5-1-2015

New Approaches for Data-mining and Classification of Mental Disorder in Brain Imaging Data

Mustafa Sinan Cetin

Mustafa Sinan Cetin

Candidate

Computer Science

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Vince D. Calhoun, Chairperson

Abdullah Mueen (co-chair)

Shuang Luan

Arvind Caprihan

New Approaches for Data-mining and Classification of Mental Disorder in Brain Imaging Data

by

MUSTAFA SİNAN ÇETİN

B.S, Computer and Control Science, Marmara University, 2004

M.S., Informatics, Istanbul University, 2007

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Computer Science

The University of New Mexico

Albuquerque, New Mexico

May, 2015

DEDICATION

To my beloved parents for their unforgettable sacrifice and devotion,

To my dear wife, Nazan, for her great support and inspiring encouragement.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Vince Calhoun, for his patience, support and guidance throughout this doctoral work. He has been a great mentor to a student like myself, who knew nothing about medical imaging.

I would also like to sincerely thank my co-advisor, Abdullah Mueen, to whom I am indebted for his significant assistance to this work, and for his kindness and support in the difficult moments I went through during this endeavor.

I would like to thank all my wonderful colleagues at Mind Research Network especially those in the Medical Image Analysis Laboratory. Special thanks to Julia Stephen, Eswar Damaraju and Jon Houck for their help and contribution. Also, special thanks to Shuang Luan and Arvind Caprihan for their support and trust.

Finally, I would like to thank my wife, Nazan, and my beloved family for their unconditional love and support. Without their support and inspiration this endeavor would have not been possible.

New Approaches for Data-mining and Classification of Mental Disorder in Brain Imaging Data

by

Mustafa Sinan Cetin

B.S., Computer and Control Science, Marmara University, 2004

M.S., Informatics, Istanbul University, 2007

PhD, Computer Science, University of New Mexico, 2015

ABSTRACT

Brain imaging data are incredibly complex and new information is being learned as approaches to mine these data are developed. In addition to studying the healthy brain, new approaches for using this information to provide information about complex mental illness such as schizophrenia are needed. Functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) are two well-known neuroimaging approaches that provide complementary information, both of which provide a huge amount of data that are not easily modelled.

Currently, diagnosis of mental disorders is based on a patient's self-reported experiences and observed behavior over the longitudinal course of the illness. There is great interest in identifying biologically based marker of illness, rather than relying on symptoms, which are a very indirect manifestation of the illness. The hope is that biological markers will lead to earlier diagnosis and improved treatment as well as reduced costs. Understanding mental disorders is a challenging task due to the complexity of brain structure and function, overlapping features between disorders, small numbers of data sets for training, heterogeneity within disorders, and a very large amount of high dimensional data.

This doctoral work proposes machine learning and data mining based algorithms to detect abnormal functional network connectivity patterns of patients with schizophrenia and distinguish them from healthy controls using 1) independent components obtained from task related fMRI data, 2) functional network correlations based on resting-state and a hierarchy of tasks, and 3) functional network correlations in both fMRI and MEG data. The abnormal activation patterns of the functional network correlation of patients are characterized by using a statistical analysis and then used as an input to classification algorithms.

The framework presented in this doctoral study is able to achieve good characterization of schizophrenia and provides an initial step towards designing an objective biological marker-based diagnostic test for schizophrenia. The methods we develop can also help us to more fully leverage available imaging technology in order to better understand the mystery of the human brain, the most complex organ in the human body.

Contents

Table of Figures

List of Tables

Chapter 1: Introduction

1.1. Motivation

Functional magnetic resonance imaging (fMRI) and Magnetoencephalography (MEG) are two well-known neuroimaging approaches that have been used to understand various mysteries related to the brain. These technologies have provided opportunities to scientists to investigate the integrity of neural circuits and map physical and cognitive actions to different regions within the brain. Also, these neuroimaging approaches enable scientists to examine the joint information between tasks that are obtained from different functional domains. Previous work has yielded information across a wide range of topics including basic sensory processing, tobacco and alcohol use, neurodegenerative diseases, longitudinal studies and neuropsychiatric illnesses.

The number of people affected by mental disorders such as schizophrenia, is substantial and has a significant long term cost both economically and in terms of human suffering. The general approach for the diagnosis of mental disorders is based on a patients self-reported experiences and observed behavior over the longitudinal course of the illness. There is great interest in identifying biologically based marker of illness, rather than relying on symptoms because the current approach may postpone the diagnosis of the disorder whereas early diagnosis can improve treatment response and reduce associated costs (Kubicki, M. et al., 2007). But understanding mental disorders is a challenging task due to the complexity of brain structure and function, overlapping features between disorders, small training subjects, the heterogeneity within disorders and a very large amount of high dimensional data.

The brain regions that show significant differences between patients and healthy individuals can be identified based on activation patterns obtained from rest-state or tasks with certain stimuli such as audio and visual stimuli. These differences motivate us to investigate the abnormal activity of patients with schizophrenia on functional regions of the brain by using machine learning and data mining algorithms. In this work, we propose several approaches to detect abnormal functional network connectivity patterns of patients with schizophrenia and improve the individual prediction of schizophrenia patients. Therefore, we may help to design an objective biological marker-based diagnostic test for schizophrenia.

1.2. Thesis Statement

This doctoral work proposes machine learning and data mining based research algorithms to detect abnormal functional network connectivity patterns of patients with schizophrenia and distinguish them from healthy controls by using time series of non-artifactual independent components obtained from task related fMRI data, functional network correlation data based on resting-state and task hierarchy fMRI experiment, functional network correlation data obtained with fMRI and MEG methods. The abnormal activation patterns of functional network correlation of patients are characterized by using a statistical analysis and these abnormal activation patterns are used as an input data to classification algorithms. The framework presented in this doctoral study lay the groundwork towards methods to achieve a better characterization of schizophrenia to design an objective biological marker-based diagnostic test for schizophrenia.

1.3. Research Goals & Perspective

The general approach for the diagnosis of schizophrenia is primarily based on a patients self-reported experiences and observed behavior over the longitudinal course of the illness. There is great interest in identifying biologically based marker of illness, rather than relying on symptom assessment because the current approach may postpone the diagnosis of the disorder, whereas early diagnosis can improve treatment response and reduce associated costs (Kubicki, M. et al., 2007). But small numbers of training subjects and high dimensional datasets make it challenging to design robust and accurate classifiers for schizophrenia. Functional connectivity shows promise in predicting individual patients. Seed-based functional connectivity approaches assess the temporal correlation between a seed region and individual brain voxels (Cordes D et al., 2002; Fox MD et al., 2005). Independent component analysis (ICA) based functional network connectivity (FNC) is a correlation value that summarizes the overall connection between independent brain maps over time (Allen EA et al., 2012; Calhoun VD et al., 2001a, 2008; Cetin, MS., et al., 2014; Jafri MJ et al., 2008). Therefore, the FNC feature gives a picture of the connectivity pattern over time between independent components.

Examination of intrinsic functional connectivity using functional MRI (fMRI) with resting-state has provided important findings regarding dysconnectivity in schizophrenia. However, exclusive reliance on fMRI to generate such networks may limit inference on dysconnectivity: Whilst the blood oxygenation-level dependent (BOLD) response measured by fMRI allows high spatial resolution maps, it is limited by being an indirect and slow physiological signal (Kim SG, et al., 1997). Neural oscillatory activity, which comprises rhythmic electrical activity in cell assemblies, is thought to underlie BOLD responses. This occurs in the ~1-900Hz band; such rapid electrical signals cannot be assessed using fMRI but can be measured directly by techniques such as magnetoencephalography (MEG). The integration of MEG and fMRI should allow us to interrogate this disconnection with high spatiotemporal resolution.

To date, most studies have focused only on the analysis of functional connectivity during performance of a single task. Changes in connectivity between extended rest and multiple tasks have not been used in the discrimination of schizophrenia patients from healthy controls. Such an approach does not take advantage of the within-subject pattern of response which likely occurs across tasks, and which can be of benefit in a number of applications (Calhoun VD and Adali T, 2009; Calhoun VD et al., 2006, 2008). Only a few studies (Arbabshirani MR et al., 2013a; Cetin, MS., et al., 2014; Repovš G and Barch DM, 2012) have investigated how cognition changes under an established progression of task manipulation, but these studies have not focused on individual subject measures in the context of classification. To gain a broader understanding of brain function and dysfunction as a dynamic process, we must examine how cognition changes under an established progression of task manipulations.

The results of this study lay important groundwork for developing clinical tools that can detect abnormal activation patterns of mental disorders such as schizophrenia. The contributions of this doctoral work are briefly listed and discussed as follows:

- Improving fast candidate evaluation algorithm using multi-length indexing scheme. As a first step of this study, we will use shapelet algorithm to investigate time series of non-artifactual independent components obtained from task related fMRI data for classification

of schizophrenia patients from healthy controls. Time series shapelets are small segments of time series that distinguish between classes based on existence of such segments in the classes. Despite numerous works on shapelet discovery, mostly on efficient algorithms, shapelets for multi-dimensional time series data are yet to be explored because of the added computational requirement for multiple dimensions of time series data. In the simplest form, if we wish to discover shapelets for each of the dimensions separately, the original exact algorithm (Lexiang, Y. and Keogh, E., 2009) takes months while the fastest approximate algorithm would take ten hours on our target neuroimage dataset. First, we plan to overcome the barrier of computational requirements by using fast candidate *generation* and *evaluation*. A typical shapelet discovery algorithm works in two phases. 1) the algorithm generates a set of candidate shapelets and 2) the candidates are evaluated for the information gain they achieve when used as classification features. Surprisingly, all of the speed-up techniques in the literature prune the candidate pool by admissible heuristics (Abdullah, M. et al., 2011) or by random projections (Rakthanmanon, T. and Keogh, E., 2013). In chapter 3, we will describe a fast candidate evaluation algorithm using multi-length indexing scheme, which can be used in conjunction with any prior algorithm.

- Determining whether cortical connectivity patterns remain stable or change across a hierarchy of sensory tasks. To the best of our knowledge there has been no study to investigate this issue in a variety of different FNC networks in a multi-task hierarchy with a relatively large number of subjects. In this study, we examined FNC across a hierarchy of sensory tasks with varying levels of sensory load in chapter 4. Data for each participant were gathered across multiple fMRI scanning sessions over the course of up to two months (1~2 months) with prospective randomization of task presentation and close monitoring of SPs to ensure clinical stability. Our goal is to track connectivity changes in schizophrenia patients and healthy controls as sensory load increased. Using multiple tasks in addition to multiple conditions within a single task allows us to recognize that individuals' reactions to sensory stimuli are conditioned by the circumstances in which such stimuli are presented and measurements at separate time points allows us to better assess state versus trait group differences. We sought to determine whether schizophrenia patients and healthy controls

showed significant FNC differences among brain regions across the task hierarchy by modeling the temporal dependency between functional networks derived from fMRI data. The tasks defined a natural hierarchy related to sensory load and included a rest task, two levels of auditory sensory gating, and two levels of multisensory perception with auditory and audio-visual stimuli. We remain skeptical of the notion that rest differences necessarily equate to characteristic differences in cognition between schizophrenia patients relative to healthy controls. We hypothesized that data collected using a sensory load task hierarchy including rest will provide evidence of both stable (static functional network connectivity effects) and state-based differences (dynamic functional network connectivity effects).

- Investigating hypothesis that the static functional network connectivity effects contain valuable trait-based information that can be used for individual prediction of mental illness such as schizophrenia. In chapter 5, we used a group ICA approach (Calhoun VD and Adali T, 2012; Erhardt EB et al., 2011a), excluded non-artifactual brain networks (Jafri MJ et al., 2008), then conducted a classification study of schizophrenia and healthy subjects using static functional network connectivity effects and compared results with dynamic functional network connectivity effects by examining across a hierarchy of sensory tasks with varying levels of sensory load (including resting-state only). Data for each participant were gathered across multiple fMRI scanning sessions and both FNC analyses were examined across a hierarchy of sensory tasks with varying levels of sensory load.

- Improving a method to use both fMRI and MEG together to investigate healthy normal volunteers and schizophrenia patients. Whilst the blood oxygenation-level dependent (BOLD) response measured by fMRI allows high spatial resolution maps, it is limited by being an indirect and slow physiological signal (Kim SG, et al., 1997). Neural oscillatory activity, which comprises rhythmic electrical activity in cell assemblies, is thought to underlie BOLD responses. This occurs in the ~1-900Hz band; such rapid electrical signals cannot be assessed using fMRI but can be measured directly by techniques such as magnetoencephalography (MEG) (Cohen D., 1968). The purpose of the chapter 6 is to use both fMRI and band limited envelope correlation metrics in MEG to interrogate functional

connectivity in the resting state in a sample of healthy normal volunteers and schizophrenia patients. Using methods based on group spatial ICA, for the first time we estimate networks from both MEG and fMRI and compare and contrast the networks and findings from the three modalities, with the hypotheses that 1) Patients and controls would differ significantly on both MEG and fMRI measures of among-network connectivity, called functional network connectivity (FNC), 2) MEG and fMRI spatial maps would show substantial overlap and 3) Using both MEG and fMRI measures of among-network connectivity would show improvement to classification of schizophrenia patients

1.4. Innovations and Contributions

This doctoral work will help researchers in better understanding abnormal activation patterns of functional network correlation and disconnection hypothesis of patients with schizophrenia. A list of the primary innovations and contributions of this dissertation includes:

- Investigate time series of non-artifactual independent components obtained from task related fMRI data for classification by using a shapelet algorithm.

- Compare and question the applicability of various frameworks for classification of schizophrenia patients and healthy control subjects based on the fMRI data obtained from a hierarchy of sensory tasks with varying levels of sensory load (including rest-state only).

- Combining both fMRI and MEG are used to interrogate functional connectivity in the resting state individual prediction of schizophrenia.

1.5. Organization

The rest of this dissertation describes concepts, techniques, and results we have implemented and analyzed in the course of developing the proposed frameworks to detect abnormal functional network connectivity patterns of patients with schizophrenia and increase the individual prediction of schizophrenia patients. This dissertation is organized as follows:

Chapter 2 provides a brief background description regarding the some of the building blocks of this thesis including fMRI mechanism, MEG mechanism, schizophrenia, independent component analysis and functional network correlation.

Chapter 3 introduces a shapelet based classification algorithm. Time series shapelets are small segments of time series that distinguish between classes based on existence of such segments in the classes. The proposed shapelet algorithm uses time series of independent components of fMRI data.

Chapter 4 describes a novel functional network connectivity analysis by examining across a hierarchy of sensory tasks with varying levels of sensory load to determine whether cortical connectivity patterns remain stable or change across a hierarchy of sensory tasks.

Chapter 5 presents a classification study of schizophrenia patients and healthy subjects using static FNC and compared results with dynamic FNC by examining across a hierarchy of sensory tasks with varying levels of sensory load (including resting-state only).

Chaplet 6 describes a novel functional network connectivity analysis by combining both fMRI and MEG methods. They are used to interrogate functional connectivity in the resting state individual prediction of schizophrenia. Chapter 6 also presents a multi-model classification study of schizophrenia patients and healthy subjects by using FNC data obtained from fMRI and MEG

Chapter 7, features conclusions, future works, and recommendations.

Chapter 2: Background

In this chapter, we will provide a brief background description regarding the some of the building blocks of this thesis including fMRI mechanism, MEG mechanism, schizophrenia, independent component analysis and functional network correlation

2.1. fMRI

Functional magnetic resonance imaging (fMRI) is a type of specialized neuroimaging MRI-related technique. It has a temporal resolution of about one second and spatial resolution of 1-3mm^3 (which is known a voxel) also it is used to capture the functional activation of brain regions (Smith, S.M., 2004). See Figure 2 - 1. Although brain activity known as a transfer of electrical and chemical energy between the different regions of the brain, fMRI measures indirect levels of brain activity associated with a physical or mental action.

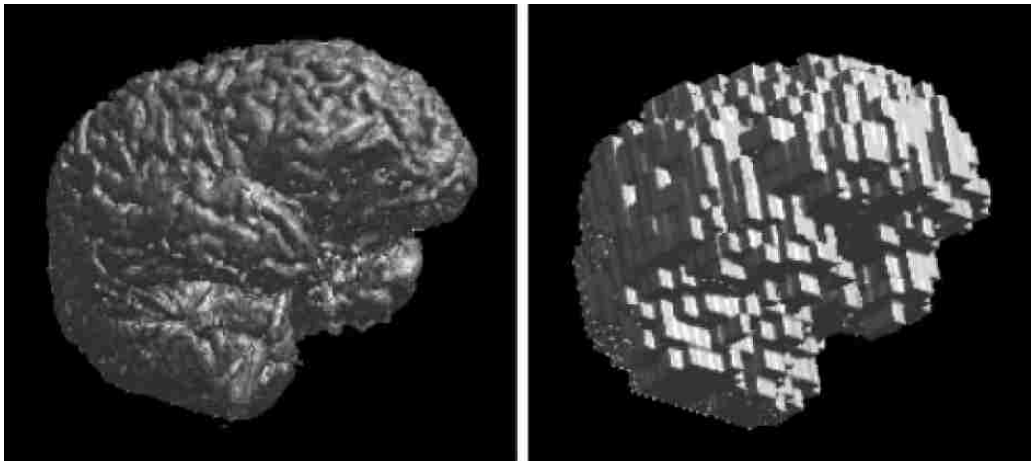

Figure 2 - 1: Surface renderings of 3D brain images. (a) High resolution image (1×1×1) voxels and (b) low-resolution (7×7×7) image of the same brain (Smith, S.M., 2004).

The hemodynamic response of a neural activity at a certain brain region that obtains additional oxygen consumption and oxygen present in the neighborhood of that brain region is imaged by fMRI. The neurovascular linkage between networks of neurons and blood vessels results in the exchange of energy that further causes change in oxygenated hemoglobin (Huettel, S., et al., 2004). Blood oxygenation-level dependent (BOLD) activations are considered an acceptable indicator of bundled neural activity by scientists.

When the brain supplies an active brain region with oxygenated blood, it causes a drop in the deoxyhemoglobin concentration around the active region and change the ratio of oxyhemoglobin and deoxyhemoglobin. See Figure 2 - 2. This change the activity of magnetic properties of the environment around the activated brain region. While a subject is asked to perform a task such as pressing a button, the scanner records BOLD changes of different brain regions. In other words, during an fMRI experiment, fMRI measures changes in deoxyhemoglobin concentrations in nearby (to neurons) blood vessels.

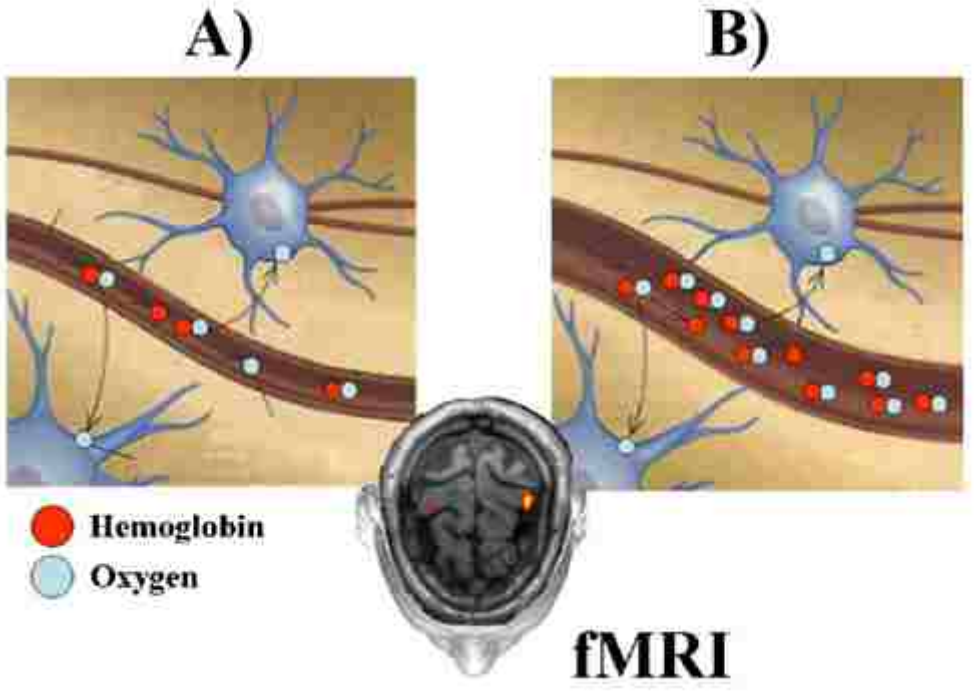

Figure 2 - 2 : Physiologic principle of fMRI signals; (a) increasing the oxygen consumption, then (b) hemodynamic response in a second scale (Astolfi, L., et al., 2004)

fMRI results have been used in the identification of neuropsychological disorders such as schizophrenia. Pathophysiology of schizophrenia investigated with fMRI, in particular to assess the disconnection hypothesis of schizophrenia (Friston KJ and Frith CD, 1995; Woodward ND, 2012). fMRI allows researchers to investigate groups of patients with schizophrenia and healthy controls and identify differential brain activation between these groups.

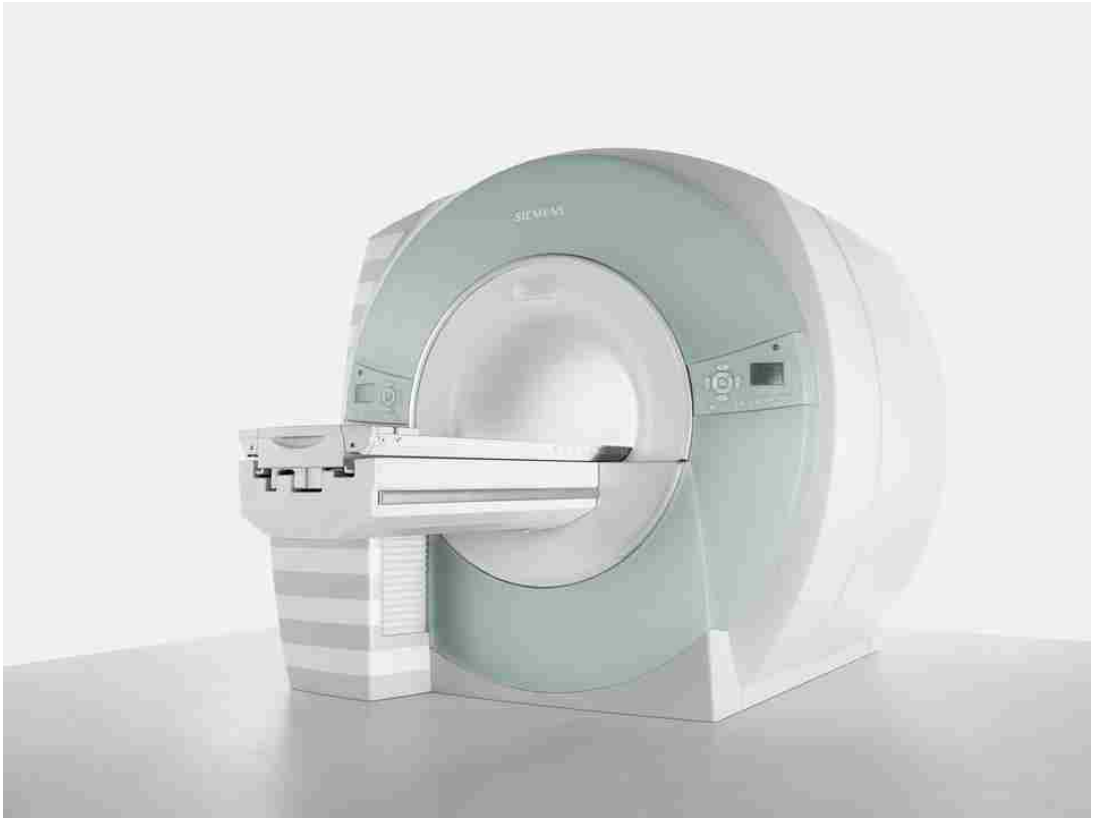

Figure 2 - 3 : An MRI scanner. Siemens magnetom trio scanner at the Mind Research Network.

2.1.1. fMRI Experiment

The goal of the fMRI experiment is recording the activity of brain regions. During an fMRI experiment, the subject is requested to lie in a MRI scanner (See Figure 2 - 3), not move his/her head since motion is one of the main artifacts in an fMRI experiment and performs a task while the scanner captures the hemodynamic response from different brain regions as neural activity. If it is necessary for the experiment, tools such as a projector, speaker, headphones, button boxes and microphones may be used together with the MR machine with special settings due to a very strong magnetic field. See Figure 2 - 4.

Generally there are two types of experiments: (1) in the resting-state experiments, participants performed a simple rest task. Their eyes were open during the scan and they gazed passively at a

central fixation cross. (2) In a task-based experiment, subjects need to attend to a task and respond if it is required. There are two main types of task-based fMRI designs: (a) in a block design task experiment, conditions are alternated in order to determine the differences among them. (b) In an event related design task experiment, conditions are randomized throughout the course of the experiment to simulate a real world experiment.

During any kind of fMRI experiment the main coil in the MRI machine makes high steady magnetic field in the chamber of the scanner. Then the gradient coil creates small changes in the steady magnetic field. In the final stage, radio frequency coil that emits radio frequency signals to excite the protons that are spinning at a certain frequency

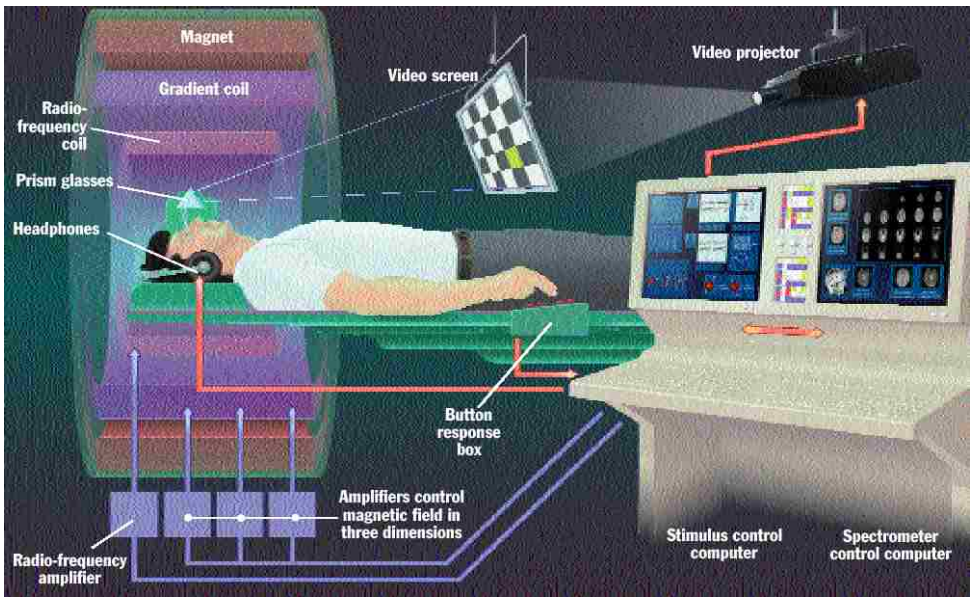

Figure 2 - 4 : Experimental fMRI setting for testing of responses to visual stimuli (http://www.ece.unm.edu/~vcalhoun/courses/fMRI_Spring12)

2.2. MEG

Magnetoencephalography (MEG) is increasingly available and is regarded as one of the most popular imaging methods. MEG is a noninvasive neuroimaging technique for detecting the

magnetic fields associated with the intracellular current flow within neurons and recording this brain activity (D. Cohen, 1972) which allows real-time investigation of cortical activity. See Figure 2 - 5.

In the MEG method, even very small magnetic field signals are amplified by using superconducting quantum interference devices. The millisecond temporal and spatial resolution of MEG, distinguish it from fMRI techniques. MEG involves the measurement of neuromagnetic signals emanating from the brain. Magnetic activity measured outside the head is produced primarily by intracellular electrical currents within the dendrites of pyramidal cells in neo- and archeo-cortical brain structures.

Although MEG provides some advantages to investigate cognitive activity such as millisecond temporal and spatial resolution, it comes with some limitations which are listed below:

- MEG measurements have to be taken in a magnetically shielded laboratory environment due to the neuromagnetic signals being very weak compared to the magnetic fields in a laboratory environment.

- In order to produce activation maps, MEG data requires to be combined with MR data into a composite image of function overlaid on anatomy. Due to lack of anatomical or structural information in MEG data.

- Obtaining reliable information regarding the subcortical sources of brain activity is one of the main challenges. Due to the existence of the spherical symmetry of the head and distance to the source.

- Localization of the sources of activity within the brain from magnetic measurement outside the head and helmed is another challenge with MEG.

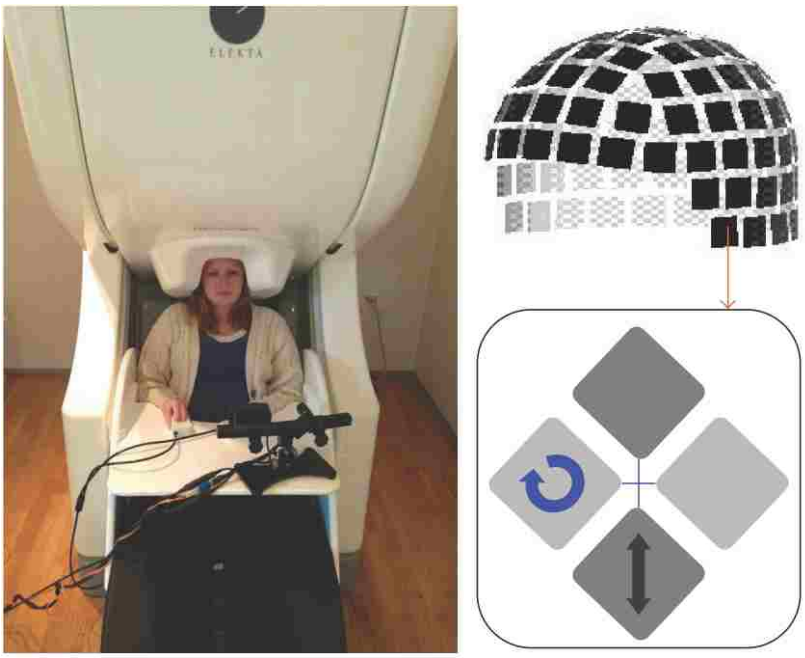

Figure 2 - 5: Experimental MEG setting for testing of responses, (a) MEG scanner, (b) helmet-shaped array and magnetometer coils (Braeutigam, S., 2013).

2.2.1. MEG Experiment

The goal of the MEG experiment is recording the activity of the brain regions. During an MEG experiment, the subject is requested to sit under the MEG scanner and performs a task while a helmet-shaped array captures the magnetic fields via 3 pick-up coils which are called a channel. There are 360 channels in a helmet-shaped array (Braeutigam, S., 2013).

If it is necessary for the experiment, tools such as a projector, speaker, headphones, button boxes and microphone may be used together with the MEG scanner with special settings due to magnetic field.

2.3. Schizophrenia

Schizophrenia is a mental disorder, affecting one's thoughts, feeling and acts. It has been widely accepted as a neurodevelopmental disorder substantially affecting the neuroanatomy and brain function, including functional networks during resting-state and task (Bullmore E and Sporns O, 2009; Millier, A., et al., 2014; Van Den Heuvel MP., and Hulshoff Pol HE., 2010; Van DKR et al., 2010). Schizophrenia is a complex disease that affects 1% of the population and across all cultures, genders and socioeconomic groups (Bhurga D., 2005; Leucht S, et al., 2007; Millier, A., et al., 2014). It was reported as the 5th leading worldwide cause of global disease burden in 2004 by The World Health Organization.

Although the root cause of schizophrenia is still unclear, many recent research studies have pointed to the combination of environmental and genetic factors as possible contributing causes to the disease. The most accepted hypothesis regarding what causes schizophrenia is the disconnection hypothesis (Friston K., 1998). It is based on disturbances in white matter connectivity between different brain regions. Fiber number or density differences of white matter (See Figure 2 - 7) between schizophrenia patients and healthy controls are possibly due to abnormalities in the myelin sheaths around the axons (Foong J, et al., 2002; Minami T, et al., 2003). Also, some early researchers have found the existence of diminished white matter anisotropy in some areas of schizophrenic brain such as the prefrontal region (Buchsbaum M, et al., 1998). See Figure 2 - 6. In conclusion, the diminished white matter differences in schizophrenia patients are not circumscribed and are characterized by a deficit of interconnections between different parts of the brain (Allen EA et al., 2012; Breakspear M, et al., 2003; Calhoun VD and Adali T, 2009; Calhoun VD et al., 2006, 2008; Friston K., 1998; Kubicki, M. et al., 2007).

General symptoms of schizophrenia include hallucinations, emotional dysregulation disorganized behavior and difficulty in separating between reality and delusional formations. Symptoms of schizophrenia are divided into 3 subgroups (Andreasen N, et al., 1998; Andreasen NC, et al., 2005).

- *Positive symptoms* are manifestations of the disorder and include delusions, auditory hallucinations, thought disorder and disorders of movement.

- *Negative symptoms* are considered the loss of normal traits and abilities and include symptoms such as lack of emotion, lack of motivation, poverty of speech and failure to experience pleasure.
- *Disorganization Symptoms* are categorized affective incongruity, attentional impairment such as problems with working memory and executive functions that needed to plan and organize.

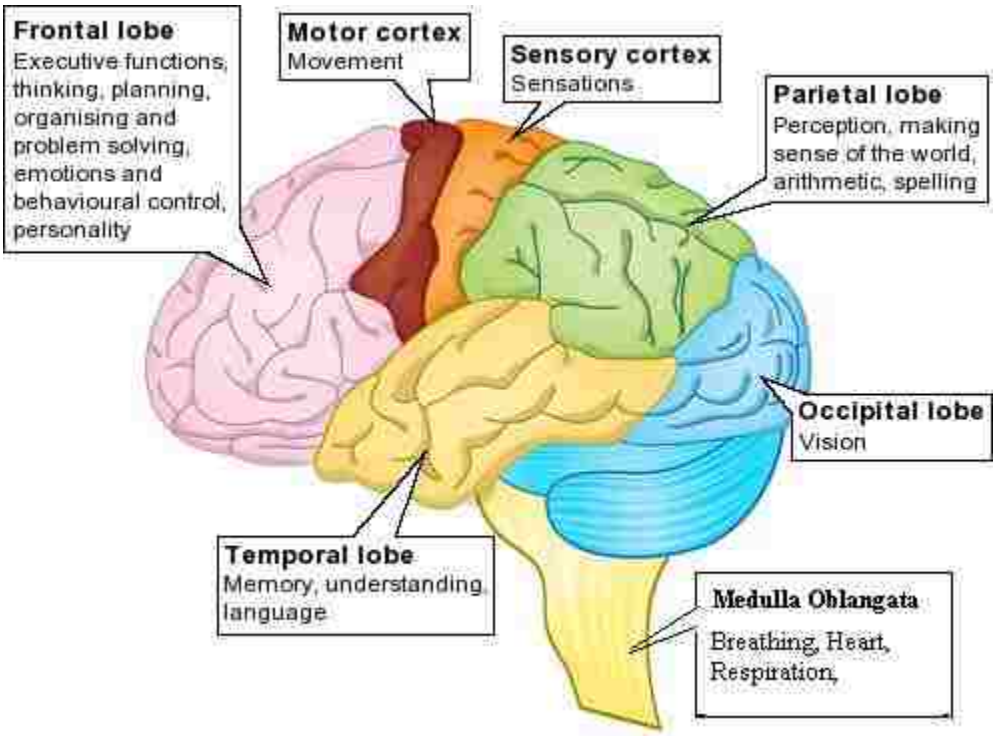

Figure 2 - 6: A lateral/sagittal view of the human brain with locations of frontal, temporal, motor, sensory, parietal, occipital, and medulla oblongata lobes. (http://www.smartketing.net/?p=49)

Studies separately analyzing structural and functional images have found that multiple brain regions appear to be affected in schizophrenia (Allen EA et al., 2012; Cetin, MS., et al., 2014; Goldstein JM, et al., 1999; Honea R, et al., 2005). Functional images studies showed that the population mean of schizophrenia patients perform at least one standard deviation below then healthy controls in various areas such as attention, memory, motor speed, executive functions, ability to acquire skills, problem solving and community functioning (Green M, et al., 2000). Also,

structural image studies reported that there is structural brain differences in schizophrenia patients and healthy controls such as larger ventricles in patients and, overall gray matter volume can be lower in patients.

There are also social and medical negative impact of this disorder such as personality changes, social isolation, occupational disability, cognitive impairment, susceptibility to suicidal behavior and poor health (Carlborg A, et al., 2010). There are currently no generally accepted curative treatments for schizophrenia (Andreasen NC, et al., 2005; Van Os J, et al., 2006).

The costs of the disease's management remain large for individuals and society. While literature on the economic impact of schizophrenia is abundant, few studies have focused on its humanistic burden. Estimated the schizophrenia-associated direct medical costs at $2.13 billion (Desai et al., 2013; McDonald et al., 2005). This does not only concern patients, but also caregivers, relatives, neighbors and others in a patient's daily life(Abouzaid S, et al., 2010; de Silva J, et al., 2012; Millier, A., et al., 2014; Willis M, et al., 2010).

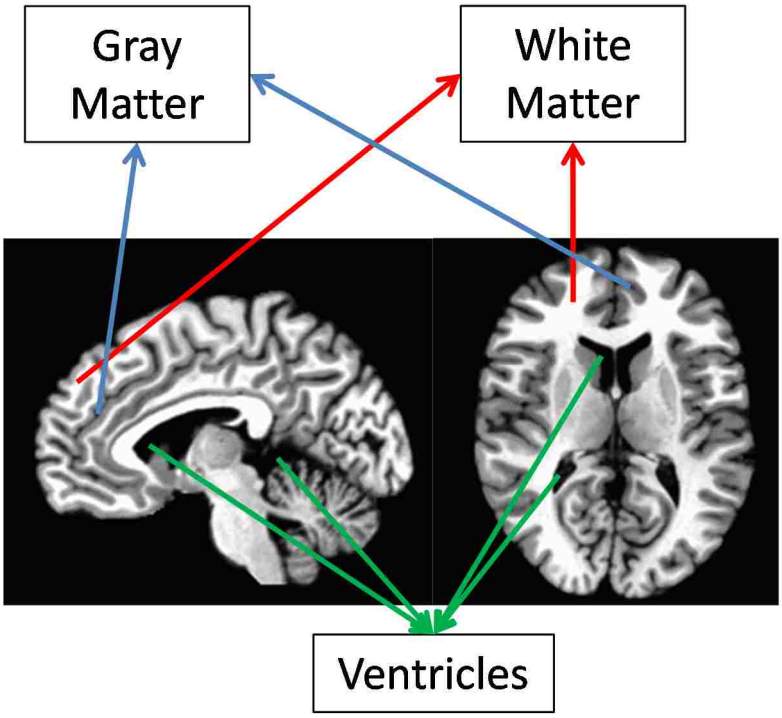

Figure 2 - 7: A lateral/sagittal and axial slice of a structural human brain map; gray matter, white matter and ventricles.

2.4. Independent Component Analysis

ICA is a technique that separates multivariate signals into statistically independent components. It is used as a data-driven approach for resting-state and task-related fMRI data (Beckmann C et al., 2005; Biswal, B., et al., 1995; Calhoun VD et al., 2001a, 2001b, 2001c; Cetin, MS., et al., 2014; Houck JM, et al., 2015; Jafri MJ et al., 2008; McKeown MJ. et al., 1998).

ICA assumes a generative model that identifies the components (cluster of voxels) from fMRI data. The observations are linear mixtures of independent sources. In fMRI data, since each component reflects brain regions which exhibit temporal coherence components are maximally independent and linearly mixed. Also, ICA is used to discover differences in temporal dynamics

and changes with respect to spatially distributed brain networks (Calhoun VD et al., 2001a, 2001b, 2001c). Mathematically ICA formulation can be written as:

$$X = AS$$

Where A is the unknown mixing matrix, S is the spatial component map whose elements are independent sources, and X is the observed vector from the fMRI data by using all subjects. The ICA technique used with fMRI data has two categories (Calhoun VD et al., 2001a) : (1) Temporal ICA (*t*ICA) that decomposes the fMRI data into independent time-series and (2) spatial ICA (*s*ICA) decomposes the data into independent spatial maps which is the most common application of ICA since it recovers independent specialized networks in the brain and their corresponding time-courses.

Also, it is possible to use ICA with a group of subjects. The intuition behind this technique is finding the activation maps of voxels that are correlated to each other while maximally independent of other sets of voxels. Then, each subject's spatial activation patterns can be found by using back-reconstruction. The details of ICA and its implementation can be found in previous studies (Calhoun VD et al., 2001a, 2001b, 2001c), they are not covered here.

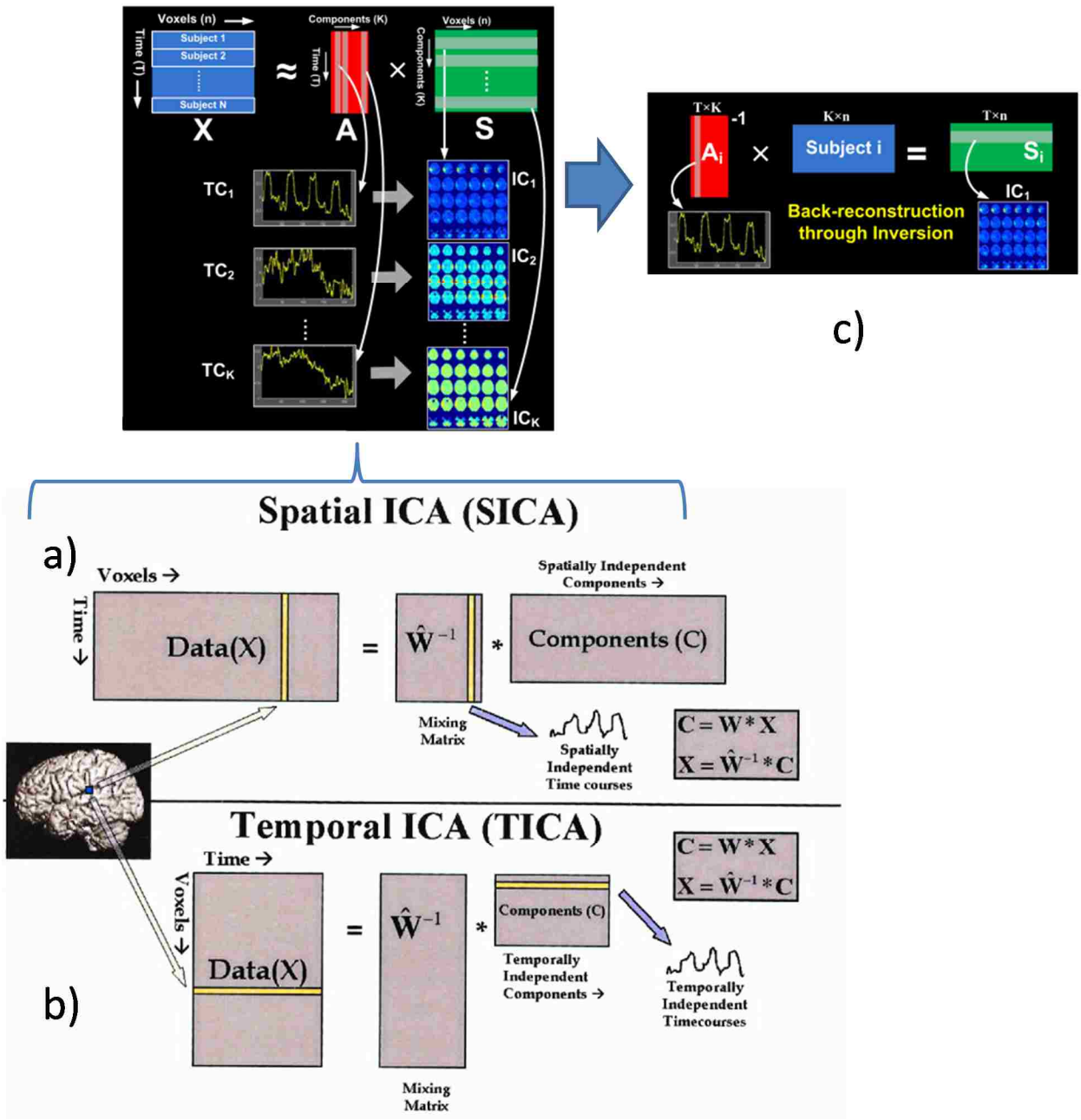

Figure 2 - 8: Illustration of two types of ICA on fMRI data: a) Spatial ICA, b) Temporal ICA and c) Back-reconstruction (Calhoun VD et al., 2001a, 2001b, 2001c).

2.5 Functional Network Connectivity

Functional connectivity shows promise in providing individual subject predictive power. Seed-based functional connectivity approaches assess the temporal correlation between a seed region and individual brain voxels (Cordes D et al., 2002; Fox MD et al., 2005). Independent component analysis based on functional connectivity, also known as functional network connectivity is considered a high level functional connectivity. Functional network connectivity (FNC) is a correlation value that summarizes the overall connection between independent brain maps over time (Arbabshirani MR et al., 2013a; Jafri MJ et al., 2008). Therefore, the FNC feature gives a picture of the connectivity pattern over time between independent components.

The provided FNC information was obtained by using fMRI and MEG methods from a set of schizophrenia patients and healthy controls, using GICA. The GICA decomposition of the preprocessed fMRI data resulted in a set of brain maps, and corresponding timecourses. These timecourses indicated the activity level of the corresponding brain map at each point in time. The FNC features were the pair-wise correlations between these timecourses, for each subject. FNC indicates a subject's overall level of 'synchronicity' between brain areas. Because this information is derived from fMRI and MEG scans, FNCs are considered a functional modality feature (i.e., they describe patterns of brain function). Figure 2 - 9 shows an example of functional network connectivity of fMRI scans among seven brain networks for 28 heathy controls (Jafri MJ et al., 2008).

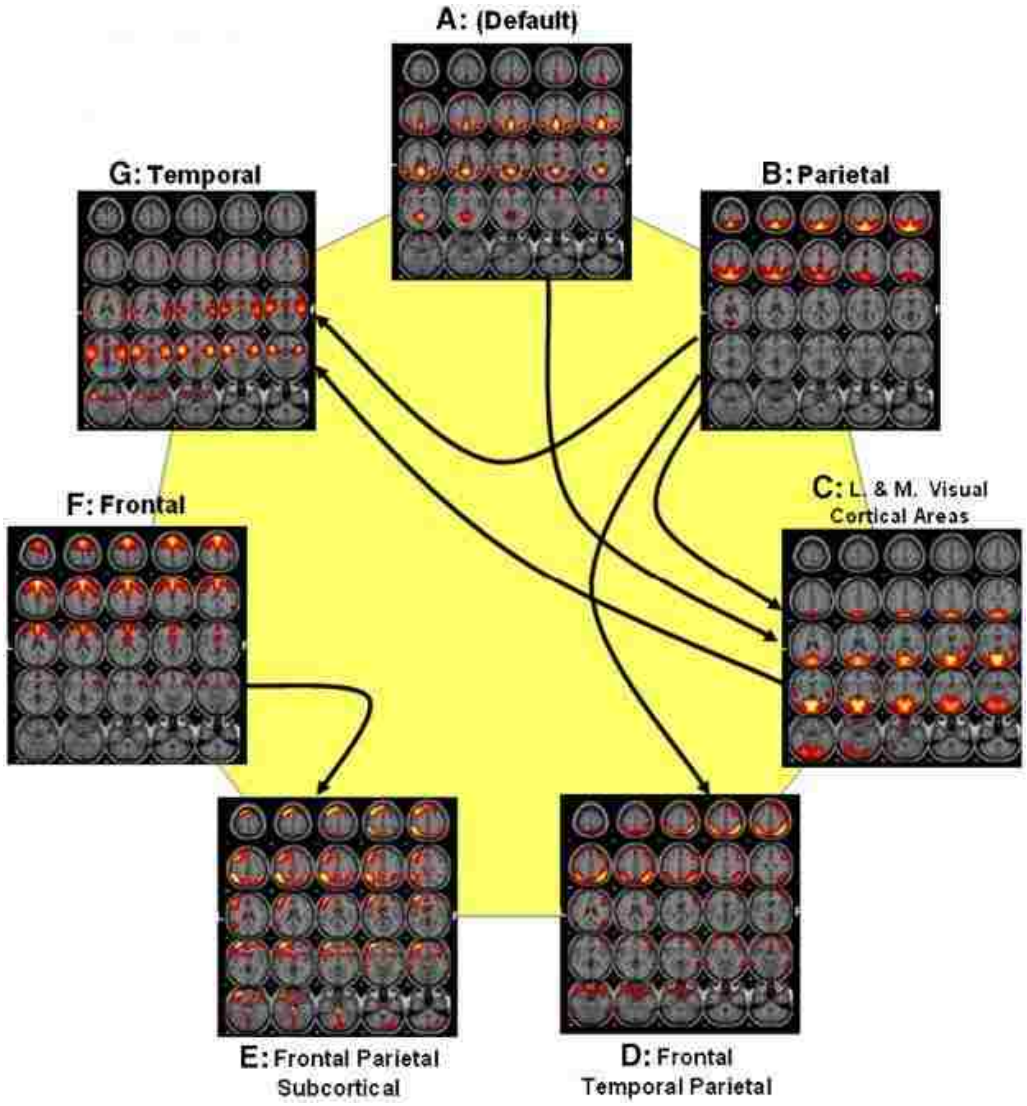

Figure 2 - 9 : Functional network connectivity of 28 heathy controls (Jafri MJ et al., 2008).

Chapter 3: Shapelet Ensemble for Multi-dimensional Time Series

Time series shapelets are small subsequences that maximally differentiate classes of time series. Since the inception of shapelets, researchers have used shapelets for various data domains including anthropology and health care, and in the process suggested many efficient techniques for shapelet discovery. However, multi-dimensional time series data poses unique challenges to shapelet discovery that are yet to be solved.

We show that an ensemble of shapelet-based decision trees on individual dimensions works better than shapelets defined over multiple dimensions. Generating a shapelet ensemble for multi-dimensional time series is computationally expensive. Most of the existing techniques *prune* shapelet candidates for speed. In this chapter, we propose a novel technique for shapelet discovery that *evaluates* remaining candidates efficiently. Our algorithm uses a multi-length approximate index for time series data to efficiently find the nearest neighbors of the candidate shapelets. We employ a simple skipping technique for additional candidate pruning and a voting based technique to improve accuracy while retaining interpretability. Not only do we find a significant speed increase, our techniques enable us to efficiently discover shapelets on datasets with multi-dimensional and long time series such as hours of brain activity recordings. We demonstrate our approach on a biomedical dataset and find significant differences between patients with schizophrenia and healthy controls.

3.1. Introduction

Time series shapelets are small segments of time series that distinguish between classes based on existence of such segments in the classes. Figure 3 - 1 shows an example of a shapelet in real ECG data. The idea of shapelet discovery has gained popularity for its intuitive classification rules. Shapelets are not just useful for classification purposes, any dataset of long time series can potentially be represented by a set of shapelets as such the shapelets perform an empirical basis (Jason Lines et al., 2012). Shapelets have also been used for unsupervised knowledge discovery such as clustering (Jesin, Z. et al., 2012).

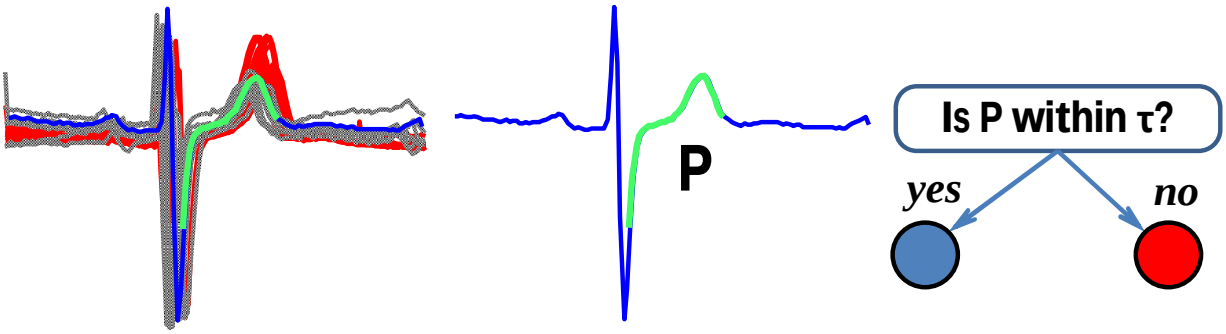

Figure 3 - 1 : An example of shapelet. (left) Heartbeats of a 67 year old male in two different days shown in blue and red. Red has a higher peak than the blue at the T wave. (middle) The shapelet (P) that distinguishes the classes most. (right) A highly accurate decision tree using the shapelet P.

Despite numerous works on shapelet discovery, mostly on efficient algorithms, shapelets for multi-dimensional time series data are yet to be explored because of the added computational requirement for multiple dimensions of time series data. In the simplest form, if we wish to discover shapelets for each of the dimensions separately, the original exact algorithm (Lexiang, Y. and Keogh, E., 2009) takes months while the fastest approximate algorithm would take ten hours on our target dataset of Electroencephalogram. In this chapter, we overcome the barrier of computational requirements by fast candidate generation and evaluation. The new algorithm allows us to generate a simple ensemble of shapelets from individual dimensions. In Figure 3 - 2, we show test accuracies on a dataset (functional MRI) of 540 dimensions recorded from 43 healthy controls and 32 schizophrenic patients. Individual shapelet classifiers perform similar to a random classifier for lack of enough information, while the ensemble achieves a significant accuracy. Our experiments show that ensemble of shapelet classifiers for individual dimensions of a multi-dimensional time series data is promising for electrical sensors such as EEG and Accelerometers.

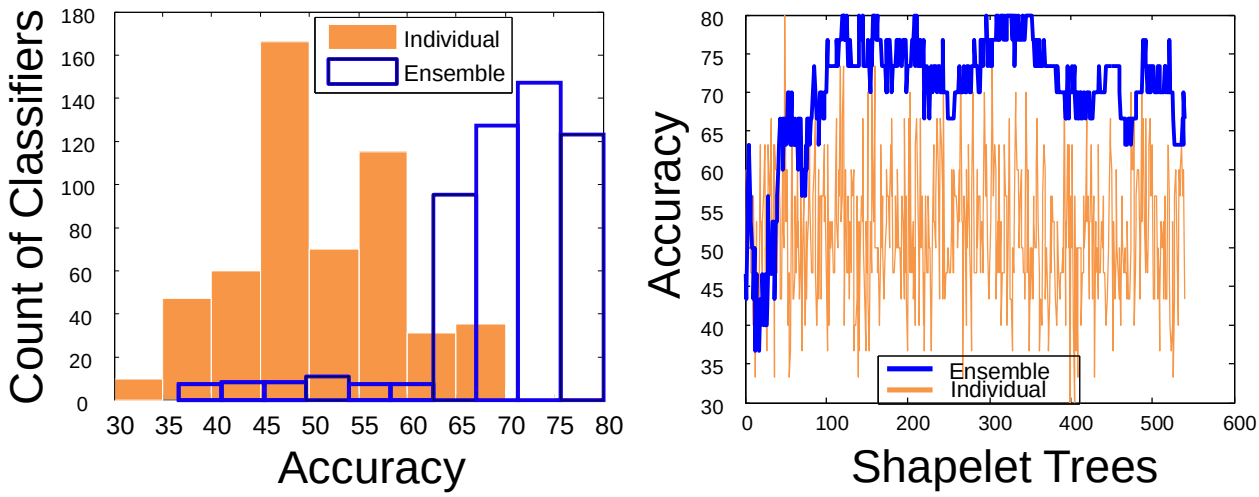

Figure 3 - 2 : Ensembling can achieve better accuracies than all individual trees on unseen test data. (left) Shift in the distribution of classifiers when shapelet trees are ensemble. (right) Each blue point is the accuracy from an ensemble of all the individual trees to its left.

A typical shapelet discovery algorithm works in two phases. First, the algorithm generates a set of candidate shapelets and next, the candidates are evaluated for the information gain they achieve when used as classification features. Surprisingly, all of the speed-up techniques in the literature prune the candidate pool by admissible heuristics (Abdullah, M. et al., 2011) or by random projections (Rakthanmanon, T. and Keogh, E., 2013). In this chapter, we describe a fast candidate evaluation algorithm using multi-length indexing scheme, which can be used in conjunction with any prior algorithm.

Our algorithm, named mc^2, uses a dynamic stepping technique for massive candidate pruning. It achieves an order of magnitude speed-up (up to 9.58x) over current state of the art algorithm (Rakthanmanon, T. and Keogh, E., 2013) and several order of magnitude speed-ups (up to 281x) over the original algorithm (Lexiang, Y. and Keogh, E., 2009) without having a significant difference in accuracy. Moreover, the increased speed allows us to train shapelet classifiers for every dimension of a multi-dimensional time series for an ensemble classifier and thus, achieve increased accuracy and reduced variance.

The chapter is structured as follows. Section 3.2 introduces the definitions and notations. In section 3.3, we described previous and related work. In section 3.4, the skeleton of the shapelet algorithm is discussed. Section 3.5-6 describes the changes we have made to the shapelet algorithm and propose a way to speed up shapelet algorithm, and section 3.7, we demonstrate the performance and accuracy of our algorithm with 44 data sets. Section 3.8 shows the case studies. Finally, in Section 3.9, we form our conclusions.

3.2. Definition and Notation

We start with defining shapelets and the other notations used in the chapter.

A *Time Series T* is a sequence of real numbers $t_1, t_2, \ldots, t_m$. A time series subsequence $S_{i,l} = t_i, t_{i+1}, \ldots, t_{i+l-1}$ is a continuous subsequence of T starting at position i and length l. A time series of length m can have $m(m+1)/2$ subsequences of all possible lengths from one to m. If we are given two time series X and Y of the same length m, we can use the Euclidean norms of their difference (i.e. $X - Y$) as a distance measure. To achieve scale and offset invariance, we must normalize the individual time series before the actual distance is computed and we can do this via *z-normalization*. Normalization is a critical step; even tiny differences in scale and offset rapidly swamp any similarity in shape (Abdullah, M. et al., 2009). In addition, we normalize the distances by dividing with the length of the time series. This allows comparability of distances for pairs of time series of various lengths. We call this *length-normalization*.

The normalized Euclidean distance is generally computed by the formula $\sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - y_i)^2}$. Thus, the time for computing a distance value is linearly related to the length of the time series. In contrast, we compute the normalized Euclidean distance between X and Y using five numbers derived from X and Y. These numbers are denoted as sufficient statistics in (Yasushi, S. et al., 2005). The numbers are $\sum x, \sum y, \sum x^2, \sum y^2$ and $\sum xy$. As it was made clear in (Abdullah, M. et al., 2011), computing the distance in this manner enables us to reuse computations and reduce the amortized time complexity from linear to constant.

The sample mean and standard deviation can be computed from these statistics as $\mu_x = \frac{1}{m}\sum x$ and $\sigma_x^2 = \frac{1}{m}\sum x^2 - \mu_x^2$, respectively. The positive correlation and the normalized Euclidean distance between *X* and *Y* can then be expressed as $dist(x,y) = \sqrt{2(1 - C(x,y))}$ where $C(x,y) = \frac{\sum xy - m\mu_x\mu_y}{m\sigma_x\sigma_y}$).

Many time series data mining algorithms (k-NN classification, clustering, density estimation, etc.) require only comparisons of time series that are of equal lengths. The major reason is because the underlying distance metrics they rely upon do not allow varying lengths. In contrast, time series shapelets require us to test if a short time series (the shapelet) is *contained within* a certain threshold somewhere inside a much longer time series. To achieve this, the shorter time series is slid against the longer one to find the best possible alignment between them. We call this distance measurement the *subsequence distance* and define it as $sdist(x,y) = \sqrt{2(1 - C_s(x,y))}$ where *x* and *y* are the two time series with lengths *m* and *n*, respectively, and for $m \leq n$.

$$C_s(x,y) = min_{0 \leq l \leq n-m} \frac{\sum_{i-1}^{m} x_i y_{i+l} - m\mu_x\mu_y}{m\sigma_x\sigma_y}$$

In the above definition μ_y and σ_y denote the mean and standard deviation of *m* consecutive values from *y* starting at position $l + 1$. Note that, *sdist* is not symmetric.

Assume that we have a dataset *D* of *n* time series from *C* different classes. Let us also assume, every class *i* ($i = 1, 2, \ldots, C$) has n_i labeled instances in the dataset where $\sum_i n_i = n$. An instance time series in *D* is also denoted by D_i for $i = 1, 2, \ldots, n$. The entropy of a dataset *D* is defined as

$$E(D) = -\sum_{i=1}^{C} \frac{n_i}{n} log(\frac{n_i}{n}).$$

If the smallest time series in *D* is of length *m*, there are at least $n\frac{m(m+1)}{2}$ subsequences in *D* that are shorter than every time series in *D*. We define a split as a tuple (s, τ) where *s* is a subsequence and τ is a distance threshold. A split divides the dataset *D* into two disjoint subsets or partitions D_{left} =

$\{x : x \in D, sdist(s, x) \leq \tau\}$ and $D_{right} = \{x : x \in D, sdist(s, x) > \tau\}$. We use two quantities to measure the goodness of a split: *information gain* and *separation gap*.

Definition 1. The information gain of a split is

$$I(s, \tau) = E(D) - \frac{|D_{left}|}{N} E(D_{left}) - \frac{|D_{right}|}{N} E(D_{right})$$

Definition 2. The separation gap of a split is

$$G(s, \tau) = \frac{1}{|D_{right}|} \sum_{x \in D_{right}} sdist(s, x) - \frac{1}{|D_{left}|} \sum_{x \in D_{left}} sdist(s, x)$$

Definition 3. The shapelet for a dataset *D* is a tuple (s, τ) of a subsequence of an instance in *D* and a distance threshold (i.e. a split, τ) that has the maximum information gain while breaking ties by maximizing the separation gap.

We visually summarize the concept of shapelets with our toy example shown in Figure 3 - 3. Shapelets can be organized in a decision tree format until the training set achieves desired level of representation.

To extend the definition of shapelets to multi-dimensional time series data, several aspects should be addressed. Do we find shapelets in *all* dimensions or in *some* of them? Will the shapelets be *time aligned*? Will they be equally *long*? And most importantly, how do we build a *classifier* using the shapelets? We answer these questions with an *independence assumption* among the dimensions of the time series. Under this assumption, it is sufficient to discover shapelets of arbitrary *size* and *location* for *every* dimension and *ensemble* them instead of one giant tree containing them.

Definition 4. Shapelets for a multi-dimensional time series data *D* is a set of shapelet-based decision trees on individual dimensions that maximizes training accuracy upon ensembling.

3.3. Previous and Related Work

Since the inception of shapelets, various shapelet discovery algorithms have been proposed to improve brute force approach. Lexiang *et.al.* (Lexiang, Y. and Keogh, E., 2009) introduced the first improvement by proposing a technique for abandoning some unfruitful entropy computations early. However, this does not improve the worst case complexity. Mueen *et al.*(Abdullah, M. et al., 2011) reduced the worst case complexity by caching distance computations for future use and using a triangular inequality based pruning strategy that achieves an order of magnitude speedup over the method of Lexiang *et al.* (Lexiang, Y. and Keogh, E., 2009). Both of these methods achieved admissible pruning and thus retained the exactness of the gain maximization.

Rakthanmanon, T. *et al.* (Rakthanmanon, T. and Keogh, E., 2013) used a random projection technique (Rakthanmanon, T. and Keogh, E., 2013; Tompa, M. and Buhler, J., 2002) using the SAX representation (Lin, J. et al., 2007; L. Wei, et al., 2006) to find potential shapelet candidates sacrificing the exactness for speed. This algorithm separates the candidate generation and evaluation process by first generating a *constant number* of candidates and then, evaluating them for an overall computation time of $O(nm^2)$ while retaining significant accuracy. Our proposed algorithm builds upon the *FastShapelet* (Rakthanmanon, T. and Keogh, E., 2013) algorithm. Our algorithm skips candidates in the candidate generation process heuristically while improving the candidate evaluation process to reduce running time significantly. We compare our algorithm with the *FastShapelet* algorithm as it is the current *state of the art* for shapelet discovery.

Chang, K.-W. *et al.* (Chang, K.-W. et al., 2012) improved a dynamic programming algorithm for highly parallel Graphics Process Units (GPUs). Results showed that the proposed GPU implementation significantly reduces the running time of the shapelet discovery algorithm. Lines, J. *et al.*(Jason Lines et al., 2012) have used several features based on shape-similarity and named it as shapelet transformation. Although shapelet transform has competitive classification accuracy, it loses the interpretability of the decision tree based classifiers over individual shapes.

To the best of our knowledge, ours is the first attempt to generalize shapelet discovery to multi-dimensional time series data. Our multi-length indexing scheme for time series data is the first of

its kind. Previous work in multi-resolution indexing for images (Ljosa, V. et al., 2006) use more space to cache different resolution while our method does not use any extra space.

Table 3 - 1 : Brute Force Algorithm

Algorithm 1 Shapelet_Discovery(D)

Require : A dataset D of time series

Ensure : Return the shapelet

1: *max_length* = maximum length of a time series in D

2: *max_gain* = 0, *min_gap* = 0, *m* = maximum shapelet length

3: **for** $j = 1$ **to** |D| **do** {every time series in D}

4: $S = D_j$

5: **for** $l = 1$ **to** m **do** {every possible length}

6: **for** $i = 1$ **to** $|S| - l + 1$ **do** {every start position}

7: **for** $k = 1$ **to** |D| **do** {compute distances of every time series to the candidate shapelet $S_{i,l}$}

8: $L_k = sdist(S_{i,l}, D_k)$

9: sort(L)

10: [*gain*,*gap*] = *CalculateInfoGain*(L)

11 **if** (*gain* > *max_gain*) or ((*gain* == *max_gain*) and (*gap* > *min_gap*))

12: **then** *max_gain* = *gain*, *min_gap* = *gap*,

13: *bestshapelet* = $S_{i,l}$, *best*τ = τ

3.4. One Dimensional Shapelet Discovery

In order to properly explain our contribution, we define the brute-force algorithm for shapelet discovery and refine the algorithm in progression.

The brute force shapelet discovery algorithm shown in Table 3 - 1 is a simple algorithm that generates and tests all possible candidates and returns the best one. The final shapelet can be of any length, all subsequences of every length in the dataset D is generated as candidate subsequences $S_{i,l}$ in the three loops in lines 3, 5, and 6 of algorithm 1. In lines 7-9, an array L is created which holds the points in D in the sorted order of their subsequence distance from the shapelet candidate. Finally, the information gain is computed in line 10 and returns the candidate with maximum information gain.

For each of the candidates, we need to compute subsequence distances to each of the n time series in D (line 8) by using a distance function *sdist* which essentially finds the nearest neighbor distance between the candidate and subsequences of the instance time series. Note that *sdist* is the inner-most statement of the above algorithm and thus, a slight improvement in efficiency will give us a large payoff. Figure 3 - 3 shows how the *sdist* between a candidate and n (=8) instances are organized in the array L.

As shown in Table 3 - 1, in line 7-9, an array L is created which holds the points in D in the sorted order of their distance from the shapelet candidate. The ideal shapelet is the one that orders the data as such all instances of one class are near the origin, and all instances of the other classes are to the far right, with no interleaving of the classes.

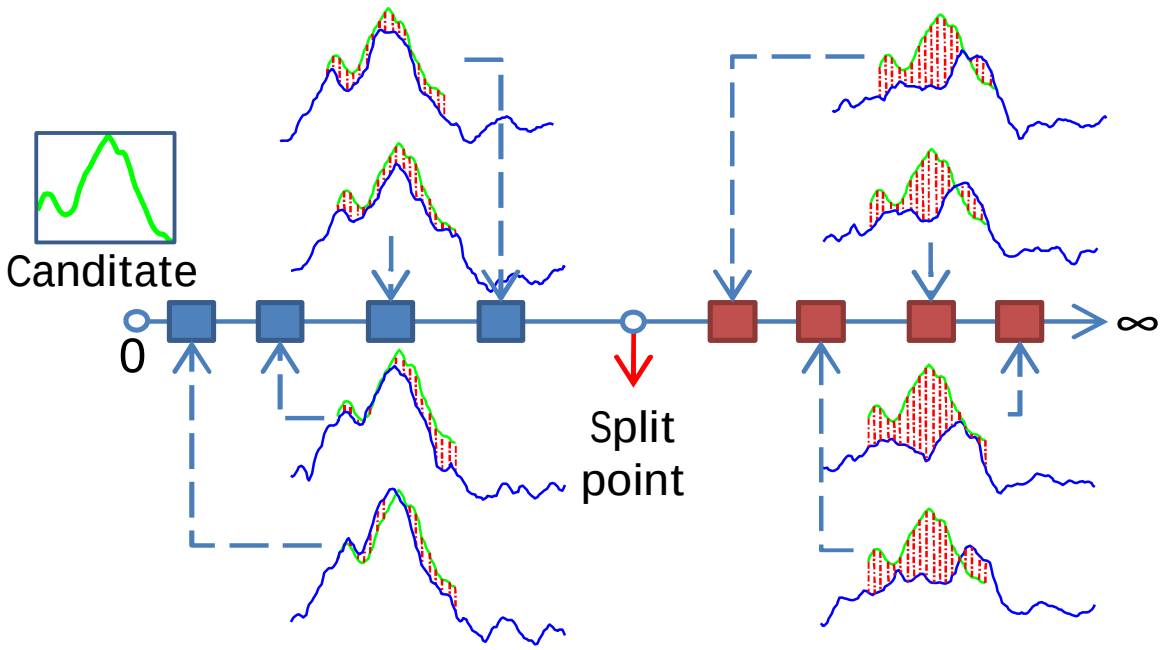

Figure 3 - 3 : Instances are sorted based on the 1NN distance (sdist) from the candidate. A split point is found to group the instances maximizing the information gain.

A distance computation between a candidate and an instance time series may take $O(m^2)$ time in the worst case and for all the instances it can take up to $O(nm^2)$ time. Once we know the distances, the computation of the information gain takes a linear scan to try different split points. Therefore, the candidate evaluation process is dominated by the $O(nm^2)$ complexity for the subsequence distance computation. Since there are at least $n\frac{m(m+1)}{2}$ shapelet candidates in the dataset, total number of all candidates in the D is $O(nm^2)$. Being brute force, the algorithm *generates* and *evaluates* all the candidates and thus, needs $O(n^2m^4)$ running time. Such computational cost makes the brute-force algorithm infeasible for long time series.

The state-of-the-art algorithm reduces the candidate generation phase to a simple linear scan with random-projection (Rakthanmanon, T. and Keogh, E., 2013) and thus, the overall complexity reduces to $O(nm^2)$.

3.5. Speed-up Techniques

In order to speed up the state of the art, we propose two techniques. First, we improve the candidate evaluation phase by variable length indexing. Second, we reduce the computation in the candidate generation part by dynamic stepping.

3.5.1. Multi-length Indexing

The candidate evaluation process needs $O(nm^2)$ time to compute the distances between the candidate and the instances of the data. A distance between a candidate and an instance is essentially identical to finding the nearest neighbor of the candidate among the subsequences of the instance. There have been numerous research studies on how to use indexing techniques to efficiently search for the nearest neighbor of a time series (Christos, F. et al., 1994; Jin S. and Eamonn, K., 2008). Yet indexing is not adopted for shapelet discovery because it requires querying nearest neighbors for different lengths of time series. There is no efficient technique to build a data structure that can serve multiple length queries in runtime and therefore, the cost of building indexes for every length becomes unrealistic. In this chapter, we introduce a very efficient multi-length indexing scheme that can be used across queries of multiple lengths. Multi-length index reduces the number of distance computation significantly from the state of the art. We start explaining the technique with the description of a very simple indexing structure, *order-line*.

3.5.1.1. Order-Line

An order-line is a sorted sequence of Euclidean distances of the subsequences of an instance from a random pivot point R. In the Figure 3 - 4, we show a schematic view. The idea behind choosing a pivot point, R and projecting the objects (time series in our case) to a 1-D line has already been used by a plethora of dimensionality reduction algorithms (Faloutsos, C and Lin, K.-I., 1995) and specifically for searching in time series data. This ordering of the objects about R provides us with some useful heuristic information to find minimum distance for a candidate shapelet. The intuition is that if two objects are close in the original space, they must also be close in this ordering or two objects can be arbitrarily close in the linear ordering but very far apart in the original space. More

precisely, the ordering has two very useful properties. First, we can use triangular inequality (centered at R) to produce a lower bound for the distance between any pair of points. Second, there is an implicit order of the objects based on lower bounds. For example, the lower bound between Q and C (i.e. $\sqrt{25} - \sqrt{18}$) is more than the lower bound between Q and D (i.e. $\sqrt{18} - \sqrt{10}$). The order lines for each instance time series can be calculated during the first candidate evaluation assuming the candidate as the pivot point and thus, it requires no additional cost.

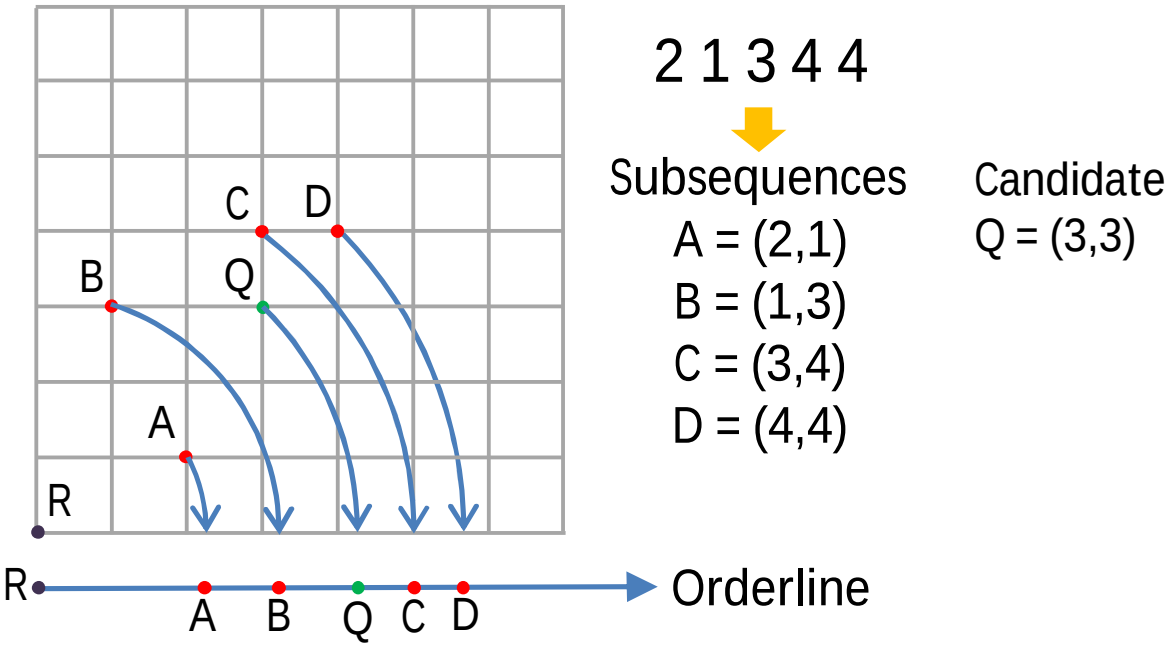

Figure 3 - 4 : Projection of subsequences of a toy time series 2 1 3 4 4 to one-dimensional ordering. Subsequences are of length two. The pivot point is at the origin for simplicity. A candidate shapelet is Q.

3.5.1.2. Finding the Minimum Distance

Given the order-line, how can we use this as an index to find the nearest neighbors quickly? The generic idea can be found in (Faloutsos, C and Lin, K.-I., 1995; Yasushi, S. et al., 2005).

Given a query, the algorithm first computes the one dimensional projection by taking a distance from R. To give an example, let's assume the projection value of the query be Q in the Figure 3 - 4. The algorithm will check points starting with the closest one in the line with an increasing order

of distances from Q. In this example, the algorithm will check C, B, D and A, in this order. For each point, the algorithm computes the Euclidean distance with early abandoning and checks if the point is better than the closest discovered so far. If we encounter a point that has a distance less than the current best-so-far, we update the best-so-far. At any time, if the distance in one dimensional space (i.e. on the line) is larger than the best-so-far distance in original space, we stop checking (break point) as we have found the nearest neighbor exactly. Table 3 - 2 shows the steps for finding nearest neighbor of Q in the high dimensional space with some hypothetical distance values.

Table 3 - 2 : Progress step of scanning

	Dist	**1D Dist**	**MD Dist**	**bsf**	**1D dist>bsf**
1	Q-C	$\sqrt{25} - \sqrt{18}$	1	1	No
2	Q-B	$\sqrt{18} - \sqrt{10}$	2	1	Yes
3	Q-D	$\sqrt{32} - \sqrt{18}$	$\sqrt{2}$	1	Yes
4	Q-A	$\sqrt{18} - \sqrt{5}$	$\sqrt{5}$	1	Yes

3.5.1.3. Bounding Distances for Longer Queries

In an order-line we have the lower-bounds of the distances between a candidate and the objects based on triangular inequality. The assumption is that the objects, R and the candidate are all in the same dimensionality or length. However, shapelet candidates can be of different lengths and for each of them, we will need its nearest neighbor distances to the instance time series as shown in Figure 3 - 4. Let's assume that we have an order-line for dimensionality *m* where the points and the reference point R, are in an *m*-dimensional space. If we have a new candidate with a larger length (i.e. more dimensionality) of *m+1*, how are we going to use the order-line to quickly find the nearest neighbor?

We take the first *m*-dimensions of the query and compute its position on the order-line. Let's assume, without losing the generality, that the *m*-prefix of the candidate is the object A in the

Figure 3 - 4. We cannot simply say, as before, the lower bound between the subsequence D and the query Q in the *m+1* dimensional space is $\sqrt{18} - \sqrt{10}$. The reason is that the normalized distance between D and A in the *m+1* dimensional space does not increase or decrease monotonically with their normalized prefix distance in the *m*-dimensional or lower dimensional space. In a recent work (Abdullah, M., 2013) it has been shown that it is possible to lower bound distances upon extension. More precisely, if *x* and *y* are two time series of length *m* and x_{+1} and y_{+1} are the two one step extensions of them, respectively, then

$d_{LB}^2(\hat{\boldsymbol{x}}_{+1}, \hat{\boldsymbol{y}}_{+1}) = \frac{1}{\sigma_m^2} d^2(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ where $\sigma_m^2 = \frac{m}{m+1} + \frac{m}{(m+1)^2} z^2$ and z = max(*abs*($\hat{\boldsymbol{X}}$), *abs*($\hat{\boldsymbol{Y}}$))

In the above equations, the *X* and *Y* are the original sequences while *x* and *y* are subsequences. The hat operator is describing the normalization operation. The variable *z* represents the highest possible value that can appear next. The bound is trivially trued for raw vectors without normalization.

This result makes it possible to use one order-line for the next query length. In our running example, we know the lower bound distance between Q and D in the m-dimensional space and we can find a lower bound in the *m+1* dimensional space by multiplying 2 with a fraction $1/\sigma_m^2$. Thus the algorithm to find the nearest neighbor can continue until this lower bound is larger than the *best-so-far*.

Note that as we increase the level of extension such as *s*-step extension for *s* > 1, the lower bounding fraction can be repeatedly applied. For example, for 2-step extension, the lower bound can be

$$d_{LB}^2(\hat{\boldsymbol{x}}_{+2}, \hat{\boldsymbol{y}}_{+2}) < \frac{1}{\sigma_{m+1}^2} d_{LB}^2(\hat{\boldsymbol{x}}_{+1}, \hat{\boldsymbol{y}}_{+1}) = \frac{1}{\sigma_{m+1}^2} \frac{1}{\sigma_m^2} d^2(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) < \frac{1}{\sigma_m^4} d^2(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$$

Thus, for queries of length *m+s*, we can use the *m*-prefix of the query to find the spot in the *order-line*. Once the spot is found, the lower bounds will be the distances on the *order-line* multiplied by the fraction as computed above.

Example: Let's take the example for lower bounding trick for longer queries without z-normalization. Let us assume a 3D candidate Q_1=(3,3,3) has a 2D prefix Q. The distance between Q and A is $\sqrt{5}$=2.2361 in Figure 3 - 4. The largest coordinate value *z*=4. Then $\sigma_m^2 = \frac{2}{3} + \frac{2*16}{9} = 4.2156$ and the lower bound between Q_1 and A=(2,1,3) is $\frac{\sqrt{5}}{\sqrt{4.2156}} = 1.089$.

The complete algorithm for finding the nearest neighbor using our multi-length indexing scheme is given in the Table 3 - 3.

Table 3 - 3 : Finding Minimum Distance Algorithm

Algorithm 2 *findMinDist*(D,*candidate*,*OrderLine*)

Require:D : A dataset of time series
candidate : Shapelet candidate
OrderLine : Distances of the subsequences toa reference
Ensure: Return the most similar subsequence to the candidate
1: *m*=dimensionality of the *OrderLine*
2: *S*=*length*(*candidate*)-*m*, *1D*=0, *bsf*=∞
3: Find the location *l* of the *m-prefix* of the *candidate*
in the *OrderLine*
4: $f_c = 1/\sigma_m^{m*S}$
5: **for each** D_i in the OrderLine in order of distances from *l*
6: *1D* = (distance_in_Orderline between *l* and D_i) * f_c
7: **if** *bsf* > *1D* **then** *bsf* = *EuclideanDistance*(*candidate*,D_i)
8: **if** *1D* > *bsf* **then return** (*bsf*)

For FECG data set, results showed that multi-length indexing (MLI) sped up 2.37 times the overall computation time which outperforms the *Fast Shapelet* algorithm. See Figure 3 - 5.

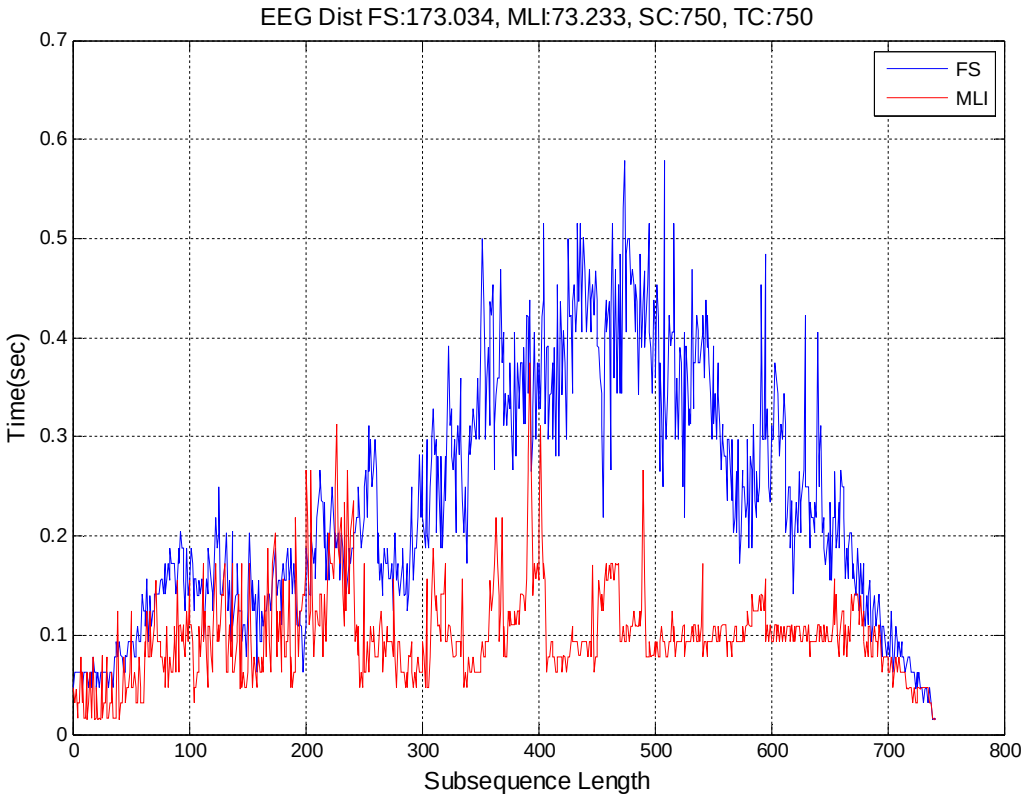

Figure 3 - 5 : Execution time of distance computation for different subsequence length, FS(blue) and MLI(red). Total running time of the FS: 173.034 and MLI: 73.233.

Even though, we were able to achieve a good speed up, performance of MLI depends on the characteristic of the data. Thus performance of MLI is not the same for all data sets. Figure 3 - 6 shows execution time of Fast Shapelet and MLI algorithms for Beef data set (Keogh, E. et al., 2012) which has 5 classes. The accuracy of our algorithm is 55.17% and Fast Shapelet achieved 55.83% accuracy for this dataset while our algorithm is slightly faster (1.2x).

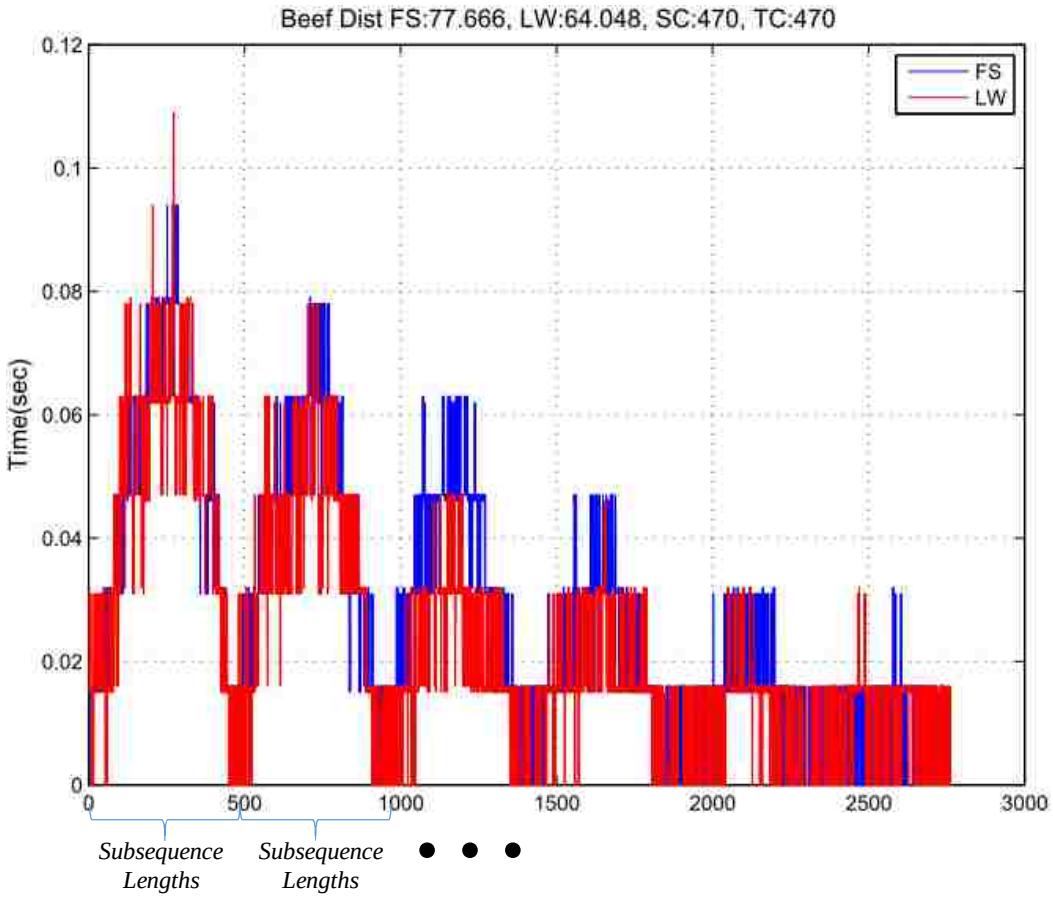

Figure 3 - 6 : *Execution time of distance computation for different subsequence length, FS(blue) and MLI(red). Total running time of the FS: 77.67 and MLI: 64.048.*

In conclusion, multi-length indexing can reduce the computation time. However it is not sufficient for the scale of the dataset we have. These results showed us that current shapelet algorithm still requires more improvements for execution time without sacrificing accuracy.

3.5.2. Dynamic Stepping

Multi-length indexing can reduce the computation time reasonably. We further improvise the algorithm with a very simple strategy without any significant accuracy degradation. Before we describe the technique, we introduce the complete mc^2 algorithm in Table 3 - 4.

The algorithm takes a time series dataset as input and three user defined parameters namely the minimum and maximum shapelet length and, the *usr* described later. The algorithm runs a loop over the possible lengths of the candidates in line 3 which is the outer-most loop. The algorithm generates a set of ten candidates for every length using the random projection technique described in the *FastShapelet* method (Rakthanmanon, T. and Keogh, E., 2013) in lines 4-7. The loop at line 8 goes over each of the ten candidates to find the best one.

Our dynamic stepping algorithm uses a certain step size for the outer-most loop running over the lengths. The step size can be constant or variable. Constant step size has a disproportionate effect on shorter shapelets. For example, a step size of 2 may generate as low as 0.5 overlap between successive candidates starting at a certain location. To have a uniform effect, we vary stepping strategy.

Before we describe the strategy, we clarify that stepping is not the same as sampling the time series down to smaller size. If we down-sample, the dimensionality of the distances would change. In dynamic stepping, we skip some of the subsequences in systematic manner so the impact on accuracy is uniform.

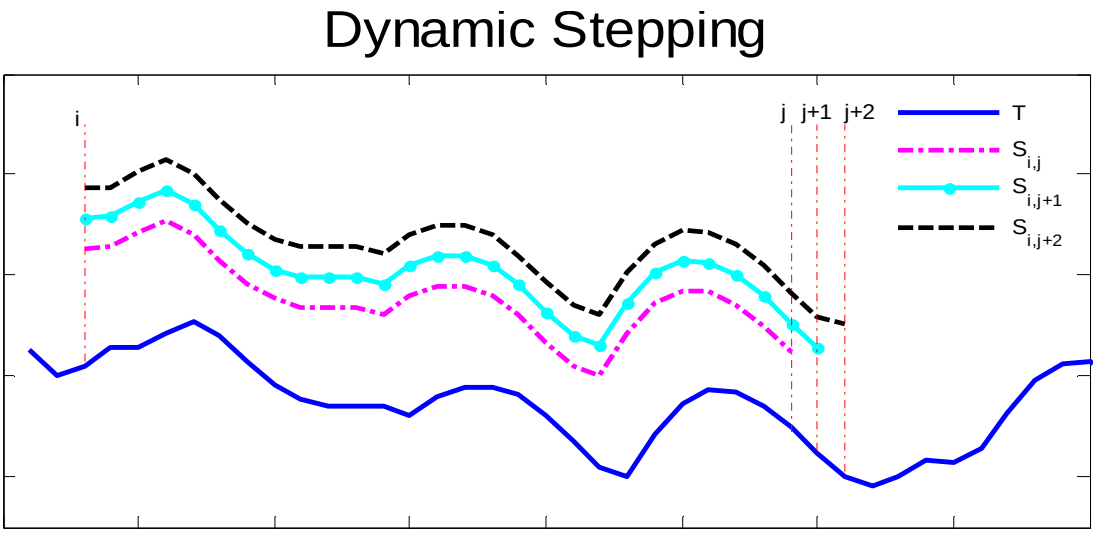

Figure 3 - 7 : Shapelet candidates of three arbitrary subsequence length j, j+1, j+2 for i.

Recall, we define a time series as a sequence of real numbers T = $t_1, t_2, \ldots, t_m$ and the subsequence of length j as the $S_{i,j}=t_i,t_{i+1},\ldots t_{i+j-1}$. It has been well observed that there is a large overlap between $S_{i,j}$ and $S_{i,j+1}$ for any i making them very similar to each other (i.e. trivial matches). See Figure 3 - 7. We define the ratio between two successive candidates starting at the same position as similarity ratio (*sr*) and use it to determine the size of the stepping. For a step size (*sz*), the similarity ratio is simply *j/(j+sz)* for any *j* within the minimum and maximum lengths specified by the user.

Table 3 - 4 : mc^2 *Shapelet_Discovery Algorithm*

Algorithm *mc²* :*Shapelet_Discovery*(D,*min_length*,*max_length*,*usr*)

Require	**:** A dataset D of time series
Ensure	**:** Return the shapelet
1: *maxGain* = 0, *maxGap* = 0, *k* = 1, *stp* = 1	
2: [*qsum*,*qsum2*] = *cumulativeSums*(D,min_length,max_length)	
3: **for** *sbsqlen* = *min_length* : *stp* : *max_length* **do**	
{for every subsequence length at steps}	
4:	*SAXList* = *CreateSAXList*(D,*sbsqlen*)
5:	*RandomProjection*(*SAXList*)
6:	*ScoreAllSAX*
7:	*top_10_cand* = *FindBestSAX*(10), *max_gain* = ∞, *min_gap* = 0
8:	**for** *i*=1:10
9:	*cand* = *top_10_cand*(*i*)
10:	**if** *i* == 1 {----------Multi Length Indexing ----------}
11:	$f_c = 1/\sigma_m^{m*S}$
12:	**if** (f_c < 0.75 **or** f_c ==1) **then**
	OL=Linear_ordering(D,*cand*), f_c=1
13:	*Q* = *EuclideanDistance*(*cand*,*referance_query*)
14:	**for** *i* = 1 **to** |D|

15: **if** (Q == 0 **and** f_c == 1) **then** L = min(OL)

16: **else** L = *findMinDist*(D,*cand*,*OL*,f_c)

{----------Multi Length Indexing----------}

17: [*gain*,*gap*] = *CalInfoGain*(*L*)

18: **if** (*gain* > *max_gain*) **or**

((*gain*== *max_gain*) **and** (*gap*>*min_gap*))

then

19: *max_gain* = *gain*, *min_gap* = *gap*,

Bestshapelet = *cand*, *best*τ = τ

20: **If** *sr* ≥ *usr* **then** *stp* = ⌊*j*(1-*usr*)/(2*usr*-1)⌋ {-Dynamic Stepping-}

The initial step size is 1 and it increases by one whenever it crosses a user defined maximum threshold (*usr*). See Figure 3 - 8. As long as *sr* < *usr*, we use the same step size. When *sr* ≥ *usr* holds, we increase step size. For example, *m=113*, *i=0*, *j=10*, *sz=1* and *usr=0.95* then *sr=j/(j+sz)=0.9091* and *sr* ≥ *usr* does not hold. Therefore we use subsequence length *j=j+sz* to create new shapelet candidates and keep on increasing *j* with *sz* until *sr* ≥ *usr.* When *j =19, sr=0.95* then we can increase *sz.* This operation can be easily implemented by setting *sz*= ⌊*j(1-usr)/(2usr-1)*⌋.

Without any skipping, the number of different lengths is *m-minLength+1* which is the number of iteration for a time series of length *m* in the original algorithm. By using skipping technique we can reduce the number of iteration. We can define optimization rate as *(m-minlng+1)* / *(total iterations in our algorithm)*.

usr can be thought of as a control for the speedup-accuracy tradeoff. Using a lower *usr* value increases optimization rate but using very low *usr* value may cause missing the best shapelet candidate, especially for short subsequence lengths. Also dynamic stepping gives better optimization rate for longer subsequence length because *sr* needs less number of iteration to satisfy *sr* ≥ *usr*.

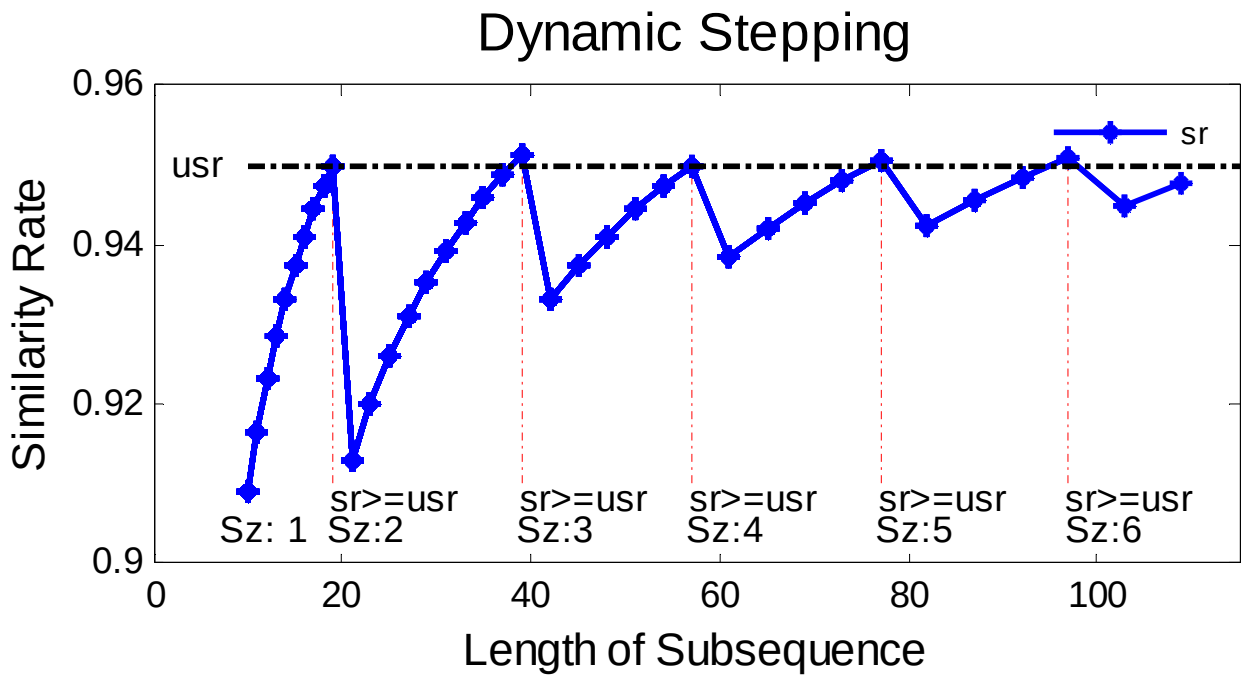

Figure 3 - 8: Increasing step size for short subsequences and more for long subsequences for maximum optimization

3.6. Voting Based Ensembling

With the speedup techniques described above, we could find a set of shapelet trees for each of the dimensions of a multidimensional time series.

Dietterich TG. *et al.* (Dietterich TG. and Kong EB., 1995) have shown that teaming up a set of decision trees based on majority voting can reduce the machine learning bias of the classifier. We employ the same technique with an assumption that individual dimensions are independent. Thus counting votes from the classifiers is the way to generalize.

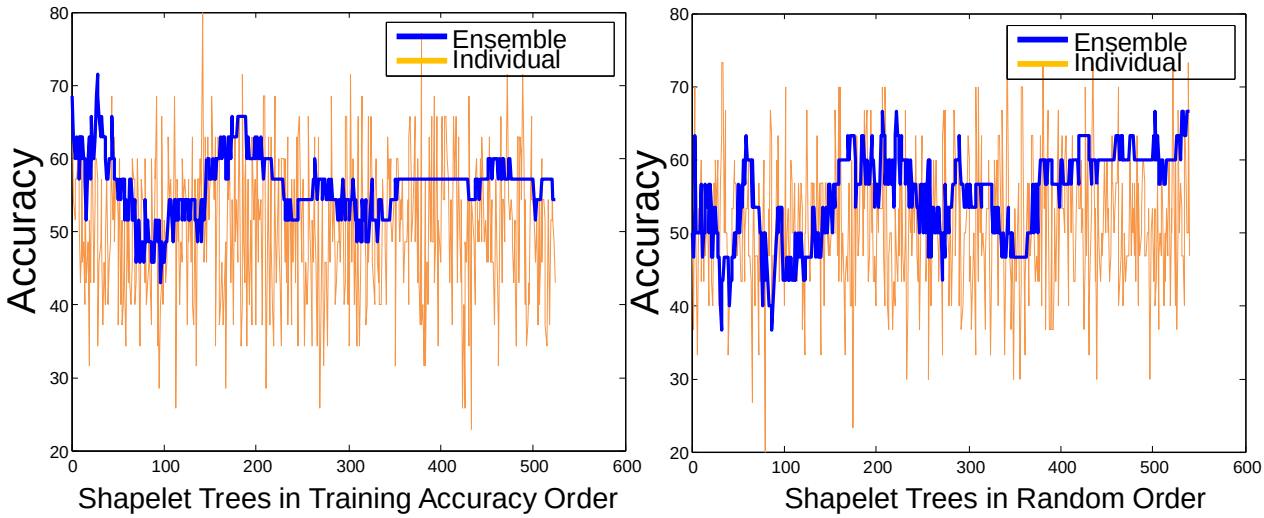

Figure 3 - 9 : Accuracy of all-ensemble method is higher on average.

An important question is "*Should we ensemble all the trees?"* As shown in Figure 3 - 2, if we ensemble all of them we converge to a lower accuracy than what could be achieved by a smaller number of trees. We experiment to test if the training accuracies can guide us to find the smallest ensemble of highest accuracy. We find there is *no strong correlation* between training accuracy of individual trees with the overall accuracy of the ensemble. Therefore, we use *random ordering* of the shapelet trees and ensemble *all* trees for an average gain over individual trees. Figure 3 - 9 shows two different orderings, where the left one achieves a tiny gain on average and misses the maximum gain. The right one achieves the maximum gain.

Other than multi-dimensional data, we can also use ensembling for one-dimensional data. The random projection part for candidate generation introduces randomness in the accuracy of the decision tree unlike the exact methods. We can run the same algorithm on the same data to generate a set of trees. We investigate the accuracy for various datasets before and after we combine the classifiers via voting. We spend roughly the same amount of time to generate decision tree(s) by both of the algorithms. Since ours is faster, it generates more trees than the state of the art algorithm. The accuracies in these two sets show a significant difference just because of the number of trees. We claim that our algorithm can generate confident classifiers than the state-of-the-art shapelet discovery algorithm by having lower variance.

There exists the method called *shapelet transform*, which finds k-shapelets in the first round to build a forest of single node trees. Shapelet transform has been shown to have larger accuracies than the exact method and it is faster because of no recursive decision tree building. However, our method is different from shapelet transform in that we produce multi-node trees, which are complete classifiers. Such trees retain the original motivation of shapelets, interpretability, while shapelet transform uses many shapelet candidates to transform the data to a new feature space, which makes it difficult to interpret in the original space.

3.6.1. Complexity of the Algorithm

Multi-length indexing reduces the time to search for nearest neighbor. Because of curse of dimensionality the worst case complexity is still $O(nm^2)$. On average, the complexity becomes better with large data making the method more scalable. Dynamic stepping does not reduce the worst case complexity either, although it speeds up the algorithm by a constant factor.

3.7. Experimental Results

In order to evaluate the performance of our algorithm, we present the results of our experiments and compare them with the current state of the art (*FastShapelet*) algorithm (a heuristic algorithm) (Lexiang, Y. and Keogh, E., 2009) and original shapelet algorithm (Lexiang, Y. and Keogh, E., 2009). We use the code from the authors' webpage (Lexiang, Y. and Keogh, E., 2009; Rakthanmanon, T. and Keogh, E., 2013) and run these algorithms and our algorithm on the same device. We use the same parameters as suggested to have maximum accuracy. All of the code and the datasets used in this chapter are available at (Cetin, MS., et al., 2015a).

3.7.1. Time Series Datasets

In this chapter, we introduce 2 new time series datasets collected at the Mind Research Network at University of New Mexico. The first dataset is from a pre-attentional sensory processing experiment using fMRI scans of 32 schizophrenia patients (SP) and 43 healthy controls (HC). The second dataset is from an experiment on multi-modal sensory integration for fMRI scan on the

same set of patients and control. To obtain the time series of components for first and second data sets, we use the GIFT Toolbox (http://mialab.mrn.org/software/gift/) and infomax algorithm (Bell AJ and Sejnowski TJ, 1995) for group independent component analysis (Calhoun VD and Adali T, 2012). We perform classification on both of these datasets to classify schizophrenia patients and healthy control.

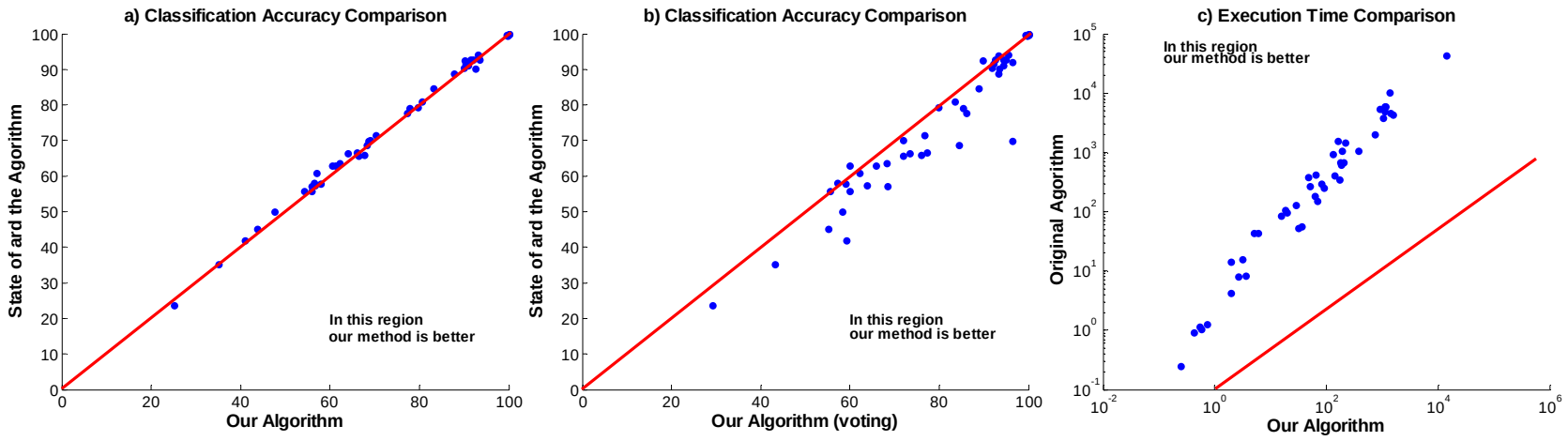

Figure 3 - 10 : *(a) Accuracy comparison between our algorithm and the current state-of-the-art algorithm (b) Accuracy comparison when we use voting based ensemble (c) Execution time comparison between our algorithm and the current state of the algorithm.*

Table 3 - 5 : Accuracy and execution time comparison between our algorithm and the current state-of-the-art algorithm

					mc^2				State-of-the-art algorithm			Speed Up
Data Sets	**Classes**	**Train**	**Test**	**TC Length**	**Execution Time**	**Accuracy**	**Accuracy (voting)**	**StDev**	**Execution Time**	**Accuracy**	**StDev**	
Beef	5	30	30	470	18.48	54.17	60	6.74	106.27	55.83	4.82	5.75
CBF	3	30	900	128	3.56	93.12	95.55	1.7	8.35	94.12	1.25	2.35
ChlorineConcentration	3	467	3840	166	183.53	57.83	58.95	1.38	620.79	57.81	1.22	3.38
CinC_ECG_torso	4	40	1380	1639	190.83	56.46	57.17	4.19	1053.3	58.18	2.58	5.52
Coffee	2	28	28	286	3.16	90.71	96.43	2.43	15.55	91.96	3.04	4.92
Cricket_X	12	390	390	300	1122.33	43.71	55.13	3.01	4941.33	45.08	4.28	4.40
Cricket_Y	12	390	390	300	1399.22	47.65	58.21	3.08	4596.29	49.95	3.49	3.28
Cricket_Z	12	390	390	300	1035.46	40.95	59.23	2.99	3798.16	41.85	2.56	3.67
DiatomSizeReduction	4	16	306	345	1.97	92.45	93.46	1.59	14.34	90.21	2.46	7.28
ECG2	2	92	89	750	51.09	100	100	0	268.23	100	0	5.25
ECG3	3	92	89	750	177.78	100	100	0	688.22	100	0	3.87
ECGFiveDays	2	23	861	136	1.95	99.82	99.42	0.4	4.22	99.61	0.23	2.16
FaceAll	14	560	1690	131	377.21	62.21	68.22	1.31	1051.71	63.57	0.92	2.79
FaceFour	4	24	88	350	15.23	90.85	94.32	2.91	86.19	91.19	1.6	5.66
FacesUCR	14	200	2050	131	91.82	68.19	84.39	3.71	249.75	68.68	1.96	2.72
fish	7	175	175	463	63.6	80.51	83.43	2.28	415.76	80.91	1.37	6.54
Gun_Point	2	50	150	150	2.62	93.33	93.33	0.3	8.07	93.77	1.64	3.08
Haptics	5	155	308	1092	214.63	35.13	43.18	2.36	1464.48	35.18	3.69	6.82
InlineSkate	7	100	550	1882	1556.7	25.16	29.27	2.41	4251.6	23.57	2.81	2.73
ItalyPowerDemand	2	67	1029	24	0.25	90.27	92.23	3.07	0.25	91.64	2.21	1.00
Lighting2	2	60	61	637	201.67	55.98	55.45	4.77	683.85	55.66	5.22	3.39
Lighting7	7	70	73	319	82.56	55.89	68.49	5.12	292.87	57.05	3.65	3.55
MALLAT	8	55	2345	1024	47.61	87.83	93.22	5.08	384.64	88.86	2.76	8.08
MedicalImages	10	381	760	99	68.71	57.06	62.24	2.06	153.1	60.86	1.67	2.23
MoteStrain	2	20	1252	84	0.53	79.69	79.79	0.46	1.16	79.27	0.92	2.19
OSULeaf	6	200	242	427	163.16	66.03	77.28	3.03	1563.46	66.69	3.24	9.58
OliveOil	4	30	30	570	5.93	70.33	76.67	3.4	43.2	71.5	4.65	7.28
Sony	2	601	20	70	31.91	93.5	95	2.86	53.62	92.75	3.43	1.68
SonyAIBORobotSurface	2	20	601	70	0.57	68.55	96.34	0	1.04	69.82	5.69	1.82
SonyAIBORobotSurfaceII	2	27	953	65	0.74	77.84	85.41	4.03	1.26	79.21	1.15	1.70
StarLightCurves	3	1000	8236	1024	1344.15	89.93	91.88	1	10068.04	90.43	1	7.49
SwedishLeaf	15	500	625	128	141.5	77.26	86.08	2.5	404.17	77.7	2.18	2.86
Symbols	6	25	995	398	5.01	83.07	88.94	3.9	43.31	84.63	2.66	8.64
synthetic_control	6	300	300	60	36.66	91.52	94.33	1.52	56.05	92.65	1.63	1.53
Trace	4	100	100	275	62.3	99.6	100	0.6	184.53	99.6	0.82	2.96
TwoLeadECG	2	23	1139	82	0.42	91.9	92.45	2.27	0.91	92.74	1.03	2.17
Two_Patterns	4	1000	4000	128	739.34	90.17	89.85	5.19	1980.66	92.39	1.37	2.68
uWaveGestureLibrary_X	8	896	3582	315	921.62	63.96	73.45	1.23	5460.62	66.35	1.12	5.93
uWaveGestureLibrary_Y	8	896	3582	315	1148.54	56.63	63.79	1.6	5934.02	57.36	1.22	5.17
uWaveGestureLibrary_Z	8	896	3582	315	1128.92	61.28	65.8	1.81	5929.55	62.82	1.63	5.25
wafer	2	1000	6000	152	171.93	99.67	99.75	0.25	348.93	99.51	0.34	2.03
yoga	2	300	3000	426	131.31	68.88	71.97	1.94	923.36	69.99	1.78	7.03
fMRI Singe Task	2	50	25	145	19.89	67.6	76	8.5	125.31	62	6.68	6.30
fMRI Multi Task	2	50	25	411	28.72	66.32	72	10.61	239	65.8	7.62	8.32
MEG Rest	2	60	31	14000	13991.2	60.5	60	5.42	43280.08	63	6	3.09

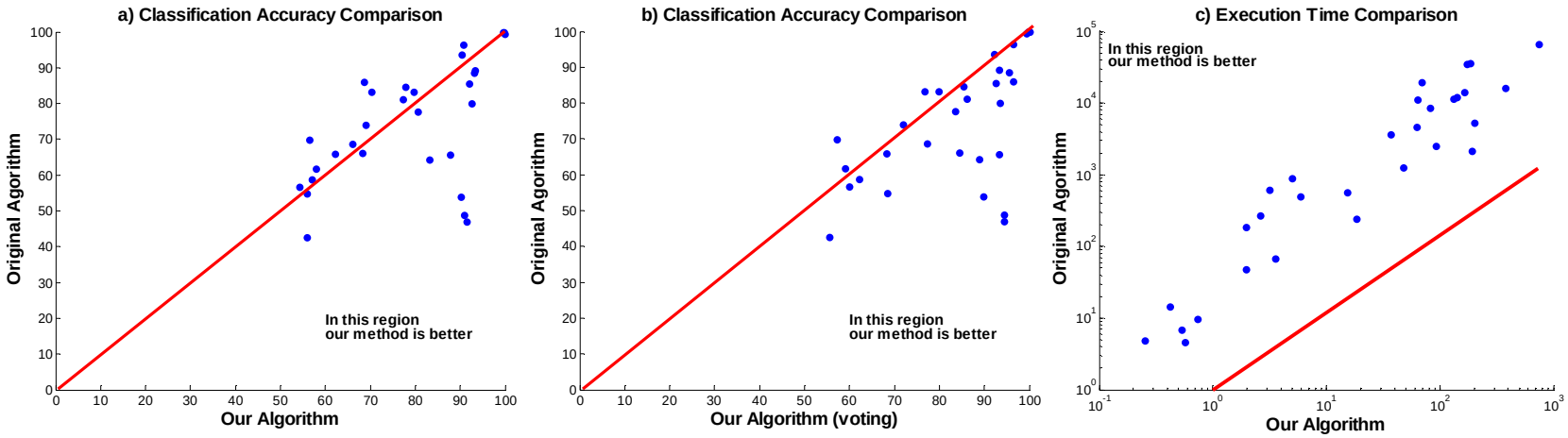

Figure 3 - 11 : (a) Accuracy comparison between our algorithm and the original algorithm (b) Accuracy comparison when we use voting based ensemble (c) Execution time comparison between our algorithm and the original algorithm.

Table 3 - 6 : Accuracy and execution time comparison between our algorithm and original algorithm

					mc^2				Original algorithm			Speed Up
Data Sets	Classes	Train	Test	TC Length	Execution Time	Accuracy	Accuracy(voting)	StDev	Execution Time	Accuracy	StDev	
Beef	5	30	30	470	18.48	54.17	60	6.74	242.29	56.67	0	13.11
CBF	3	30	900	128	3.56	93.12	95.55	1.7	66.86	88.56	0	18.78
ChlorineConcentration	3	467	3840	166	183.53	57.83	58.95	1.38	36402.28	61.85	0	198.35
CinC_ECG_torso	4	40	1380	1639	190.83	56.46	57.17	4.19	2149.95	69.86	0	11.27
Coffee	2	28	28	286	3.16	90.71	96.43	2.43	621.91	96.43	0	196.81
Cricket_X	12	390	390	300	1122.33	43.71	55.13	3.01	Not completed			
Cricket_Y	12	390	390	300	1399.22	47.65	58.21	3.08	Not completed			
Cricket_Z	12	390	390	300	1035.46	40.95	59.23	2.99	Not completed			
DiatomSizeReduction	4	16	306	345	1.97	92.45	93.46	1.59	184.34	80.07	0	93.57
ECG2	2	92	89	750	51.09	100	100	0	Not completed			
ECG3	3	92	89	750	177.78	100	100	0	Not completed			
ECGFiveDays	2	23	861	136	1.95	99.82	99.42	0.4	47.64	99.42	0	24.43
FaceAll	14	560	1690	131	377.21	62.21	68.22	1.31	16255.48	65.86	0	43.09
FaceFour	4	24	88	350	15.23	90.85	94.32	2.91	561.18	48.86	0	36.85
FacesUCR	14	200	2050	131	91.82	68.19	84.39	3.71	2528.52	66.24	0	27.54
fish	7	175	175	463	63.6	80.51	83.43	2.28	11153.03	77.71	0	175.36
Gun_Point	2	50	150	150	2.62	93.33	93.33	0.3	266.1	89.33	0	101.56
Haptics	5	155	308	1092	214.63	35.13	43.18	2.36	Not completed			
InlineSkate	7	100	550	1882	1556.7	25.16	29.27	2.41	Not completed			
ItalyPowerDemand	2	67	1029	24	0.25	90.27	92.23	3.07	4.92	93.59	0	19.68
Lighting2	2	60	61	637	201.67	55.98	55.45	4.77	5297.6	42.62	0	26.27
Lighting7	7	70	73	319	82.56	55.89	68.49	5.12	8619.35	54.79	0	104.40
MALLAT	8	55	2345	1024	47.61	87.83	93.22	5.08	1254.91	65.63	0	26.36
MedicalImages	10	381	760	99	68.71	57.06	62.24	2.06	19325.2	58.68	0	281.26
MoteStrain	2	20	1252	84	0.53	79.69	79.79	0.46	6.87	83.23	0	12.96
OSULeaf	6	200	242	427	163.16	66.03	77.28	3.03	14186.53	68.6	0	86.95
OliveOil	4	30	30	570	5.93	70.33	76.67	3.4	502.27	83.33	0	84.70
Sony	2	601	20	70	31.91	93.5	95	2.86	Not completed			
SonyAIBORobotSurface	2	20	601	70	0.57	68.55	96.34	0	4.56	86.02	0	8.00
SonyAIBORobotSurfaceII	2	27	953	65	0.74	77.84	85.41	4.03	9.76	84.58	0	13.19
StarLightCurves	3	1000	8236	1024	1344.15	89.93	91.88	1	Not completed			
SwedishLeaf	15	500	625	128	141.5	77.26	86.08	2.5	11953.61	81.28	0	84.48
Symbols	6	25	995	398	5.01	83.07	88.94	3.9	894.26	64.32	0	178.50
synthetic_control	6	300	300	60	36.66	91.52	94.33	1.52	3667.43	47	0	100.04
Trace	4	100	100	275	62.3	99.6	100	0.6	4626.86	100	0	74.27
TwoLeadECG	2	23	1139	82	0.42	91.9	92.45	2.27	14.29	85.6	0	34.02
Two_Patterns	4	1000	4000	128	739.34	90.17	89.85	5.19	65783.11	53.9	0	88.98
uWaveGestureLibrary_X	8	896	3582	315	921.62	63.96	73.45	1.23	Not completed			
uWaveGestureLibrary_Y	8	896	3582	315	1148.54	56.63	63.79	1.6	Not completed			
uWaveGestureLibrary_Z	8	896	3582	315	1128.92	61.28	65.8	1.81	Not completed			
wafer	2	1000	6000	152	171.93	99.67	99.75	0.25	34653.13	99.88	0	201.55
yoga	2	300	3000	426	131.31	68.88	71.97	1.94	11388.99	74	0	86.73
fMRI Singe Task	2	50	25	145	19.89	67.6	76	8.5	Not completed			
fMRI Multi Task	2	50	25	411	28.72	66.32	72	10.61	Not completed			
MEG Rest	2	60	31	14000	13991.2	60.5	60	5.42	Not completed			

In addition to the novel datasets, we use 41 datasets from the UCR time series archive (Keogh, E. et al., 2012) to compare the efficiency of our algorithm to the state-of-the-art algorithm. See Figure 3 - 10 and Table 3 - 5. We test on all the 43 datasets on original algorithm; however, we abandoned the experiments (14 data sets) in which the original algorithm (Lexiang, Y. and Keogh, E., 2009) had not finished after 18 hours. Comparison with the original algorithm is presented in Figure 3 - 11 and Table 3 - 6. For UCR time series data sets, we used a train and test split ratio as suggested in their webpage (Keogh, E. et al., 2012). For the two new datasets, we used 66% for training (66% SP and 66% HC) and 33% for testing (33% SP and 33% HC). Subjects in the training and testing sets are chosen randomly.

3.7.2. Experimental Settings

For all the experiments, we used the same hardware (Intel(R) Core(TM) i7 CPU 860@ 2.80 GHz, 1.59GHz, 2.96 GB of RAM). The parameters, minimum length and maximum length, are set to 10 and 250, respectively (If the length of the time series is less than 250, maximum shapelet length is set to the length) and the upper bound of similarity rate (*usr*) is 0.95.

3.7.3. Performance Comparison on One Dimensional Data

We use accuracy and execution time parameters to compare performance of our algorithm. We used the average of 20 runs for all data sets and for both algorithms. More detailed results are available at (Cetin, MS., et al., 2015a).

3.7.3.1. Accuracy

We perform the experiment on the 43 datasets for *FastShapelet* (current state of the art) algorithm and 30 data sets for the original algorithm. Our speedup techniques do not target improvement in accuracy and we do not expect any significant change in accuracy due to skipping some of the lengths. Figure 3 - 10(a) shows the comparison between our algorithm and the *FastShapelet* algorithm.

We use paired t-tests on the accuracy results. The cut off P-value for all of the tests is set at $P < 0.05$. The results confirm no significant difference in accuracy.

We perform an experiment to validate the goodness of the voting based ensembling we propose. In Figure 3 - 10(b), the comparison is shown and the paired t-tests on the accuracy results showed significant increase in accuracy at $P \geq 0.05$ for *FastShapelet* algorithm and the original algorithm.

3.7.3.2. Execution time

On all of the 43 data sets, we record the running time of the algorithms and compare them in the Figure 3 - 10(c) with log scale. As claimed before, we achieved an order of magnitude (up to 9.58x) speeds up over the current state of the art algorithm and several order of magnitude speeds up (up to 281x) over the original algorithm.

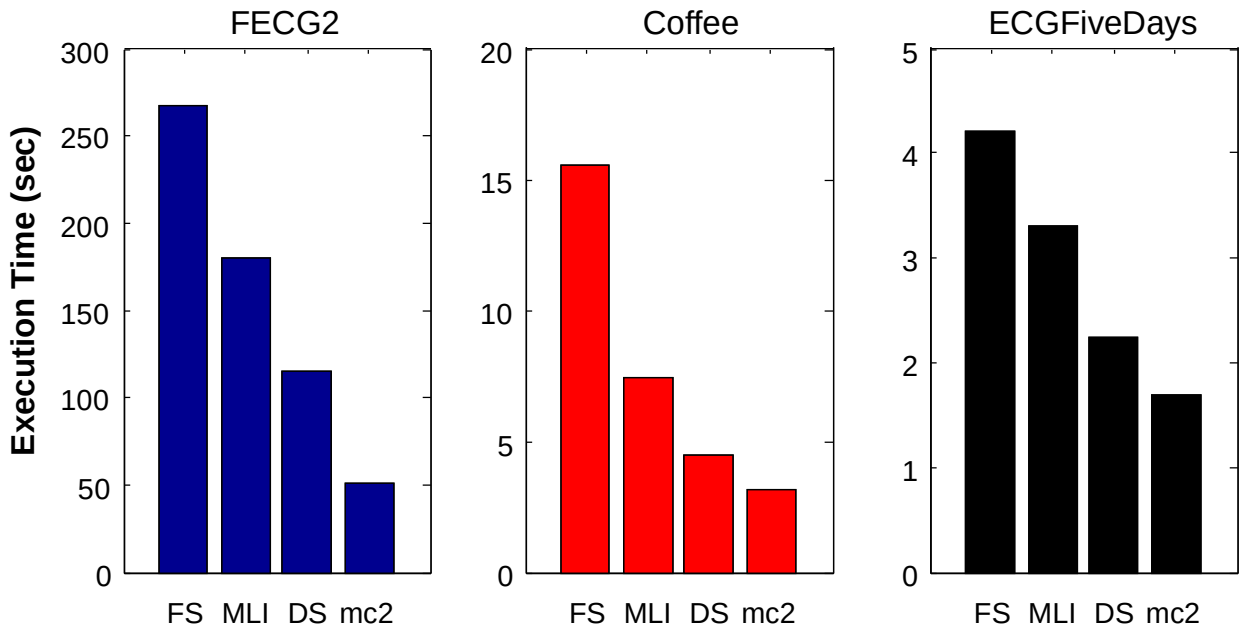

Figure 3 - 12 : Individual execution time (sec) of three different data sets for the current state of art algorithm and our algorithm with just Multi-length indexing (MLI), Dynamic Stepping (DS) and union of MLI and DS (mc2).

In addition to the comparison on total running time, we analyze the speedup achieved by the individual techniques, multi-length indexing and dynamic stepping, over the current state of the art algorithm. We measured the individual speedup for each of the techniques while deactivating the other. The results for three datasets are shown in Figure 3 - 12.

The multi-length indexing technique and the dynamic stepping technique sped up overall running time for all data sets individually. The union of these 2 techniques showed maximum speed up performance for all the data sets.

Table 3 - 7 : Accuracy comparison between concatenation based and ensemble based method.

Dataset	Concatenation	Ensembling (max individual)	Merge and Ensembling
Single-Task	66%	76% (72%)	80%
Multi-Task	64%	72% (64%)	

3.7.4. Accuracy on Multi-dimensional Data

We test our shapelet ensemble in comparison to the method described in (Abdullah, M. et al., 2011) which suggests a concatenation of the dimensions to convert a multi-dimensional data into a one-dimensional data. We use the two fMRI datasets described above. Shapelet ensemble achieves an unprecedented accuracy on these datasets. See Table 3 - 7. Note that the ensembling can achieve more accuracy than the maximum individual accuracy of the participating trees. This shows a huge potential of shapelet ensembles for sensory data from high frequency electrical sensors.

3.8. Case Studies

We have studied four datasets in the medical domains. Case 3 and Case 4 are multi-dimensional datasets where shapelet ensemble achieve higher accuracy than current state of art algorithm. We detail the cases in this section.

3.8.1. Fetal Electrocardiogram (2 sensor)

A non-invasive fetal electrocardiogram (FECG) dataset has been collected from Physionet.org (Goldberger AL et al., 2000). It contains a series of 2 multichannel abdominal FECG recordings, taken from a single subject between 21 to 40 weeks of pregnancy. The records have variable durations, and were taken weekly. The training set contains a balanced (42/50) mix of 92 time series of two abdominal channels. The testing set also contains a balanced (45/44) mix of time series. The length of each time series is 750.

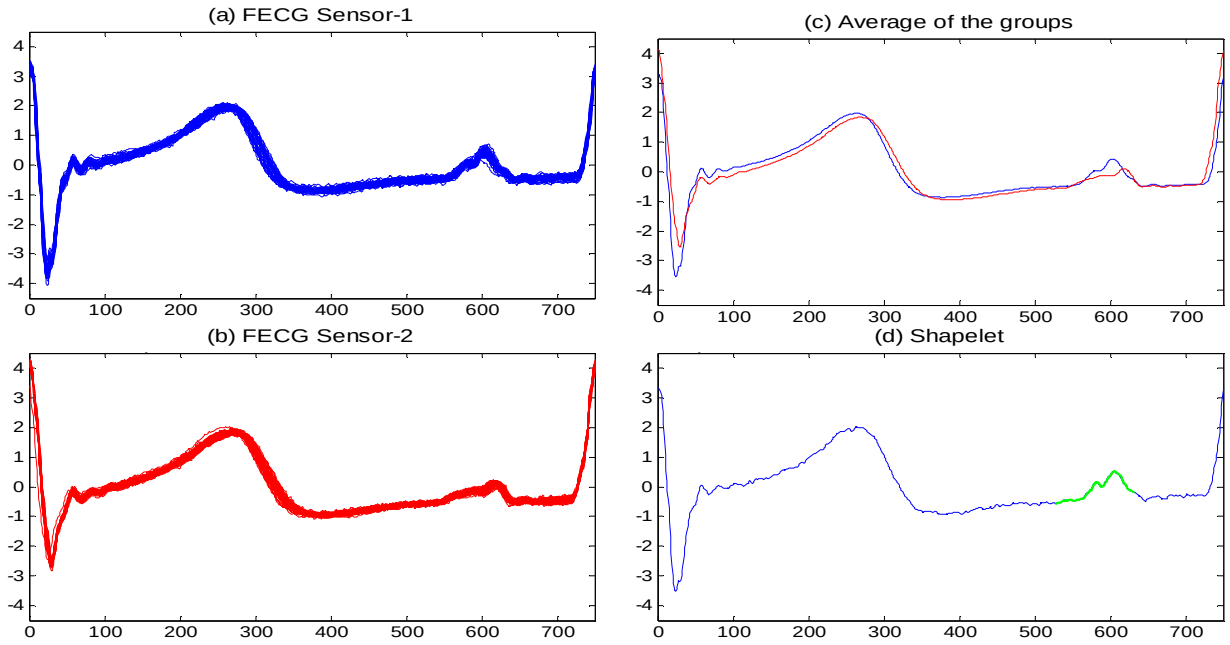

Figure 3 - 13 : (a-b) Training set of FECG data for all classes. (c) Average of the classes (d) Shapelet is shown in green. It classifies the sesnsor-1 and sesnsor-2

The accuracy of both *FastShapelet* and our algorithm for FECG data set is 100%. However for the same data set, we sped up 5.25 times the overall computation time which outperforms the *FastShapelet* algorithm. The shapelet classified the sensor-2 that is shown in green in Figure 3 - 13d.

3.8.2. Fetal Electrocardiogram (3 sensor)

Also we used 3 multichannel abdominal FECG recordings, The training set contains a balanced (45/44/46) mix of 135 time series of three abdominal channels. The testing set also contains a balanced (42/50/42) mix of 134 time series. The length of each time series is 750.

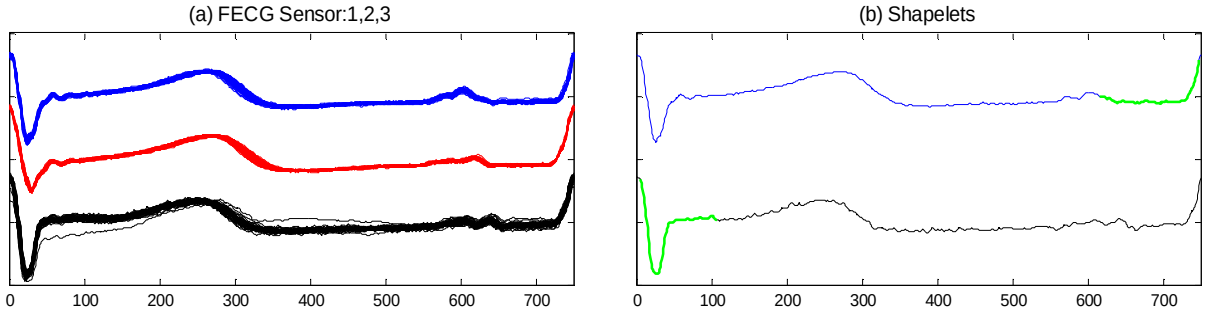

Figure 3 - 14 : (a) Training set of FECG data for all classes. (b) Shapelets are shown in green. First shapelet classifies the sesnsor-1, Second shapelet classifies the sesnsor-3.

The accuracy of both *FastShapelet* and our algorithm for FECG data set is 100%. However for the same data set, we sped up 3.87 times the overall computation time which outperforms the *FastShapelet* algorithm. The First shapelet classified the sensor-1 and the second shapelet classified the sensor-3. They are shown in green in Figure 3 - 14b.

3.8.3. Sensory Gating task for functional MRI (fMRI)

The dataset contains the results of pre-attentional sensory processing experiment for fMRI scan of the schizophrenia patients (SP=32) and healthy controls (HC=43). The task represented cognitive paradigm that was analyzed separately in (Mayer AR et al., 2012). The original study hypothesizes that the SP would have deficits at multiple levels including pre-attentive sensory processing, integration of sensory information across modalities and impaired working memory performance.

The task was a variant of the paired click paradigm for testing sensory gating. The participants were presented with paired tones that are either identical 2 kHz tones or a 2 kHz tone followed by a 3 kHz tone. The task was passive in that a response was not required.

We use the time series of the sensory motor (putamen) component. There are 50 (SP=23, HC=27) time series for training data sets and 25 (SP=9, HC=16) time series for testing data sets. The length of each time series is 145.

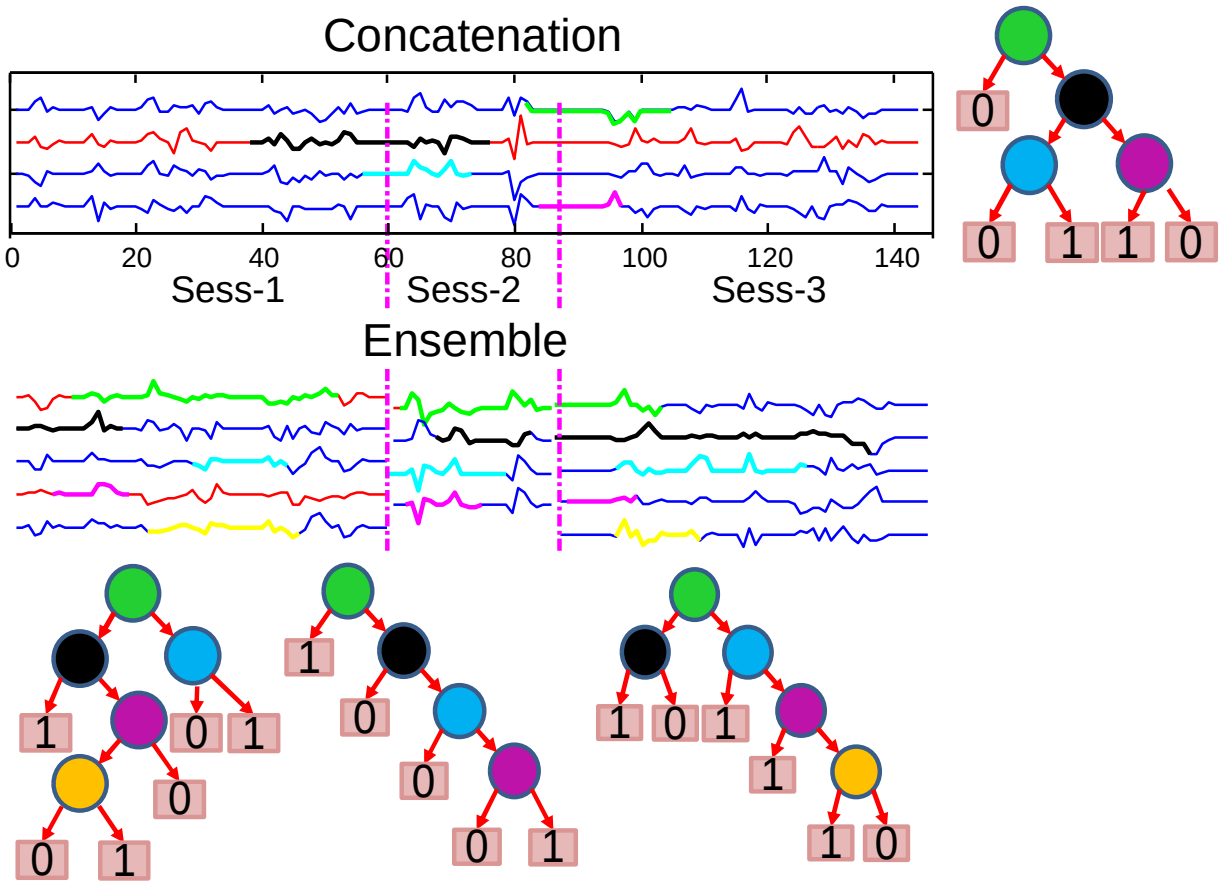

Figure 3 - 15 : The decision tree from concatenated signals is shown in the top. Three trees from the three sessions are shown in the bottom. Shapelets and nodes in the trees have matching colors. Label 0 represents controls and 1 represents patients. Signals in red are from patients and in blue are from healthy subjects (best viewed in color).

If we concatenate dimensions and run FastShapelet we achieve a shapelet tree shown in Figure 3 - 15, which gives us 66% accuracy. Note the shapelets that span two successive sessions are meaningless. When we treat each dimension independently, the algorithm provides one tree for each of the three sessions. The ensemble achieves 76% accuracy. Interestingly, no significant class difference is reported in Mayer et al. (Mayer AR et al., 2012) for sensory gating task (same data) while shapelet ensemble achieves a significant accuracy.

If we were to generate the same ensemble using FastShapelet algorithm, we would need 6.3x more time.

3.8.4. Multi-modal sensory integration task for functional MRI

The data set contains the results of multi-modal sensory integration task for fMRI scan of the schizophrenia patients (SP=32) and healthy controls (HC=43). The task represented a cognitive paradigm that was analyzed separately in (Stone DB et al., 2011).

Task was designed to test multisensory integration and employed a simple forced choice behavioral task. A perspective digital drawing was used as a visual background with a fixation point, and participants were instructed to maintain fixation throughout the task (Stone DB et al., 2011). An auditory stimulus (500 Hz tone) was presented to participants at two different volumes, 80dB to simulate a sound near the participant (NEAR) and 64dB to simulate it being farther away (FAR). Also, the auditory stimulus was presented synchronously with a visual stimulus - an image of a soccer ball appeared in one of two possible positions (NEAR or FAR) in the participant's lower visual field with size and position consistent with the perspective drawing. The NEAR visual stimulus was presented in the participant's peripheral visual field and the FAR stimulus was presented closer to fixation in the central visual field. The participants underwent fMRI scanning while deciding whether the stimuli presented were NEAR or FAR with a button press.

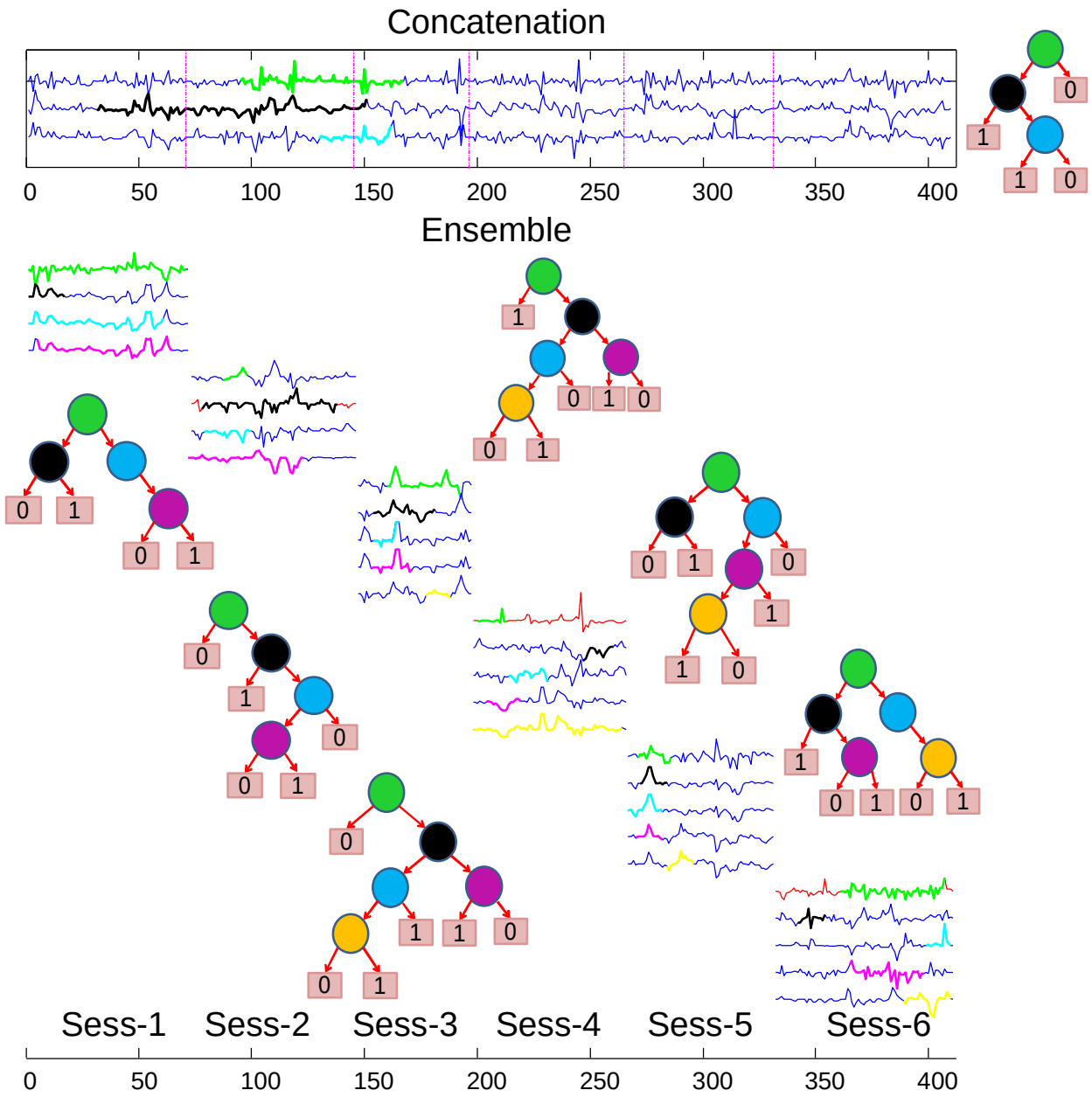

Figure 3 - 16 : The decision tree from concatenated signals is shown in the top. Six trees from the six sessions are shown in the bottom. Shapelets and nodes in the trees have matching colors. Label 0 represents controls and 1 represents patients. Signals in red are from patients and in blue are from healthy subjects (best viewed in color).

We use the time series of frontal network component. There are 50 (SP=23, HC=27) time series for training datasets and 25 (SP=9, HC=16) time series for test data sets. The length of each time series is 411.

On the concatenated signals, the algorithm finds a very compact tree with 62% accuracy (Figure 3 - 16 (top)). The shapelets span over multiple sessions and mostly focus on the first three sessions and thus, one cannot learn from all the dimensions. In contrast, shapelet ensemble algorithm finds six trees from six sessions and ensemble them using majority voting. The ensemble achieves 72% accuracy. However, for the same data set, we decreased by 8.32x the overall computation time over the *FastShapelet* algorithm. Note that, individual trees combine information from both red (SP) and blue (HC) subjects while the tree from concatenated signals only contains blue signals. The group differences shown in these case studies provide motivation for understanding how connectivity patterns differ in response to these different stimulus conditions and the current analysis approach may provide enhanced sensitivity to identify group differences of SPs and HCs.

3.9. Conclusion

In this chapter, we proposed an algorithm to speed up the current shapelet algorithms without decreasing accuracy, especially for multi-dimensional and longer time series. The experiments showed that our algorithm is significantly faster for all data sets that we tested which makes our shapelet discovery algorithm more suitable for real life problems.

Chapter 4: Inter-network connectivity at rest and across sensory paradigms in schizophrenia

Although a number of recent studies have examined functional connectivity at rest, few have assessed differences between connectivity both during rest and across active task paradigms. Therefore, the question of whether cortical connectivity patterns remain stable or change with task engagement continues to be unaddressed. We collected multi-scan fMRI data on healthy controls (N = 53) and schizophrenia patients (N = 42) during rest and across paradigms arranged hierarchically by sensory load. We measured functional network connectivity among 45 non-artifactual distinct brain networks. Then, we applied a novel analysis to assess cross paradigm connectivity patterns applied to healthy controls and patients with schizophrenia. To detect these patterns, we fit a group by task full factorial ANOVA model to the group average functional network connectivity values. Our approach identified both stable (static effects) and state-based differences (dynamic effects) in brain connectivity providing a better understanding of how individuals' reactions to simple sensory stimuli are conditioned by the context within which they are presented. Our findings suggest that not all group differences observed during rest are detectable in other cognitive states. In addition, the stable differences of heightened connectivity between multiple brain areas with thalamus across tasks underscore the importance of the thalamus as a gateway to sensory input and provide new insight into schizophrenia.

4.1. Introduction

Functional connectivity is an approach that helps to assess the integrity of neural circuits by examining the covariance in activity across brain regions and can be assessed using a seed-based analysis approach or independent component analysis (ICA) (Calhoun VD and Adali T, 2012; Erhardt EB et al., 2011b). Seed-based approaches assess the temporal correlation between a seed region and individual brain voxels (Cordes D et al., 2002; Fox MD et al., 2005) whereas ICA is a data-driven approach which identifies spatially distinct but temporally related brain networks (Calhoun VD et al., 2001c).

To date, most studies have focused only on the analysis of functional connectivity during performance of a single task. Such an approach does not take advantage of the within-subject pattern of response which likely occurs across tasks, and which can be of benefit in a number of applications (Calhoun VD and Adali T, 2009; Calhoun VD et al., 2006, 2008). However, one of

the challenges associated with studying the resting state is that connectivity changes could reflect differences dependent on the cognitive states of the individual's brain, rather than consistent structural or functional-based differences in brain connectivity (Repovš G and Barch DM, 2012). The results reported by previous studies limit our ability to understand whether observed cortical connectivity deficits in schizophrenia represent consistent characteristics rather than differences in cognitive state or task response.

Functional MRI (fMRI) results have been used to better understand the pathophysiology of schizophrenia, in particular to assess the disconnection hypothesis of schizophrenia (Friston KJ and Frith CD, 1995; Woodward ND, 2012). These differences include a link between prefrontal cortex activation and vulnerability to psychosis (Fusar-Poli P, 2007), reduced network activation during executive task performance (Minzenberg MJ, 2009), and abnormal activation patterns in working memory tasks (Glahn DC et al., 2005). There has been growing interest in investigating the integrity of the neural circuits in schizophrenia that work together to support sensory, cognitive, and emotional processes (Calhoun VD et al., 2009; Liu H. et al., 2012; Yu Y. et al., 2013).

Previous seed-based and ICA studies (Allen EA et al., 2011a; Bassett DS et al., 2011; Cole MW et al., 2011; Garrity AG et al., 2007; Woodward ND et al., 2011) that examined functional connectivity in schizophrenia during rest found reduced connectivity for schizophrenia patients (SPs) within the default mode network, frontal network, cingulo-opercular network and cerebellar network. Several other studies (Anticevic A et al., 2011; Diaconescu AO et al., 2011; Fornito A et al., 2011) examined task-related functional connectivity in schizophrenia that largely focused on specific brain networks. These studies have also provided evidence for alterations in functional connectivity across a range of tasks where each task was studied separately. ICA provides measures of functional connectivity (within component coherence) as well as functional network connectivity (FNC) which measures changes in connectivity across networks (Jafri MJ et al., 2008).

To gain a broader understanding of brain function and dysfunction as a dynamic process, we must examine how cognition changes under an established progression of task manipulations. Dynamic

changes across tasks have been investigated with FNC and functional connectivity. Arbabshirani et al. (Arbabshirani MR et al., 2013a) compared dynamic FNC changes across two tasks including resting state and an auditory oddball task in 28 healthy controls (HCs). Results of this chapter showed decreased FNC during task relative to rest among numerous network pairs. Also, Repovš et al. (Repovš G and Barch DM, 2012) examined differences in functional connectivity during rest and a working memory task with increasing memory loads; SPs and their siblings had reduced connectivity between the frontal network and cingulo-opercular network with cerebellar network relative to HCs and their healthy siblings demonstrating network differences related to genetic risk. These group differences did not change as a function of task state or memory load. Although Arbabshirani et al. (Arbabshirani MR et al., 2013a) compared FNC across two task in HC and Repovš et al. (Repovš G and Barch DM, 2012) identified group differences in functional network across tasks, none of these studies evaluated changes in FNC across a hierarchy of tasks between the HC and SP groups.

The goal of this chapter is to determine whether cortical connectivity patterns remain stable or change across a hierarchy of sensory tasks. To the best of our knowledge there has been no study to investigate this issue in a variety of different FNC networks in a multi-task hierarchy with a relatively large number of subjects. This present chapter examined FNC across a hierarchy of sensory tasks with varying levels of sensory load. Data for each participant were gathered across multiple fMRI scanning sessions over the course of up to two months (1~2 months) with prospective randomization of task presentation and close monitoring of SPs to ensure clinical stability. Our goal was to track connectivity changes in SPs and HCs as sensory load increased. Using multiple tasks in addition to multiple conditions within a single task allows us to recognize that individuals' reactions to sensory stimuli are conditioned by the circumstances in which such stimuli are presented and measurements at separate time points allows us to better assess state versus trait group differences. We sought to determine whether SPs and HCs showed significant FNC differences among brain regions across the task hierarchy by modeling the temporal dependency between functional networks derived from fMRI data. The tasks defined a natural hierarchy related to sensory load and included a rest task, two levels of auditory sensory gating, and two levels of multisensory perception with auditory and audio-visual stimuli. We remained

skeptical of the notion that rest differences necessarily equate to characteristic differences in cognition between SPs relative to HCs. We hypothesized that data collected using a sensory load task hierarchy including rest will provide evidence of both stable (static effects) and state-based differences (dynamic effects).

4.2. Methods and Materials

4.2.1. Participants

The study in this chapter combined existing data from 95 subjects. Informed consent was obtained from all subjects according to institutional guidelines at the University of New Mexico Human Research Protections Office, and all data were anonymized prior to group analysis. Inclusion criteria for patient selection included diagnosis of schizophrenia or schizoaffective disorder between 18 to 65 years of age. Each SP completed the Structured Clinical Interview for DSM-IV Axis I Disorders (First MB et al., 2002a) for diagnostic confirmation and evaluation for co-morbidities. The imaging sessions (three cumulative hours) were completed in 1-2 sessions within up to two months (1~2 months) to reduce subject fatigue. SP had to demonstrate retrospective and prospective clinical stability to be included in this investigation. The Clinical Core (COBRE Stability Clinic) affiliated with this project determined retrospective stability from relevant psychiatric records documenting no change in symptomatology or type/dose of psychotropic medications occurred during the three months prior to the referral. The Clinical Core assessed prospective stability during three consecutive weekly visits and during each imaging assessment. Prospective stability was defined as no change in clinical symptoms > 2 points from the positive symptom items on the Positive and Negative Syndrome Scale (Kay SR et al., 1987). No score of "worse" or "much worse" on the Clinical Global Impression (Guy W, 1976) no suicidal or violent ideation, and no psychiatric or medical hospitalizations. The doses of antipsychotic medications were converted to olanzapine equivalents (Gardner DM et al., 2010). SPs with a history of neurological disorders including head trauma (loss of consciousness > 5 minutes), mental retardation, or history of active substance dependence or abuse (except for nicotine) within the past year were excluded. All SPs had a negative toxicology screen for drugs of abuse at the start of the study. HCs were recruited from the same geographic location and completed the Structured

Clinical Interview for DSM-IV Axis I Disorders – Non-Patient Edition to rule out Axis I conditions (First MB et al., 2002a). SPs and HCs were matched on parental educational level ($p<0.05$), a less biased estimate of pre-morbid educational attainment potential (Saykin, AJ et al., 1991).We assessed symptom variability among each item of the PANSS positive symptom scores. Consistent with our inclusion criteria, the included subjects had minimal variance (< 2) associated with each symptom measure. Table 4 - 1 provides demographic characteristics of the participants and Table 4 - 2 lists the medications of the patient group.

Table 4 - 1 : Demographic and clinical variables for SPs and HCs. Abbreviations: PANSS= Positive and Negative Syndrome Scale. CGI = Clinical Global Impression. PCEL: Primary caregiver education level.CODEM-6: Highest Level of Education for Primary Caretaker until 18 years old. CODEM-7: Highest Level of Education for Secondary Caretaker until subject was 18 years old. Educational levels as follows 1: grade 6 or less, 2: grade 7-12, 3: graduated high school, 4: part college, 5: graduated 2 year college, 6: graduated 4 year college, 7: graduate or professional school, 8: completed graduate or professional school.

	SP (SD) (n=42)	**HC(SD) (n=53)**	***t or x2 (p-value)***
Demographics			
Age	37.38 (13.44)	35.92 (11.97)	-0.56 (0.57)
Gender (M/F)	33/9	39/14	0.32 (0.57)
Age on onset	20.04 (8.03)		
Illness duration	16.22 (12.91)		
Calgary	3.25 (1.01)		
Depression CGI			
PCEL			
CODEM-6	4.4 (2.11)	4.75 (1.89)	0.082*(0.41)*
CODEM-7	4.93 (2.24)	4.97 (2.27)	0.082*(0.93)*
PANSS			
Positive	14.35 (5.27)		
Negative	13.95 (5.28)		

General	28.71 (9.63)		
Medications			
OE(mg/day)	16.78 (12.95)		

*Table 4 - 2 : Medication list for the patient group. *Eight of the 28 patients were treated with multiple antipsychotics. This table lists either the long acting injection or the antipsychotic with the higher olanzapine equivalents. ** mg/day or dose of long acting injection*

Antipsychotics*	Dosage Range**
Aripiprazole (n = 3)	10 - 15 mg
Clozapine (n = 4)	50 - 400 mg
Fluphenazine (oral, n = 1)	10 mg
Haloperidol (oral, n = 1)	5 mg
Haloperidol decanoate (n = 2)	50 mg
Olanzapine (n = 1)	20 mg
Perphenazine (n = 1)	8 mg
Quetiapine (n = 1)	200 - 800 mg
Risperidone (oral, n =6)	1 - 4 mg
Risperidoneconsta (n =6)	12.5 - 50 mg
Thiothixene (n = 1)	60 mg
Ziprasidone (n = 1)	160

4.2.2. Task Hierarchy

The tasks represented cognitive paradigms that were analyzed separately (Mayer AR et al., 2012; Stone DB et al., 2011). Each task represented a different cognitive demand: resting state, pre-attentional sensory processing and multisensory processing. These studies were embedded in a larger study that explicitly proposed that SP would have deficits at multiple levels including pre-

attentive sensory processing, integration of sensory information across modalities and impaired working memory performance. For the present investigation, we arranged the sensory tasks into five levels according to the amount of sensori-motor processing required by participants during the task. A subset of the subjects (SPs:22/42, HCs:23/53) were reported previously in Mayer et al. (Mayer AR et al., 2012) based on the original hypotheses of the sensory gating response. The Stone et al. (Stone DB et al., 2011) paper describes the paradigm but reports on event related potential results and does not include the fMRI results presented here.

- Resting state fMRI: On the first level, participants performed a simple rest task. Subjects were instructed to keep their eyes open during the scan and stare passively at a central fixation cross, as this is suggested to facilitate network delineation compared to eyes-closed conditions (Mayer AR et al., 2012).

- Sensory gating: Tasks on the second and third levels were variants of the paired click paradigm for testing sensory gating (Van DKR et al., 2010). On the second level, participants were presented with a 5ms tone at either 2 kHz or 3 kHz. On the third level, participants were presented with paired tones - either identical 2 kHz tones or a 2 kHz tone followed by a 3 kHz tone. Both of these tasks were passive in that a response was not required.

- Multi-modal sensory integration: Tasks on the fourth and fifth levels were designed to test multisensory integration and employed a simple forced choice behavioral task (Stone DB et al., 2011). A perspective digital drawing was used as a visual background with a fixation point, and participants were instructed to maintain fixation throughout the task. On the fourth level, an auditory stimulus (500 Hz tone) was presented to participants at two different volumes, 80dB to simulate a sound near the participant (NEAR) and 64dB to simulate it being farther away (FAR). On the fifth level, the above auditory stimulus was presented synchronously with a visual stimulus - an image of a soccer ball appeared in one of two possible positions (NEAR or FAR) in the participant's lower visual field with size and position consistent with the perspective drawing. The NEAR visual stimulus was

presented in the participant's peripheral visual field and the FAR stimulus was presented closer to fixation in central visual field. Participants underwent fMRI scanning while deciding whether the stimuli presented were NEAR or FAR with a button press. Participants all performed the task at a high level of performance more than 88% percent correct in both groups across task conditions.

Since data acquisition across the tasks takes 3 hours, fMRI data collection occurred across 1-2 scanning sessions completed within up to two months. During this time participants underwent clinical, functional, and neuropsychological assessment as well as magnetoencephalography data acquisition across 1-3 visits. As mentioned above, a brief clinical stability questionnaire was obtained at each study visit to ensure stability across the multiple study visits. Task order for the MRI scan sessions was randomized and there is no significant group differences in the task randomization at $p<0.05$ level.

4.2.3. Data Acquisition

All images were collected on a single 3-Tesla Siemens Trio scanner with a 12-channel radio frequency coil. High resolution T1-weighted structural images were acquired with a five-echo MPRAGE sequence with TE = 1.64, 3.5, 5.36, 7.22, 9.08 ms, TR = 2.53 s, TI = 1.2 s, flip angle = 7°, number of excitations = 1, slice thickness = 1 mm, field of view = 256 mm, resolution = 256×256. T2*-weighted functional images were acquired using a gradient-echo EPI sequence with TE = 29 ms, TR = 2 s, flip angle = 75°, slice thickness = 3.5 mm, slice gap = 1.05 mm, field of view 240 mm, matrix size = 64×64, voxel size = 3.75 mm×3.75 mm×4.55 mm. Resting state scans consisted of 149 volumes. Tasks on the second (19 trials/condition) and third levels (20 trials/condition) acquired 111 volumes and tasks on the fourth (17 trials/condition) and fifth levels (17 trials/condition) acquired 181 volumes per run.

4.2.4. Data Preprocessing

For preprocessing the MATLAB-based (www.mathworks.com) SPM-5 toolbox (www.fil.ion.ucl.ac.uk/spm/software/spm5) was used. To remove T1 equilibration effects, the first

four volumes were discarded. Next, the INRIalign algorithm (Freire L et al., 2002) was used to realign the images, and slice-timing correction was applied using the middle slice as the reference frame in the functional data pipeline. The data was then spatially normalized to the standard Montreal Neurological Institute space (Friston KJ et al., 1995) using a nonlinear (affine + low frequency direct cosine transform basis functions) registration, resampled to 3 mm × 3 mm × 3 mm voxels, and smoothed using a Gaussian kernel (FWHM=5mm) with a full-width at half-maximum of 10 mm. The preprocessed time series data was scaled to a mean of 100. This intensity normalization improves the test-retest reliability of the Group Independent Component Analysis (GICA) (Allen EA et al., 2011b). See Figure 4 - 1, step 1.

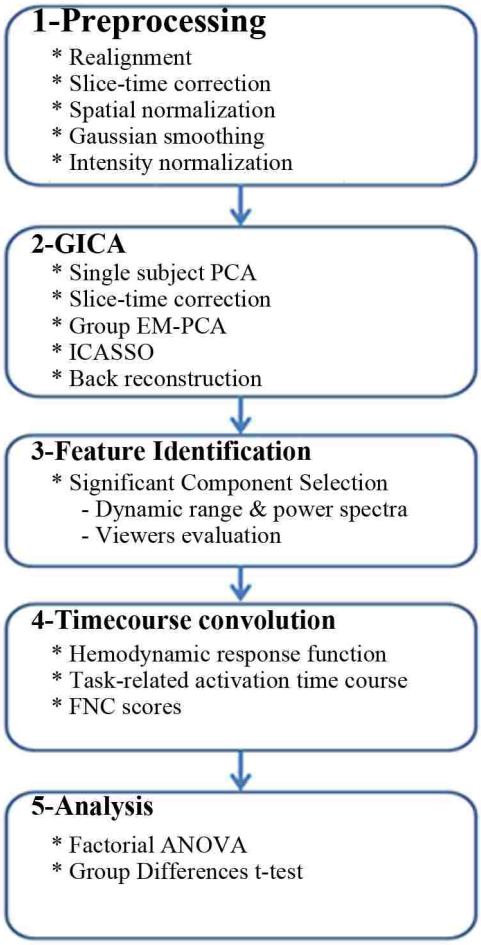

Figure 4 - 1 : Schematic of the analysis pipeline

4.2.5. Group Independent Component Analysis (GICA)

We used the GIFT Toolbox (http://mialab.mrn.org/software/gift/) and infomax algorithm (Bell AJ and Sejnowski TJ, 1995) for GICA. We performed a subject-specific data reduction principle component analysis retaining 100 principal components (PC) using a standard economy-size decomposition (Allen EA et al., 2011a). The relatively large number of subject-specific PCs has been shown to stabilize subsequent back-reconstruction (Erhardt EB et al., 2011b). Then reduced data from all subjects and all sessions were concatenated together and put through another reduction step. To use memory more efficiently, further group data reduction was performed using an expectation maximization (EM) principle component analysis algorithm (Roweis S., 1998) and 75 PCs were retained.

We used a relatively high model order ICA (number of components, C = 75), since such models yield refined components that correspond to known anatomical and functional segmentation (Abou-Elseoud A et al., 2010; Kiviniemi V et al., 2009). In order to estimate the reliability of the decomposition (Himberg J et al., 2004), the Infomax ICA algorithm was applied repeatedly in Icasso (http://research.ics.aalto.fi/ica/icasso/) and resulting components were clustered.

The ICA results, once estimated and fixed, can be considered as a set of weighted seed maps (Joel SE et al., 2011). The ICA algorithm is trying to identify maximally independent sets of maps (which can overlap) each of which are represented by a strongly coherent (correlated) time-course. In the case of a distributed set of regions, there are multiple locations which are highly correlated to one another, and thus can be considered a node in this sense. Indeed, this is a major strength of multivariate approaches like ICA, as one knows that all the voxels with strong weights in a given component are highly correlated, and thus it makes sense to consider them a node and use FNC to evaluate the inter-relationship among these nodes. In contrast, for a seed-based approach, one knows the correlation to the seed region, but the correlation between any two voxels which are correlated to the seed may not be correlated with one another (Erhardt EB et al., 2011a). Thus, biologically, we would argue it is more interpretable to work with ICA-defined regions than it does seed-derived regions. Another benefit to the ICA approach is the artifact components have some

overlap, and this provides spatial filtering or ‘cleaning’ of the remaining variance which can improve our ability to, e.g. classify groups based on the results (Erhardt EB et al., 2011a).

All rest and task data were analyzed in one group ICA instead of separate ICAs so that a tighter comparison between rest and tasks could be performed without introducing additional variability from the required matching of components if separate ICA analyses were performed. It has been shown in multiple previous papers that group ICA characterizes individual variation such as might occur across sessions quite well (Allen EA et al., 2011a; Calhoun VD and Adali T, 2012; Erhardt EB et al., 2011b). Since the ICA model constrains the fluctuations of each voxel in a given component to have the same time course, each ICA component can be considered a temporally coherent network (Erhardt EB et al., 2011a). Next we calculate the within session cross-correlation among ICA timecourses (called FNC) and subsequently model the cross-session effects as described in the paper. Therefore, comparing the time-courses of different components in the resting state and during the tasks provides insights into the change in dynamics across the task hierarchy between the independent component networks. See Figure 4 - 1, step 2.

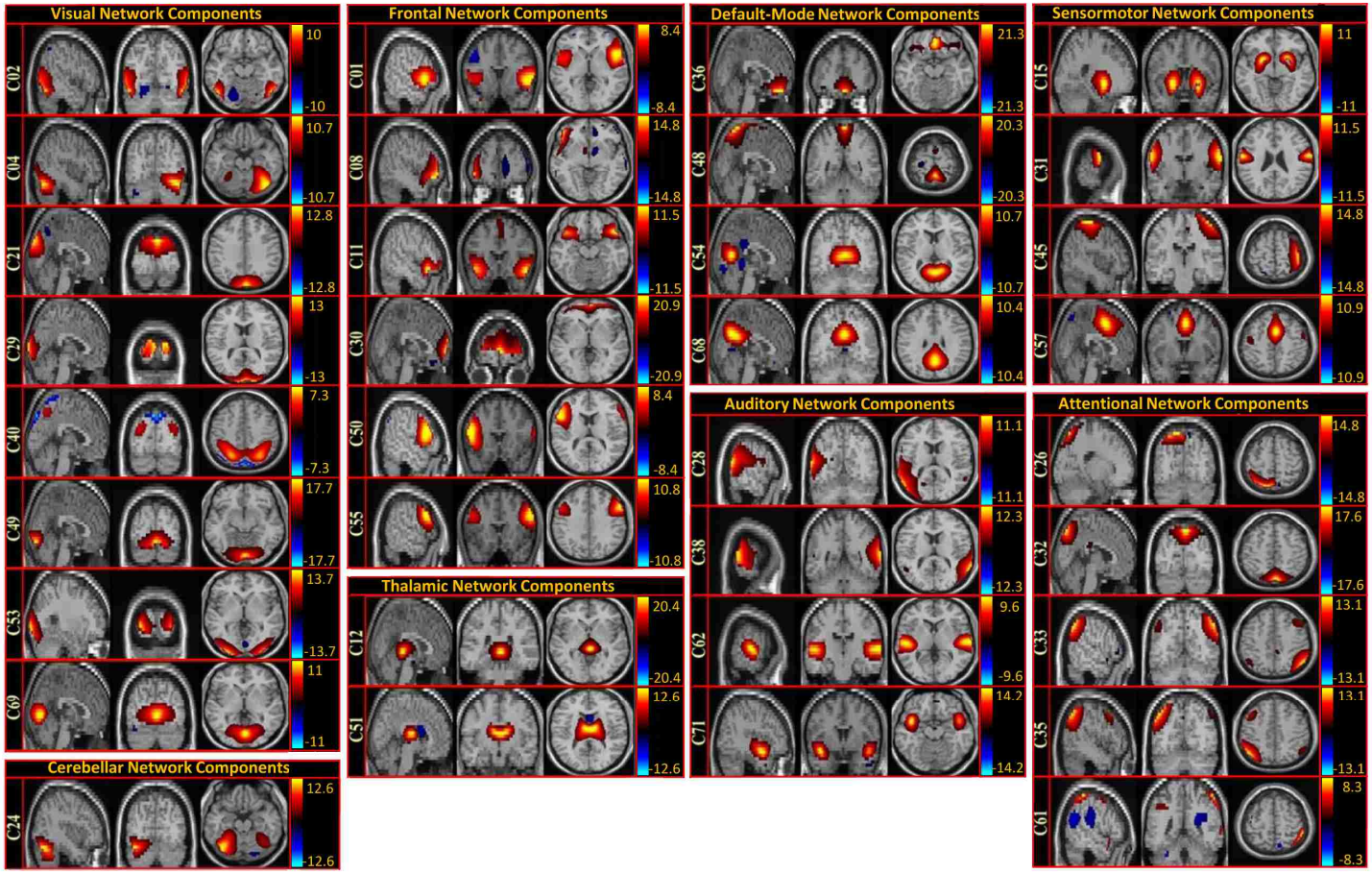

Figure 4 - 2 : Maps of the components identified as non-artifactual in static FNC or dynamic FNC analysis: Of the 75 components returned by the GICA, 45 were identified as non-artifactual components. Only 34 of these non-artifactual components showed static FNC or dynamic FNC effects. 34 non-artifactual components are divided into groups based on their anatomical and functional properties and include visual network, thalamic network, cerebellar network, frontal network, attentional network, default mode network, sensory motor network, and auditory networks.

4.2.6. Feature Identification

We used a two-step process to identify non-artifactual components that contain features associated with resting state networks and task performance on a sensory task (Robinson S et al., 2009). See Figure 4 - 1, step 3. In the first step we examined the power spectra with two criteria in mind: dynamic range and low frequency/high frequency ratio. See Figure 4 - 3a .Dynamic range refers to the difference between the peak power and minimum power at frequencies to the right of the peak in the power spectra. Low frequency to high frequency power ratio is the ratio of the integral of spectral power below 0.10 Hz to the integral of power between 0.15 and 0.25 Hz (Allen EA et al., 2011a). For the second step, three expert reviewers evaluated the components for functional relevance. See Figure 4 - 3b .In this evaluation, if a component exhibited peak activation in gray matter, low spatial overlap with known vascular, ventricular, motion, and susceptibility artifacts, and time courses dominated by low frequency fluctuations (Cordes D et al., 2000), it was classified as a non-artifactual component. At the end of the evaluation, the components were separated into two broad classes: artifactual and non-artifactual components. Of the 75 components returned by the GICA, 45 were identified as non-artifactual components. See Figure 4 - 2.

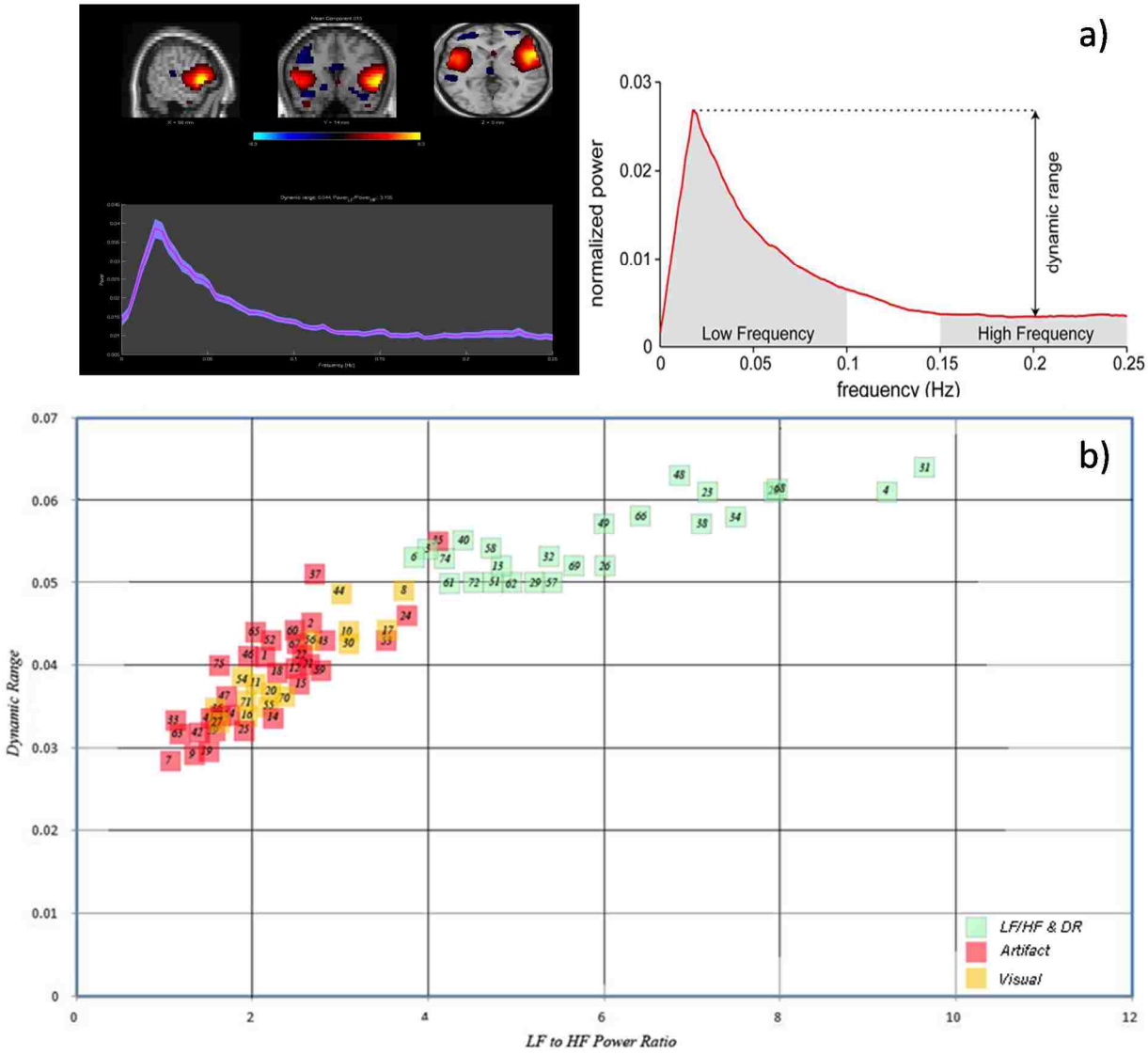

Figure 4 - 3 : Two-step processing to identify non-artifactual components; a) visual inspection, b) dynamic range and low frequency/high frequency ratio.

4.2.7. Timecourse convolution

To get FNC scores during the resting state scan, time courses (TC) of separate components in the resting state were correlated with one another by using a cosine similarity measure that can be computed as follows:

$$\text{Corr}_{(\text{Cx,Cy})} = \frac{TC_{Cx} \,.\, TC_{Cy}}{\|TC_{ICx}\| \,.\, \|TC_{ICy}\|} = \frac{\sum_{i=1}^{n} TC_{Cxi}\, x\, TC_{Cyi}}{\sqrt{\sum_{i=1}^{n} TC_{Cxi}{}^{2}}\; x \sqrt{\sum_{i=1}^{n} TC_{Cyi}{}^{2}}}$$

Where TC_{Cx} and TC_{Cy} are time courses of two separate components and *n* is length of time courses. For other tasks in the analysis, we isolated activations related to particular tasks within an fMRI scanning session. Figure 4 - 4 shows timecourse convolution of fMRI task data. The design matrix denoting stimulus presentation (when the stimuli occur for each task) during fMRI scanning sessions were convolved with a hemodynamic response function. The resulting function was normalized on a zero-to-one scale. These functions were termed hemodynamic predictor functions. A hemodynamic predictor function models the expected pattern of activation associated with a task and can be thought of as a weight expressing the degree to which component activation at a particular time would associate with a given task. Each task's hemodynamic predictor function was then convolved with the component time courses from the GICA to yield a task-related component time course. A task-related component time course indicates the activation of a particular GICA component solely as it pertains to a given task performed in the fMRI scanner, and is zero where the task does not influence activity. Task-related component time courses for separate components within a task were then correlated with one another exclusively over non-zero areas of the hemodynamic predictor function using a cosine similarity measure to yield task-related FNC scores for pairs of components. See Figure 4 - 1, step 4. The statistical tests described below were performed on these FNC scores.

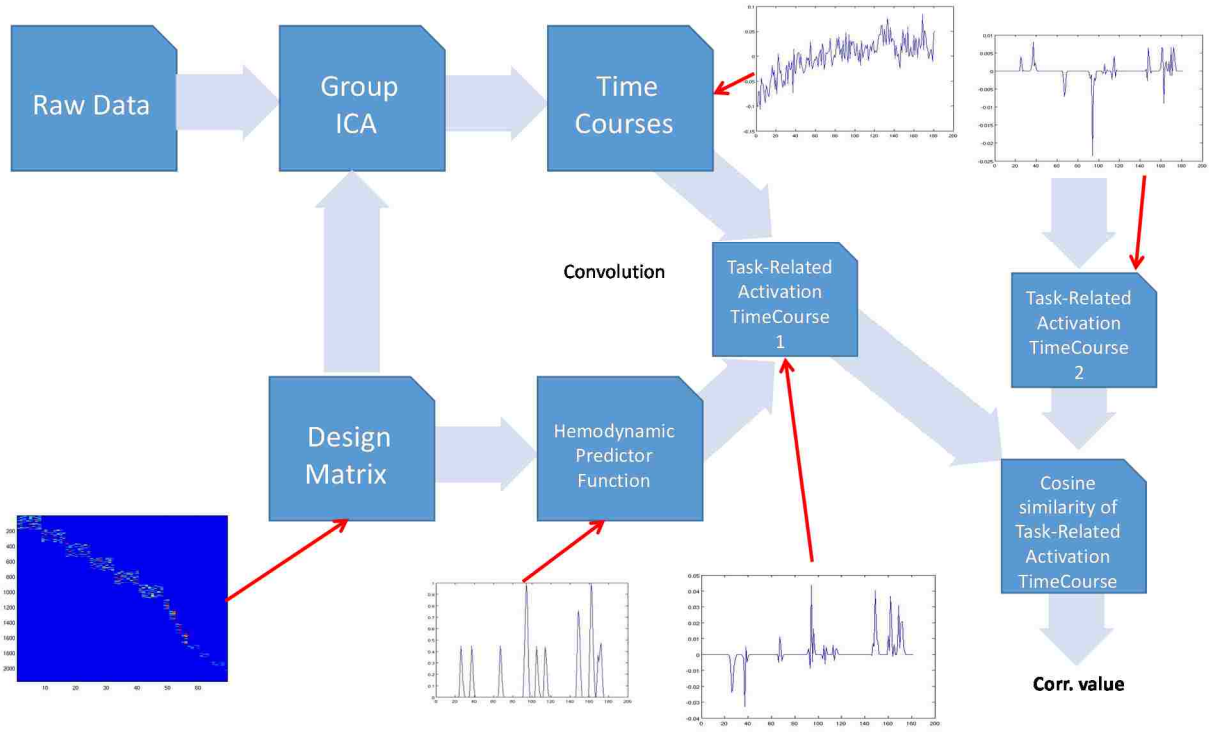

Figure 4 - 4 : Timecourse convolution of fMRI task data.

4.2.8. Data Structure

For each pair of components identified by the GICA, a vector of FNC results was created with values for every task performed by every subject. This allowed us to address questions about FNC effects between SPs and HCs at distinct levels of the hierarchy. We evaluated effects in two FNC categories. First, *static* FNC component pairs (see Figure 4 - 5A) showed *consistency* between SP and HC groups across levels of the task hierarchy (see Figure 4 - 5C). Second, *dynamic* FNC showed differences in connectivity between SP and HC groups at *different* levels of the task hierarchy (see Figure 4 - 5D). By using these two categories, we were able to identify static and dynamic group differences for SPs and HCs across task.

4.2.9. Data Analysis

We maintained an interest in where we observed static and dynamic connectivity effects, and how this analysis approach may provide insights about current findings on connectivity in schizophrenia. To detect differential (state-dependent) connectivity effects, we fit a 2x5 (Group x Task) full factorial ANOVA model to the group average FNC values. To assess medication effects, we repeated the analysis for significant component pairs from the static FNC and dynamic FNC effects with a median split of the olanzapine equivalents. See Figure 4 - 1, step 5. With 45 non-artifactual components in our data set, 990 pairwise comparisons were performed.

We examined component pairs that showed static FNC offset between groups throughout the hierarchy of tasks by using a factorial ANOVA model at $\alpha > 0.001$ level. The retained pairs demonstrated a main effect of group but did not show signs of a diagnosis-by-task interaction. We then averaged FNC values across tasks to control for individual subject effects and performed two-sample t-tests to identify those component pairs that showed significant static FNC effects ($p<0.001$) (see Figure 4 - 5C).

Decisions about whether a particular component pair showed significant dynamic FNC effects was based on an F-test of the model including the task-by-diagnosis interaction term (factorial ANOVA model at $\alpha \leq 0.001$ level)(see Figure 4 - 5D). We defined the results as a significant diagnosis-by-task interaction. According to our hierarchy of tasks these component pairs have FNC differences between SPs and HCs as a function of task level.

To determine the proportion of false discoveries in the set of declared discoveries, we used a method as described in Soric et al. (Soric B, 1989). For n number of experiments with r discoveries, α is the probability of wrong null-hypothesis rejections where $r/n > \alpha$ and for r rejections of null hypotheses the proportion of fallacies Q has least upper bound: $Q_{max} = (n/r-1)\alpha/(1-\alpha) < 1$. In our study, Q_{max} (dynamic FNC) = 0.06 and Q_{max}(static FNC) = 0.05. Based on this analysis, *at least* 94% of dynamic FNC and 95% of static FNC discoveries are true.

4.3. Results

4.2.1. Static Functional Network Connectivity Effects

We obtained 18 significant pairs (see Figure 4 - 5A lower-left portion of the connectivity image) that showed static FNC effects. The significant static effects with a main effect of group and no interaction between group and task are shown in Figure 4 - 5C, D denoted by the solid blue or red squares. We present our component pair results by first specifying the anatomic representation followed by the independent component(IC) number.

Our study found higher static FNC in SP relative to HC (solid red squares) between the following thalamic network pairs (see solid red blocks in Figure 4 - 5A): thalamus (IC12) / auditory networks (IC28), thalamus (IC51) / auditory networks (IC28), thalamus (IC12) / auditory networks (IC38), and the thalamus (IC51) / auditory networks (IC62), thalamus (IC51) / sensory motor network (putamen)(IC15) and the thalamus(IC51) / visual network (extrastriate cortex)(IC53). The thalamic FNC are the only network pairs that show increased static FNC in SP relative to HC with only one exception, attentional network (IC33) / visual network (IC21).

The remaining static FNC results describe higher connectivity in HC relative to SP (solid blue squares). These static FNC effects were observed in the default mode network, in particular, connectivity relating to the posterior cingulate cortex (IC54) / cingulate cortex (IC68). While we did not find overall static FNC patterns governing the interaction of the default mode network with other component networks, we found a number of individual brain regions showing stable differences between SPs and HCs. The posterior cingulate cortex showed reduced connectivity for SPs in the following pairs: default mode network(IC54) / auditory networks (IC62), dorsal precuneus, default mode network(IC68) /attentional network (IC32), post-central gyrus; default mode network (IC54) / sensory motor network (IC31), default mode network (IC54) / visual network (IC69) and lingual gyrus; default mode network(IC68) / visual network (IC49). We also observed static FNC effects in other parts of the default mode network (IC48) / frontal network (IC30) and default mode network (IC48) / visual network (IC29).

Other component pairs showing static FNC effects at the $\alpha > 0.001$ level are summarized in Figure 4 - 5A.

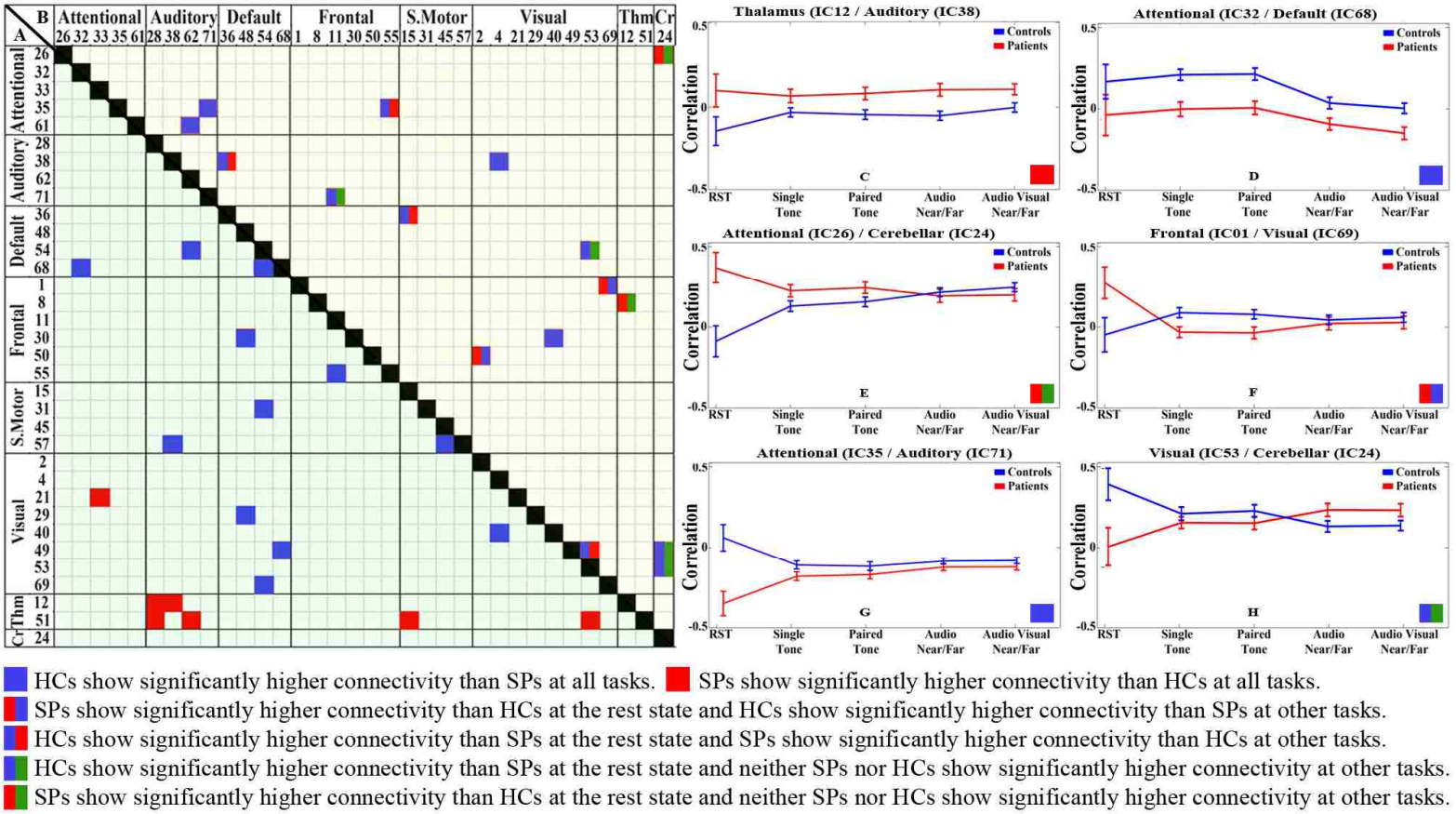

Figure 4 - 5 : A) static FNC matrix(lower part). Pairwise correlations of component pairs showed static FNC effects at the α> 0.001 level. B) dynamic FNC matrix(upper part). Pairwise correlations of component pairs showed dynamic FNC effects at the α≤ 0.001 level. C-D) Samples for static FNC; thalamus (IC12) / auditory networks (IC38), attentional network (IC32) / default mode network (IC68).E-F-G-H) Sample for dynamic FNC; attentional network (IC26) / cerebellar network (IC24), frontal network (IC01) / visual network (IC69), attentional network (IC35) / auditory network (IC71), visual network (IC35) /cerebellar network(IC24)

4.2.2. Dynamic Functional Network Connectivity Effects

We obtained 16 significant pairs (see Figure 4 - 5B) that showed dynamic FNC effects. Component pairs that yielded a significant dynamic FNC effect showed different patterns of interaction (and often negligible FNC differences) across the active tasks (See Figure 4 - 5D, F, G, H) and resting state measures also showed considerable differences.

We found significant dynamic FNC in the thalamus, posterior cingulate cortex, and posterior temporal area for the following pairs: thalamus (IC12) / frontal network (IC08); default mode network (cingulate cortex) (IC54) / visual network (extrastriate cortex) (IC53); auditory networks (posterior temporal area) (IC38) /default mode network (orbito frontal cortex) (IC36); and auditory networks (posterior temporal area) (IC38) /visual network (extrastriate cortex)(IC04). SPs showed significantly higher FNC than HCs in resting state on the thalamus (IC12) and frontal network (IC08) pair, while HCs FNC was higher with the task-related paradigms.

Unlike the observed static FNC effects, SPs had higher FNC than HCs in the resting state relative to task activated conditions in the following pairs of networks: thalamus (IC12) / frontal network (IC08), attentional network (IC26) / cerebellar network (IC24), frontal network (IC01) / visual network (IC69) and frontal network (Broca's area) (IC50) / visual network (IC02).

Some visual network components had the opposite pattern. For example, two visual network components: visual networks (extrastriate cortex) (IC53) / visual networks (IC49) which is located close to the calcarine fissure. Also, they showed significant dynamic FNC effects with the cerebellar network (IC24)/ visual networks (IC49) and cerebellar network (IC24)/ visual networks (IC53). Both showed positive FNC among SPs and HCs throughout the hierarchy, although HCs showed more FNC than SPs at resting state. Other component pairs showing dynamic FNC effects at the $\alpha \leq 0.001$ level are summarized in Figure 4 - 5B.

We, also, used a paired t-test to check the effect of antipsychotic dose (olanzapine equivalents), PANSS score, age and gender to the FNC values at $p<0.05$ and did not detect any significant effect on the significant static or dynamic effects reported in this chapter.

4.3. Discussion

The method used in this study identifies static and dynamic differences in cortical networks by determining whether SPs and HCs showed significant FNC differences among brain regions across the task hierarchy. To detect these differences, we fit a group by task full factorial ANOVA model to the group average FNC values. This method provides a better approach of recognizing the reactions of SPs and HCs to simple sensory stimuli that are conditioned by the circumstances. The present study highlights a number of FNC differences between SPs and HCs, indicative of both stable abnormalities (static FNC) and task-dependent effects (dynamic FNC). Static FNC differences in SPs impacted multiple networks involving the thalamus, posterior cingulate, and posterior temporal cortices. Dynamic FNC differences in SPs affected similar anatomic locations, but the affected number of components was less.

Among the regions where we observed significant static FNC effects, the thalamus and posterior temporal area (centered on posterior superior temporal gyrus but extending into middle temporal gyrus) deserve particular mention (see cluster of red boxes denoting Thalamus/Auditory FNC in Figure 4 - 5A). The temporal lobe has long been implicated in schizophrenia including multiple studies demonstrating auditory-related deficits (Hamm JP et al., 2011; Meda SA et al., 2012) in addition to structural changes in posterior temporal regions (Cullen AE et al., 2012; Mathiak K et al., 2011). This is consistent with differences identified in the posterior temporal lobe region in response to unisensory auditory stimuli during the multisensory integration paradigm; only the response to multisensory stimuli showed enhanced responses in SP compared to HC as reported in Stone et al. (Stone DB et al., 2011). These group differences provide motivation for understanding how connectivity patterns differ in response to these different stimulus conditions. Interestingly, no significant group differences were reported in Mayer et al. (Mayer AR et al., 2012) for sensory gating, indicating that the current analysis approach may provide enhanced sensitivity to identify group differences through the comparison across task thereby identifying both stable abnormalities (static FNC) and task-dependent effects (dynamic FNC).

Furthermore, thalamic abnormalities in schizophrenia are well-documented (Goff DC and Coyle JT, 2001; Shenton ME et al., 2001; Woodward ND, 2012). Recent research suggests that these

regions play a role related to production of auditory verbal hallucinations (Hoffman RE. and Hampson M., 2012) along with the putamen, which also showed a significant static thalamic connectivity effect.

The present study is consistent with previous results suggesting that in addition to cortical deficits, the thalamus plays an important role in the cortical abnormalities of schizophrenia. We found static FNC effects between the thalamus and posterior temporal area, the putamen, and extrastriate cortex. In all cases, SPs had significantly higher connectivity than HCs, across all levels of the sensory load hierarchy.

In this chapter, we used a high noise model of schizophrenia instead of a direct measure of brain signal variability. The use of the high noise model (Daniel PK et al., 2006; Rolls ET et al., 2008; Susan WG et al., 2009; Tononi G and Edelman GM, 2000) here is referring to the dampened stimulus evoked responses, similar to the previous descriptions that indicate default mode activity is not sufficiently dampened during stimulus presentation in SZ relative to HC. The group differences in static FNC may represent an inability for brain connectivity to be modulated by external stimuli consistent with the high noise model of schizophrenia (see Figure 4 - 5D). Interestingly, these areas are not only a part of the default mode network they may also represent broader group differences signifying a lack of flexibility in SP. Our data supports the conclusion that these effects, along with other differences involving the posterior cingulate and the default mode network, represent stable characteristics underlying schizophrenia.

We also found significant static FNC differences involving elements of the default mode network, particularly the posterior cingulate. This provides support for the hypothesis (Dosenbach NUF et al., 2007; Garrity AG et al., 2007; Jafri MJ et al., 2008) that aberrant default mode network connectivity in resting state reflects the existence of a stable static FNC characteristic that does not depend on cognitive state or task response in a particular default mode network sub-region. Future studies can also focus on changes in the dynamics of functional connectivity, a new area of investigation which provides interesting results in the healthy brain (Allen EA et al., 2012).

Our study identified dynamic FNC group differences across the active tasks between both the Fronto-Parietal networks (frontal network (IC08)) / cingulo-opercular network (thalamus (IC12) and Fronto-Parietal networks (attentional network (IC26)) / cerebellar network (IC24). These results show differences in connectivity in a similar network as the previous study by Repovš et al. (Repovš G and Barch DM, 2012) showing group differences in connectivity with increasing working memory load between the Fronto Parietal networks. Furthermore, connectivity patterns varied across components including in some cases differences in connectivity with increasing sensory load. While this study did not employ a working memory paradigm, the similar pattern of dynamic connectivity changes suggests that the dynamic connectivity patterns form a more general pattern across sensory and cognitive paradigms. In addition, we also found other dynamic FNC effects which are summarized in Figure 4 - 5B. Although, 990 comparisons limit our ability to fully describe the results, we feel this approach provides a novel method for assessing trait-based and state-based differences between clinical groups.

The majority of the previous studies that examined functional connectivity in schizophrenia during rest, found reduced connectivity for SPs while some other studies found stronger connectivity during the performance of the task (Harrison et al., 2008; Shirer et al., 2012). Most of these studies were limited to a few preselected brain regions using a seed-based approach. In this chapter, we used an ICA approach (FNC) to extract all measurable non-artifactual brain networks (Jafri MJ et al., 2008). The results of this study support the previous studies described above by identifying group differences in commonly reported networks including: the attentional network, auditory networks, default mode network, frontal network, sensory motor network, visual network, thalamus and cerebellar network. Also, our results provide additional information about how each of these networks responds to different sensory loads through testing both static FNC and dynamic FNC changes. Our study extends the previous results by identifying group differences in 18 static FNC and 16 dynamic FNC in 8 different networks in a multi task hierarchy with a large number of subjects.

There are likely some spatial differences in the networks across tasks. For ICA analyses, we have two options. The first option is using separate ICA analyses for each task which then requires

component matching to make any statements regarding task hierarchy. This is a time consuming and error prone process (especially for large numbers of subjects) which is less robust to noise (Du, Y. et al., 2014). The second option is incorporating all rest and task data in one group ICA which eliminates the need to match components across ICA analyses due to a common analytic model and thereby allows for tighter comparisons between rest and tasks. It has been shown in multiple previous papers that group ICA characterizes individual variation such as might occur across sessions quite well (Allen EA et al., 2011b; Calhoun VD and Adali T, 2012; Erhardt EB et al., 2011a). Next we calculate the within subject/within session (task) cross-correlation among ICA timecourses (called FNC) and subsequently model the cross-session effects as described in the paper. Such an approach thus allows for each component to have a distinct timecourse for each task/subject.

4.4. Limitations and Future Work

We consider the current results a first step in developing more sophisticated models to map hierarchical patterns across different brain regions to better understand how information content influences brain function at all levels of cortical functioning. Furthermore, the task hierarchy defined in the current study was based on a study designed to assess function across the sensory to cognitive hierarchy by sampling differences at pre-attentional sensory levels, multisensory integration with motor response, and working memory performance. In the current study we limited the analysis to the sensory tasks based in part on the recent results linking sensory deficits to cognitive impairments in schizophrenia. Furthermore, the current design attempts to systematically assess the perceptual/arousal state of participants which is included as a sub-domain of the Research Domain Criteria established to recognize the variability within diagnostic categories and the similarities across diagnostic categories. Additional task hierarchies are likely to yield different results within the dynamic FNC framework. In this section we discuss limitations of the current study and future work.

One challenge with performing a multi-task analysis is to determine the expected hierarchal relationship between these tasks. Here, we chose the two simplest assumptions: static versus dynamic along the preselected hierarchical arrangement. There are a number of possible reasons

that this assumption may fail with the current dataset, including the following examples: (a) levels 4 and 5 require a response, whereas levels 2 and 3 do not. So from the broader cognitive perspective, levels 4 and 5 clearly require higher cognitive load than levels 2 and 3 suggesting that higher order brain areas may follow a relationship across the established hierarchical framework. (b) Similarly, the change from level 4 to 5 represents an addition of a visual stimulus with presentation of the exact same auditory stimulus in both levels which, again, one might argue, would add complexity at the cognitive level, but the auditory cortex response might be expected to remain constant from level 4 to 5. Despite these considerations, we identified a natural hierarchy(linearly increased) in some of the IC pairs, such as, static FNC effects: attentional network (IC33) / visual network (IC21), thalamus (IC51) / auditory networks (IC62), default mode network (IC48) / visual network (IC29) and dynamic FNC effects: visual network (IC49) / visual network (IC53), auditory networks (IC38) / visual network (IC04), auditory networks (IC62) / attentional network(IC61). 990 comparisons limit our ability to fully describe the results, but we feel this approach provides a novel method for assessing trait-based and state-based differences between clinical groups.

Only subjects able to complete all tasks were retained in the final analysis. Such a rigorous research protocol inevitably excluded more symptomatic subjects who were unable to complete the entire protocol, as confirmed by comparing our final sample with the entire sample that began this study. Thus, our results may have limited generalizability to the entire spectrum of symptom severity across schizophrenia. Second, all of our SPs were taking antipsychotic medications (as noted in our inclusion criteria) whereas HCs were not. Antipsychotic dose and type did not change for the duration of the research protocol (up to two months). As mentioned above we found no relationship between medication dose and FNC measures, however, a main effect of medication cannot be eliminated in this study.

Another challenge is performing a timecourse convolution with a multi-task analysis. While the entire resting-state task time course was analyzed, for other tasks, a hemodynamic predictor function was convolved with the component time courses from the GICA to yield a task-related component time course. Although, this causes some loss of data in the time course, it helps to

reduce noise and improve the correctness of the correlation results between the component pairs (Arbabshirani MR et al., 2013a; Mayer AR et al., 2012). Based on the different questions that are being asked in rest versus task-based activations, it is unclear how to better address these differences.

In future studies, we intend to explore aspects that may improve the current methodology. An important extension to the current study is to include cognitive tasks with established pathophysiology in schizophrenia, such as working memory, delayed match-to-sample, reinforcement learning, or Go/No-Go tasks.

4.5. Conclusion

In conclusion, we used a novel analysis to investigate cross paradigm connectivity patterns applied to a large dataset of patients with schizophrenia contrasted with healthy controls. The results provided evidence of both static FNC differences between SPs and HCs and differences modulated by the rest and task hierarchy. This suggests that FNC differences observed only at rest or during performance of particular tasks are not necessarily indicative of the fundamental characteristics of cognition in schizophrenia.

Chapter 5: Sensory load hierarchy-based classification of schizophrenia patients

Schizophrenia is currently diagnosed by physicians through evaluation of their clinical assessment and a patient's self-reported experience over the longitudinal course of the illness. There is great interest in identifying biologically based markers at illness onset, rather than relying on the evolution of symptoms across time. Functional connectivity shows promise in providing individual subject predictive power. However, the majority of previous studies considered only the analysis of functional connectivity during resting-state or performance of a single task. Changes in connectivity between rest and multiple tasks have not been used in the discrimination of schizophrenia patients from healthy controls. In this work, we propose a framework for classification of schizophrenia patients and healthy control subjects based on functional network component pairs which show consistency between patients and controls across levels of the resting-state data and task hierarchy. Our results show that these functional network components as a function of task contain valuable information for individual prediction of schizophrenia patients. Such information is useful for training and replicates in testing. Performance was improved significantly (up to ~20%) relative to a single FNC (resting-state) measure.

5.1. Introduction

The general approach for the diagnosis of schizophrenia is based on a patients self-reported experiences and observed behavior over the longitudinal course of the illness. There is great interest in identifying biologically based marker of illness, rather than relying on symptom assessment because the current approach may postpone the diagnosis of the disorder, whereas early diagnosis can improve treatment response and reduce associated costs (Kubicki, M. et al., 2007). But small numbers of training subjects and high dimensional datasets make it challenging to design robust and accurate classifiers for schizophrenia. Functional connectivity shows promise in providing individual subject predictive power. Seed-based functional connectivity approaches assess the temporal correlation between a seed region and individual brain voxels (Cordes D et al., 2002; Fox MD et al., 2005). Independent component analysis based functional network connectivity (FNC) is a correlation value that summarizes the overall connection between independent brain maps over time (Allen EA et al., 2012; Calhoun VD et al., 2001c, 2008; Cetin, MS., et al., 2014; Jafri MJ et al., 2008). Therefore, the FNC feature gives a picture of the connectivity pattern over time between independent components.

Most of the previous studies have focused only on the analysis of functional connectivity during performance of a resting-state or single task for classification. Changes in connectivity between extended rest and multiple tasks have not been used in the discrimination of schizophrenia patients from healthy controls. Such an approach does not take advantage of the within-subject pattern of response which likely occurs across tasks, and which can be of benefit in a number of applications (Calhoun VD and Adali T, 2009; Calhoun VD et al., 2006, 2008). Only a few studies (Arbabshirani MR et al., 2013a; Cetin, MS., et al., 2014; Repovš G and Barch DM, 2012) have investigated how cognition changes under an established progression of task manipulation, but these studies have not focused on individual subject measures in the context of classification.

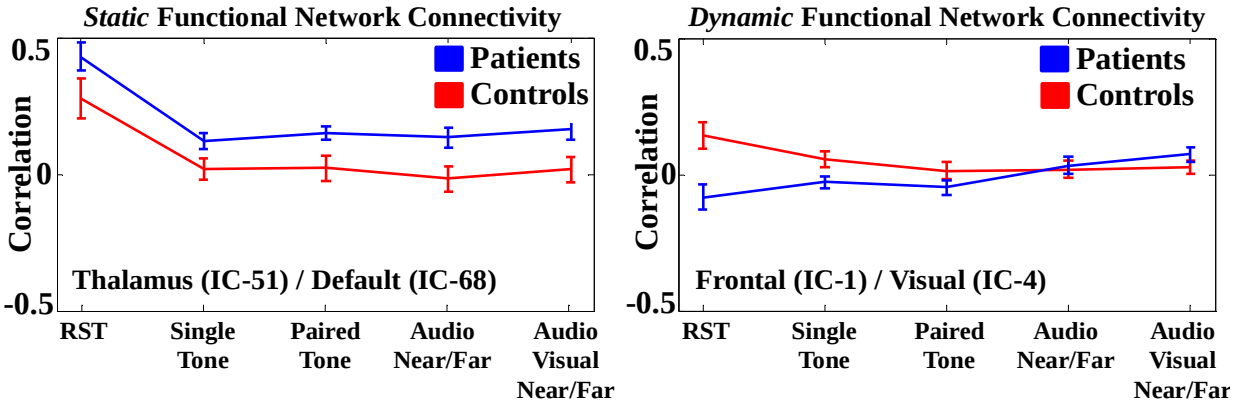

Figure 5 - 1 : Static FNC pair (left), dynamic FNC pair (right)

Some studies (Dosenbach NUF et al., 2007; Garrity AG et al., 2007; Jafri MJ et al., 2008) suggest that aberrant default mode network connectivity during the resting state reflects the existence of a stable FNC characteristic that does not depend on cognitive state or task response in a particular default mode network sub-region. Results of previous chapter supports this hypothesis by showing the existence of significant FNC patterns that remain stable between schizophrenia patients (SPs) and healthy control (HCs) across a hierarchy of sensory tasks. We reported these differences in two categories (See Figure 5 - 1) (1) s*tatic* functional network connectivity (*s*FNC) pairs showed *consistency* between SPs and HCs groups across different tasks/levels and (2) *dynamic* FNC showed differences in connectivity between SPs and HCs groups for *different* tasks/levels.

Based on this, we hypothesized that the *s*FNC components contain valuable trait-based information that can be used for individual prediction of mental illness such as schizophrenia. In this chapter, we used a group ICA approach (Calhoun VD and Adali T, 2012; Erhardt EB et al., 2011a), excluded non-artifactual brain networks (Jafri MJ et al., 2008), then conducted a classification study of schizophrenia and healthy subjects using *s*FNC and compared results with *d*FNC by examining across a hierarchy of sensory tasks with varying levels of sensory load (including resting-state only). Data for each participant were gathered across multiple fMRI scanning sessions and both FNC analyses were examined across a hierarchy of sensory tasks with varying levels of sensory load. To the best of our knowledge there has been no study to investigate this issue in a variety of different FNC networks in a multi-task hierarchy with a relatively large number of subjects (n=95).

5.2. Materials and methods

Details of demographic characteristics and clinical variables of the participants can be found in previous chapter.

5.2.1. Data Acquisition and Preprocessing

All images were collected on a single 3-Tesla Siemens Trio scanner with a 12-channel radio frequency coil with a repetition time of 2 sec. After initial standard preprocessing (Allen EA et al., 2012; Cetin, MS., et al., 2014) the imaging data was decomposed into functionally homogeneous cortical and subcortical regions exhibiting temporally coherent activity using a high model order (100) group-level spatial independent component analysis (GICA).

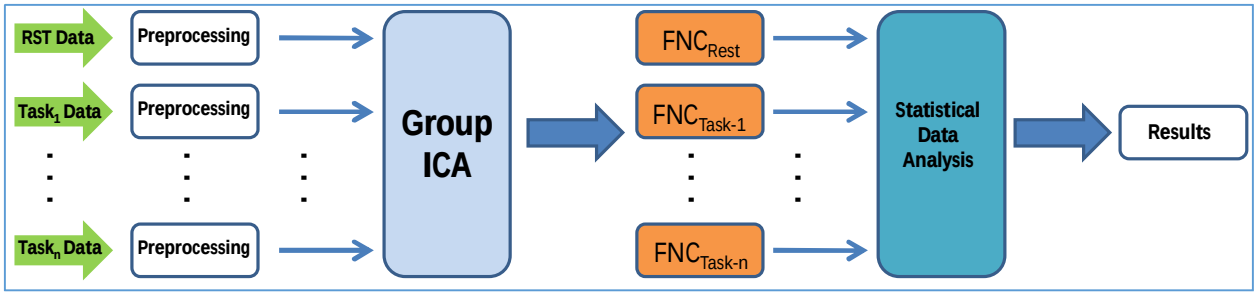

Figure 5 - 2 : Schematic of the analysis pipeline

Of the 75 components returned by the GICA, 45 were identified as non-artifactual (see Figure 5 - 3) independent components (IC) using a combination of two methods (Allen EA et al., 2011a). In the first method we examined the power spectra with two criteria in mind: dynamic range and low frequency/high frequency ratio. Dynamic range refers to the difference between the peak power and minimum power at frequencies to the right of the peak in the power spectra. Low frequency to high frequency power ratio is the ratio of the integral of spectral power below 0.10 Hz to the integral of power between 0.15 and 0.25 Hz. To verify the results, three expert reviewers evaluated the components for functional relevance. In this evaluation, if a component exhibited 1) peak activation in gray matter, 2) low spatial overlap with known vascular, ventricular, motion, and susceptibility artifacts, and 3) TCs dominated by low frequency fluctuations, it was classified as a non-artifactual component. Schematic of the analysis is presented in Figure 5 - 2.

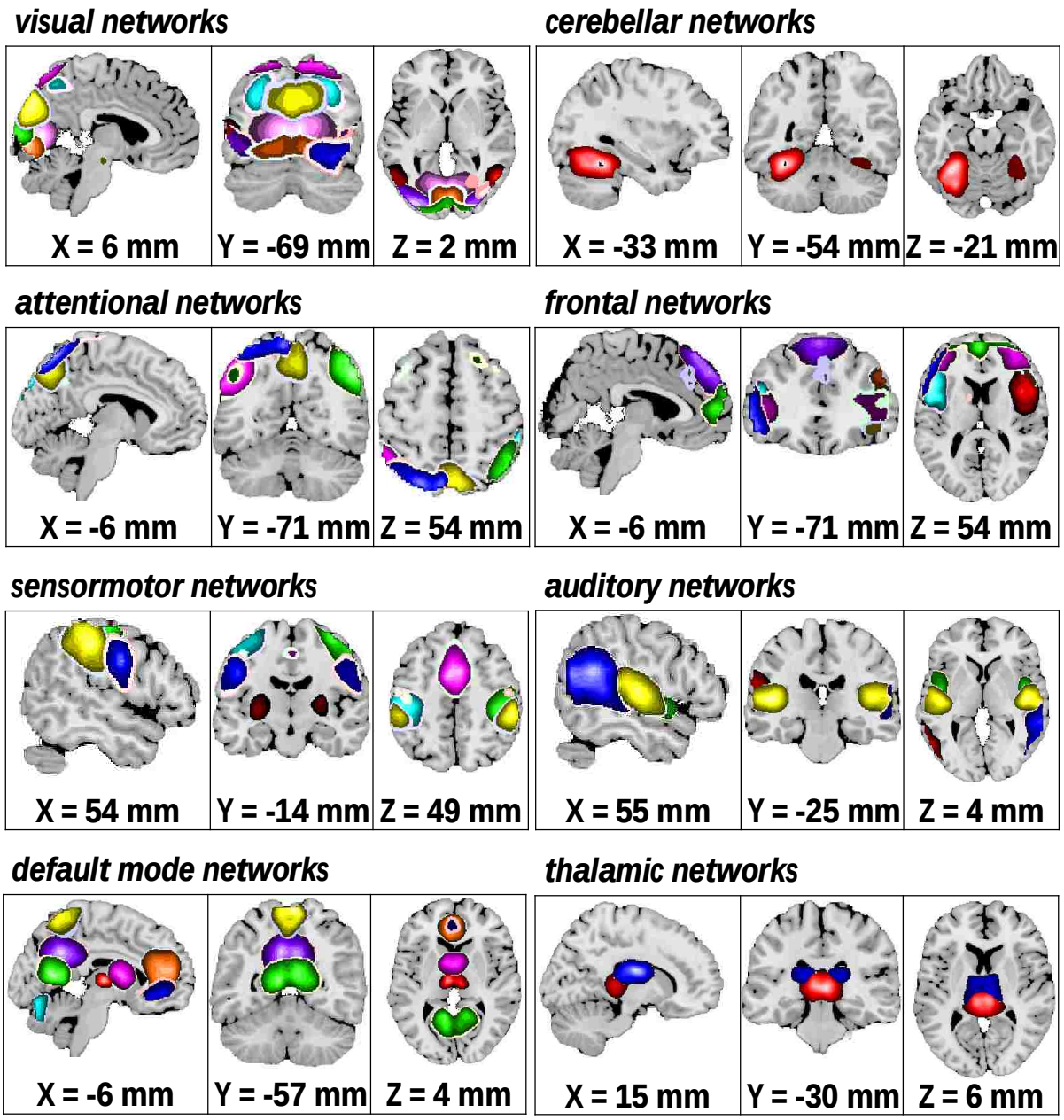

Figure 5 - 3 : Thresholded group mean spatial maps of 45 non-artifactual independent components

Subject specific time courses (TCs) and spatial maps (SMs) were obtained using back reconstruction (Erhardt EB et al., 2011a). The FNC for each subject was estimated from the TC matrix as a C×C sample covariance matrix by using a cosine similarity measure. For tasks in the analysis, we isolated activations related to particular tasks within an fMRI scanning session. Task-related component time courses for separate components within a task were then correlated with

one another exclusively over non-zero areas of the hemodynamic predictor function using a cosine similarity measure to yield task-related FNC scores for pairs of components.

5.2.2. Task Hierarchy

The tasks represented cognitive paradigms that were analyzed separately (Mayer AR et al., 2012; Stone DB et al., 2011) and together (Cetin, MS., et al., 2014). Each task represented a different cognitive demand: resting state, pre-attentional sensory processing and multisensory processing.

Level 1: Resting state fMRI: On the first level, participants performed a simple rest task. Their eyes were open during the scan and they gazed passively at a central fixation cross.

Level 2-3: Sensory gating: Tasks on the second and third levels were variants of the paired click paradigm for testing sensory gating (Van DKR et al., 2010). On the second level, participants were presented with a 5ms tone at either 2 kHz or 3 kHz. On the third level, participants were presented with paired tones - either identical 2 kHz tones or a 2 kHz tone followed by a 3 kHz tone. Both of these tasks were passive in that a response was not required.

Level 4-5: Multi-sensory integration: Tasks on the fourth and fifth levels were designed to test multisensory integration and employed a simple forced choice behavioral task (Stone DB et al., 2011). On the fourth level, an auditory stimulus (500 Hz tone) was presented to participants at two different volumes, 80dB to simulate a sound near the participant (NEAR) and 64dB to simulate it being farther away (FAR). On the fifth level, the above auditory stimulus was presented synchronously with a visual stimulus. Participants underwent fMRI scanning while deciding whether the stimuli presented were NEAR or FAR with a button press. Participants all performed the task at a high level of performance more than 88% percent correct in both groups across task conditions.

A brief clinical stability questionnaire was obtained at each study visit to ensure stability across the multiple study visits. Task order for the MRI scan sessions was randomized and there were no significant group differences in the task randomization at $p<0.05$ level. For rest of the chapter "*task*

hierarchy" refers to the fMRI scans in the following order: resting-state, sensory gating, multi-sensory integration.

5.2.3. Static Functional Network Connectivity

To find *d*FNC and *s*FNC, we followed the steps of our previous chapter (Cetin, MS., et al., 2014). We fit a 2 × 5 (Group × Task) full factorial ANOVA model to the group average FNC values hence we detect differential (state-dependent) connectivity effects. With 45 non-artifactual components in our data set, 990 pairwise comparisons were performed. We examined component pairs that showed *d*FNC diagnosis by task interaction throughout the hierarchy of tasks by using a factorial ANOVA model at $\alpha \leq 0.001$ level. The retained pairs that demonstrated a main effect of group but did not reveal a diagnosis-by-task interaction were then evaluated further by averaging FNC values across tasks to control for individual subject effects and performed two-sample t-tests to identify those component pairs that showed significant *s*FNC effects ($p<0.001$). For the rest of the chapter, IC pairs that show *d*FNC or *s*FNC properties will be called *significant* IC pairs.

To determine the proportion of false discoveries in the set of declared discoveries, we used a method as described in Soric et al. (Soric B, 1989). For *n* number of experiments with *r* discoveries, α is the probability of wrong null-hypothesis rejections where $r/n > \alpha$ and for *r* rejections of null hypotheses the proportion of fallacies *Q* has least upper bound: $Q_{max} = (n/r-1)\alpha/(1-\alpha) < 1$ which hold when we have at least 18 *significant* IC pairs across levels of the task hierarchy.

5.2.4. Classification

We performed 3 well known classification algorithms; linear discriminant classifier (LDC), Naïve Bayes classifier (NBC) (Duda RO, et al., 2001) and non-linear SVM (nSVM) with Gaussian radial bases function kernel (Burges, C. J. C., 1998) to test the initial hypothesis.

80% of the SPs and HCs datasets were used for classifier training, which returned at least 18 IC pairs across levels of the task hierarchy and rest of the data were used for the testing data set.

Members of each data set were selected randomly. Each subject has a feature vector containing 990 × 5 (IC pairs × tasks) elements and each IC pair has 5 features.

We used a repeated random sub-sampling validation (Kohavi, Ron., 1995) strategy to estimate the accuracy scores for each of the IC pairs. In each validation run, 80% of the SP and HC in the **training set** were used to train *significant* IC pairs and each of them were tested individually on validation data which is 20% of the SPs and HCs in **training set** and their accuracy scores were recorded. Members of each data set were selected randomly. We repeated a random sub-sampling validation process 1000 times and used the average scores of each IC pair to select the best half of the IC pairs. Then each IC pair in the best half was used to classify the testing data set individually. Majority voting of these individual classifications was used to compute the estimated accuracy scores of classification (Es Acc).

The entire process was repeated 100 times and the average accuracy score was recorded as a final classification accuracy score.

5.3. Results

In order to assess the advantages of using *s*FNC for classification of SPs and HCs, we examined metrics such as: comparison of the estimated accuracy scores with real accuracy scores (Real Acc) of the IC pairs, examination across a hierarchy of sensory tasks with varying levels of sensory load and comparison of *s*FNC results with *d*FNC.

By using repeated random sub-sampling validation we estimated the accuracy score of each *significant* IC pair and sorted them based on this score and used the best half of these pairs. We compared our accuracy scores with maximum accuracy scores (Max Acc) that used the best half of IC pairs based on real accuracy scores obtained from testing data set. Figure 5 - 4 shows that repeated random sub-sampling validation was successfully used to estimate the individual accuracy sores of IC pairs. The most accurate half of pairs identified in the training set (**Best half**) were used to classify the testing data set individually then majority voting of these individual classifications was used to compute the final classification score.

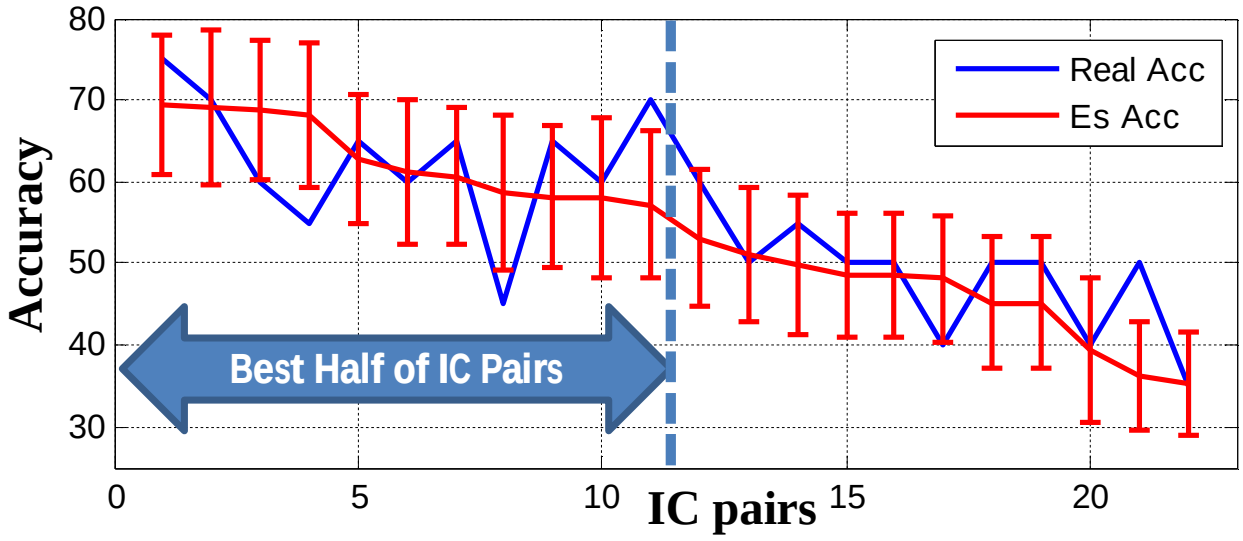

Figure 5 - 4 : Estimated order of significant IC pairs based on accuracy scores obtained from training data set (labelled Es Acc - and colored red) and real order of IC pairs based on accuracy scores obtained from testing data set (labelled Real Acc - and colored blue). Correlation score of Es Acc and Real Acc of all significant IC pairs is 0.67.

As mentioned before, we used 3 different classification algorithms to exam the success of the described method. We repeated the proposed method across varying levels of sensory load; resting state (R), resting state - sensory gating (S) and resting state - sensory gating – multisensory integration (M).

Table 5 - 1 summarizes the average accuracy scores of 100 runs on the testing dataset for *estimated accuracy (Es Acc), maximum accuracy scores (Max Acc)* and *average correlation scores (Corr)* of estimated order of significant IC pairs based on accuracy scores obtained from training data set and real order of IC pairs based on accuracy scores obtained from testing data set. Also we compared accuracy of our method with the *accuracy (Acc)* obtained from all significant IC pairs.

*Table 5 - 1: s*FNC performance for classification

***s*FNC**		**Best half of IC pairs**			**All IC pairs**
		Es Acc	**Max Acc**	**Corr**	**Acc**
LDC	**R**	66.12	73.45	0.42	60.5
	R+S	70.00	76.05	0.54	65.25
	R+S+M	72.90	77.25	0.53	66.45
NBC	**R**	72.75	76.72	0.47	61.8
	R+S	77.12	80.25	0.55	65.75
	R+S+M	80.10	83.37	0.6	69.25
***n*SVM**	**R**	64.05	74.85	0.36	62.6
	R+S	74.50	79.28	0.54	65.55
	R+S+M	80.93	84.45	0.62	65.75

Results showed that using all levels of the task hierarchy improve the accuracy scores significantly ($p < 0.05$, FDR corrected for multiple comparisons) compared to resting-state and resting state - sensory gating sensory loads. Also, we obtained improvement by using estimated best half of the significant IC pairs instead of using all significant IC pairs ($p < 0.05$, FDR corrected for multiple comparisons) for all hierarchy level. The best accuracy score was obtained by using *n*SVM algorithm with best half (estimated) of the features and all levels of the task hierarchy (R+S+M).

Also, we repeated the same experiment with significant IC pairs that were obtained from *d*FNC analysis. See Table 5 - 2. The results showed a similar pattern as *s*FNC but were overall less accurate. Moreover, accuracy results were not changed significantly when we added a new hierarchy level.

Table 5 - 2 : d*FNC performance for classification*

***s*FNC**		**Best half of IC pairs**			**All IC pairs**
		Es Acc	**Max Acc**	**Corr**	**Acc**
LDC	**R**	63.55	69.80	0.50	54.00
	R+S	63.60	70.40	0.52	53.35
	R+S+M	63.72	70.40	0.53	54.95
NBC	**R**	64.15	72.46	0.50	54.80
	R+S	64.67	73.13	0.52	56.07
	R+S+M	66.13	77.60	0.54	56.00
***n*SVM**	**R**	63.55	70.30	0.59	56.80
	R+S	64.43	70.50	0.61	55.85
	R+S+M	64.80	71.96	0.62	56.10

5.4. Discussion

In this chapter, we present the initial exploratory effort to investigate the classification performance of *s*FNC based method that identifies static differences in cortical networks by determining *significant* IC pairs among all measurable non-artifactual brain networks (Jafri MJ et al., 2008) across the task hierarchy. We detected 8 different networks in a multi task hierarchy with 95 subjects. Results showed that *s*FNC entails valuable information for individual prediction of SPs and can be used to improve the accuracy of classification.

The present study is consistent with previous results identifying group differences in networks commonly associated with altered FNC in SPs relative to HCs, including: the attentional network, auditory networks, default mode network, frontal network, sensory motor network, visual network, thalamus and cerebellar network.

Our results showed that using *significant* IC pairs from *s*FNC obtained from all hierarchical task structure improves the classification results compared to resting-state and resting state - sensory gating sensory loads. On the other hand, accuracy results obtained using *significant* IC pairs from *d*FNC were not changed significantly when we added a new hierarchy level. This may be due to the main cognitive differences between SPs and HCs being driven by one or more but not all levels of the hierarchy as shown in Figure 5 - 1. This shows that clearly some valuable information was ignored or missed when we used *d*FNC instead *s*FNC IC pairs. These results support our initial hypothesis.

A potential explanation for the difference in sensitivity of the sFNC versus the dFNC IC pairs is that the sFNC ICs identified trait-based characteristics that were stable across both task presentation as well as across time. As mentioned above, data collection was performed across multiple days to maintain the attentional state of participants throughout the multiple tasks. We suggest that the features that show persistent group differences across these conditions represent stable characteristics associated with schizophrenia and may be less sensitive to variations in symptoms across time.

In this study, we also used random sub-sampling validation to improve the generalizability and accuracy of the results. Results showed that such an approach helps to estimate accuracy success of each IC pair obtained from *s*FNC and *d*FNC. These estimation results were used to select the best half of the *significant* IC pairs. As reported in Table 5 - 1 and Table 5 - 2 this selection method significantly improved classification accuracy.

The task hierarchy defined in the current study was based on a study designed to assess function across the sensory to cognitive hierarchy by sampling differences at pre-attentional sensory levels, multisensory integration with motor response, and working memory performance. In future studies, we intend to explore aspects that may improve the classification accuracy scores with current methodology by including cognitive tasks with established pathophysiology in schizophrenia.

5.5. Conclusion

In conclusion, we used a novel method for prediction of SPs by using the best half of the *significant* IC pairs observed from *s*FNC which also showed *consistency* between SP and HC groups across levels of the task hierarchy. The results provided evidence that *s*FNC differences capture important information for classification that is missed in a *d*FNC approach. This suggests that *s*FNC differences provide valuable information that captures fundamental characteristics of brain network connectivity in schizophrenia across performance of a multi task hierarchy that increases the classification accuracy of the SPs and HCs and may help to design an objective biological marker-based diagnostic test for schizophrenia.

Chapter 6: Magnetoencephalographic and functional MRI connectomics in schizophrenia via intra- and inter- network connectivity

Examination of intrinsic functional connectivity using functional MRI (fMRI) has provided important findings regarding dysconnectivity in schizophrenia. Extending these results using a complementary neuroimaging modality (magnetoencephalography (MEG)) we present the first direct comparison of functional connectivity between schizophrenia patients and controls, using these two modalities combined. We develop a novel MEG approach for estimation of networks using MEG that incorporates spatial independent component analysis (ICA) and pairwise correlations between independent component timecourses, to estimate within- and among-network connectivity. This analysis enables group-level inference and testing of between-group differences. Resting state MEG and fMRI data were acquired from a large sample of healthy controls (n=45) and schizophrenia patients (n=46). Group spatial ICA was performed on fMRI and MEG data to extract intrinsic fMRI and MEG networks. Results: Similar, but not identical spatial independent components were detected for MEG and fMRI. Analysis of functional network connectivity (FNC) revealed a differential between-modalities pattern, with greater connectivity among occipital networks in fMRI and among frontal networks in MEG. Most importantly, significant differences between controls and patients were observed in both modalities. MEG FNC results in particular indicated dysfunctional hyperconnectivity within frontal and temporal networks in patients, while in fMRI FNC was always greater for controls than for patients. Results suggest that combining these two neuroimaging modalities reveals additional disease-relevant patterns of connectivity that were not detectable with fMRI or magnetoencephalography alone.

6.1. Introduction

More than 2000 neuroimaging papers examining the human “resting state” have been published since the first fMRI study (Biswal, B., et al., 1995; Calhoun VD et al., 2001b; Raichle ME, et al., 2001). Analysis of resting state data has yielded information across a wide range of topics including basic sensory processing (Shostak VI, 1968), tobacco and alcohol use (Brown BB, 1968), neurodegenerative diseases (Rosadini G, et al., 1974), and neuropsychiatric illnesses (Reeve A, et al., 1993). Resting state protocols are particularly advantageous for the study of disease states where patients may have difficulty responding or performing behavioral tasks due to compromised cognitive and/or physiological functions. Connectivity methods such as independent component analysis (ICA) are a set of powerful analysis techniques used to analyze

resting brain activity. Such approaches are particularly relevant to disease states such as schizophrenia, in which dysfunctional connectivity ("dysconnectivity") is hypothesized to underlie patient symptoms (Bullmore ET, et al., 1997; Stephan KE, et al., 2006). Because this dysconnectivity is thought to be driven by aberrant synaptic plasticity (Stephan KE, et al., 2009), characterizing functional connectivity may be essential to understanding the disorder.

Group independent component analysis (gICA) is an effective means of interrogating functional dysconnectivity in schizophrenia (Calhoun VD and Adali T, 2012). As typically applied to resting fMRI data, this technique identifies and reconstructs temporally-coherent, spatially-independent networks in groups of subjects (Calhoun VD et al., 2001a), where "networks" are defined as components identified via ICA. Spatial independence facilitates the comparison of topographies, or maps, between groups, while the property of temporal coherence permits the assessment of inter-regional, connectivity between spatially independent networks. However, exclusive reliance on fMRI to generate such networks may limit inference on dysconnectivity: Whilst the blood oxygenation-level dependent (BOLD) response measured by fMRI allows high spatial resolution maps, it is limited by being an indirect and slow physiological signal (Bandettini PA, et al., 1993; Kim SG, et al., 1997). Neural oscillatory activity, which comprises rhythmic electrical activity in cell assemblies, is thought to underlie BOLD responses. This occurs in the ~1-900Hz band; such rapid electrical signals cannot be assessed using fMRI but can be measured directly by techniques such as MEG (Cohen D., 1968), a noninvasive neuroimaging technique used to infer the cortical current distribution via assessment of the induced extra-cranial magnetic fields. Measurement of resting state brain activity using both fMRI and MEG, within a common sample of subjects, combines the strengths of each modality by allowing comparison of haemodynamic and electrophysiological effects. In this way we provide significant insight into functional connectivity, with special relevance for the study of schizophrenia and similar conditions. Significant progress towards integrating MEG and fMRI has been made in the past decade. MEG inverse solutions such now permit functional connectivity analysis in the same brain space as fMRI (Brookes MJ, et al., 2005). This approach has already been used to evaluate intrinsic connectivity networks (ICNs) in MEG in a similar way to that typically used in fMRI (Brookes MJ, et al., 2011a). Neural oscillations are implicated strongly in this approach and in particular, multiple studies (Brookes

MJ, et al., 2011b; Luckhoo H, et al., 2012) have now shown that assessment of temporal correlation between the amplitude envelopes of oscillatory activity facilitates elucidation of distributed network structure that bears reasonable resemblance to fMRI.

The purpose of the present chapter is to use both fMRI and band limited envelope correlation metrics in MEG to interrogate functional connectivity in the resting state in a sample of healthy normal volunteers and schizophrenia patients. Using methods based on group spatial ICA, for the first time we estimate networks from both MEG and fMRI and compare and contrast the networks and findings from the two modalities, with the hypotheses that 1) Patients and controls would differ significantly on both MEG and fMRI measures of among-network connectivity, called functional network connectivity (FNC), 2) MEG and fMRI spatial maps would show substantial overlap, and 3) Using both MEG and fMRI measures of among-network connectivity would show improvement to classification of schizophrenia patients.

6.2. Materials and Methods

6.2.1. Participants

This investigation combined existing data from 91 participants, 46 schizophrenia patients and 45 healthy controls. Informed consent was obtained from all participants according to institutional guidelines at the University of New Mexico Human Research Protections Office (HRPO). All participants were compensated for their participation. Patients with a diagnosis of schizophrenia or schizoaffective disorder were invited to participate. Each patient completed the Structured Clinical Interview for DSM-IV Axis I Disorders (First M, et al., 1997) for diagnostic confirmation and evaluation of co-morbidities. Patients with a history of neurological disorders including head trauma (loss of consciousness > 5 minutes), mental retardation, history of substance dependence, or active substance abuse (except for nicotine) within the past year were excluded, as were patients who were clinically unstable (e.g., in the previous month were discharged from the hospital or had any changes in their psychotropic medications). Stability was also monitored throughout the study to confirm that patients had no clinically meaningful symptom changes. All patients had a negative urine toxicology for drugs of abuse at the time of enrollment in the study. Patients were treated

with a variety of antipsychotic medications. The doses of antipsychotic medications were converted to olanzapine equivalents (Gardner DM et al., 2010) (see Table 6 - 1). Healthy controls were recruited from the same geographic location and completed the SCID – Non-Patient Edition to rule out Axis I conditions (First MB et al., 2002a). All participant smokers were instructed not to use tobacco during the two hours prior to each scan to minimize acute effects. This was confirmed via a breath carbon monoxide measure of less than 8 ppm. Each participant completed resting MEG and MRI scans. Scans were collected in counterbalanced order, with a median time between scans of approximately 22 days.

Table 6 - 1: Demographic information. A/A: American Indian/Alaska Native, PCE: Primary caregiver education, SCE: Secondary caregiver education

	Mean (SD)		***t or x^2***
	Schizophrenia (n=44)	**Control (n=47)**	***(p-value)***
Demographics			
Age	37. 28 (13.86)	35.18 (11.83)	0.78 (0.44)
Gender (M/F)	37/7	34/13	0.27 (0.78)
Ethnicity (H/NH)	23/21	26/21	
Race			
A/A	2	2	
Asian	2	0	
African American	1	4	
Pacific Islander	1	0	
White	38	41	
PCE	4.24 (2.11)	4.53 (1.18)	
SCE	4.72 (1.83)	4.72 (1.84)	
PANSS			
Positive	15.13 (5.136)		
Negative	15.15 (5.013)		
General	29.79 (8.108)		

Medications			
OLZ(mg/day)	14.02 (12.39)		

6.2.2. fMRI Data Acquisition

All fMRI data were collected on a 3-Tesla Siemens Trio scanner with a 12-channel radio frequency coil. High-resolution T1-weighted structural images were acquired with a five-echo MPRAGE sequence with TE = 1.64, 3.5, 5.36, 7.22, 9.08 ms, TR=2.53 s, TI=1.2 s, flip angle=7°, number of excitations=1, slice thickness=1 mm, field of view=256 mm, resolution=256×256. T2*-weighted functional images were acquired using a gradient-echo EPI sequence with TE=29 ms, TR=2 s, flip angle=75°, slice thickness=3.5 mm, slice gap=1.05 mm, field of view 240 mm, matrix size=64x64, voxel size=3.75 mm×3.75 mm×4.55 mm. Resting-state scans consisted of 149 volumes per run.

6.2.3. fMRI Data Preprocessing

An automated preprocessing pipeline and neuroinformatics system developed at MRN (Scott A, 2011) was used to preprocess the fMRI data. The first four volumes were discarded to remove T1 equilibration effects. Images were realigned and slice-timing correction was applied using the middle slice as the reference frame in the functional data pipeline. The data were then spatially normalized to the standard MNI space, resampled to 3×3× 3 mm voxels, and smoothed using a Gaussian kernel with a full-width at half-maximum (FWHM) of 10 mm. The preprocessed time series data were scaled to a mean of 100.

6.2.4. fMRI Group Spatial Independent Component Analysis (gsICA)

Following Allen et al. (Allen EA et al., 2011b), we performed a subject-specific data reduction PCA retaining 100 principal components (PC). In order to use memory more efficiently, group data reduction was performed using an EM-based PCA algorithm and C = 75 PCs were retained. The infomax algorithm (cf. 50) was used for GICA. It was performed using the GIFT Toolbox (http://mialab.mrn.org/software/gift/). We use this high model order ICA (number of components,

C = 75), since such models yield refined components that correspond to known anatomical and functional segmentation (Allen EA et al., 2011b). In order to estimate the reliability of the decomposition, the Infomax ICA algorithm was applied repeatedly via ICASSO (Himberg J, and Hyvarinen A., 2003) and resulting components were clustered.

6.2.5. fMRI Feature Identification

To identify non-artifactual components that contain features associated with resting state networks a combination of two methods was used (Allen EA et al., 2011b). In the first method we examined the power spectra with two criteria in mind: dynamic range and low frequency/high frequency ratio. See Figure 6 - 1. Dynamic range refers to the difference between the peak power and minimum power at frequencies to the right of the peak in the power spectra. Low frequency to high frequency power ratio is the ratio of the integral of spectral power below 0.10 Hz to the integral of power between 0.15 and 0.25 Hz. To verify the results, three expert reviewers evaluated the components for functional relevance. In this evaluation, if a component exhibited 1) peak activation in gray matter, 2) low spatial overlap with known vascular, ventricular, motion, and susceptibility artifacts, and 3) TCs dominated by low frequency fluctuations, it was classified as a non-artifactual component. Of the 75 components returned by the GICA, 39 were identified as BOLD-related component; see Figure 6 - 9.

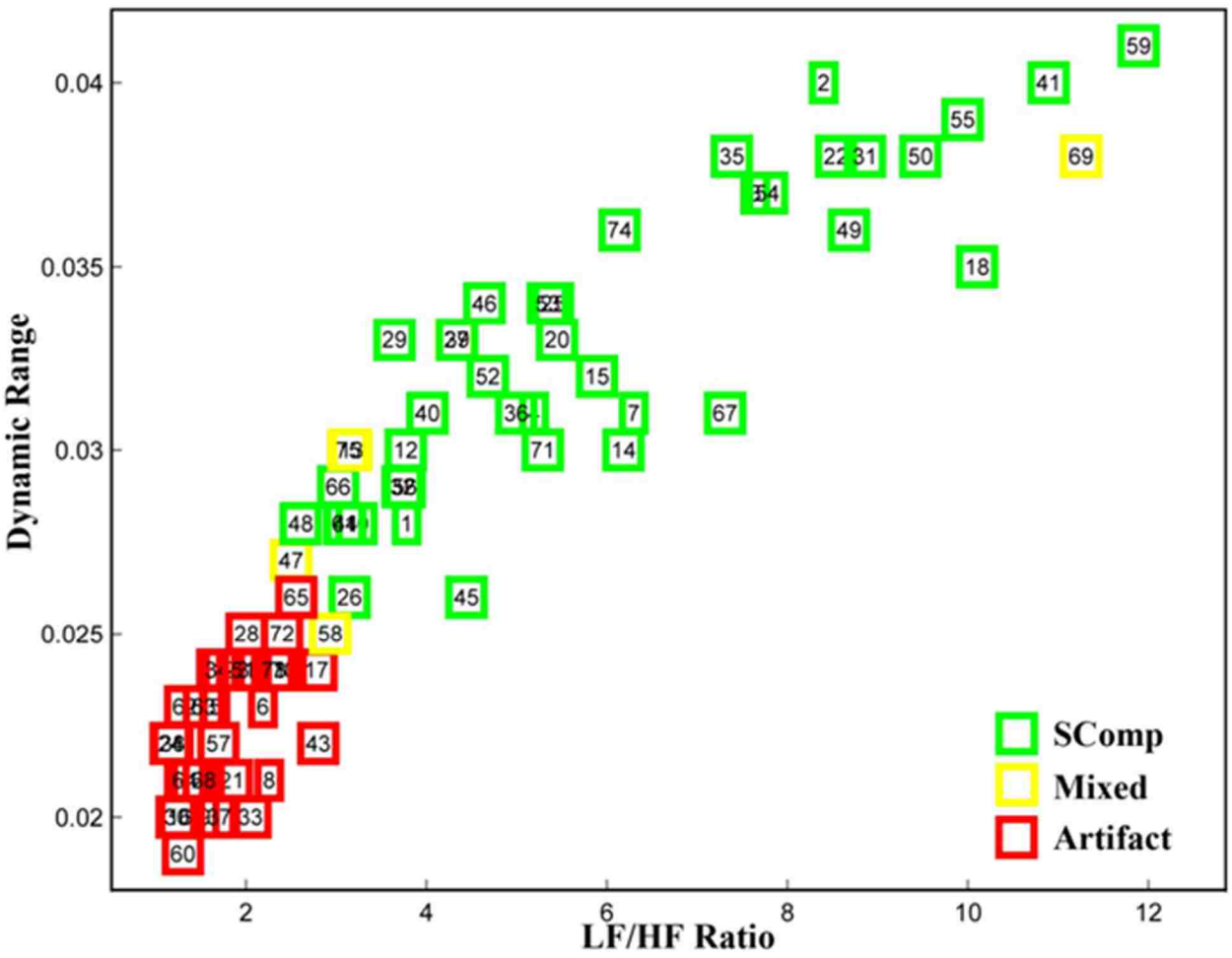

Figure 6 - 1: fMRI component quality measures

6.2.6. MEG data acquisition

MEG data were collected in a magnetically and electrically shielded room (VAC Series Ak3B, Vacuumschmelze GmbH) using a whole-cortex 306-channel MEG array (Elekta Neuromag^{TM}) at the Mind Research Network. Before positioning the participant in the MEG, four coils were affixed to the participant's head: two on the forehead and one behind each ear. These coils allow determination of the position of the participant's head relative to the position and orientation of the MEG sensors. Additional positioning data were collected using a head position device (Polhemus Fastrak) in order to permit co-localization of MEG activity with each participant's anatomical MRI. Two channels of electro-oculogram (EOG), one vertical and one horizontal, and one channel of electrocardiogram (ECG) were collected simultaneously with MEG. MEG data

were sampled at a rate of 1000 Hz, with a bandpass filter of 0.10 to 330 Hz. Head position was monitored continuously throughout the MEG session. Raw data were collected and stored. Participants were instructed to keep their eyes open and maintain fixation during the 6-minute scan to minimize occipital alpha rhythm (Cohen D., 1968).

6.2.7. MEG data preprocessing

Artifact removal, correction for head movement, and downsampling to 250 Hz were conducted offline using Elekta Maxfilter software, with 123 basis vectors, a spatiotemporal buffer of 10 s, and a correlation limit of 0.95. To facilitate comparison with previous research (Brookes MJ, et al., 2011b), data were bandpass filtered into four frequency ranges of interest: delta (1-4 Hz), theta (5-9 Hz), alpha (10-15 Hz), and beta (16-29 Hz). For classification purpose, we used all MEG frequencies.

6.2.8. MEG beamformer projection

Covariance matrices were generated independently for each subject and frequency band, using all recorded data. Covariance matrices were regularized using a value of 4 times the minimum singular value of the unregularized matrix. Voxels were placed on a regular 6-mm grid spanning the brain image. Source orientation at each voxel was based on a nonlinear search for maximum projected signal-to-noise ratio. The forward solution was based on a dipole model (Sarvas J, 1987) and a single-shell boundary element model (Hamalainen MS, and Sarvas J, 1989). Beamformer projection was performed separately for each subject and frequency range. After beamformer projection, source-space signals were normalized by an estimate of projected noise (Hall EL, et al., 2013) and transformed to standard (MNI) space using FLIRT in FSL. A Hilbert transform was applied to the time course at each voxel time to derive the analytic signal. The absolute value of this analytic signal was computed to yield the Hilbert envelope, an amplitude envelope of oscillatory power. The Hilbert envelope at each voxel was downsampled to an effective sampling rate of 1 Hz (Brookes MJ, et al., 2011b). Source space envelope data were smoothed spatially (6 mm at full-width half-maximum), and the voxel size was resampled to 3×3×3 mm to facilitate comparison with the fMRI data. Note that while strong and sustained correlations between brain

regions can lead to beamformer failure, this requires correlations that persist through 30-40% of the period analyzed (Hadjipapas A, et al., 2005), unlikely in resting data (Brookes MJ, et al., 2011b).

6.2.9. MEG Group Spatial Independent Component Analysis (gsICA)

Group spatial ICA was applied to the individual subject data using the GIFT toolbox. The gsICA approach was selected over group temporal ICA (gtICA) for two reasons: 1) Because components produced by gsICA are not temporally independent, relations among network timecourses can be evaluated; and 2) gtICA of participant timecourses carries the assumption of temporal consistency, limiting its utility in group analysis of resting data. Each frequency range was treated as a session in GIFT to permit analysis of each band, as well as the mean across bands. MEG ICA processing generally followed the procedures applied to the fMRI. Reduction steps were applied using principal component analysis. First, subject-specific data reduction was applied, retaining 100 principal components. Next, group level data reduction was applied to reduce the dataset to 75 principal components. Infomax ICA was applied 20 times in ICASSO and the resulting components were clustered. Spatial maps were generated by decomposing the mixed MEG timecourses to yield a set of spatially independent and temporally coherent networks. As with fMRI, FNC was computed as the zero-lag cross-correlations among reconstructed timecourses.

6.2.10. MEG feature identification

Consistent with standard practice (Allen EA et al., 2011b; Robinson S et al., 2009), component quality was assessed both qualitatively, to remove components situated in white matter and ventricles, and quantitatively, using assessments of dynamic range and the ratio of low-frequency to high-frequency power in each component. Components were separated into artifactual and non-artifactual components. Of the 75 components requested from the group ICA, 29 were retained as non-artifactual components; see Figure 6 - 2. In the present context these criteria, applied previously to fMRI ICA components (Allen EA et al., 2011b; Robinson S et al., 2009), also appeared to perform well for MEG.

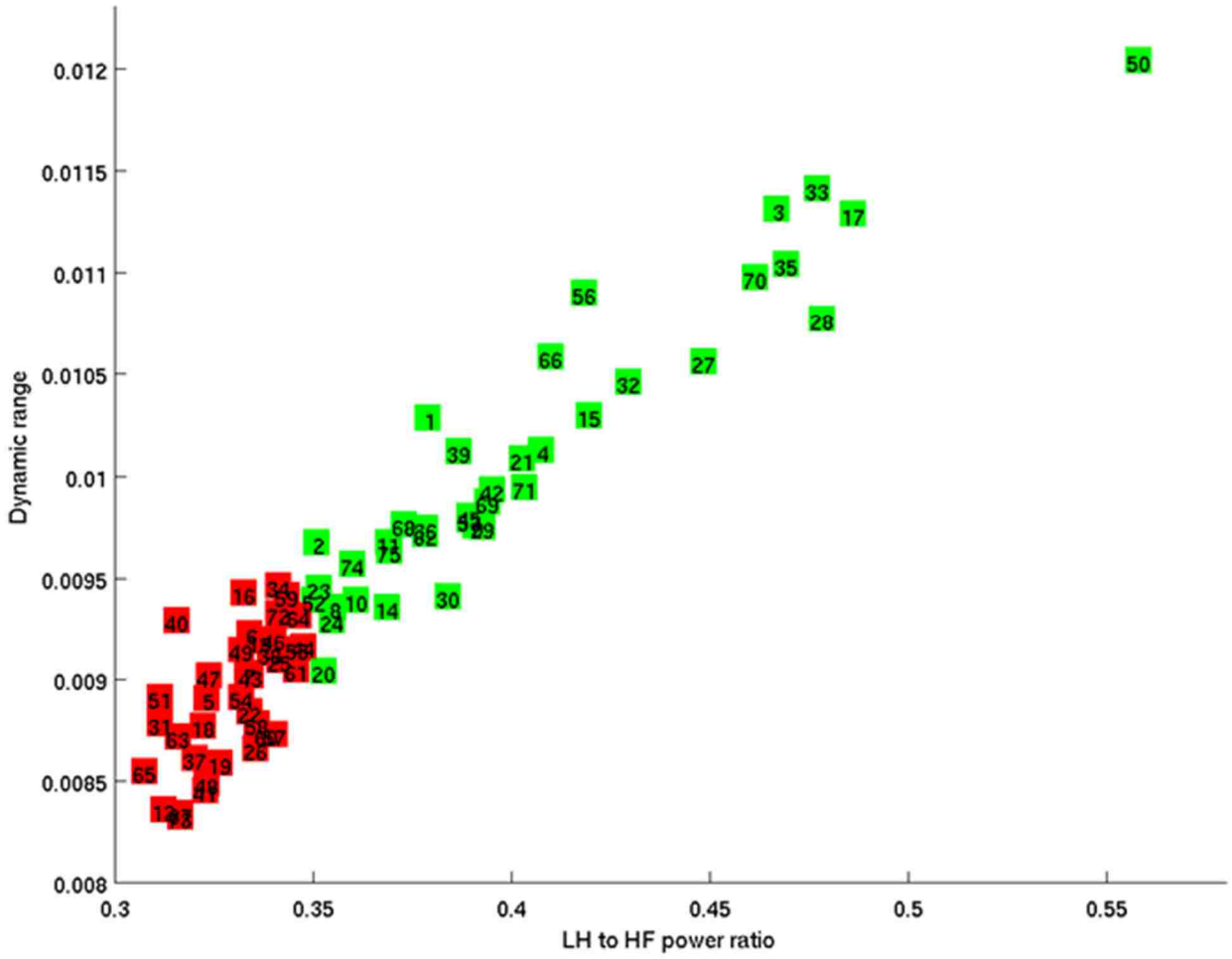

Figure 6 - 2: MEG component quality measures

6.2.11. Classification

Determining a reliable biological feature for a mental disorder is an important step for developing a more accurate and reliable framework for diagnosis, and ultimately treatment (Keshavan, M. S., et al., 2013). Resting-state fMRI connectivity has been used in determining the differences based on biological features of mental disorders including schizophrenia. (Arbabshirani MR et al., 2013b; Barnaly Rashid et al., 2014, 2015; Sakoğlu, Ünal et al., 2010). Using methods based on group spatial ICA, we estimate networks from both MEG and fMRI and findings from these two

modalities, with the hypotheses that using both MEG and fMRI measures of among-network connectivity would show improvement to classification of schizophrenia patients. In this chapter, we used dynamic FNC to determine reliable differences based on dynamic FNC differences of schizophrenia.

6.2.11.1. Dynamic Functional Network Connectivity and Clustering

Recent studies (Allen EA et al., 2012; Barnaly Rashid et al., 2014, 2015; Damaraju, E., et al., 2014; Sakoğlu, Ünal et al., 2010) show that connectivity dynamics can capture repetitive patterns of interactions among intrinsic networks during a rest or task related experiments that cannot be detected with FNC (static functional connectivity analyses). These repetitive patterns of interactions contain valuable information for individual prediction of schizophrenia patients. Such information is useful for training and replicates in testing.

In order to obtain connectivity dynamics we follow the steps of a previous study (Damaraju, E., et al., 2014). First, we computed correlations between non-artifactual components' time courses using a sliding window approach with a rectangular window of 25 TR (in steps of 1TR) convolved with Gaussian of sigma 3 TRs to obtain tapering along the edges. To characterize the full covariance matrix, we estimated covariance from regularized inverse covariance matrix (ICOV) (Smith S.M., et al., 2011) by using the graphical LASSO framework (Friedman J., et al., 2008). Then we placed a penalty on the L1 norm of the precision matrix to enforce sparsity. The regularization parameter was optimized for each subject separately by evaluating the log-likelihood of unseen data of the subject in a cross-validation framework. Second, we selected group centrotypes by using *k*-means clustering algorithm from all of the dynamic windowed FNC matrices for each group. Then for each FNC time point, we regressed out the dynamic FNC matrix against these $2\times k$ states and obtained the corresponding beta coefficients. We used the mean of these beta coefficients and finalized $2\times k$ features for each subject for classification. See Figure 6 - 3 for schematic description of dynamic FNC, clustering and regression of dynamic FNC matrices.

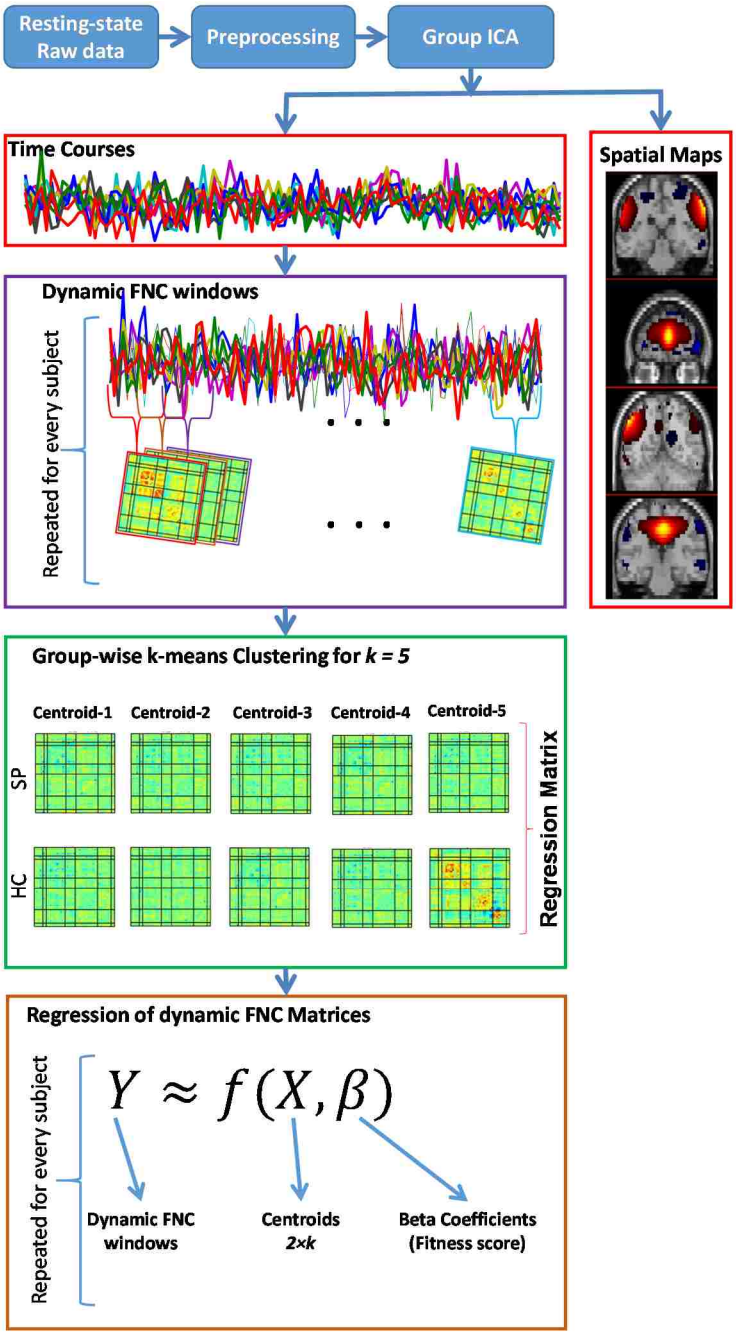

Figure 6 - 3: Schematic description of dynamic FNC, clustering and regression of dynamic FNC matrices

In order to compute the most efficient cluster number, we used the elbow criterion of the cluster validity index, which is computed as the ratio between within-cluster distances to between-cluster distance.

6.3. Results

6.3.1. Analytic approach

Resting MEG and fMRI data were acquired in 45 healthy controls and 46 schizophrenia patients. MEG data were source space projected using beamforming, enabling subsequent processing in brain space, equivalent to fMRI. Source space MEG and fMRI data were decomposed based upon a standard group spatial ICA analysis. This approach produces two form of output: 1) network spatial maps and 2) network timecourses. Spatial maps reflect within-network connectivity (i.e., the extent to which the regions in a network tend to co-activate), while timecourses are used to assess among-network connectivity. Timecourses were Hilbert transformed in order to compute the analytic signal. The absolute analytic signal was then calculated generating an amplitude envelope timecourse. Functional network connectivity (FNC) among networks was then derived which here we define to be pairwise zero-lag correlation between amplitude envelope timecourses from spatially independent networks.

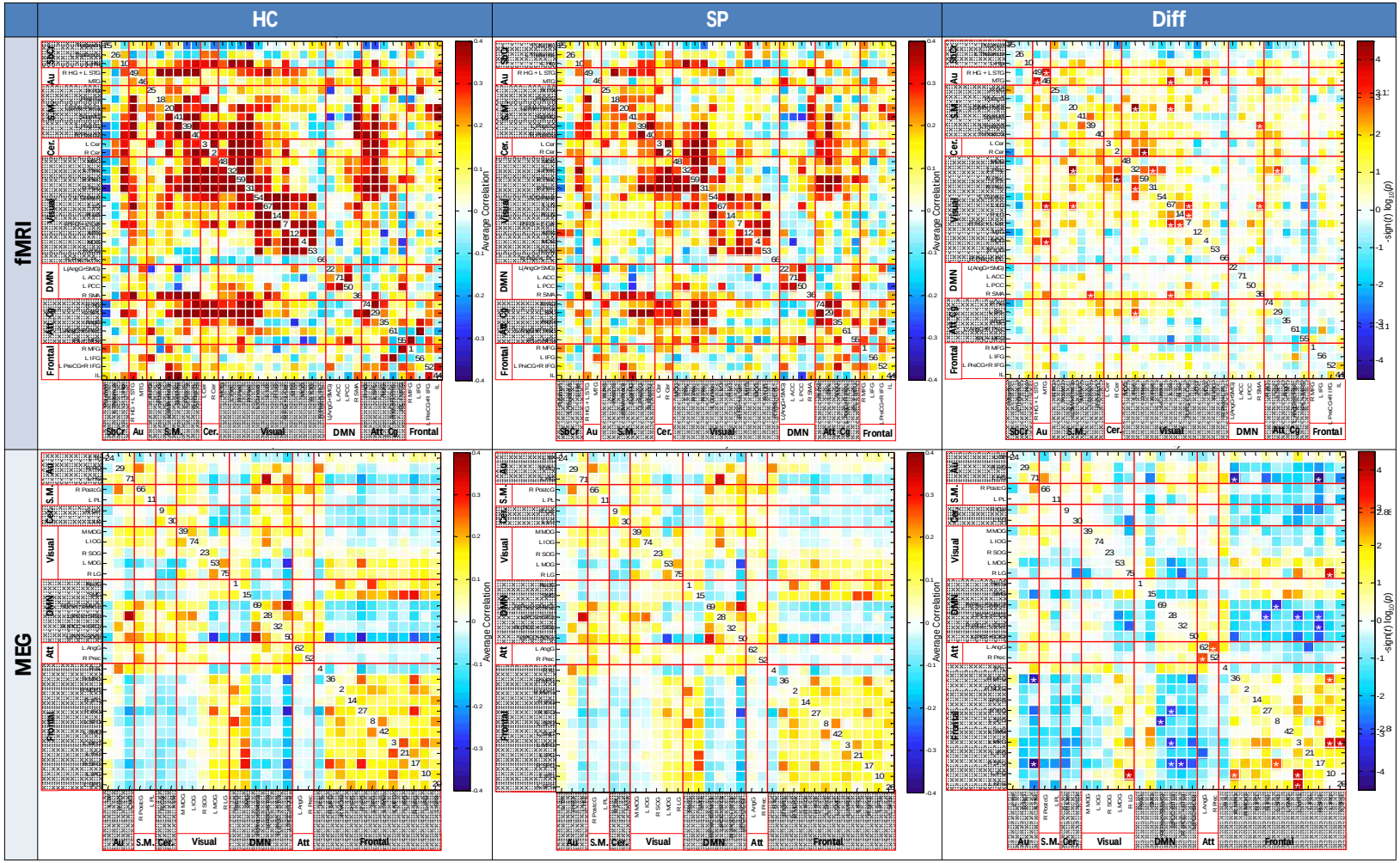

Figure 6 - 4: Functional network connectivity (FNC) for fMRI (top) and concatenation of MEG frequencies (bottom), for healthy controls (left column), Schizophrenia patients (center column), and FDR-corrected group differences (right column). ICA component numbers are on the diagonal.

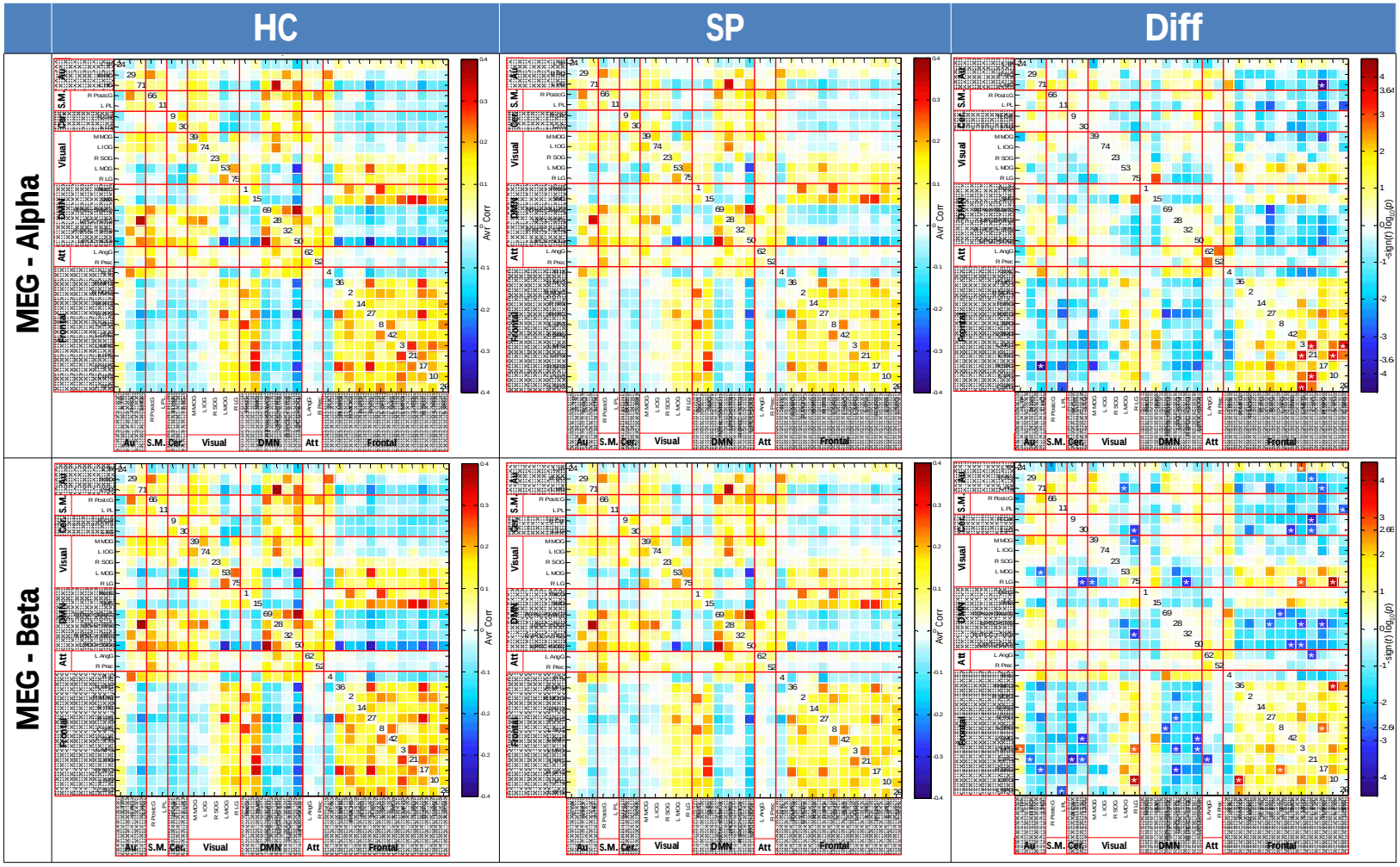

Figure 6 - 5: Functional network connectivity of MEG-Alpha (top) and Beta frequencies (bottom), for healthy controls (left column), Schizophrenia patients (center column), and FDR-corrected group differences (right column). ICA component numbers are on the diagonal.

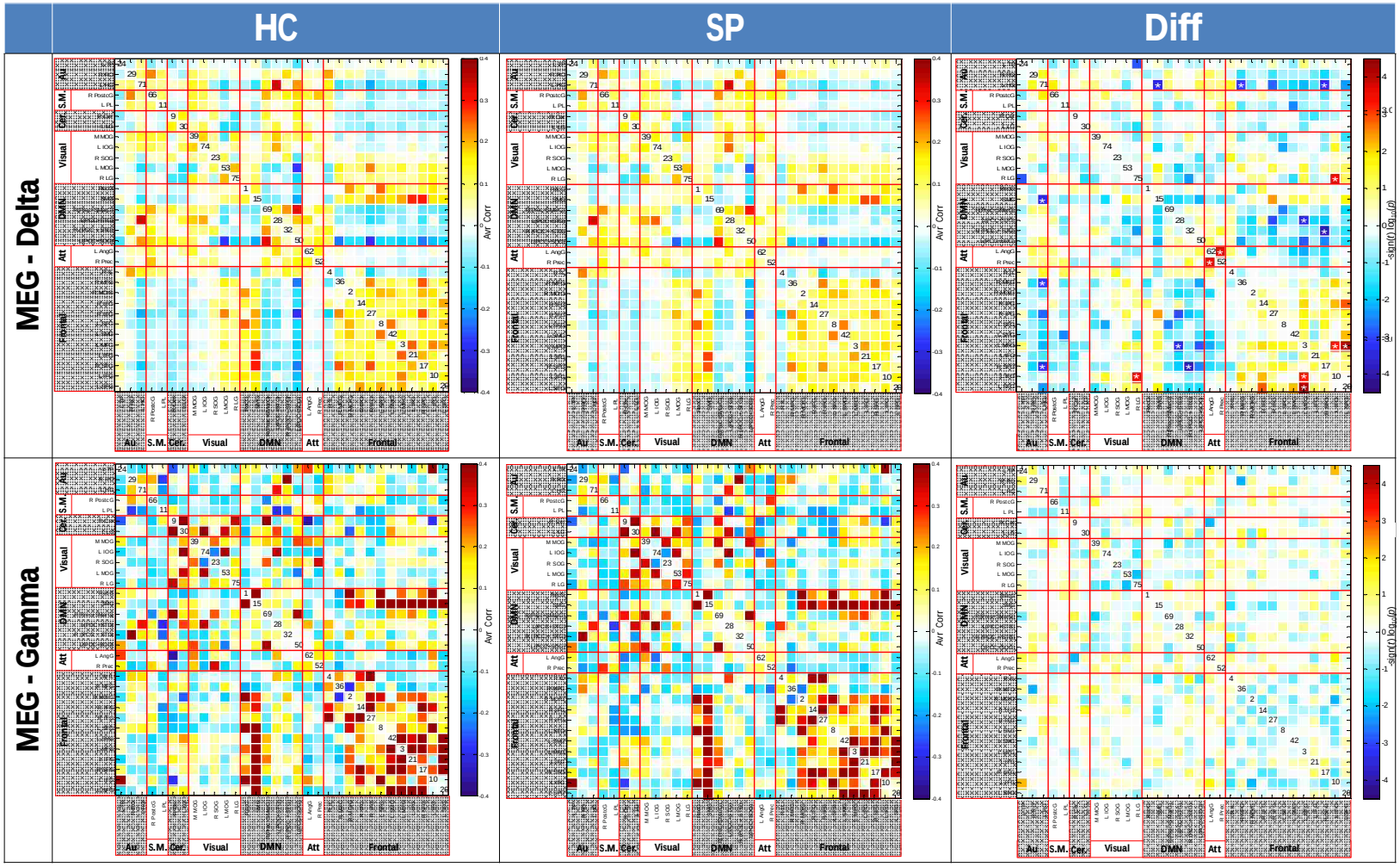

Figure 6 - 6: Functional network connectivity of MEG-Delta (top) and Gamma frequencies (bottom), for healthy controls (left column), Schizophrenia patients (center column), and FDR-corrected group differences (right column). ICA component numbers are on the diagonal.

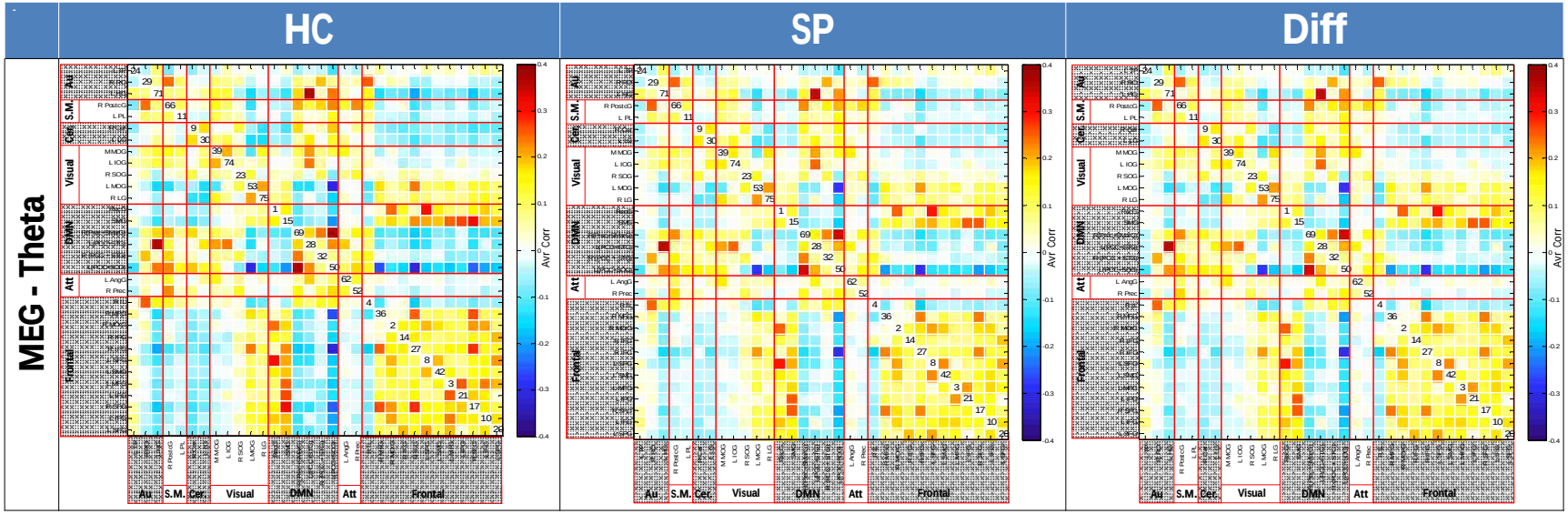

Figure 6 - 7: Functional network connectivity of MEG-Theta for healthy controls (left column), Schizophrenia patients (center column), and FDR-corrected group differences (right column). ICA component numbers are on the diagonal.

6.3.2. Functional network connectivity (FNC)

We hypothesized that patients and controls would differ significantly on MEG and fMRI measures of FNC. FNC was assessed for each modality. Note that in MEG, we analyse FNC using a mean timecourse encompassing all frequencies. Generally, we observed greater FNC in visual networks for fMRI components, and greater FNC in frontal networks for MEG components. Significant FNC differences between patients and controls are shown in Figure 6 - 4 and Figure 6 - 8. The images show the spatial signature of the FNC differences observed between groups. Those differences were evaluated using *t*-tests with a false discovery rate (FDR) correction. All results are rendered on a white matter surface. Within- and across-frequency FNC was also assessed for MEG in order to facilitate comparison with the fMRI networks. Figure 6 - 5 shows the detailed of MEG FNC matrices for alpha - beta frequencies, Figure 6 - 6 shows the detailed of MEG FNC matrices for delta - theta frequencies and Figure 6 - 7 shows the detailed of MEG FNC matrices for theta frequency.

For fMRI, most group differences in among-network connectivity were detected among temporal-occipital and frontal-occipital networks, and within the occipital networks (see Figure 6 - 4 and Figure 6 - 8). For MEG, fewer FNC group differences were detected, with the majority in frontal-DMN networks and within the frontal networks. While fMRI FNC revealed no hyperconnectivity in SZ, approximately half of the MEG FNC relationships indicated higher connectivity in SZ than in HC, suggesting dysfunctional network connectivity in SZ networks revealed with MEG.

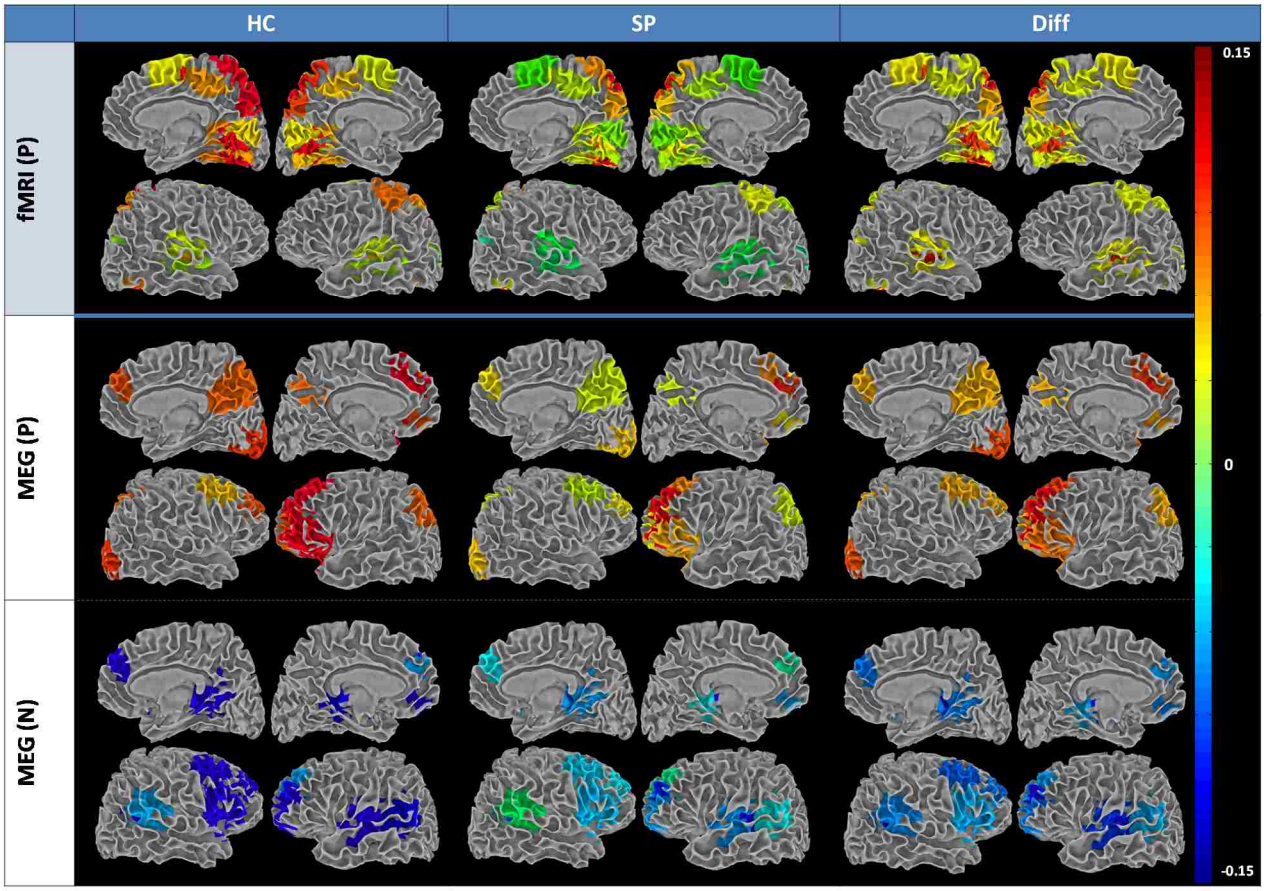

Figure 6 - 8: Summary of functional network connectivity (FNC) group averages and group differences for fMRI and MEG rendered on white matter surface. Only those regions involved in significant group differences are included. For networks showing a significant group difference, the rendered values represent the weighted sum of the five strongest correlations with label network.

6.3.3. Spatial maps

We hypothesized that there would be considerable spatial overlap in network maps between MEG and fMRI. Spatial maps for each modality were assessed across groups via one-sample *t*-tests of back-reconstructed subject maps. Identified networks included temporal, sensorimotor, parietal, occipital, frontal, subcortical, and DMN regions (See Figure 6 - 9).

Commonalities and differences can be seen across participant groups for fMRI and MEG. Spatial overlap between MEG and fMRI was initially assessed via visual inspection and subsequently verified quantitatively using spatial correlation. Individual subject-level spatial correlation among these optimal matches ranged from 0.12 to 0.50, with a mean of 0.31. Substantial overlap was detected across multiple networks including DMN as well as frontal, parietal, temporal, and occipital regions (See Figure 6 - 10).

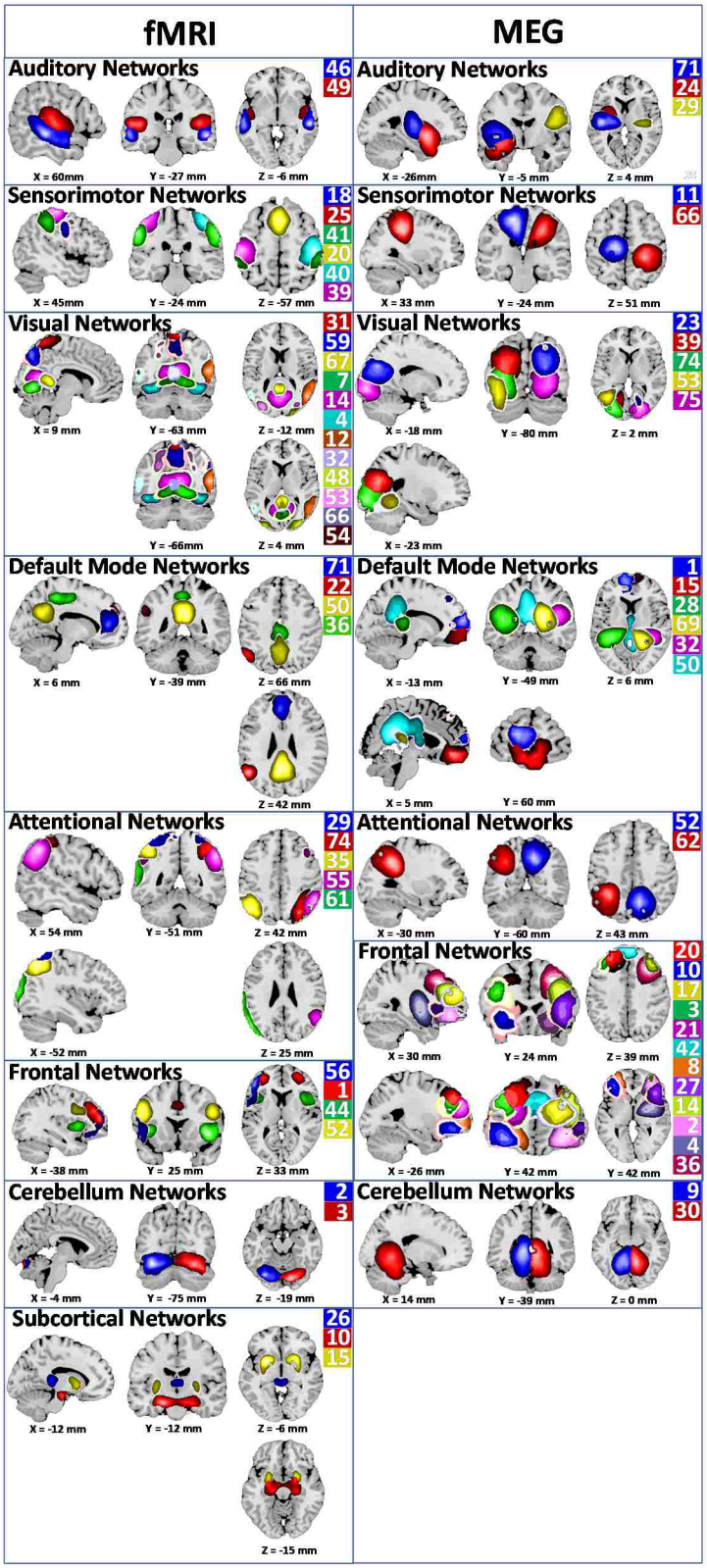

Figure 6 - 9: fMRI (left) and MEG (right) network spatial maps

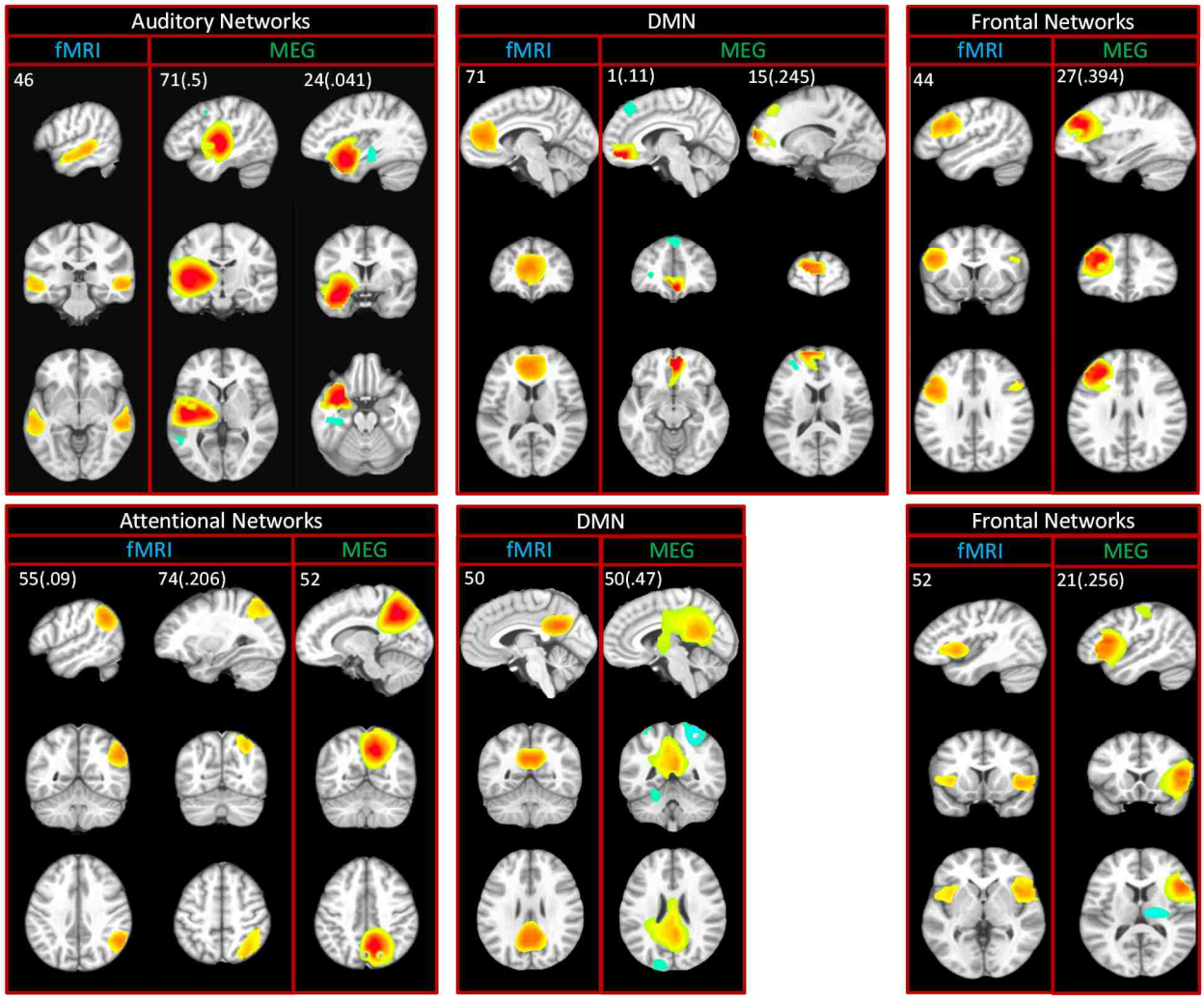

Figure 6 - 10: Spatial overlap in spatial maps detected using MEG and fMRI.

6.3.4. Multi-Model Classification

We evaluated the performance improvement of classification based on dynamic FNC and combination estimated networks from both MEG and fMRI methods. Our main focus was to extract reliable features from the dynamic FNC matrices and combine these features to perform the best classification results.

First, we used fMRI dynamic FNC matrixes and MEG dynamic FNC matrixes separately (for each frequency) for classification (See Table 6 - 2) then we combined fMRI (subject × time × FNC = 91 × 119 × 703) and MEG (frequency × subject × time × FNC = 5 × 91 × 270 × 496) dynamic FNC matrixes as a data set for classification (See Table 6 - 3). And we compared results to show the improvement of combining fMRI and MEG methods for classification.

We used leave-one-out cross validation method. One subject for testing and the rest of the data (90 subjects) were used as a training data set. And this process is repeated for each subject. In each cross-validation run, we obtained 5 cluster centroids for each group and regressed out the dynamic FNC matrix against these 10 centroids (5 centroids for each group) and computed the corresponding beta coefficients for all dynamic FNC for each subject. Then, we used the mean of these beta coefficients across the subjects and finalized 10 features for each subject for classification.

Table 6 - 2: Classification accuracy obtained from fMRI data, MEG data for each frequency and combination of all MEG data frequencies by using majority voting method

	fMRI	**MEG Alpha**	**MEG Beta**	**MEG Delta**	**MEG Gamma**	**MEG Theta**	**MEG MJ Voting**
NBC	82.42%	65.93%	71.43%	71.43%	51.65%	53.85%	65.93%
nSVM	83.52%	69.23%	72.53%	69.23%	51.65%	58.24%	69.23%
LDF	82.42%	65.93%	68.13%	71.43%	52.75%	53.85%	65.93%

We performed leave one out method with three well known classification algorithms; linear discriminant classifier (LDC), Naïve Bayes classifier (NBC) (Duda RO, et al., 2001) and non-

linear SVM (nSVM) with Gaussian radial bases function kernel (Burges, C. J. C., 1998) to test the hypothesis.

Table 6 - 2 reports the classification results that were obtained from fMRI data, MEG data for each frequency and combination of all MEG data frequencies by using majority voting method. Results show that the classification accuracy obtained from fMRI data (nSVM - 83.52%) provides better classification performance than MEG data for all frequencies and combination of all MEG data frequencies by using majority voting method. Comparison of internal MEG frequencies shows that beta (nSVM– 72.53%) frequency has better performance than other frequencies and combination of all MEG data frequencies. Similarly, FDR-corrected group differences of MEG-beta and MEG – delta frequencies show more significant differences than other frequencies.

Table 6 - 3: Classification accuracy obtained from the combination of fMRI data and MEG data for each frequency and the combination of all by using majority voting method.

	fMRI	**fMRI**	**fMRI**	**fMRI**	**fMRI**	
	MEG Alpha	**MEG Beta**	**MEG Delta**	**MEG Gamma**	**MEG Theta**	**MJ Voting**
NBC	83.52%	87.91%	86.81%	83.52%	85.71%	90.11%
nSVM	82.42%	85.71%	84.62%	82.42%	81.32%	87.91%
LDF	82.42%	83.52%	83.52%	82.42%	83.52%	85.71%

Table 6 - 3 summarized the classification accuracy obtained from the combination of fMRI data and MEG data for each frequency and the combination of all by using majority voting method. Combination of features obtained from dynamic FNC of fMRI and MEG-Beta frequency provided better results (NBC – 87.91%) than other frequencies. Best performance is provided by the combination of all by using majority voting method (NBC – 90.11%).

We repeated the clustering method by using different distance functions such as Euclidian, correlation, cosine similarities. We did not find any performance differences.

6.4. Discussion

FNC group differences evident for MEG shows greater functional connectivity across widespread frontal and temporal regions. Three interesting network patterns are seen in the MEG FNC results (See Figure 6 - 4 and Figure 6 - 8). First, within the DMN we see that the dorsal anterior cingulate/superior frontal region appears hyperconnected for SZ compared to HC (blue regions) while posterior cingulate/precuneus is hypoconnected in SZ (red/orange regions). For HC, bilateral posterior cingulate is hyperconnected with an adjacent parietal component. Other recent studies of schizophrenia similarly showed hyperconnectivity within anterior cingulate for SZ (Jafri MJ et al., 2008; Skudlarski P, et al., 2010). Hyperconnectivity has also been detected within one subcomponent of the DMN (i.e., anterior cingulate and portions of posterior cingulate) while the other subcomponent of the DMN revealed hypoconnectivity (bilateral parietal and dorsolateral prefrontal) for SZ (Skudlarski P, et al., 2010). These investigators identified the posterior cingulate as being the focus of decoupling found between anatomical (DTI) and functional (fMRI) connectivity. They emphasized that the DMN should not be viewed as a single unit; it is composed of substructures that all contribute to resting state activation but vary substantially in connectivity patterns.

For MEG, we see widespread hyperconnectivity between perisylvian and frontal regions (see Figure 6 - 4 and Figure 6 - 8) which closely resemble that seen in fMRI resting state of SZ patients who reportedly experience auditory hallucinations (Diederen KMJ, et al., 2013; Sommer IE, et al., 2012). These regions include bilateral superior temporal gyri (i.e., auditory cortex) and middle temporal gyrus, along with supramarginal gyrus, and the right hemisphere homologue of Wernicke's area (Jardri R, et al., 2010; Sommer IE, et al., 2012). Stephen and colleagues (Stone DB, et al., 2014) have examined multisensory integration in SZ and find that SZ benefit from multisensory (auditory/visual) integration more than do HC. They relate these results to the "high noise" theory where increased activity during rest is attributed to impaired responsiveness to external stimuli. This widespread hyperconnectivity involving auditory and speech-perception regions may provide important support for the "high noise" theory; this widespread network appears to be hyperconnected for SZ, compared to HC, during normal resting state. Therefore, additional stimulus intensity and/or additional attention resources may be required for SZ to direct

their attention to the external environment. In support of this interpretation, the attentional parietal regions are hyperconnected in HC. These patterns of functional connectivity lead us to speculate that SZ are directed inward more during resting state while HC are directed toward the external environment, ready to respond. Altogether our results indicate that multimodal methods are essential to understanding the mechanisms of inter-regional brain connectivity (Brookes MJ, et al., 2011b).

Recent meta-analytic work on fMRI of the resting state in schizophrenia has revealed that schizophrenia patients tend to show hyperactivation bilaterally in lingual gyrus and broad hypoactivity elsewhere, with decreases in resting state activity observed in VMPFC, left hippocampus, PCC, and precuneus (Kühn S, and Gallinat J., 2013). Other implicated regions with lower connectivity include paracingulate cortex, bilateral thalamus, fusiform, left caudate, and left thalamus (Argyelan M, et al., 2013), with greater hypoactivity generally related to worse functioning. Larger studies of functional network connectivity in schizophrenia have indicated dysfunction across a range of networks, with schizophrenia-specific deficits in midbrain/cerebellar and fronto-temporal paralimbic networks (Khadka S, et al., 2013). Our data suggest that patterns of connectivity involving hippocampus, fusiform, and middle frontal regions are particularly relevant to the level of functioning within the sample patient group, a finding consistent with the observed patient-control differences in frontal and temporal networks. Notably, in each neuroimaging modality we examined contributed both common and unique findings.

Our MEG analysis of within-frequency FNC showed multiple group differences in inter-regional connectivity in the beta (16-29 Hz) range (See Figure 6 - 5), particularly in the frontal-cerebellar, frontal-DMN, and frontal-auditory networks. Nearly all FNC group differences in the beta range suggested hyperconnectivity in patients. Beta band has previously been implicated in long-range cortical synchrony (Stein A von, et al., 1999; Tallon-Baudry C, et al., 2004; Thatcher RW, et al., 2008), notably in visual processing (Sehatpour P, et al., 2008) and working memory (Piantoni G, et al., in press) networks observed in the present data. Consistent with the present study, research in schizophrenia has indicated abnormal synchronization in the beta and gamma ranges (Uhlhaas PJ, and Singer W, 2010, 2011), particularly in the beta band (Siebenhühner F, et al., 2013).

Synchrony between the two hippocampi, regions with particular relevance for schizophrenia (Hanlon FM, et al., 2011, 2012), has also been linked to the beta band (Lee H, et al., 2014).

What do we learn about functional connectivity in schizophrenia from this dual modality study? By combining both MEG and fMRI we are able to interrogate both network structure (maps) and network dynamics (FNC) in schizophrenia, revealing patterns of connectivity impossible to detect with either modality alone. In particular, MEG appeared more sensitive to hyperconnectivity in frontal and temporal networks among patients. On the whole, our resting fMRI FNC findings converge with the schizophrenia literature, which reports hypoactivation across multiple regions, including the prefrontal cortex (Kühn S, and Gallinat J., 2013). However, prefrontal FNC with MEG was increased in our patient group. This suggests abnormally increased synchronous firing from neuronal populations in prefrontal networks in our chronically ill schizophrenia subjects. Whether these hyper-synchronous networks underlie core deficits of the illness or represent compensation to overcome other primary functional defects, we cannot say. However, we did not find correlations with positive symptoms, which have been reported to be associated with fMRI functional hyper-connectivity (Ford JM, et al., 2014). Our results suggesting hyper-synchronous prefrontal networks are consistent with the dysconnectivity model of schizophrenia (Stephan KE, et al., 2006) and suggest that these represent a deficit at the synaptic/neuronal level and not just in the coupling of vascular/neuronal function.

One recent study also applied beamformer analysis and group spatial ICA to resting MEG data collected from a very small sample (n=9) of healthy volunteers (Ramkumar P, et al., in press). This work varied somewhat from other recent studies (Brookes MJ, et al., 2005, 2011b) in that a Hilbert transform was not applied; instead, Fourier-transformed sensor data were inverted using a cortically-constrained minimum norm solution. This work also did not directly assess component quality; instead, two MEG scans were performed for each subject, and components that emerged in both scan sessions were retained. In addition, fMRI data were not collected, prohibiting any direct comparisons. These methods indicated similar spatial patterns and detected both within- and cross-frequency networks, complementing the present effort.

In the present study, the combination of data from multiple modalities, collected at different times, conveys additional confidence in our results. For instance, one critique of resting fMRI data from populations with mental illness or disease is that patients may respond differentially to auditory scanner noise (Skouras S, et al., 2013). We observed bilateral temporal components in fMRI, which has substantial background noise during scans, and unilateral temporal components in MEG, which is silent, providing some support for this critique. Similarly, schizophrenia patients have well-known autonomic nervous system dysregulation (Bär K-J, et al., 2007; Rachow T, et al., 2011; Toichi M, et al., 1999) related to variability in heart rate and respiration (Paterson AS, 1935; Whitehorn JC, and Richter H, 1937; Wittkower E, 1934) which can directly affect the BOLD response (Cohen ER, et al., 2002). However, the electromagnetic signal detected by MEG is less affected, particularly when the cardiac signal has been removed as in the present study. Overlapping MEG-fMRI components can reasonably be assumed to be free of such influences, revealing only the underlying dysregulation. Finally, eye movements also differ between controls and patients (Clementz BA, and Sweeney JA, 1990), again directly affecting the BOLD signal, where changes in the flow of vitreous humor during eye movement increase signal variance from nearby regions (Beauchamp MS, 2003). The corneo-retinal potential can affect electrophysiological signals (Kolder H, and North AW, 1966), but generally would appear as a source between the eyes. Frontal and occipital components that are in common for MEG and fMRI are arguably free of such modality-specific artifacts.

Also, our results provided evidence that the combination of fMRI and MEG modalities captures important information for classification that is missed by using only one modality. This suggests that the combination of these two methods provides valuable information that captures fundamental characteristics of brain network connectivity in schizophrenia. These results may help to design an objective biological marker-based diagnostic test for schizophrenia.

6.5. Conclusion

The present study employed a novel approach to estimate intrinsic connectivity networks from group spatial ICA of fMRI and MEG data to evaluate spatial patterns and functional connectivity in a sample of schizophrenia patients and healthy volunteers. This is the first study to use group

spatial ICA with resting MEG data, and also the first to apply these methods to directly compare a patient population to healthy volunteers. We observed substantial spatial overlap in multiple intrinsic connectivity networks as assessed using spatial correlation. We also observed intermodality differences in functional network connectivity of ICNs, with more instances of high frontal FNC for MEG and high occipital FNC for fMRI. In addition, while group differences in network spatial topography were observed primarily in frontal regions for fMRI, in MEG these differences were observed more broadly in frontal and temporal networks. The results suggest hyper-synchronous prefrontal networks in schizophrenia with deficits at the synaptic/neuronal level and not merely in neurovascular coupling. The combination of data from MEG and fMRI, collected on different days, also allows us to rule out multiple alternative explanations for the observed results, including scanner noise, artefacts from autonomic nervous system activity, motion, and eye movements. Also, the combination of data from MEG and fMRI increase the discrimination of schizophrenia patients from healthy controls. Results suggest that the application of group spatial ICA to multimodal neuroimaging using MEG and fMRI provides important information about complex mental illnesses such as schizophrenia that would have been missed otherwise.

Chapter 7: Conclusion and Future Works

7.1. Conclusion

In this doctoral dissertation, we developed and presented machine learning and data mining algorithms to detect abnormal functional network connectivity patterns of patients with schizophrenia and distinguish them from healthy controls using functional network correlation data based on resting-state and task hierarchy fMRI data, time series of functional networks during the resting state fMRI and functional network correlation data obtained with fMRI and MEG methods.

In this thesis, we introduced a shapelet algorithm that can be used for individual prediction of schizophrenia patients by using multi-dimensional time series in chapter - 3. We showed that an ensemble of shapelet-based decision trees on individual dimensions work better than shapelets defined over multiple dimensions. Generating a shapelet ensemble for multi-dimensional time series is computationally expensive. Most of the existing techniques *prune* shapelet candidates for speed. In chapter 3, we proposed a novel technique for shapelet discovery that *evaluates* remaining candidates efficiently. Our algorithm uses a multi-length approximate index for time series data to efficiently find the nearest neighbors of the candidate shapelets. We employed a simple skipping technique for additional candidate pruning and a voting based technique to improve accuracy while retaining interpretability. Not only did we find a significant speed increase, our techniques enabled us to efficiently discover shapelets on datasets with multi-dimensional and long time series such as hours of brain activity recordings. We demonstrated our approach on a biomedical dataset and found significant differences between patients with schizophrenia and healthy controls.

Furthermore, we introduced a novel approach which identified both stable (static effects) and state-based differences (dynamic effects) in brain connectivity. Results of chapter-4 provided a better understanding of how individuals' reactions to simple sensory stimuli are conditioned by the context within which they are presented. Our findings suggest that not all group differences observed during rest are detectable in other cognitive states. In addition, the stable differences of heightened connectivity between multiple brain areas with the thalamus across tasks underscore the importance of the thalamus as a gateway to sensory input and provide new insight into schizophrenia.

We proposed a framework for classification of schizophrenia patients and healthy control subjects based on functional network component pairs which show consistency between patients and controls across levels of the resting-state data and task hierarchy in chapter 4. Our results showed that these functional network components as a function of task contain valuable information for individual prediction of schizophrenia patients. Such information is useful for training and replicates in testing. Performance was improved significantly (up to ~20%) relative to a single FNC (resting-state) measure.

We developed a novel MEG approach for estimation of networks using MEG in chapter 5 that incorporates spatial independent component analysis (ICA) and pairwise correlations between independent component timecourses, to estimate within- and among-network connectivity. This analysis enables group-level inference and testing of between-group differences. Resting state MEG and fMRI data were acquired from a large sample of healthy controls (n=45) and schizophrenia patients (n=46). Group spatial ICA was performed on fMRI and MEG data to extract intrinsic fMRI and MEG networks. Results: Similar, but not identical spatial independent components were detected for MEG and fMRI. Analysis of functional network connectivity (FNC) revealed a differential between-modality patterns, with greater connectivity among occipital networks in fMRI and among frontal networks in MEG. Most importantly, significant differences between controls and patients were observed in both modalities. MEG FNC results in particular indicated dysfunctional hyperconnectivity within frontal and temporal networks in patients, while in fMRI FNC was always greater for controls than for patients. Results suggest that combining these two neuroimaging modalities reveals additional disease-relevant patterns of connectivity that were not detectable with fMRI or MEG alone.

It is also important to remark that combining these two neuroimaging modalities provides better classification scores that were not achievable with fMRI or MEG alone. This suggests that the combination of these two methods provides valuable information that captures fundamental characteristics of brain network connectivity in schizophrenia.

The published results (Cetin, MS., et al., 2014, 2015a, 2015b, 2015c; Houck JM, Cetin, et al., 2015), presented in this dissertation showed considerable improvement over existing methods and have provided new and important biomarkers. We sincerely hope that this doctoral work contributes to detecting abnormal functional network connectivity patterns of patients with schizophrenia and distinguishes them from healthy controls and decodes the mysteries related to the human brain as the most complex self-sustaining system.

7.2. Future work

Future work and suggestions were mentioned individually at the end of the relevant chapters (chapter – 3, 4, 5, and 6). In this section, I briefly discuss some additional directions of future research.

One of the suggested future works of this study is applying higher resolution imaging (7 Tesla scanner) to smaller brain regions. It is possible that detailed investigation of the certain regions of the brain can provide researchers more detailed information regarding the abnormal functional network connectivity patterns of patients with schizophrenia.

The shapelet algorithm was successfully applied to the time series of non-artifact components and results were published (Cetin, MS., et al., 2015a). Dynamic FNC time series provide researchers a higher dimensional time series domain. Using the shapelet algorithm mentioned in chapter 3 may allow us to achieve higher accuracy.

Sensory load task hierarchy experiment with fMRI provided valuable information regarding how individuals' reactions to simple sensory stimuli are conditioned by the context within which they are presented. An important extension to the current study is to include cognitive tasks with established pathophysiology in schizophrenia, such as working memory, delayed match-to-sample, reinforcement learning, or Go/No-Go tasks. This extension would also increase the accuracy classification and diagnosis of patients with schizophrenia.

Lastly, results obtained from resting state MEG and fMRI data suggest that combining these two neuroimaging modalities reveal additional disease-relevant patterns of connectivity that were not detectable with fMRI or MEG alone. In addition to using just resting state data, it also may be helpful to use sensory load task hierarchy experiment to discover the details of schizophrenia.

References

Abdullah, M. (2013). Enumeration of Time Series Motifs of All Lengths. ICDM 547–556.

Abdullah, M., Eamonn, K., Qiang, Z., Sydney, C., and M. Brandon, W. (2009). Exact Discovery of Time Series Motifs. SDM 473–484.

Abdullah, M., Keogh, E., and Neal, Y. (2011). Logical-Shapelets: An Expressive Primitive for Time Series Classification. KDD *August 21–24*, 1154–1162.

Abou-Elseoud A, Starck T, Remes J, Nikkinen J, Tervonen O, and Kiviniemi V (2010). The effect of model order selection in group PICA. Hum Brain Mapp *31*, 1207–1216.

Abouzaid S, Jutkowitz E, Foley KA, Pizzi LT, Kim E, and Bates J. (2010). Economic impact of prior authorization policies for atypical antipsychotics in the treatment of schizo- phrenia. Popul Health Manag *13(5)*, 247e54.

Allen EA, Erhardt EB, Wei Y, Eichele T, and Calhoun VD (2011a). Capturing inter-subject variability with group independent component analysis of fMRI data: a simulation study. NeuroImage *59*, 4141–4159.

Allen EA, Erhardt EB, Damaraju E, Gruner W, Segall JM, Silva RF, Havlicek M, Rachakonda S, Fries J, and Kalyanam, R. (2011b). A baseline for the multi- variate comparison of resting-state networks. Front Syst Neurosci *5:2*.

Allen EA, Damaraju E, Plis SM, Erhardt EB, and Calhoun VD (2012). Tracking whole-brain connectivity dynamics in the resting state. Cereb. Cortex *24*, 663–676.

Andreasen N, Paradiso S, and O'Leary D. (1998). "Cognitive dysmetria" as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? Schizophr Bull *24(2)*, 203–218.

Andreasen NC, Carpenter Jr WT, Kane JM, Lasser RA, Marder SR, and Weinberger DR. (2005). Remission in schizophrenia: proposed criteria and rationale for consensus. Am J Psychiatry *162(3):*, 441e9.

Anticevic A, Repovs G, and Barch DM (2011). Emotion effects on attention, amygdale activation, and functional connectivity in schizophrenia. Schizophr Bull.

Arbabshirani MR, Havlicek M, Kiehl KA, Pearlson GD, and Calhoun VD (2013a). Functional Network Connectivity During Rest and Task Conditions: A Comparative Study. Hum. Brain Mapp. *34*, 2959–2971.

Arbabshirani MR, Kiehl KA, Pearlson GD, and Calhoun VD (2013b). Classification of schizophrenia patients based on resting-state functional network connectivity. Front. Brain Imaging Methods *7*, 1–16.

Argyelan M, Ikuta T, DeRosse P, Braga RJ, Burdick KE, and John M. (2013). Resting-State fMRI Connectivity Impairment in Schizophrenia and Bipolar Disorder. Schizophr Bullsbt *092*.

Astolfi, L., Cincotti, F., Mattia, D., Ding, L., and He, B., Salinari, S., & Babiloni, F. (2004). Estimating Causality among Cortical Areas of the Human Brain: A Study on the Application of Directed Transfer Function and Structural Equation Modeling to High Resolution EEG. Int. J. Bioelectromagn. *6(1)*.

Bandettini PA, Jesmanowicz A, Wong EC, and Hyde JS. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med J Soc Magn Reson Med Soc Magn Reson Med *30*, 161–173.

Bär K-J, Boettger MK, Koschke M, Schulz S, Chokka P, Yeragani VK, and Voss A. (2007). Non-linear complexity measures of heart rate variability in acute schizophrenia. Clin Neurophysiol *118*, 2009–2015.

Barnaly Rashid, Eswar Damaraju, Godfrey D. Pearlson, and Vince D. Calhoun (2014). Dynamic connectivity states estimated from resting fMRI Identify differences among Schizophrenia, bipolar disorder, and healthy control subjects. Front. Hum. Neurosci. *8*.

Barnaly Rashid, Mohammad Reza Arbabshirani, Eswar Damaraju, Robyn Miller, Mustafa S. Cetin, Godfrey Pearlson, and Vince Calhoun (2015). Classification of Schizophrenia and Bipolar Patients Using Static and Time-Varying Resting-State Fmri Brain Connectivity.

Bassett DS, Nelson BG, Mueller BA, Camchong J, and Lim KO (2011). Altered resting state complexity in schizophrenia. Neuroimage *59*, 2196–2207.

Beauchamp MS (2003). Detection of eye movements from fMRI data. Magn Reson Med *49*, 376–380.

Beckmann C, DeLuca M, Devlin J, and Smith S (2005). Investiga- tions into resting-state connectiv- ity using independent component analysis. Philos Trans R Soc B Biol Sci *360*, 1001.

Bell AJ and Sejnowski TJ (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Comput *7*, 1129–1159.

Bhurga D. (2005). The Global Prevalence of Schizophrenia. PLoS Med.

Biswal, B., Yetkin, F., Haughton, V., and Hyde, J. (1995). Functional connectivity in the motor cortex of resting human brain using echo planar MRI. Magn Reson Med *34*, 537–541.

Braeutigam, S. (2013). Magnetoencephalography: Fundamentals and Established and Emerging Clinical Applications in Radiology. ISRN Radiol. *2013*.

Breakspear M, Terry JR, Friston KJ, Harris AW, Williams LM, Brown K, Brennan J, and Gordon E. (2003). A disturbance of nonlinear interdependence in scalp EEG of subjects with first episode schizophrenia. NeuroImage *20(1)*, 466–478.

Brookes MJ, Gibson AM, Hall SD, Furlong PL, Barnes GR, and Hillebrand A . (2005). GLM-beamformer method demonstrates stationary field, alpha ERD and gamma ERS co-localisation with fMRI BOLD response in visual cortex. NeuroImage *26*, 302–308.

Brookes MJ, Hale JR, Zumer JM, Stevenson CM, Francis ST, and Barnes GR. (2011a). Measuring functional connectivity using MEG: Methodology and comparison with fcMRI. NeuroImage *56*, 1082–1104.

Brookes MJ, Woolrichb M, Luckhoob H, Pricea D, Halea JR, Stephensona MC, Barnesc GR, Smithd SM, and Morrisa PG (2011b). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Neuroimage *26*, 302–308.

Brown BB (1968). Some characteristic EEG differences between heavy smoker and non-smoker subjects. Neuropsychologia *6*, 381–388.

Buchsbaum M, Tang C, Peled S, Gudbjartsson H, Lu D, Hazlett E, Downhill J, Haznedar M, Fallon J, and Atlas S. (1998). MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. Neuroreport *9(3)*, 425–430.

Bullmore E and Sporns O (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. *10(3)*, 186–198.

Bullmore ET, Frangou S, and Murray RM (1997). The dysplastic net hypothesis: an integration of developmental and dysconnectivity theories of schizophrenia. Schizophr Res *28*, 143–156.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc *2*, 121–167.

Calhoun VD, and Adali T (2009). Feature-based Fusion of Medical Imaging Data. IEEE Trans. Info Tech Biomed *13*, 1–10.

Calhoun VD, and Adali T (2012). Multi-subject Independent Component Analysis of fMRI: A Decade of Intrinsic Networks, Default Mode, and Neurodiagnostic Discovery. IEEE Rev. Biomed. Eng. *5*, 60–73.

Calhoun VD, Adali T, Pearlson G, and Pekar J. (2001a). A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. *14*, 140–151.

Calhoun VD, Adali T, McGinty V, Pekar J, and Watson T, and Pearlson G. (2001b). fMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis. Neuroimage *14*, 1080–1088.

Calhoun VD, Adali T, and Pearlson G, and Pekar J. (2001c). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp. *13*, 43–53.

Calhoun VD, Adali T, Kiehl KA, Astur R, Pekar J, and Pearlson GD (2006). A method for multitask fMRI data fusion applied to schizophrenia. Hum Brain Mapp *27*, 598–610.

Calhoun VD, Kiehl K, and Pearlson G (2008). Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. Hum. Brain Mapp. *29(7)*, 828–838.

Calhoun VD, Liu J, and Adali T (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and erp data. Neuroimage *45*, 1–10.

Carlborg A, Winnerbäck K, Jönsson EG, Jokinen J, and Nordström P. (2010). Suicide in schizophrenia. Expert Rev Neurother *10(7)*, 1153e64.

Cetin, MS., Christensen, F., Abbott, CC., Stephen, JM., Mayer, AR., Cañive, JM., and Calhoun, VD. (2014). Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. NeuroImage *97*, 117–126.

Cetin, MS., Mueen, A., and Calhoun VD. (2015a). Shapelet Ensemble for Multi-dimensional Time Series. SIAM SDM.

Cetin, MS., Houck JM, and Calhoun VD. (2015b). Multi-model based classification for schizophrenia patients. Submitt. IEEE EMBC Milano Italy.

Cetin, MS., Stephen, J., and Calhoun VD. (2015c). Sensory load hierarchy based classification of schizophrenia patients. Submitt. IEEE ICIP.

Chang, K.-W., Deka, B., Hwu, W.-M. W., and Roth, D. (2012). Efficient Pattern-Based Time Series Classification on GPU.

Christos, F., M. Ranganathan., and Yannis-M. (1994). Fast subsequence matching in time-series databases. pp. 419–429.

Clementz BA, and Sweeney JA (1990). Is eye movement dysfunction a biological marker for schizophrenia? A methodological review. Psychol Bull *108*, 77–92.

Cohen D. (1968). Magnetoencephalography: evidence of magnetic fields produced by alpha rhythm currents. Science 161:784–786.

Cohen ER, Ugurbil K, and Kim S-G (2002). Effect of Basal Conditions on the Magnitude and Dynamics of the Blood Oxygenation Level–Dependent fMRI Response. J Cereb Blood Flow Metab *22*, 1042–1053.

Cole MW, Anticevic A, Repovs G, and Barch D (2011). Variable global dysconnectivity and individual differences in schizophrenia. Biol Psychiatry *70*, 43–50.

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, and Moritz CH (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR Am J Neuroradiol *21*, 1636–1644.

Cordes D, Haughton V, Carew JD, Arfanakis K, and Maravilla K (2002). Hierarchical clustering to measure connectivity in fMRI resting-state data. MagnReson Imaging *4*, 305–317.

Cullen AE, De Brito SA, Gregory SL, Murray RM, Williams SC, and Hodgins S (2012). Temporal Lobe Volume Abnormalities Precede the Prodrome: A Study of Children Presenting Antecedents of Schizophrenia. Schizophr Bull *39*, 1318–1327.

Damaraju, E., E. A. Allen, A. Belger, J. M. Ford, S. McEwen, D. H. Mathalon, and B. A. Mueller et al. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. NeuroImage Clin. *5*, 298–308.

Daniel PK, Elizabeth R, and Eric C (2006). Failing to deactivate: resting functional abnormalities in autism. Proc. Natl. Acad. Sci. U. S. A. *103*, 8275–8280.

D. Cohen (1972). Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer. Sci- Ence Vol *175*, 664–666.

Desai, P.R., Lawson, K.A., Barner, J.C., and Rascati, K.L. (2013). Identifying patient characteristics associated with high schizophrenia-related direct medical costs in community-dwelling patients. J Manag Care Pharm *19*, 468–477.

de Silva J, Hanwella R, and de Silva VA. (2012). Direct and indirect cost of schizophrenia in outpatients treated in a tertiary care psychiatry unit. Ceylon Med J *57(1)*, 14e8.

Diaconescu AO, Jensen J, Wang H, Willeit M, Menon M, Kapur S, and McIntosh AR (2011). Aberrant effective connectivity in schizophrenia patients during appetitive conditioning. Front Hum Neurosci 4239 Doi 103389fnhum201000239.

Diederen KMJ, Neggers SFW, de Weijer AD, van Lutterveld R, Daalman K, and Eickhoff SB (2013). Aberrant resting-state connectivity in non-psychotic individuals with auditory hallucinations. Psychol Med *43*, 1685–1696.

Dietterich TG., and Kong EB. (1995). Machine Learning Bias, Statistical Bias, and Statistical Variance of Decision Tree Algorithms. (San Francisco), pp. 313–321.

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, and Dosenbach RAT (2007). Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci USA *104*, 11073–11078.

Duda RO, Hart PE, and Stork DG. (2001). Pattern classification (New York: Wiley).

Du, Y., Allen EA., He, H., Sui, J., and Calhoun, VD. (2014). Comparison of ICA based fMRI artifact removal: single subject and group approaches. In Proceeding. (Hamburg, Germany),.

Erhardt EB, Allen EA, Damaraju E, and Calhoun VD (2011a). On network derivation, classification, and visualization: a response to Habeck and Moeller. Brain Connect. *1*, 1–19.

Erhardt EB, Rachakonda S, Bedrick EJ, Allen EA, Adali T, and Calhoun VD (2011b). Comparison of multi-subject ICA methods for analysis of fMRI data. Hum Brain Mapp *32*, 2075–2095.

Faloutsos, C, and Lin, K.-I. (1995). FastMap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. (San Jose, California), pp. 163–174.

First M, Spitzer R, Gibbon M, and Williams J (1997). Structured clinical interview for DSM-IV axis I disorders-clinician version (SCID-CV). Am Psychiatr Assoc Press Wash DC.

First MB, Spitzer RL, Gibbon M, and Williams JBW (2002a). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition. *New York State Psychiatric Institute, Biomedical Research.*

Foong J, Symms MR, Barker GJ, Maier M, Miller DH, and Ron MA. (2002). Investigating regional white matter in schizophrenia using diffusion tensor imaging. Neuroreport *13(3),* 333–336.

Ford JM, Palzes VA, Roach BJ, Potkin SG, Erp TGM van, and Turner JA. (2014). Visual Hallucinations Are Associated With Hyperconnectivity Between the Amygdala and Visual Cortex in People With a Diagnosis of Schizophrenia. Schizophr Bull *31*.

Fornito A, Yoon J, Zalesky A, Bullmore ET, and Carter CS (2011). General and specific functional connectivity disturbances in first-episode schizophrenia during cognitive control performance. Biol Psychiatry *70*, 64–72.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, and Raichle ME (2005). The human brain is intrinsically organized into dynamic, anti correlated functional networks. Proc Natl Acad Sci USA *102:9673–9678.*

Freire L, Roche A, and Mangin JF (2002). What is the best similarity measure for motion correction in fMRI time series? IEEE Trans Med Imaging *21*, 470–484.

Friedman J., Hastie T., and Tibshirani R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics *9*, 432–441.

Friston K. (1998). The disconnection hypothesis. Schizophr Res *30(2)*, 115–125.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, and Frackowiak RSJ (1995). Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp *2*, 189–210.

Friston KJ and Frith CD (1995). Schizophrenia: a disconnection syndrome? Clin. Neurosci. *3(2)*, 89–97.

Fusar-Poli P (2007). Neurofunctional correlates of vulnerability to psychosis: a systematic review and meta-analysis. Neuro Sci Behav *31*, 465–484.

Gardner DM, Murphy AL, O'Donnell H, Centorrino F, and Baldessarini RJ (2010). International consensus study of antipsychotic dosing. Am. J. Psychiatry *167*, 686–693.

Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, and Calhoun VD (2007). Aberrant default mode functional connectivity in schizophrenia. Am J Psychiatry *164*, 450–457.

Glahn DC, Ragland JD, Abramoff A, Barrett J, Laird AR, and Bearden CE (2005). Beyond hypofrontality: a quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. Hum Brain Mapp *25:60-69*.

Goff DC, and Coyle JT (2001). The emerging role of glutamate in the pathophysiology and treatment of schizophrenia. Am J Psychiatry *158*, 1367–1377.

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, and Mark RG et. al. (2000). Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. Circulation *101*, e215–e220.

Goldstein JM, Seidman LJ, Kennedy DN, Makris N, Lee H, Tourville J, Caviness JVS, Faraone SV, and Tsuang MT. (1999). Cortical Abnormalities in Schizophrenia Identified by Structural Magnetic Resonance Imaging. Am Med Assoc 537–547.

Green M, Kern R, Braff D, and Mintz J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the "right stuff"? Schizophr Bull *26(1)*, 119–136.

Guy W (1976). ECDEU Assessment Manual for Psychopharmacology. US Dept Health Educ. Welf. ADM 76–338.

Hadjipapas A, Hillebrand A, Holliday IE, Singh KD, and Barnes GR (2005). Assessing interactions of linear and nonlinear neuronal sources using MEG beamformers: a proof of concept. Clin Neurophysiol *116*, 1300–1313.

Hall EL, Woolrich MW, Thomaz CE, Morris PG, and Brookes MJ (2013). Using variance information in magnetoencephalography measures of functional connectivity. NeuroImage *67*, 203–212.

Hamalainen MS, and Sarvas J (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Trans Biomed Eng *36*, 165–171.

Hamm JP, Gilmore CS, Picchetti NA, Sponheim SR, and Clementz BA (2011). Abnormalities of neuronal oscillations and temporal integration to low- and high-frequency auditory stimulation in schizophrenia. Biol Psychiatry *69*, 989–996.

Hanlon FM, Houck JM, Pyeatt CJ, Lundy SL, Euler MJ, and Weisend MP. (2011). Bilateral hippocampal dysfunction in schizophrenia. NeuroImage *58*, 1158–1168.

Hanlon FM, Houck JM, Klimaj S d., Caprihan A, Mayer AR, and Weisend MP. (2012). Frontotemporal anatomical connectivity and working-relational memory performance predict everyday functioning in schizophrenia. Psychophysiology *49*, 1340–1352.

Harrison, B.J., Pujol, J., Lopez-Sola, M., Hernandez-Ribas, R., Deus, J., Ortiz, H., Soriano-Mas, C., Yucel, M., Pantelis, C., and Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. Proc. Natl. Acad. Sci. U. S. A. *105*, 9781–9786.

Himberg J, and Hyvarinen A. (2003). Icasso: software for investigating the reliability of ICA estimates by clustering and visualization. IEEE 13th Workshop Neural Netw. Signal Process. 259–268.

Himberg J, Hyvärinen A, and Esposito F (2004). Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage *22*, 1214–1222.

Hoffman RE., and Hampson M. (2012). Functional connectivity studies of patients with auditory verbal hallucinations. Front. Hum. Neurosci. *6*.

Honea R, Crow TJ, Passingham D, and Mackay CE. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. Am.J.Psychiatry *162(12)*, 2233–2245.

Houck JM, Cetin, MS., Mayer, AR., Bustillo, JR., Stephen, JM., Aine, CJ, and Cañive, JM., et al. (2015). Magnetoencephalographic and functional MRI connectomics in schizophrenia via intra- and inter- network connectivity. Biol. Psychiatry.

Huettel, S., Song, A., and McCarthy, G. (2004). Functional Magnetic Resonance Imaging. Sinauer Associates Sunderland, MA.

Jafri MJ, Pearlson GD, Stevens M, and Calhoun VD (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. Neuroimage *39*, 1666–1681.

Jardri R, Pouchet A, Pins D, and Thomas P. (2010). Cortical Activations During Auditory Verbal Hallucinations in Schizophrenia: A Coordinate-Based Meta-Analysis. Am J Psychiatry *168*, 73–81.

Jason Lines, Luke M. Davis, Jon Hills, and Anthony Bagnall (2012). A Shapelet Transform for Time Series Classification. KDD *August 12–16*.

Jesin, Z., Sarah, R., Abdullah, M., Khaleel, R., and Keogh, E. (2012). Mining massive archive of mice sounds with symbolized representations. Siam Int. Conf. Data Min.

Jin S., and Eamonn, K. (2008). iSAX: indexing and mining terabyte sized time series. pp. 623–631.

Joel SE, Caffo BS, Van Zijl PC, and Pekar JJ (2011). On the relationship between seed-based and ICA-based measures of functional connectivity. MagnReson Med *66*, 644–657.

Kay SR, Fiszbein A, and Opler LA (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull *13*, 261–276.

Keogh, E., Zhu, Q., Hu, B., Hai Y., Xi, X., Wei, L., and Ratanamahatana, C. (2012). The UCR Time Series Classification / Clustering Homepage:

Keshavan, M. S., Clementz, B. A., Pearlson, G. D., Sweeney, J. A., and Tamminga, C.A. (2013). Reimagining psychoses: an agnostic approach to diagnosis. Schizophr Res *146*, 10–16.

Khadka S, Meda SA, Stevens MC, Glahn DC, Calhoun VD, and Sweeney JA. (2013). Is Aberrant Functional Connectivity A Psychosis Endophenotype? A Resting State Functional Magnetic Resonance Imaging Study. Biol Psychiatry *74*, 458–466.

Kim SG, Richter W, and Uğurbil K. (1997). Limitations of temporal resolution in functional MRI. Magn Reson Med J Soc Magn Reson Med Soc Magn Reson Med *37*, 631–636.

Kiviniemi V, Starck T, Remes J, Long X, Nikkinen J, and Haapea M, et al. (2009). Functional segmentation of the brain cortex using high model order group PICA. Hum Brain Mapp *30*, 3865–3886.

Kohavi, Ron. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI. *14*.

Kolder H, and North AW (1966). Oscillations of the Corneo-Retinal Potential in Animals. Ophthalmologica *152*, 149–160.

Kubicki, M., Mccarley, R., Westin, C.-F., Park, H.-J., Maier, S., Kikinis, R., and Jolesz, F.A., and Shenton, M.E. (2007). A review of diffusion tensor imaging studies in schizophrenia. J. Psychiatr. Res. *41*, 15–30.

Kühn S, and Gallinat J. (2013). Resting-State Brain Activity in Schizophrenia and Major Depression: A Quantitative Meta-Analysis. Schizophr Bull *39*, 358–365.

Lee H, Dvorak D, and Fenton AA (2014). Targeting Neural Synchrony Deficits is Sufficient to Improve Cognition in a Schizophrenia-Related Neurodevelopmental Model. Front Psychiatry 5.

Leucht S, Burkard T, Henderson J, and Maj M, Sartorius N. (2007). Physical illness and schizophrenia: a review of the literature. Acta Psychiatr Scand *116(5)*, 317e 33.

Lexiang, Y., and Keogh, E. (2009). Time Series Shapelets: A New Primitive for Data Mining. KDD *June 29–July 1*.

Lin, J., Keogh, E, Wei, L., and Lonardi, S, (2007). Experiencing SAX: a novel symbolic representation of time series. In DMKD, pp. 107–144.

Liu H., Kaneko Y., Ouyang X., Li L., Hao Y., and Chen EYH. (2012). Schizophrenic Patients and Their Unaffected Siblings Share Increased Resting-State Connectivity in the Task-Negative Network but Not Its Anticorrelated Task-Positive Network. Schizophr. Bull. *38*, 285–294.

Ljosa, V., Bhattacharya, A., and Singh, AK. (2006). LB-Index: A Multi-Resolution Index Structure for Images. p. 144.

Luckhoo H, Hale JR, Stokes MG, Nobre AC, Morris PG, Brookes MJ, and Woolrich MW. (2012). Inferring task-related networks using independent component analysis in magnetoencephalography. NeuroImage *62*, 530–541.

L. Wei, Keogh, E., and Xiaopeng Xi (2006). SAXually Explicit Images: Finding Unusual Shapes. pp. 711–720.

Mathiak K, Ackermann H, Rapp A, Mathiak KA, Shergill S, and Riecker A, et al. (2011). Neuromagnetic oscillations and hemodynamic correlates of P50 suppression in schizophrenia. Psychiatry Res *194*, 95–104.

Mayer AR, Ruhl D, Merideth F, Ling J, Hanlon FM, Bustillo J, and Cañive J (2012). Functional imaging of the hemodynamic sensory gating response in schizophrenia. Hum Brain Mapp *34*, 2302–2312.

McDonald, M., Hertz, R.P., Lustik, M.B., and Unger, A.N. (2005). Healthcare spending among community-dwelling adults with schizophrenia. Am. J. Manag. Care *11*, S242–S247.

McKeown MJ., Makeig S., Brown GG., Jung TP., and Kindermann SS. et. al (1998). Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp *6*, 160–188.

Meda SA, Gill A, Stevens MC, Lorenzoni RP, Glahn DC, and Calhoun VD, et al. (2012). Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. Biol Psychiatry *71*.

Millier, A., Schmidt, U., Angermeyer, M. C., and Chauhan, D., Murthy, V., Toumi, M., & Cadi-Soussi, N. (2014). Humanistic burden in schizophrenia: A literature review. J. Psychiatr. Res. *54*, 85–93.

Minami T, Nobuhara K, Okugawa G, Takase K, Yoshida T, Sawada S, Ha-Kawa S, and Ikeda K and Kinoshita T (2003). Diffusion tensor magnetic resonance imaging of disruption of regional white matter in schizophrenia. Neuropsychobiology *47(3)*, 141–145.

Minzenberg MJ (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. Arch Gen Psychiatry *66*, 811–822.

Paterson AS (1935). The Respiratory Rhythm in Normal and Psychotic Subjects. J Neurol Psychopathol *16*, 36–53.

Piantoni G, Van Der Werf YD, Jensen O, and Van Someren EJW (in press). Memory traces of long-range coordinated oscillations in the sleeping human brain. Hum Brain Map.

Rachow T, Berger S, Boettger MK, Schulz S, Guinjoan S, and Yeragani VK (2011). Nonlinear relationship between electrodermal activity and heart rate variability in patients with acute schizophrenia. Psychophysiology *48*, 1323–1332.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, and Shulman GL. (2001). A Default Mode of Brain Function. Proc Natl Acad Sci *98*, 676–682.

Rakthanmanon, T., and Keogh, E. (2013). Fast Shapelets: A Scalable Algorithm for Discovering Time Series Shapelets. In SIAM SDM, (Austin, USA),.

Ramkumar P, Parkkonen L, and Hyvärinen A (in press). Group-level spatial independent component analysis of Fourier envelopes of resting-state MEG data. NeuroImage.

Reeve A, Knight J, Maclin E, Lewine J, and Orrison W (1993). Resting-state magnetoencephalography in schizophrenia. Abstr Int Congr Schizophr Res Colo Springs.

Repovš G, and Barch DM (2012). Working memory related brain network connectivity in individuals with schizophrenia and their siblings. Front. Hum. Neurosci. *6*.

Robinson S, Basso G, Soldati N, Sailer U, Jovicich J, and Bruzzone L, et al. (2009). A resting state network in the motor control circuit of the basal ganglia. BMC Neurosci 10 137 Doi1011861471-2202-10-137.

Rolls ET, Loh M, Deco G, and Winterer G (2008). Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nat Rev Neurosci *9*, 696–709.

Rosadini G, Rodriguez G, and Siani C. (1974). Acute alcohol poisoning in man: An experimental electrophysiological study. Psychopharmacologia *35*, 273–285.

Roweis S. (1998). EM algorithms for PCA and SPCA. Adv Neural Inf Process Syst *10*, 626–632.

Sakoğlu, Ünal, Godfrey D. Pearlson, Kent A. Kiehl, Y. Michelle Wang, Andrew M. Michael, and Vince D. Calhoun (2010). A method for evaluating dynamic functional network connectivity

and task-modulation: application to schizophrenia. Magn. Reson. Mater. Phys. Biol. Med. *23*, 351–366.

Sarvas J (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys Med Biol *32*, 11.

Saykin, AJ, Gur RC, Gur RE, Mozley PD, Mozley LH, and Resnick SM et. al. (1991). Neuropsychological function in schizophrenia. Selective impairment in memory and learning. Arch Gen Psychiatry *48:*, 618–624.

Scott A, (2011). COINS: An Innovative Informatics and Neuroimaging Tool Suite Built for Large Heterogeneous Datasets. Front Neuroinformatics *5*, 33.

Sehatpour P, Molholm S, Schwartz TH, Mahoney JR, Mehta AD, and Javitt DC (2008). A human intracranial study of long-range oscillatory coherence across a frontal–occipital–hippocampal brain network during visual object processing. Proc Natl Acad Sci *105*, 4399–4404.

Shenton ME, Dickey CC, Frumin M, and McCarley RW (2001). A review of MRI findings in schizophrenia. Schizophr. Res *49*, 1–52.

Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M.D. (2012). Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. Cereb. Cortex *22*, 158–165.

Shostak VI (1968). Reflection of excitability of the visual centers in the spontaneous electroencephalogram in man. Zhurnal Vysshei Nervn Deyatelnosti *18*, 880–885.

Siebenhühner F, Weiss SA, Coppola R, Weinberger DR, and Bassett DS. (2013). Intra- and Inter-Frequency Brain Network Structure in Health and Schizophrenia. PLoS ONE *8*, 1–13.

Skouras S, Gray M, Critchley H, and Koelsch S. (2013). fMRI Scanner Noise Interaction with Affective Neural Processes. PLoS ONE *8*.

Skudlarski P, Jagannathan K, Anderson K, Stevens MC, Calhoun VD, Skudlarska BA, and Pearlson G. (2010). Brain Connectivity Is Not Only Lower but Different in Schizophrenia: A Combined Anatomical and Functional Approach. Biol Psychiatry Schizophr. N-Methyl--Aspartate Recept. Dysfunct. Cortical Connect. *68*, 61–69.

Smith, S.M. (2004). Overview of fMRI analysis. Br J Radiol *77*, 167–175.

Smith S.M., Miller K.L., Salimi-Khorshidi G., Webster M., Beckmann C.F., Nichols T.E., Ramsey J.D., and Woolrich M.W. (2011). Network modelling methods for FMRI. Neuroimage. *54*, 875–891.

Sommer IE, Clos M, Meijering AL, Diederen KMJ, and Eickhoff SB. (2012). Resting State Functional Connectivity in Patients with Chronic Hallucinations. PLoS ONE *7*, e43516.

Soric B (1989). Statistical “Discoveries” and Effect-Size Estimation. J. Am. Stat. Assoc. *84*, 608–610.

Stein A von, Rappelsberger P, Sarnthein J, and Petsche H. (1999). Synchronization Between Temporal and Parietal Cortex During Multimodal Object Processing in Man. Cereb Cortex *9*, 137–150.

Stephan KE, Baldeweg T, and Friston KJ. (2006). Synaptic Plasticity and Dysconnection in Schizophrenia. Biol Psychiatry *59*, 929–939.

Stephan KE, Friston KJ, and Frith CD (2009). Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. Schizophr Bull *35*, 509–527.

Stone DB, Urrea LJ, Aine CJ, Bustillo JR, Clark VP, and Stephen JM (2011). Unisensory processing and multisensory integration in schizophrenia: a high-density electrical mapping study. Neuropsychologia *49*, 3178–3187.

Stone DB, Coffman BA, Bustillo JR, Aine CJ, and Stephen JM (2014). Multisensory stimuli elicit altered oscillatory brain responses at gamma frequencies in patients with schizophrenia. Front Hum Neurosci *8*, 788.

Susan WG, Heidi WT, Snezana M, Ming TT, Stephen VF, Robert WM, Martha ES, Alan IG, Alfonso NC, Peter LV, et al. (2009). Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. Proc. Natl. Acad. Sci. *106*, 1279–1284.

Tallon-Baudry C, Mandon S, Freiwald WA, and Kreiter AK. (2004). Oscillatory Synchrony in the Monkey Temporal Lobe Correlates with Performance in a Visual Short-term Memory Task. Cereb Cortex *14*, 713–720.

Thatcher RW, North DM, and Biver CJ. (2008). Development of cortical connections as measured by EEG coherence and phase delays. Hum Brain Mapp *29*, 1400–1415.

Toichi M, Kubota Y, Murai T, Kamio Y, Sakihama M, and Toriuchi T (1999). The influence of psychotic states on the autonomic nervous system in schizophrenia. Int J Psychophysiol *31*, 147–154.

Tompa, M., and Buhler, J. (2002). Finding motifs using random projections. J Comput Biol *9*, 225–242.

Tononi G, and Edelman GM (2000). Schizophrenia and the mechanisms of conscious integration. Brain Res Rev *31*, 391–400.

Uhlhaas PJ, and Singer W (2010). Abnormal neural oscillations and synchrony in schizophrenia. Nat Rev Neurosci *11*, 100–113.

Uhlhaas PJ, and Singer W (2011). The Development of Neural Synchrony and Large-Scale Cortical Networks During Adolescence: Relevance for the Pathophysiology of Schizophrenia and Neurodevelopmental Hypothesis. Schizophr Bull *37*, 514–523.

Van Den Heuvel MP., and Hulshoff Pol HE. (2010). A review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. *20(8)*, 519–534.

Van DKR, Hedden T, Venkataraman A, Evans KC, Lazar SW, and Buckner RL (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. Neurophysiol *103*, 297–321.

Van Os J, Burns T, Cavallaro R, Leucht S, Peuskens J, and Helldin L, et al. (2006). Standardized remission criteria in schizophrenia. Acta Psychiatr Scand *113(2):*, 91e5.

Whitehorn JC, and Richter H (1937). Unsteadiness of the heart rate in psychotic and neurotic states. Arch Neurol Psychiatry *38*, 62–70.

Willis M, Svensson M, Löthgren M, Eriksson B, Berntsson A, and Persson U. (2010). The impact on schizophrenia-related hospital utilization and costs of switching to long acting risperidone injections in Sweden. *Eur J Health Econ*, 585e94.

Wittkower E (1934). Further Studies in the Respiration of Psychotic Patients. Br J Psychiatry *80*, 692–704.

Woodward ND (2012). Thalamocortical dysconnectivity in schizophrenia. Am. J. Psychiatry *1*, 1092–1099.

Woodward ND, Rogers B, and Heckers S (2011). Functional resting state networks are differentially affected in schizophrenia. Schizophr Res *130*, 86–93.

Yasushi, S., Spiros, P., and Christos, F. (2005). BRAID: Stream mining through group lag correlations. SIGMOD Conf. 599–610.

Yu Y., Shen H., Zhang H., Zeng L., Xue Z., and Hu D. (2013). Functional connectivity-based signatures of schizophrenia revealed by multiclass pattern analysis of resting-state fMRI from schizophrenic patients and their healthy siblings. Biomed. Eng. Online *12:10*.