Understanding the relationship between Functional and Structural Connectivity of Brain Networks

Sashank J. Reddi SJAKKAMR@CS.CMU.EDU
Machine Learning Department, Carnegie Mellon University

Abstract

Background. Functional magnetic resonance imaging (fMRI) and diffusion spectrum imaging (DSI) are two popular methods used for analyzing brain networks. While one measures the neuronal activity (functional connectivity) in the brain, the other measures its anatomical structure (structural connectivity). While several researchers have studied the properties of fMRI and DSI independently, the relationship between them is largely unexplored.
Aim. In this paper, we study the relationship between these two fundamental measurements of brain networks. In particular, we answer the following fundamental questions: (1) Does there exist any dependence between structural and functional connectivity? If yes, what is the nature of this dependence? (2) Is structural information useful for predictive analysis of the functional data? (3) Is it possible to make predictions about the structural data based on the functional data?
Data. The data used for our study consists of resting state fMRI and DSI data from 55 subjects. The fMRI data is recorded for 210 time steps (in a span of 10 minutes).
Methods. We use various independence tests for detecting and understanding the relationship between these two sources of information. We also analyze the strength and the nature of the relationship and its variability across different regions of the brain. We use a regression setup for studying the prediction questions of our interest.
Results. Our experiments clearly demonstrate that there is a high dependence between functional and structural connectivity. For example, independence tests yield p-values less than 0.01 — indicating high dependence. We also show that the nature of the relationship is intricate and non-uniform across different regions of the brain. Furthermore, our analysis shows that it is possible to predict functional data from structural data and vice-versa with R^2 value as high as 0.7 in certain regions of the brain.
Conclusions. We show that functional and structural connectivity share a strong relationship with each other. Our analysis provides insights into the nature of functional and structural connectivity relationship across different regions of the brain. We also identify the key regions of the brain where this relationship is strong, thus providing a potentially valuable tool for inference problems involving functional and structural data.

1. Introduction

Magnetic resonance imaging (MRI) has revolutionized the field of neuroscience. MRI scanners use radio waves and magnetic fields to obtain images of the body without exposure to ionizing radiation and hence, provides a fast, non-destructive and non-invasive means of imaging the body. Functional magnetic resonance imaging (fMRI) and diffusion spectrum imaging (DSI) are two fundamental and popular brain imaging techniques. fMRI is a technique for measuring the activity of the brain. fMRI takes advantage of the differences in magnetic fields of oxygenated and deoxygenated blood in order to measure the *functional* activity of the brain. More specifically, it measures the localized

increases in blood oxygenation level, which in turn reflects the increase in neuronal activity. This method is called blood-oxygenation-level-dependent (BOLD) imaging. On the other hand, DSI is an important technique for studying the anatomy of the brain. DSI measures the diffusion behavior of water molecules in the brain. The difference in diffusion of water molecules along the axonal bundles and perpendicular direction is key to estimating the anatomical organization of the brain. This provides us with rich anatomical information about the white matter, the component of the brain that is composed of axons. Hence, fMRI and DSI measure two different properties of the brain. While one measures the activity of the brain, the other measures the anatomical connectivity of the brain network.

A natural and fundamental issue in neuroscience is to understand the relationship between different imaging techniques. A question of special interest to us is: what is the relationship between measurements obtained from fMRI and DSI? As explained above, these techniques are designed to measure different aspects of the brain but it is intuitive that anatomical connectivity of the brain plays a role in neuronal activity of the brain. Our main focus in this paper is to understand the dependence (or independence) between the fMRI and DSI data. In other words, we are interested in the question: Is there any relationship between functional connectivity (as measured through fMRI data) and structural connectivity (as measured through DSI data) of the brain and, if yes, how does it vary across different regions of the brain?

To this end, we investigate various principled approaches for quantitatively and qualitatively detecting the dependence between the functional and structural data. Furthermore, we also characterize the variability of this dependence across different regions in the brain network. For this purpose, we use various inference tools, independence tests, and canonical correlation analysis to give a comprehensive view of the relationship.

2. Related Work

In this section, we briefly review the related work for the problem of our interest. The relationship between functional and structural data of brain networks has received considerable interest in the neuroscience community. While there has been some evidence confirming the relationship between the two types of data that are of our interests, the results are neither compelling nor complete. Earlier works (see, for example, (Koch et al., 2002)) argue that the relationship between fMRI and DSI data is not direct and straightforward. They, further, argue that the *indirect* structural connectivity has much higher relationship with fMRI than direct structural connectivity.

More recent works attempt to demonstrate a direct relationship between functional and structural connectivity. The notion used to establish this dependence is that of functional correlation graphs (Honey et al., 2009). In particular, the functional correlation graph is shown to have direct relationship with the structural connectivity. Honey et al. (2009) mainly focus on the relationship within subjects. Our experiments reveal that the structural connectivity notion used in Honey et al. (2009) is much weaker. We define a new notion of structural connectivity which shows much stronger relationship with the functional connectivity than that obtained in Honey et al. (2009). Furthermore, our experiments across subjects reveal very interesting functional and structural connectivity relationship within various regions of the brain.

Recently, there also has been some interest in inferring functional connectivity from structural connectivity and vice versa (Robinson, 2012; Robinson et al., 2014; Deligianni et al., 2010). While such an inference demonstrates relationship between function and structural connectivity, it does not

allow direct examination of this relationship. Our methodology, on the other hand, allows us to infer the relationship in a more direct manner. This, in turn, provides a compelling case for understanding the relationship between functional and structural data.

2.1 Data

In this section, we describe the datasets used in the paper. We acquired resting state fMRI and DSI data for 55 subjects. The exact data acquisition process for fMRI (functional data) and DSI (structural data) is given in Jarbo and Verstynen (2015). We use region of interest (ROI) analysis in this paper. ROI analysis involves selecting clusters of voxels or brain regions a priori while analyzing the brain. ROI analysis has several advantages over whole brain (or voxel level) analysis. We refer the reader to Poldrack (2006) for a more comprehensive treatment of ROI analysis. Our dataset consists of 625 ROIs. The description for generating the ROIs can be seen in Jarbo and Verstynen (2015)[1].

The fMRI data for a subject consists of functional activity (time series) at 210 time steps, one for each of the 625 ROIs. The fMRI data is measured in resting state, i.e., the participants were not given any explicit task during the experiment. The DSI (structural data) for a subject is a 625 x 625 matrix where the ij$^{\text{th}}$ entry of the matrix represents some structural feature between the i^{th} and j^{th} ROI. In this paper, we are mainly concerned with the quantitative anisotropy (QA) and mean length for each pair of ROIs (see (Jarbo and Verstynen, 2015) and references therein for more description of these quantities). The normalized QA between two ROIs is defined as the QA between the corresponding ROIs normalized by the mean length between them. This is referred to as structural connectivity throughout this paper.

3. Problem Setup

We formally define the problem statement in this section. We assume that the number of ROIs in our problem is d. The number of subjects in our experiments is denoted by n. The functional input of our concern is a time series, one for each subject. In general, the length of the time series can be different for subjects. However, for simplicity, we assume that it is same across subjects and is denoted by T. We represent the functional input as a tensor $X \in \mathbb{R}^{T \times d \times n}$. Here $X[t, j, i]$ represents the value of j^{th} ROI at time step t for the i^{th} subject. This data forms the functional input to the problem of our interest.

Additionally, we are also provided with structural information for all subjects in the experiment. Throughout this paper, we represent the structural data as a tensor in $\mathbb{R}^{d \times d \times n}$, where $S[j, k, i]$ represents a structural aspect involving ROIs j and k of the i^{th} subject. There are two primary and important aspects of the data at our disposal: (a) it is inherently high dimensional (b) small sample size because the experiments are carried on a small set of subjects. More specifically, we are interested in the case where $T, n < d$. In our dataset, n = 55, T = 210 and d = 625.

Three Fundamental Questions

In this paper, we aim to explore the following three important questions concerning the relationship between structural and functional connectivity:

1. The datasets used in this paper are not publicly available.

1. Does there exist any dependence between structural and functional connectivity for the subjects in the dataset? Or are they independent? If a dependence indeed exists, what kind of relationship exists between them?

2. Is structural information useful for predictive analysis of the functional data? More specifically, can structural data be used in making or improving predictions concerning functional data? If yes, how this benefit vary across different regions of the brain?

3. Finally, we also attempt to explore the inverse of (2) i.e., is it possible to make predictions on the structural data based on the functional data? Is it possible to recover the support of the structural data from the functional data?

We will explore each of these questions in great detail in Section 4.7. Note that each of these questions is focused on the relationship between the structural and functional properties of the brain and attempt to uncover different aspects of this underlying relationship, if any. While the first question focuses on detecting dependence between functional and structural connectivity, the other two allow us to understand the nature of this dependence.

4. Preliminaries

We provide a brief overview of the approaches used in this paper. This section is mainly to review the necessary background for the paper and is not intended to be an exhaustive or complete treatment of these approaches. We refer the interested readers to the relevant literature for more details.

4.1 Functional Correlation Graph

The functional correlation graph (FCG) is a popular tool used in analysis of brain networks. We describe the procedure used for constructing the functional correlation graph for a subject. The adjacency matrix of the graph is constructed in the following manner: For each pair of ROIs, the weight of the edge between them is the Pearson correlation between the corresponding time series for that pair of ROIs. We first note that the graph is undirected since the adjacency matrix will be symmetric. We denote the resultant adjacency matrix as Φ_i. Furthermore, we also observe that weights can be negative since the ROIs can have negative correlation. The reasons for negative correlation is not fully understood in the neuroscience community. In many of our experiments, we use the absolute values of the correlation while measuring the metrics based on functional data.

The FCG and structural connectivity matrix share interesting features. Figure 1 shows the functional correlation graph and structural connectivity for a subject. Though FCG is dense in comparison to the structural connectivity matrix, the two indeed share some structural similarities. This will, in fact, be one of our major focus in the later sections, where we explore the relationship in a systematic fashion.

4.2 MMD & HSIC

We briefly review the theory behind Maximum Mean Discrepancy (MMD) & HSIC in this section. These measures will be useful in defining distances between distributions and testing for independence between random variables. Both these measures are useful in identifying non-linear relationships between random variables. Suppose X is a random variable on $\mathcal{X}$. Let $\mathcal{H} = \{f : \mathcal{X} \to \mathbb{R}\}$

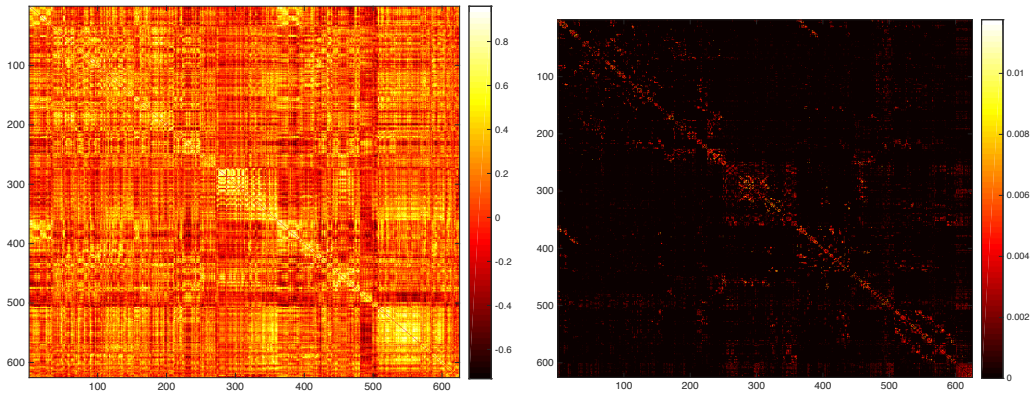

Figure 1: Functional correlation graph using the function data (left) and structural connectivity graph (right) between ROIs for a subject. Note the structural similarity between the two heat maps.

be a reproducing kernel Hilbert Space (RKHS) associated with $\mathcal{X}$ with feature map $\phi(x) \in \mathcal{H}$ ($x \in \mathcal{X}$) and kernel $k(x, y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}}$ ($x, y \in \mathcal{X}$). The kernel satisfies the property $f(x) = \langle f, k(x, .) \rangle_{\mathcal{H}_X}$ for $f \in \mathcal{H}_X$, which is called the reproducing property of the kernel.

The mean embedding of a distribution P_X on $\mathcal{X}$ is defined as $\mu_X = \int \phi(x) dP_X$ (Gretton et al., 2007). The MMD distance between two distributions P_X and P_Y on $\mathcal{X}$ is defined as:

$$\text{MMD}(P_X, P_Y) = \|\mu_X - \mu_Y\|^2_{\mathcal{H}_X}$$

This quantity can, in general, be estimated using empirical mean embeddings (see (Gretton et al., 2007)) and is useful for measuring distances between distributions. In particular, suppose $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are samples from distribution P_X and P_Y respectively, then one popular estimator of MMD is the following:

$$\widehat{\text{MMD}}(P_X, P_Y) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{m} k(X_i, X_j) + \frac{1}{m(m-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{m} k(Y_i, Y_j) - \frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} k(X_i, Y_j) \tag{1}$$

The Hilbert Schmidt Independence Criterion (HSIC) is useful for testing independence between random variables. Suppose (X, Y) is a random variable on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{H}_X = \{f : \mathcal{X} \to \mathbb{R}\}$ be a RKHS associated with $\mathcal{X}$ with feature map $\phi(x) \in \mathcal{H}_X$ ($x \in \mathcal{X}$) and kernel k_X. We can similarly define RKHS $\mathcal{H}_Y$ and kernel k_Y associated with $\mathcal{Y}$. HSIC between set of random variables is defined as the following:

$$\text{HSIC}(X, Y) = \text{MMD}\left(P_{(X,Y)}, P_X \times P_Y\right),$$

where $P_{(X,Y)}, P_X, P_Y$ represent the joint distribution and corresponding marginals of random variables X and Y respectively (Gretton et al., 2005). Here $P_X \times P_Y$ denotes the product distribution

of X and Y. An alternative way of writing HSIC measure is the following:

$$\text{HSIC}(X, Y) = \|\Sigma_{XY}\|^2_{HS},$$

where Σ_{XY} denotes the cross-covariance operator and $\|.\|_{HS}$ denotes the Hilbert-Schmidt norm. The estimation of HSIC can be done efficiently (both statistically and computationally) through samples from the joint distribution on (X, Y) (see (Gretton et al., 2005; Reddi and Póczos, 2013)). The key property of HSIC is that, under certain conditions, it is non-zero iff X and Y are independent. We will use this property for our independence tests. One of the remarkable features of HSIC is that it allows one to measure the dependence between very complex structures. For example, $\mathcal{X}$ and $\mathcal{Y}$ can be the space of trees with n nodes and strings of length l respectively. In Section 5.1, we will exploit this remarkable property by using kernels over graphs to measure the dependence between functional and structural data. The key ingredients for calculating HSIC are the kernels k_X and k_Y. Note that we define two different kernels for independence tests since $\mathcal{X}$ and $\mathcal{Y}$ can be different i.e., X and Y can be over different domains. Throughout the paper, we use the Gaussian kernel for our experiments i.e., the following kernel :

$$k(x, y) = \exp(-\|x - y\|^2/\sigma^2)$$

as the kernel k whenever X is an Euclidean space. For $\text{MMD}(X, Y)$, σ is chosen through a popular heuristic approach called median heuristic and is essentially the median of pairwise distances between samples of X and Y respectively. In the case of HSIC, σ_X and σ_Y are each chosen using median heuristic on samples for X and Y respectively.

4.3 Independence Tests

An important component of our paper is to test for dependence between functional and structural data. Independence tests are very useful for this purpose. Let $I : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a test statistic for measuring the dependence of random variables. We assume that $I(X, Y)$ is 0 if and only if random variables X and Y are independent. For testing independence, we define the following hypothesis test:

$$H_0 : I(X, Y) = 0 \text{ v.s. } H_1 : I(X, Y) \neq 0.$$

A low p-value indicates dependence between random variables X and Y. We do not reject the null hypothesis when $p > \alpha$ and reject it otherwise. Here, α is the significance level. Throughout the paper, the null distribution, and thereby p-value, is calculated using permuting the samples of either X or Y (this gives us samples from the product distribution i.e., empirical distribution under the null hypothesis). We either use mutual information (likelihood ratio test) or HSIC as I in all our experiments.

4.4 Lasso

Lasso is a popular linear regression model that promotes sparsity (Tibshirani, 1994). It essentially solves the least square problem with an additional l_1-regularization to encourage sparsity of the solution. In particular, suppose $\{(Y_i, Z_i)\}$ are samples from a distribution such that $Z_i = \beta^\top Y_i + \eta_i$, where η_i is a random variable with mean zero and bound variance, then lasso solves the following

optimization problem for estimating β:

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^{n} \|Z_i - \beta^\top Y_i\|^2 + \lambda \|\beta\|_1.$$

The l_1 penalty promotes sparsity in β. This is particularly relevant to our setting since we are dealing with reasonably high dimensional data and require (or expect) the solution to be sparse. Many fast solvers have been developed for solving the optimization problem. We refer interested readers to Friedman et al. (2010); Tibshirani et al. (2011) for additional details. In Section 6, we will use lasso to understand the relationship between functional and structural connectivity.

4.5 Canonical Correlation Analysis (CCA)

Let Z_1 and Z_2 be random vectors in $\mathbb{R}^d$. Quite often we are interested in finding relationship between components of Z_1 and Z_2. CCA is a natural approach to tackle this problem. The main goal in CCA is to find linear combination of components of Z_1 and Z_2 that have high correlation. Formally, this amounts to finding vectors $a, b \in \mathbb{R}^d$ such that correlation between $a'Z_1$ and $b'Z_2$ is maximized i.e.,

$$\arg\max_{a,b} \text{corr}(a'Z_1, b'Z_2),$$

where corr denotes the Pearson correlation coefficient. The resultant vectors a and b are called the first pair of canonical variables. The second set of canonical variables can be obtained by maximizing correlation subject to the additional constraint that the variables are uncorrelated with the first pair of variables. This process can be continued up to d times (given at least d samples of the random variables) to yield the canonical variables. The computation of the these variables can be computed using singular value decomposition (SVD). More details on CCA can be found in the tutorial by Borga on this subject (Borga, 2001).

4.6 Notation

We denote the set $\{1, \ldots, n\}$ by $[n]$. Suppose Y is a tensor on $\mathbb{R}^{d_1 \times d_2 \times d_3}$ then use $Y[:, \backslash j, i]$ to denote matrix M on $\mathbb{R}^{d_1 \times (d_2 - 1)}$ which represents the matrix $Y[:, :, i]$ without the j^{th} column. For a tensor Y, we use Y_i to denote the matrix $Y[:, :, i]$. With slight abuse of notation, we use Φ_i and S_i for the FCG and structural connectivity graphs of the i^{th} subject respectively. Suppose $\{(P_i, Q_i)\}_{i=1}^{n_{\text{test}}}$ where $P_i \in \mathbb{R}^p, Q_i \in \mathbb{R}$ for $i \in [n_{\text{test}}]$ represents the test data, then R^2 of a prediction function $f : \mathbb{R}^p \to \mathbb{R}$ is defined as follows:

$$R^2[f] = 1 - \frac{\sum_{i=1}^{n_{\text{test}}} (Q_i - f(P_i))^2}{\sum_{i=1}^{n_{\text{test}}} (Q_i - \frac{1}{n_{\text{test}}} \sum_{k=1}^{n_{\text{test}}} Q_k)^2}. \tag{2}$$

We will drop the f from the above notation whenever f is clear from the context.

4.7 Implementation Details

The implementation for all the experiments in this paper is in Matlab 2015a. We use matlab implementation of Glmnet (Qian et al., 2013) for lasso and ridge regression. We use the matlab implementation of canonical correlation analysis for CCA.

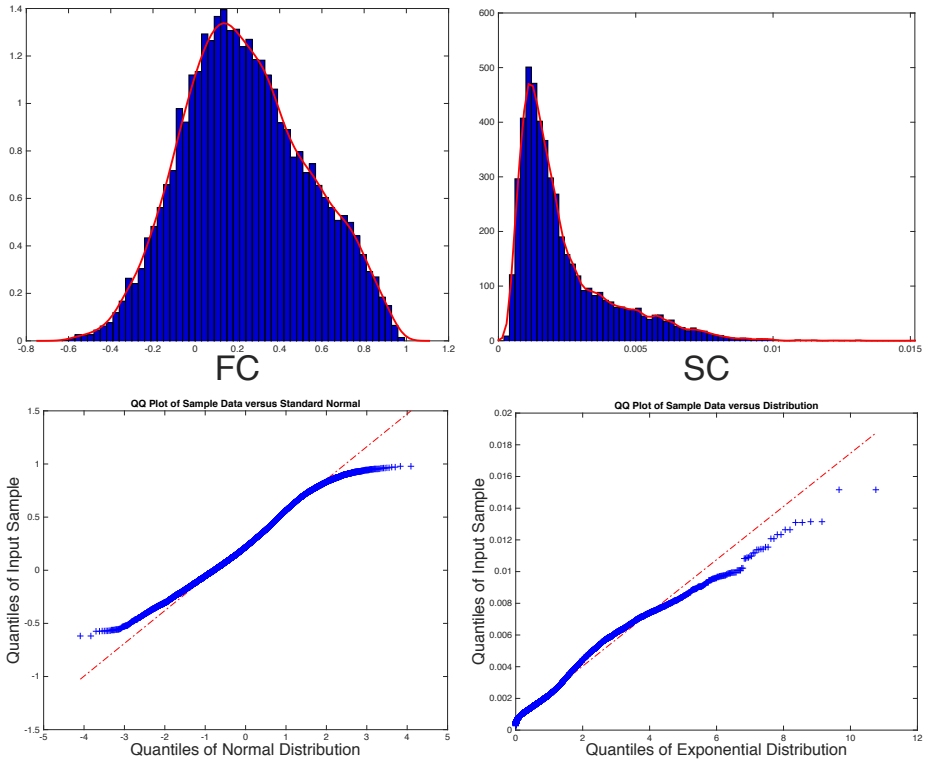

Figure 2: Distribution of the weights of edges in functional correlation graph (FC) using the function data and structural connectivity graph (SC) between ROIs for a subject. The top row figures show the histogram plots for the distribution. The quantile-quantile (qq) plots for the corresponding distributions are shown in the bottom row. The distributions have characteristics similar to a normal and an exponential distributions respectively.

Data Analysis

We will look at various analyses that are devoted to addressing the questions raised in Section 3. We first look at the distributions of the functional and structural data. Figure 2 shows the distribution of the entries of the functional correlation graph and structural connectivity matrix for a subject. The histogram and quantile-quantile (qq) plots for the subject are shown in the figure. The distributional characteristics are similar for all other subjects. It can be seen that the distribution of functional correlation graph looks very similar to a normal distribution (this is much more evident in the qq-plot). The distribution of the entries in the structural connectivity matrix is very close to an exponential distribution (again this is more evident in the qq-plot, though the tail distributions are slightly different). Similar distribution behavior was observed across subjects.

5. Analysis of dependence between Functional Data and Structural Data

Our primary goal in this section is to explore the dependence relationship between functional and structural data. We investigate various methods to further our understanding about this relationship.

We take the route of hypothesis testing to measure the dependence between the functional and structural data.

5.1 Direct Independence Test using Graph Kernels

In this approach we would like to directly use the samples $\{(\Phi_1, S_1), \ldots, (\Phi_n, S_n)\}$ to perform independence test. Recall $\Phi_i \in \mathbb{R}^{d\times d}$ and $S_i \in \mathbb{R}^{d\times d}$ denote the FCG of i^{th} subject and $S[:,:,i]$ respectively. We use kernel HSIC based test for testing independence. Note that the number of samples for test, being equal to the number of subjects, is considerably small. As mentioned earlier in Section 4.2, HSIC can be used to measure dependence between complex structures. Here, we would like to measure the dependence between functional correlation graph and structural connectivity. Hence, we need to use kernels over graphs in this hypothesis test. For the purpose of this experiment, we use the random walk graph kernels. The random walk kernel is based on the number of matching walks obtained by performing random walks on the given pair of graphs. We refer the reader to Vishwanathan et al. (2010) for more details on the random walk graph kernel. We use 10,000 random shuffles for calculating the null distribution of the hypothesis test. The hypothesis test using other graph kernels like shortest path kernel yielded similar result.

Test	p-value
HSIC based Test	0.01

Table 1: Direct Independence Test for functional and structural data.

We delay the analysis of the result until discussion of the following hypothesis test for measuring dependence between the functional and structural data.

5.2 Similarity-based Hypothesis Test

For the purpose of this test, we need a distance function between sample sets. In particular, we need a semimetric function $\rho_f : \mathbb{R}^{T\times d} \times \mathbb{R}^{T\times d} \to \mathbb{R}$. Furthermore, we also need a distance function on matrices on $\mathbb{R}^{d\times d}$. We use $\rho_s : \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d} \to \mathbb{R}$ to denote the corresponding semimetric on the connectivity matrices. We consider the random variables Y and Z that represent $\rho_f(X_i, X_j)$ and $\rho_s(S_i, S_j)$ respectively for randomly selected subjects i and j. Recall that X_i denotes the functional data $X[:,:,i]$. We first note that if the structural and function data are generated independently, then $I(Y, Z) = 0$. Therefore, we use the following hypothesis test:

$$H_0 : I(Y, Z) = 0 \text{ v.s. } H_1 : I(Y, Z) > 0.$$

We call this test a pairwise independence test. The key intuition for this test is to test whether subjects who share structural similarity exhibit functional similarity as well. The null hypothesis is that the test statistic is 0 while the alternative is that the test statistic is non-zero indicating dependence. Therefore, rejecting the null hypothesis and low $p-$value are reasonable indicators of dependence between functional and structural data.

Note that Y and Z can be calculated for all pairs of subjects in our dataset. However, the samples are no longer i.i.d. To mitigate this issue, we use bootstrapping where we sample entries these pairs uniformly with replacement across all pairs of subjects. However, in our experiments, the random variables Y still share the same datasets across subjects. Hence, the results of this independence test should be treated with caution.

Test	p-value
Likelihood Ratio Test	0.036
HSIC based Test	0.012

Table 2: Pairwise independence Test for functional and structural data.

Table 2 shows the results for hypothesis test. For this experiment, we use MMD (Equation 1) for ρ_f and frobenius norm of difference as ρ_s. The null distribution was calculated using 10,000 random shuffles. The results indicate that there is a dependence between functional and structural data. It is important to note that all these tests can detect non-linear relationships between the Y and Z.

The p-values in both the aforementioned hypothesis tests are considerably low. For example, with a significance threshold of $\alpha = 0.05$, we would reject the null hypothesis (i.e., the functional and structural data are independent). Hence, our statistical analysis clearly assert dependence between the functional and structural connectivity. However, we have not yet explored the strength and nature of this dependence. In the next few sections, we tackle this problem by demonstrating the nature of this relationship from various angles.

5.3 Measuring dependence within subjects

We would like to test if there is correlation between functional and structural data across all the ROIs within subjects. Our analysis is divided into two components: (a) dependence over edges of the structural network; (b) dependence over ROIs of the brain.

- **Dependence over edges:** Recall we define structural connectivity (SC) between two ROIs as the normalized QA between them. Functional connectivity (FC) between two ROIs is the Pearson correlation between the time series obtained for the ROIs i.e., entry in FCG corresponding to the pair of ROIs. More formally, suppose non-zero entries of S_i and corresponding entries in Φ_i are samples from a joint distribution. Recall that Φ_i represents the FCG for the i^{th} subject. Let random variables A and B represent the (non-zero) entry in S_i and corresponding entry in Φ_i between two randomly selected ROIs. Then SC-FC correlation for a subject represents the correlation corr(A, B). We report the SC against FC results for two subjects in Figure 3. Figure 4 shows the histogram of SC-FC correlation for all the subjects. The mean and standard deviation across subjects are 0.42 and 0.03 respectively. The correlations we report have p-value $\ll 1e-3$. It is clear from the figures that SC and FC exhibit statistically significant relationship. Furthermore, it can also be seen that this relationship is consistent across subjects.

- **Dependence over ROIs:** In this case, we define SC (ROI) of an ROI as the sum of normalized QA over all the edges connected to that particular ROI. Similarly, FC (ROI) of a particular ROI is defined the sum of absolute values of the Pearson correlations of the structurally connected edges at that ROI. The ROI notation is used in order distinguish it from SC and FC

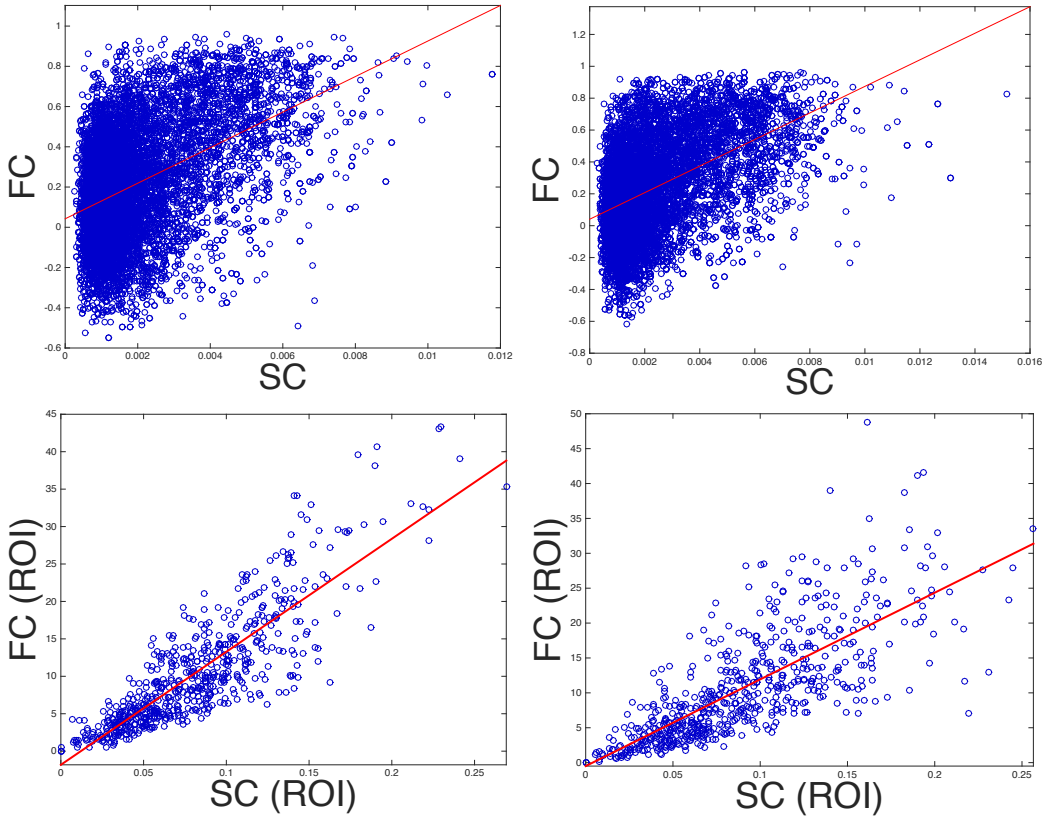

Figure 3: (Top plots) Structural connectivity (SC) vs functional connectivity (FC) for two subjects across pairs of ROIs. Each point in the top plots represents an edge between two ROIs. (Bottom plots) SC vs FC for two subjects across different ROIs. Each point in the bottom plots represent an ROI. The correlation between SC(ROI) and FC(ROI) is significantly higher than that between SC and FC.

defined above. More formally, let RS_i and RF_i be vectors in $\mathbb{R}^d$ defined as follows:

$$\text{RS}_i[j] = \sum_{k=1}^{d} S_i(k,j)$$

$$\text{RF}_i[j] = \sum_{k=1, S(k,j,i)>0}^{d} |\Phi_i(k,j)|$$

Assuming that the entries of vectors RS_i an RF_i are samples from a joint distribution with corresponding random variables A and B, SC-FC (ROI) correlation represents the correlation $\text{corr}(A, B)$. Figure 3 shows the results of SC (ROI) against FC (ROI) for the two subjects used in the previous experiment. Figure 4 shows the histogram of SC-FC (ROI) correlation for all the subjects. The mean and standard deviation across subjects are 0.81 and 0.04 respectively. The correlations we report have p-value $\ll 1e{-}3$. The results clearly show that there is strong

relationship between FC and SC over ROIs. High correlation was observed consistently over all the subjects.

Our results indicate that the dependence over ROIs is much stronger than the dependence over edges. More specifically, the relationship between the statistics of FC and SC over ROI is much higher than that obtained from individual edges. We hypothesize that this stronger relationship is due to the fact that FC at a particular edge does not solely depend on the SC at that particular edge. However, the aggregate of structural connectivity over the ROI (represented by SC(ROI) in our experiments) is a better reflection of the functional activity at a particular ROI. Hence, we believe, that this yields much stronger SC-FC relationship than that obtained over edges.

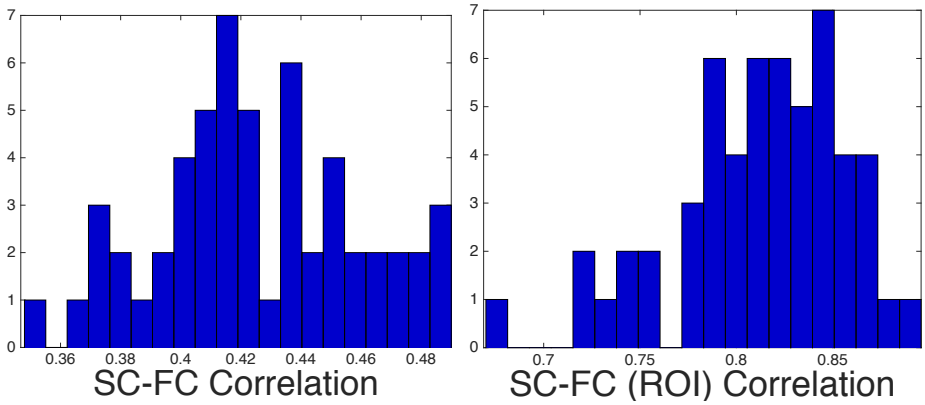

Figure 4: Histogram of correlation within subjects (for all subjects) to measure correlation over edges (left plot) and ROIs (right plot). The correlation between SC(ROI) and FC(ROI) is significantly higher than that between SC and FC.

5.4 Measuring dependence across subjects

In the previous experiment, we analyzed the SC-FC relationship within a subject. However, this does not provide any information about this relationship across subjects. Our goal in this section is to investigate the relationship between functional and structural data across subjects for different ROIs. In particular, we would like to identify the regions of the brain that exhibit strong SC-FC relationship and those with weaker relationship. For the purpose of this experiment, similar to the previous approach, we divide our analysis over edges and ROIs.

- **Dependence over edges:** Similar to the previous section, we use SC and FC over edges of the structural graph. However, instead of using all edges within a subject, we measure dependence over each edge across all the subjects. Hence, for each edge, we will have n (the number of subjects) samples since the samples across subjects are assumed to be from a distribution. Figure 5 shows the results for the experiment.
- **Dependence over ROIs:** Here, similar to the previous section, we use SC (ROI) and FC (ROI) over all the ROIs. Unlike the previous section, however, we measure dependence over each ROI across all the subjects i.e., samples across subjects are assumed to be from a

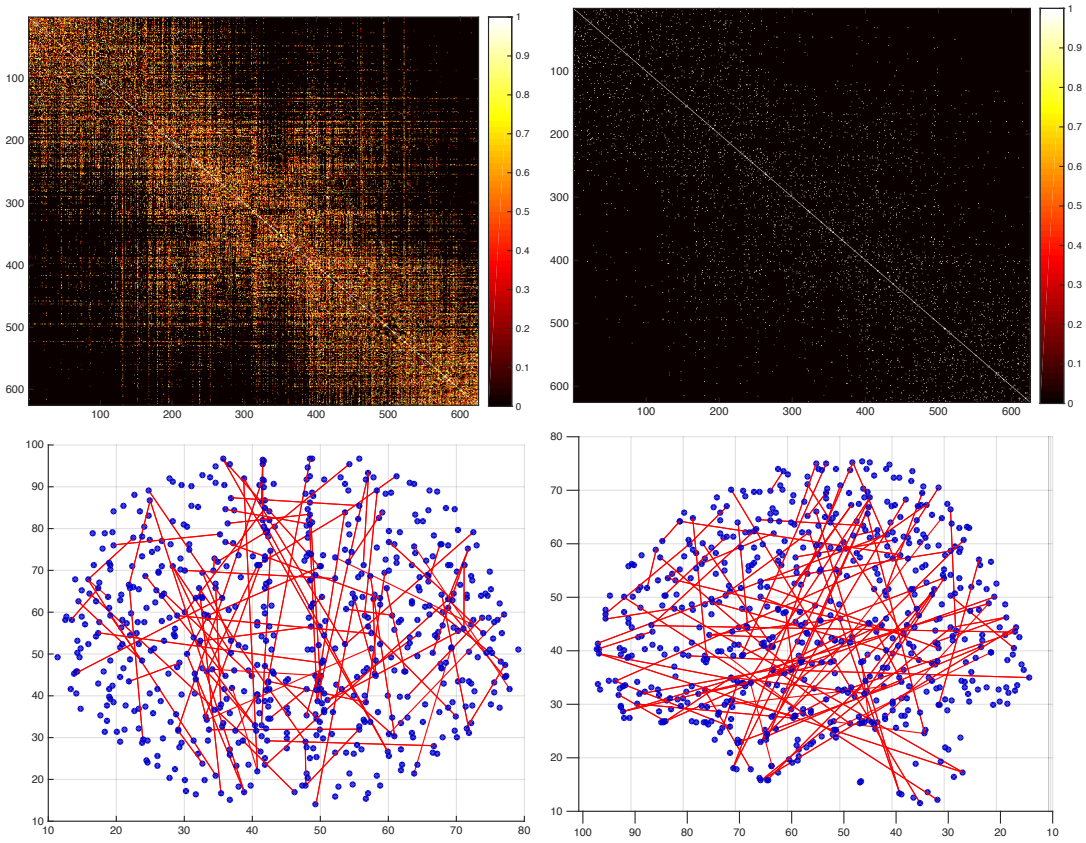

Figure 5: Independence tests for measuring dependence over edges. The top left plot shows a heat map for 1-p value for the independence tests over edges across subjects. Higher value indicates greater dependence in this plot. The top right plot shows the heat map for edges that have low p-value less than that obtained by false discovery rate (Genovese et al., 2002). The bottom plots show top and lateral views of the connections with very low p-value (top 200). The plots show the regions of the brain where the relationship between functional and structural connectivity is the highest.

distribution. Note that similar to the analysis over edges, we have n samples. Figure 5 shows the results for the experiment. The mean and standard deviation of correlation observed across all ROIs are 0.72 and 0.13 respectively.

Our analysis for measuring dependence across subjects uncovers several interesting relationship in the network. First, Figure 5 shows several interesting SC-FC relationships over edges. We observe that the the relationship is strong inside the hemispheres (represented by the diagonal quadrants). The inter hemispherical relationships can seen in the off-diagonal quadrants. The inter hemispherical SC-FC relationships are mostly in the prefrontal lobe. We also observed strong SC-FC relationship in the visual cortex region of the brain (see Figure 12 for a reference plot of the brain with marked regions). We also report the edges with high SC-FC connectivity. Further investigation is needed to validate and interpret the results.

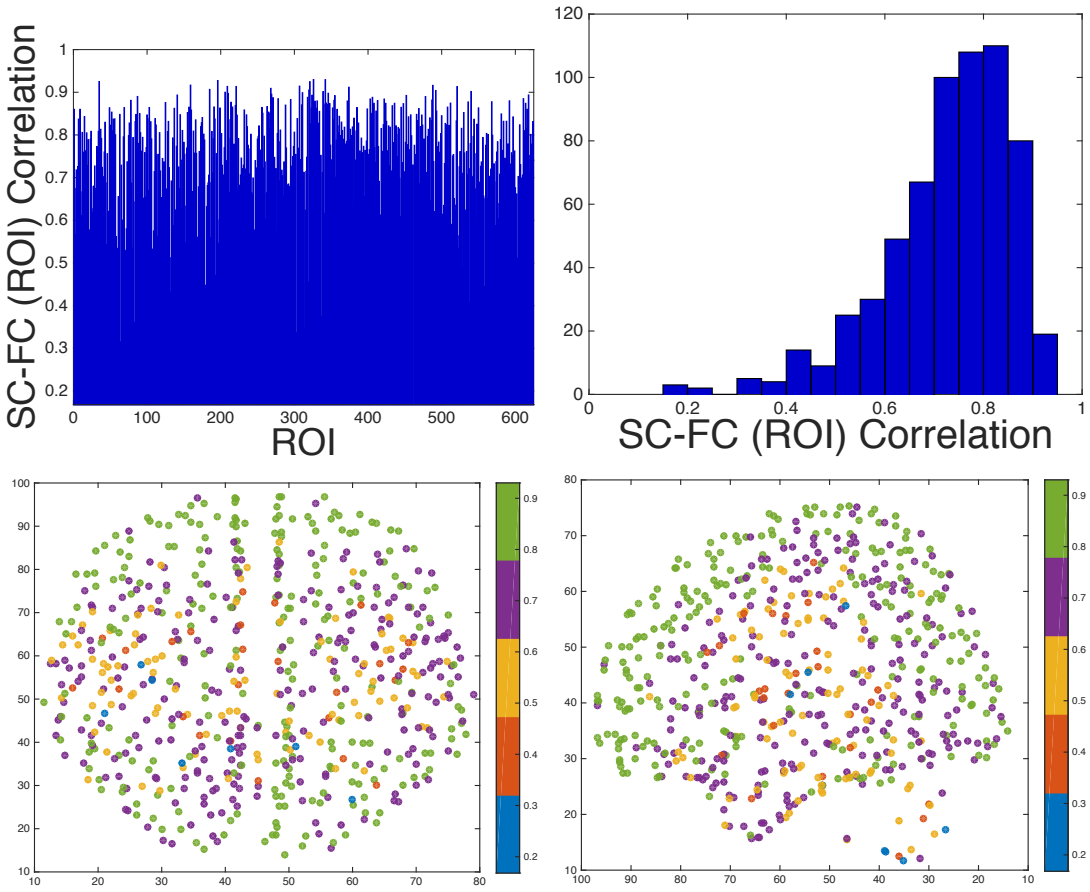

Figure 6: Correlation for SC-FC relationship across 625 ROIs. The top left plot shows the correlation across ROIs and the top right plot shows the histogram of the correlation values. The bottom plots show the variation of correlation across the brain. The bottom left and bottom right plots show the top and lateral views of the brain respectively. The correlation between SC (ROI) and FC (ROI) is high in many ROIs. The regions colored in green have the highest correlation.

We also observe interesting relationships through SC-FC (ROI). Similar to our experiment within a subject, the mean of the SC-FC (ROI) correlation is considerably high in this experiment as well. Figure 6 shows the SC-FC relationship across different parts of the brain. It can be seen that the relationship is prominent in the prefrontal, motor cortex, parietal lobe and visual cortex. It can also be observed that the relationship is slightly weaker in the cerebellum.

Interhemispherical Differences

We also analyzed differences in dependence of left and right hemispheres of the brain. It is generally believed that the two hemispheres of the brain, though serving different functions, are largely symmetric in terms of the structural and functional relationship. Our experiments provide a similar

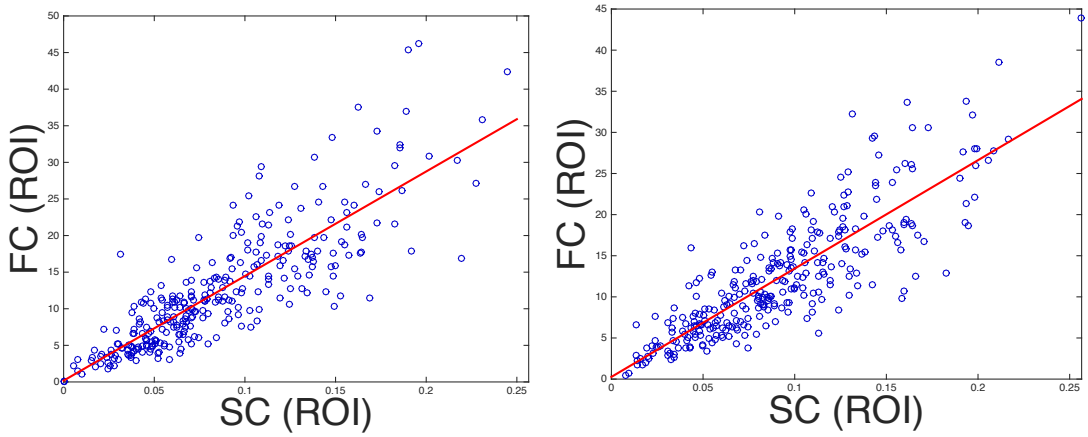

Figure 7: SC-FC (ROI) relationship for ROIs in left hemisphere (left plot) and right hemisphere (right plot) for a subject. The correlation between SC(ROI) and FC(ROI) is similar in both the hemispheres.

conclusion. For example, the plots for SC-FC ROI relationship for the left and right hemisphere, on same subjects as in Figure 3, is given in Figure 7. The correlation in the two hemisphere is similar, with correlation of SC-FC being slightly more in the left hemisphere than right hemisphere. The mean of SC-FC correlation across subjects for left and right hemispheres are 0.82 and 0.80 respectively. The corresponding standard deviation is 0.06 in both the cases. Further investigation is required to check if the SC-FC relationship across hemispheres is different in other aspects.

The key question that remains is whether this dependence can be utilized in a reasonable manner for any predictive analysis based on the data at hand. This will be our focus in the next few sections.

6. Analysis of Functional Data using Structural Data

In the previous section, we showed through various analysis that there is dependence between structural and functional data of the brain. We now shift our focus to address the second goal: Is structural information useful for predictive analysis of functional data. A simple approach for tackling this problem is to use regression analysis. Our goal is to predict the functional data at each ROI based on the rest of the structural and functional data. In particular, we consider the problem of regression on the data: $(F_{ij}, Y_{ij}) = (X[:, \backslash j, i], X[:, j, i])$ where $F_{ij} \in \mathbb{R}^{T \times d-1}$ and $Y_{ij} \in \mathbb{R}^T$. Recall that $X[:, \backslash j, i]$ represents the matrix $X[:, :, i]$ without the j^{th} column. Each row of F_{ij} represents an observation with corresponding entry of Y_{ij} as the regressed value. More formally, we assume the following noise model:

$$Y_{ij} = F_{ij}\beta_{ij} + \eta_{ij} \text{ for } j \in [d], i \in [n],$$

where η_{ij}'s are zero-mean independent random variables with variance σ^2. Our goal is to solve the regression problem and estimate the β_{ij}'s used for generating the samples, one for each subject per ROI. We use lasso for regression since our problem is high dimensional and we are interested in sparse solutions. We solve the following optimization problem to accomplish the aforementioned

task:

$$\min_{\beta_{ij}} \frac{1}{2}\|Y_{ij} - F_{ij}\beta_{ij}\|^2 + \lambda_{ij}\|\beta_{ij}\|_1 .$$

Simply using this approach for every pair of ROI and subject generally yields poor results. This is due to the fact that the problem is high dimensional and the length of the time series is quite small. Instead, we aggregate the sample across subjects and learn a parameter β_j for each ROI j. This intuition for the approach is that there is inherent relationship between functional data of different ROI, independent of subjects. This will be the primary approach throughout our analysis. More formally, we solve the following optimization problem:

$$\min_{\beta_j} \sum_{i=1}^{n} \frac{1}{2}\|Y_{ij} - F_{ij}\beta_j\|^2 + \lambda_j\|\beta_j\|_1 . \tag{3}$$

We observe that this approach does not take into account the structural information available to us. We incorporate the structural information in the form of regularization. To this end, we consider several approaches directly incorporating this information. The key goal of these approaches is to show that structural connectivity plays a role in functional predictive analysis. We use the following methods for this purpose.

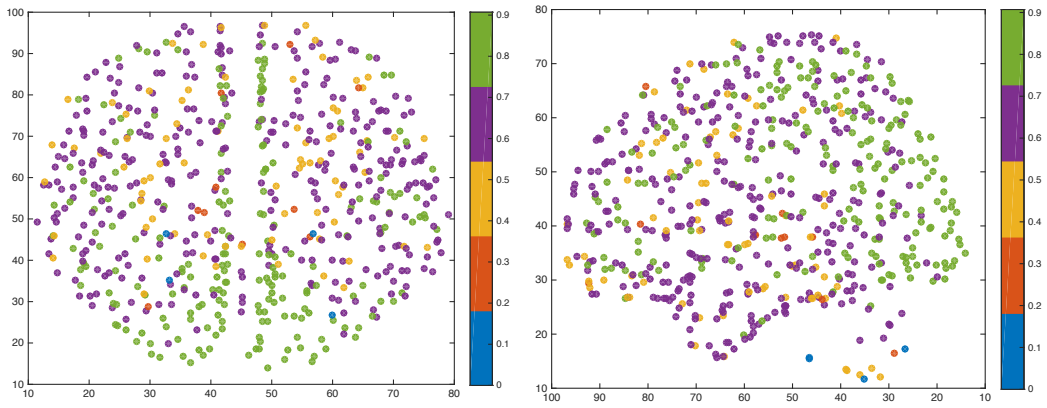

Figure 8: R^2 value across different ROIs of the brain obtained by using lasso. The left and right plot show the top and lateral view of the brain respectively. The regions colored in green have the highest R^2 value.

Adaptive Lasso

One of the natural ways to incorporate structural information for functional predictive analysis is to use a weighted l_1-regularization rather the standard l_1 regularization used in Equation (3). Formally, this is equivalent to solving the following optimization problem:

$$\min_{\beta_j} \sum_{i=1}^{n} \frac{1}{2}\|Y_{ij} - F_{ij}\beta_j\|^2 + \lambda_j \left\|\beta_j \circ \frac{1}{w_j}\right\|_1 \tag{4}$$

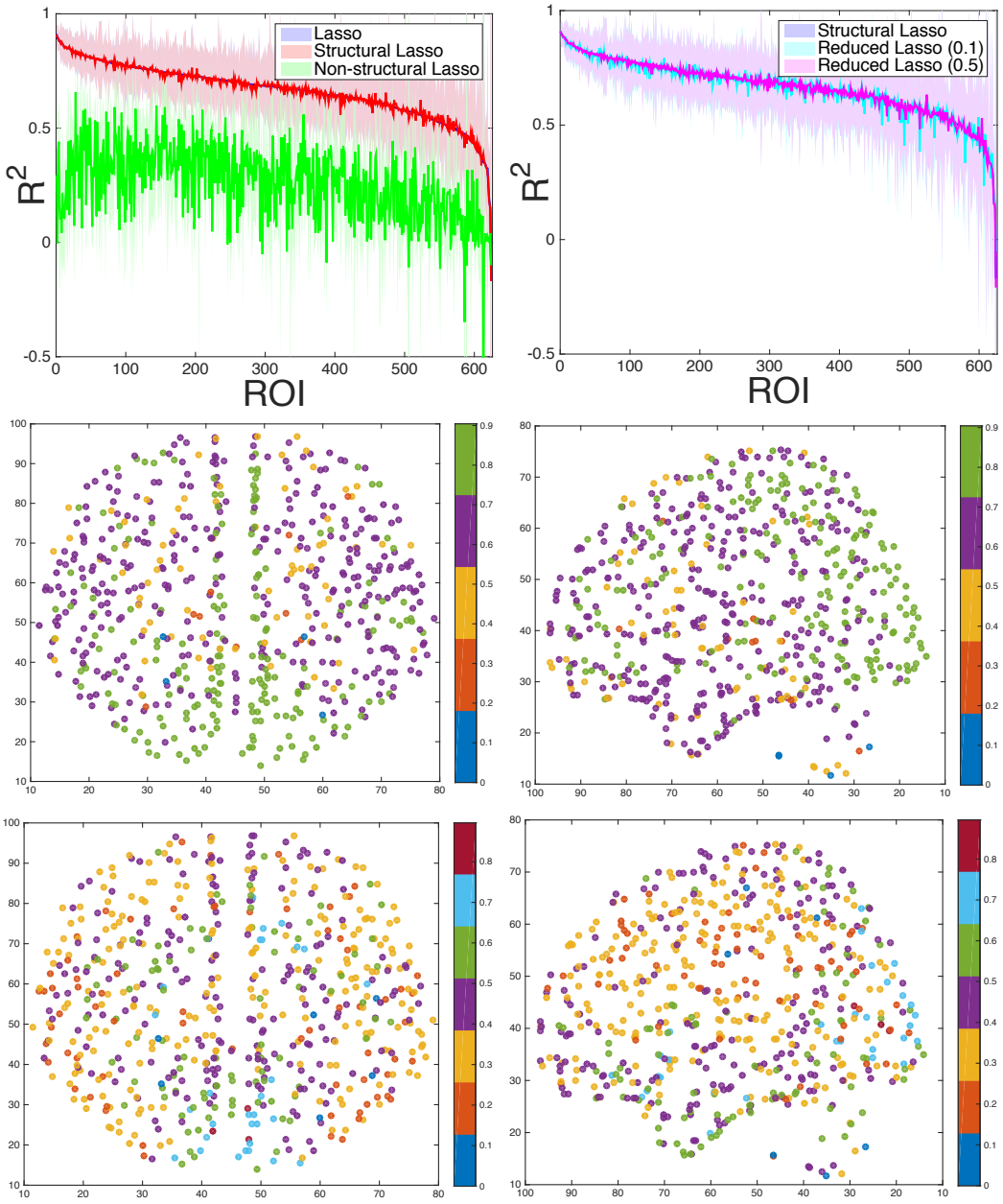

Figure 9: (Top Row Plots) R^2 value across different ROIs of the brain obtained by using lasso. The ROIs are sorted according to the R^2 value for better presentation. (Middle Row Plots) R^2 value obtained from structural lasso. (Bottom Row Plots) Difference between R^2 of structural and non-structural lasso (positive values indicate structural lasso has higher R^2 value than non-structural lasso). The left and right plot show the top and lateral view of the brain respectively. Prediction with structural information (structural lasso) is significantly better compared to the one without structural information (non-structural lasso).

where w_j represents the j^{th} thresholded row of matrix $M = [S(:, \backslash j, i) > 0]$ (where $[A > 0]$ of matrix A represents the matrix whose ij$^{\text{th}}$ entry represents the sign function of A_{ij}) averaged over all subjects i. More formally, $w_j[k] = h_t(M(j,k))$ where h_t is defined as follows:

$$h_t(x) = \begin{cases} x & \text{if } x \geq t \\ t & \text{if } x < t \end{cases}$$

We call this approach ADAPTIVELASSO. The threshold parameter t in this method will be explained in more detail later.

Reduced Lasso

The next approach is to run lasso, i.e., solve problem in Equation 3, but on the reduced support set containing only the structural edges incident on the j^{th} ROI while solving for β_j. The key motivation for using this approach is to check if the predictive performance is competitive even after restricting the support set to the structurally connected ROIs. If the performance is not significantly affected, it provides some evidence that structure has a significant role in the functional data of the brain.

In this approach, we use a parameter δ, which represents the fraction of the edges we use for reduced support set. For example, $\delta = 0.5$ indicates that only half of the highest weighted edges are considered for the reduced support. We use REDUCEDLASSO(δ) to denote this approach with parameter δ. We refer to REDUCEDLASSO(1) as structural lasso.

Non-structural Lasso

Finally, to verify the structure indeed is essential element for functional prediction, we use regression analysis where the support set is restricted to components without any edges. In other words, the structurally connected edges are explicitly removed from the support set. A low performance in this approach is, again, an indicator that of structural and functional dependence in the brain. We use NONSTRUCTURALLASSO to refer to this approach.

For all the above approaches, we choose the regularization parameter λ for each method by cross-validation. We use 75% of the subjects for training and report the R^2 on the remaining subjects. The error bars for the test subjects are also reported. For testing the adaptive lasso, we results reported are for threshold of $t = 0.1$. This was chosen by cross-validation.

Figure 8 and 9 show the results for the various lasso algorithms. The R^2 values obtained through lasso is similar to the correlations obtained in Section 5.4 (see Figure 6 and 8). The most interesting inference from this section can be seen in Figure 9. We first observe that lasso and structural lasso have roughly similar performance. Furthermore, the R^2 of lasso and structural lasso is considerably better than the obtained through non-structural lasso. This provides a strong evidence that structure indeed plays a critical role in predictive of functional data. Furthermore, perhaps surprisingly, we also observe that reduced lasso with just one-tenth fraction of edges performs as well as the full lasso. We believe that this also provides a key evidence to support relationship between functional and structural data. We omit the results of adaptive lasso because they were similar to those obtained from lasso.

6.1 Inferring FC from SC

Before ending our discussion on analysis of functional data, we consider the problem of inferring FC from SC. In particular, we would like to infer FC (ROI) of a subject from its corresponding

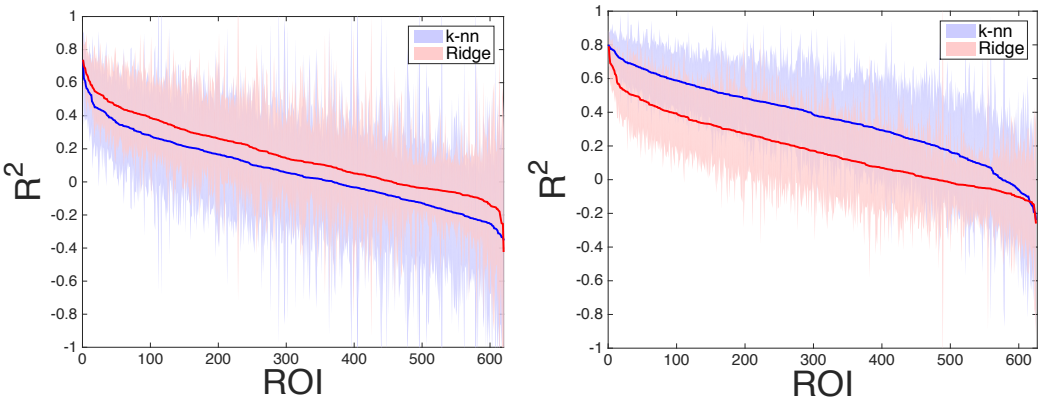

Figure 10: R^2 for inference of FC (ROI) from SC (ROI) and SC (ROI) from FC (ROI) over 625 ROIs using k-nn regression and ridge regression. The ROIs are sorted for better presentation. The error bars for each method are shown using the corresponding transparent colors. These are obtained over 50 repetitions of training on 75% of the subjects and testing on rest of them. R^2 value is considerably high across many ROIs, indicating good predictive power.

SC (ROI) information. To this end, we learn a regression function $f_j : \mathbb{R} \to \mathbb{R}$ for each ROI j, that regresses from its structural connectivity SC to the functional connectivity FC. We use two regression models for this purpose: (a) k-nn regression (a non-parametric model) (b) linear ridge regression (a parametric model). The value of k in k-nn regression and the regularization parameter in linear ridge regression were chosen by 5-fold cross-validation. Figure 10 shows the results for k-nn regression and ridge regression. The training is performed on 75% of the subjects and R^2 is reported for rest of the subjects (test data). The ROIs are sorted according to the R^2 value for clear presentation. Error bars are reported for 50 repetitions for each ROI.

7. Analysis of Structural Data using Functional Data

7.1 Inferring SC from FC

In the previous section, we looked at the approach of inferring FC (ROI) from SC (ROI). Here, we consider the problem in the other way i.e., analysis of structural data using the functional data. We use the same regression models as used in Section 6.1. Similar to the previous section, the value of k in k-nn regression and the regularization parameter in linear ridge regression were chosen by 5-fold cross-validation. Figure 10 shows the results for both the algorithms (k-nn regression and linear ridge regression). The training and testing process is similar to the one reported in Section 6.1.

From Figure 10, it can be seen that we are able to predict both SC (ROI) from FC (ROI) and FC (ROI) from SC (ROI) with considerable accuracy. It is possible to further increase the performance by using more sophisticated inference procedures. However, our main goal was to demonstrate that it is possible to infer SC and FC from each other with considerable accuracy for a large portion of the brain networks. One interesting observation is that inference of SC from FC is considerably better than that of FC from SC. We believe this is due to the fact that SC misses long range inter-

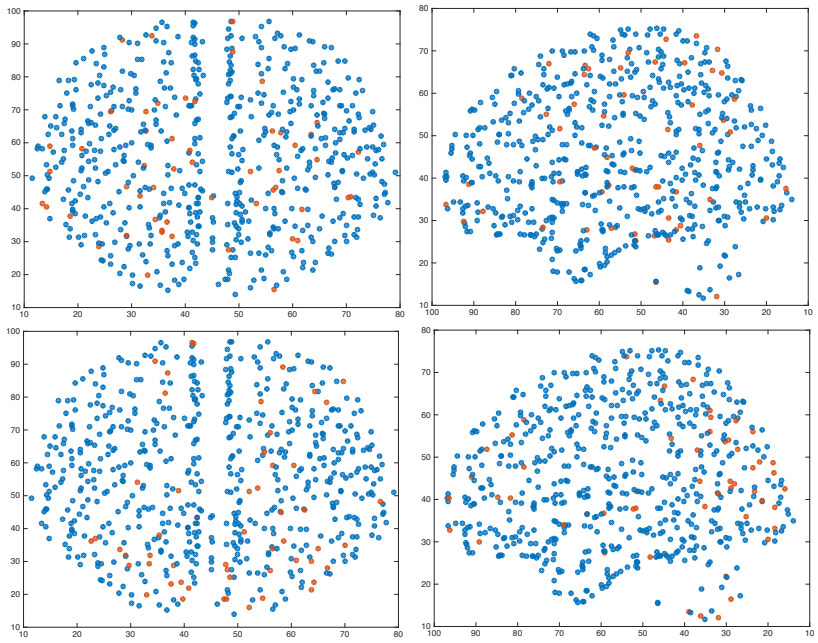

Figure 11: First pair of canonical correlation variables for FC (ROI) (top plots) and SC (ROI) (bottom plots). The red and blue portions represent the non-zero and zero components of the canonical correlation variables respectively. The regions represented in red is the group of ROIs that have high SC-FC relationship.

hemispherical connections due to nature of DSI measurements while FC inherently captures them. Hence, inference of FC from SC might be slightly more difficult.

8. Canonical Correlation Analysis (CCA)

Throughout this paper, we restricted our attention to analysis of SC-FC relationship over ROIs (within and across subjects). A natural question to consider is if there is correlation amongst groups of ROIs. This is particularly relevant in brain networks where we expect groups of ROIs to exhibit correlation amongst themselves. Here we are interested in the SC-FC relationship between groups of ROIs. In particular, the random variables used in CCA are FC(ROI) and SC(ROI) (over all the ROIs) respectively. Figure 11 shows the non-zero elements of the first pair of canonical correlation variables (see Section 4.5). The red portions of the plot represents the non-zero components of canonical variables corresponding to functional (top plots) and structural (bottom plots) data. We observe that the non-zero components are similar in both the functional and structural canonical variables. Further analysis is required to validate and understand the relationship between the ROIs obtained through our analysis.

9. Conclusion

In this paper, we examined the problem of understanding the dependence between functional and structural connectivity of the brain networks. We looked at various approaches for establishing and

understanding the nature of this relationship. We used independence test for showing that these two sources of data are indeed related. We used graph kernel based measures for performing the independence tests. To understand this relationship, we used various correlation analysis approaches. Our study shows that SC-FC relationship over ROI is much stronger than that obtained over edges. We also highlighted the variation of the SC-FC relationship over different regions of the brain by using correlation analysis across subjects. Our results show SC-FC relationship in prominent regions of the brain and is consistent with the neuroscience literature. We have also investigated techniques for predictive analysis of FC from SC and vice-versa. Our results in this context demonstrate that FC and SC can indeed be inferred from each other with considerable accuracy.

While we made several important strides in understanding the relationship between structural and functional connectivity, there are still many interesting open problems and future directions. In this paper, we mostly focused on direct anatomical connections in the brain. A natural and more precise approach to characterize structural connectivity between ROIs is by resorting to indirect connections in the brain. It will be interesting to define the notion of indirect connectivity and apply our techniques to measure dependence between functional and indirect structural connectivity. Our purpose of inferring SC and FC from each other was to demonstrate the relationship between them. We believe that the inference can be improved by using more sophisticated inference techniques. Furthermore, an important problem to tackle is exact recovery of structural connectivity from functional connectivity and vice-versa (see, for example, (Sarkar et al., 2015)). We believe that our analysis can provide some insights in this context.

Similar to many prior works in the neuroscience community, a major limitation of this study is the size of the dataset. This issue is especially exacerbated in our case because of our goal to understand the variability of SC-FC relationship across different regions of the brain. Hence, interpretation of the results reported in this paper should be exercised with caution. It is an interesting future work to apply our methods to a larger dataset.

Acknowledgement

I thank Barnabás Póczos, Aarti Singh, Roy Maxion, Timothy Verstynen and Alex Smola for discussions throughout the course of the project. This work is supported by NSF grant IIS-1247658.

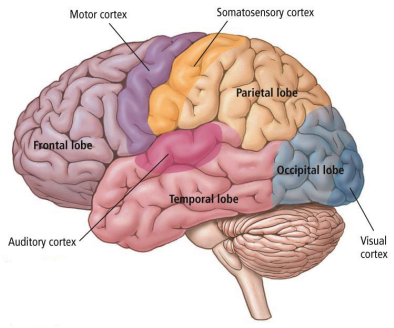

Figure 12: Reference plot for regions of the brain.

References

Magnus Borga. Canonical correlation analysis: A tutorial, January 2001. URL `http://people.imt.liu.se/magnus/cca`.

Fani Deligianni, Emma C. Robinson, Christian F. Beckmann, David Sharp, A. David Edwards, and Daniel Rueckert. Inference of functional connectivity from structural brain connectivity. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Institute of Electrical & Electronics Engineers (IEEE), 2010.

Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 2010.

Christopher R. Genovese, Nicole A. Lazar, and Thomas Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878, 2002.

A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory*, pages 63–77, 2005.

A. Gretton, K. Borgwardt, M. Rasch, B. Schölkopf, and A. Smola. A kernel approach to comparing distributions. In *In Proceedings of the 22th AAAI Conference on Artificial Intelligence*, pages 1637–1641, 2007.

C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040, 2009.

K. Jarbo and T. D. Verstynen. Converging structural and functional connectivity of orbitofrontal, dorsolateral prefrontal, and posterior parietal cortex in the human striatum. *Journal of Neuroscience*, 35(9):3865–3878, 2015.

Martin A. Koch, David G. Norris, and Margret Hund-Georgiadis. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage*, 16(1):241–250, 2002.

R. A. Poldrack. Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, 2(1):67–70, 2006.

J. Qian, J. Hastie T., Friedman, R. Tibshirani, and N. Simon. Glmnet for matlab. 2013. URL `http://www.stanford.edu/~hastie/glmnet_matlab/`.

S. J. Reddi and B. Póczos. Scale invariant conditional dependence measures. In *Proceedings of the 30th International Conference on Machine Learning, (ICML)*, pages 1355–1363, 2013.

P. A. Robinson. Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. *Physical Review E*, 85(1), 2012.

P. A. Robinson, S. Sarkar, Grishma Mehta Pandejee, and J. A. Henderson. Determination of effective brain connectivity from functional connectivity with application to resting state connectivities. *Physical Review E*, 90(1):012707, 2014.

Somwrita Sarkar, Sanjay Chawla, and Donna Xu. On inferring structural connectivity from brain functional-mri data. *Computing Research Repository (CoRR)*, abs/1502.06659, 2015.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1994.

Robert Tibshirani, Jacob Bien, Jerome Friedman, Trevor Hastie, Noah Simon, Jonathan Taylor, and Ryan J. Tibshirani. Strong rules for discarding predictors in lasso-type problems. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(2):245–266, 2011.

S.V.N. Vishwanathan, Nicol N. Schraudolph, Risi Kondor, and Karsten M. Borgwardt. Graph kernels. *Journal of Machine Learning Research*, 11:1201–1242, 2010.