[31] M. Razavi, B. Eaton, S. Paradiso, M. Mina, A. Hudetz, and L. Bolinger, "Source of low-frequency fluctuations in functional mri signal," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 891–897, 2008.

[32] C. Beckmann, M. D. Luca, J. Devlin, and S. Smith, "Investigations into resting-state connectivity using independent components analysis," *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, vol. 360, no. 1457, pp. 1001–1003, 2005.

[33] M. D. Luca, C. Beckmann, N. D. Stefano, P. Matthews, and S. Smith, "fMRI resting state networks define distinct modes of long-distance interactions in the human brain," *Neuroimage*, vol. 29, no. 4, pp. 1359–1367, 2006.

[34] J. Damoiseaux, S. Rombouts, F. Barkhof, P. Scheltens, C. Stam, S. Smith, and C. Beckmann, "Consistent resting-state networks across healthy subjects," *Proceedings of the National Academy of the Sciences USA*, vol. 103, no. 37, pp. 13 848–13 853, 2006.

[35] R. Salvador, J. Suckling, M. Coleman, J. Pickard, D. Menon, and E. Bullmore, "Neurophysiological architecture of functional magnetic resonance images of human brain," *Cerebral Cortex*, vol. 15, no. 9, pp. 1332–1342, 2005.

[36] M. V. den Huevel, R. Mandl, and H. H. Pol, "Normalized cut group clustering of resting-state fMRI data," *PLoS ONE*, vol. 3, no. 4, 2008.

[37] D. Cole, S. Smith, and C. Beckmann, "Advances and pitfalls in the analysis and

interpretation of resting-state fmri data," *Frontiers in Systems Neuroscience*, vol. 4, no. 8, pp. 1–15, 2010.

[38] W. Seeley, V. Menon, A. Schatzberg, J. Keller, G. Glover, H. Kenna, A. Reiss, and M. Greicius, "Dissociable intrinsic connectivity networks for salience processing and executive control," *Journal of Neuroscience*, vol. 27, pp. 2349–2356, 2007.

[39] S. Smith, P. Fox, K. Miller, D. Glahn, P. Fox, C. Mackay, N. Filippini, K. Watkins, R. Toro, A. Laird, and C. Beckmann, "Correspondence of the brain's functional architecture during activation and rest," *Proceedings of the National Academy of the Sciences USA*, vol. 106, pp. 13 040–13 045, 2009.

[40] P. Vemuri, D. Jones, and J. CR Jack, "Resting state functional MRI in alzheimer's disease," *Alzheimer's Research and Therapy*, vol. 4, no. 1, p. 2, 2012.

[41] L. Sang, J. Zhang, L. Wang, J. Zhang, Y. Zhang, P. Li, J. Wang, and M. Qui, "Alteration of brain functional networks in early-stage parkinson's disease: A resting-state fMRI study," *PLoS ONE*, vol. 10, no. 10.

[42] W. Liu, J. Yang, K. Chen, C. Luo, M. Burgunder, Q. Gong, and H. Shang, "Resting-state fmri reveals potential neural correlates of impaired cognition in huntington's disease," *Parkinsonism and Related Disorders*, vol. 27, pp. 41–46, 2016.

[43] A. Demertzi, F. Gòmez, J. Crone, A. Vanhaudenhuyse, L. Tshibanda, Q. Noirhomme, M. Thonnard, V. Charland-Verville, M. Kirsch, S. Laureys, and A. Soddu, "Multiple fMRI system-level abseline connectivity is disrupted in patients with consciousness alterations," *Cortex*, vol. 52, pp. 35–46, 2014.

[44] M. Greicius, G. Srivastava, A. Reiss, and V. Menon, "Default-mode network activity distinguishes alzheimer's disease from healthy aging: Evidence from functional MRI."

[45] B. B. Biswal, J. V. Kylen, and J. S. Hyde, "Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps," *NMR in Biomedicine*, vol. 10,

pp. 165–170, 1997.

[46] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai, "Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis," *Magnetic Resonance Imaging*, vol. 18, no. 1, pp. 89–94, January 2000.

[47] A. Baune, F. T. Sommer, M. Erb, D. Wildgruber, B. Kardatzki, G. Palm, and W. Grodd, "Dynamical cluster analysis of cortical fMRI activation," *NeuroImage*, vol. 9, no. 5, pp. 477–489, May.

[48] C. Goutte, P. Toft, E. Rostrup, F. Å. Nielsen, and L. K. Hansen, "On clustering fMRI time series," *NeuroImage*, vol. 9, no. 3, pp. 298–310, March 1999.

[49] P. Filzmoser, R. Baumgartner, and E. Moser, "A hierarchical clustering method for analyzing functional MR images," *Magnetic Resonance Imaging*, vol. 17, no. 6, pp. 817–826, July 1999.

[50] E. Allen, E. Damaraju, S. Plis, E. Erhardt, T. Eichele, and V. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral cortex*, vol. 24, pp. 663–676, 2014.

[51] D. S. Margulies, J. Böttger, X. Long, Y. Lv, C. Kelly, A. Schäfer, D. Goldhahn, A. Abbushi, M. P. Milham, G. Lohmann, and A. Villringer, "Resting developments: a review of fmri post-processing methodologies for spontaneous brain activity," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, no. 5, pp. 289–307, 2010.

[52] M. Lee, C. Smyser, and J. Shimony, "Resting-state fmri: A review of methods and clinical applications," *American Journal of Neuroradiology*, vol. 34, pp. 1866–1872, 2013.

[53] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of

structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, pp. 186–198, 2009.

[54] E. Tagliazucchi, M. Siniatchkin, H. Laufs, and D. R. Chialvo, "The voxel-wise functional connectome can be efficiently derived from co-activations in a sparse spatio-temporal point process," *Frontiers in Neuroscience*, vol. 10, no. 381, 2016.

[55] C. Chang and G. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *Neuroimage*, vol. 50, no. 1, pp. 81–98, March 2010.

[56] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, 2014.

[57] R. M. Hutchison, T. Womelsdorf, J. S. Gati, S. Everling, and R. S. Menon, "Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques," *Human Brain Mapping*, vol. 34, no. 9, pp. 2154–2177, September 2013.

[58] Z. Shi, B. Rogers, L. Chen, V. Morgan, A. Mishra, D. Wilkes, and J. Gore, "Realistic models of apparent dynamic changes in resting-state connectivity in somatosensory cortex," *Human Brain Mapping*, vol. 37, pp. 3897–3910, 2016.

[59] V. D. Calhoun and T. Adali, "Time-varying brain connectivity in fMRI data: whole-brain data-driven approaches for capturing and characterizing dynamic states," *IEEE Signal Processing Magazine*, pp. 52–66, May 2016.

[60] Z. Rajna, J. Kananen, A. Keskinarkaus, T. Seppänen, and V. Kiviniemi, "Detection of short-term activity avalanches in human brain default mode network with ultrafast mr encephalography," *Frontiers in Human Neuroscience*, vol. 9, pp. 1–12, August 2015.

[61] M. G. Preti, T. A. Bolton, and D. V. D. Ville, "They dynamic functional connectome: State-of-the-art and perspectives," *Neuroimage*, 2016.

[62] R. Betzel, L. Byrge, Y. He, J. Goñi, X. Zuo, and O. Sporns, "Changes in structural and functional connectivity among resting-state networks across the human lifespan," *Neuroimage*, vol. 102, pt. 2, pp. 345–357, November 2014.

[63] Z. Shehzad, A. M. C. Kelly, P. T. Reiss, D. G. Gee, K. Gotimer, L. Q. Uddin, S. H. Lee, D. S. Margulies, A. K. Roy, B. B. Biswal, E. Petkova, F. X. Castellanos, and M. P. Milham, "The resting brain: Unconstrained yet reliable," *Cerebral Cortex*, vol. 19, no. 10, pp. 2209–2229, Oct 2009.

[64] N. Leonardi and D. V. D. Ville, "On spurious and real fluctuations of dynamic functional connectivity during rest," *NeuroImage*, vol. 104, pp. 430–436, 2015.

[65] R. Poldrack, "Region of interest analysis for fmri," *Social Cognitive and Affective Neuroscience*, vol. 2, no. 1, pp. 67–70, 2007.

[66] W. Sohn, K. Yoo, Y.-B. Lee, S. Seo, D. Na, and Y. Jeong, "Influence of ROI selection on resting state functional connectivity: an individualized approach for resting state fmri analysis," *Frontiers in Neuroscience*, vol. 9, no. 280, 2015.

[67] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9673–9678, 2005.

[68] P. Fransson, "Spontaneous low frequency bold signal fluctuations: an fmri investigation of the resting-state default model of the brain function hypothesis," *Human Brain Mapping*, no. 26, pp. 15–29, 2005.

[69] C. Keller, S. Nickel, J. Honey, D. Groppe, L. Entz, R. Craddock, F. Lado, C. Kelly, M. Milham, and A. Mehta, "Neurophysiological investigation of spontaneous correlated and anticorrelated fluctuations of the bold signal," *The Joural of Neuroscience*, vol. 33, no. 15, pp. 6333 – 6342, April 2013.

[70] A. Iraji, V. D. Calhoun, N. M. Wiseman, E. Davoodi-Bojd, M. R. Avanaki, E. M. Haacke, and Z. Kou, "The connectivity domain: Analyzing resting state fmri data using feature-based data-driven and model-based methods," *NeuroImage*, vol. 134, pp. 494 – 507, 2016.

[71] H. Liu, R. Buckner, T. Talukdar, N. Tanaka, J. Madsen, and S. Stufflebeam, "Task-free presurgical mapping using functional magnetic resonance imaging intrinsic activity: Laboratory investigation," *Journal of Neurosurgery*, vol. 111, no. 4, pp. 746–754, 2009.

[72] N. Kucukboyaci, N. Kemmotsu, C. Cheng, M. Girard, E. Tecoma, V. Iragui, and C. McDonald, "Functional connectivity of the hippocampus in temporal lobe epilepsy: Feasibility of a task-regressed seed-based approach," *Brain Connectivity*, vol. 3, no. 5, pp. 464–474, 2013.

[73] R. Bharath, S. Sinha, R. Panda, K. Raghavendra, L. George, G. Chaitanya, A. Gupta, and P. Satishchandra, "Seizure frequency can alter brain connectivity: Evidence from resting-state fmri," *American Journal of Neuroradiology*, vol. 36, pp. 1890–1898, 2015.

[74] J. Andrews-Hanna, A. Snyder, J. Vincent, C. Lustig, D. Head, M. Raichle, and R. Buckner, "Disruption of large-scale brain systems in advanced aging," *Neuron*, vol. 56, pp. 924–935, 2007.

[75] M. Brier, J. Thomas, A. Snyder, T. Benzinger, D. Zhang, M. Raichle, D. Holtzman, J. Morris, and B. Ances, "Loss of intra- and inter-network resting state functional connections with alzheimer's disease progression," *Journal of Neuroscience*, vol. 32, no. 26, pp. 8890–8899, 2012.

[76] S. Noonan, F. Haist, and R. Muller, "Aberrant functional connectivity in autism: Evidence from low-frequency BOLD signal fluctuations," *Brain Research*, vol. 1262, pp. 48–63, 2009.

[77] A. D. Martino, C.-G. Yan, Q. Li, E. Denio, F. Castellanos, K. Alaerts, J. Ander-

son, M. Assaf, S. Bookheimer, M. Dapretto, B. Deen, S. Delmonte, I. Dinstein, B. Ertl-Wagner, D. Fair, L. Gallagher, D. Kennedy, C. Keown, C. Keysers, J. Lainhart, C. Lord, B. Luna, V. Menon, N. Minshew, C. Monk, S. Mueller, R.-A. Muller, M. Nebel, J. Nigg, K. O'Hearn, K. Pelphrey, S. Peltier, J. Rudie, S. Sunaert, M. Thioux, J. Tyszka, L. Uddin, J. Verhoeven, N. Wenderoth, J. Higgins, S. Mostofsky, and M. Hilham, "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," *Molecular Psychiatry*, vol. 19, pp. 659–667, 2014.

[78] Y. Sheline, J. Price, Z. Yan, and M. Mintun, "Resting-state functional mri in depression unmasks increased connectivity between networks via the dorsal nexus," *Proceedings of the National Academy of Science*, no. 24, pp. 11 020–11 025, 2010.

[79] X. Chai, S. Whitfield-Gabrieli, A. Shinn, J. Gabrieli, A. Castanon, J. McCarthy, B. Cohen, and D. Ongur, "Abnormal medial prefrontal cortex resting-state connectivity in bipolar disorder and schizophrenia," *Neuropsychopharmacology*, vol. 36, pp. 2009–2017, 2011.

[80] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. Chialvo, "Brain resting state is disrupted in chronic back pain patients."

[81] M. J. McKeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fmri data by blind source separation into independent spatial components," *Human Brain Mapping*, vol. 6, pp. 160–188, 1998.

[82] V. Calhoun, T. Adali, G. Pearlson, and J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Human Brain Mapping*, vol. 13, pp. 43–53, 2001.

[83] C. Beckmann and S. Smith, "Probabilistic independent components analysis for functional magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 23,

pp. 137–152, February 2004.

[84] S. Storti, E. Formaggio, R. Nordio, P. Manganotti, A. Fiaschi, A. Bertoldo, and G. Toffolo, "Automatic selection of resting-state networks with functional magnetic resonance imaging," *Frontiers in Neuroscience*, vol. 7, no. 72, 2013.

[85] M. DeLuca, C. Beckmann, T. Behrens, S. Clare, N. DeStefano, P. Matthews, M. Woolrich, and S. Smith, "Low frequency signals in FMRI - "resting state networks" and the "intensity normalisation problem"," in *Proceedings of the International Society of Magnetic Resonance Imaging in Medicine*, vol. 10, 2002.

[86] L. Zhang, M. Guindani, and M. Vannucci, "Bayesian models for functional magnetic resonance imaging data analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 7, no. 1, pp. 21–41, 2015.

[87] X. Zheng and J. C. Rajapakse, "Learning functional structure from fMR images," *NeuroImage*, vol. 31, no. 4, pp. 1601 – 1613, 2006.

[88] J. C. Rajapakse and J. Zhou, "Learning effective brain connectivity with dynamic bayesian networks," *NeuroImage*, vol. 37, no. 3, pp. 749 – 760, 2007.

[89] R. Patel, F. Bowman, and J. Rilling, "A bayesian approach to determining connectivity of the human brain," *Human Brain Mapping*, vol. 27, no. 3, pp. 267–276, 2006.

[90] X. Wu, R. Li, A. S. Fleisher, E. M. Reiman, X. Guan, Y. Zhang, K. Chen, and L. Yao, "Altered default mode network connectivity in alzheimer's disease: a resting functional MRI and bayesian network study," *Human Brain Mapping*, vol. 32, no. 11, pp. 1868–1881, 2011.

[91] J. Li, Z. J. Wang, S. J. Palmer, and M. J. McKeown, "Dynamic bayesian network modeling of fMRI: A comparison of group-analysis methods," *Neuroomage*, vol. 41, no. 2, pp. 398 – 407, 2008.

[92] X. Wu, X. Yu, L. Yao, and R. Li, "Bayesian network analysis revealed the connectivity

difference of the default mode network from the resting-state to task-state," *Frontiers in Computational Neuroscience*, vol. 8, p. 118, 2014.

[93] H. Stark and J. W. Woods, *Probability, Statistics, and Random Processes for Engineers*. Publishing House of Electronics Industry.

[94] M. R. Chernick and R. A. LaBudde, *An Introduction to Bootstrap Methods with Applications to R*. John Wiley and Sons, Inc.

[95] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping*, vol. 17, no. 3, pp. 143–155, November 2002.

[96] M. Jenkinson, M. Pechaud, and S. Smith, "BET2: MR-based estimation of brain, skull and scalp surfaces," in *In Eleventh Annual Meeting of the Organization for Human Brain Mapping (2005)*.

[97] S. M. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, H. Johansen-Berg, P. Bannister, M. D. Luca, I. Drobnjak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. D. Stefano, J. Brady, and P. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," *Neuroimage*, vol. 23, no. S1, pp. 208–219, 2004.

[98] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, and S. Jbabdi, "The minimal preprocessing pipelines for the human connectome project," *Neuroimage*, vol. 80, pp. 105–124, October 2013.

[99] X. Liu and J. Duyn, "Time-varying functional network information extracted from brief instances of spontaneous brain activity," *Proceedings of the National Academy of Sciences USA*, vol. 110, no. 11, pp. 4392–4397, 2013.

[100] K. R. V. Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner, "Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization," *Journal of Neurophysiology*, vol. 103, no. 1, pp. 297–321,

2010.

[101] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[102] J. L. Lancaster and M. J. Martinez, "Multi-image analysis graphical user interface (mango) [computer software]." http://ric.uthscsa.edu/mango/index.html.

[103] M. Grecius, B. Flores, V. Menon, G. Glover, H. Solvason, H. Kenna, A. Reiss, and A. Schatzberg, "Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus," *Biological Psychiatry*, 2007.

[104] M. Grecius, K. Supekar, V. Menon, and R. Dougherty, "Resting-state functional connectivity reflects structural connectivity in the default mode network," *Cerebral Cortex*, January 2009.

[105] D. Fair, A. Cohen, N. Dosenbach, J. Church, F. Miezin, D. Barch, M. Raichle, S. Petersen, and B. Schlaggar, "The maturing architecture of the brain's default mode network," *Proceedings of the National Academy of the Sciences USA*, vol. 105, no. 10, pp. 4028–4032, March 2008.

[106] D. Tomasi and N. Volkow, "Aging and functional brain networks," *Molecular Psychiatry*, vol. 17, pp. 549–558, 2012.

[107] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. R. Chialvo, "Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis," *Frontiers in Physiology*, vol. 3, no. 15, pp. 1–12, February 2012.

[108] J. Beggs and D. Plenz, "Neuronal avalanches in neocortical circuits," *Journal of Neuroscience*, vol. 23, no. 35, pp. 11 167–11 177, 2003.

[109] D. Chialvo, "Emergent complex neural dynamics," *Nature Physics*, vol. 6, pp. 744–750, 2010.

[110] L. de Arcangelis and H. Herrmann, "Activity-dependent neuronal model on complex networks," *Frontiers in physiology*, vol. 3, no. 62, 2012.

[111] X. Liu and J. H. Duyn, "Time-varying functional network information extracted from brief instances of spontaneous brain activity," *Proceedings of the National Academy of Science (PNAS)*, vol. 110, no. 11, pp. 4392–4397, march 2013.

[112] X. Liu, C. Chang, and J. H. Duyn, "Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns," *Frontiers in Systems Neuroscience*, vol. 7, no. 101, pp. 1–11, December 2013.

[113] R. Desikan, F. Ségonne, B. Fischl, B. Quinn, B. Dickerson, D. Blacker, R. Buckner, A. Dale, R. Maguire, B. Hyman, M. Albert, and R. Killiany, "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[114] S. Wakana, A. Caprihan, M. Panzenboeck, J. Fallon, M. Perry, R. Gollub, K. Hua, J. Zhang, H. Jiang, P. Dubey, A. Blitz, P. van Sijl, and S. Mori, "Reproducibility of quantitative tractography methods applied to cerebral white matter," *Neuroimage*, vol. 36, no. 3, pp. 630–644, 2007.

[115] B. Biswal, M. Mennes, X. Zuo, S. Gohel, C. Kelly, S. Smith, C. Beckmann, J. Adelstein, R. Buckner, and et. al, "Towards discovery science of human brain function," *Proceedings of the National Academy of the Sciences USA*, vol. 107, no. 10, pp. 4734–4739, March 2010.