Statistical Methods for Connectome Genetics

Research Proposal for Advancement to Candidacy

Dustin Pluta

March 28th, 2018

Contents

1 Research Objectives

In recent years, the set of the dynamic and structural relationships between brain regions, referred to as the *connectome*, has been shown to play a crucial role in cognition and behavior. Advances in neuroimaging technologies have facilitated a more thorough and detailed study of the connectome at all scales, while developments in neurophysiology and cognitive psychology have motivated new theories and new questions regarding the structure of the connectome. Concurrent with the development of connectomics, the field of imaging genetics has emerged as an important area of research for deepening our understanding of the many factors that influence cognitive functions. Given the extreme complexity of both the brain and the genome, there is a pressing need for more refined and powerful statistical methods tailored to imaging genetics data.

We here propose the development of a collection of statistical methods and models for high-dimensional inference in settings relevant to connectome genetics. The estimates and inferences of these methods will yield scientifically interpretable results on the role of particular genetic influences on the human connectome, and may lead to a greater understanding of the underlying factors that determine complex cognitive traits such as learning styles, memory performance, and decision-making. Moreover, the proposed methods are designed to accommodate a wide variety of high-dimensional data types, and are suitable for many applications beyond imaging genetics.

Aim 1. To develop **flexible global association testing methods** suitable for inference with multi-modal data produced by current imaging genetics studies. These methods will provide a means to quickly explore relationships in these high-dimensional data sets as part of a preliminary analysis. While many such methods currently exist, a rigorous methodology for effectively applying and understanding the results of these methods in high-dimensional settings is still needed, and moreover, current methods are often too underpowered for practical use. Our preliminary work suggests that penalized kernels methods for association testing may yield much more powerful tests compared to existing approaches. The novel adaptive Mantel test developed here reduces the burden of tuning parameter selection, and provides a practical means of conducting association testing with kernel similarities.

Aim 2. To develop **multi-subject dynamic connectivity models** for fMRI and EEG data. Many approaches to modeling dynamic connectivity have been proposed in recent years, but the majority of these models are not able to jointly model multiple subjects, and many require unreasonable assumptions regarding the scale and structure of relevant features. The models developed in this aim will refine previously developed dynamic connectivity models in order to address some of the numerous statistical challenges in estimating common connectivity states across multiple subjects. Motivated by the scientific goal of understanding the genetic influences on dynamic connectivity features, these models will be designed to allow for covariate adjustment, and the estimation and interpretation of covariate effects. In particular, the developed models will be applied to assess the influence of genetics on various features of dynamic connectivity across a variety of experimental settings and imaging modalities.

Aim 3. To develop **manifold regression models for heritability analysis of connectivity**. Building on the methods and challenges in Aim 1, a mixed-effects model for heritability analysis of manifold-valued phenotypes will be developed to increase the power and efficiency in testing and estimating the effects of genetics on high-dimensional phenotypes, such as brain connectivity measurements. By utilizing the inherent structure of connectivity matrices measured from neuroimaging data, the developed models will reduce dimensionality in a more theoretically principled manner compared to standard regression methods. A practical mixed-effects model for heritability analysis of manifold-valued data would have important applications for many types of neuroimaging and connectomic data.

Aim 4. **For broad impact on science:** The above developed models will be applied to thoroughly analyze imaging genetics data from a study of 1000 healthy Chinese college students, collected by collaborators at Beijing Normal University. The data includes measures from a variety of psychological tasks related to learning and memory, and EEG, fMRI, and DTI imaging, as well as genetic information in the form of approximately 5×10^5 SNPs. The goals of the study include the identification of genetic factors that influence both neurological activity and behavioral outcomes. We will also apply our methods to analyze data from the Human Connectome Project (HCP), which includes a variety of fMRI scans of both related and unrelated

subjects (including SNP data for unrelated subjects). The HCP data is open access, which will allow us to directly compare our results to other analyses of the same data, and from which we will provide a publicly available and entirely reproducible analysis using our developed methods. Code for the applied methods will also be made available in a set of `R` packages.

2 Background

2.1 Measuring Brain Connectivity

Commonly used technologies in current neuroimaging studies include structural and functional magnetic resonance imaging (MRI and fMRI), electroencephalogram (EEG), magnetoencephalogram (MEG), and diffusion tensor imaging (DTI). Since each of these imaging modalities measures a different set of neural features or phenomena, they may differ in their effectiveness in answering particular scientific questions. For instance, fMRI measures changes in the blood oxygen level over time in over 5×10^5 voxels throughout the brain (Lindquist (2008)), with very high spatial resolution (typically less than 3mm), but poor temporal resolution (around 2s). In contrast, EEG measures the collective cortical activity of populations of neurons via 64-256 electrodes placed on the scalp (Ombao et al. (2016)), allowing for a high temporal resolution (1 kHz) at the cost of spatial resolution. Instead of measuring dyanimc brain activity, strucutral imaging methods provide information about the physical relationship of brain regions, such as DTI, which measures the structure and orientation of the brain's white matter fiber tracks via the diffusion of water molecules through the entire brain volume (Basser, Mattiello, and LeBihan (1994a)).

It is now generally believed that higher-level cognitive processing (e.g., in memory retrieval, decision making) critcally depends on the interaction and transfer of information between many localized regions. Numerous different methods for calculating brain connectivity have been proposed, but the reliability, interpretation, and relationship of these different measures is not well established (Fiecas et al. (2013)). To date, there are three concepts of brain "connectivity" that have been of primary interest, namely *structural*, *functional*, and *effective* connectivities. Structural connectivity refers to the anatomical connections between brain regions, measured using DTI. Functional connectivity is a symmetric and undirected measure of concordant activity between brain regions, commonly calculated as Pearson's correlation or coherence between the activation signals of two regions (Fiecas and Ombao (2011)). By contrast, effective connectivity is a directed measure of how past activity in one region influences the future activity of another region. Effective connectivity is closely related to Granger causality and is often estimated with a vector autoregressive (VAR) model (Gorrostieta et al. (2012), Gorrostieta et al. (2013), S. Chiang et al. (2017)). Modalities used for functional and effective connectivity studies include fMRI, EEG, and MEG. Despite the continued use of functional and effective connectivity measures in neuroimaging studies over the past two decades, there remain open questions regarding the rigor and underlying statistical assumptions of these methods, and the sensitivity of these measures to measurement noise and choice of pre-processing pipelines.

In studying the etiology of neurological diseases such as Alzheimer's, schizophrenia, and autism spectrum disorders, significant links with these forms of brain connectivity have been consistently shown (Woodward and Cascio (2015)). There is also evidence that dynamic characteristics of brain connectivity, and not just global "static" measures of connectivity over an entire experiment, are also important for understanding brain function. Formally defining and modeling dynamic connectivity has motivated the development of new statistical methods. Current approaches for modeling dynamic connectivity include sliding window methods (e.g., Chang and Glover (2010)), graphical Bayesian modeling (Warnick et al. (2017), allen2014tracking, zalesky2014time, lindquist2014evaluating, havlicek2010dynamic), change-point detection via VAR models (Kirch, Muhsal, and Ombao (2015a)), hidden Markov switching-VAR models (Samdin et al. (2017)), and the recently proposed "instantaneous" phase coherence (Cabral et al. 2018). As an example of dynamic connectivity estimation, a brief description of a three-step procedure for fitting he regime switching factor model (sf-VAR) (Ting et al. (2017)) is as follows. (1) Compute initial estimates of connectivity subspaces shared across regimes using a stationary factor model; (2) apply a factor SVAR model, and identify regime boundaries with a switching Kalman filter to partition the neural signal into a small number of distinct

states; (3) finally, use the low-dimensional factor representations for each regime to estimate within-regime effective connectivity. This model is able to detect abrupt changes in mental state, such as might occur in an experiment with varied cognitive demands, and is also able to estimate recurring states over the course of the experiment.

2.2 Models and Tests for Imaging Genetics

Human traits, including neural, cognitive, and behavioral traits, are shaped by genetics and environmental factors through a number of complex biological processes. In the past two decades, single nucleotide polymosphisms (SNPs) have become one of the most commonly measured forms of genetic data, as they are abundant (about 10 million in the human genome), stable, and easy to measure ((“What Are Single Nucleotide Polymorphisms (Snps)? - Genetics Home Reference,” n.d.)). Each SNP measurement contains the allele at a particular base pair location in a subject's DNA, typically coded as a 0 for the common homozygotic allele, 1 for the heterozygotic allele, and 2 for the uncommon homozygotic allele. Many complex traits are believed to depend on higher-order interactions between collections of SNPs, but determining these relationships is difficult, and in practice it is common to restrict attention to only the additive effects of SNPs. The proportion of variance of an observed trait that is accounted for by the additive effect of genome-wide SNPs is referred to as narrow-sense heritability, denoted h^2.

For many cognitive phenotypes, there is currently little prior information on heritability estimates, important genetic factors, and functional relationships. Given this lack of knowledge regarding the structure and role of genetic influences, global association testing is a convenient method to evaluate the strength and significance of genetic influences on a particular trait. In genome-wide association studies (GWAS), the random effects (or variance components) model is a standard tool to estimate the heritability of complex traits (Goeman, Van De Geer, and Van Houwelingen (2006); Liu, Lin, and Ghosh (2007); Boyle, Li, and Pritchard (2017)), popularized in recent years by the freely available Genome-wide Complex Trait Analysis (GCTA) software (Yang et al. (2011)). In the variance components model, the effect of an individual genetic variant is treated as a random quantity, rather than a fixed parameter, so that the genetic effects of a large number of SNPs can be parameterized by a small number of variance components. As a result of this, the number of genetic variants studied is not required to be less than the sample size. This method is closely related to semi-parametric kernel regression methods, which have received increasing attention in this area as a means to flexibly model non-additive effects of SNPs (Schaid (2010a)). The variance components model and related kernel methods have also been extended to model the heritability of multi-dimensional phenotypes, such as brain activation or shape (Ge et al. (2016)).

The variance components model commonly used for genetic analysis is a linear random effects model, from which one can estimate the narrow-sense (or additive) heritability of a phenotype, denoted h^2. Formally, let Y be an $n \times 1$ vector of observed continuous phenotype measurements from n subjects, and X be the $n \times p$ SNP data matrix for p SNPs. The heritability model is

$$Y = Xb + \varepsilon,$$

where b is a $p \times 1$ vector of random effects with $b \sim \mathcal{N}_p(0, \sigma_b^2 \otimes I_p)$, and $\varepsilon \sim \mathcal{N}_n(0, \sigma_\varepsilon^2 \otimes I_n)$. Since the target of inference is the total genetic contribution, this model is often rewritten by aggregating the individual SNP effects $Xb = g$, yielding

$$Y = g + \varepsilon,$$

where $g \sim \mathcal{N}(0, \sigma_g^2 \otimes G)$, with $G = XX^T/p$, referred to as the *genetic relationship matrix*. Estimates of the variance components of the model $\hat{\sigma}_g^2, \hat{\sigma}_\varepsilon^2$ can be computed using standard methods for mixed effects models, with the narrow-sense heritability calculated as

$$\hat{h}^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_\varepsilon^2}.$$

Distance-based testing methods, which include the Mantel test (Mantel (1967a)), RV coefficient (Robert and Escoufier (1976)), and the distance-covariance test (Székely et al. (2007)), are a common class of methods to test for global association between two sets of features using the pairwise distance or similarity between observations in each feature set. These methods can easily accommodate data of arbitrary dimension, and are generally straightforward to implement and interpret. In our recent work, we have found that Mantel's test (using the L_2 inner product to measure similarity between subject) is equivalent to the global score test for h^2 from the random effects model (Pluta et al. (2017)). By adopting other similarity measures, such as weighted inner products or non-linear kernels, the Mantel test can also be used to conduct global association tests corresponding to other methods, e.g. multivariate distance matrix regression, pseudo F-tests, and kernel-based tests. The flexibility of this approach makes it possible to test a wide variety of relationships and data types in a common framework. For details on generalizations and applications of distance-based methods, see Schaid (2010a); Pan (2011); Zapala and Schork (2012); Ge et al. (2016); Pluta et al. (2017).

Although distance-based methods have many attractive features, there is evidence that they can be severely underpowered in high-dimensional settings, making them impractical for many applications of interest. Alternative methods, such as the "sum of powered correlation" (SPC) and "sum of powered score" (SPU) tests, have been proposed as more powerful for imaging genetics data. Versions of these tests have been shown to perform better than distance-based methods in some simulations (Xu, Xu, and Pan (2017)). In Pluta et al. (2017), we have proposed the adaptive Mantel test as an extension of the classical Mantel test (or the RV coefficient test) that utilizes penalized regularization as in ridge regression in order to achieve higher power in high-dimensional settings compared to the other association testing methods mentioned above. By following the "adaptive" algorithm from Xu, Xu, and Pan (2017), the adaptive Mantel test is able to simultaneously test a set of candidate penalization terms while still maintaining the correct type I error, and without suffering a severe loss of power relative to a univariate Mantel test, thus making it a practical alternative for association testing in imaging genetics.

Inferential modeling for imaging genetics is challenging, in part due to the extremely high-dimension of the feature spaces. The most commonly used data-driven methods in these settings are principal component analysis (PCA), and independent component analysis (ICA) (Liu et al. (2009)). Although these methods have proven effective for many applications, they have the drawback that extracted leading components do not necessarily capture the most significant associations. In contrast, simultaneous dimension reduction methods, such as partial least squares (PLS), canonical correlation analysis (CCA), reduced rank regression (RRR), and parallel ICA (Ahn et al. (2015)), perform dimension reduction jointly on responses and covariates, and may be preferred for analyzing imaging genetics data. Regularization methods developed from ridge regression and the LASSO have been widely used in these contexts as well (Tibshirani (1996)).

Manifold regression methods seek to use the inherent structure of the data to improve modeling efficiency. For instance, covariance matrices are (often) positive definite, and so belong to the Riemannian manifold of positive definite matrices. This is directly relevant to modeling many types of brain connectivity, although the development of these models is still in its early stages, and many theoretical and computational challenges remain. In previous work, manifold-value regression models have been proposed as analogs to the fixed effects (Zhu et al. (2009)) and mixed effects models (Kim et al. (2017)). Applications to DTI data show some promise with these methods, although they can be technically and computationally challenging to implement.

2.3 Review of Previous Results in Imaging Genetics

Some hallmark successes in genetics have come from the study of prominent neurological and psychological diseases, such as Parkinson's disease, Alzheimer's disease, and schizophrenia. For each of these diseases, many significant genetic risk factors of relatively large effect have been found, which have in turn led to improved understanding of the mechanisms of action for these diseases, and have even resulted in promising gene therapies for Parkinson's disease (Palfi et al. (2014)). Initially motivated by these promising results, the

field of imaging ge- netics has expanded to now consider a number of different neural disorders and many facets of cognition in healthy-individuals. We mention a few recent results to illustrate the breadth of current research. The study by Bohlken et al. (2016) found that global reductions in white matter, which underpin structural connectivity, are largely explained by genetic risk factors for schizophrenia. In very recent work by Sudre et al. (2017), a study of families with histories of attention-deficit/hyperactivity disorder (ADHD) found significant genetic heritability for the default mode, cognitive control, and ventral attention networks. An analysis of 161 twin pairs found significant heritability of anatomical connectivity measured with DTI, and identified significant differences in heritability across subnetworks (with av- erage heritability over all white matter tracts estimated at around 30% (Shen et al. (2014)). With data from 1320 unrelated, young, healthy adults, Ge et al. (2016) demonstrated that several characteristics of brain structure are genetically heritable, such as brain volume and neuroanatomical shape. Using resting-state fMRI data from twins and non-twin siblings, Vidaurre et al. (2017) found evidence of significant heritability for certain characteristics of dynamic connectivity patterns; this is one of the few studies to date to have considered the heritability of dynamic connectivity.

3 Proposed Research

3.1 Aim 1: Development of the Adaptive Mantel Test

To develop **flexible global association testing methods** suitable for inference with high-dimenesional multi-modal data produced by current neuroimaging studies. These methods will provide a means to quickly explore possible significant relationships in these high-dimensional data sets as part of a preliminary analysis. While many such methods currently exist, a rigorous methodology for effectively applying these methods in high-dimensional settings is still needed, and moreover, current methods are often too underpowered for practical use. Our preliminary work suggests that adapting penalized kernels methods for association testing may yield much more powerful tests compared to existing approaches. The developed methods will target applications to imaging genetics studies, with a focus on meaningful interpretation of results in terms of standard concepts in the genetics literature, such as phenotype heritability.

Overview. Current methods for association testing tend to have low power in high-dimensional settings and when there are numerous noisy covariates present. This is problematic for applications in imaging genetics, since the signal in heritability testing is often weak and distributed across numerous SNPs. To improve the power of the Mantel test, we propose the use of penalized kernels. By examining the Mantel test with the ridge kernel $X(X^TX + \lambda I)^{-1}X^T$, we derive connections between the test and the score tests for the linear fixed effects, random effects, and ridge regression models, which provide context to understand the most powered alternatives for the test. In simulation studies, the kernel Mantel test consistently outperforms competing methods such as the RV coefficient test, distance covariance, and the sum of powered score test. We further propose the adaptive Mantel test as an efficient method for simultaneous testing across a set of tuning parameters in the kernel Mantel test. As part of our preliminary work, we have applied the adaptive Mantel test to test the association of a small set of SNPs with EEG coherence during a working memory task.

Proposed Method for Multi-modal Association Testing.

If two statistics T_1 and T_2 produce equivalent test results when calculated on the same size n sample, we refer to these statistics as *testing equivalent*, and write $T_1 \asymp T_2$. When the tests from T_1 and T_2 are asymptotically equivalent as $n \to \infty$, we say T_1 is *asymptotically testing equivalent* to T_2, and write $T_1 \doteq T_2$. Consider n independent observations each measured on $\mathbf{X} \times \mathbf{Y}$, where X is the $n \times p$ design matrix from the p-dimensional feature space $\mathbf{X}$ and Y the $n \times 1$ response vector from a univariate feature space $\mathbf{Y}$, and further suppose X is column-centered, and Y is centered. Let K be an $n \times n$ similarity matrix that measures the similarity of observations in $\mathbf{X}$. For $\mathcal{K}^{\mathbf{X}}(\cdot, \cdot)$ a similarity or dissimilarity metric on $\mathbf{X} \times \mathbf{X}$, let K be the corresponding Gram matrix with $K_{ij} = \mathcal{K}^{\mathbf{X}}(X_i, X_j)$. Similarily, for some metric $\mathcal{K}^{\mathbf{Y}}$ defined on $\mathbf{Y} \times \mathbf{Y}$, let H be the corresponding Gram matrix of Y. The Mantel test statistic is defined as

$$T(X,Y) = \langle H, K \rangle := \sum_i \sum_j H_{ij} K_{ij} = tr(HK).$$

The reference distribution under the null hypothesis of no association between the $\mathcal{K}^{\mathbf{X}}$ measures and $\mathcal{K}^{\mathbf{Y}}$ measures, can be obtained from the observed features X and Y by permuting the observation labels for one set of features and calculating the empirical reference distribution for T. Equivalently, one can hold one matrix fixed, say K, and simultaneously permute the rows and columns of H.

We now introduce three Gram matrices calculated from weighted inner products, which we denote K_E, K_M, and K_λ, which use the Euclidean inner product, Mahalanobis inner product, and ridge kernel respectively. For a positive semi-definite weight matrix $\mathcal{W}$, the corresponding weighted inner product is calculated as $\langle X_i, X_j \rangle_{\mathcal{W}} = X_i^T \mathcal{W} X_j$. Choosing $\mathcal{W}_E = \mathbf{I}$ gives the standard Euclidean inner product, with Gram matrix

$$K_E := XX^T$$

where $K_{E,ij} = \langle X_i, X_j \rangle$, $K_{E,ii} = \|X_i\|^2$. Another natural choice for weight matrix is $\mathcal{W}_M = (X^T X)^{-1}$ (assuming the inverse exists), which is the projection matrix into $\mathcal{C}(X)$, the column space of X. The Gram matrix is

$$K_M := X(X^T X)^{-1} X^T.$$

K_M is also recognizable as the "hat matrix" from the fixed effects model, and is related to the Mahalanobis distance. When $X^T X$ is not full rank, such as when $n > p$, we can replace $(X^T X)^{-1}$ with a generalized inverse $(X^T X)^-$, since the similarity matrix is invariant to the choice of inverse. Alternatively, we can pre-condition the weight matrix by adding a positive constant λ to the diagonal, i.e. $\mathcal{W}_\lambda = (X^T X + \lambda I_p)^{-1}$, which has Gram matrix

$$K_\lambda := X(X^T X + \lambda I_p)^{-1} X^T.$$

Linear Model Score Tests In general, the score test is defined as follows. For a stochastic model with parameters $\theta = (\beta, \alpha)$ and observations (X, Y), and with likelihood $\mathcal{L}(\theta)$, score vector $\mathcal{U}(\theta) = \frac{\partial \log \mathcal{L}}{\partial \theta}$, and Fisher information $\mathcal{I}(\theta) = -\mathbb{E}\frac{\partial}{\partial \theta}\mathcal{U}(\theta)$, the score test statistic for testing $H_0 : \beta = 0$ is

$$S = [\mathcal{U}(\theta)^T (\mathcal{I}(\theta))^{-1} \mathcal{U}(\theta)]|_{\beta=0}.$$

The classical fixed effects model is robust and broadly applicable for simple association testing, but requires $n > p$, and so is not feasible for high-dimensional settings. It does however serve as a useful theoretical comparison to the random effects and ridge regression models to be discussed next. The fixed effects model can be written

$$Y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n).$$

The global score test statistic can be written as

$$S_F = \frac{1}{\sigma^2} \text{tr}([X(X^T X)^{-1} X^T][YY^T]).$$

In practice, the nuisance parameter σ^2 can be replaced with $\hat{\sigma}^2 = Y^T Y/(n-1)$, which is the REML estimator when there are no adjustment covariates. This is a scalar that is fixed under permutations of Y. Therefore,

$$S_F \asymp Y^T X(X^T X)^{-1} X^T Y = \text{tr}(HK_M),$$

where $H = YY^T$.

For the random effects model

$$Y = Xb + \epsilon,$$

where $b \sim N(\mathbf{0}, \sigma_b^2 I_p)$, $\epsilon \sim N(\mathbf{0}, \sigma^2 I_n)$. Under the null hypothesis of no association between X and Y, we have $\sigma_b^2 = 0$. Similar to Liu, Lin, and Ghosh (2007), we first calculate the score and use the term that involves both X and Y as our score test statistic. Thus a testing equivalent form for the random effect score test statistic is

$$S_R \asymp \text{tr}(HK_E)/\sigma^2 \asymp \text{tr}(HK_E).$$

The ridge regression model is a penalized form of the fixed effects model, for which the estimator of β is defined as the minimizer of the penalized residual sum of squares:

$$\hat{\beta}_\lambda = \arg\min_b \left\{ ||Y - Xb||^2 + \lambda ||b||^2 \right\}.$$

From the penalized likelihood, we can derive the score statistic

$$S_\lambda = \frac{1}{\sigma^2} Y_\lambda^T X_\lambda (X_\lambda^T X_\lambda)^{-1} X_\lambda^T Y_\lambda = \frac{1}{\sigma^2} \text{tr}(HK_\lambda) \asymp \text{tr}(HK_\lambda).$$

Connecting the fixed-effects and random-effects models through ridge regression. The sample correlation of the similarity measures of all subjects is defined as

$$r(H, K) = \frac{tr(HK)}{\sqrt{tr(HH)tr(KK)}} \tag{1}$$

Since $r(H, K)$ is a correlation, we have $-1 \leq r(H, K) \leq 1$ in general, and for H and K p.s.d., this is further restricted to $0 \leq r(H, K) \leq 1$. When permutations are used to assess the strength of association, using T is equivalent to testing the significance of $r(H, K)$, since the denominator of $r(H, K)$ is fixed when simultaneously permuting the rows and columns of either H or K. The sample correlation of similarities for the three choices of K above can be conveniently related through the singular value decomposition (SVD) of X.

Proposition. Let X be an $n \times p$ column centered matrix of covariates with $\text{rank}(X) = r$ and singular value decomposition $X = U_{n\times r} D_{r\times r} V_{p\times r}^T$, with squared singular values $\eta_i, i = 1, \cdots, r$. Let Y be an $n \times 1$ centered vector of scalar responses, and let $H = YY^T$ and $Z = U^T Y$.

1. Fixed Effects
$$r(H, K_M) = \frac{\sum_{i=1}^r z_i^2}{\sqrt{p} \sum_{i=1}^n y_i^2}. \tag{2}$$

2. Random Effects
$$r(H, K_E) = \frac{\sum_{i=1}^r \eta_i z_i^2}{\sqrt{\sum_{i=1}^r \eta_i^2} \sum_{i=1}^n y_i^2} \tag{3}$$

3. Ridge Regression
$$r(H, K_\lambda) = \frac{\sum_{i=1}^r \frac{\eta_i}{\lambda+\eta_i} z_i^2}{\sqrt{\sum_{i=1}^r \left(\frac{\eta_i}{\eta_i+\lambda}\right)^2} \sum_{i=1}^n y_i^2}. \tag{4}$$

4. Relation to Correlation of X and Y
$$r(H, K_M) = \frac{1}{\sqrt{p}} R^2(X, Y)$$

5. Asymptotic Equivalences

$$\lim_{\lambda \to 0} r(H, K_\lambda) = r(H, K_M)$$
$$\lim_{\lambda \to \infty} r(H, K_\lambda) = r(H, K_E).$$

Corollary. Connections between the Mantel test statistics

Let $V_i \sim \chi^2_1, i = 1, \cdots, r$ and $V \sim \chi^2_p$.

1. Fixed-effects model.

$$S_F \asymp \text{tr}(HK_M) = Z^T D(D^T D)^{-1} D^T Z \ \sim \ cV$$

2. Random-effects model.

$$S_R \asymp \text{tr}(HK_E) = Z^T D D^T Z \ \sim \ \sum_{i=1}^{r} \eta_i V_i,$$

3. Ridge regression.

$$S_\lambda \asymp \text{tr}(HK_\lambda) = \sum_{i=1}^{r} \frac{\eta_i}{\lambda + \eta_i} z_i^2 \propto \lambda \sum_{i=1}^{r} \frac{\eta_i}{\lambda + \eta_i} z_i^2$$
$$\overset{\lambda \to \infty}{\to} \sum_{i=1}^{r} \eta_i z_i^2 \sim \ \sum_{i=1}^{r} \eta_i \chi^2_1$$

These results provide a description of the ridge regression score test as a natural intermediary between the fixed and random effects score tests. A crucial difference of the ridge test is the presence of the tuning parameter λ. It is obvious that when $\lambda = 0$, S_λ reduces to S_F, and $r(H, K_\lambda)$ reduces to $r(H, K_0)$. On the other hand, when $\lambda \to \infty$, the test based on a ridge-penalty converges to that based on the random-effect model. Note that in permutation based methods, multiplying a constant does not change the permuted p-value. Therefore, $S_\lambda \propto \lambda \text{tr}(HK_\lambda)$, which converges to $\sum_{i=1}^{r} \eta_i z_i^2$, i.e., S_R.

The correlation formulas ((2)) – ((4)) give some insight into the relationship of the three models in terms of $Z = U^T Y$, the image of Y after transformation by the left eigenvectors of X. From (**??**), we see that the fixed effects score test is equivalent to testing the Euclidean norm of Z. Whereas from (**??**), the random effects score test statistic is equivalent to testing the *weighted* norm of Z, where the jth component (corresponding to the jth eigenvector) is weighted by η_j. This has the effect of emphasizing the influence of directions in $\mathbf{X}$ for which X has large variance and reducing the influence of directions with small variance. The ridge regression score test is a compromise between the fixed and random effects, with small λ yielding a test close to the fixed effects (or identical at $\lambda = 0$), and large λ yielding a test close to the random effects score test, and identical tests for $\lambda \to \infty$. Geometrically, the ridge test weights the z_j proportional to the η_j as in the random effects test, but flattens each weight by a factor of $\frac{1}{\lambda + \eta_j}$.

Mantel test for multivariate outcomes. Given the connection between fixed-effect and random-effect models, the preceding results can be extended to the relationship between two sets of multivariate measurements. This extension is essential for many applications, such as heritability analysis of multi-dimensional traits. Suppose now that each subject i is measured on a q-variate phenotype Y_i so that Y is an $n \times q$ matrix. Let X be an $n \times p$ covariate matrix as before.The fixed-effect model assumes $Y \sim N(XB, \Sigma_e)$, where Σ_e is a $q \times q$ positive definite matrix. The score statistic can also be expresses as a Mantel statistic where the similarity matrices for both X and Y are their corresponding projection matrices:

$$S_F \asymp \text{tr}(H_M K_M),$$

where $K_M = X(X^T X)^{-1} X^T$ and $H_M = Y(Y^T Y)^{-1} Y^T$.

Note that both similarities used in S_F are based on the projection matrices. In the literature of multivariate linear models, several correlations have been proposed to quantify the overall correlation between X and Y. For example, Hooper's trace correlation (Hooper (1959)) is defined as

$$r_T^2 = \frac{1}{q}\text{tr}\left((Y^TY)^{-1}Y^TX(X^TX)^{-1}X^TY)\right)$$

The relationship between Hooper's trace correlation and $r(H_M, K_M)$ follows immediately:

$$r_T^2 = \sqrt{\frac{p}{q}}r(H_M, K_M).$$

}

Random effects model for multivariate response. In the random-effect model for univariate outcomes, the p coefficients are assumed to be i.i.d. random variables from a univariate normal distribution with a mean of zero. Extending this assumption to multivariate outcomes, we assume the p vectors of coefficients, each is of length q are i.i.d. random vectors from a multivariate normal distribution. In other words, the coefficients matrix follows a matrix normal distribution and the outcomes can be modeled as

$$Y = XB + \epsilon,$$

where $B \sim N_{p\times q}(\mathbf{0}, \Sigma_b, I_p)$, $\epsilon \sim N_{n\times q}(\mathbf{0}, \Sigma_e, I_n)$ where Σ_b and Σ_e are positive semidefinite matrices. In genetic studies, it is often interesting to evaluate the "overall" heritability of the multivariate outcomes. One heritability measure is based on the traces of the total variance in Y, which is summarized by $tr(\Sigma_b \otimes XX^T + \Sigma_e \otimes I_n)$, and the amount explained by X, which is $tr(\Sigma_b \otimes XX^T)$:

$$h^2 = \frac{tr(\Sigma_b \otimes XX^T)}{tr(\Sigma_b \otimes XX^T + \sigma_e \otimes I_n)} = \frac{tr(K)tr(\Sigma_b)}{tr(K)tr(\Sigma_b) + tr(\Sigma_e)}$$

Using moment matching, [**?**] provided the following MOM estimate:

$$\hat{h}^2_{MOM} = \frac{tr(K_R)tr(\hat{\Sigma}_b)}{tr(H_E)} = \frac{tr(K_E)}{tr(H_E)}\frac{tr(H_EK_E) - tr(H_E)tr(K_E)/n}{tr(K_E^2) - tr^2(K_E)/n},$$

which is approximately

$$\frac{tr(K_E)}{tr(H_E)}\sqrt{\frac{tr(H_E^2)}{tr(K_E^2)}}r(H_E, K_E).$$

When both X and Y are column-standardized and assumed that their ranks are the same as their number of columns, we have $tr(H_E) = np$ and $tr(K_E) = np$. It can be shown that

$$\hat{h}^2_{MOM} \in \left[\frac{1}{\sqrt{q}}r(H_E, K_E), \sqrt{p}r(H_E, K_E)\right]$$

Ridge-penalized multivariate linear models. It has been found that, when taking the dependence structure in Y into consideration, ridge regressions can be very useful tool for simultaneous prediction of multiple outcome variables ([**?**, **?**]). In the setting where both the independent and the outcome variables are high dimensional, we can use two tuning parameters, λ_x and λ_y when defining the similarity measures used in Mantel test:

$$H_{\lambda_y} = Y(Y^TY + \lambda_y I_q)^{-1}Y^T \ , \ K_{\lambda_x} = X(X^TX + \lambda_x I_p)^{-1}X^T \tag{5}$$

Similar to the univariate case, the ridge-penalize regression multivariate regression also connects fixed- and random-effect models through correlations. **Corollary. Connections between the correlations**.

$$\lim_{\lambda_x\to 0, \lambda_y \to 0} r(H_{\lambda_y}, K_{\lambda_x}) = r(H_M, K_M)$$

$$\lim_{\lambda_x \to \infty, \lambda_y \to \infty} r(H_{\lambda_y}, K_{\lambda_x}) = r(H_E, K_E)$$

To understand the underlying statistical implication of the measures, we consider data augmentation. Data augmentation has been widely used to ease the computation in several problems, such as the expectation and maximization algorithms, and Gibbs or MCMC sampling ((**???**), (**???**), (**???**), (**???**), (**???**), (**???**)). In particular, (Hodges (1998)) described how data augmentation can be used as a "computing device" so that hierarchical models can be written in the format of linear models. It can be verified that the data augmentation in corresponds to the similarity measures in Equation @ref(eqn: HK_lambda} is

$$\tilde{Y} = \begin{pmatrix} Y \\ \mathbf{0} \\ \sqrt{\lambda_y} I_q \end{pmatrix}, \tilde{X} = \begin{pmatrix} X \\ \sqrt{\lambda_x} I_p \\ \mathbf{0} \end{pmatrix} \tag{6}$$

Let $\tilde{H} = \tilde{Y}(\tilde{Y}^T\tilde{Y})^{-1}\tilde{Y}^T$, $\tilde{K} = \tilde{X}(\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T$. The least squares estimate (based on the augmented data) of B is

$$\hat{B} = (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\tilde{Y} = (X^TX + \lambda_x I)^{-1}X^TY,$$

which only depends on the tuning parameter for X. To see the effect of the tuning parameter λ_Y, we examine the log-likelihood based on the augmented data when the assumptions of the fixed-effect model is used:

$$l = -\frac{(n+p+q)q}{2}log(2\pi) - \frac{n+p+q}{2}log|\Sigma_e| - \frac{1}{2}tr[(Y-XB)\Sigma_e^{-1}(Y-XB)^T] - \frac{\lambda_x}{2}tr[B\Sigma_e^{-1}B^T] - \frac{\lambda_y}{2}tr[\Sigma_e^{-1}] \tag{7}$$

At first glance, the two tuning parameters seem to have different effects. While λ_x shrinks B toward zero, λ_y encourages smaller sums of the reciprocals of the eigenvalues of Σ_e. As discussed in (***?), shrinking B by the ridge penalization induced by λ_x is equivalent replace $X(X^T + \lambda_x I_p)^{-1}X^T$ with XX^T when λ_x is large. The penalization through λ_y encourages a small sum of the reciprocals of the eigenvalues of Σ_e; equivalently, it is encouraged that Σ_e has large and equally-sized eigenvalues. When λ_y is large, we expect that Σ_e is proportional to an identity matrix, which means we can replace $Y(Y^TY + \lambda_y)^{-1}Y^T$ with YY^T (up to some scalar). Thus, large values of tuning parameters favors the use of Euclidean distance, which ignores the dependence structure with X or Y.

Penalized inference with the adaptive Mantel test. Effectively using kernel methods requires an appropriate selection of the kernel function and tuning parameters for the particular setting. Selection methods have been extensively considered in the context of prediction problems, with cross-validation as the *de facto* standard. Cross-validation is a straight-forward and practical selection method for prediction, but may be difficult to implement for hypothesis testing, since the type I error rate needs to be controlled. Furthermore, the tuning parameter selected by minimizing the CV MSE may not necessarily yield the highest powered test. This section introduces the adaptive Mantel test (AMT), which extends the classical MT to simultaneously test across a set of tuning parameters and kernels without the need to directly apply adjustments for multiple comparisons.

The "adaptive" procedure used here is similar to the adaptive sum of powered score test algorithm described in Xu, Xu, and Pan (2017). The procedure receives as input a list of pairs of metrics/kernels $\{(\delta_m^{\mathbf{X}}, \delta_m^{\mathbf{Y}}) | m = 1, \cdots M\}$ from which the matrices $K_m = \delta_m^{\mathbf{X}}(X)$ and $H_m = \delta_m^{\mathbf{Y}}(Y)$ are computed for each metric pair, $m = 1, \cdots, M$. These metrics may be from a single family with varying tuning parameters, such as ridge kernels with different penalization terms, or may include kernels from different families.

For each $m = 1, \cdots, M$, P_m is calculated as the P-value of the Mantel test with metrics $\delta_m^{\mathbf{X}}$ and $\delta_m^{\mathbf{Y}}$ for X and Y respectively. The AMT test statistic is defined as the minimum of these values,

$$P^{(0)} := \min_{m=1,\cdots,M} P_m.$$

A permutation procedure can be used to calculate the reference distribution for $P^{(0)}$. For each m, and $b = 1, \cdots, B$, $H_m^{(b)}$ is generated by permuting rows and columns of H simultaneously, and the corresponding test statistic $P^{(b)}$ is calculated. The AMT P-value is then calculated as

$$P_{ADA} = \frac{1}{B+1} \sum_{b=0}^{B} I\left(P^{(0)} \leq P^{(b)}\right).$$

Algorithm 1: Adaptive Mantel Algorithm

Input: $X, Y, \Lambda = \{\lambda_j\}_{j=1}^m$.
Output: Adaptive Mantel P-value.

$H \leftarrow YY^T$
for $j = 1, \ldots, m$ **do**
  $K_{\lambda_j} \leftarrow X(X^TX + \lambda I_p)^{-1}X^T$
  $T_j^{(0)} \leftarrow \mathbf{tr}(HK_{\lambda_j})$
Generate B permutations of H, labeled $H^{(b)}, b = 1, \ldots, B$
$T_j^{(b)} \leftarrow \mathbf{tr}(H^{(b)}K_{\lambda_j}), \quad \forall b = 1, \ldots, B; j = 1, \ldots, m$
$P_j^{(b)} \leftarrow \frac{1}{B+1}\sum_{b'=0}^{B} \mathbb{1}(T_j^{(b)} \geq T_j^{(b')}), \quad \forall b = 1, \ldots, B; j = 1, \ldots, m$
$P^{(b)} \leftarrow \min_j P_j^{(b)}, \quad \forall b = 1, \ldots, B$
$P_{ADA} \leftarrow \frac{1}{B+1}\sum_{b=0}^{B} \mathbb{1}(P^{(0)} \leq P^{(b)})$

Extensions of the Adaptive Mantel Test Given the recent success of kernel methods for regression and prediction, there are likely many application settings for which a Mantel test with a non-linear similarity or distance metric will be more appropriate and effective than the linear metrics previously discussed. Kernel methods have been thoroughly developed for prediction problems, but, as with ridge regression, the use of kernels in inference settings have received relatively little attention. Through the use of the adaptive Mantel test, the difficulties of tuning parameter selection for inference can be partially mitigated, making it practical to test for the association of two sets of features for which the similarity of subjects is measured via a kernel similarity function rather than the typical Euclidean metric.

We here mention two kernels that are of interest for imaging genetics, namely the Gaussian radial basis function (RBF) and the identity by descent (IBD) kernels. The RBF kernel is defined as

$$\mathcal{K}(x, x') = \exp\left(-\lambda \|x - x'\|^2\right),$$

where $\|x - x\|^2$ is the squared Euclidean distance and λ is a bandwidth tuning parameter.

When a observational modality has a smooth manifold structure, it is also possible to define an inner product on the manifold to use as a similarity measure. In general. this approach is distinct from the kernel approach, in that the intrinsic manifold metric will induce curvature on the manifold, and will not be representable as a transformation of a Euclidean measure, whereas a kernel-based inner product will a representation of a weighted Euclidean inner product. The manifold approach may offer some advantage by utilizing the natural structure of the observed data, thus achieving some dimensionality reduction relative to using the Euclidean metric.

Yet another approach, particularly relevant for analysis of connectivity data, is the use of graph-theoretic similarity measures, which measure the similarity of graph topology and structure according to different properties, such as average degree, cut-norm, girth, eccentricity, average clustering coefficient, small-worldness, etc. Each of these metrics emphasize different aspects of graph structure, thus through the combination of these measures, one can develop similarity measures on graphs that are more interpretable than using the Frobenius norm or similar on the adjacency matrix.

While the above metrics can easily be used as part of a Mantel test, the theoretical properties of using these metrics has not yet been developed, and so it would be difficult to apply and interpret the results of such a test in a rigorous fashion. Building on the results of our work on the adaptive Mantel test, we propose to assess the properties of the Mantel test when different metrics are chosen, and determine which choice of metrics is most appropriate for answering particular scientific questions in imaging genetics studies.

3.1.1 References

- Kwee et al. (2008)
- Schaid (2010a)
- Schaid (2010b)
- Mantel (1967a)
- Robert and Escoufier (1976)
- Xu, Xu, and Pan (2017)

3.2 Aim 2: Models for Dynamic Connectivity

To develop **multi-subject dynamic connectivity models** for fMRI and EEG data. Many approaches to modeling dynamic connectivity have been proposed in recent years, but the majority of these models have not been fully developed to jointly model common connectivity states across multiple subjects, and many require strong assumptions regarding the scale and structure of relevant features. The developed models will extend the switching factor VAR model to multiple subjects, which will allow for the estimation of recurring connectivity states that are common across subjects, without requiring the specification of a time scale or the temporal alignment of subject data. The current switching factor VAR models will be enhanced with the methods developed in Aims 1 and 2 in order to identify common connectivity subspaces across subjects to facilitate multi-subject modeling. Motivated by the scientific goal of understanding the genetic influences on dynamic connectivity features, these models will be further developed to allow for covariate adjustment, and the estimation and interpretation of covariate effects. In particular, the developed models will be applied to assess the influence of genetics on various features of dynamic connectivity as measured by fMRI or EEG across a variety of experimental settings.

3.2.1 Aim 2.1: Heritability Analysis for Dynamic Functional Connectivity

A natural approach to estimating dynamic connectivity is through the use of a *sliding-window* method. For a $q \times T$ multivariate time series $Y(t)$, a connectivity measure $\gamma : \mathbb{R}^{q\times T} \to \mathbb{R}^{q\times q}$, and a window length L, the dynamic connectivity series of Y, denoted DC_Y, is the $q \times q \times T$ array of connectivity matrices, where $DC_Y(t) = \gamma\{Y(t : (t + L))\}$ is the connectivity matrix computed from applying γ to the length ℓ multivariate series starting at time t. While simple to describe and implement, sliding-window methods have received criticism for requiring a choice of window length, which limits the temporal resolution of the resulting estimates, and may obscure important dynamics at finer temporal scales.

A recent dynamic connectivity analysis by Cabral et al. (2018) instead applied a measure of *phase coherence connectivity* to analyze the dynamic functional connectivity from resting state fMRI for 98 healthy adults aged 50+. The phase coherence connectivity at time t is a "quasi-instantaneous" measure of functional connectivity, computed as

$$dFC(q, q', t) = \cos(\theta(q, t) - \theta(q', t)),$$

where $\theta(q, t)$ is the phase of the BOLD signal in region q at time t. Using this measure avoids the need to specify a window size, and ideally yields the finest temporal resolution allowed by the data.

After computing the dFC for $t = 1, \cdots, T$, the leading eigenvector (i.e. the eigenvector corresponding to the largest eigenvalue) for each matrix is extracted, denoted $V_1(t)$, and the rank 1 approximation to the connectivity matrix is constructed as $V_1(t)V_1(t)^T$. Cabral et al. justify the use of only the leading eigenvector by appealing to the common use of the leading eigenvector to detect community structures in the network analysis literature. From these reconstructed rank 1 connectivity matrices, a time-versus-time matrix to represent functional connectivity dynamics at times t_x and t_y is calculated using Pearson correlation or cosine similarity

$$FCD(t_x, t_y) = \frac{V_1(t_x) \cdot V_1(t_y)}{\|V_1(t_x)\|V_1(t_y)\|} \in [-1, 1],$$

where $FCD(t_x, t_y)$ measures the resemblance between the dFC at the two time points.

To determine discrete FC states, k-means clustering is applied for the leading eigenvectors $V_1(t)$ across all time points and subjects. In the analysis, Cabral et al choose the total clusters k from 2 to 20, and evaluate the cluster solution with Dunn's score. Finally, after separating the subjects into two groups from cognitive assessment results, the group differences in FC are tested using a permutation-based paired t-test and with the Kolmogorov-Smirnov (KS) distance between the FDC probability densities of the two groups.

We propose an extension of this method to test for differences in heritability and state switching patterns across groups of genetically similar individuals. Given a clustering of individuals into two groups according to genetic similarity, the KS distance between the groups can be tested directly. A test of the overall genetic influence can also be carried out using the Mantel test on the state transition matrices for all subjects, which would not require clustering subjects into two groups. Further, using the method of moments estimator for heritability of multivariate phenotypes from Aim 1, an estimate of the heritability of state transition probabilities can be calculated for either all states globally, or for each state separately, which could yield some insight into which states are strongly genetically linked.

A potential criticism of the dynamic connectivity analysis described here is the observation that the leading eigenvector may not capture those network dynamics that are most strongly associated with the covariates of interest. As part of our proposed research, we will further explore the advantages and challenges of using higher-rank approximations of the estimated dFC matrices, as well as bi-modal approximation methods that account for the strength of association with the covariates, such as CCA and PLS.

3.2.2 Aim 2.2: Multi-subject Switching Factor VAR

Functional connectivity measures are currently the most popular choice in the analsis of the connectome, due in part to their simplicity and ease of estimation. However, functional connectivity is an undirected measure of correspondence between two brain regions, and are not capable of capturing directed relationships, such as when the activity in one brain region modulates activity in another. Effective connectivity is instead an asymmetric, directed measure of the relationship between two brain regions. Static effective connectivity is typically measured with a VAR model, from which the partial directed coherence can be estimated to determine how fluctuations in a particular frequency band are related across brain regions. Our collaborators have proposed the use of a switching factor VAR (sf-VAR) to estimate the dynamic effective connectivity between regions ((**???**)).

Figure @fig:samdin (Samdin 2017) gives the steps of the dynamic connectivity modeling for a single subject, which we briefly describe here. Let $Y(t)$ be a $q \times T$ multivariate time series as above, and let L be a pre-selected window length as in the sliding-window methods. The sf-VAR estimation procedure extracts initial features by fitting a VAR at each time t for the series $Y\,(t : (t + L))$. The model can be written

$$Y_{it} = \sum_{\ell=1}^{L} \Phi_{i\ell t} Y_{i,t-\ell} + v_{it}$$

where $\Phi_{i\ell t}$ is the $Q \times Q$ matrix of VAR coefficients for lag ℓ at time t and observational noise $v_{it} \sim N(0, \Sigma_{iv})$.

State-space form and model estimation. For estimation, it is convenient to work with this model in a state-space form, as follows. Define $a_{it} = vec([\Phi_{it1}, \cdots, \Phi_{itL}]')$

$$a_{it} = a_{i,t-1} + w_{it}$$

where a_{it} is the hidden state, assumed to follow a first-order Gauss-Markov process, $w_{it} \sim N_{LQ^2}(0, \Sigma_{wi})$.

$$Y_{it} = C_{it} a_{it} + v_{it},$$

with $C_{it} = I_q \otimes X'_{it}$, where $X_{it} = [Y'_{i,t-1}, \cdots, Y'_{i,t-L}]$ is a $qL \times 1$ vector of past observations.

We can estimate $\hat{a}_{it} = \mathbb{E}(a_{it}|Y_i)$ for each $i = 1, \cdots, N$ using the Kalman Filter to adaptively extract the sequence of estimated TV-VAR coefficients. From the estimated TV-VAR coefficient matrices, we then calculate the estimated PDC from region q to region q' as

$$\hat{\pi}_{qq'}(f) = \frac{|\Phi_{qq'}(f)|}{\sqrt{\sum_{j=1}^{Q} |\Phi_{jq'}(f)|^2}}.$$

Clustering into distinct regimes. Let $\hat{\Pi}$ be the $q \times q \times nT$ tensor of estimated PDC matrices, concatenating the arrays for all subjects.

Apply K-means tensor clustering via higher-order singular value decomposition (HOSVD) by minimizing

$$\min_{\{C_k\}} \sum_{i=1}^{N} \min_{1 \le k \le K} \|M_i - C_k\|^2 = \sum_{k=1}^{K} \sum_{c \in C_k} \|M_i^{(k)} - C_k\|^2,$$

where C_k is the centroid of cluster k and $M^{(i)}$ is the loading matrix for subject i.
By Huang 2008, this is equivalent to solving

$$\min_{U,V,M} J = \|\hat{\Pi} - U \otimes V \otimes M\|^2 = \sum_{i=1}^{n} \|\hat{\Pi}_i - U \otimes V \otimes M_i\|^2.$$

This results in K clusters which can then be identified with K latent brain states that are common across subjects. This can also be written as

$$\hat{\Pi}_i = U \otimes V \otimes M_i$$

where U and V are fixed for all subjects, and M_i is the subject-specific "core" matrix.

If desired, it is also possible to do dimension reduction in the channel-channel space simultaneously with the K-means clustering step. This requires the choice of reduction rank R, which could possibly be chosen usinig AIC/BIC or similar.

Covariate effects on state transition probabilities. After segmenting the signals for model initialization with a state sequence $\hat{S}_t^{KM}$, the estimates within each regime can be refined using a Markov switching-VAR model:

$$Y_{it} = \sum_{\ell=1}^{L} \Phi_{i\ell}^{[S_t]} Y_{i,t-\ell} + v_{it}.$$

Assuming the state transitions for each subject i follow a hidden Markov chain, the state transition matrix $Z_i = [z_{ijk}], 1 \le j, k \le K$ can be estimated, where

$$z_{ijk} = P(S_{it} = j | S_{i,t-1} = k)$$

denotes the probability of transition from state k at time $t-1$ to state j at time t. Having estimated these state transition probability matrices for all subjects, the differences in state transitions between groups can be tested using the paired t-test or KS distance, similar to the method used for dFC described above. Heritability of state transition probabilities can be also be assessed following existing methods for multivariate heritability estimation.

References:

- Samdin et al. (2017)
- Ting, Ombao, and Salleh (2017)
- Taghia et al. (2017)
- S. Chiang et al. (2017)

- Hutchison et al. (2013)
- Havlicek et al. (2010)

3.3 Aim 3: Manifold Regression Models for Imaging Genetics

3.3.1 Aim 3.1: Manifold Regression for Connectivity Differences

Although the foundational ideas for manifold regression models are very old, reaching back to Fisher's development of regression models for circular data, there has been a renewed interest in models for manifold-valued data in recent years. Initially spurred by impressive progress in predictive settings with machine learning methods, these ideas have been developed into tools for the analysis of covariance in time series, and neuroimaging applications. The work by Zhu et. al (Zhu et al. (2009)) presents a semiparametric regression model for manifold-valued responses, for application to diffusion tensor imaging data. To map the structural connectivity of the brain, DTI measures a 3×3 positive-definite matrix at each voxel, which captures the degree and direction that water molecules tend to diffuse along white matter tracts in the brain. If the goal is to estimate differences in connectivity structure adjusting for covariates, classical multivariate regression can be applied, but this introduces technical and theoretical difficulties, since the resulting estimates must be modified in some way to ensure that the intrinsic constraints of the observation space are respected when making predictions, i.e., one has to somehow guarantee that predicted values are in the space of positive-definite matrices. This is a common difficulty in modeling and estimating characteristics of positive-definite matrices. A theoretically natural alternative then is to model the data on the PD manifold, denoted instead of in Euclidean space.

As an example, the following model is proposed by Zhu et al. as one possibility. Suppose a $q \times q$ PD matrix $S_i \in \text{Sym}^+(q)$ and a $k \times 1$ vector of covariates are observed for each subject $i = 1, \cdots, n$. Let $\beta \in \mathbb{R}^p$ be a $p \times 1$ vector of regression coefficients, and $\Sigma(\cdot, \cdot) : \mathbb{R}^{\urcorner} \times \mathbb{R}^{\shortmid} \to \text{Sym}^+(q)$. We are interested in modeling the "conditional mean" of S_i given x_i, denoted $\Sigma_i(\beta) = \Sigma(x_i, \beta) \in \text{Sym}^+(q)$. To state the "geodesic" model for Σ, let $D \in \text{Sym}^+(q)$ be the intercept matrix $D = \Sigma(0, \beta)$, with $D = BB^T$ for some $B \in GL(q)$, let $Y_D(x_i, \beta) = Y_{D,i} \in \text{Sym}(q)$ be a "directional" matrix, and let $C_i(\beta)$ be a Cholesky square root $\Sigma(x_i, \beta) = C_i(\beta)C_i(\beta)^T$. The geodesic model then assumes that

$$\Sigma(x_i, \beta) = B \exp(B^{-1} Y_{D,i}(\beta) B^{-T}) B^T = C_i(\beta) C_i(\beta)^T.$$

Following a result in Schwartzman (2006), $\Sigma(x_i, \beta)$ is related to the geodesic through D in the direction of $Y_{D,i}$. The residual of S_i with respect to $\Sigma_i(\beta)$ from this model is defined as

$$\mathcal{E}_i(\beta) = \log(C_i(\beta)^{-1} S_i C_i(\beta)^{-T}).$$

The intrinsic regression model is then specified by $\mathbb{E}[\mathcal{E}_i(\beta)|x_i] = 0$, where the intrinsic least squares estimate of β is defined as

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} \text{tr}\left(\mathcal{E}_i(\beta)^2\right).$$

3.3.2 Aim 3.2: Mixed Effects Manifold Regression for Heritability of Brain Connectivity

The classical mixed effects model has become a popular and powerful tool for analyzing longitudinal measurements of real-valued response data. In the case of n subjects measured at T regularly spaced times, with response y_{it} for subject i at time t, fixed-effects covariates X_i, and random effects covariates Z_i which link the subject specific random effects with the response, the linear mixed effects model can be written as

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i,$$

where $b_i \sim N(0, D)$ is a vector of subject specific random effects, and $\varepsilon_i \sim N(0, R)$ is a vector of error terms that are assumed to be independent and identically distributed across subjects, and with b_i and ε_i assumed independent (Laird and Ware (1982)).

Kim et al. (2017) have proposed an extension of the linear mixed effects model for manifold-valued responses. Let $Y_{[ij]}, B, B_i \in \mathcal{M}, V \in T_B\mathcal{M}^p, U_i \in T_{h[ij]}\mathcal{M}^q, x_{[ij]} \in \mathbb{R}^p, z_{[ij]} \in \mathbb{R}^q$, and let $\Gamma_{B\to B_i}V$ be the parallel transport of V from B to B_i. A simplified form of the model can be stated as

$$Y_{[ij]} = \mathrm{Exp}(\mathrm{Exp}(B_i, \Gamma_{B\to B_i}(V)x_{[ij]}, \varepsilon_{[ij]}))$$
$$B_i = \mathrm{Exp}(B, U_i).$$

Similar to the classical setting, the random effects U_i are assumed to follow a normal distribution,

$$\Gamma_{B\to I}U_i \ \sim \ \mathcal{N}_{SYM}(0, \sigma_U^2),$$

where $\mathcal{N}_{SYM}$ is the normal distribution over symmetric positive definite matrices.

Analagous to the definition of the random effects model for heritability analysis, we propose a random effects model for manifold-valued responses measured from a set of unrelated subjects with SNP matrix X and genetic relationship matrix $G = XX^T/p$. For this model, we instead consider the random effects U_i to have a common aggregate genetic effect σ_U^2, such that the contribution of this effect to each subject's response is proportional to the genetic relationship matrix, giving random effect distribution $\Gamma_{B\to I}U \ \sim \ \mathcal{N}_{SYM}(0, \sigma_U^2 G)$.

3.3.3 References

- Kim et al. (2017)
- Yang et al. (2011)
- Laird and Ware (1982)

3.4 Aim 4: Broad Scientific Impact: Data Analysis and Software Development

3.4.1 Aim 4.1: Identifying Significantly Heritable Features of EEG Coherence

3.4.1.1 Materials and Methods

Our collaborators at the Beijing Normal University (BNU) Center for Brain and Learning Sciences have conducted a large imaging genetics study, which collected data from approximately 1000 healthy college Chinese students participated. The study participants were given a number of standard psychological tasks to measure cognitive traits such as working memory performance, learning rates, and risk-taking behavior. During each of these tasks, fMRI or EEG data was recorded. Structural MRI, DTI, and SNP data was also collected from the subjects.

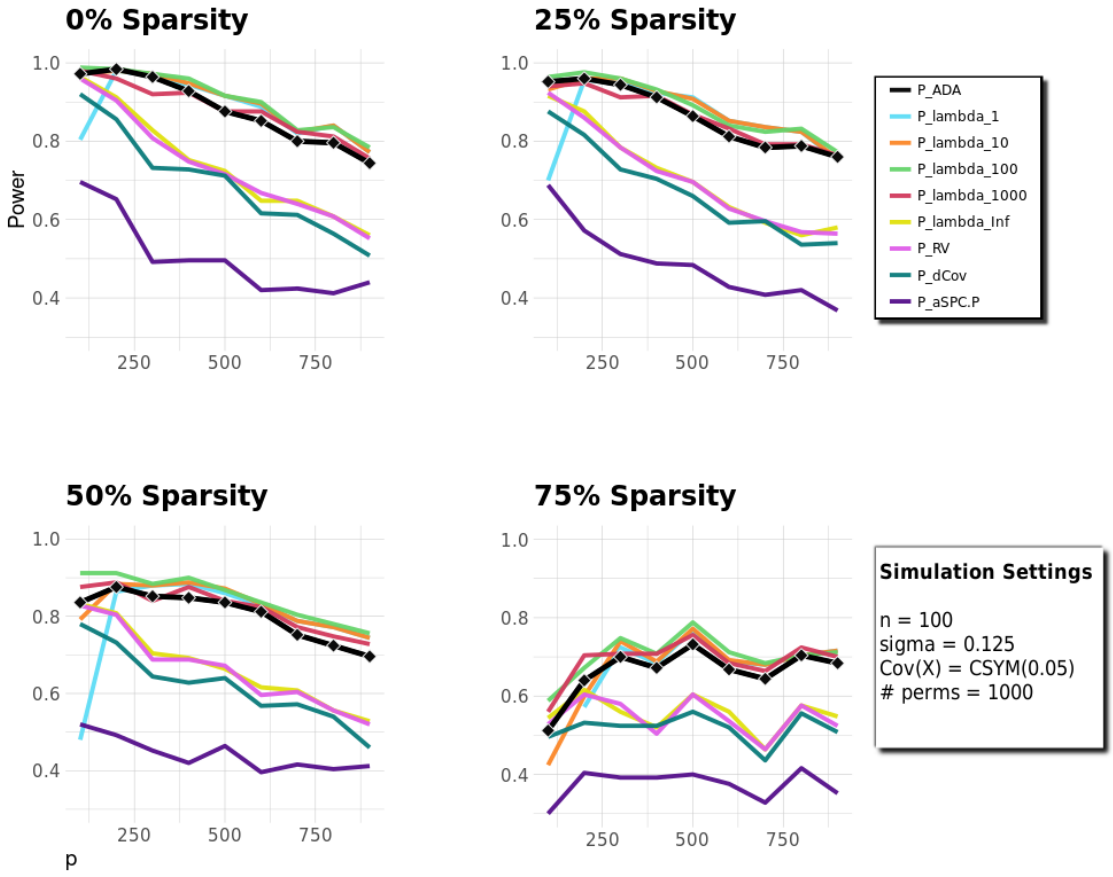
For our preliminary analysis, we have considered data from 350 BNU subjects who participated in a visual working memory task, during which 64-channel EEG was recorded at 1 kHz. The total duration of the experiment was 5-10 minutes for each subject. Approximately 5×10^5 SNPs were also measured for each subject. Standard pre-processing and quality control steps were applied to both the EEG and genetic data. The goal of the analysis is to determine the association of theta, alpha, and beta band coherence with a group of 13 SNPs that have been identified as potentially related to Alzheimer's Disease.

This results in 2080 distinct features when using all 64 channels, and 300 distinct features for the 25 selected frontal channels. The adaptive Mantel test was performed with $\lambda \in \{0.5, 1, 5, 10, 100, 1000, \infty\}$ and using 1000 permutations. Genetic similarity of subjects was calculated as the L_2 inner product of the centered standardized SNP data for all tests.

The coherence between two EEG channels at a particular frequency ω is a measure of the oscillatory concordance of the the two signals at ω. The pairwise coherence for q EEG channels is a $q \times q$ symmetric matrix, from which we extract the upper triangle and vectorize to form the $n \times \binom{q}{2}$ matrix X. This results in 2080 distinct features when using all 64 channels, and 300 distinct features for the 25 selected frontal channels. The adaptive Mantel test was performed with $\lambda \in \{0.5, 1, 5, 10, 100, 1000, \infty\}$ and using 1000 permutations. Genetic similarity of subjects was calculated as the L_2 inner product of the centered standardized SNP data for all tests. For the alpha band, all 64 channels gave $P = 0.065$, and the selected frontal channels gave $P = 0.381$. Results for the theta band were $P = 0.416$ and $P = 0.085$ for all 64 channels and frontal channels respectively. Since the adaptive Mantel test was used, these P-values already take into account testing across multiple λ. In the case of the alpha band, the test results suggest that coherence involving channels outside of the selected frontal channels may be associated with genetic similarity determined by the 11 AD SNPs, whereas for the theta band, the SNP association appears stronger when considering only the frontal channels. For a better sense of the significance of these 11 SNPs, the AMT P-values for 200 sets of 11 randomly selected SNPs were calculated for each of the four tests considered here. The boxplot of these values are given in Figure **??**. For the alpha – all channels test and the theta – frontal channels test, the P-values from the AD SNPs (0.065 and 0.085 respectively) are outside the ranges of P-values from the randomly selected SNPs.

3.5 Preliminary Results from Simulations

Preliminary simulation results have indicated the feasibility of the Adaptive Mantel as an alternative to multiple comparison adjustments for a collection of penalized Mantel tests in terms of statistical power for the simulation settings considered. Moreover, relative to competing methods, including the classical distance Mantel test, RV-coefficient and distance-covariance tests, and the adaptive sum of powered score tests, the Adaptive Mantel test has sufficiently high power to make it a practical and easily implemented alternative.



3.6 Preliminary Results from BNU Data: Association of EEG Coherence and Selected SNPs

For the alpha band, all 64 channels gave $P = 0.065$, and the selected frontal channels gave $P = 0.381$. Results for the theta band were $P = 0.416$ and $P = 0.085$ for all 64 channels and frontal channels respectively. Since the adaptive Mantel test was used, these P-values already take into account testing across multiple λ. In the case of the alpha band, the test results suggest that coherence involving channels outside of the selected frontal channels may be associated with genetic similarity determined by the 13 AD SNPs, whereas for the theta band, the SNP association appears stronger when considering only the frontal channels.

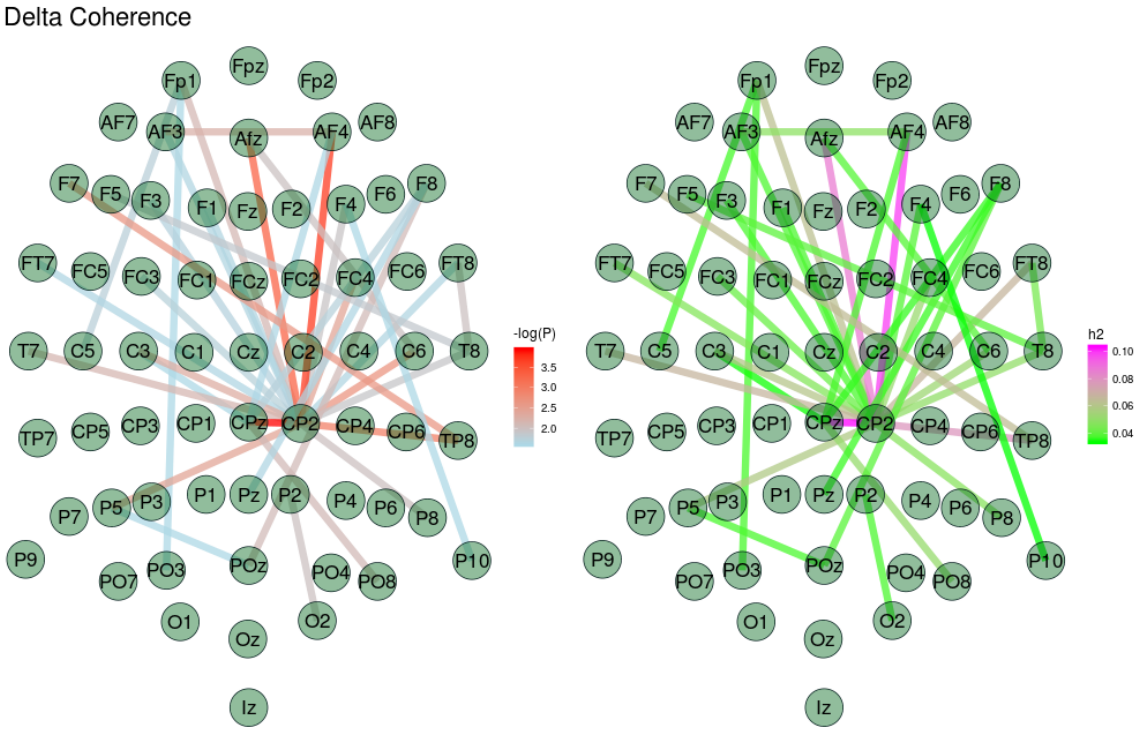
To assess possible contributions of individual channel pairs, each channel pair was separately tested for significance with the AD SNPs with the GCTA toolkit (Yang et al. (2011)). Figure **??** shows the most significant channel pairs across all channel pairs for the alpha band. The TP8 – CP2 connection is the most significant at $P < 0.0001$ (with no adjustment for multiple comparisons). Other top pairs are C4 – PO4, C4 – AF7, and T8 – FT7. The most significant connections are mostly between the right temporal regions with the left frontal regions. For the theta band, the most significant channel pairs were P9 – Cz, P9 – CP3, CP3 – AF3, and P8 – F3. The P9 channel also had relatively significant connections with many other channels in the frontal left hemisphere. These results are supported by a number of previous studies that have established links between working memory performance and features measured by EEG. For instance, Onton, Delorme, and Makeig (2005) found increases in frontal midline theta power with increasing memory load during a verbal-working memory task; Sauseng et al. (2005) also found that alpha coherence plays a significant role in "top-down" control during working memory tasks; and Simons and Spiers (2003) identified important

Band	Channels	P-value
β	All	0.619
β	Frontal	0.517
α	All	0.075
α	Frontal	0.381
θ	All	0.416
θ	Frontal	0.081
δ	All	0.015
δ	Frontal	0.088

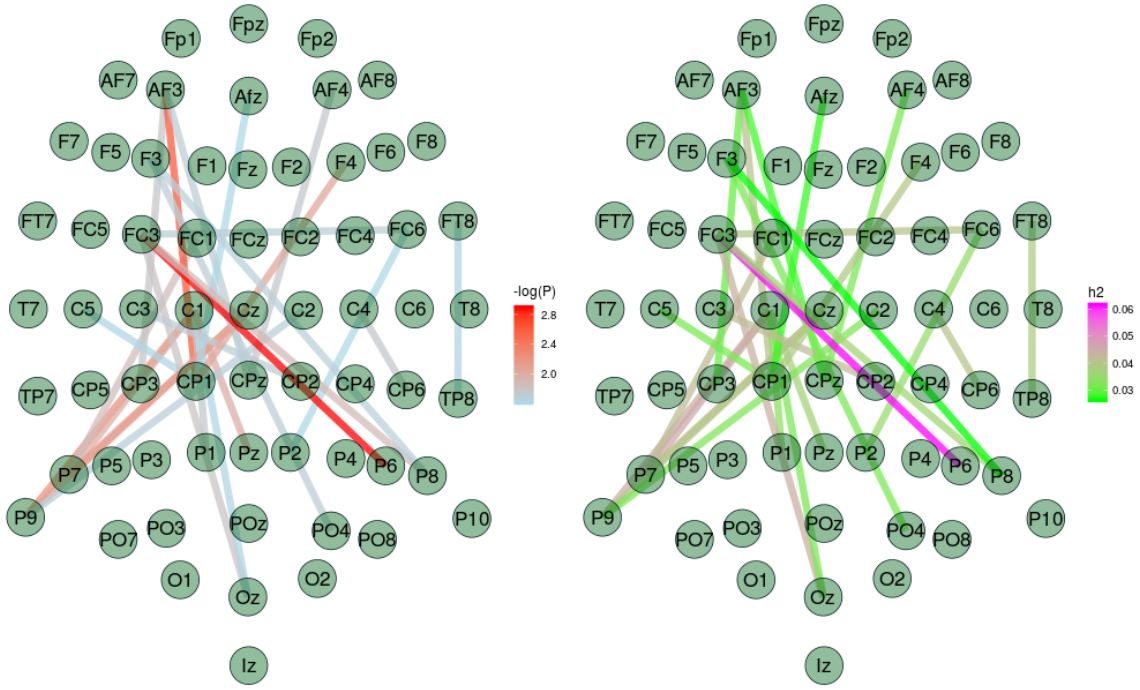
Table 1: Testing results for associations of EEG coherence and AD SNPs. Previous studies have found evidence that power in the θ, α, and β bands is important for working-memory performance. The delta band results are also presented for comparison.

interactions between the prefrontal and medial temporal lobes for the processing of long-term memory.

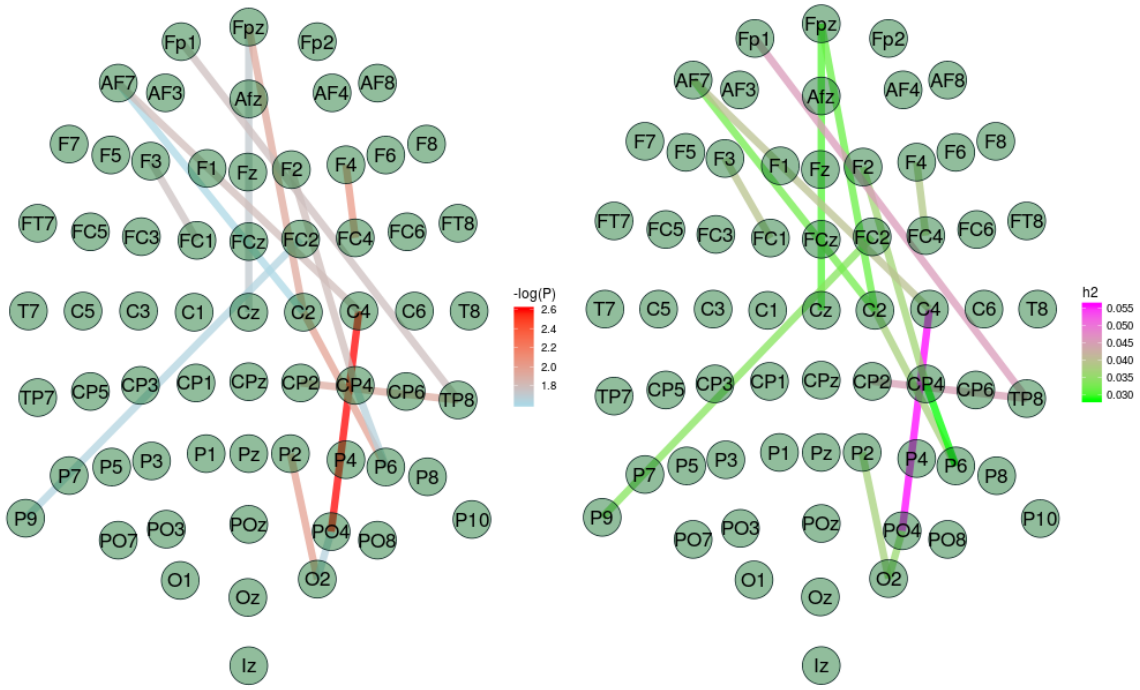
Similarly, the association of each individual SNP with EEG coherence was assessed with AMT. For alpha coherence with all channels, the most significant SNP ($P = 0.02$) is rs2227564, a functional polymorphism within plasminogen activator urokinase (PLAU) gene. An allele of this SNP has been linked to significantly high plaque counts in AD, although its role is not well-established. The second most significant SNP from the individual tests is rs3851179, a SNP upstream of the PICALM gene. This SNP has been repeatedly implicated as a factor in AD, as well as Parkinson's Disease and schizophrenia, although there are dissenting results regarding its significance in particular Chinese populations. In the genetics literature on cognitive function in healthy subjects, polymorphisms in neurotransmitter genes, such as those in the dopamine pathway, have been shown to be significantly associated with increased neuronal activity in the prefrontal cortex during working memory tasks, as measured by fMRI Bertolino et al. (2006). Vogler et al. analyzed data from the n-back memory task for 2298 subjects, and estimated genome-wide heritability of working memory accuracy to be 41% (95% CI: 0.13, 0.69) Vogler et al. (2014).

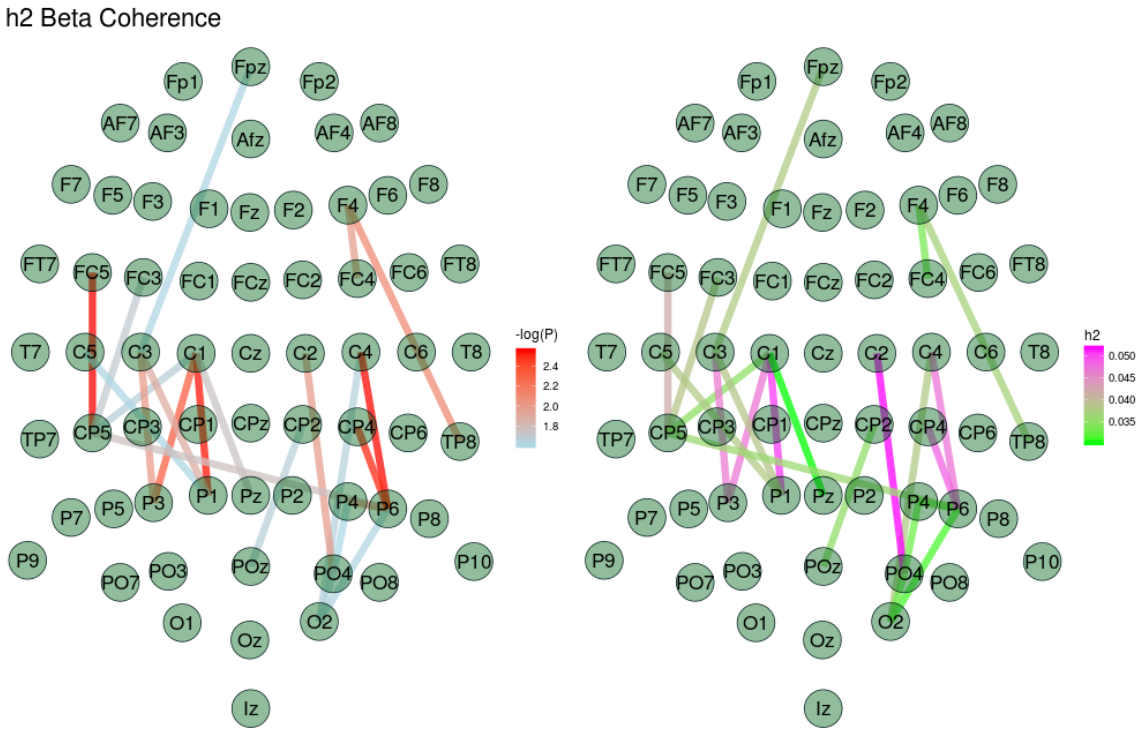


h2 Theta Coherence



h2 Alpha Coherence





3.6.1 Aim 4.3: R Package for Adaptive Mantel Test

3.6.1.1 AdaMant Package

- Current version: https://github.com/dspluta/adamant

3.6.1.1.1 Installation

1. Install using the `devtools` package:

```
install.packages(c("devtools", "tidyverse"))
devtools::install_github("github.com/dspluta/adamant")
```

2. After installation, load the package as usual:

3.6.1.1.2 Example 1: $n = 50, p = 10$

```
# Generate Example Data: n = 50, p = 10
set.seed(1234)
n <- 50
p <- 10
X <- matrix(rnorm(n * p), nrow = n, ncol = p)
Y <- X %*% rep(c(0, 0.6), p / 2) + rnorm(n, 0, 2)

# Apply Adaptive Mantel Test
adamant(X, Y, lambdas_X = c(0, 1, 10, Inf),
        n_perms = 2000, P_val_only = TRUE)
```

```
## ------------------------------
## Adaptive Mantel Output
```

```
## -------------------------------
## P_val         = 0.0055
## n             = 50
## p             = 10
## rank(X^TX)    = 10
## kappa         = 1.982
## Best Lambda   = 10
## n_perms       = 2000
## time          = 0.1 secs
## -------------------------------
```

```
## [1] 0.0055
```

3.6.1.1.3 Example 2: $n = 500, p = 100$

```
# Generate Example Data: n = 500, p = 100
set.seed(1234)
n <- 500
p <- 100
X <- matrix(rnorm(n * p), nrow = n, ncol = p)
Y <- X %*% rep(c(0, 0.06), p / 2) + rnorm(n, 0, 2)

adamant(X, Y, lambdas_X = c(0, 1, 10, Inf),
        n_perms = 2000, P_val_only = TRUE)
```

```
## -------------------------------
## Adaptive Mantel Output
## -------------------------------
## P_val         = 0.149
## n             = 500
## p             = 100
## rank(X^TX)    = 100
## kappa         = 2.458
## Best Lambda   = Inf
## n_perms       = 2000
## time          = 2.285 secs
## -------------------------------
```

```
## [1] 0.149
```

3.6.1.1.4 Example 3: $n = 500, p = 1000$

```
# Generate Example Data: n = 500, p = 1000
set.seed(1234)
n <- 500
p <- 1000
X <- matrix(rnorm(n * p), nrow = n, ncol = p)
Y <- X %*% rep(c(0, 0.03), p / 2) + rnorm(n, 0, 2)

adamant(X, Y, lambdas_X = c(0.1, 1, 10, Inf),
        n_perms = 2000, P_val_only = TRUE)
```

```
## -------------------------------
## Adaptive Mantel Output
## -------------------------------
```

```
## P_val         = 0.116
## n             = 500
## p             = 1000
## rank(X^TX)    = 499
## kappa         = 3121075104362398
## Best Lambda   = Inf
## n_perms       = 2000
## time          = 9.727 secs
## -------------------------------
```

```
## [1] 0.116
```

4 Appendix

4.1 Appendix: Aim 1

4.1.1 Computational Methods

If the feature space is very high-dimensional or if n is large, a straightforward implementation of Algorithm 1 may be computationally impractical. However, when only ridge kernels with varying values of λ are included in AMT, there are two approaches that can be used to greatly reduce the computational cost.

The first approach utilizes the SVD $X = UDV^T$. The computational complexity needed for finding the SVD for X is $O(np^2)$. Once the SVD is computed, we can compute $Z = U^TY$ and $S_\lambda = \sum_{i=1}^{r} \frac{\eta_i}{\lambda+\eta_i} z_i^2$, which has a total complexity of $O(nr)$. Note that when $p \gg n$, the rank r is often the same as n; as a result, the cost needed for calculating S_λ is $O(nr) = O(n^2)$. Calculating the test statistics for B permutations requires $O(Bn^2)$, for a total computational complexity of $O(np^2 + Bn^2)$.

Alternatively, when p is very large relative to n, we can use the identity $X(X^TX + \lambda I)^{-1} = (XX^T + \lambda I)^{-1}X$, so that the matrix inverse is applied to an $n \times n$, rather than $p \times p$, matrix. From this identity, K_λ can be rewritten as

$$\begin{aligned} K_\lambda &= X(X^TX + \lambda I_p)^{-1}X^T \\ &= (XX^T + \lambda I_n)^{-1}XX^T \end{aligned}$$

Note that calculating K_λ involves multiplying the $n \times p$ matrix X and the $p \times n$ matrix X^T, multiplying two $n \times n$ matrices, and inverting an $n \times n$ matrix. When $p \gg n$, the computation cost is dominated by calculating XX^T, which has a complexity of $O(n^2p)$. The Mantel test statistic can be calculated as

$$S_\lambda = \mathrm{tr}(YY^TK_\lambda) = Y^TK_\lambda Y,$$

which has a complexity of $O(n^2)$. With B permutations, the total computational complexity is $O(n^2p + Bn^2)$, which is less than the required computational complexity using SVD. Thus, switching from the feature space to the subject space (i.e., from a $p \times p$ similarity matrix of the features to an $n \times n$ similarity matrix of the subjects), has a computational advantage. Additionally, in some situations the SVD computation may be unstable, thus the matrix identity method may be recommended as the more robust approach.

4.1.2 Variance Explained and the Ridge Penalty

By allowing for simultaneous testing over a set of tuning parameter values, AMT lessens the challenge of parameter selection, but does not completely resolve it. Although it does not require an overly conservative adjustment for multiple testing, the power of AMT does decrease as the number of metrics considered increases. Conversely, the test results are highly sensitive to the choice of parameters to test. Consequently, one must still take some care in the selection of the included parameters, balancing the desire to use a wide range of parameter values, with the gains of using a small set of parameters. When only ridge kernels are included in AMT, previous results on the role of the ridge penalty term in predictive modeling can help with the identification of a reasonable set of values to test. Specifically, it has been shown that when the ridge penalty λ is chosen to be the noise to signal ratio, the resulting shrinkage estimator $\hat{\beta}_\lambda$ is identical to the best linear unbiased predictor (BLUP) $\hat{b}$ for the random effects model, and moreover, for a new observation with unknown response value the predictions using the ridge and random effects models are the same (Campos et al. (2013)). Thus, when using ridge regression for prediction, it is recommended that the penalty should reflect the relevant level of noise versus signal, i.e., $\lambda = \sigma_e^2/\sigma^2$.

To apply this result to practical settings, if one can determine *a priori* a likely range for the noise to signal ratio or a related quantity, this will determine a reasonable range of penalty terms. For instance, in assessing the genetic influence on observed phenotypes, the noise to signal ratio is related to what is known as the

heritability of the phenotype, which can be understood as the proportion of variance in the observed trait explained by the genetic data. Formally, for (standardized) genetic data matrix X consisting of alleles for p single nucleotide polymorphisms (SNPs), and observed phenotype vector Y, the random effects model given in Eq. @ref(eq:random} is commonly used to estimate the genetic heritability of the trait (Yang et al. (2011), liu2007semiparametric). From this model, the *narrow-sense heritability* h^2 of the phenotype is defined as

$$h^2 := \frac{p\sigma^2}{p\sigma^2 + \sigma_\varepsilon^2}.$$

Plugging in $Var(Y) = \sigma^2 + \sigma_\varepsilon^2 XX^T$ gives

$$h^2 = \frac{\sigma^2 tr(XX^T/n)}{\sigma^2 tr(XX^T/n) + \sigma_\varepsilon^2 tr(I_n/n)} = \frac{tr(XX^T/n)}{tr(XX^T/n) + \sigma_\varepsilon^2/\sigma^2}.$$

We see then that if h^2 is known, the optimal penalty $\tilde{\lambda}$ (for prediction) can be found by solving for the noise to signal ratio. In practice, the scientific interpretation and range of plausible values of h^2 will depend on the specific modalities of $\mathbf{X}$ and $\mathbf{Y}$. For instance, in the genetics literature, $h^2 > 0.5$ would generally indicate high heritability, while a heritibility of $h^2 < 0.1$ is probably not scientifically interesting. As a point of reference, most estimated heritability in the UK Biobank data is between 0.1 to 0.4 (Ge et al. (2017)).

4.1.3 Asymptotic Results for the Penalized Mantel Test

Assume true model is the random effects

$$Y = g + \varepsilon,$$

where $g \sim \mathcal{N}(0, \sigma^2 XX^T), \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Assume the SVD of X is $X = UDV^T$, with singular values d_j and eigenvalues $\delta_j = d_j^2, j = 1, \cdots, p$. Further assume that X and Y have been centered and scaled, so that $\text{tr}(XX^T) = np$, and $\text{tr}(X(X^TX)^{-1}) = p$.

4.1.3.1 Random Effects Test

Let $K_R = XX^T$.

$$T_R = \text{tr}(XX^T) = \sum_{j=1}^{p} \delta_j z_j^2 \tag{8}$$

$$\mathbb{E}T_R = \text{tr}(\sigma^2 K_R^2) + \sigma_\varepsilon^2 np = \sigma^2 \sum \delta_j^2 + \sigma_\varepsilon^2 np \tag{9}$$

$$T_R^* = \frac{T_R - \sigma_\varepsilon^2 np}{\sigma_\varepsilon^2 \sqrt{2\sum \delta_j^2}} \overset{H_0}{\dot{\sim}} \mathcal{N}(0,1) \tag{10}$$

$$\mathbb{E}T_R^* = \frac{\mathbb{E}T_R - \sigma_\varepsilon^2 np}{\sigma_\varepsilon^2 \sqrt{2\sum \delta_j^2}} = \frac{\sigma^2}{\sigma_\varepsilon^2} \sqrt{\frac{1}{2} \sum \delta_j^2} \tag{11}$$

4.1.3.2 Fixed Effects Test

Let $K_F = X(X^TX)^{-1}X^T$.

$$T_F = \text{tr}(K_F) = \sum_{j=1}^{p} z_j^2 \tag{12}$$

$$\mathbb{E}T_F = \text{tr}(\sigma^2 K_F K_R) + \sigma_\varepsilon^2 p = \sigma^2 np + \sigma_\varepsilon^2 p \tag{13}$$

$$T_F^* = \frac{T_F - \sigma_\varepsilon^2 p}{\sigma_\varepsilon^2 \sqrt{2p}} \overset{H_0}{\sim} \mathcal{N}(0,1) \tag{14}$$

$$\mathbb{E}T_F^* = \frac{\mathbb{E}T_F - \sigma_\varepsilon^2 p}{\sigma_\varepsilon^2 \sqrt{2p}} = \frac{\sigma^2}{\sigma_\varepsilon^2} \cdot n\sqrt{\frac{p}{2}} \tag{15}$$

4.1.3.3 Ridge Kernel Test

Let $K_\lambda = X(X^T X + \lambda I_p)^{-1} X^T$.

$$T_\lambda = \text{tr}(K_\lambda) = \sum_{j=1}^{p} \frac{\delta_j}{\delta_j + \lambda} z_j^2 \tag{16}$$

$$\mathbb{E}T_\lambda = \text{tr}(\sigma^2 K_\lambda K_R) + \sigma_\varepsilon^2 \sum \frac{\delta_j}{\delta_j + \lambda} \tag{17}$$

$$T_\lambda^* = \frac{T_\lambda - \sigma_\varepsilon^2 \sum \frac{\delta_j}{\delta_j+\lambda}}{\sigma_\varepsilon^2 \sqrt{2 \sum \left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}} \overset{H_0}{\sim} \mathcal{N}(0,1) \tag{18}$$

$$\mathbb{E}T_\lambda^* = \frac{\mathbb{E}T_\lambda - \sigma_\varepsilon^2 \sum \frac{\delta_j}{\delta_j+\lambda}}{\sigma_\varepsilon^2 \sqrt{2 \sum \left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}} = \frac{\sigma^2}{\sigma_\varepsilon^2} \cdot \frac{\text{tr}(K_\lambda K_R)}{\sqrt{2 \sum \left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}} \tag{19}$$

$$\mathbb{E}T_\lambda^* = \frac{\sigma^2}{\sigma_\varepsilon^2} \cdot \frac{\sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda}}{\sqrt{2 \sum \left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}} \tag{20}$$

(Trace identity for $\text{tr}(K_\lambda K_R)$ shown below.)

4.1.3.4 Asymptotic Equivalences

For n sufficiently large, we can assume the expected values of the test statistics are

$$\mathbb{E}T_F^* = \frac{\mathbb{E}T_F - \sigma_\varepsilon^2 p}{\sigma_\varepsilon^2 \sqrt{2p}} = \frac{\sigma^2}{\sigma_\varepsilon^2} \cdot n\sqrt{\frac{p}{2}}$$

$$\mathbb{E}T_R^* = \frac{\mathbb{E}T_R - \sigma_\varepsilon^2 np}{\sigma_\varepsilon^2 \sqrt{2 \sum \delta_j^2}} = \frac{\sigma^2}{\sigma_\varepsilon^2} \sqrt{\frac{1}{2} \sum \delta_j^2}$$

$$\mathbb{E}T_\lambda^* = \frac{\sigma^2}{\sigma_\varepsilon^2} \cdot \frac{\sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda}}{\sqrt{2 \sum \left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}}$$

Of course for $\lambda = 0$ we have $\mathbb{E}T_\lambda^* = \mathbb{E}T_F^*$ as usual. Considering the limit as $\lambda \to \infty$ of the square of the second quotient in T_λ^* (which is the only part that depends on λ):

$$\lim_{\lambda\to\infty} \frac{\left(\sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda}\right)^2}{2\sum\left(\frac{\delta_j}{\delta_j+\lambda}\right)^2} = \lim \frac{\frac{1}{1/\lambda}}{\frac{1}{1/\lambda}} \frac{\left(\sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda}\right)^2}{2\sum\left(\frac{\delta_j}{\delta_j+\lambda}\right)^2}$$

$$= \lim \frac{\left(\sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j/\lambda+1}\right)^2}{2\sum\left(\frac{\delta_j}{\delta_j/\lambda+1}\right)^2}$$

$$= \frac{\left(\sum_{j=1}^{p} \delta_j^2\right)^2}{2\sum \delta_j^2} = (\mathbb{E}T_R^*)^2$$

$$\Rightarrow \lim_{\lambda\to\infty} \mathbb{E}T_\lambda^* = \mathbb{E}T_R^*$$

Thus the expected value of the ridge test statistic T_λ^* converges to that of the random effects test statistics T_R^*, and so the two tests have approximately equal power for sufficiently large n and λ.

4.1.3.5 Computing $\mathbf{tr}(K_R K_\lambda)$

Claim $\mathrm{tr}(K_R K_\lambda) = \sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda}$

Proof

We use the SVD $X = UDV^T$. We know that

$$K_R = XX^T = UDD^TU^T \tag{21}$$

$$K_\lambda = UD(D^TD + \lambda I)^{-1}D^TU^T. \tag{22}$$

Plug these identities into the trace:

$$\mathrm{tr}(K_R K_\lambda) = \mathrm{tr}\left([UDD^TU^T][UD(D^TD+\lambda I)^{-1}D^TU^T]\right) \tag{23}$$

$$= \mathrm{tr}\left(D^TU^TUDD^TU^TUD(D^TD+\lambda I)^{-1}\right) \tag{24}$$

$$= \mathrm{tr}\left(D^TDD^TD(D^TD+\lambda I)^{-1}\right) \tag{25}$$

$$= \sum_{j=1}^{p} \frac{\delta_j^2}{\delta_j+\lambda} \tag{26}$$

5 References

Abdelnour, Farras, Michael Dayan, Orrin Devinsky, Thomas Thesen, and Ashish Raj. 2018. "Functional Brain Connectivity Is Predictable from Anatomic Network's Laplacian Eigen-Structure." *NeuroImage*. Elsevier.

Ahn, Mihye, Haipeng Shen, Weili Lin, and Hongtu Zhu. 2015. "A Sparse Reduced Rank Framework for Group Analysis of Functional Neuroimaging Data." *Statistica Sinica* 25 (1). NIH Public Access: 295.

Allen, E. A., E. Damaraju, T. Eichele, L. Wu, and V. D. Calhoun. 2018. "EEG Signatures of Dynamic Functional Network Connectivity States." *Brain Topography* 31 (1): 101–16. doi:10.1007/s10548-017-0546-2.

Allen, Elena A, Eswar Damaraju, Sergey M Plis, Erik B Erhardt, Tom Eichele, and Vince D Calhoun. 2014. "Tracking Whole-Brain Connectivity Dynamics in the Resting State." *Cerebral Cortex* 24 (3). Oxford University Press: 663–76.

Basser, Peter J, James Mattiello, and Denis LeBihan. 1994a. "Estimation of the Effective Self-Diffusion Tensor from the Nmr Spin Echo." *Journal of Magnetic Resonance, Series B* 103 (3). Elsevier: 247–54.

———. 1994b. "MR Diffusion Tensor Spectroscopy and Imaging." *Biophysical Journal* 66 (1). Elsevier: 259–67.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 289–300.

Bertolino, Alessandro, Giuseppe Blasi, Valeria Latorre, Valeria Rubino, Antonio Rampino, Lorenzo Sinibaldi, Grazia Caforio, et al. 2006. "Additive Effects of Genetic Variation in Dopamine Regulating Genes on Working Memory Cortical Activity in Human Brain." *Journal of Neuroscience* 26 (15). Society for Neuroscience: 3918–22.

Bohlken, Marc M, Rachel M Brouwer, René CW Mandl, Martijn P Van den Heuvel, Anna M Hedman, Marc De Hert, Wiepke Cahn, René S Kahn, and Hilleke E Hulshoff Pol. 2016. "Structural Brain Connectivity as a Genetic Marker for Schizophrenia." *JAMA Psychiatry* 73 (1). American Medical Association: 11–19.

Bowman, F DuBois, Lijun Zhang, Gordana Derado, and Shuo Chen. 2012. "Determining Functional Connectivity Using fMRI Data with Diffusion-Based Anatomical Weighting." *NeuroImage* 62 (3). Elsevier: 1769–79.

Boyle, Evan A, Yang I Li, and Jonathan K Pritchard. 2017. "An Expanded View of Complex Traits: From Polygenic to Omnigenic." *Cell* 169 (7). Elsevier: 1177–86.

Breiman, Leo, and Jerome H Friedman. 1997. "Predicting Multivariate Responses in Multiple Linear Regression." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 59 (1). Wiley Online Library: 3–54.

Brown, Philip J, and James V Zidek. 1980. "Adaptive Multivariate Ridge Regression." *The Annals of Statistics*. JSTOR, 64–74.

Browning, Sharon R, and Elizabeth A Thompson. 2012. "Detecting Rare Variant Associations by Identity-by-Descent Mapping in Case-Control Studies." *Genetics* 190 (4). Genetics Soc America: 1521–31.

Calhoun, Vince D, and Tulay Adali. 2016. "Time-Varying Brain Connectivity in fMRI Data: Whole-Brain Data-Driven Approaches for Capturing and Characterizing Dynamic States." *IEEE Signal Processing Magazine* 33 (3). IEEE: 52–66.

Calhoun, Vince D, Robyn Miller, Godfrey Pearlson, and Tulay Adali. 2014. "The Chronnectome: Time-Varying Connectivity Networks as the Next Frontier in fMRI Data Discovery." *Neuron* 84 (2). Elsevier: 262–74.

Campbell, Trevor, Miao Liu, Brian Kulis, Jonathan P How, and Lawrence Carin. 2013. "Dynamic Clustering via Asymptotics of the Dependent Dirichlet Process Mixture." In *Advances in Neural Information Processing*

Systems, 449–57.

Campos, Gustavo de los, Ana I Vazquez, Rohan Fernando, Yann C Klimentidis, and Daniel Sorensen. 2013. "Prediction of Complex Human Traits Using the Genomic Best Linear Unbiased Predictor." *PLoS Genetics* 9 (7). Public Library of Science: e1003608.

Cassidy, Clifford M, Jared X Van Snellenberg, Caridad Benavides, Mark Slifstein, Zhishun Wang, Holly Moore, Anissa Abi-Dargham, and Guillermo Horga. 2016. "Dynamic Connectivity Between Brain Networks Supports Working Memory: Relationships to Dopamine Release and Schizophrenia." *Journal of Neuroscience* 36 (15). Soc Neuroscience: 4377–88.

Chang, Catie, and Gary H Glover. 2010. "Time–frequency Dynamics of Resting-State Brain Connectivity Measured with fMRI." *Neuroimage* 50 (1). Elsevier: 81–98.

Chau, Joris, Hernando Ombao, and Rainer von Sachs. 2017. "Data Depth and Rank-Based Tests for Covariance and Spectral Density Matrices." *arXiv Preprint arXiv:1706.08289*.

Chekouo, Thierry, Francesco C Stingo, Michele Guindani, Kim-Anh Do, and others. 2016. "A Bayesian Predictive Model for Imaging Genetics with Application to Schizophrenia." *The Annals of Applied Statistics* 10 (3). Institute of Mathematical Statistics: 1547–71.

Chiang, S., M. Guindani, H.J. Yeh, Z. Haneef, J.M. Stern, and M. Vannucci. 2017. "A Bayesian Vector Autoregressive Model for Multi-Subject Effective Connectivity Inference Using Multi-Modal Neuroimaging Data." *Human Brain Mapping* 38: 1311–32.

Chiang, Sharon, Michele Guindani, Hsiang J Yeh, Zulfi Haneef, John M Stern, and Marina Vannucci. 2017. "Bayesian Vector Autoregressive Model for Multi-Subject Effective Connectivity Inference Using Multi-Modal Neuroimaging Data." *Human Brain Mapping* 38 (3). Wiley Online Library: 1311–32.

Choe, Ann S, Mary Beth Nebel, Anita D Barber, Jessica R Cohen, Yuting Xu, James J Pekar, Brian Caffo, and Martin A Lindquist. 2017. "Comparing Test-Retest Reliability of Dynamic Functional Connectivity Methods." *Neuroimage* 158. Elsevier: 155–75.

Cohen, Jonathan D, William M Perlstein, Todd S Braver, Leigh E Nystrom, Douglas C Noll, John Jonides, and Edward E Smith. 1997. "Temporal Dynamics of Brain Activation During a Working Memory Task." *Nature* 386 (6625). Nature Publishing Group: 604.

Crobe, Alessandra, Matteo Demuru, Luca Didaci, Gian Luca Marcialis, and Matteo Fraschini. 2016. "Minimum Spanning Tree and K-Core Decomposition as Measure of Subject-Specific Eeg Traits." *Biomedical Physics & Engineering Express* 2 (1). IOP Publishing: 017001.

Cule, Erika, Paolo Vineis, and Maria De Iorio. 2011. "Significance Testing in Ridge Regression for Genetic Data." *BMC Bioinformatics* 12 (1). BioMed Central: 372.

Dahlhaus, Rainer. 2000. "A Likelihood Approximation for Locally Stationary Processes." *The Annals of Statistics* 28 (6). JSTOR: 1762–94.

Dahlhaus, Rainer, and others. 1997. "Fitting Time Series Models to Nonstationary Processes." *The Annals of Statistics* 25 (1). Institute of Mathematical Statistics: 1–37.

Dai, Zhongxiang, Joshua De Souza, Julian Lim, Paul M Ho, Yu Chen, Junhua Li, Nitish Thakor, Anastasios Bezerianos, and Yu Sun. 2017. "EEG Cortical Connectivity Analysis of Working Memory Reveals Topological Reorganization in Theta and Alpha Bands." *Frontiers in Human Neuroscience* 11. Frontiers: 237.

David, Olivier, Stefan J Kiebel, Lee M Harrison, Jérémie Mattout, James M Kilner, and Karl J Friston. 2006. "Dynamic Causal Modeling of Evoked Responses in Eeg and Meg." *NeuroImage* 30 (4). Elsevier: 1255–72.

Delaney, Nancy Jo, and Sangit Chatterjee. 1986. "Use of the Bootstrap and Cross-Validation in Ridge Regression." *Journal of Business & Economic Statistics* 4 (2). Taylor & Francis: 255–62.

Di, Xin, and Bharat B Biswal. 2015. "Dynamic Brain Functional Connectivity Modulated by Resting-State

Networks." *Brain Structure and Function* 220 (1). Springer: 37–46.

Diggle, Peter J, and Ibrahim Al Wasel. 1997. "Spectral Analysis of Replicated Biomedical Time Series." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 46 (1). Wiley Online Library: 31–71.

Edin, Fredrik, Torkel Klingberg, Pär Johansson, Fiona McNab, Jesper Tegnér, and Albert Compte. 2009. "Mechanism for Top-down Control of Working Memory Capacity." *Proceedings of the National Academy of Sciences* 106 (16). National Acad Sciences: 6802–7.

Egan, Michael F, Richard E Straub, Terry E Goldberg, Imtiaz Yakub, Joseph H Callicott, Ahmad R Hariri, Venkata S Mattay, et al. 2004. "Variation in Grm3 Affects Cognition, Prefrontal Glutamate, and Risk for Schizophrenia." *Proceedings of the National Academy of Sciences of the United States of America* 101 (34). National Acad Sciences: 12604–9.

Eklund, Anders, Thomas E Nichols, and Hans Knutsson. 2016. "Cluster Failure: Why fMRI Inferences for Spatial Extent Have Inflated False-Positive Rates." *Proceedings of the National Academy of Sciences*. National Acad Sciences, 201602413.

Elston, Robert C, Sarah Buxbaum, Kevin B Jacobs, and Jane M Olson. 2000. "Haseman and Elston Revisited." *Genetic Epidemiology* 19 (1): 1–17.

Euan, Carolina, Hernando Ombao, and Joaquin Ortega. 2016. "The Hierarchical Spectral Merger Algorithm: A New Time Series Clustering Procedure." *arXiv Preprint arXiv:1609.08569*.

Euan, Carolina, Ying Sun, and Hernando Ombao. 2017. "Coherence-Based Time Series Clustering for Brain Connectivity Visualization." *arXiv Preprint arXiv:1711.07007*.

Fiecas, Mark, and Hernando Ombao. 2011. "The Generalized Shrinkage Estimator for the Analysis of Functional Connectivity of Brain Signals." *Annals of Applied Statistics* 5: 1102–25.

———. 2016a. "Modeling the Evolution of Dynamic Brain Processes During an Associative Learning Experiment." *Journal of the American Statistical Association* 111 (516). Taylor & Francis: 1440–53.

———. 2016b. "Modeling the Evolution of Dynamic Brain Processes During an Associative Learning Experiment." *Journal of the American Statistical Association* 111: 1440–53.

Fiecas, Mark, Hernando Ombao, Dan van Lunen, Richard Baumgartner, Alexandre Coimbra, and Dai Feng. 2013. "Quantifying Temporal Correlations: A Test–retest Evaluation of Functional Connectivity in Resting-State fMRI." *NeuroImage* 65. Elsevier: 231–41.

Finn, Emily S, Xilin Shen, Dustin Scheinost, Monica D Rosenberg, Jessica Huang, Marvin M Chun, Xenophon Papademetris, and R Todd Constable. 2015. "Functional Connectome Fingerprinting: Identifying Individuals Using Patterns of Brain Connectivity." *Nature Neuroscience* 18 (11). Nature Research: 1664–71.

Friedman, Jerome, Trevor Hastie, Saharon Rosset, Robert Tibshirani, and Ji Zhu. 2004. "[Consistency in Boosting]: Discussion." *The Annals of Statistics* 32 (1). JSTOR: 102–7.

Friston, Karl J, CD Frith, PF Liddle, and RSJ Frackowiak. 1991. "Comparing Functional (Pet) Images: The Assessment of Significant Change." *Journal of Cerebral Blood Flow & Metabolism* 11 (4). SAGE Publications Sage UK: London, England: 690–99.

Fu, Zening, Shing-Chow Chan, Xin Di, Bharat Biswal, and Zhiguo Zhang. 2014. "Adaptive Covariance Estimation of Non-Stationary Processes and Its Application to Infer Dynamic Connectivity from fMRI." *IEEE Transactions on Biomedical Circuits and Systems* 8 (2). IEEE: 228–39.

Gao, Xu, Babak Shahbaba, and Hernando Ombao. 2017. "Modeling Binary Time Series Using Gaussian Processes with Application to Predicting Sleep States." *arXiv Preprint arXiv:1711.05466*.

Gao, Xu, Babak Shahbaba, Norbert Fortin, and Hernando Ombao. 2016. "Evolutionary State-Space Model and Its Application to Time-Frequency Analysis of Local Field Potentials." *arXiv Preprint arXiv:1610.07271*.

Ge, Tian, Chia-Yen Chen, Benjamin M Neale, Mert R Sabuncu, and Jordan W Smoller. 2017. "Phenome-Wide

Heritability Analysis of the Uk Biobank." *PLoS Genetics* 13 (4). Public Library of Science: e1006711.

Ge, Tian, Jianfeng Feng, Derrek P Hibar, Paul M Thompson, and Thomas E Nichols. 2012. "Increasing Power for Voxel-Wise Genome-Wide Association Studies: The Random Field Theory, Least Square Kernel Machines and Fast Permutation Procedures." *Neuroimage* 63 (2). Elsevier: 858–73.

Ge, Tian, Martin Reuter, Anderson M Winkler, Avram J Holmes, Phil H Lee, Lee S Tirrell, Joshua L Roffman, Randy L Buckner, Jordan W Smoller, and Mert R Sabuncu. 2016. "Multidimensional Heritability Analysis of Neuroanatomical Shape." *Nature Communications* 7. Nature Publishing Group: 13291.

Gelfand, Alan E, Athanasios Kottas, and Steven N MacEachern. 2005. "Bayesian Nonparametric Spatial Modeling with Dirichlet Process Mixing." *Journal of the American Statistical Association* 100 (471). Taylor & Francis: 1021–35.

Gelman, Andrew, and Iain Pardoe. 2006. "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models." *Technometrics* 48 (2). Taylor & Francis: 241–51.

Genovese, Christopher R, Nicole A Lazar, and Thomas Nichols. 2002. "Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate." *Neuroimage* 15 (4). Elsevier: 870–78.

Goeman, Jelle J, Sara A Van De Geer, and Hans C Van Houwelingen. 2006. "Testing Against a High Dimensional Alternative." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68 (3). Wiley Online Library: 477–93.

Goeman, Jelle J, Sara A Van De Geer, Floor De Kort, and Hans C Van Houwelingen. 2004. "A Global Test for Groups of Genes: Testing Association with a Clinical Outcome." *Bioinformatics* 20 (1). Oxford University Press: 93–99.

Golub, Gene H, Michael Heath, and Grace Wahba. 1979. "Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter." *Technometrics* 21 (2). Taylor & Francis: 215–23.

Gonzalez-Castillo, Javier, and Peter A Bandettini. 2017. "Task-Based Dynamic Functional Connectivity: Recent Findings and Open Questions." *Neuroimage*. Elsevier.

Gorfine, Malka, Sonja I Berndt, Jenny Chang-Claude, Michael Hoffmeister, Loic Le Marchand, John Potter, Martha L Slattery, Nir Keret, Ulrike Peters, and Li Hsu. 2017. "Heritability Estimation Using a Regularized Regression Approach (Herra): Applicable to Continuous, Dichotomous or Age-at-Onset Outcome." *PloS One* 12 (8). Public Library of Science: e0181269.

Gorrostieta, Cristina, Mark Fiecas, Hernando Ombao, Erin Burke, and Steven C. Cramer. 2013. "Hierarchical Vector Auto-Regressive Models and Their Applications to Multi-Subject Effective Connectivity." *Frontiers in Computational Neuroscience* 7 (159): 1–11.

Gorrostieta, Cristina, Hernando Ombao, and Rainer von Sachs. 2016. "Time-Dependent Dual Frequency Coherency in Multivariate Non-Stationary Time Series." Submission.

Gorrostieta, Cristina, Hernando Ombao, Patrick Bedard, and Jerome.N. Sanes. 2012. "Investigating Stimulus-Induced Changes in Connectivity Using Mixed Effects Vector Autoregressive Models." *NeuroImage* 59: 3347–55.

Gower, John C. 1966. "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis." *Biometrika* 53 (3-4). Oxford University Press: 325–38.

———. 1971. "A General Coefficient of Similarity and Some of Its Properties." *Biometrics*. JSTOR, 857–71.

Guhaniyogi, Rajarshi, Shaan Qamar, and David B. Dunson. 2016. "Bayesian Tensor Regression." *Arxiv*, no. 1509.06490.

Harris, Sarah E, Helen Fox, Alan F Wright, Caroline Hayward, John M Starr, Lawrence J Whalley, and Ian J Deary. 2007. "A Genetic Association Analysis of Cognitive Ability and Cognitive Ageing Using 325 Markers

for 109 Genes Associated with Oxidative Stress or Cognition." *BMC Genetics* 8 (1). BioMed Central: 43.

Harush, Uzi, and Baruch Barzel. 2017. "Dynamic Patterns of Information Flow in Complex Networks." *Nature Communications* 8 (1). Nature Publishing Group: 2181.

Havlicek, Martin, Jiri Jan, Milan Brazdil, and Vince D Calhoun. 2010. "Dynamic Granger Causality Based on Kalman Filter for Evaluation of Functional Network Connectivity in fMRI Data." *Neuroimage* 53 (1). Elsevier: 65–77.

Hindriks, Rikkert, Mohit H Adhikari, Yusuke Murayama, Marco Ganzetti, Dante Mantini, Nikos K Logothetis, and Gustavo Deco. 2016. "Can Sliding-Window Correlations Reveal Dynamic Functional Connectivity in Resting-State fMRI?" *Neuroimage* 127. Elsevier: 242–56.

Hodges, James S. 1998. "Some Algebra and Geometry for Hierarchical Models, Applied to Diagnostics." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 60 (3). Wiley Online Library: 497–536.

Hoerl, Arthur E, Robert W Kannard, and Kent F Baldwin. 1975. "Ridge Regression: Some Simulations." *Communications in Statistics-Theory and Methods* 4 (2). Taylor & Francis: 105–23.

Hooper, John W. 1959. "Simultaneous Equations and Canonical Correlation Theory." *Econometrica: Journal of the Econometric Society*. JSTOR, 245–56.

Hua, Wen-Yu, and Debashis Ghosh. 2015. "Equivalence of Kernel Machine Regression and Kernel Distance Covariance for Multidimensional Phenotype Association Studies." *Biometrics* 71 (3). Wiley Online Library: 812–20.

Hua, Wen-Yu, Thomas E Nichols, Debashis Ghosh, and Alzheimer's Disease Neuroimaging Initiative. 2014. "Multiple Comparison Procedures for Neuroimaging Genomewide Association Studies." *Biostatistics* 16 (1). Oxford University Press: 17–30.

Huang, Meiyan, Thomas Nichols, Chao Huang, Yang Yu, Zhaohua Lu, Rebecca C Knickmeyer, Qianjin Feng, Hongtu Zhu, Alzheimer's Disease Neuroimaging Initiative, and others. 2015. "FVGWAS: Fast Voxelwise Genome Wide Association Analysis of Large-Scale Imaging Genetic Data." *Neuroimage* 118. Elsevier: 613–27.

Huh, Myung-Hoe, and Myoungshic Jhun. 2001. "Random Permutation Testing in Multiple Linear Regression." *Communications in Statistics-Theory and Methods* 30 (10). Taylor & Francis: 2023–32.

Hung, Hung, and Chen-Chien Wang. 2012. "Matrix Variate Logistic Regression Model with Application to Eeg Data." *Biostatistics* 14 (1). Oxford University Press: 189–202.

Hutchison, R Matthew, Thilo Womelsdorf, Elena A Allen, Peter A Bandettini, Vince D Calhoun, Maurizio Corbetta, Stefania Della Penna, et al. 2013. "Dynamic Functional Connectivity: Promise, Issues, and Interpretations." *Neuroimage* 80. Elsevier: 360–78.

Jiang, Zheng-yan. 2005. "Study on Eeg Power and Coherence in Patients with Mild Cognitive Impairment During Working Memory Task." *Journal of Zhejiang University-Science B* 6 (12). Springer: 1213–9.

Kang, Hakmook, Hernando Ombao, Christopher Fonnesbeck, Zhaohua Ding, and Victoria L Morgan. 2017. "A Bayesian Double Fusion Model for Resting-State Brain Connectivity Using Joint Functional and Structural Data." *Brain Connectivity* 7. Mary Ann Liebert, Inc. r: 219–27.

Kherif, Ferath, Jean-Baptiste Poline, Sébastien Mériaux, Habib Benali, Guillaume Flandin, and Matthew Brett. 2003. "Group Analysis in Functional Neuroimaging: Selecting Subjects Using Similarity Measures." *NeuroImage* 20 (4). Elsevier: 2197–2208.

Kim, Hyunwoo J, Nagesh Adluru, Heemanshu Suri, Baba C Vemuri, Sterling C Johnson, and Vikas Singh. 2017. "Riemannian Nonlinear Mixed Effects Models: Analyzing Longitudinal Deformations in Neuroimaging." In *Proceedings of Ieee Conference on Computer Vision and Pattern Recognition (Cvpr)*.

Kim, Junghi, Jeffrey R Wozniak, Bryon A Mueller, and Wei Pan. 2015. "Testing Group Differences in Brain Functional Connectivity: Using Correlations or Partial Correlations?" *Brain Connectivity* 5 (4). Mary Ann

Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: 214–31.

Kirch, Claudia, Birte Muhsal, and Hernando Ombao. 2015a. "Detection of Changes in Multivariate Time Series with Application to Eeg Data." *Journal of the American Statistical Association* 110: 1197–1216.

———. 2015b. "Detection of Changes in Multivariate Time Series with Application to Eeg Data." *Journal of the American Statistical Association* 110 (511). Taylor & Francis: 1197–1216.

Kitzbichler, Manfred G, Richard NA Henson, Marie L Smith, Pradeep J Nathan, and Edward T Bullmore. 2011. "Cognitive Effort Drives Workspace Configuration of Human Brain Functional Networks." *Journal of Neuroscience* 31 (22). Soc Neuroscience: 8259–70.

Klingberg, Torkel, Brendan T O'Sullivan, and Per E Roland. 1997. "Bilateral Activation of Fronto-Parietal Networks by Incrementing Demand in a Working Memory Task." *Cerebral Cortex (New York, NY: 1991)* 7 (5): 465–71.

Kwee, Lydia Coulter, Dawei Liu, Xihong Lin, Debashis Ghosh, and Michael P Epstein. 2008. "A Powerful and Flexible Multilocus Association Test for Quantitative Traits." *The American Journal of Human Genetics* 82 (2). Elsevier: 386–97.

Laird, Nan M, and James H Ware. 1982. "Random-Effects Models for Longitudinal Data." *Biometrics*. JSTOR, 963–74.

Lan, Shiwei, Andrew Holbrook, Norbert J Fortin, Ombao Hernando, and Babak Shahbaba. 2017. "Flexible Bayesian Dynamic Modeling of Covariance and Correlation Matrices." *arXiv Preprint arXiv:1711.02869*.

Langer, Nicolas, Claudia C von Bastian, Helen Wirz, Klaus Oberauer, and Lutz Jäncke. 2013. "The Effects of Working Memory Training on Functional Brain Network Efficiency." *Cortex* 49 (9). Elsevier: 2424–38.

Lasky-Su, Jessica. 2017. "Chapter 19 - Statistical Techniques for Genetic Analysis." In *Clinical and Translational Science (Second Edition)*, edited by David Robertson and Gordon H. Williams, Second Edition, 347–62. Academic Press.

Lee, Jason D, Dennis L Sun, Yuekai Sun, Jonathan E Taylor, and others. 2016. "Exact Post-Selection Inference, with Application to the Lasso." *The Annals of Statistics* 44 (3). Institute of Mathematical Statistics: 907–27.

Leonardi, Nora, and Dimitri Van De Ville. 2015. "On Spurious and Real Fluctuations of Dynamic Functional Connectivity During Rest." *Neuroimage* 104. Elsevier: 430–36.

Li, Hao, Sally Wetten, Li Li, Pamela L St Jean, Ruchi Upmanyu, Linda Surh, David Hosford, et al. 2008. "Candidate Single-Nucleotide Polymorphisms from a Genomewide Association Study of Alzheimer Disease." *Archives of Neurology* 65 (1). American Medical Association: 45–53.

Li, Hongzhe. 2012. "U-Statistics in Genetic Association Studies." *Human Genetics* 131 (9). Springer: 1395–1401.

Li, Hongzhe, Zhi Wei, and John Maris. 2009. "A Hidden Markov Random Field Model for Genome-Wide Association Studies." *Biostatistics* 11 (1). Oxford University Press: 139–50.

Li, Zeran, Jorge L. Del Aguila, Umber Dube, John Budde, Rita Martinez, Kathlee Black, Qingli Xiao, et al. 2018. "Genetic Variants Associated with Alzheimers Disease Confer Different Cerebral Cortex Cell-Type Population Structure." *bioRxiv*. Cold Spring Harbor Laboratory. doi:10.1101/266296.

Lin, Dahua, Eric Grimson, and John W Fisher. 2010. "Construction of Dependent Dirichlet Processes Based on Poisson Processes." In *Advances in Neural Information Processing Systems*, 1396–1404.

Linden, David EJ, Robert A Bittner, Lars Muckli, James A Waltz, Nikolaus Kriegeskorte, Rainer Goebel, Wolf Singer, and Matthias HJ Munk. 2003. "Cortical Capacity Constraints for Visual Working Memory:

Dissociation of fMRI Load Effects in a Fronto-Parietal Network." *Neuroimage* 20 (3). Elsevier: 1518–30.

Lindquist, Martin. 2008. "The Statistical Analysis of fMRI Data." *Statistical Science*. JSTOR, 439–64.

Lindquist, Martin A, Yuting Xu, Mary Beth Nebel, and Brain S Caffo. 2014. "Evaluating Dynamic Bivariate Correlations in Resting-State fMRI: A Comparison Study and a New Approach." *Neuroimage* 101. Elsevier: 531–46.

Liu, Dawei, Xihong Lin, and Debashis Ghosh. 2007. "Semiparametric Regression of Multidimensional Genetic Pathway Data: Least-Squares Kernel Machines and Linear Mixed Models." *Biometrics* 63 (4). Wiley Online Library: 1079–88.

Liu, Jingyu, Godfrey Pearlson, Andreas Windemuth, Gualberto Ruano, Nora I Perrone-Bizzozero, and Vince Calhoun. 2009. "Combining fMRI and Snp Data to Investigate Connections Between Brain Function and Genetics Using Parallel Ica." *Human Brain Mapping* 30 (1). Wiley Online Library: 241–55.

Lu, Zhao-Hua, Zakaria Khondker, Joseph G Ibrahim, Yue Wang, Hongtu Zhu, Alzheimer's Disease Neuroimaging Initiative, and others. 2017. "Bayesian Longitudinal Low-Rank Regression Models for Imaging Genetic Data from Longitudinal Studies." *NeuroImage* 149. Elsevier: 305–22.

Madhyastha, Tara M, Mary K Askren, Peter Boord, and Thomas J Grabowski. 2015. "Dynamic Connectivity at Rest Predicts Attention Task Performance." *Brain Connectivity* 5 (1). Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: 45–59.

Mahyari, Arash Golibagh, David M Zoltowski, Edward M Bernat, and Selin Aviyente. 2017. "A Tensor Decomposition-Based Approach for Detecting Dynamic Network States from Eeg." *IEEE Transactions on Biomedical Engineering* 64 (1). IEEE: 225–37.

Maity, Arnab, and Xihong Lin. 2011. "Powerful Tests for Detecting a Gene Effect in the Presence of Possible Gene–Gene Interactions Using Garrote Kernel Machines." *Biometrics* 67 (4). Wiley Online Library: 1271–84.

Maity, Arnab, Patrick F Sullivan, and Jun-ing Tzeng. 2012. "Multivariate Phenotype Association Analysis by Marker-Set Kernel Machine Regression." *Genetic Epidemiology* 36 (7). Wiley Online Library: 686–95.

Maldonado, Yolanda Munoz. 2009. "Mixed Models, Posterior Means and Penalized Least-Squares." *Lecture Notes-Monograph Series*. JSTOR, 216–36.

Mantel, Nathan. 1967a. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27 (2 Part 1). AACR: 209–20.

———. 1967b. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27 (2 Part 1). AACR: 209–20.

Martins-Filho, Carlos, and Feng Yao. 2006. "A Note on the Use of V and U Statistics in Nonparametric Models of Regression." *Annals of the Institute of Statistical Mathematics* 58 (2). Springer: 389–406.

McIntosh, Anthony Randal, and Nancy J Lobaugh. 2004. "Partial Least Squares Analysis of Neuroimaging Data: Applications and Advances." *Neuroimage* 23. Elsevier: S250–S263.

Meda, Shashwath A, Balaji Narayanan, Jingyu Liu, Nora I Perrone-Bizzozero, Michael C Stevens, Vince D Calhoun, David C Glahn, et al. 2012. "A Large Scale Multivariate Parallel Ica Method Reveals Novel Imaging–genetic Relationships for Alzheimer's Disease in the Adni Cohort." *Neuroimage* 60 (3). Elsevier: 1608–21.

Murphy, Kevin P. 1998. "Switching Kalman Filters." Citeseer.

Nadarajah, Saralees, and Samuel Kotz. 2008. "Exact Distribution of the Max/Min of Two Gaussian Random Variables." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 16 (2). IEEE: 210–12.

Nathoo, Farouk S, Linglong Kong, and Hongtu Zhu. 2017. "A Review of Statistical Methods in Imaging Genetics." *arXiv Preprint arXiv:1707.07332*.

Nguyen, Vinh T, Michael Breakspear, and Ross Cunnington. 2014. "Fusing Concurrent Eeg–fMRI with

Dynamic Causal Modeling: Application to Effective Connectivity During Face Perception." *Neuroimage* 102. Elsevier: 60–70.

Nichols, Thomas E. 2012. "Multiple Testing Corrections, Nonparametric Methods, and Random Field Theory." *Neuroimage* 62 (2). Elsevier: 811–15.

Nichols, Thomas E, and Andrew P Holmes. 2002. "Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples." *Human Brain Mapping* 15 (1). Wiley Online Library: 1–25.

Nielsen, Søren FV, Mikkel N Schmidt, Kristoffer H Madsen, and Morten Mørup. 2017. "Predictive Assessment of Models for Dynamic Functional Connectivity." *NeuroImage*. Elsevier.

Niethammer, Martin, Chris C Tang, Peter A LeWitt, Ali R Rezai, Maureen A Leehey, Steven G Ojemann, Alice W Flaherty, et al. 2017. "Long-Term Follow-up of a Randomized Aav2-Gad Gene Therapy Trial for Parkinson's Disease." *JCI Insight* 2 (7). American Society for Clinical Investigation.

Olson, Chester L. 1976. "On Choosing a Test Statistic in Multivariate Analysis of Variance." *Psychological Bulletin* 83 (4). American Psychological Association: 579.

Ombao, Hernando, and Sebastien Van Bellegem. 2008. "Coherence Analysis: A Linear Filtering Point of View." *IEEE Transactions on Signal Processing* 56 (6): 2259–66.

Ombao, Hernando, Mark Fiecas, Chee-Ming Ting, and Yin Fen Low. 2017. "Statistical Models for Brain Signals with Properties That Evolve Across Trials." *NeuroImage*. Elsevier.

Ombao, Hernando, Martin Lindquist, Wesley Thompson, and John Aston. 2016. *Handbook of Statistical Methods for Neuroimaging*. CRC Press.

Omelka, Marek, and Šárka Hudecová. 2013. "A Comparison of the Mantel Test with a Generalised Distance Covariance Test." *Environmetrics* 24 (7). Wiley Online Library: 449–60.

Onton, Julie, Arnaud Delorme, and Scott Makeig. 2005. "Frontal Midline Eeg Dynamics During Working Memory." *Neuroimage* 27 (2). Elsevier: 341–56.

Palfi, Stéphane, Jean Marc Gurruchaga, G Scott Ralph, Helene Lepetit, Sonia Lavisse, Philip C Buttery, Colin Watts, et al. 2014. "Long-Term Safety and Tolerability of Prosavin, a Lentiviral Vector-Based Gene Therapy for Parkinson's Disease: A Dose Escalation, Open-Label, Phase 1/2 Trial." *The Lancet* 383 (9923). Elsevier: 1138–46.

Palva, J Matias, Simo Monto, Shrikanth Kulashekhar, and Satu Palva. 2010. "Neuronal Synchrony Reveals Working Memory Networks and Predicts Individual Memory Capacity." *Proceedings of the National Academy of Sciences* 107 (16). National Acad Sciences: 7580–5.

Palva, Satu, Simo Monto, and J Matias Palva. 2010. "Graph Properties of Synchronized Cortical Networks During Visual Working Memory Maintenance." *Neuroimage* 49 (4). Elsevier: 3257–68.

Pan, Wei. 2011. "Relationship Between Genomic Distance-Based Regression and Kernel Machine Regression for Multi-Marker Association Testing." *Genetic Epidemiology* 35 (4). Wiley Online Library: 211–16.

Park, Timothy, Idris A Eckley, and Hernando C Ombao. 2014. "Estimating Time-Evolving Partial Coherence Between Signals via Multivariate Locally Stationary Wavelet Processes." *IEEE Transactions on Signal Processing* 62 (20). IEEE: 5240–50.

Pavlov, Yuri G, and Boris Kotchoubey. 2017. "EEG Correlates of Working Memory Performance in Females." *BMC Neuroscience* 18 (1). BioMed Central: 26.

Pillai, KCS. 1955. "Some New Test Criteria in Multivariate Analysis." *The Annals of Mathematical Statistics*. JSTOR, 117–21.

Pluta, Dustin, Tong Shen, Gui Xue, Chuansheng Chen, Zhaoxia Yu, and Hernando Ombao. 2017. "Adaptive Mantel Test for Penalized Inference, with Applications to Imaging Genetics." *UC Irvine Statistics Research*

Paper.

Rashid, Barnaly, Eswar Damaraju, Godfrey D Pearlson, and Vince D Calhoun. 2014. "Dynamic Connectivity States Estimated from Resting fMRI Identify Differences Among Schizophrenia, Bipolar Disorder, and Healthy Control Subjects." *Frontiers in Human Neuroscience* 8. Frontiers: 897.

Raz, Hui, Jonathanand Zheng, Hernando Ombao, and Bruce Turetsky. 2003. "Statistical Test for fMRI Based on Experimental Randomization." *NeuroImage* 19: 226–32.

Reiss, Philip T, M Henry H Stevens, Zarrar Shehzad, Eva Petkova, and Michael P Milham. 2010. "On Distance-Based Permutation Tests for Between-Group Comparisons." *Biometrics* 66 (2). Wiley Online Library: 636–43.

Robert, Paul, and Yves Escoufier. 1976. "A Unifying Tool for Linear Multivariate Statistical Methods: The Rv-Coefficient." *Applied Statistics*. JSTOR, 257–65.

Robinson, George K. 1991. "That Blup Is a Good Thing: The Estimation of Random Effects." *Statistical Science*. JSTOR, 15–32.

Rogers, Jeffrey, Peter Kochunov, Jack Lancaster, Wendy Shelledy, David Glahn, John Blangero, and Peter Fox. 2007. "Heritability of Brain Volume, Surface Area and Shape: An Mri Study in an Extended Pedigree of Baboons." *Human Brain Mapping* 28 (6). Wiley Online Library: 576–83.

Salimi-Khorshidi, Gholamreza, Stephen M Smith, John R Keltner, Tor D Wager, and Thomas E Nichols. 2009. "Meta-Analysis of Neuroimaging Data: A Comparison of Image-Based and Coordinate-Based Pooling of Studies." *Neuroimage* 45 (3). Elsevier: 810–23.

Sambataro, Fabio, Eleonora Visintin, Nadja Doerig, Janis Brakowski, Martin Grosse Holtforth, Erich Seifritz, and Simona Spinelli. 2017. "Altered Dynamics of Brain Connectivity in Major Depressive Disorder at-Rest and During Task Performance." *Psychiatry Research: Neuroimaging* 259. Elsevier: 1–9.

Samdin, S Balqis, Chee-Ming Ting, Hernando Ombao, and Sh-Hussain Salleh. 2017. "A Unified Estimation Framework for State-Related Changes in Effective Brain Connectivity." *IEEE Transactions on Biomedical Engineering* 64 (4). IEEE: 844–58.

Sarnthein, J, Hellmuth Petsche, P Rappelsberger, GL Shaw, and A Von Stein. 1998. "Synchronization Between Prefrontal and Posterior Association Cortex During Human Working Memory." *Proceedings of the National Academy of Sciences* 95 (12). National Acad Sciences: 7092–6.

Sauseng, Paul, Wolfgang Klimesch, Manuel Schabus, and Michael Doppelmayr. 2005. "Fronto-Parietal Eeg Coherence in Theta and Upper Alpha Reflect Central Executive Functions of Working Memory." *International Journal of Psychophysiology* 57 (2). Elsevier: 97–103.

Schaid, Daniel J. 2010a. "Genomic Similarity and Kernel Methods I: Advancements by Building on Mathematical and Statistical Foundations." *Human Heredity* 70 (2). Karger Publishers: 109–31.

———. 2010b. "Genomic Similarity and Kernel Methods Ii: Methods for Genomic Information." *Human Heredity* 70 (2). Karger Publishers: 132–40.

Schwartzman, Armin. 2006. "Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data." PhD thesis, Stanford University.

Sejdinovic, Dino, Bharath Sriperumbudur, Arthur Gretton, and Kenji Fukumizu. 2013. "Equivalence of Distance-Based and Rkhs-Based Statistics in Hypothesis Testing." *The Annals of Statistics*. JSTOR, 2263–91.

Shehzad, Zarrar, Clare Kelly, Philip T Reiss, R Cameron Craddock, John W Emerson, Katie McMahon, David A Copland, F Xavier Castellanos, and Michael P Milham. 2014. "A Multivariate Distance-Based Analytic Framework for Connectome-Wide Association Studies." *Neuroimage* 93. Elsevier: 74–94.

Shen, Kai-Kai, Stephen Rose, Jurgen Fripp, Katie L McMahon, Greig I de Zubicaray, Nicholas G Martin, Paul M Thompson, Margaret J Wright, and Olivier Salvado. 2014. "Investigating Brain Connectivity Heritability

in a Twin Study Using Diffusion Imaging Data." *NeuroImage* 100. Elsevier: 628–41.

Shu, Hai, Bin Nan, and Robert Koeppe. 2015. "Multiple Testing for Neuroimaging via Hidden Markov Random Field." *Biometrics* 71 (3). Wiley Online Library: 741–50.

Simons, Jon S, and Hugo J Spiers. 2003. "Prefrontal and Medial Temporal Lobe Interactions in Long-Term Memory." *Nature Reviews Neuroscience* 4 (8): 637–48.

Smilde, Age K, Henk AL Kiers, S Bijlsma, CM Rubingh, and MJ Van Erk. 2008. "Matrix Correlations for High-Dimensional Data: The Modified Rv-Coefficient." *Bioinformatics* 25 (3). Oxford University Press: 401–5.

Sokal, Robert R. 1979. "Society of Systematic Biologists Testing Statistical Significance of Geographic Variation Patterns." *Source: Systematic Zoology* 28 (2): 227–32. http://www.jstor.org/stable/2412528 http://about.jstor.org/terms.

Sporns, Olaf, and Rolf Kötter. 2004. "Motifs in Brain Networks." *PLoS Biology* 2 (11). Public Library of Science: e369.

Stein, Jason L, Xue Hua, Suh Lee, April J Ho, Alex D Leow, Arthur W Toga, Andrew J Saykin, et al. 2010. "Voxelwise Genome-Wide Association Study (vGWAS)." *Neuroimage* 53 (3). Elsevier: 1160–74.

Stingo, Francesco C, Michele Guindani, Marina Vannucci, and Vince D Calhoun. 2013. "An Integrative Bayesian Modeling Approach to Imaging Genetics." *Journal of the American Statistical Association* 108 (503). Taylor & Francis: 876–91.

Sudre, Gustavo, Saadia Choudhuri, Eszter Szekely, Teighlor Bonner, Elanda Goduni, Wendy Sharp, and Philip Shaw. 2017. "Estimating the Heritability of Structural and Functional Brain Connectivity in Families Affected by Attention-Deficit/Hyperactivity Disorder." *JAMA Psychiatry* 74 (1). American Medical Association: 76–84.

Sun, Felicia, L Miller, and Mark D'Esposito. 2004. "Measuring Interregional Functional Connectivity Using Coherence and Partial Coherence Analyses of fMRI Data." *NeuroImage* 21: 647–58.

Sun, Wenguang, Brian J Reich, T Tony Cai, Michele Guindani, and Armin Schwartzman. 2015. "False Discovery Control in Large-Scale Spatial Multiple Testing." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 77 (1). Wiley Online Library: 59–83.

Székely, Gábor J, Maria L Rizzo, Nail K Bakirov, and others. 2007. "Measuring and Testing Dependence by Correlation of Distances." *The Annals of Statistics* 35 (6). Institute of Mathematical Statistics: 2769–94.

Taghia, Jalil, Srikanth Ryali, Tianwen Chen, Kaustubh Supekar, Weidong Cai, and Vinod Menon. 2017. "Bayesian Switching Factor Analysis for Estimating Time-Varying Functional Connectivity in fMRI." *NeuroImage* 155. Elsevier: 271–90.

Tan, Hui-Ru, Chee-Ming Ting, Sh-Hussain Salleh, I Kamarulafizam, and AM Noor. 2016. "Shrinkage Estimation of High-Dimensional Vector Autoregressions for Effective Connectivity in fMRI." In *Biomedical Engineering and Sciences (Iecbes), 2016 Ieee Embs Conference on*, 121–26. IEEE.

Tang, Wei, Hesheng Liu, Linda Douw, Mark A Kramer, Uri T Eden, Matti S Hämäläinen, and Steven M Stufflebeam. 2017. "Dynamic Connectivity Modulates Local Activity in the Core Regions of the Default-Mode Network." *Proceedings of the National Academy of Sciences* 114 (36). National Acad Sciences: 9713–8.

Thompson, Paul M, Tian Ge, David C Glahn, Neda Jahanshad, and Thomas E Nichols. 2013. "Genetics of the Connectome." *Neuroimage* 80. Elsevier: 475–88.

Thompson, William Hedley, and Peter Fransson. 2017. "A Common Framework for the Problem of Deriving Estimates of Dynamic Functional Brain Connectivity." *NeuroImage*. Elsevier.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*. JSTOR, 267–88.

Ting, Chee-Ming, Hernando Ombao, and Sh-Hussain Salleh. 2017. "Multi-Scale Factor Analysis of High-

Dimensional Brain Signals." *arXiv Preprint arXiv:1705.06502*.

Ting, Chee-Ming, Hernando Ombao, S Balqis Samdin, and Sh-Hussain Salleh. 2017. "Estimating Dynamic Connectivity States in fMRI Using Regime-Switching Factor Models." *IEEE Transactions on Medical Imaging*. IEEE.

Tzeng, Jung-Ying, Daowen Zhang, Sheng-Mao Chang, Duncan C Thomas, and Marie Davidian. 2009. "Gene-Trait Similarity Regression for Multimarker-Based Association Analysis." *Biometrics* 65 (3). Wiley Online Library: 822–32.

Tzeng, Jung-Ying, Daowen Zhang, Monnat Pongpanich, Chris Smith, Mark I McCarthy, Michele M Sale, Bradford B Worrall, Fang-Chi Hsu, Duncan C Thomas, and Patrick F Sullivan. 2011. "Studying Gene and Gene-Environment Effects of Uncommon and Common Variants on Continuous Traits: A Marker-Set Approach Using Gene-Trait Similarity Regression." *The American Journal of Human Genetics* 89 (2). Elsevier: 277–88.

Uludağ, Kâmil, and Alard Roebroeck. 2014. "General Overview on the Merits of Multimodal Neuroimaging Data Fusion." *Neuroimage* 102. Elsevier: 3–10.

Vergara, Victor M, Alvaro Ulloa, Vince D Calhoun, David Boutte, Jiayu Chen, and Jingyu Liu. 2014. "A Three-Way Parallel Ica Approach to Analyze Links Among Genetics, Brain Structure and Brain Function." *Neuroimage* 98. Elsevier: 386–94.

Vidaurre, Diego, Stephen M Smith, and Mark W Woolrich. 2017. "Brain Network Dynamics Are Hierarchically Organized in Time." *Proceedings of the National Academy of Sciences* 114 (48). National Acad Sciences: 12827–32.

Visscher, Peter M, Gibran Hemani, Anna AE Vinkhuyzen, Guo-Bo Chen, Sang Hong Lee, Naomi R Wray, Michael E Goddard, and Jian Yang. 2014. "Statistical Power to Detect Genetic (Co) Variance of Complex Traits Using Snp Data in Unrelated Samples." *PLoS Genetics* 10 (4). Public Library of Science: e1004269.

Vogler, C, L Gschwind, D Coynel, V Freytag, A Milnik, T Egli, A Heck, D Jf De Quervain, and A Papassotiropoulos. 2014. "Substantial Snp-Based Heritability Estimates for Working Memory Performance." *Translational Psychiatry* 4 (9). Nature Publishing Group: e438.

Vounou, Maria, Thomas E Nichols, Giovanni Montana, Alzheimer's Disease Neuroimaging Initiative, and others. 2010. "Discovering Genetic Associations with High-Dimensional Neuroimaging Phenotypes: A Sparse Reduced-Rank Regression Approach." *Neuroimage* 53 (3). Elsevier: 1147–59.

Wager, Tor D, Martin Lindquist, and Lauren Kaplan. 2007. "Meta-Analysis of Functional Neuroimaging Data: Current and Future Directions." *Social Cognitive and Affective Neuroscience* 2 (2). Oxford University Press: 150–58.

Wang, Xuefeng, Seunggeun Lee, Xiaofeng Zhu, Susan Redline, and Xihong Lin. 2013. "GEE-Based Snp Set Association Test for Continuous and Discrete Traits in Family-Based Association Studies." *Genetic Epidemiology* 37 (8). Wiley Online Library: 778–86.

Warnick, Ryan, Michele Guindani, Erik Erhardt, Elena Allen, Vince Calhoun, and Marina Vannucci. 2017. "A Bayesian Approach for Estimating Dynamic Functional Network Connectivity in fMRI Data." *Journal of the American Statistical Association* – (just-accepted). Taylor & Francis.

Wei, Zhi, Mingyao Li, Timothy Rebbeck, and Hongzhe Li. 2008. "U-Statistics-Based Tests for Multiple Genes in Genetic Association Studies." *Annals of Human Genetics* 72 (6). Wiley Online Library: 821–33.

Wessel, Jennifer, and Nicholas J Schork. 2006. "Generalized Genomic Distance–based Regression Methodology for Multilocus Association Analysis." *The American Journal of Human Genetics* 79 (5). Elsevier: 792–806.

"What Are Single Nucleotide Polymorphisms (Snps)? - Genetics Home Reference." n.d. *U.S. National Library*

of Medicine. National Institutes of Health. https://ghr.nlm.nih.gov/primer/genomicresearch/snp.

Wieringen, Wessel N van. 2015. "Lecture Notes on Ridge Regression." *arXiv Preprint arXiv:1509.09169*.

Winkler, Anderson M, Gerard R Ridgway, Matthew A Webster, Stephen M Smith, and Thomas E Nichols. 2014. "Permutation Inference for the General Linear Model." *Neuroimage* 92. Elsevier: 381–97.

Woo, Choong-Wan, Anjali Krishnan, and Tor D Wager. 2014. "Cluster-Extent Based Thresholding in fMRI Analyses: Pitfalls and Recommendations." *Neuroimage* 91. Elsevier: 412–19.

Woodward, Neil D, and Carissa J Cascio. 2015. "Resting-State Functional Connectivity in Psychiatric Disorders." *JAMA Psychiatry* 72 (8). American Medical Association: 743–44.

Worsley, Keith J, Alan C Evans, S Marrett, and P Neelin. 1992. "A Three-Dimensional Statistical Analysis for Cbf Activation Studies in Human Brain." *Journal of Cerebral Blood Flow & Metabolism* 12 (6). SAGE Publications Sage UK: London, England: 900–918.

Worsley, Keith J, CH Liao, J Aston, V Petre, GH Duncan, F Morales, and AC Evans. 2002. "A General Statistical Analysis for fMRI Data." *Neuroimage* 15 (1). Elsevier: 1–15.

Wu, Michael C, Peter Kraft, Michael P Epstein, Deanne M Taylor, Stephen J Chanock, David J Hunter, and Xihong Lin. 2010. "Powerful Snp-Set Analysis for Case-Control Genome-Wide Association Studies." *The American Journal of Human Genetics* 86 (6). Elsevier: 929–42.

Wu, Michael C, Seunggeun Lee, Tianxi Cai, Yun Li, Michael Boehnke, and Xihong Lin. 2011. "Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test." *The American Journal of Human Genetics* 89 (1). Elsevier: 82–93.

Xu, Gongjun, Lifeng Lin, Peng Wei, and Wei Pan. 2016. "An Adaptive Two-Sample Test for High-Dimensional Means." *Biometrika* 103 (3). Oxford University Press: 609–24.

Xu, Junhai, Xuntao Yin, Haitao Ge, Yan Han, Zengchang Pang, Baolin Liu, Shuwei Liu, and Karl Friston. 2016. "Heritability of the Effective Connectivity in the Resting-State Default Mode Network." *Cerebral Cortex*, 1–9.

Xu, Zhiyuan, Chong Wu, Wei Pan, Alzheimer's Disease Neuroimaging Initiative, and others. 2017. "Imaging-Wide Association Study: Integrating Imaging Endophenotypes in Gwas." *NeuroImage*. Elsevier.

Xu, Zhiyuan, Gongjun Xu, and Wei Pan. 2017. "Adaptive Testing for Association Between Two Random Vectors in Moderate to High Dimensions." *Genetic Epidemiology*. Wiley Online Library.

Xue, Gui, Zhonglin Lu, Irwin P Levin, Joshua A Weller, Xiangrui Li, and Antoine Bechara. 2008. "Functional Dissociations of Risk and Reward Processing in the Medial Prefrontal Cortex." *Cerebral Cortex* 19 (5). Oxford University Press: 1019–27.

Xue, Wenqiong, F DuBois Bowman, Anthony V Pileggi, and Andrew R Mayer. 2015. "A Multimodal Approach for Determining Brain Networks by Jointly Modeling Functional and Structural Connectivity." *Frontiers in Computational Neuroscience* 9. Frontiers Media SA.

Yaesoubi, Maziar, Robyn L Miller, and Vince D Calhoun. 2017. "Time-Varying Spectral Power of Resting-State fMRI Networks Reveal Cross-Frequency Dependence in Dynamic Connectivity." *PloS One* 12 (2). Public Library of Science: e0171647.

Yang, Jian, Beben Benyamin, Brian P McEvoy, Scott Gordon, Anjali K Henders, Dale R Nyholt, Pamela A Madden, et al. 2010. "Common Snps Explain a Large Proportion of the Heritability for Human Height." *Nature Genetics* 42 (7). Nature Research: 565–69.

Yang, Jian, S Hong Lee, Michael E Goddard, and Peter M Visscher. 2011. "GCTA: A Tool for Genome-Wide Complex Trait Analysis." *The American Journal of Human Genetics* 88 (1). Elsevier: 76–82.

Yang, Ying, Elissa Aminoff, Michael Tarr, and Kass E Robert. 2016. "A State-Space Model of Cross-Region Dynamic Connectivity in Meg/Eeg." In *Advances in Neural Information Processing Systems 29*, edited by D.

D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, 1234–42. Curran Associates, Inc. http://papers.nips.cc/paper/6593-a-state-space-model-of-cross-region-dynamic-connectivity-in-megeeg.pdf.

Yu, Qingbao, Erik B Erhardt, Jing Sui, Yuhui Du, Hao He, Devon Hjelm, Mustafa S Cetin, et al. 2015. "Assessing Dynamic Brain Graphs of Time-Varying Connectivity in fMRI Data: Application to Healthy Controls and Patients with Schizophrenia." *Neuroimage* 107. Elsevier: 345–55.

Yu, Qingbao, Lei Wu, David A Bridwell, Erik B Erhardt, Yuhui Du, Hao He, Jiayu Chen, et al. 2016. "Building an Eeg-fMRI Multi-Modal Brain Graph: A Concurrent Eeg-fMRI Study." *Frontiers in Human Neuroscience* 10. Frontiers: 476.

Yu, Zhe, Raquel Prado, Erin Burke, Steven C. Cramer, and Hernando Ombao. 2016. "A Hierarchical Bayesian Model for Studying the Impact of Stroke on Brain Motor Function." *Journal of the American Statistical Association* 111: 549–63.

Zalesky, Andrew, Alex Fornito, Luca Cocchi, Leonardo L Gollo, and Michael Breakspear. 2014. "Time-Resolved Resting-State Brain Networks." *Proceedings of the National Academy of Sciences* 111 (28). National Acad Sciences: 10341–6.

Zapala, Matthew A, and Nicholas J Schork. 2012. "Statistical Properties of Multivariate Distance Matrix Regression for High-Dimensional Data Analysis." *Frontiers in Genetics* 3. Frontiers Media SA.

Zhang, Daowen, and Xihong Lin. 2003. "Hypothesis Testing in Semiparametric Additive Mixed Models." *Biostatistics (Oxford)* 4 (1): 57–74.

Zhang, Tingting, Minh Pham, Jianhui Sun, Guofen Yan, Huazhang Li, Yinge Sun, Marlen Z Gonzalez, and James A Coan. 2017. "A Low-Rank Multivariate General Linear Model for Multi-Subject fMRI Data and a Non-Convex Optimization Algorithm for Brain Response Comparison." *NeuroImage*. Elsevier.

Zhang, Ying, Alessandro Bertolino, Leonardo Fazio, Giuseppe Blasi, Antonio Rampino, Raffaella Romano, Mei-Ling T Lee, et al. 2007. "Polymorphisms in Human Dopamine D2 Receptor Gene Affect Gene Expression, Splicing, and Neuronal Activity During Working Memory." *Proceedings of the National Academy of Sciences* 104 (51). National Acad Sciences: 20552–7.

Zhou, Hua, Lexin Li, and Hongtu Zhu. 2013. "Tensor Regression with Applications in Neuroimaging Data Analysis." *Journal of the American Statistical Association* 108 (502). Taylor & Francis Group: 540–52.

Zhu, Dajiang, Tuo Zhang, Xi Jiang, Xintao Hu, Hanbo Chen, Ning Yang, Jinglei Lv, Junwei Han, Lei Guo, and Tianming Liu. 2014. "Fusing Dti and fMRI Data: A Survey of Methods and Applications." *NeuroImage* 102. Elsevier: 184–91.

Zhu, Hongtu, Yasheng Chen, Joseph G Ibrahim, Yimei Li, Colin Hall, and Weili Lin. 2009. "Intrinsic Regression Models for Positive-Definite Matrices with Applications to Diffusion Tensor Imaging." *Journal of the American Statistical Association* 104 (487). Taylor & Francis: 1203–12.

Zhu, Hongtu, Linglong Kong, Runze Li, Martin Styner, Guido Gerig, Weili Lin, and John H Gilmore. 2011. "FADTTS: Functional Analysis of Diffusion Tensor Tract Statistics." *NeuroImage* 56 (3). Elsevier: 1412–25.