Quantifying and visualizing intra-regional connectivity in resting state fMRI with correlation densities

Alexander Petersen*
Department of Statistics and Applied Probability
University of California Santa Barbara
Santa Barbara, CA 93106, USA
petersen@pstat.ucsb.edu
(805)-893-2129

Chun-Jui Chen
Department of Statistics
University of California Davis
Davis, CA 95616, USA
cjuchen@ucdavis.edu

Hans-Georg Müller
Department of Statistics
University of California Davis
Davis, CA 95616, USA
hgmueller@ucdavis.edu

July 29, 2018

Keywords: Functional connectivity; Density functions; Correlation; Cognitive performance; Functional linear model; Transformation

*Corresponding Author

Abstract

The use of correlation densities is introduced to quantify and provide visual interpretation for intra-regional functional connectivity in the brain. For each brain region, all pairwise correlations are computed between voxels within the region, and the distribution of the ensemble of these correlation values is represented as a probability density, the correlation density. The correlation density can be estimated by kernel smoothing. It provides an intuitive and comprehensive representation of subject-specific functional connectivity strength at the local level for each region. To address the challenge of interpreting and utilizing this rich connectivity information when multiple regions are considered, methods from functional data analysis are implemented, including a recently developed method of dimensionality reduction specifically tailored to the analysis of probability distributions. To illustrate the utility of these methods in neuroimaging, experiments were carried out to identify associations between local functional connectivity and a battery of neurocognitive scores. These experiments demonstrate that correlation densities facilitate the discovery and interpretation of specific region-score associations.

1 Introduction

The availability of neuroimaging data consisting of time-varying signals across spatial locations in the brain, as provided by functional magnetic resonance imaging (fMRI) or magneto- and electro-encephalography (MEG and EEG), has made it possible to study *functional connectivity*, or spatial patterns of time course similarity, in the resting human brain. These patterns are interesting in their own right, and have been used for the discovery of so-called network hubs (Buckner et al., 2009) and other structural properties of brain networks such as small-worldness (Bassett and Bullmore, 2006). An important scientific goal is to relate connectivity to other variables of interest such as age or cognitive status, which conveys external validation that brain connectivity as derived from fMRI is associated with cognitive performance. While various studies have shown the effect of aging on connectivity, a relatively unexplored topic has been to relate, at the individual level, local connectivity patterns with cognitive ability, as measured by cognitive scores. This leads to the challenge of efficiently quantifying local, intra-regional connectivity as opposed to inter-regional connectivity, where the latter is a standard topic in connectivity. In this paper we demonstrate a promising approach to achieve such quantification.

Subject-specific connectivity patterns can be investigated on two scales. At the global scale, connections are established by measuring temporal Pearson correlations between summary signals of pairs of brain regions that are spatially far apart and often analyzed using graph-theoretic approaches, after applying a threshold to the Pearson correlations (Worsley et al., 2005; Achard et al., 2006; Bassett and Bullmore, 2006; van den Heuvel et al., 2008; Buckner et al., 2009; Tomasi and Volkow, 2011; Zalesky et al., 2012). At the local scale, correlations between pairs of nearby voxels can be used to quantify the strength of local connectivity of a voxel or local region. Two popular methods for summarizing the strength of this local connectivity are functional connectivity density mapping (Tomasi and Volkow, 2010) and regional homogeneity (Zang et al., 2004). The relevance of local connectivity has been established in a vast number of studies that linked local connectivity patterns to age, gender and various diseases (Wu et al., 2009; Shukla et al., 2010; Lopez-Larson et al., 2011; Tomasi and Volkow, 2012; Zalesky et al., 2012; Han et al., 2015; Qi et al., 2015).

While intuitive, these current measures of local connectivity sacrifice potentially valuable information as they only represent scalar numeric summaries of a large number of correlations. Another intuitive approach, demonstrated in Petersen and Müller (2016), is to

assemble all correlations within a region to obtain a probability density function, the *correlation density*, for which one can use kernel estimation (Rosenblatt, 1956; Parzen, 1962; Wand and Jones, 1995) or smoothing of histograms (Silverman, 1986). Characteristics such as the mode, spread, and shape of these correlation densities contain potentially valuable information that cannot be quantified by any single numeric summary.

We demonstrate a method to efficiently incorporate information from multiple brain regions to visualize and interpret intra-regional connectivity, as well as to investigate associations between local resting state connectivity and cognitive ability. The proposed approach is based on statistical methodology from functional data analysis (FDA) (Ramsay and Silverman, 2005; Wang et al., 2016). The key tool is a recent method of dimensionality reduction specifically tailored to probability density functions (Petersen and Müller, 2016), where it was also shown in an exploratory fashion that subject-specific correlation densities for a single region, corresponding to a functional network hub, are useful in predicting executive function test scores. We illustrate the full utility of this approach in connectivity analysis by considering associations between correlation densities from mutiple functional connectivity hubs and a battery of four cognitive test scores, as well as methods of inference and visualization to discover and interpret these associations (Chen and Lynch, 2017). An important point not considered in Petersen and Müller (2016) is that the different correlation densities of a single subject that are associated with different functional connectivity hubs are statistically dependent, so that it is necessary to consider them jointly as predictors in a regression model, implementing appropriate model selection methods. While FDA methods have been utilized previously in functional connectivity studies (Bassett et al., 2012; Hosseini et al., 2012), another important aspect of the approach studied here is the use of subject-specific connectivity information, rather than the usual group based approach, where one works with averages across the subjects belonging to a group.

2 Materials and Methods

2.1 Participants, fMRI acquisition and preprocessing

The example data used here to demonstrate our methods are from a study of elderly participants in a longitudinal study of cognitive impairment that has been described previously in Hinton et al. (2010). Participants were evaluated within the research program of the University of California, Davis Alzheimer's Disease Center as described in He et al. (2012), where

also a description of the clinical evaluation of this cohort and the neuropsychological test battery is provided. Included in our analyses are a group of 164 cognitively normal subjects and a second group of 63 subjects diagnosed with Alzheimer's disease.

As described previously (He et al., 2012), fMRI scans were performed at the UC Davis Imaging Research Center on a 1.5 T GE Signa Horizon LX Echospeed system, along with an 8-minute axial echo-planar imaging (EPI) BOLD fMRI scan. Subjects were provided with no specific instructions prior to the acquisition other than to keep their eyes open. The scan parameters were: TR 2.0s, TE 40ms, FOV 22 cm, Flip angle 90, 24 5 mm thick contiguous slices with bandwidth 62.5 KHz, and 64 x 64 matrix with R-L frequency encode direction. This sequence provided 240 time points of data at each voxel.

The standard preprocessing steps of slice timing and head motion correction, followed by coregistration to the subject's 3DT1 MRI scan, were performed. Multiple linear regression was applied to the signal at each voxel to remove global linear trends to account for signal drift, as well as global cerebral spinal fluid and white matter signals. Finally, a band-pass filter was applied, preserving frequency components between 0.01 and 0.08 Hz. These steps were performed in MATLAB, using the Statistical Parametric Mapping (SPM8, `http://www.fil.ion.ucl.ac.uk/spm`) and Resting-State fMRI Data Analysis Toolkit V1.8 (REST1.8, `http://restfmri.net/forum/?q=rest`). The first four time points were discarded to eliminate non-equilibrium effects of magnetization. Time points with large head motion, defined as translation greater than 1.5mm and/or rotation greater than 1.5° were then identified, and participants with any such time points were excluded.

2.2 Functional Principal Component Analysis

The statistical analysis of a random sample $X_1, \ldots, X_n$, where $X_i : [0,1] \to \mathbb{R}$ are functions with a common domain, is known as functional data analysis (Ramsay and Silverman, 2005; Wang et al., 2016), or simply FDA. Due to the infinite dimensionality of function spaces, dimension reduction is a key FDA technique. Many FDA methodologies rely on the Karhunen-Loève expansion

$$X_i(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(t), \tag{2.1}$$

where $\mu(t) = E(X(t))$ is the mean function, ϕ_k are eigenfunctions associated with the covariance kernel $G(s,t) = E\left[(X_i(s) - \mu(s))\, X_i(t)\right]$ with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq$

$\cdots \geq 0$, and

$$\xi_{ik} = \int_0^1 (X_i(t) - \mu(t))\phi_k(t)\ \mathrm{dt} \tag{2.2}$$

are the uncorrelated functional principal component scores with mean zero and variance λ_k. Expansion (2.1) provides a linear representation of the data, the functional principal component analysis (FPCA). These representations are the infinite-dimensional analog of principal components analysis (PCA) for multivariate data.

Dimensionality reduction is then obtained by truncating the sum to a finite number of components, as well as visualization of the average behavior of the process via the mean function μ and the dominant modes of variation $\mu \pm \alpha\phi_k$ (Jones and Rice, 1992), which lend interpretability to the scores ξ_{ik}.

2.3 Functional Data Analysis for Density Functions

To perform dimension reduction for a sample of univariate probability distributions, which are often best represented and visualized through their density functions $f_1, \ldots, f_n$, direct application of standard FDA methodology (Kneip and Utikal, 2001) has proven suboptimal (Petersen and Müller, 2016). This is due to the nonlinearity of the space of densities implied by the constraints that a density be positive and integrate to one. A more promising approach is to transform the densities into a Hilbert space via a nonlinear functional transformation ψ, yielding a sample $X_i = \psi(f_i)$, to which the standard FDA methodology can then be suitably applied, as the transformed functions are not subject to any restrictions.

To illustrate the utility of such a nonlinear transformation, consider a basic setup in finite dimensions. Here, one observes

$$Y_i = (\sin(Z_{i1} + Z_{i2}), \cos(Z_{i1} - Z_{i2})),$$

where $Z_i = (Z_{i1}, Z_{i2})$ are a bivariate normal sample with $\mathrm{Var}(Z_{i1}) = 1$, $\mathrm{Var}(Z_{i2}) = 0.2$ and correlation 0.95. The left panel of Figure 1 shows the data and first PCA direction (i.e., the orthogonal least squares line) for a sample $Y_1, \ldots, Y_{40}$, which is clearly unsatisfactory due to the nonlinear nature of data. Alternatively, one can transform the data and obtain the first PCA loading (orthogonal least squares line) of the Gaussian data $Z_1, \ldots, Z_{40}$ (right panel of Figure 1), which is appropriate due to the linear nature of these data, and then map it back to the original space, obtaining the curved dashed line in the left panel, and thus a much more informative analysis.

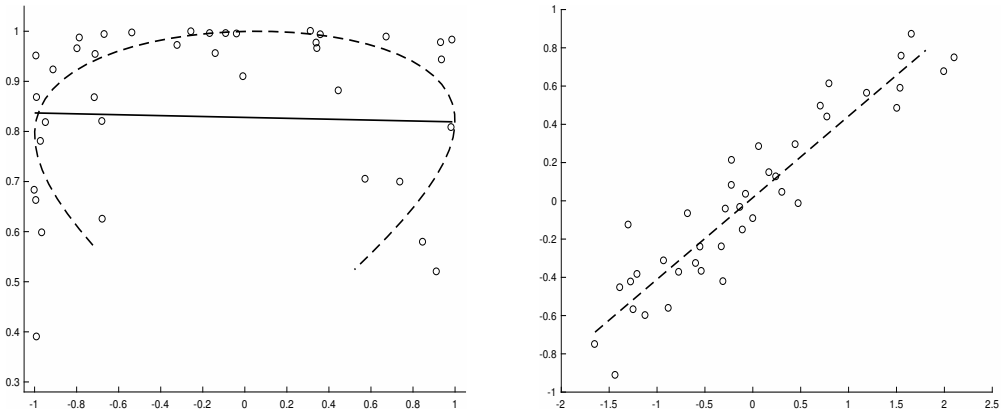

Figure 1: (Left) Scatterplot of nonlinear data Y_i, with linear PCA loading (orthogonal least squares line) and nonlinear loading (dashed line) obtained by transforming the loading obtained for the data X_i (dashed curve); (Right) Original data X_i with first PCA loading (orthogonal least squares; dashed line).

Of course, in the case of density functions, one cannot so easily transform the densities so that the transformed versions have a known distribution, such as a Gaussian process. However Petersen and Müller (2016) discussed several generic transformations that transform density functions into unrestricted square integrable functions and thus improve upon the naive FPCA directly applied to the density functions. The most promising transformation utilizes the so-called quantile density function $q(t) = Q'(t)$, where Q is the quantile function corresponding to a density f. The log-quantile density (LQD) transformation of f is

$$X(t) = \psi(f)(t) = \log(q(t)) = -\log(f(Q(t))), \quad t \in [0,1]. \tag{2.3}$$

If one starts with a sample $f_1, \ldots, f_n$ of density functions, this transformation then gives rise to a sample of unrestricted square integrable functions $X_1, \ldots X_n$ on the domain $[0,1]$, for which all available FDA methods can be applied, including functional regression or FPCA. By computing the elements of the decomposition in (2.1), the dominant modes of variation are most usefully visualized as densities by means of the inverse transformation. The transformation modes are obtained across a range of values α as

$$\psi^{-1}\left(\mu \pm \alpha\phi_j\right). \tag{2.4}$$

2.4 Quantifying Intra-regional Connectivity by Correlation Densities

Our experiments focused on the intra-regional connectivity properties of ten functional hubs, which are listed together with their seed voxels in Buckner et al. (2009) (see Table 3 therein) and correspond to the following regions: left/right parietal lobules (LPL/RPL),

left/right middle frontal (LMF/RMF), left/right middle temporal (LMT/RMT), medial superior frontal (MSF), medial prefrontal (MP), posterior cingulate/precuneus (PCP) and right supramarginal (RS). Since only the seed voxels, and not the region boundaries, were well-defined, we took the following approach to contruct the correlation density for each region at the subject level. First, for a given region, an $11 \times 11 \times 11$ cube of voxels was isolated, centered at the seed voxel, and a mask was then used to extract the signals of all gray matter voxels within this cube. Second, letting m be the number of identified gray matter voxels including the seed voxel, the preprocessed fMRI signals for each were used to obtain pairwise Pearson correlations ρ_k, $k = 1, \ldots, m-1$, between the seed voxel signal and the remaining $m-1$ gray matter signals. Alternative similarity measures besides Pearson correlation could also be used. As a final step, only gray matter voxels for which ρ_k is positive are considered to be part of the corresponding region, so that negative correlations are effectively discarded.

The elimination of negative correlations is similar to other approaches in the analysis of local functional connectivity (Tomasi and Volkow, 2010), and is also in line with **?** THIS REFERENCE IS MISSING which specified functional regions of interest as "spatially coherent regions of homogeneous functional connectivity." Moreover, we found that negative correlations were rare for voxels nearby the seed voxel and mostly concentrated on the boundary of the cube.

The correlation density for a specific region and a specific individual is then defined as the probability density function of the distribution of the correlations ρ_k that are positive. The domain of the correlation densities is the interval $[0, 1]$. The estimated correlation densities f_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, 10$, for each of n subjects are obtained from the sample of correlations for the jth region by applying a density estimation method such as kernel density estimation (Rosenblatt, 1956) or smoothing histograms (Petersen and Müller, 2016). The correlation densities provide useful visualizations of intra-regional connectivity, for each region considered and each individual in the sample; see Section 3.

Once the correlation densities f_{ij} have been obtained, we apply transformations (2.3) to obtain corresponding unrestricted functions X_{ij} and then FPCA for dimension reduction, applying (2.1) and truncating the expansion at K_j expansion terms. Here we select K_j so that 95% of the variation in the functions X_{ij} across individuals is explained. This dimension reduction step then results in a vector of dimension K_j of functional principal components that represent the functions X_{ij} for each individual and thus the correlation densities f_{ij}. These vectors can then be used as components of various statistical models, e.g., as predictors

in a regression model. We demonstrate this approach in Section 3.

2.5 Correlation Densities as Predictors in a Functional Linear Model

We aim here at regression models where the subject-specific vectors of FPCs of the transformed density functions for the regions considered are the predictors and the responses are the cognitive scores. This allows to quantify the relations between subject-specific intra-regional connectivities and cognitive performance. Specifically, in Section 3 we utilize a subject's local connectivity characteristics, as quantified by the distributions of correlations for several brain regions, to predict each of four scalar cognitive scores. If J distinct regions are considered for n subjects in a sample, our regression model features predictors f_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, J$, where each of these is a density representing the distribution of correlations for the ith individual and the jth region, and Y_i is the corresponding response of the ith subject, a cognitive score. Rather than using the raw distributions as predictors, we employ the corresponding log-quantile density functions $X_{ij}(t) = \psi(f_{ij})(t)$, given in (2.3).

In the analysis one must account for the fact that age is highly associated with both connectivity (Ferreira and Busatto, 2013; Betzel et al., 2014) and cognitive functioning. We address this by adjusting the response variable Y_i, for a given cognitive score, to be the residual from the regression of the respective cognitive score on age. We then apply the functional linear regression model (Cai and Hall, 2006; Hall and Horowitz, 2007) for predicting Y_i from the X_{ij}, while accounting for age, which is

$$Y_i = \beta_0 + \sum_{j=1}^{J} \int_0^1 X_{ij}(t)\beta_j(t)\mathrm{d}t + \varepsilon_i, \tag{2.5}$$

where β_0 is a scalar parameter, the ε_i are independent and identically distributed errors with zero mean, and the β_j, $j \geq 1$, are square-integrable functional parameters to be estimated.

A well-known issue with (2.5) is that regularization is needed due to the functional nature of predictors, akin to ordinary multiple linear regression when the number of predictors exceeds the number of observations. One common tool (Hall and Horowitz, 2007) is to implement so-called spectral truncation regression by reducing each functional predictor X_{ij} to its first K_j functional principal component scores ξ_{ijk}, $k = 1, \ldots, K_j$ (see (2.2)). If the ϕ_{jk} are the eigenfunctions corresponding to the sample $X_{1j}, \ldots, X_{nj}$ of log-quantile transformed intra-regional correlation densities, one can also represent the functional parameters β_j in

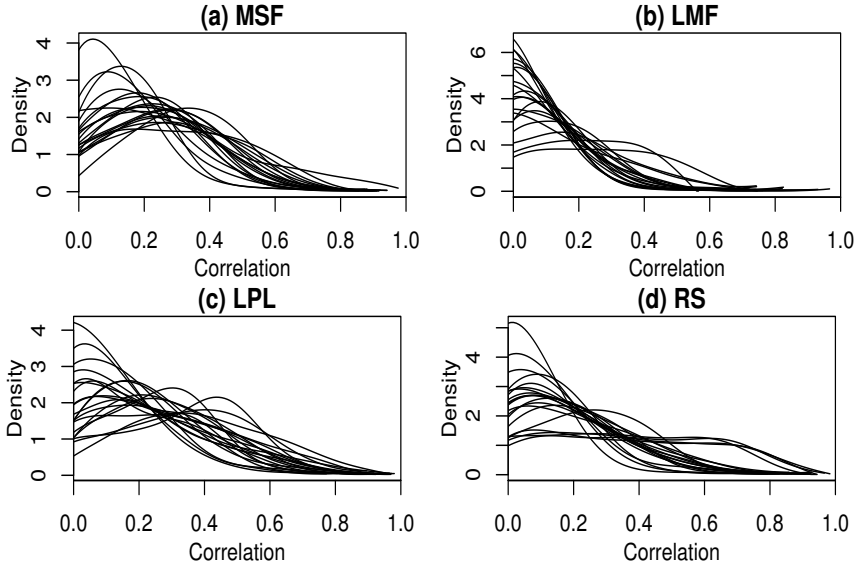

Figure 2: Local connectivity densities for 10 randomly chosen normal subjects for (a) medial superior frontal, (b) left middle frontal, (c) left parietal lobe, and (d) right supramarginal regions. These regions correspond to the the most relevant active predictors chosen in the regression analyses.

this basis, using the coefficients

$$B_{jk} = \int_0^1 \beta_j(t)\phi_{jk}(t)\mathrm{d}t, \quad k = 1, \ldots, K_j.$$

This results in a simplified linear multiple regression model,

$$Y_i = \beta_0 + \sum_{j=1}^{J} \sum_{k=1}^{K_j} B_{jk}\xi_{ijk} + \varepsilon_i. \tag{2.6}$$

As mentioned above, we will choose the truncation parameters K_j as the minimum number of components needed to explain at least 95% of the total variance, analogous to a cumulative scree plot approach in multivariate PCA.

To identify regions that are most useful in predicting each cognitive score, we must choose a subset of active predictors from (2.6) or, equivalently, identify the coefficients B_{jk} that are nonzero. The coefficients corresponding to the same functional predictor are naturally grouped, so it makes sense to set all or none of $B_{j1}, \ldots, B_{jK_j}$ to zero simultaneously. Hence, a group forward selection procedure was implemented using AIC as a selection criterion. At each step, the group of FPC scores that resulted in the lowest AIC was added to the model, and the selection procedure was halted once AIC increased in two successive iterations.

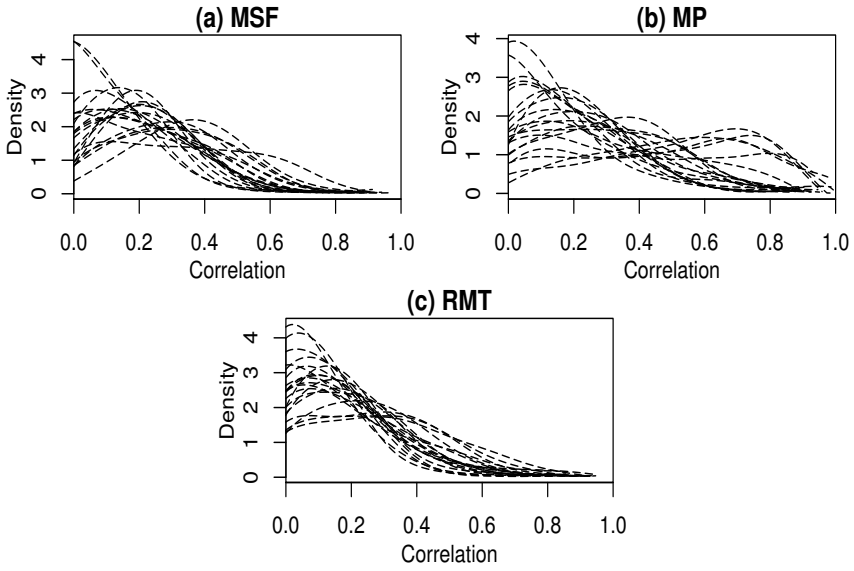

Figure 3: Local connectivity densities for 20 randomly chosen Alzheimer's subjects for (a) medial superior frontal, (b) medial prefrontal, and (c) right middle temporal regions. These regions correspond to the the most relevant active predictors chosen in the regression analyses.

3 Results

3.1 Patterns of Intra-regional Connectivity

We begin by examining the local functional connectivity distributions, as defined in Section 2.2. The intra-regional connectivity densities for normal subjects are visualized in Figure 2 for the medial superior frontal (MSF), left middle frontal (LMF), left parietal (LPL), and right supramarginal (RS) regions and for Alzheimer's subjects in Figure 3 for the medial superior frontal (MSF), medial prefrontal (MP), and right middle temporal (RMT) regions. These regions were the most relevant predictors in the optimal models discovered in the regression analyses below; densities for the other regions showed similar patterns. To facilitate visualization, these densities are plotted for a randomly chosen subset of 20 subjects in each group. The local connectivity distributions fall into one of three classes: (1) unimodal with peak near zero, (2) unimodal with peak at moderate to high correlation values and (3) flat over a wide range of correlations.

For both groups, application of the transformation methodology in (2.3) followed by FPCA resulted in the selection of three functional principal component scores $(\xi_{ij1}, \xi_{ij2}, \xi_{ij3})$ to represent the densities f_{ij}. The most common patterns of variability in these collections of density curves are represented by the transformation modes of variation in (2.4), which are illustrated in Figures 4 and 5. It emerges that the first mode of variation embodies the

shift from highly peaked connectivity densites near zero (red extreme, high FPC score) to relatively flatter densities with substantial fractions of correlations above 0.5 (blue extreme, low FPC scores). The second mode of variation indicates the level of concentration of the distribution, that is, the extent to which the correlations concentrate near 0.25, while the third mode of variation quantifies concentration around 0.15. However, this last mode only accounts for a small fraction of the overall variability.

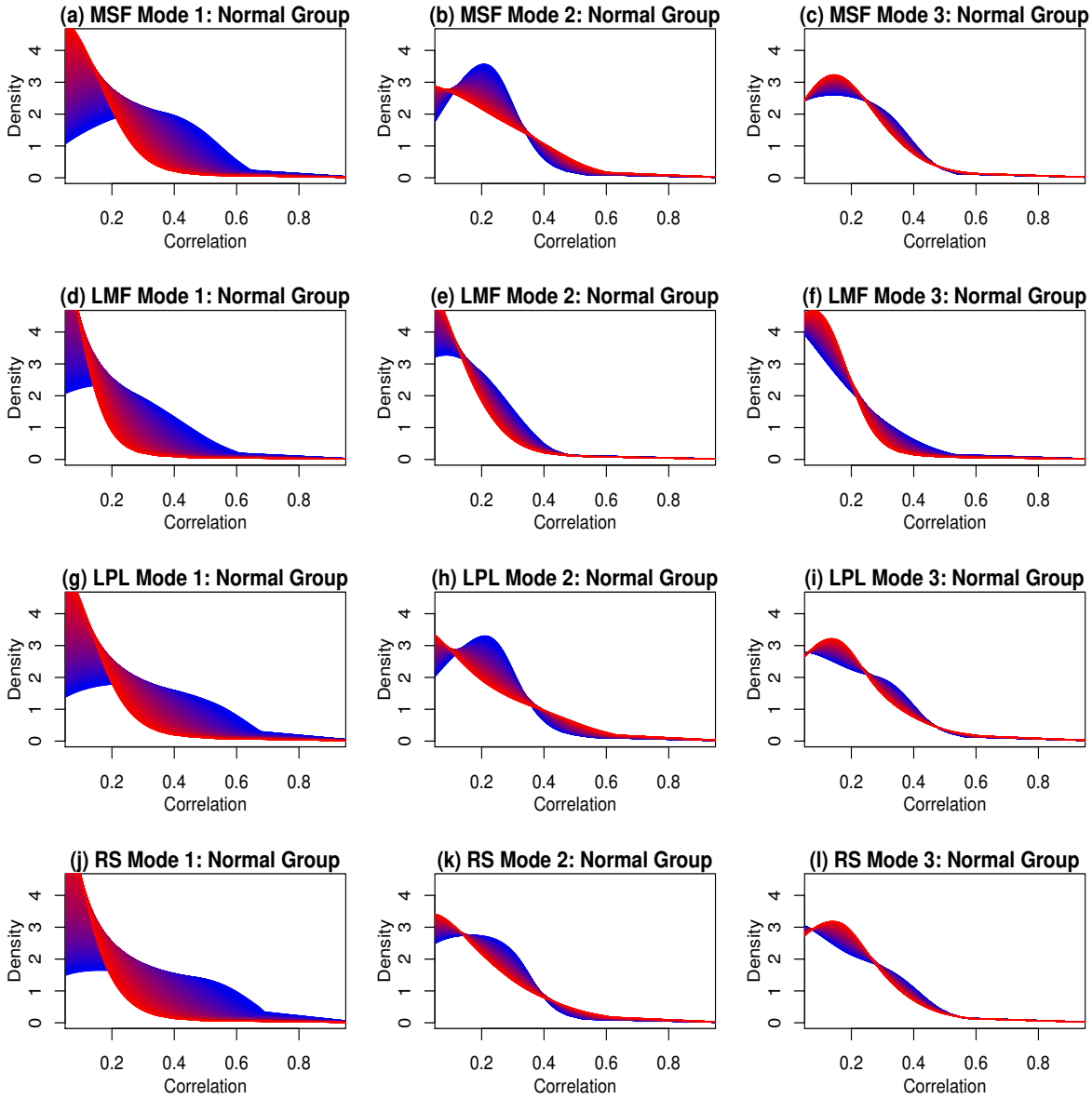

Figure 4: Mode of variation plots for the first three FPC scores of the MSF, LMF, LPL, and RS regions in the cognitively normal group. Blue (red) corresponds to low (high) values of the corresponding FPC scores, used as predictors in the regression models. For all regions, the first mode of variation indicates how much connectivity densities concentrate around 0, as well as how large it is around 0.6, while the second mode emphasizes the size of a mode around 0.25, and the third mode the size of a mode around 0.15.

3.2 Identifying Connectivity-Cognition Associations

The response variables in our analysis are the following standardized measures of cognitive function: episodic memory, executive function, spatial memory, and semantic memory. Figure 6 shows the distribution of these values for the two samples of cognitively normal and Alzheimer's subjects. As expected, scores are generally higher for normal subjects.

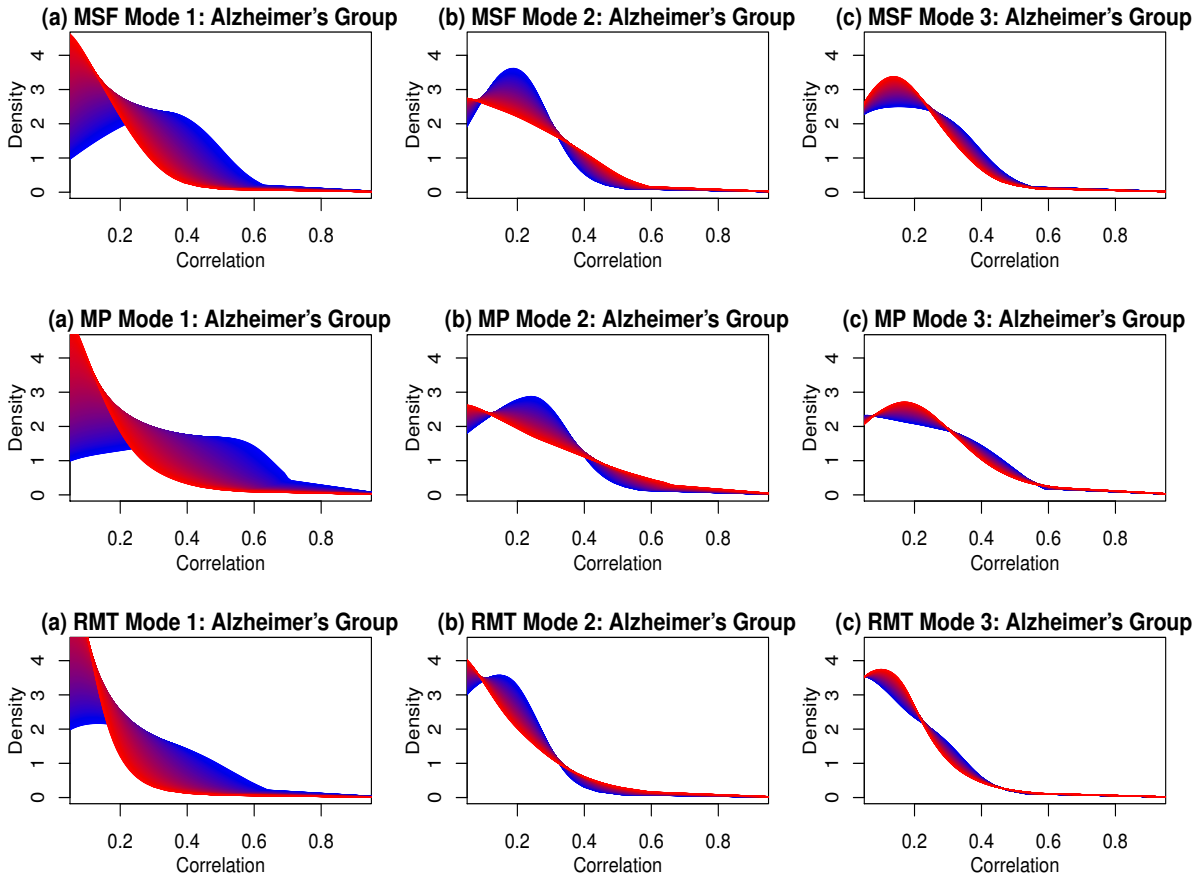

Figure 5: Mode of variation plots for the first three FPC scores of the MSF, MP, and RMT regions in the Alzheimer's group. Blue (red) corresponds to low (high) values of the corresponding FPC score used in the regression models, where the interpretation of the three modes of variation is similar to that for the three modes of variation for the regions of the normal group.

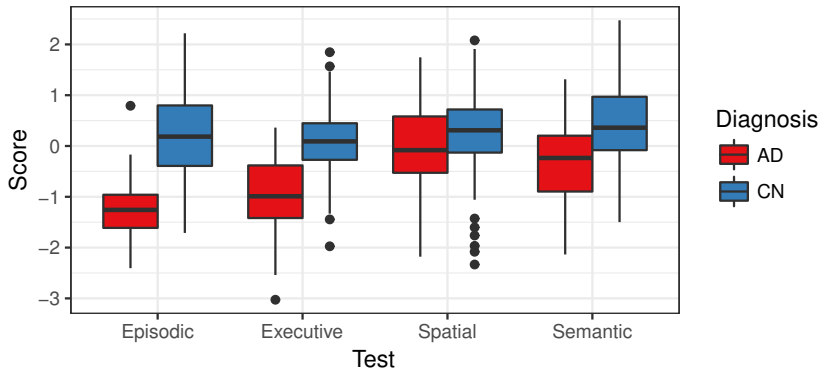

Figure 6: Boxplots for the distributions of cognitive scores within each subject group. 'AD' refers to the Alzheimer's disease group and 'CN' to the cognitively normal group.

Table 1: Post-Selection p-values for all 4 age-corrected cognitive scores using a 2-step stopping criterion based on AIC, for cognitively normal subjects. Rows indicate the order in which the regions are added to the model, with corresponding region-specific p-values in parentheses. Each p-value corresponds to the null hypothesis that the region can be removed from the model, assuming previously added regions are included.

	Episodic $N = 152$	Executive $N = 153$	Spatial $N = 139$	Semantic $N = 149$
1	MSF(0.328)	LPL(0.830)	MSF(0.352)	LPL(0.394)
2	RMF(0.968)	**LMF(0.071)**	LMF(0.848)	**LMF(0.089)**
3	**LPL(0.021)**	LMT(0.706)	RPL(0.813)	-
4	LMT(0.373)	**MSF(0.065)**	**RS(0.085)**	-
5	RMT(0.549)	-	-	-

Table 2: Post-Selection p-values for all 4 age-corrected cognitive scores using a 2-step stopping criterion based on AIC, for Alzheimer's subjects. Rows indicate the order in which the regions are added to the model, with corresponding region-specific p-values in parentheses. Each p-value corresponds to the null hypothesis that the region can be removed from the model, assuming previously added regions are included.

	Episodic $N = 60$	Executive $N = 61$	Spatial $N = 40$	Semantic $N = 59$
1	**MSF(0.073)**	RPL(0.937)	PCP(0.215)	RMF(0.248)
2	**MP(0.086)**	MSF(0.205)	MP(0.278)	**RMT(0.064)**
3	-	-	RMF(0.973)	**MSF(0.010)**
4	-	-	-	LMF(0.795)

For the regression fits obtained from model (2.6), we report the sequentially selected regions for the different subject groups and each of the four cognitive scores in Table 1 for cognitively normal and in Table 2 for Alzheimer's subjects. As some subjects had missing test scores, the total number of subjects used in each model is also indicated. Each group of included predictors was tested for significance using the post-selection inferential technique described in Loftus and Taylor (2015), and p-values less than 0.1 are shown in bold.

3.3 Interpretation

To further elucidate the association of the connectivity densities with cognitive scores, we indicate in Tables 3 and 4 the direction of the association between each cognitive score and the first and second FPC scores for the regions with small p-values in Tables 1 and 2. These are based on the sign of the coefficient estimate in the fitted model after forward selection.

Table 3: Directional associations between first FPC score and cognitive score - ↑: The higher the density peak near zero, the higher the test scores; ↓: the higher the density peak near zero, the lower the test scores.

	Group	
	Normal	Alzheimer's
Episodic	LPL(↓)	MSF(↓), MP(↑)
Executive	LMF(↑), MSF(↓)	-
Spatial	RS(↓)	-
Semantic	LMF(↑)	RMT(↓), MSF(↓)

Table 4: Directional associations between second FPC score and cognitive score - ↑: The higher the peak near 0.25, the higher the test scores; ↓: the higher the density peak near 0.25, the lower the test scores.

	Group	
	Normal	Alzheimer's
Episodic	LPL(↑)	MSF(↑), MP(↓)
Executive	LMF(↑), MSF(↑)	-
Spatial	RS(↑)	-
Semantic	LMF(↓)	RMT(↓), MSF(↓)

For normal subjects, we find that densities with high peaks near zero in the LPL region, i.e., more concentrated lower connectivity levels, which are the densities in red in Figures 4 (g)-(i), are associated with lower episodic performance, and the same pattern holds between MSF densities and executive scores, as well as RS densities and spatial scores. However, higher density peaks near zero in the LMF region are found to be positively associated with both executive and semantic scores. For the Alzheimer's group, having lower density peaks near zero in the MSF region is positively associated with both episodic and semantic memory, and the same relationship holds between RMT connectivity and semantic scores. The reverse is true for the MP region and episodic score.

Regarding the second FPC scores in Table 4 for normal subjects, we find that episodic, executive, and spatial scores are negatively associated with densities that have distinctive modes around 0.25. That is, a high concentration of moderate local functional connectivity values is related to performance decline. This is reversed for semantic score, where higher semantic performance is associated with concentrated correlations around 0.25 in the LMF region for normal subjects and in the RMT and MSF regions for the Alzheimer's group.

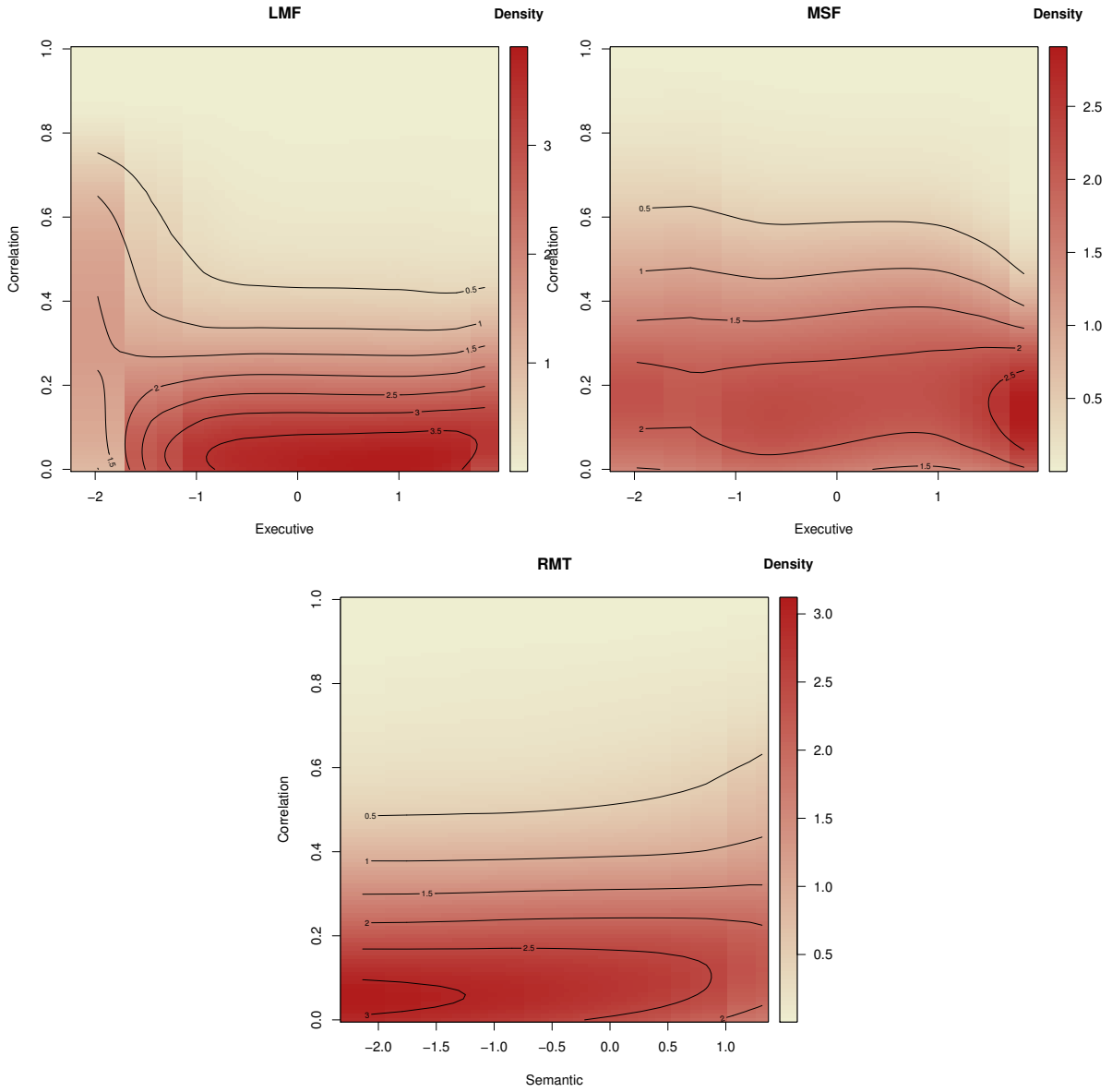

Figure 7: Smoothed contour plots of cognitive score/intraregional density relationships. (a) LMF region and executive score (normal group), (b) MSF region and executive score (normal group), and (c) RMT region and semantic score (Alzheimer's group).

Further visual interpretations of these findings are provided by contour plots based on the observed connectivity density/score pairs (f_{ij}, Y_i). Here, the x and y axes correspond to the cognitive score and correlation level, respectively, and value of the connectivity density is indicated by the color. These plots are shown in Figure 7 for three density/score combinations, two for the normal subjects and one for the Alzheimer's group. To more clearly discern the patterns in these plots, a smoothing step was applied along the cognitive score axis to reduce the noise in the data. The main relationship seen between LMF intra-regional connectivity densities and executive score is that subjects with high executive performance tend

to have higher density modes near zeros, while for MSF connectivity densities, densities with modes near 0.2 are associated with higher executive scores. In the Alzheimer's group, higher density peaks at low connectivity values for intra-regional connectivity densities in the RMT region are associated with lower semantic performance. These simple visual interpretations are confirmed by the numeric findings in Tables 3 and 4.

4 Discussion and Conclusions

We introduce connectivity densities to quantify intra-regional connectivity, using functional data analysis methodology adapted to density functions and demonstrate that this quantification of intra-regional connectivity can lead to specific findings that complement inter-regional connectivity studies and traditional inter-regional network analysis. As we demonstrate, quantification of connectivity density functions can be complemented by group forward selection and accompanying post-selection inference to discover relationships with other relevant covariates at the level of individual subjects.

In particular, our results demonstrate the value of utilizing entire distributions as predictors, where we use a nonlinear transformation method followed by functional principal component analysis to summarize each density in a vector of principal component scores. While these functional predictors contain a wealth of information, FPCA combined with the mode of variation plots simultaneously provides dimension reduction and interpretability. We conclude that intra-regional connectivity contains valuable information to assess brain function. Our methodology allows to efficiently quantify, visualize and interpret key aspects of intra-regional connectivity. Application of these novel methods revealed that intra-regional connectivity is associated with cognitive scores in specific ways. These findings are validated by related previous findings in the literature.

Our observation of a positive association between lower intra-regional connectivity with connectivity densities more concentrated around 0 and executive function for the left middle frontal cortex is supported by previous studies. Lower activation of LMF has been found to be correlated with executive function (Kirova et al., 2015; Possin et al., 2014) and with better attention performance (Murphy et al., 2014). A negative correlation between grey matter volume in LMF and executive function has also been previously discussed (Breukelaar et al., 2017). Our methodology thus highlights the specific role played by intra-regional connectivity in LMF.

It also has been previously observed that during episodic memory retrieval, several parietal regions are active, including inferior and superior parietal cortex (Wagner et al., 2005; Sajonz et al., 2010; Hutchinson et al., 2015) as well as the left inferior and superior parietal gyrus plays a role (Sajonz et al., 2010; Hutchinson et al., 2015), suggesting that activation of these regions is positively correlated with memory retrieval.

Frontal lobe deficit or dysfunction has been previously related to executive malfunction. Subjects with lesions on the frontal lobes, including in the medial superior frontal cortex, show lower executive function test performance (Roca et al., 2009). Especially subjects with damage in medial superior frontal regions are reported to perform particularly poorly in tests related to speed response in Stuss (2011), who hypothesize that the slower executive function is due to the failure of initiating or sustaining activity in MSF. This hypothesis suggests a positive correlation between executive function and activity in MSF.

Arnold et al. (2014) describe a connection between spatial function scores and activity in the right supramarginal, right precentral cortex and left hippocampus. Similarly, Bähner et al. (2015) report that the magnitude of the interaction between bilateral dorsolateral, right supramaginal cortex and hippocampus may predict spatial working memory performance.

Group analysis in Wierenga et al. (2011) has shown that thinning of the right middle temporal cortex in Alzheimer's disease patients is correlated with semantic test performance, where also activation of the right middle temporal cortex is reported when Alzheimer's disease subjects were performing certain semantic tasks. Lastly, Seidenberg et al. (2009) observe that compared to a normal control group, subjects at high risk for Alzheimer's disease show greater activity in the right middle temporal cortex.

While the emphasis of this paper is a novel quantification of intra-regional connectivity, the resulting analysis indicates that certain intra-regional correlations are associated to some degree with the cognitive status of a subject. The direction of these associations varies across regions and also differs according to disease status.

Acknowledgements

This work was supported by National Science Foundation grants DMS-1811888 and DMS-1712864.

Author Disclosure Statement

No competing financial interests exist.

References

Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006), "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *The Journal of Neuroscience*, 26(1), 63–72.

Arnold, A. E., Protzner, A. B., Bray, S., Levy, R. M., and Iaria, G. (2014), "Neural network configuration and efficiency underlies individual differences in spatial orientation ability," *Journal of cognitive neuroscience*, 26(2), 380–394.

Bähner, F., Demanuele, C., Schweiger, J., Gerchen, M. F., Zamoscik, V., Ueltzhöffer, K., Hahn, T., Meyer, P., Flor, H., Durstewitz, D. et al. (2015), "Hippocampal–dorsolateral prefrontal coupling as a species-conserved cognitive mechanism: a human translational imaging study," *Neuropsychopharmacology*, 40(7), 1674–1681.

Bassett, D. S., and Bullmore, E. (2006), "Small-world brain networks," *The Neuroscientist*, 12(6), 512–523.

Bassett, D. S., Nelson, B. G., Mueller, B. A., Camchong, J., and Lim, K. O. (2012), "Altered resting state complexity in schizophrenia," *Neuroimage*, 59(3), 2196–2207.

Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X.-N., and Sporns, O. (2014), "Changes in structural and functional connectivity among resting-state networks across the human lifespan," *NeuroImage*, 102, 345–357.

Breukelaar, I. A., Antees, C., Grieve, S. M., Foster, S. L., Gomes, L., Williams, L. M., and Korgaonkar, M. S. (2017), "Cognitive control network anatomy correlates with neurocognitive behavior: A longitudinal study," *Human brain mapping*, 38(2), 631–643.

Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., Andrews-Hanna, J. R., Sperling, R. A., and Johnson, K. A. (2009), "Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease," *The Journal of Neuroscience*, 29(6), 1860–1873.

Cai, T., and Hall, P. (2006), "Prediction in functional linear regression," *Annals of Statistics*, 34, 2159–2179.

Chen, K., and Lynch, B. (2017), "Weak Separablility for Two-way Functional Data: Concept and Test," *arXiv preprint arXiv:1703.10210*, .

Ferreira, L. K., and Busatto, G. F. (2013), "Resting-state functional connectivity in normal brain aging," *Neuroscience & Biobehavioral Reviews*, 37(3), 384–400.

Hall, P., and Horowitz, J. L. (2007), "Methodology and convergence rates for functional linear regression," *Annals of Statistics*, 35, 70–91.

Han, L., Pengfei, Z., Zhaohui, L., Fei, Y., Ting, L., Cheng, D., and Zhenchang, W. (2015), "Resting-state functional connectivity density mapping of etiology confirmed unilateral pulsatile tinnitus patients: Altered functional hubs in the early stage of disease," *Neuroscience*, 310, 27–37.

He, J., Carmichael, O., Fletcher, E., Singh, B., Iosif, A.-M., Martinez, O., Reed, B., Yonelinas, A., and DeCarli, C. (2012), "Influence of functional connectivity and structural MRI measures on episodic memory," *Neurobiology of Aging*, 33(11), 2612–2620.

Hinton, L., Carter, K., Reed, B. R., Beckett, L., Lara, E., DeCarli, C., and Mungas, D. (2010), "Recruitment of a community-based cohort for research on diversity and risk of dementia," *Alzheimer Disease and Associated Disorders*, 24(3), 234.

Hosseini, S., Hoeft, F., and Kesler, S. R. (2012), "GAT: a graph-theoretical analysis toolbox for analyzing between-group differences in large-scale structural and functional brain networks," *PloS one*, 7(7), e40709.

Hutchinson, J. B., Uncapher, M. R., and Wagner, A. D. (2015), "Increased functional connectivity between dorsal posterior parietal and ventral occipitotemporal cortex during uncertain memory decisions," *Neurobiology of learning and memory*, 117, 71–83.

Jones, M. C., and Rice, J. A. (1992), "Displaying the important features of large collections of similar curves," *The American Statistician*, 46, 140–145.

Kirova, A.-M., Bays, R. B., and Lagalwar, S. (2015), "Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimers disease," *BioMed research international*, 2015.

Kneip, A., and Utikal, K. J. (2001), "Inference for density families using functional principal component analysis," *Journal of the American Statistical Association*, 96(454), 519–542.

Loftus, J. R., and Taylor, J. E. (2015), "Selective inference in regression models with groups of variables," *arXiv preprint arXiv:1511.01478*, .

Lopez-Larson, M. P., Anderson, J. S., Ferguson, M. A., and Yurgelun-Todd, D. (2011), "Local brain connectivity and associations with gender and age," *Developmental cognitive neuroscience*, 1(2), 187–197.

Murphy, C. M., Christakou, A., Daly, E. M., Ecker, C., Giampietro, V., Brammer, M., Smith, A. B., Johnston, P., Robertson, D. M., Consortium, M. A. et al. (2014), "Abnormal functional activation and maturation of fronto-striato-temporal and cerebellar regions during sustained attention in autism spectrum disorder," *American Journal of Psychiatry*, 171(10), 1107–1116.

Parzen, E. (1962), "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, 33(3), 1065–1076.

Petersen, A., and Müller, H.-G. (2016), "Functional data analysis for density functions by transformation to a Hilbert space," *Annals of Statistics*, 44(1), 183–218.

Possin, K. L., LaMarre, A. K., Wood, K. A., Mungas, D. M., and Kramer, J. H. (2014), "Ecological validity and neuroanatomical correlates of the NIH EXAMINER executive composite score," *Journal of the International Neuropsychological Society*, 20(1), 20–28.

Qi, R., Zhang, L. J., Chen, H. J., Zhong, J., Luo, S., Ke, J., Xu, Q., Kong, X., Liu, C., and Lu, G. M. (2015), "Role of local and distant functional connectivity density in the development of minimal hepatic encephalopathy," *Scientific reports*, 5.

Ramsay, J. O., and Silverman, B. W. (2005), *Functional Data Analysis*, Springer Series in Statistics, second edn, New York: Springer.

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes, F., and Duncan, J. (2009), "Executive function and fluid intelligence after frontal lobe lesions," *Brain*, 133(1), 234–247.

Rosenblatt, M. (1956), "Remarks on some nonparametric estimates of a density function," *Annals of Mathematical Statistics*, 27, 832–837.

Sajonz, B., Kahnt, T., Margulies, D. S., Park, S. Q., Wittmann, A., Stoy, M., Ströhle, A., Heinz, A., Northoff, G., and Bermpohl, F. (2010), "Delineating self-referential processing from episodic memory retrieval: common and dissociable networks," *Neuroimage*, 50(4), 1606–1617.

Seidenberg, M., Guidotti, L., Nielson, K. A., Woodard, J. L., Durgerian, S., Antuono, P., Zhang, Q., and Rao, S. M. (2009), "Semantic memory activation in individuals at risk for developing Alzheimer disease," *Neurology*, 73(8), 612–620.

Shukla, D. K., Keehn, B., and Müller, R. A. (2010), "Regional homogeneity of fMRI time series in autism spectrum disorders," *Neuroscience letters*, 476(1), 46–51.

Silverman, B. W. (1986), *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability, London: Chapman & Hall.

Stuss, D. T. (2011), "Functions of the frontal lobes: relation to executive functions," *Journal of the international neuropsychological Society*, 17(5), 759–765.

Tomasi, D., and Volkow, N. D. (2010), "Functional connectivity density mapping," *Proceedings of the National Academy of Sciences*, 107(21), 9885–9890.

Tomasi, D., and Volkow, N. D. (2011), "Association between functional connectivity hubs and brain networks," *Cerebral Cortex*, 21(9), 2003–2013.

Tomasi, D., and Volkow, N. D. (2012), "Aging and functional brain networks," *Molecular psychiatry*, 17(5), 549–558.

van den Heuvel, M., Mandl, R., and Pol, H. H. (2008), "Normalized cut group clustering of resting-state fMRI data," *PloS one*, 3(4), e2001.

Wagner, A. D., Shannon, B. J., Kahn, I., and Buckner, R. L. (2005), "Parietal lobe contributions to episodic memory retrieval," *Trends in cognitive sciences*, 9(9), 445–453.

Wand, M. P., and Jones, M. C. (1995), *Kernel smoothing*, Vol. 60 of *Monographs on Statistics and Applied Probability*, London: Chapman and Hall Ltd.

Wang, J.-L., Chiou, J.-M., and Müller, H.-G. (2016), "Functional Data Analysis," *Annual Review of Statistics and its Application*, 3, 257–295.

Wierenga, C. E., Stricker, N. H., McCauley, A., Simmons, A., Jak, A. J., Chang, Y.-L., Nation, D. A., Bangen, K. J., Salmon, D. P., and Bondi, M. W. (2011), "Altered brain response for semantic knowledge in Alzheimer's disease," *Neuropsychologia*, 49(3), 392–404.

Worsley, K. J., Chen, J.-I., Lerch, J., and Evans, A. C. (2005), "Comparing functional connectivity via thresholding correlations and singular value decomposition," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457), 913–920.

Wu, T., Long, X., Zang, Y., Wang, L., Hallett, M., Li, K., and Chan, P. (2009), "Regional homogeneity changes in patients with Parkinson's disease," *Human brain mapping*, 30(5), 1502–1510.

Zalesky, A., Fornito, A., and Bullmore, E. (2012), "On the use of correlation as a measure of network connectivity," *Neuroimage*, 60(4), 2096–2106.

Zang, Y., Jiang, T., Lu, Y., He, Y., and Tian, L. (2004), "Regional homogeneity approach to fMRI data analysis," *Neuroimage*, 22(1), 394–400.