High-dimensional Inference for Imaging Genetics with the Adaptive Mantel Test

Dustin Pluta

28 Mar 2018

Acknowledgements

- **Gui Xue**, PI and data provider, Beijing Normal University, Center for Brain and Learning Sciences

- **Chuansheng Chen**, UCI, Dept. of Psychology and Social Behavior

- **Hernando Ombao**, KAUST, Dept. of Statistics

- **Zhaoxia Yu**, UCI, Dept. of Statistics

- **Tong Shen**, UCI, Dept. of Statistics (PhD Student)

Overview of Talk

1. The scientific setting of **connectome genetics**.

2. **Mantel test** and distance-based association testing.

3. The **adaptive Mantel test** for penalized inference.

4. Future Work: **Manifold mixed effects model** and **methods for dynamic connectivity**.
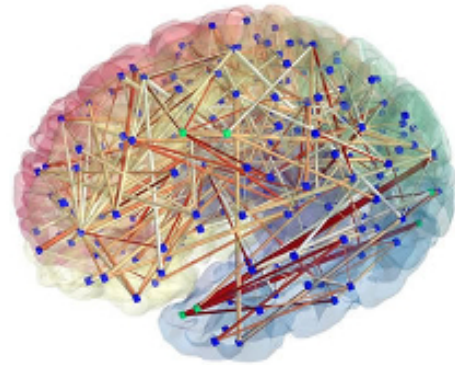

Links

- **Adaptive Mantel Test Paper:** arxiv.org/pdf/1712.07270.pdf

- **Slides available:** github.com/dspluta/Presentations/

- **Adaptive Mantel R Package:** github.com/dspluta/adamant

Scientific Background of Connectome Genetics

Scientific Background

> **Connectome**
>
> The total set of physical and functional relationships between all brain regions.

Types of Connectivity

Functional: Measures the degree of correspondence between activity in two regions, typically quantified by Pearson's correlation or coherence.

Effective: Measures the predictability of activity in one region given the previous activity in another region.

Structural: Measures the strength/quantity of physical connections between brain regions.

Functional Connectivity: Coherence
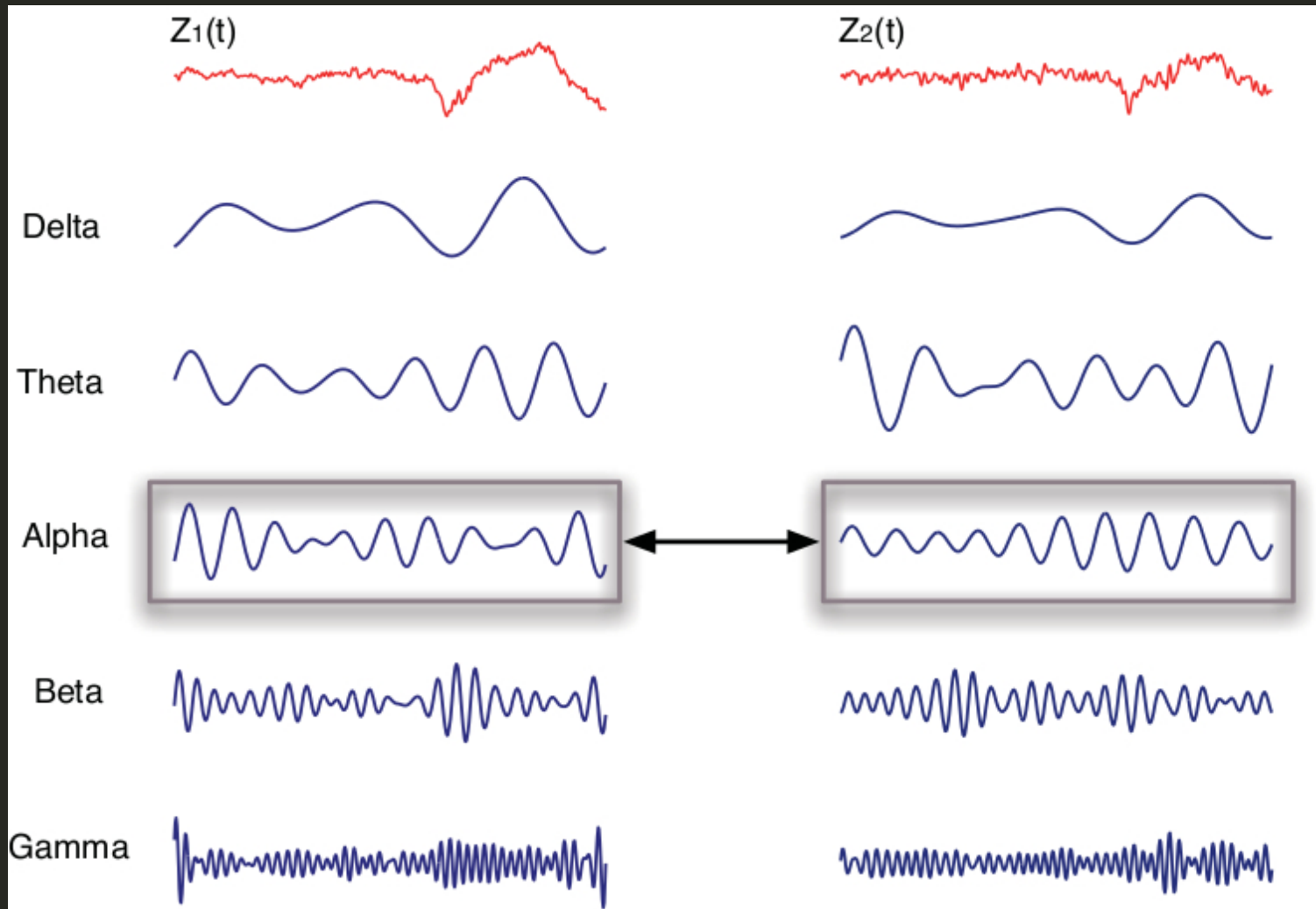
Scientific Background

> **Genome**
>
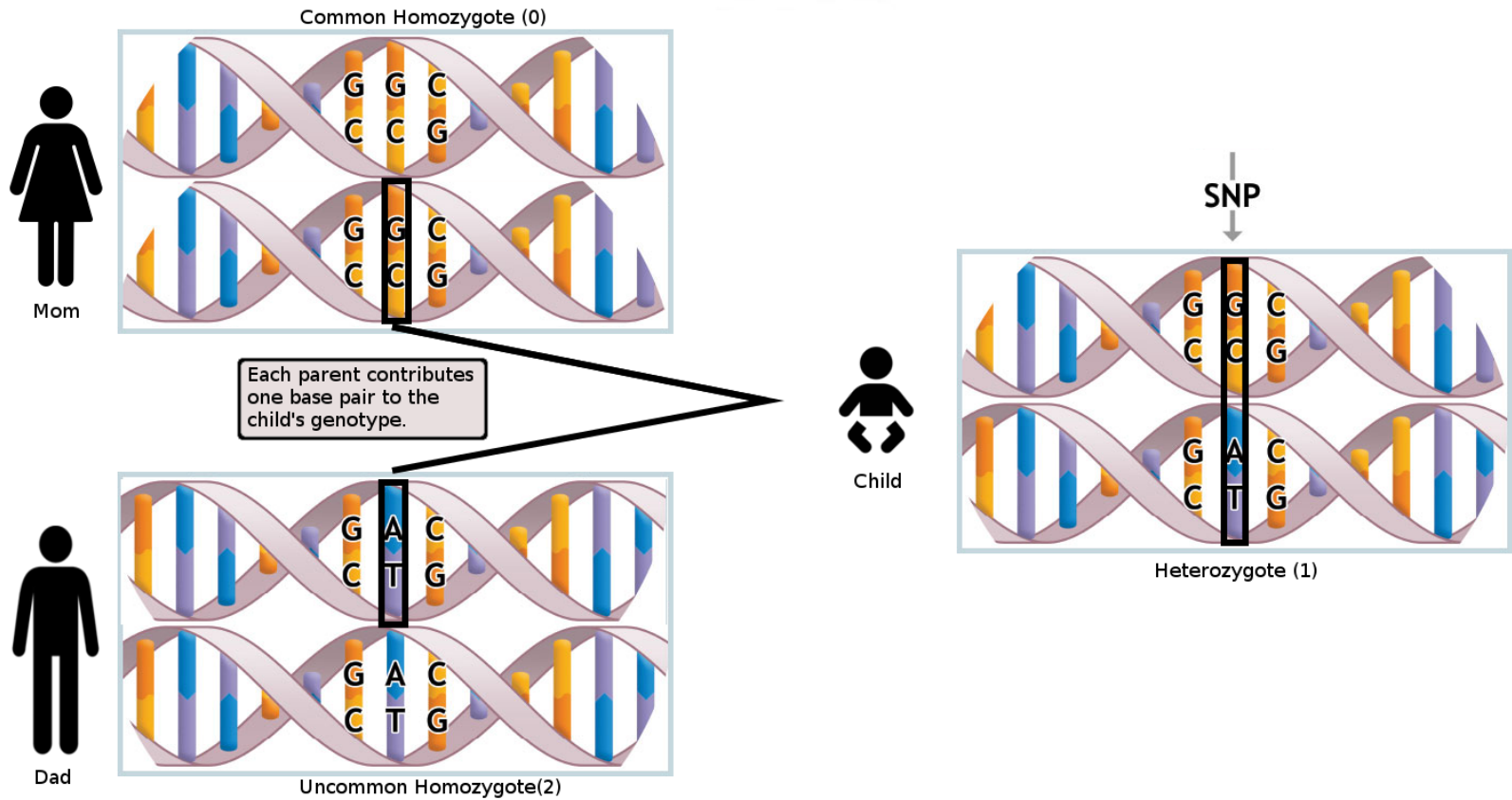> The set of all genes and genetic material in an organism.

Some Types of Genetic Data

Single Nucleotide Polymorphisms (SNPs) Are locations in the DNA that exhibit particularly high variation across individuals (in humans).

Gene Expression Data Measures the strength of expression of a particular gene.

Copy Number Variations Large DNA segments that vary in copy number relative to a reference genome.

Common Homozygote (0)

Mom

Each parent contributes one base pair to the child's genotype.

SNP

Child

Heterozygote (1)

Dad

Uncommon Homozygote(2)

Behavior/ Disease Status

Memory Tests, Decision-making Alzheimer's Disease, Schizophrenia

Genotype

Single Nucleotide Polymorphisms (SNPs)

Neuroimaging Phenotype

sMRI, fMRI, EEG, DTI, Functional Connectivity, Coherence

Scientific Background

> **Heritability** of a phenotypic trait refers to the amount of variation of the trait that is explained by genetic effects.

- **Broad-sense heritability** (H^2) is defined as the phenotypic variation explained by *all possible* genetic effects.

- **Narrow-sense heritability** (h^2) refers to phenotypic variation explained just by *additive* genetic effects only.

Scientific Background

Estimating Heritability with Variance Components Model

- A popular approach to estimate narrow-sense heritability of a phenotype is the *variance components* (*random effects*) model.

$$Y = Xb + \varepsilon,$$

with

- $\mathrm{Var}(Y) = \sigma_b^2 X X^T + \sigma_\varepsilon^2 I_n,$

- $b \sim N(0, \sigma_b^2 I_p)$ is a random vector of SNP effects,

- $\varepsilon \sim N(0, \sigma_\varepsilon^2 I_n)$ residual vector,

- X is the SNP data matrix (column centered and scaled).

Scientific Setting

Estimating Heritability with Variance Components Model

- The above random effects model can be rewritten as

$$Y = g + \varepsilon,$$

where $g \sim N(0, \sigma_g^2 G)$, for $G = XX^T/p$ and $\sigma_g^2 = p\sigma_b^2$.

Narrow-sense heritability of the phenotype measured by Y can then be estimated as

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

$$\hat{h}^2 = \frac{\text{tr}\left(\hat{\Sigma}_g\right)}{\text{tr}\left(\hat{\Sigma}_g\right) + \text{tr}\left(\hat{\Sigma}_\varepsilon\right)},$$

where $\hat{\Sigma}_g = \hat{\sigma}_g^2 G$, $\hat{\Sigma}_\varepsilon = \hat{\sigma}_\varepsilon^2 I_n$.

Mantel Test and Distance-based Association Testing

Distance-based Association Testing

> **The Inference Goal**
>
> Given observations of n subjects across two data modalities **X** and **Y**, is distance (or similarity) in **X** significantly associated with distance (or similarity) in **Y**?

Setup

- In our application, $X \in \mathrm{X}^n$ is an $n \times p$ matrix of SNP measurements, and $Y \in \mathrm{Y}$ is an $n \times 1$ vector of scalar phenotype measurements.

- Assume X and Y have been column centered and scaled.

Association Testing Methods

- **Mantel's test** (Mantel 1967) uses the inner product of the pairwise distance/similarity matrices from X and Y.

- The **RV coefficient** (Escoufier 1976) uses a test statistic based on the multivariate correlation between X and Y.

- The **distance covariance** (dCov) test (Szekely, Rizzo, Bakirov, 2007) is defined as the covariance of distances between X and Y.

- **Adaptive sum of powered score test** (Xu et. al 2017).

Challenges

- These approaches suffer from **low power** in high-dimensions,

- Generally **lack a rigorous methodology** for chooosing the appropriate metrics and testing parameters and interpreting results,

- Powered alternatives can be difficult to determine.

Mantel Test

- Given **similarity functions** $\mathcal{K}_X : \mathbb{R}^P \times \mathbb{R}^P \to \mathbb{R}$ and $\mathcal{K}_Y : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, we can form two $n \times n$ **Gram matrices** K and H, where

$$K_{ij} = \mathcal{K}_X(X_i, X_j)$$

$$H_{ij} = \mathcal{K}_Y(Y_i, Y_j).$$

- The **correlation** of these distance matrices is

$$r(H, K) := \frac{\langle K, H \rangle}{\|K\| \cdot \|H\|},$$

Mantel Test

How should we test the significance of the correlation?

Mantel's original approach (1967) is to **permute** rows and columns of one of the pairwise distance matrices to generate the reference distribution.

That is, for test statistic

$$T = \langle K, H \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} K_{ij} H_{ij} = \text{tr}(KH),$$

we compute the **permutation P-value** by permuting H to approximate the reference distribution.

GRM

$$\begin{matrix} K_{11} & K_{12} & \ldots & K_{1N} \\ K_{21} & K_{22} & \ldots & K_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ K_{N1} & K_{N2} & \ldots & K_{NN} \end{matrix}$$

Connectivity Similarity

$$\begin{matrix} H_{11} & H_{12} & \ldots & H_{1N} \\ H_{21} & H_{22} & \ldots & H_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ H_{N1} & H_{N2} & \ldots & H_{NN} \end{matrix}$$
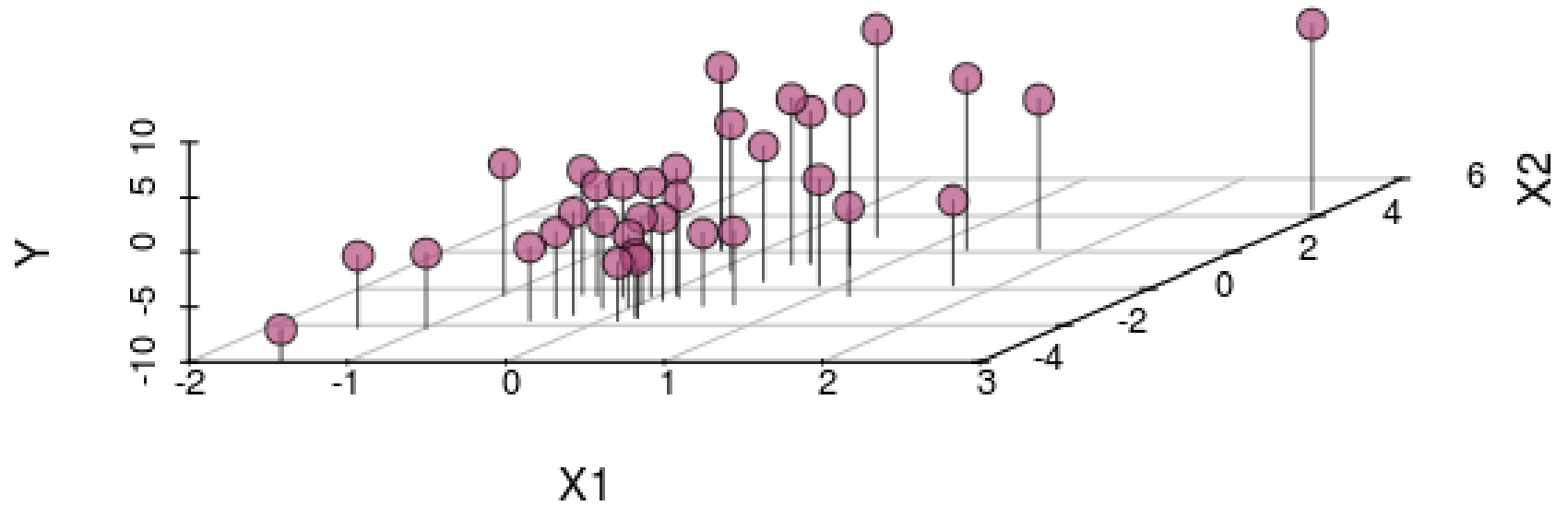
$$\begin{matrix} K_{11} \\ K_{12} \\ \vdots \\ K_{N(N-1)} \\ K_{NN} \end{matrix}$$
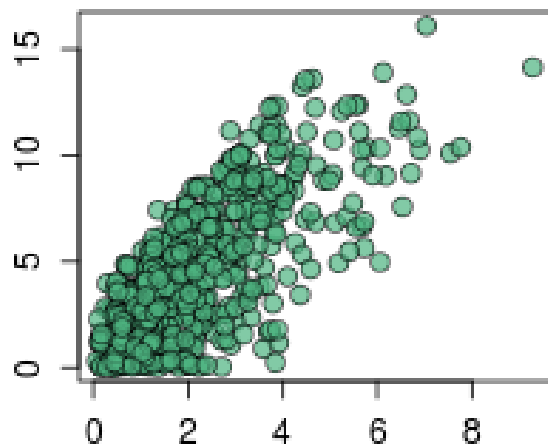
$$\begin{matrix} H_{11} \\ H_{12} \\ \vdots \\ H_{N(N-1)} \\ H_{NN} \end{matrix}$$
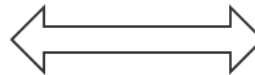
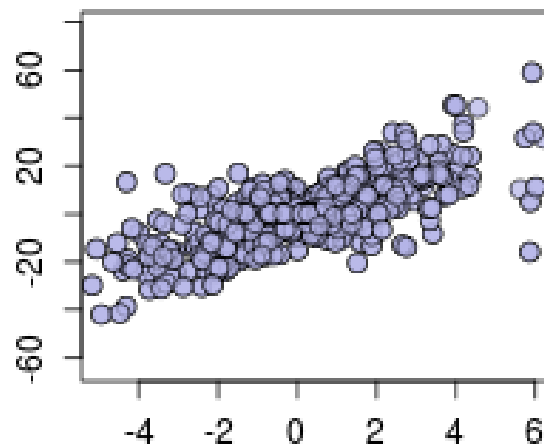$$Z^* = \langle K, H \rangle \equiv \sum_i \sum_j K_{ij} H_{ij}$$
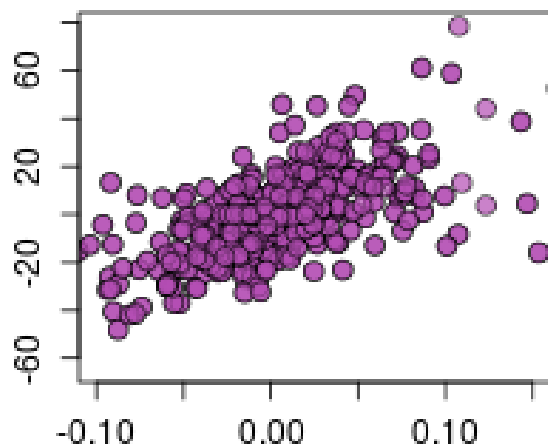
Simulated Data
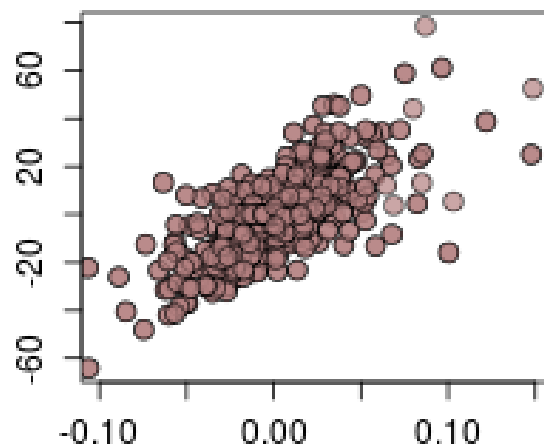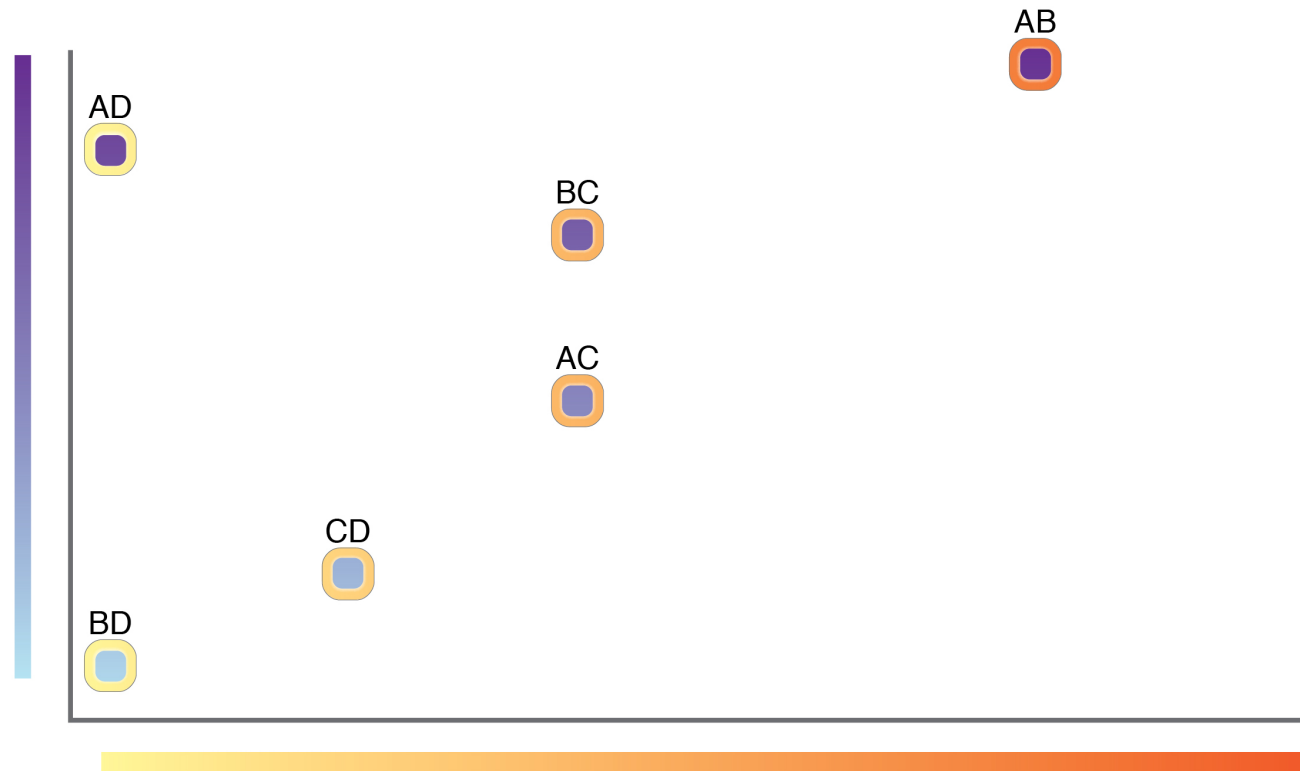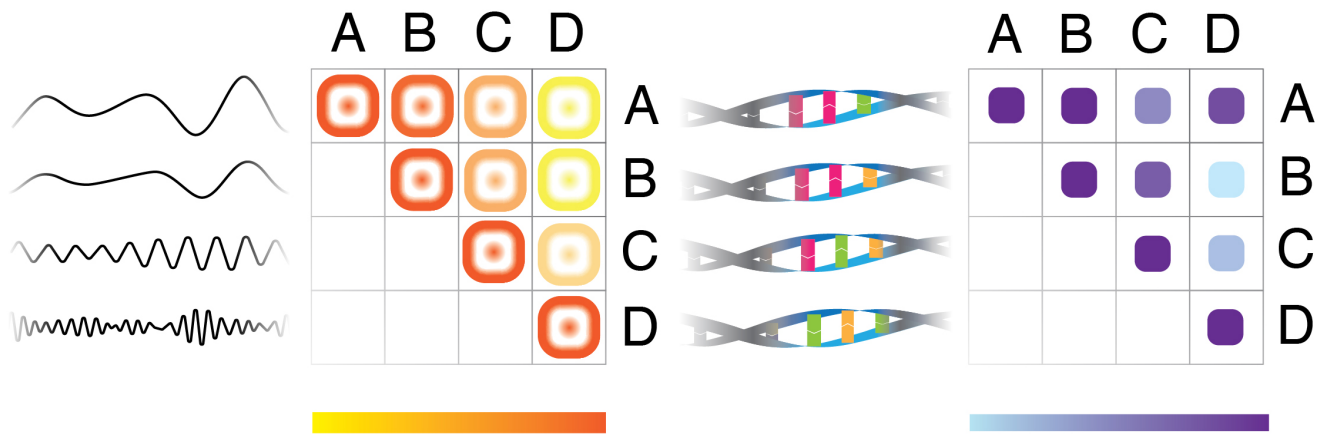
Mantel Test

Similarity with Weighted Inner Products

For two vectors $u, v \in \mathbb{R}^p$, the **weighted inner product** $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ for some positive semi-definite matrix $\mathcal{W}$, is defined as

$$\langle u, v \rangle_{\mathcal{W}} = u \mathcal{W} v^T.$$

The **Mantel Test Statistic** for similarity $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ is

$$T_{\mathcal{W}} = \mathrm{tr}(X \mathcal{W} X^T Y Y^T) = Y^T X \mathcal{W} X^T Y.$$

Mantel Test

Similarity with Weighted Inner Products

For two vectors $u, v \in \mathbb{R}^p$, the **weighted inner product** $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ for some positive semi-definite matrix $\mathcal{W}$, is defined as

$$\langle u, v \rangle_{\mathcal{W}} = u \mathcal{W} v^T.$$

The **Mantel Test Statistic** for similarity $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ is

$$T_{\mathcal{W}} = \text{tr}(X \mathcal{W} X^T Y Y^T) = Y^T X \mathcal{W} X^T Y.$$

Weight Matrices

Metric	Gram Matrix	Mantel Stat.
Euclidean	$K_R = XX^T$	$T_R = \text{tr}(K_R H)$
Mahalanobis	$K_F = X(X^T X)^{-1} X^T$	$T_F = \text{tr}(K_F H)$
Ridge Kernel	$K_\lambda = X(X^T X + \lambda I)^{-1} X^T$	$T_\lambda = \text{tr}(K_\lambda H)$

Mantel Test

Contributions

Derive testing properties of the ridge kernel in the Mantel test.

Link the random effects, fixed effects, and ridge regression score tests through the kernel Mantel test framework.

Develop the adaptive Mantel test for simultaneous testing across a range of tuning parameters.

Advantages of Proposed Methods

- Provides an extremely flexible and computationally practical framework for testing a wide variety of relationships between different modalities.

- Simulations show the kernel Mantel test is often more powerful than competing methods.

Mantel Test

Correlation of Similarities

Assume $\text{rank}(X) = r$ with singular value decomposition $X = U_{n \times r} D_{r \times r} V_{p \times r}^T$, where $\eta_j, j = 1, \cdots, r$ are the squared singular values. Let $H = YY^T$ and $Z = U^T Y$.

Euclidean Metric
$$r(H, K_R) = \frac{\sum_{j=1}^{r} \eta_j z_j^2}{\sqrt{\sum_{j=1}^{r} \eta_j^2 \sum_{i=1}^{n} y_i^2}},$$

Mahalanobis Metric
$$r(H, K_F) = \frac{\sum_{j=1}^{r} z_j^2}{\sqrt{p} \sum_{i=1}^{n} y_i^2},$$

Ridge Similarity
$$r(H, K_\lambda) = \frac{\sum_{j=1}^{r} \frac{\eta_j}{\lambda + \eta_j} z_j^2}{\sqrt{\sum_{j=1}^{r} \left(\frac{\eta_j}{\eta_j + \lambda} \right)^2 \sum_{i=1}^{n} y_i^2}}.$$

Mantel Test

Correlation of Similarities

Assume $\text{rank}(X) = r$ with singular value decomposition $X = U_{n \times r} D_{r \times r} V_{p \times r}^T$, where $\eta_j, j = 1, \cdots, r$ are the squared singular values. Let $H = YY^T$ and $Z = U^T Y$.

Euclidean Metric $\qquad r(H, K_R) \asymp \sum_{j=1}^r \eta_j z_j^2 \quad = \text{tr}(HK_R) = T_R,$

Mahalanobis Metric $\qquad r(H, K_F) \asymp \sum_{j=1}^r z_j^2 \quad = \text{tr}(HK_F) = T_F,$

Ridge Similarity $\qquad r(H, K_\lambda) \asymp \sum_{j=1}^r \frac{\eta_j}{\lambda + \eta_j} z_j^2 \; = \text{tr}(HK_\lambda) = T_\lambda.$

Mantel Test

Linear Model Definitions

Model Name	Definition	
Random Eff.	$Y \sim N(0, \sigma_b^2 G + \sigma_\varepsilon^2 I_N),$	$G = XX^T/p$
Fixed Eff.	$Y \sim N(X\beta, \sigma_\varepsilon^2 I_N)$	
Ridge Reg.	$Y \sim N(X\beta, \sigma_\varepsilon^2 I_N),$	$\|\beta\|_2^2 < c(\lambda)$

Linear Model Score Tests

Model	Score Stat.	Equivalent Stat.	Null Distribution
Random Eff.	$S_F = Z^T D (D^T D)^{-1} D^T Z$	$T_F = \operatorname{tr}(K_F H)$	$c\chi_p^2$
Fixed Eff.	$S_R = Z^T D D^T Z$	$T_R = \operatorname{tr}(K_R H)$	$\sum_{j=1}^{r} \eta_j \chi_1^2$
Ridge Reg.	$S_\lambda = Z^T D (D^T D + \lambda)^{-1} D^T Z$	$T_\lambda = \operatorname{tr}(K_\lambda H)$	$\sum_{j=1}^{r} \dfrac{\eta_j}{\lambda + \eta_j} \chi_1^2$

Mantel Test

Limiting Relationship

From the previous results, we get the following limiting relationships between the ridge test, and tests for the fixed effects and random effects models.

$$T_{\lambda=0} = T_F$$

$$T_\lambda \asymp \left\{ \lambda \sum_{j=1}^{r} \frac{\eta_j}{\lambda + \eta_j} z_j^2 \right\} \overset{\lambda \to \infty}{\to} T_R$$

Similarly, for the matrix correlations

$$r(H, K_{\lambda=0}) = r(H, K_F)$$

$$\lim_{\lambda \to \infty} r(H, K_\lambda) = r(H, K_R)$$

Mantel Test

Proportion of Variance Explained

$$R^2(X, Y) = \sqrt{\bar{p}} \cdot r(H, K_F)$$

For large n and assuming that $\text{rank}(X) = p$,

$$\hat{h}^2_{MOM} = \frac{tr(HG) - n}{tr(G^2) - n} \approx p\sqrt{\frac{tr(H^2)}{tr(K_R^2)}} r(H, K_R) \in [r(H, K_R), \sqrt{\bar{p}} \cdot r(H, K_R)]$$

Linear Model Score Tests

Geometric Interpretation

Consider $Z = U^T Y$, as the projection of Y into the column space of X.

1. The *Random Effects* model tests the **weighted Euclidean norm** of Z, where the jth component is weighted by the jth eigenvalue η_j.

2. The *Fixed Effects* model tests the **Euclidean norm** of Z

3. The *Ridge Penalization* **weights the Euclidean norm of Z proportional to the eigenvalues**, but these weights are now **flattened** by a factor of $(\lambda + \eta_j)^{-1}$.

Multivariate Mantel

- For the case of **multivariate response**, i.e. Y is an $n \times q$ response matrix, we can define the similarity matrices

$$K = X(X^T X)^{-1} X^T$$

$$H = Y(Y^T Y)^{-1} Y^T.$$

- When both X and Y are high dimensional it may make sense to use penalties for both, giving Gram matrices

$$H_{\lambda_y} = Y(Y^T Y + \lambda_y I_q)^{-1} Y^T \,,$$

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

- We can also use the L_2 similarity for both:

$$H = YY^T$$

$$K = XX^T$$

Multivariate Mantel

Hooper's trace correlation (Hooper 1959) is defined as

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

The relationship between Hooper's trace correlation and $r(H_F, K_F)$ follows from $\text{tr}(H_F) = q, \text{tr}(K_F) = p$:

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

Multivariate Mantel

Multivariate Heritability

To connect this to a multivariate h^2, we can use moment matching:

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

When both X and Y are column-standardized and full column rank it can be shown that

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

Multivariate Mantel

Interpretation of Tuning Parameters

In general, a large tuning parameter **reduces the adjustment for the correlation between features**, and so approaches the use of **Euclidean distance** in that modality as the parameter increases.

Limiting Relationships

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

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

Penalized Likelihood

$$-\ell \propto c_1 log|\Sigma_e| + tr[(Y - XB)\Sigma_e^{-1}(Y - XB)^T] + \lambda_x tr[B\Sigma_e^{-1}B^T] + \lambda_y tr[\Sigma_e^{-1}] + c_2$$

Adaptive Mantel Test

Adaptive Mantel Test

Choosing a good penalty term for inference can be difficult, since we must control the type I error.

Heuristic to choose λ

- The best linear unbiased predictors for the regression coefficients in the random effects model result from $\lambda = \frac{\sigma_\varepsilon^2}{\sigma_b^2}$ as a ridge penalty term.

- Since h^2 can be calculated from this noise to signal ratio, we can choose a "reasonable" λ by:

    1. Propose a reasonable value for h^2 from the application context.
    2. Solve for $\lambda = \frac{p(1-h^2)}{h^2}$.

- In practice, understanding h^2 as "variance in Y explained by X" can help in identifying a range of reasonable values for h^2

Adaptive Mantel Test

Idea

To simultaneously test a set of tuning parameters, use the **minimum** *P*-**value** across all parameters as the test statistic, and approximate the reference distribution using permutations.

Algorithm

- **Input:**

    - $X, n \times p$ covariates, column centered and scaled
    - $Y, n \times 1$ response, centered and scaled
    - $\left\{ \left(\mathcal{K}_m^{\mathrm{X}}, \mathcal{K}_m^{\mathrm{Y}} \right) \right\}, m = 1, \cdots, M$

- **Output:** P_{ADA} = adaptive Mantel *P*-value for global test of significant association

Adaptive Mantel Test

Algorithm

1: **for** $m = 1, \ldots, M$ **do**

2: $\quad K_m \leftarrow \delta_m^{\mathbf{X}}(X)$

3: $\quad H_m \leftarrow \delta_m^{\mathbf{Y}}(Y)$

4: $\quad$ Calculate $Z_m^{(0)} \leftarrow Z_m := \mathrm{tr}(K_m H_m)$

5: **end for**

6: Generate B permutations of H_m, labeled $H_m^{(b)} \quad \forall \, m = 1, \ldots, M; b = 1, \ldots, B.$

7: $Z_m^{(b)} \leftarrow \mathrm{tr}(K_m H_m^{(b)}) \quad \forall \, m = 1, \ldots, M; b = 1, \ldots, B$

8: $P_m^{(b)} \leftarrow \frac{1}{B+1} \sum_{b=0}^{B} I\left(Z_m^{(b)} \leq Z_m^{(b')}\right) \quad \forall \, m = 1, \ldots, M; b = 1, \ldots, B$

9: $P^{(b)} \leftarrow \min_{m=1,\cdots,M} P_m^{(b)} \quad \forall \, b = 1, \ldots, B$

10: $P_{AMT} \leftarrow \frac{1}{B+1} \sum_{b=0}^{B} I\left(P^{(0)} \leq P^{(b)}\right)$

AdaMant Package

Minimal package available at github.com/dspluta/adamant

- Install with:

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

- Then load just like any other package:

```
library(tidyverse) # Required for adamant
library(adamant)
```

AdaMant Package

```
set.seed(1234)
X <- matrix(rnorm(500), nrow = 50, ncol = 10)
Y <- X %*% rep(c(0, 0.6), 5) + rnorm(50, 0, 2)

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

```
## --------------------------------
## Adaptive Mantel Output
## --------------------------------
## P_val         = 0.0055
## n             = 50
## p             = 10
## rank(X^TX)    = 10
## kappa         = 1.982
## Best Lambda   = 10
## n_perms       = 2000
## time          = 0.092 secs
## --------------------------------

## [1] 0.0055
```

0% Sparsity
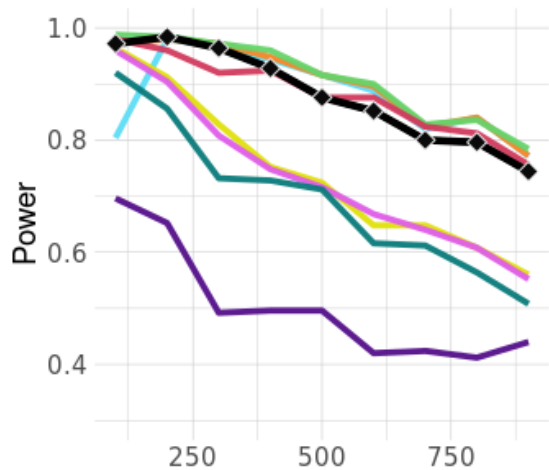
25% Sparsity

50% Sparsity

75% Sparsity

Legend:
- P_ADA
- P_lambda_1
- P_lambda_10
- P_lambda_100
- P_lambda_1000
- P_lambda_Inf
- P_RV
- P_dCov
- P_aSPC.P
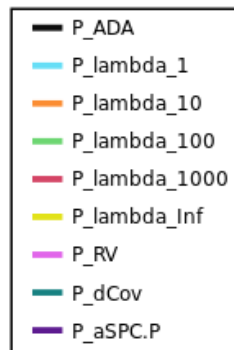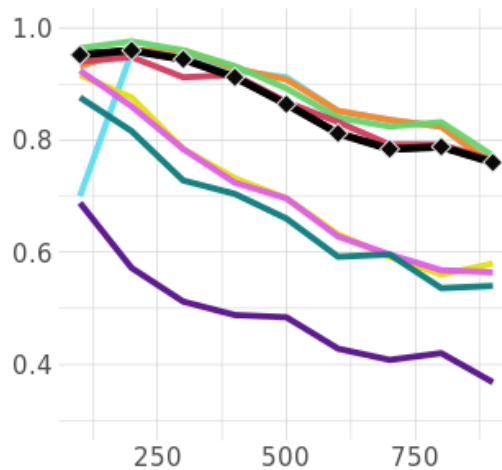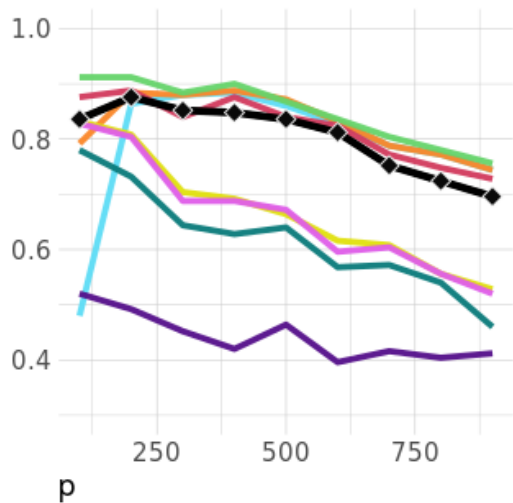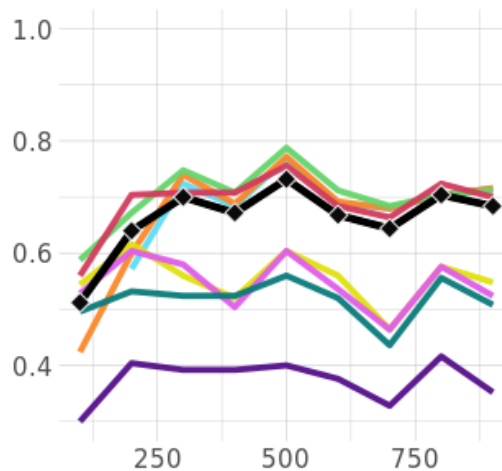
Simulation Settings

n = 100
sigma = 0.125
Cov(X) = CSYM(0.05)
# perms = 1000

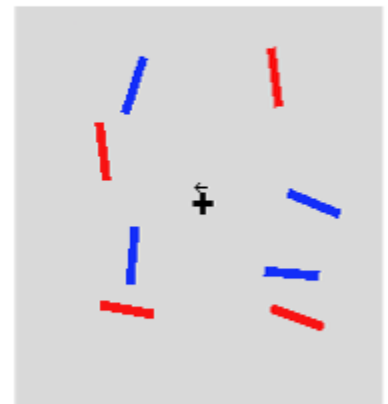# Applications of the Adaptive Mantel Test

Visual Working Memory Experiment
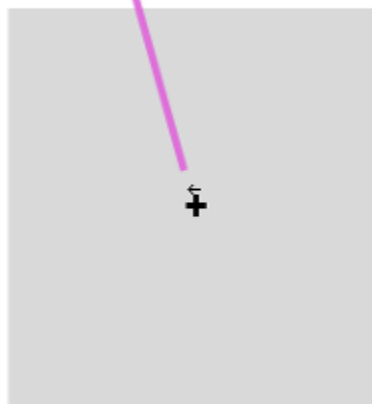

Fixation
(200ms)
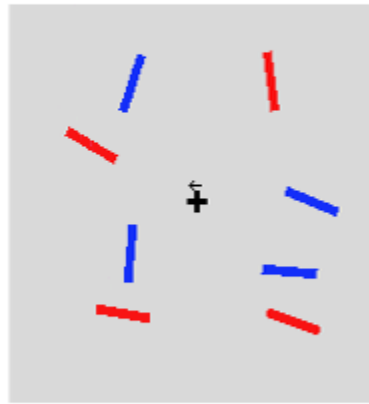
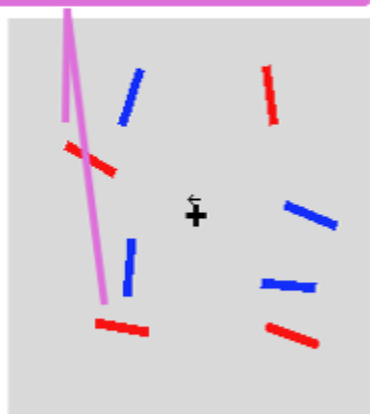**Encoding**
(100ms)

Maintenace
(900ms)

Comparison
(<2000ms)

Visual Working Memory Experiment

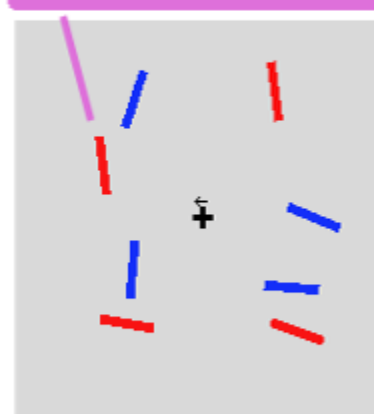Arrow Indicates Side to Remember

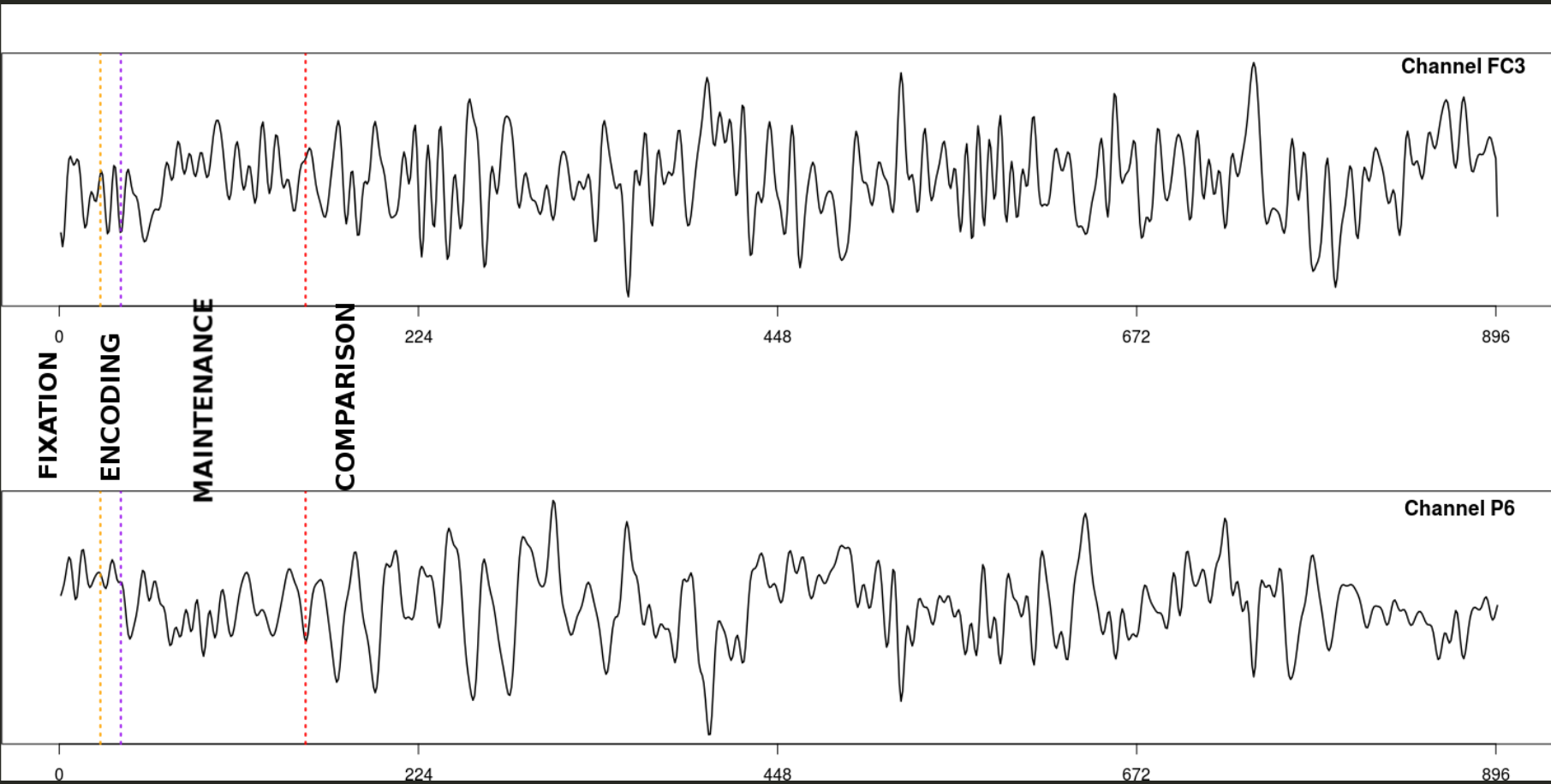Have to Remember Orientation of Red Bars

Remember during Maintenance Phase.

This red bar has a different orientation, so this is a *mismatch*.

Fixation
(200ms)

Encoding
(100ms)

Maintenace
(900ms)
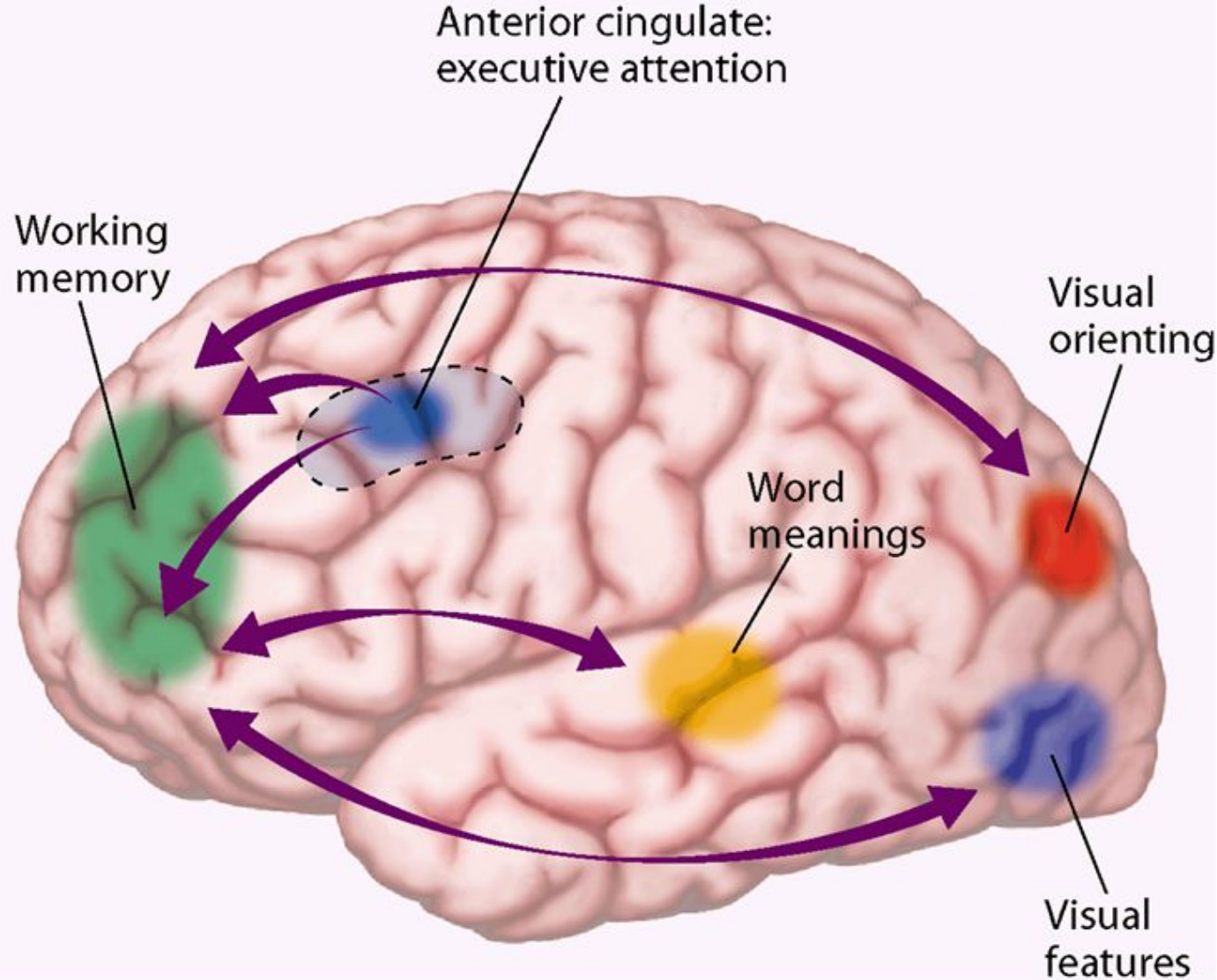
Comparison
(<2000ms)

Related Previous Results

- *Sauseng et al. (2005)* also found that **theta** and **alpha coherence** play a significant role in "top-down" control during working memory tasks

- *Jiang et al. (2005)* A study of EEG power and coherence in patients with mild cognitive impairment revealed differences in **theta, alpha, and beta band** power between MCI and normal controls during working memory tasks.

- *Zhang et al. (2007)* Found that **polymorphisms in a dopamine receptor gene affect neuronal activity** during working memory.

- *Vogler et al. (2014)* **Estimated SNP-based heritability for working memory performance** to be in the range of $h^2 = 0.31$ to $0.41 (P = 0.0008)$.

- *Cassidy et al. (2016)* detected relationships between **dynamic connectivity** during **working memory** tasks and **dopamine release in schizophrenics**.

- *Dai et al. (2017)* found topological reorginzation of EEG cortical **connectivity in the theta and alpha bands** during working memory tasks.

Anterior cingulate: executive attention

Working memory

Visual orienting

Word meanings

Visual features

The anterior cingulate cortex has been hypothesized to operate as an executive attention system. This system serves to ensure that processing in other brain regions is most efficient given the current task demands. Interactions with the PFC may select working memory buffers; interactions with the posterior cortex can serve to amplify activity in one perceptual module over others (Posner and Raichle, 1994; Gazaniga, 2002).

SNPs Related to Alzheimer's Disease

Info on SNPs from SNPedia:

- *rs2227564*: A functional polymorphism within plasminogen activator urokinase (PLAU) which some studies have shown to be associated with Alzheimer's disease. *(Riemenschneider et al., 2006)*

- *rs3851179*: A slight protective effect of the (A) allele of this SNP was found in *(Carasquillo et al., 2010)*

- *rs3818361*: A SNP associated with the complement component (3b/4b) receptor 1 CR1 gene. *(Carasquillo et al., 2010)*

- *rs9886784*: An intergenic SNP on chromosome 9, is reported to influence the risk for Alzheimer's disease; the odds ratio is 3.23 (CI: 1.79 - 5.84). *(Li et al., 2007)*
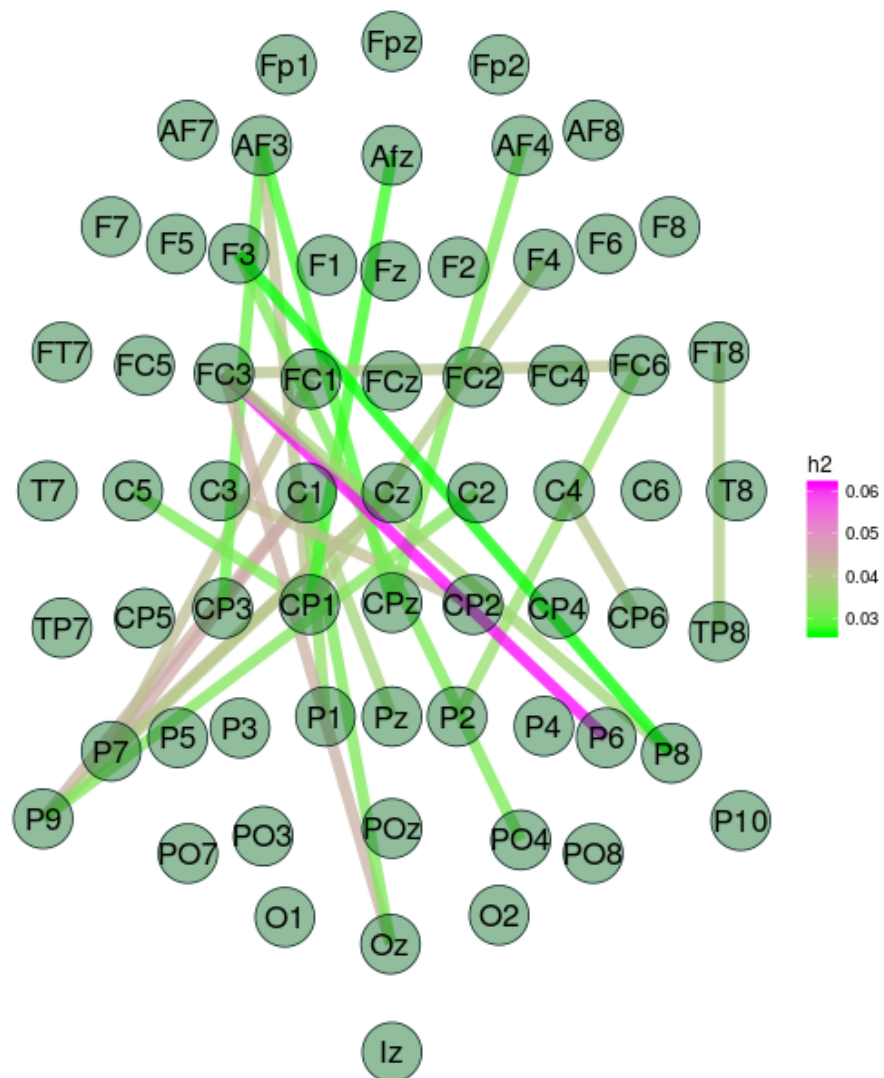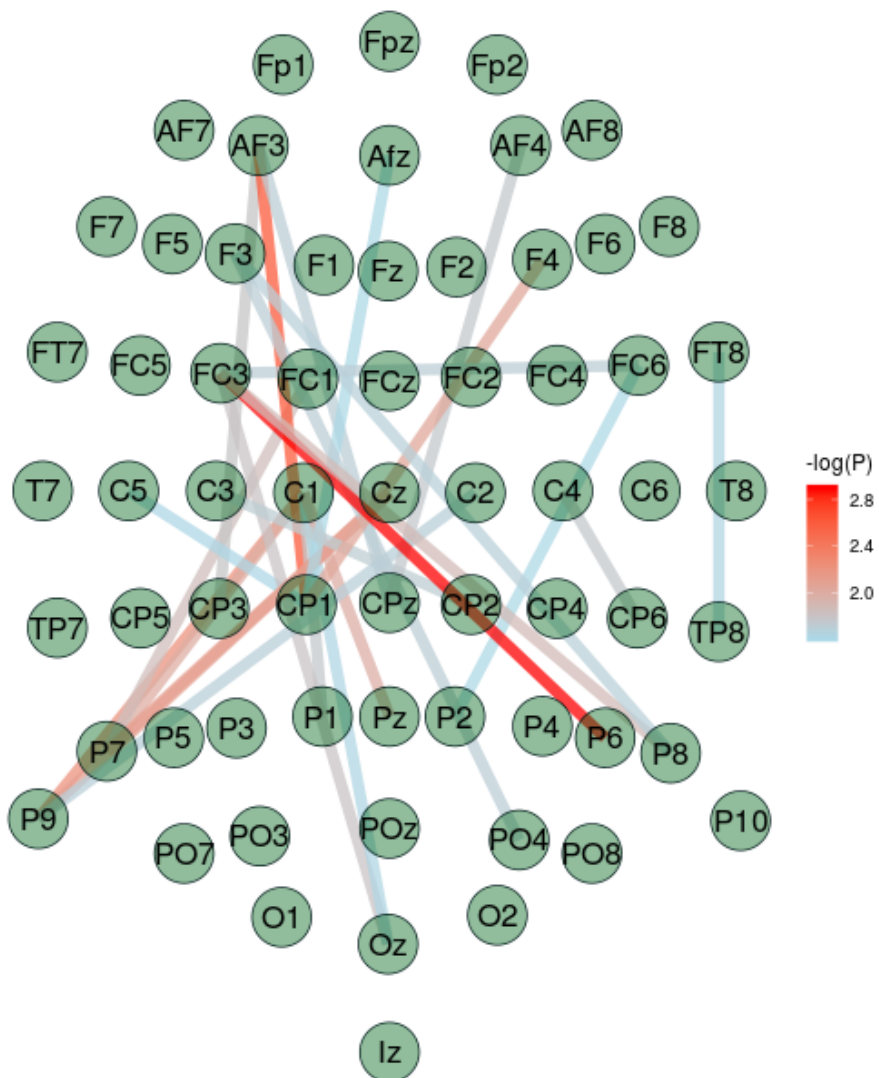
Data Description

- 350 Subjects from the BNU data set

- ~10 minute 64 channel EEG recording during VWM task

    - Preprocessed according to standard pipeline

    - Coherence measures for each channel pair was calculated by the FFT, and grouped into five frequency bands (in Hz): $\delta\,(1-4), \theta\,(4-8), \alpha\,(8-16), \beta\,(16-32), \gamma\,(32+)$

- 13 SNPs selected for analysis, previously identified as potential factors for Alzheimer's disease risk

    - All 13 SNPs passed standard MAF and HWE quality control checks

Adaptive Mantel Test Results

- Results of adaptive Mantel test for association of AD SNPs and EEG Coherence at particular frequency bands

- Used L_2 similarity for SNPs, and ridge kernel similarity for coherence, with penalty terms $\Lambda = \{0.5, 1, 5, 10, 100, 1000, \infty\}$

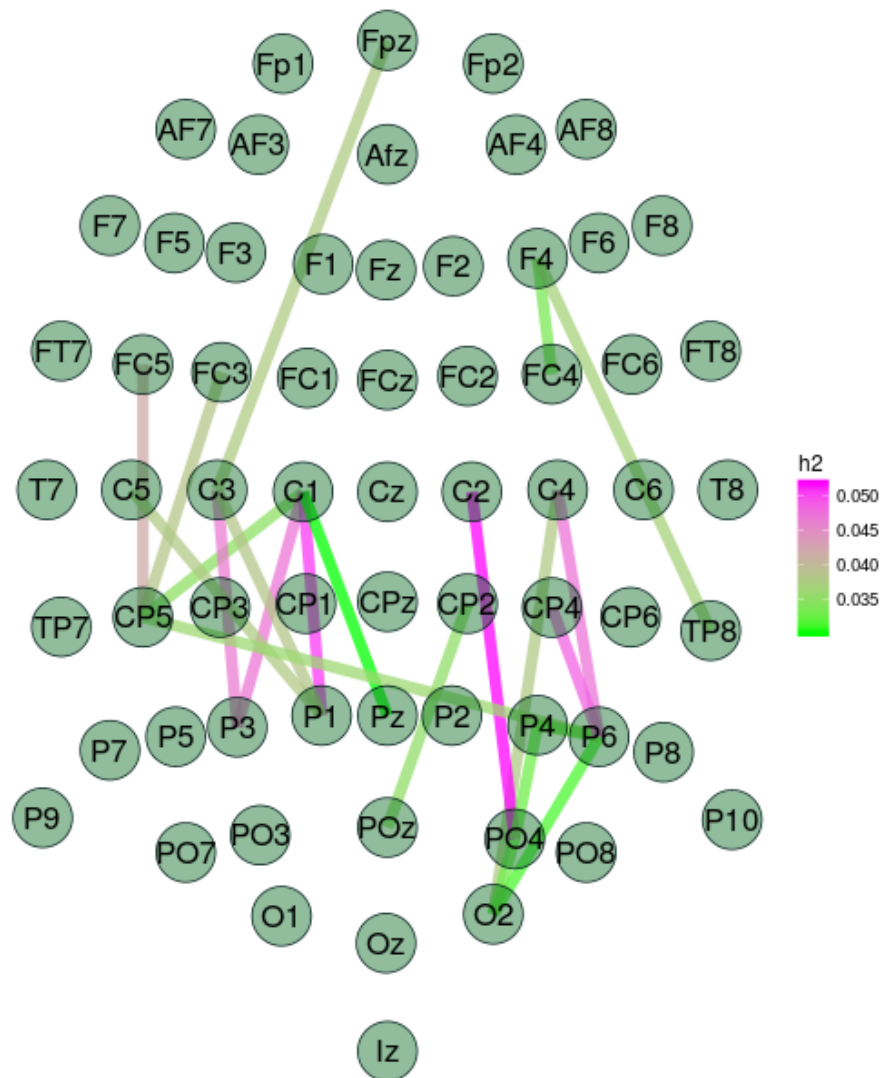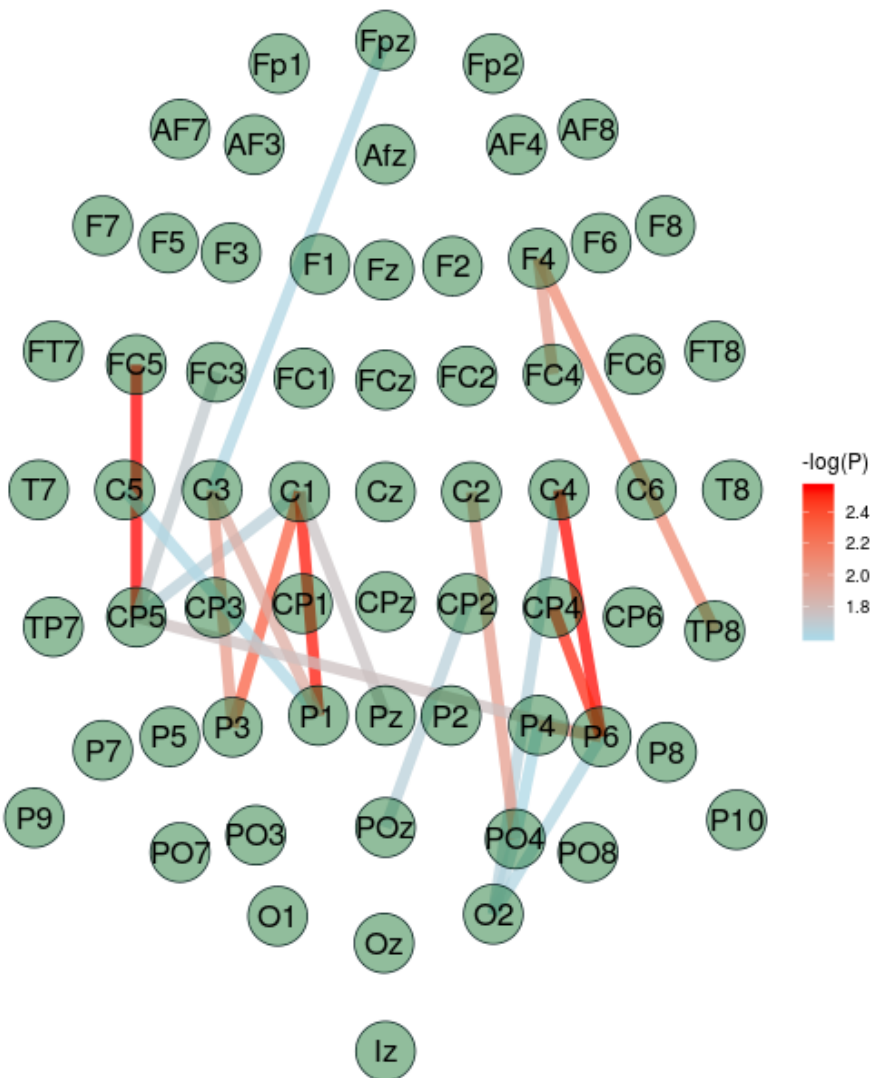| Band	Channels	$P-$value
α	All	0.075
α	Frontal	0.381
θ	All	0.416
θ	Frontal	0.081
β	All	0.619
β	Frontal	0.517
δ	All	0.015
δ	Frontal	0.088

h2 Theta Coherence

h2 Alpha Coherence
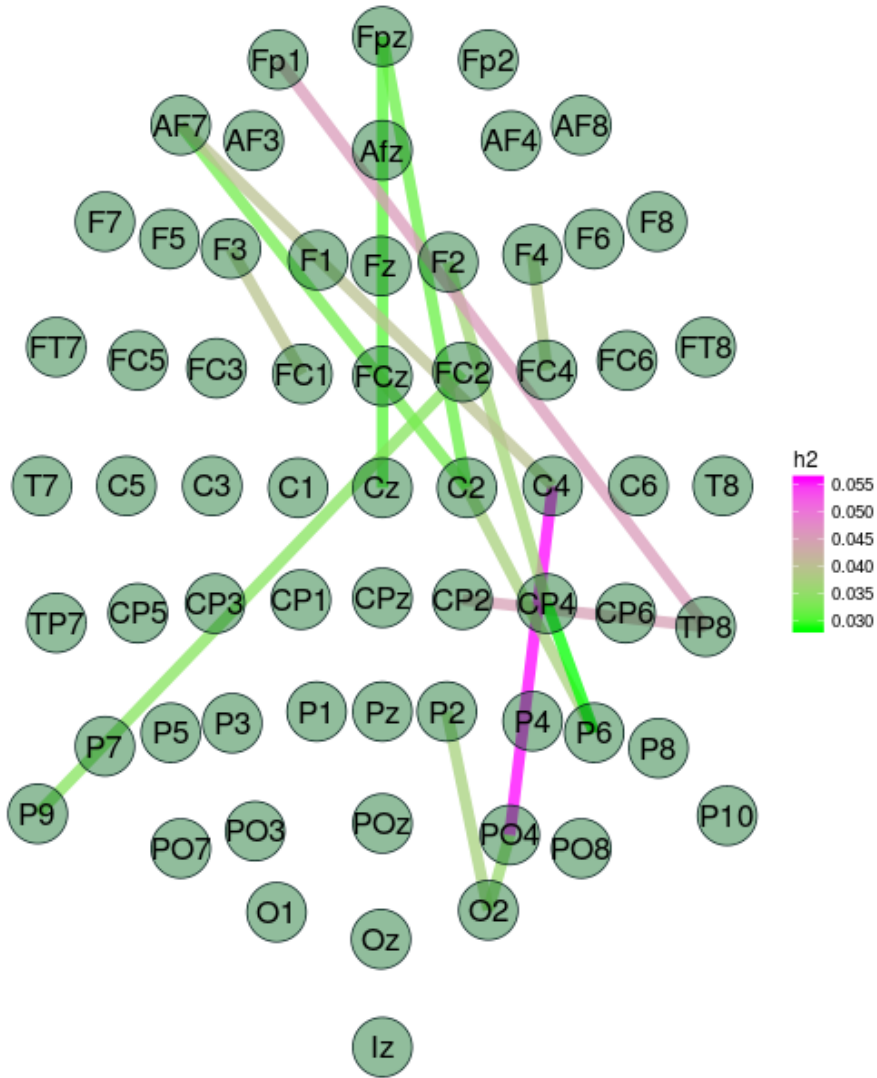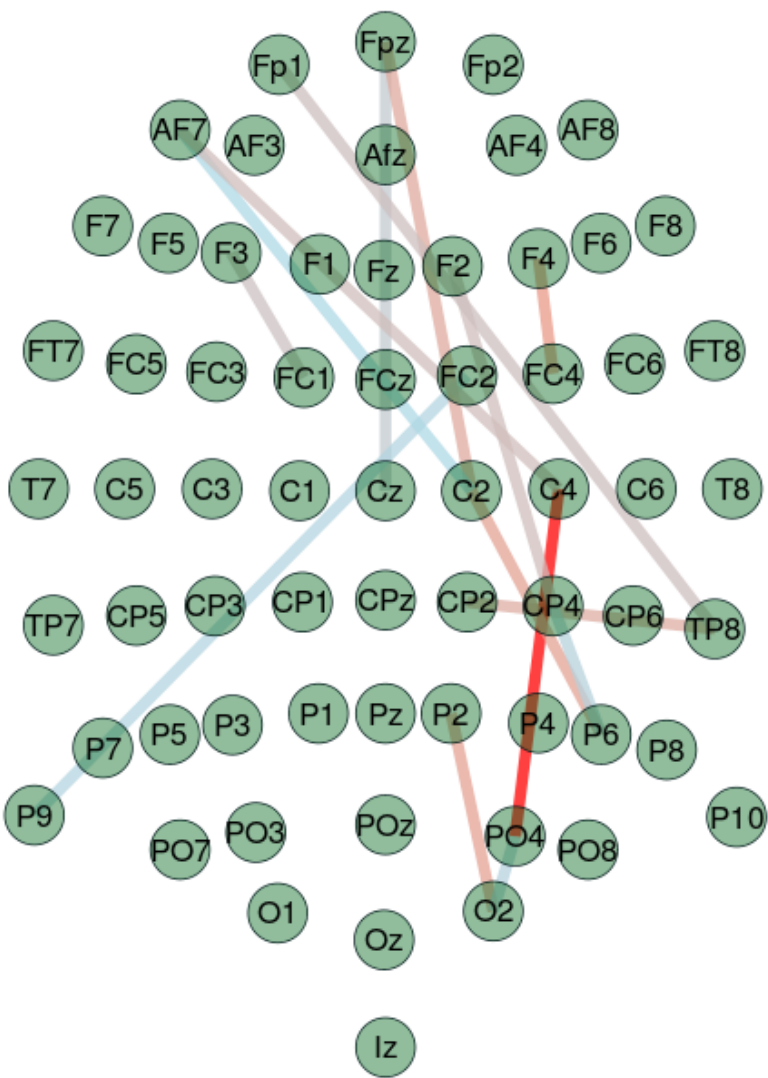
h2 Beta Coherence

Methods for Dynamic Connectivity Analysis
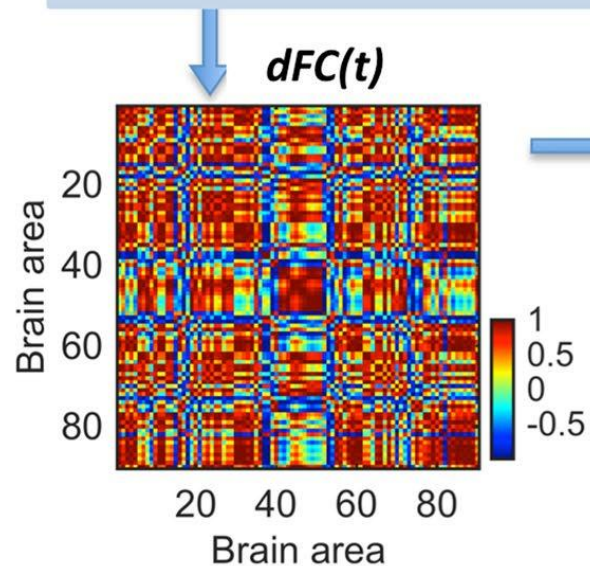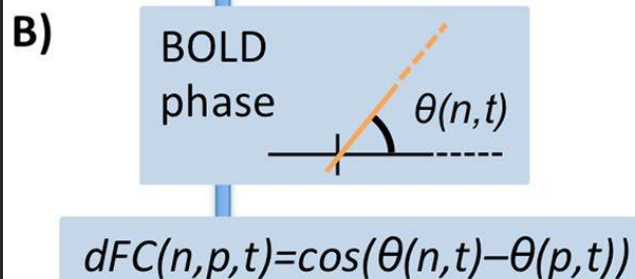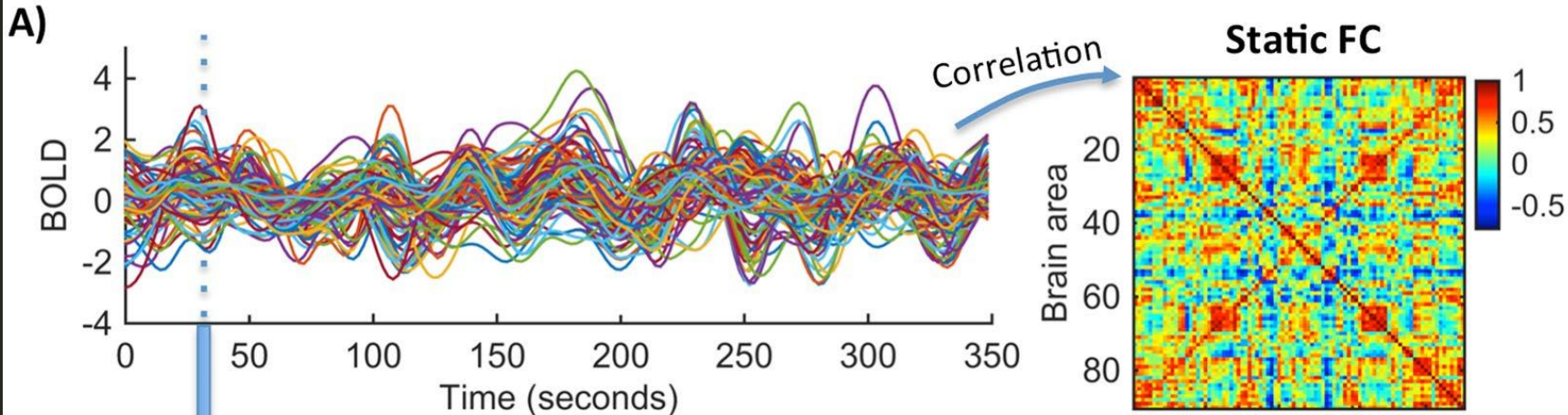
Dynamic Connectivity

DFC for fMRI with Time-varying BOLD Phase Coherence

Calculate BOLD Phase Coherence Connectivity to obtain a $q \times q \times T$ dFC matrix.

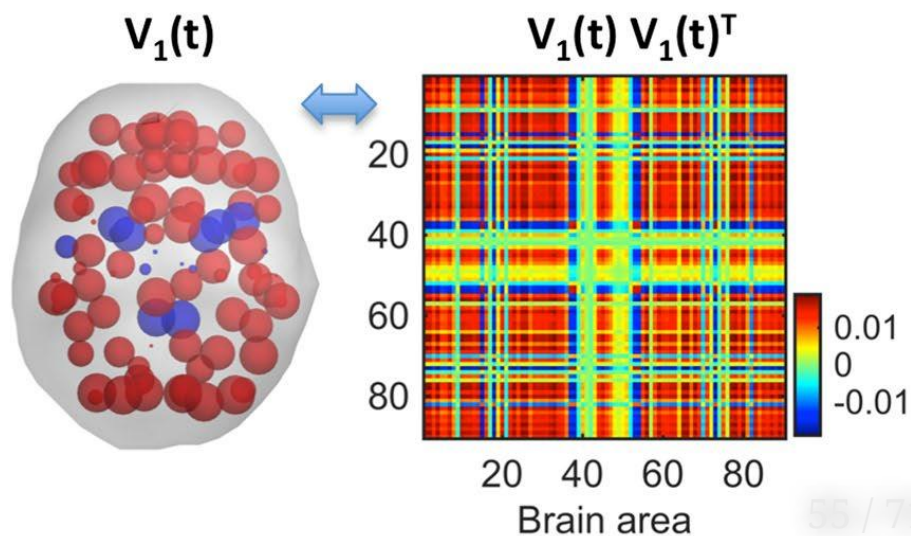- Computing Phase Coherence:

  - Estimate phase $\theta(j, t)$ of BOLD signals in area j at time t using Hilbert transform

  - Given phase of the BOLD signals, phase coherence between areas q and q' at time t (denoted $dFC(q, q', t)$) is defined as

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

- *Leading Eigenvector* of each $dFC(t)$ is used to capture connectivity structures.

A)

BOLD

Time (seconds)

Static FC

Brain area

Brain area

Correlation

B)

BOLD phase $\theta(n,t)$

$dFC(n,p,t)=cos(\theta(n,t)-\theta(p,t))$

dFC(t)

Brain area

Brain area

C) Leading Eigenvector of *dFC(t)*

$V_1(t)$

$V_1(t)\,V_1(t)^T$

Brain area

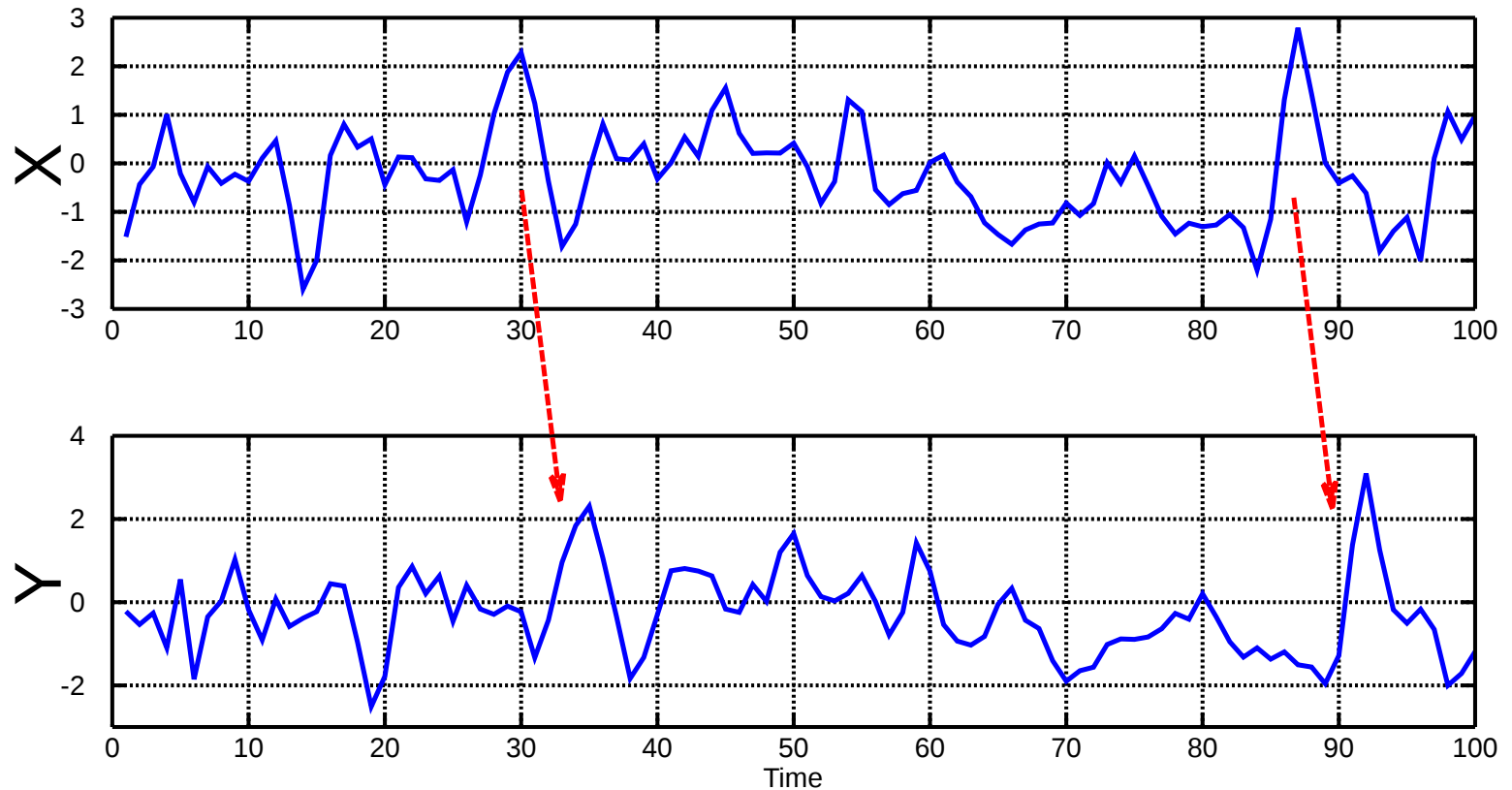Brain area

Dynamic Connectivity

Studying FC Dynamics

- Compute time-versus-time matrix to represent functional connectivity dynamics (FCD), where each $FCD(t_x, t_y)$ measures the resemblance between the dFC at times t_x and t_y using Pearson correlation or cosine similarity:

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

- *FC States:* detect discrete FC states with k-means clustering on the leading eigenvectors $V_1(t)$ across all time points and subjects.

- *Between group Comparisons:* Group differences are tested using permutation-based paired t-test

Fig. 1. Unified framework for estimating dynamic effective connectivity states in brain signals. The connectivity state estimation is initialized by K-means clustering of time-evolving connectivity features extracted using a TV-VAR process, and then refined by the SKF and EM algorithm based on a SVAR process.
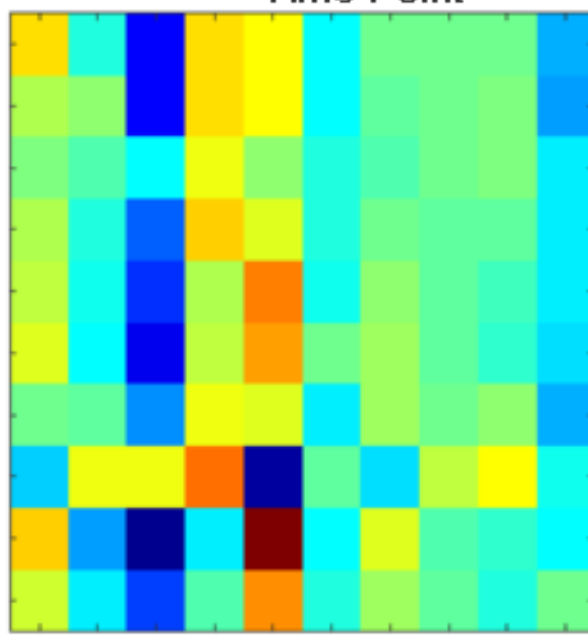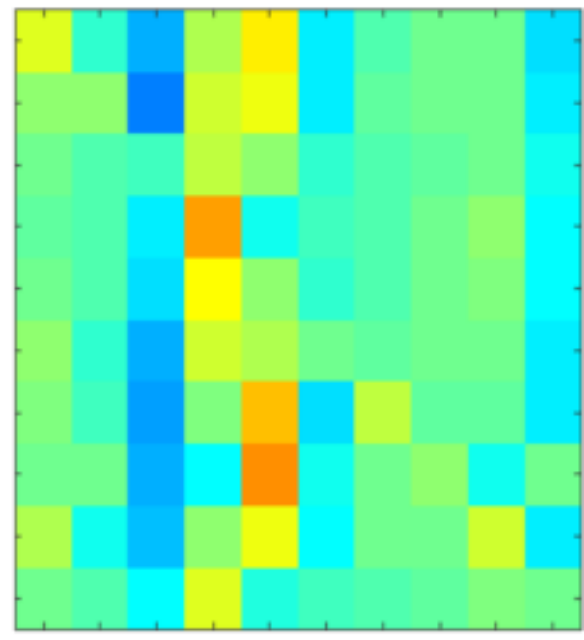
STATE 1   STATE 2   STATE 3

Multi-subject Dynamic Connectivity Modeling

Possible Extension of switching factor VAR Model

1. Fit TV-VAR for each subject and compute time series of PDC matrices.
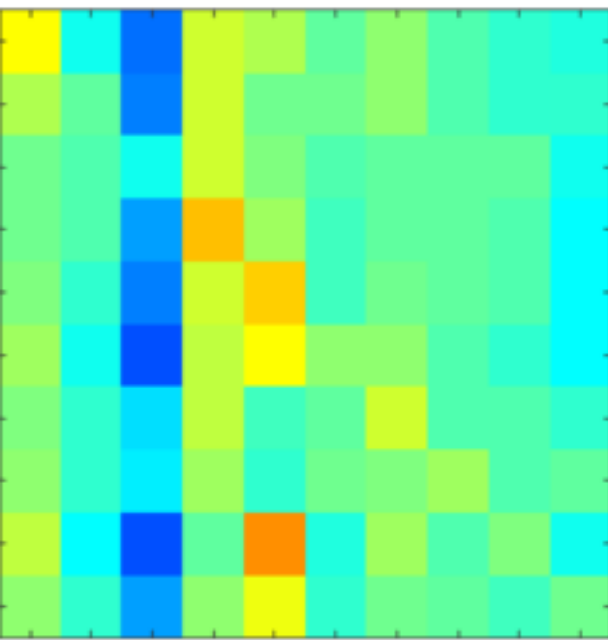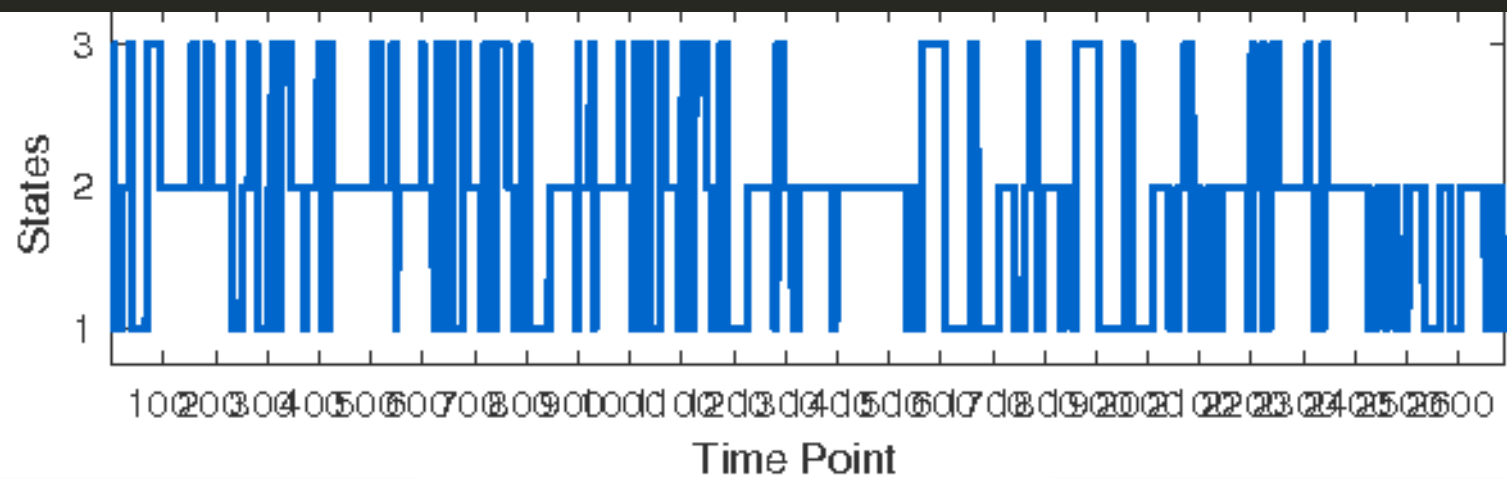
2. Apply K-means or HOSVD to concatenated PDC time series of all subjects.

3. Possibly refit and smooth with SVAR as before (or consider alternative ways of encouraging smoothness)

4. Conduct heritability analysis on things like:

   - transition probabilities between states
   - proportion of time spent in a particular state
   - frequency of a state during certain phases of the experiment

Manifold Regression

Manifold Regression

> **Riemannian Manifold**
>
> A smooth surface such that each point has a locally Euclidean neighborhood.

$\Omega_0 > 1$

$\Omega_0 < 1$

$\Omega_0 = 1$

MAP990006
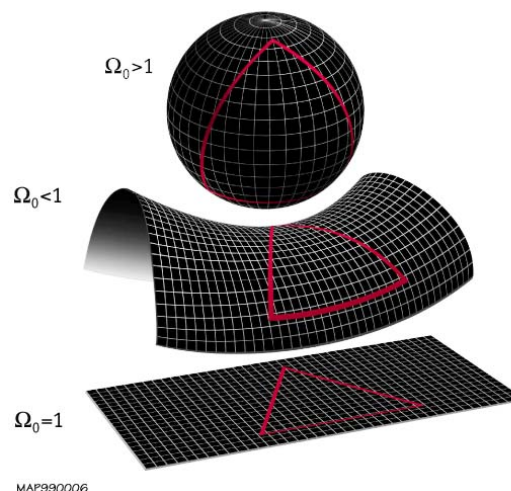
Some Manifolds of Interest

1. $\mathrm{Sym}^+(q)$, $q \times q$ **positive definite matrices**

2. $O(q)$ and $SO(q)$, the groups of **orthogonal** and **special orthogonal** $q \times q$ matrices

3. **Stiefel Manifold**, $q \times k$ matrices of orthonormal vectors.

Positive Definite Cone

$$\bar{X} = \begin{pmatrix} 0.75 & 0 \\ 0 & 0.75 \end{pmatrix}$$

$$\hat{X} = \begin{pmatrix} 0.65 & 0 \\ 0 & 0.65 \end{pmatrix}$$

$$X_1 = \begin{pmatrix} 1.3 & 0 \\ 0 & 0.2 \end{pmatrix}$$

$$X_2 = \begin{pmatrix} 0.2 & 0 \\ 0 & 1.3 \end{pmatrix}$$

c

a

b
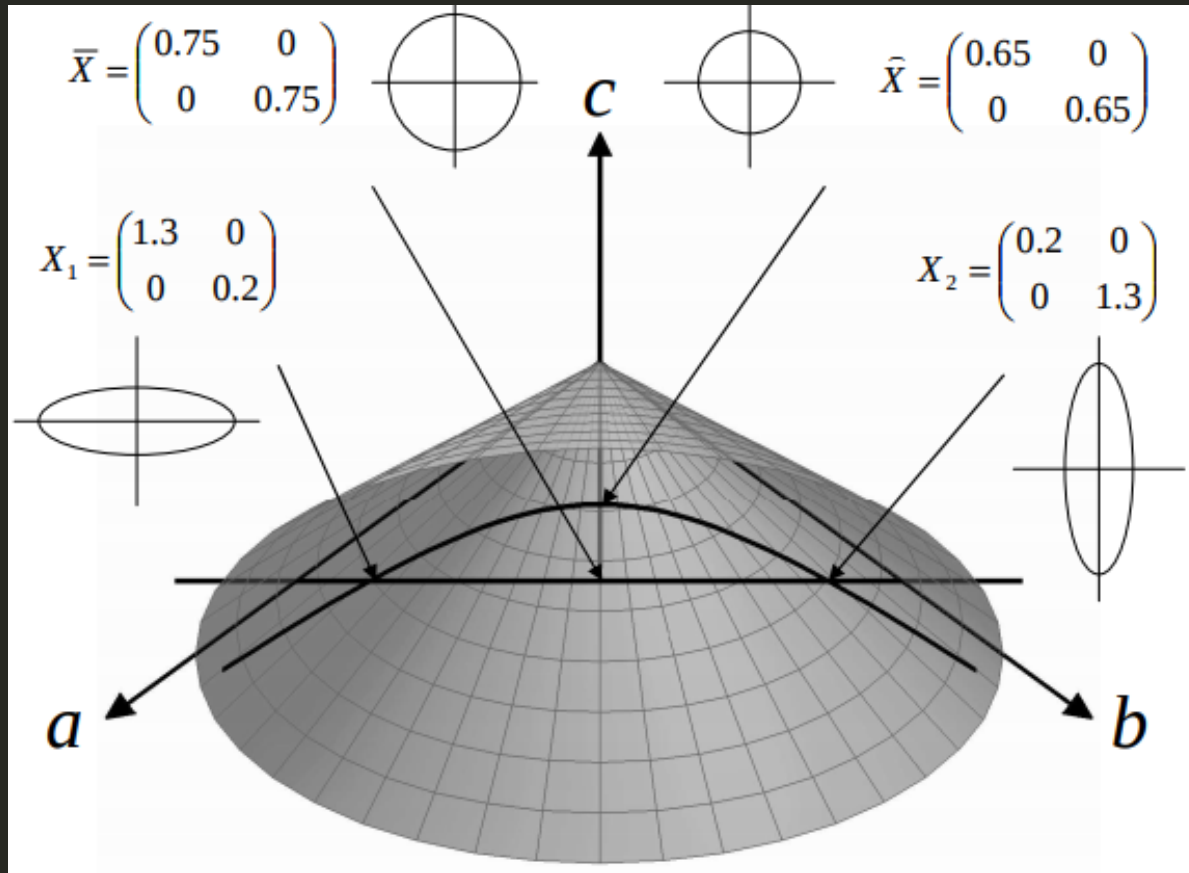
Manifold Regression

Specification of the Geodesic Model

- Suppose a $q \times q$ PD matrix $S_i \in \mathrm{Sym}^+(q)$ 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

- $\Sigma(\cdot, \cdot) : \mathbb{R}^{\Bbbk} \times \mathbb{R}^{\mathrm{p}} \to \mathrm{Sym}^+(q)$.

- We are interested in modeling the **conditional mean** of S_i given x_i, denoted

$$\Sigma_i(\beta) = \Sigma(x_i, \beta) \in \mathrm{Sym}^+(q).$$

Manifold Regression

Specification of the Geodesic Model

- Further let:

  - $D \in \operatorname{Sym}^+(q)$ be the intercept matrix $D = \Sigma(0, \beta)$, with $D = BB^T$ for some $B \in GL(q)$
  - $Y_D(x_i, \beta) = Y_{D,i} \in \operatorname{Sym}(q)$ be a "directional" matrix
  - $C_i(\beta)$ be a Cholesky square root $\Sigma(x_i, \beta) = C_i(\beta)C_i(\beta)^T$.

- The **geodesic model** is given by

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

Manifold Regression

- The **geodesic model** is given by

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

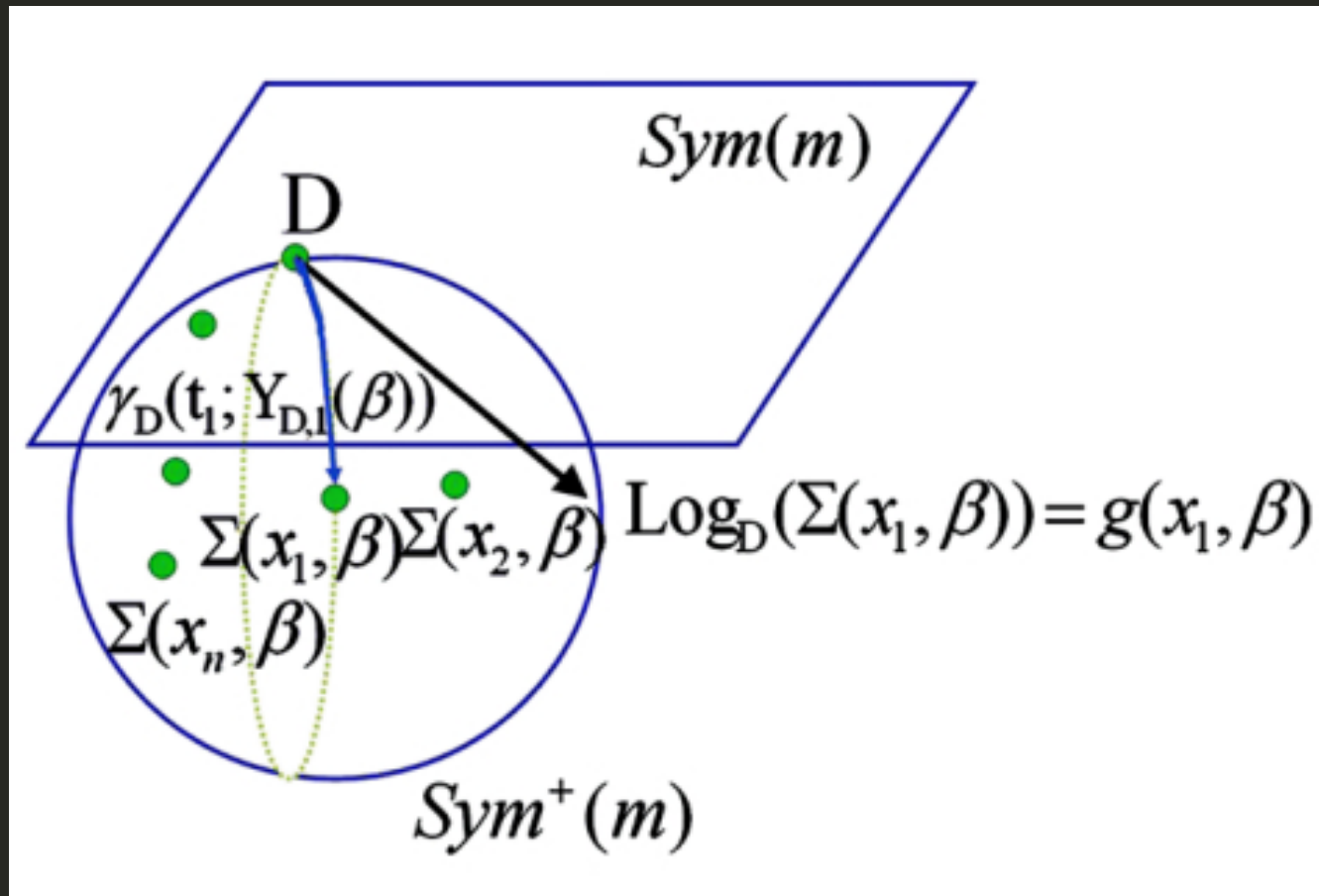- The **residuals** are defined as

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

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

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

Manifold Regression

Geodesic Model

Manifold Regression

Manifold Mixed Effects Model

- Kim et al. (2017) have proposed an extension of the linear mixed effects model to 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 **manifold mixed effects model** can be stated as

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

$$B_i = \mathrm{Exp}(B, U_i).$$

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

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

Manifold Regression

Heritability Estimation for Manifold-valued Phenotype

- Analagous to the definition of the variance components model for heritability analysis, we propose a random effects model for manifold-valued phenotypes

- For this model, we instead consider the random effects U_i to have a common aggregate genetic effect σ_U^2

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

Links

- **Adaptive Mantel Test Paper:** arxiv.org/pdf/1712.07270.pdf

- **Slides available:** github.com/dspluta/Presentations/

- **Adaptive Mantel R Package:** github.com/dspluta/adamant

References

- Wu M., et al. Kernel Machine SNP-set Testing under Multiple Candidate Kernels. Genetic Epideomiology. 2013. 37(3): 267-275.

- Cai T., et al., Kernel Machine Approach to Testing the Significance of Multiple Genetic Markers for Risk Prediction. Biometrics. 2011. 67(3): 975-986.

- Ge T, et al. Massively Expedited Genome-Wide Heritability Analysis (MEGHA). PNAS. 2015. 112, 2479-2484.

- Xue G, et al. Functional Dissociations of Risk and Reward Processing in the Medial Prefrontal Cortex. Cerebral Cortex. 2009. 19, 1019-1027.

- Yang J, et al. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics. (2011) 88, 76-82.

- Tzeng et al. (2009) Biometrics 65, 822.

- Visscher et al. (2014) Statistical power to detect genetic (co)variance of complex traits using SNP data. PLoS Genetics.

- GCTA Power, http://cnsgenomics.com/shiny/gctaPower/

Appendix

Adaptive Mantel Test

Computing the adaptive Mantel test can be done efficiently using either the SVD or a linear algebra trick, depending on the relative sizes of n and p.

SVD

- Computing the SVD $X = UDV^T$ can be completed in $O(np^2)$.

- When $rank(X) = r \leq n$, the Mantel statistic can be then be computed in $O(n^2)$:

$$T = \sum_{i=1}^{r} \eta_i z_i^2$$

- Using B permutations gives a total complexity of $O(np^2 + Bn^2)$.

Adaptive Mantel Test

Linear Algebra Trick

When $p \gg n$, it is better to instead use the following reformulation for K:

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

Calculating K_λ with this alternative form can be done in $O(n^2 p)$, giving a total computational cost of $O\left(n^2(p + B)\right)$.

- The computation for the adaptive test scales this cost linear relative the number of tuning parameters included.

- The computations can be easily parallelized.

EEG Pre-processing

- **EEG pre-processing:**

    1. Downsample from 1024 Hz to 128 Hz
    2. Remove bad channels
    3. Band-pass filter from 1 Hz to 45 Hz
    4. Interpolate/re-reference bad channels
    5. ICA to remove eyeblinks and motion artifacts
    6. Remove remaining bad trials. Exclude subjects if > 5% of trials removed.

- **Calculate coherence** for all subjects and all channels using the FFT, and compute mean coherence by frequency band.

Most Significant Channel Pairs for h^2

Estimating Heritability with GCTA

- While GCTA is a sensible and feasible approach to estimating heritability, it may prove impractical due to sample size limitations.

Power calculations for GCTA model.

Sample Size	Heritability `$h^2$`	Power	$SE(\hat h^2)$
1000	0.2	0.097	0.316
2000	0.2	0.24	0.158
3000	0.2	0.475	0.105
1000	0.5	0.353	0.316
2000	0.5	0.885	0.158
3000	0.5	0.997	0.105
400	1	---	0.79

http://cnsgenomics.com/shiny/gctaPower/