

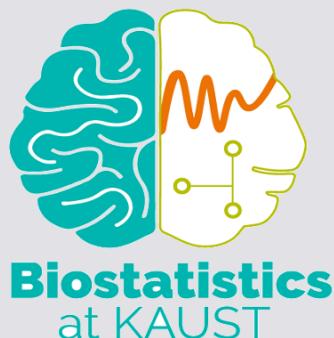
Analysis of Genetic Correlates of EEG Coherence During a Visual Working Memory Task

Dustin Pluta, Hernando Ombao, Zhaoxia Yu

22 Feb 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)



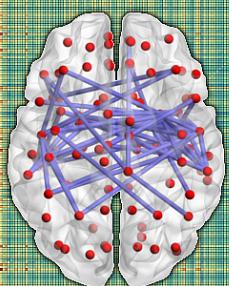
Overview of Talk

1. **Mantel test** and related distance-based approaches to multivariate association testing.
2. Introduce the **Adaptive Mantel test**, and present simulation results.
3. Application of the Adaptive Mantel test to analyze the **association of genetics and EEG coherence** from a working memory task.
4. **Current work** on EEG coherence and Genetics.
5. **Future Projects**

Resources

- **Adaptive Mantel Test Paper:** arxiv.org/pdf/1712.07270.pdf
- **Slides available:** github.com/dspluta/Presentations/
- **Adaptive Mantel R Package:** github.com/dspluta/adamant

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, $X \in \mathbb{X}^n$ is an $n \times p$ matrix of SNP measurements, and $Y \in \mathbb{Y}$ is an $n \times q$ matrix of EEG coherence values for a particular frequency band.
- Assume X and Y have been column centered and scaled.

Distance-based Association Testing

- **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 the random variables X and Y , but reduces to the Mantel test and RV coefficient test for certain choices of metrics.
- Many other classical statistics and tests can be reformulated in terms of pairwise distances or similarities, including the **multiple R^2 for linear regression, canonical correlation analysis, Hooper's trace correlation, and Pillai's trace**.
- We focus on the **similarity Mantel test** here, which is equivalent to the RV coefficient after normalization.

Mantel Test

- Assume X is an $n \times p$ column-centered matrix and Y is an $n \times 1$ centered vector.
- Given similarity metrics $d_X : \mathbb{R}^P \times \mathbb{R}^P \rightarrow \mathbb{R}$ and $d_Y : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, we can form two $n \times n$ similarity matrices K and H , where

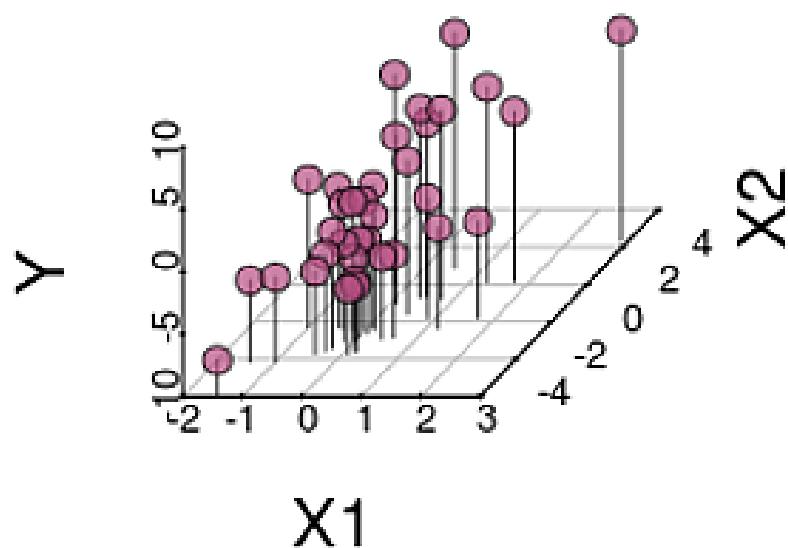
$$K_{ij} = d_X(X_i, X_j)$$

$$H_{ij} = d_Y(Y_i, Y_j).$$

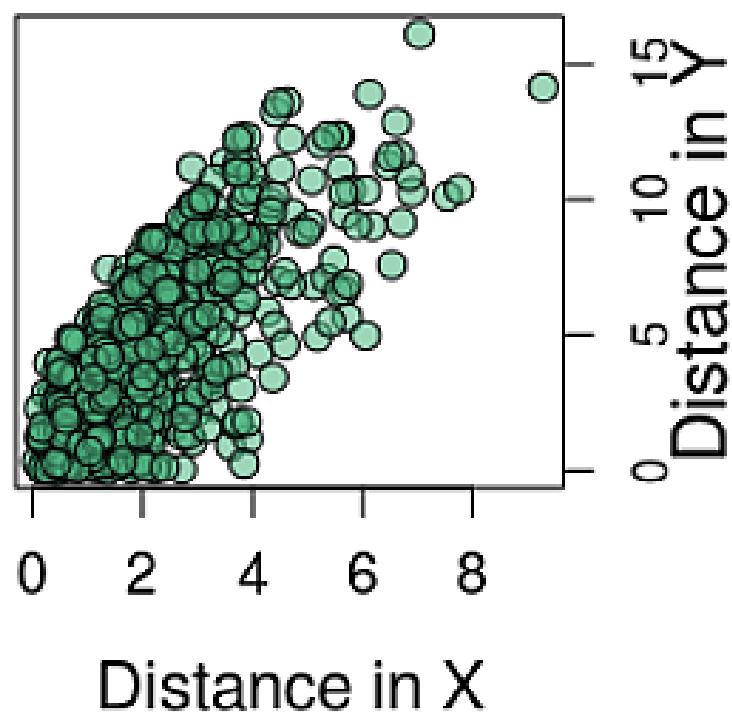
- The **correlation** of these distance matrices is

$$r(H, K) := \frac{\langle K, H \rangle}{\|K\|_2 \|H\|_2},$$

Simulated Data



Subject Pairwise Distances



Mantel Test

Statistical Question

How should we test the significance of the correlation?

- The approach originally suggested by Mantel (1967) is to **permute** rows and columns of one of the pairwise distance matrices to generate the reference distribution.
- That is, with K and H as **similarity matrices**, we calculate test statistic

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

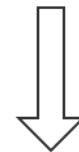
and calculate the **permutation P-value** by simultaneously permuting rows and columns of H to generate the reference distribution.

GRM

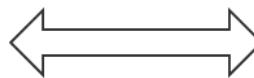
Connectivity Similarity

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

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



$$\begin{matrix} K_{11} \\ K_{12} \\ \vdots \\ K_{N(N-1)} \\ K_{NN} \end{matrix}$$



$$\begin{matrix} H_{11} \\ H_{12} \\ \vdots \\ H_{N(N-1)} \\ H_{NN} \end{matrix}$$

$$Z^* = \langle K, H \rangle \equiv \sum_i \sum_j K_{ij} H_{ij}$$

Mantel Test

Similarity with Weighted Inner Products

The **weighted inner product** $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ for some positive semi-definite matrix \mathcal{W} , is defined as

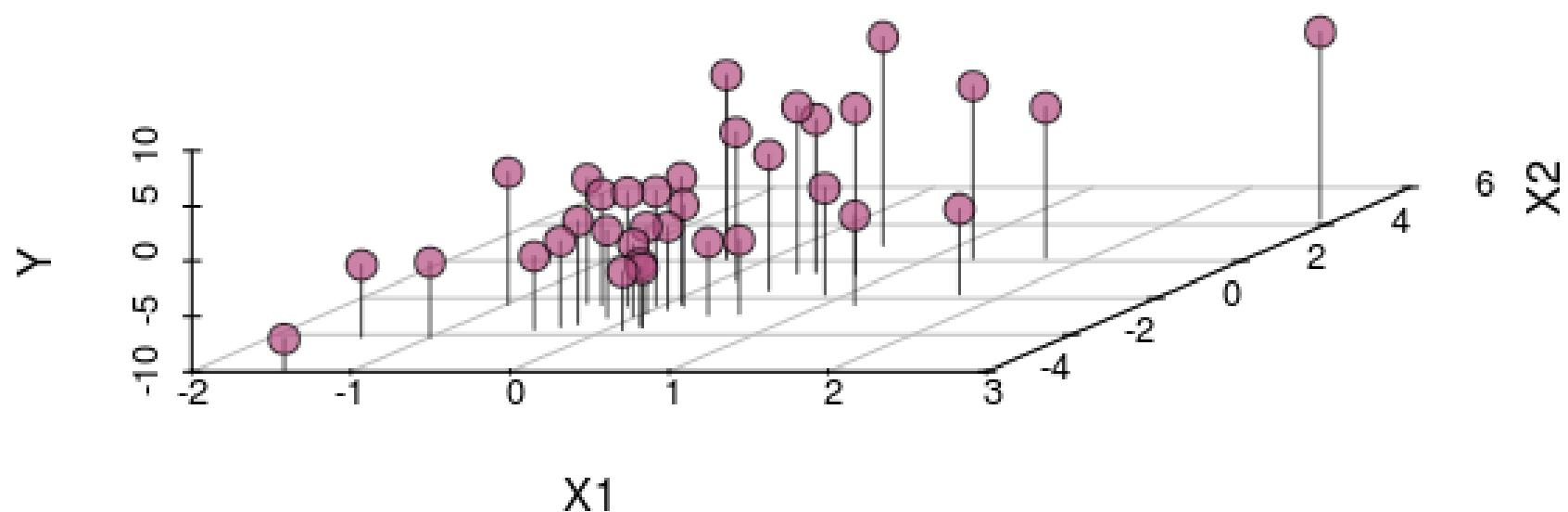
$$\langle u, v \rangle_{\mathcal{W}} = u \mathcal{W} v^T$$

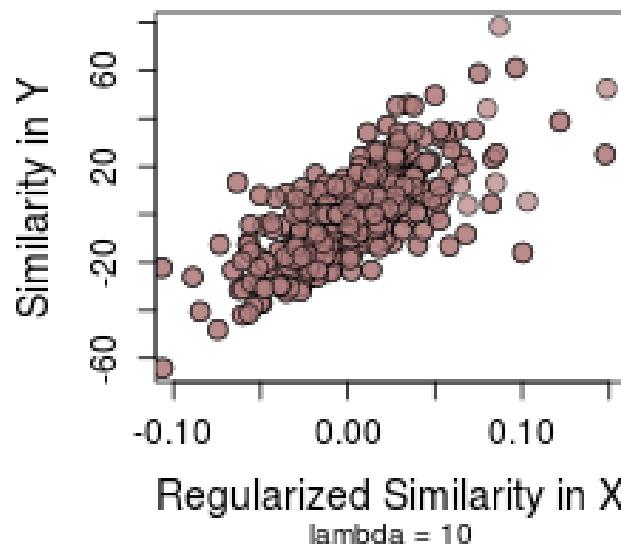
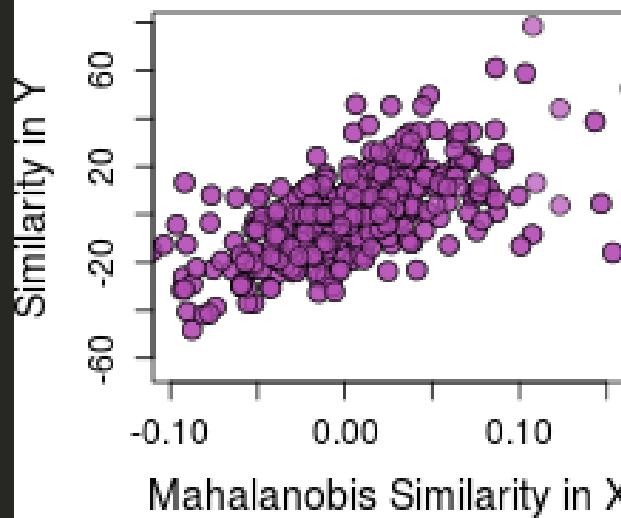
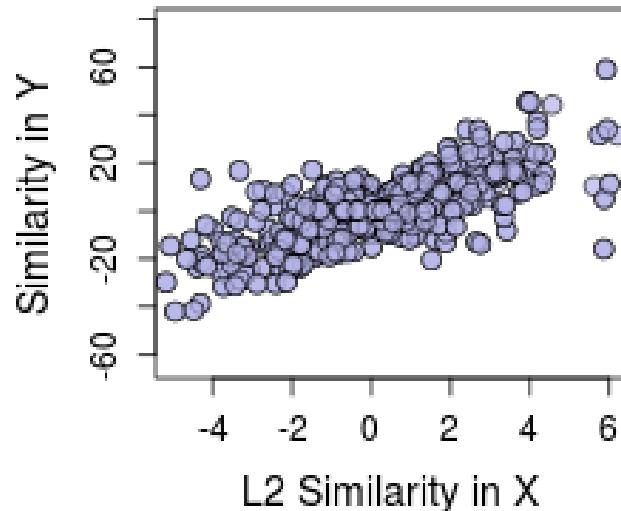
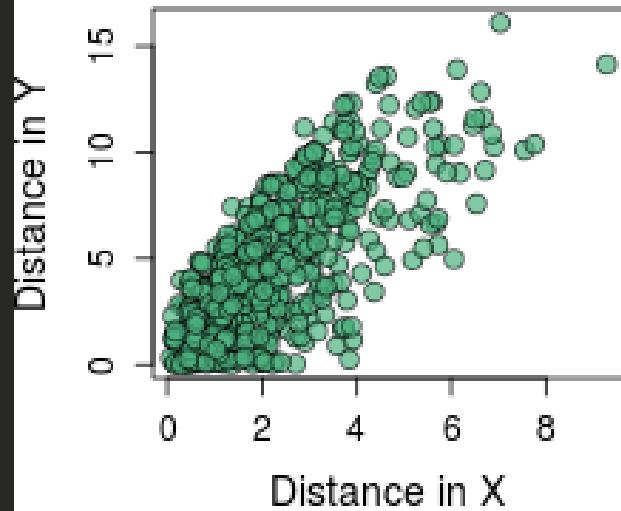
for two vectors $u, v \in \mathbb{R}^p$.

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

$$T_{\mathcal{W}} = \text{tr}(X \mathcal{W} X^T Y Y^T) = Y^T X \mathcal{W} X^T Y = \|X^T Y\|_{\mathcal{W}}^2.$$

Simulated Data





Mantel Test

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

Fixed Effects

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

Random Effects

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

Ridge Regression

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

Linear Model Score Tests

Model Definitions

Fixed Effects

$$Y \sim N(X\beta, \sigma_\varepsilon^2 I_N)$$

Ridge Regression

$$Y \sim N(X\beta, \sigma_\varepsilon^2 I_N), \quad \|\beta\|_2^2 < c(\lambda)$$

Random Effects

$$Y \sim N(0, \sigma_b^2 K + \sigma_\varepsilon^2 I_N), \quad K = XX^T$$

Linear Model Score Tests

Fixed Effects

$$S_F \asymp \text{tr}(HK_F) = Z^T D(D^T D)^{-1} D^T Z \sim c\chi_p^2,$$

Random Effects

$$S_R \asymp \text{tr}(HK_R) = Z^T D D^T Z \sim \sum_{j=1}^r \eta_j \chi_1^2,$$

Ridge Regression

$$S_\lambda \asymp \text{tr}(HK_\lambda) = \sum_{j=1}^r \frac{\eta_j}{\lambda + \eta_j} z_j^2 \sim ??$$

Limit Relationship

$$S_\lambda \asymp \text{tr}(HK_\lambda) \propto \left\{ \lambda \sum_{j=1}^r \frac{\eta_j}{\lambda + \eta_j} z_j^2 \right\} \xrightarrow{\lambda \rightarrow \infty} \sum_{j=1}^r \eta_j z_j^2 \sim \sum_{j=1}^r \eta_j \chi_1^2$$

Linear Model Score Tests

Geometric Interpretation

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

1. The *Fixed Effects* model tests the **Euclidean norm** of Z
2. The *Random Effects* model tests the **weighted Euclidean norm** of Z , where the j th component is weighted by the j th eigenvalue η_j .
3. The *Ridge Penalization* **weights the Euclidean norm** of Z proportional to the eigenvalues, **but these weights are now flattened** by a factor of $(\lambda + \eta_j)^{-1}$.

Linear Models

Summarizing these results:

- The **Similarity Mantel Test** is equivalent to the **Score test** of a linear model whose form depends on the choice of inner product.
- Thus, the **Mantel Test** implies an underlying parametric model through the choice of similarity kernel.
- The **Ridge Regression Score Test** converges to the **Random Effects Score Test** as $\lambda \rightarrow \infty$.
- The Fixed effects, random effects, and ridge penalized score tests form a single family of tests parameterized by λ .
- The differences in these tests can be understood geometrically as testing the norm of the projection of Y into $\mathcal{C}(X)$, with coordinate weights given by the eigenvalues of X .

Multivariate Mantel

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

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

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

- The Mantel test procedure can be performed exactly the same as before with test statistic $T = \text{tr}(KH)$.
- Assuming $\text{rank}(K) = p$ and $\text{rank}(H) = q$, we can write

$$\text{tr}(KK) = \text{tr}(K) = \text{rank}(K) = p$$

$$\text{tr}(HH) = \text{tr}(H) = \text{rank}(H) = q$$

$$r(K, H) = \frac{1}{\sqrt{pq}} \text{tr}(KH) = \frac{1}{\sqrt{pq}} T$$

Adaptive Mantel Test

Idea: using a variation of the adaptive sum of powered score tests from Xu et al. 2017, we can test a range of tuning parameters simultaneously and calculate an overall P -value without inflating the type I error.

Input & Output

- **Input:**
 - $X, n \times p$ covariates, column centered and scaled
 - $Y, n \times 1$ response, centered and scaled
 - Λ = list of regularization terms
- **Output:** P_{AM} = adaptive Mantel P -value for test of significant linear association

Adaptive Mantel Test

Algorithm

1. Calculate similarity matrices, $H := YY^T$ and $K_\lambda = X(X^TX + \lambda I)^{-1}X^T$.
2. Generate B permutations of H , labeled $H^{(b)}$.
3. Calculate $Z_\lambda^{(0)} := Z_\lambda := \text{tr}(K_\lambda H)$ from original data, and
 $Z_\lambda^{(b)} := \text{tr}(K_\lambda H^{(b)}), \forall b = 1, \dots, B; \lambda \in \Lambda.$
4. Compute permutation penalized P -values: $P_\lambda^{(b')} = \frac{1}{B+1} \sum_{b=0}^B I(Z_\lambda^{(b')} \leq Z_\lambda^{(b)})$.
5. Compute permutation adaptive P -values: $P^{(b)} = \min_{\lambda \in \Lambda} P_\lambda^{(b)}.$
6. Compute adaptive Mantel test P -value: $P_{AM} = \frac{1}{B+1} \sum_{b=0}^B I(P^{(0)} \leq P^{(b)}).$

AdaMant Package

Minimal package available at github.com/dspluta/adamant

- Can be installed using:

```
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

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

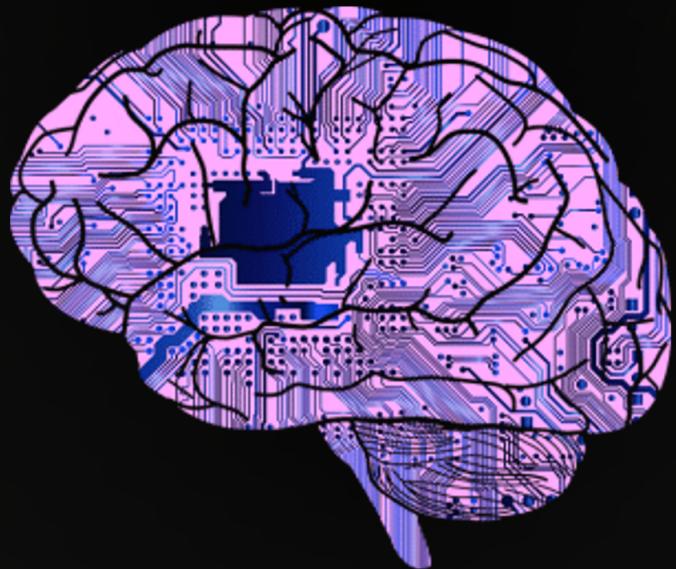
# Apply Adaptive Mantel Test
adamant(X, Y, lambdas_X = c(0, 1, 10, Inf),
        n_perms = 2000, P_val_only = TRUE)

## -----
## 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          = 0.088 secs
## -----
## [1] 0.0055
```

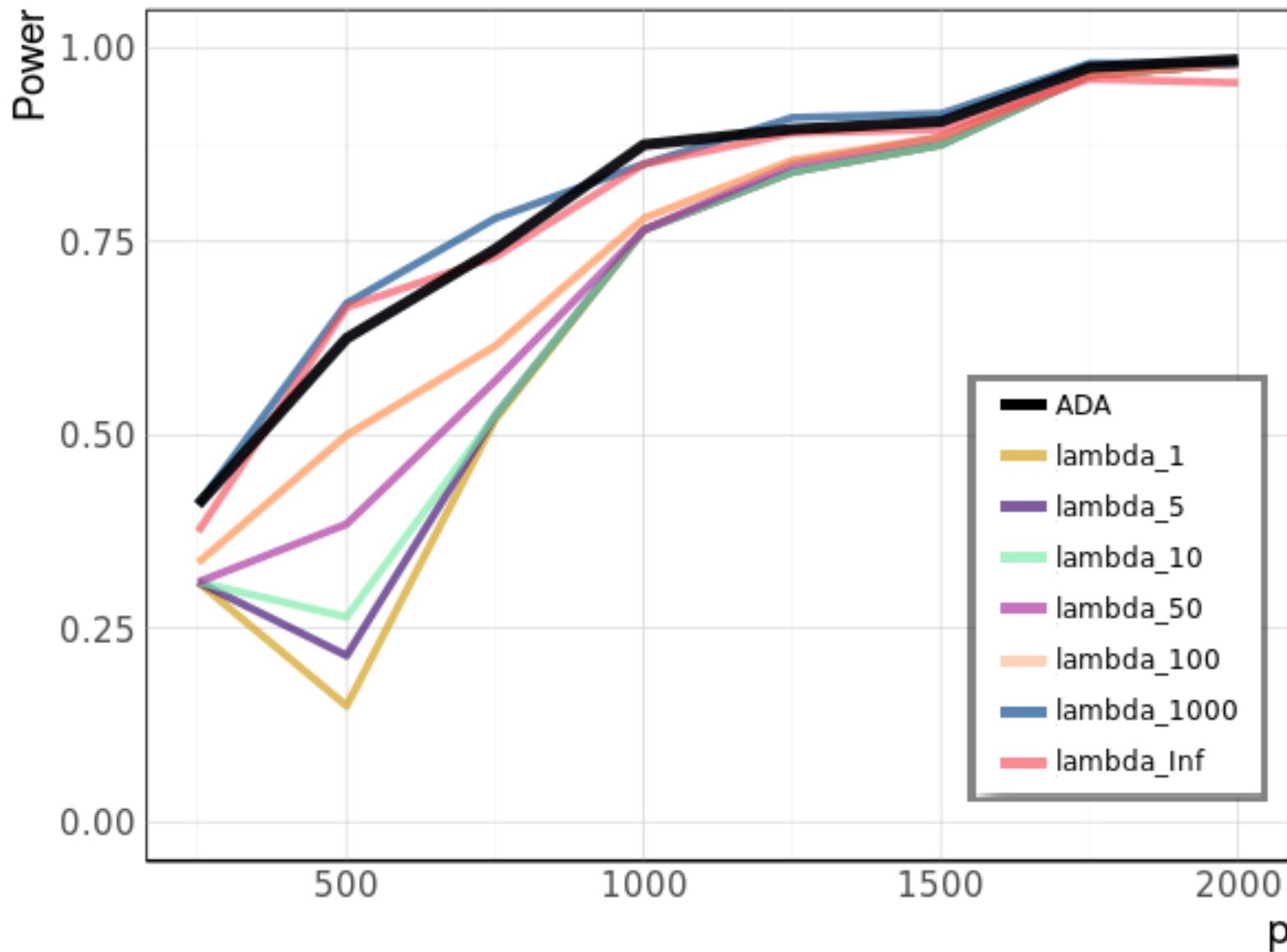
Heuristic to Choose λ

- The best linear unbiased predictors for the regression coefficients in the random effects model result from $\lambda = \frac{\sigma_e^2}{\sigma_b^2}$ as a ridge penalty term.
- Since h^2 can be calculated from this noise to signal ratio, we can choose a "reasonable" λ with the following simple rule of thumb:
 1. Guess at 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 choose a guess for h^2
- However, the formula for h^2 applies only for homogenous independent effects. Sparsity and correlation in the data require a different calculation for h^2

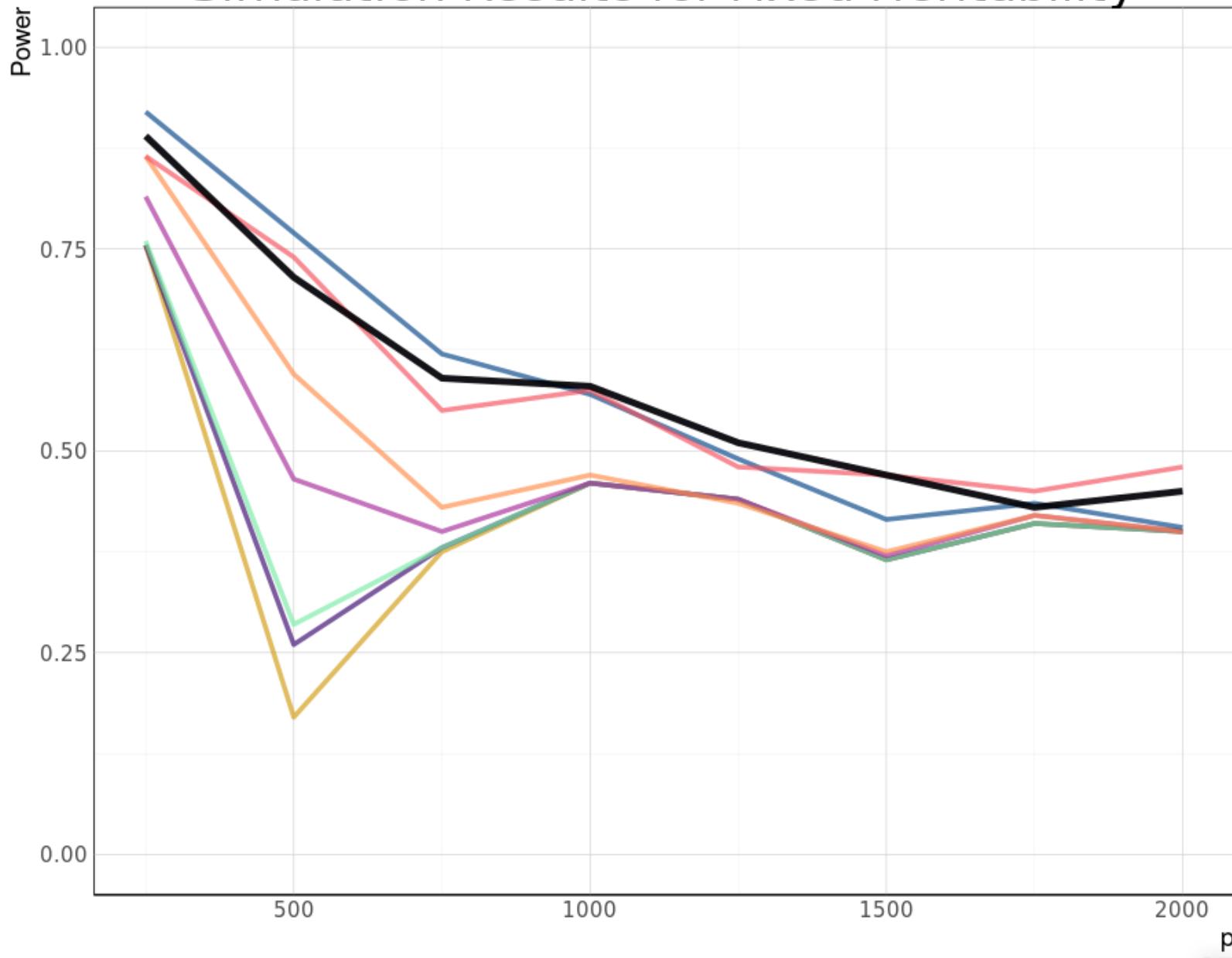
Simulations



Simulation Results for Fixed SNP Effect Size



Simulation Results for Fixed Heritability



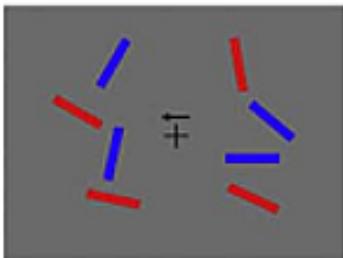
STAND BACK



I'M GOING TO TRY SCIENCE

A

2 targets +
2 distractors



2 targets



4 targets

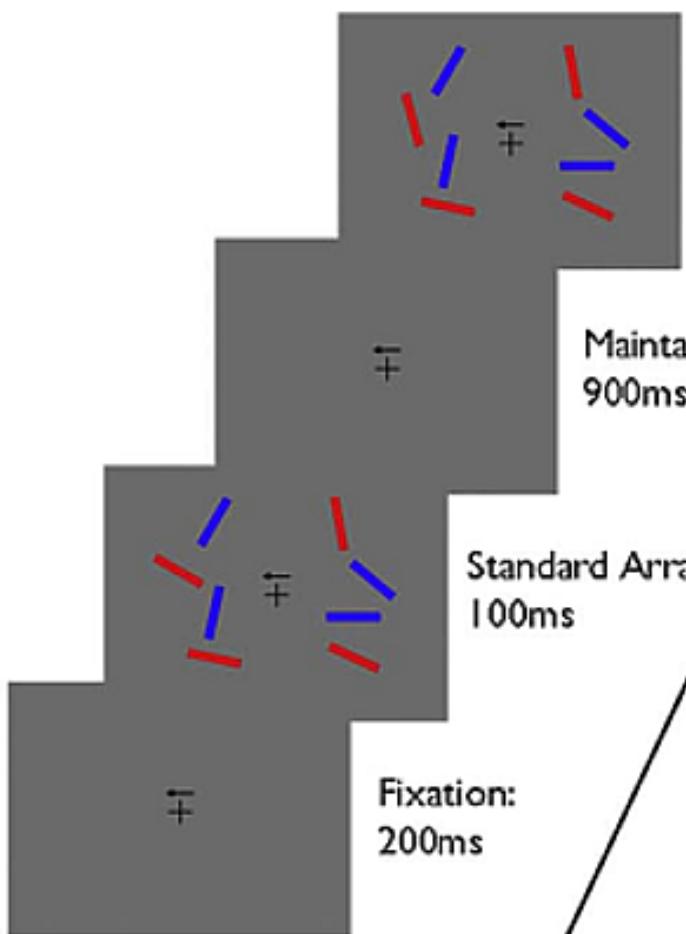
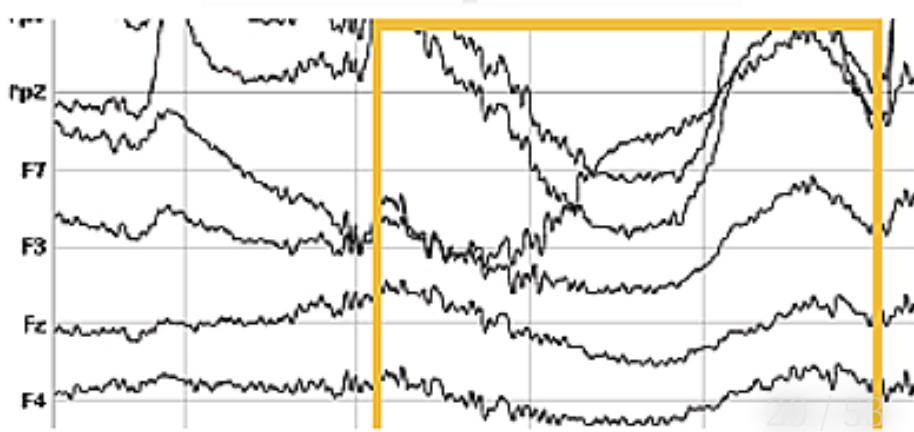
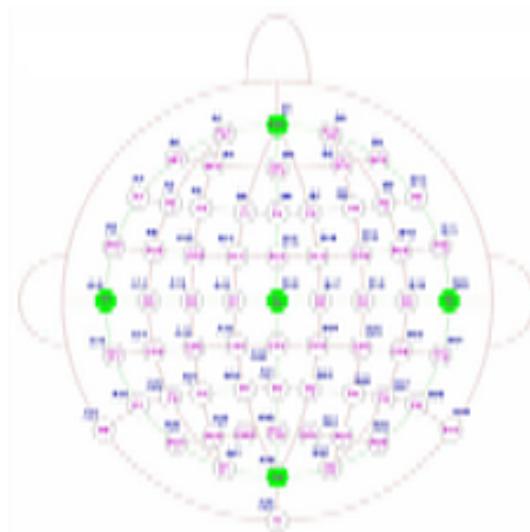
**B**

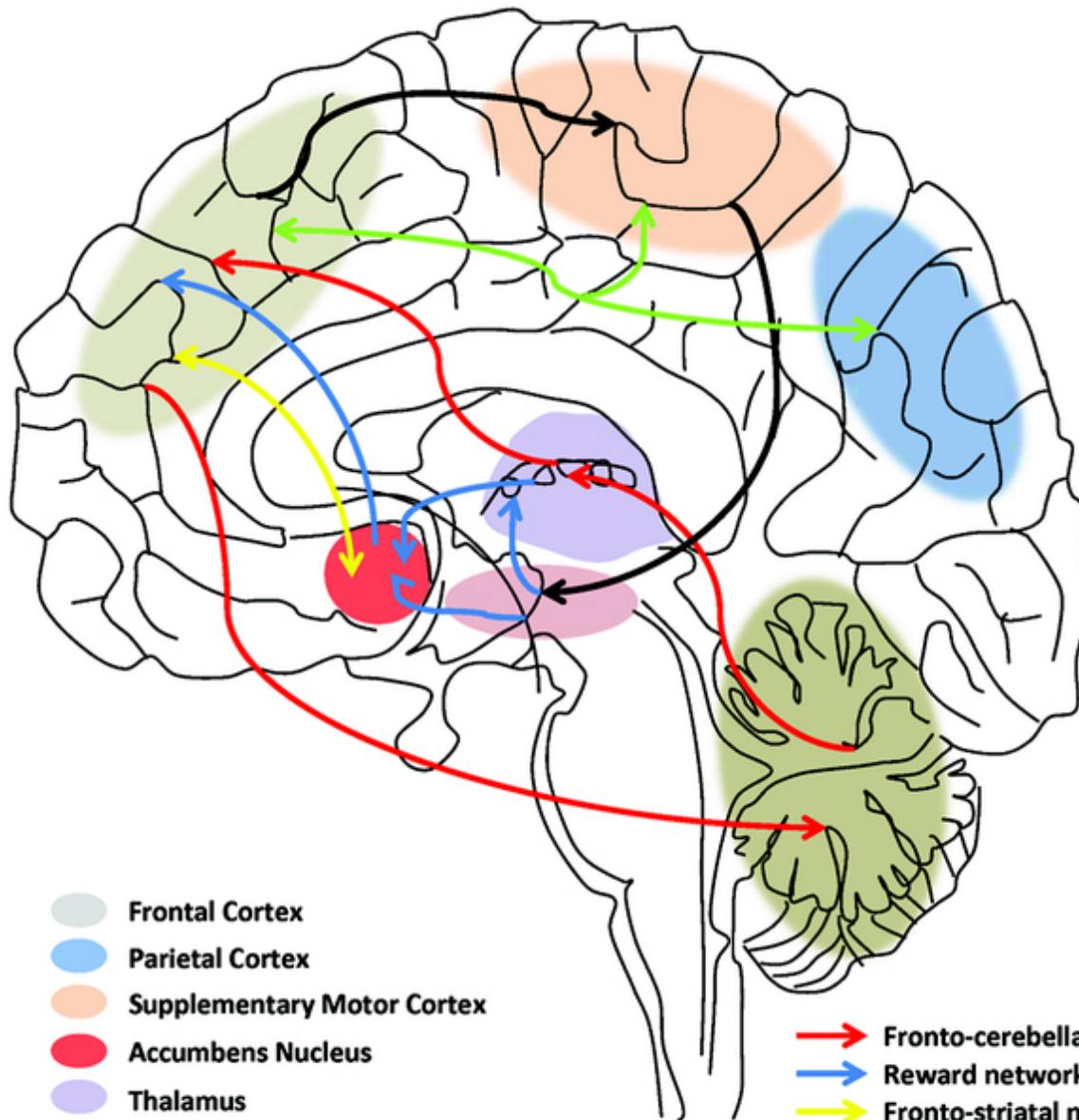
Comparison
Array:
 $<2000\text{ms}$

Maintanence:
900ms

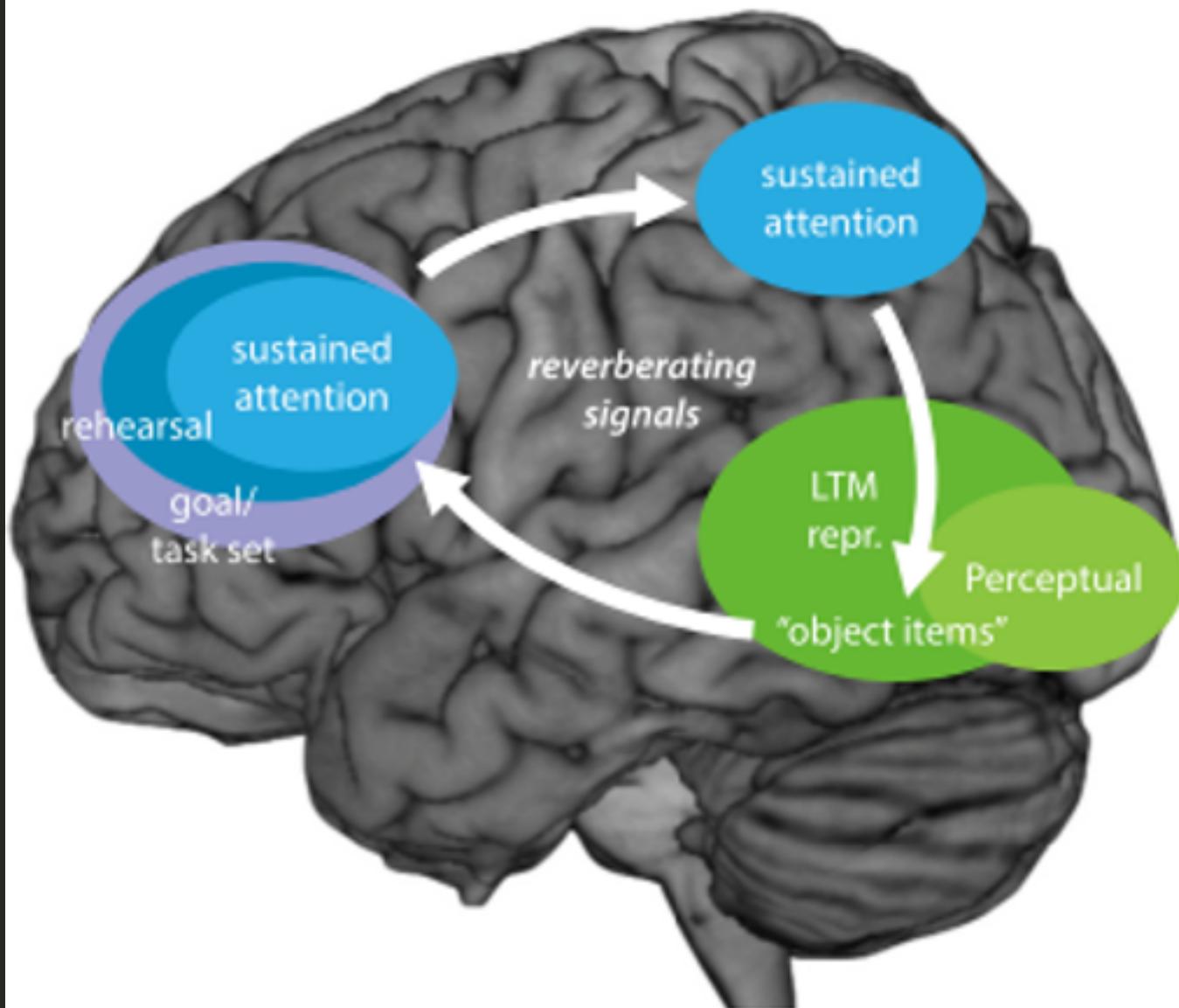
Standard Array:
100ms

Fixation:
200ms

**C**



Schematic mapping of processes/representations
to brain networks during working-memory maintenance



Related Previous Results

EEG Coherence During Working Memory Tasks

- *Onton et al. (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
- *Simons and Spiers (2003)* identified important interactions between the **prefrontal and medial temporal lobes** for the processing of long-term memory

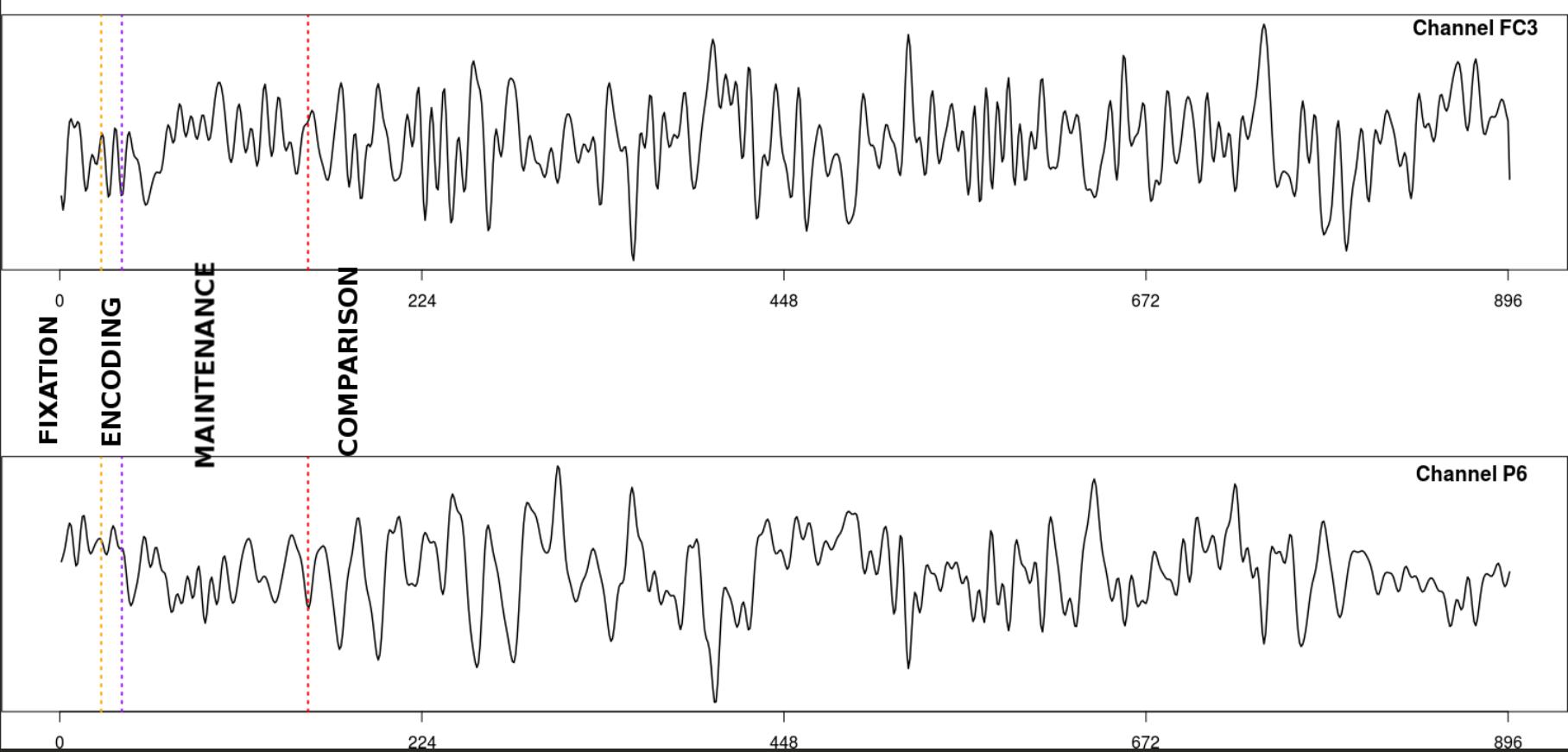
SNPs Related to Alzheimer's Disease

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 study of over 5,000 Alzheimer's disease patients (and over 10,000 controls) found a slight protective effect of the (A) allele of this SNP... The association of this SNP and late-onset Alzheimer's disease was replicated over a total of another 1829 cases (and 2576 controls). (*Carasquillo et al., 2010*)
- *rs3818361*: A SNP associated with the complement component (3b/4b) receptor 1 CR1 gene. The association of this SNP and late-onset Alzheimer's disease was replicated over a total of another 1829 cases (and 2576 controls). (*Carasquillo et al., 2010*)
- *rs9886784*: An intergenic SNP on chromosome 9, is reported to influence the risk for Alzheimer's disease based on a study of ~1100 Canadian patients. The risk allele is (A); the odds ratio is 3.23 (CI: 1.79 - 5.84). (*Li et al., 2007*)

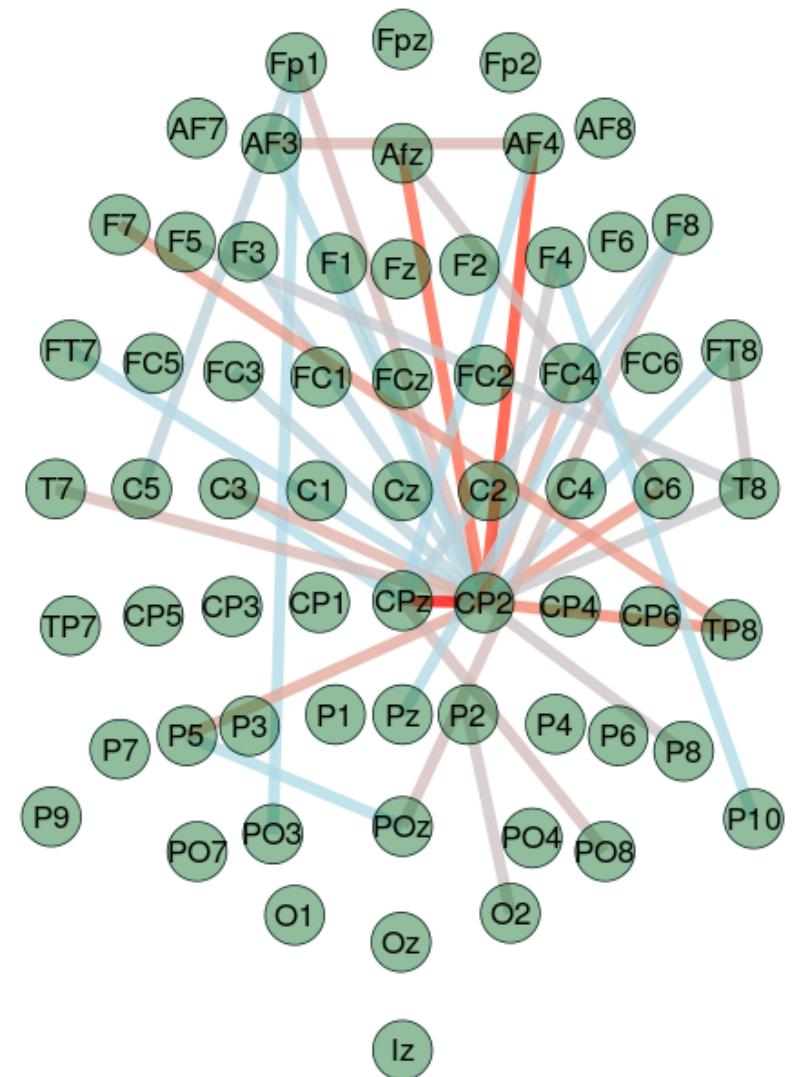
Data Description

- 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):
 δ (1 – 4), θ (4 – 8), α (8 – 16), β (16 – 32), γ (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



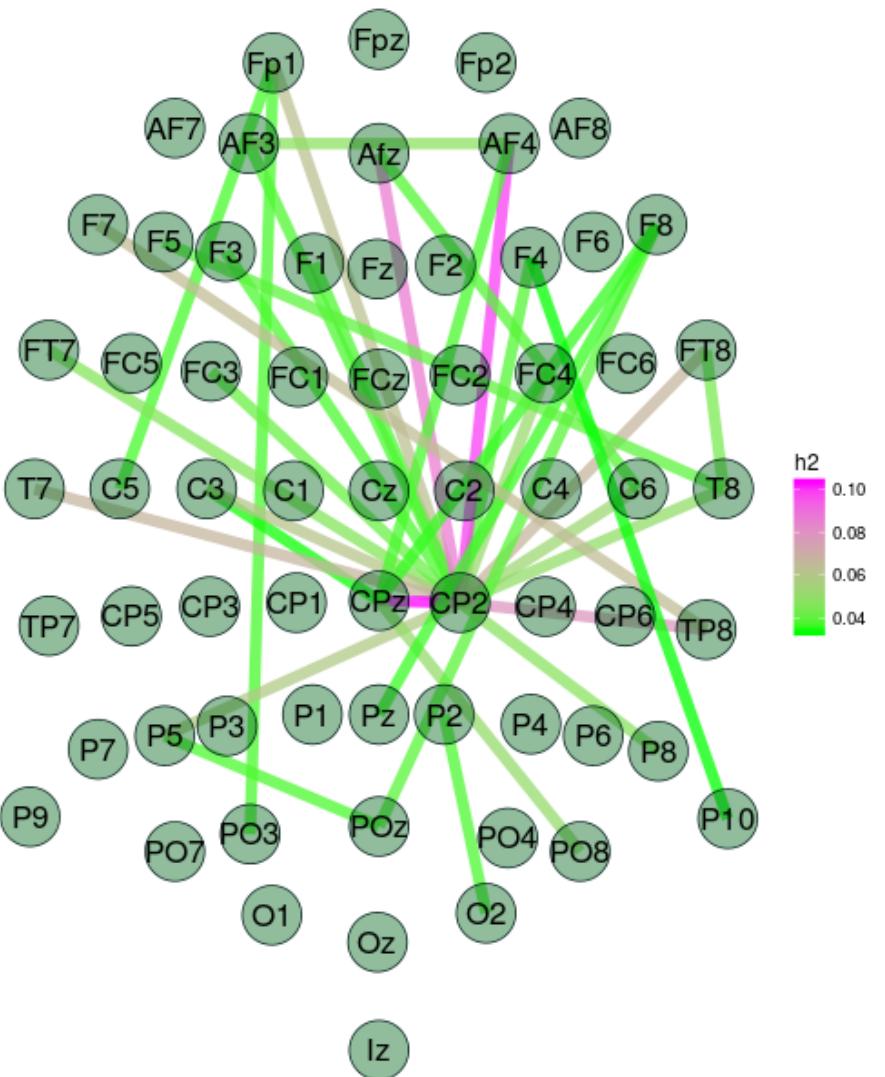
Test	P -value (Best λ)
α , All Channels	0.065 (5)
α , Frontal Channels	0.381 (1)
θ , All Channels	0.416 (0.5)
θ , Frontal Channels	0.085 (0.5)

Delta Coherence



$-\log(P)$

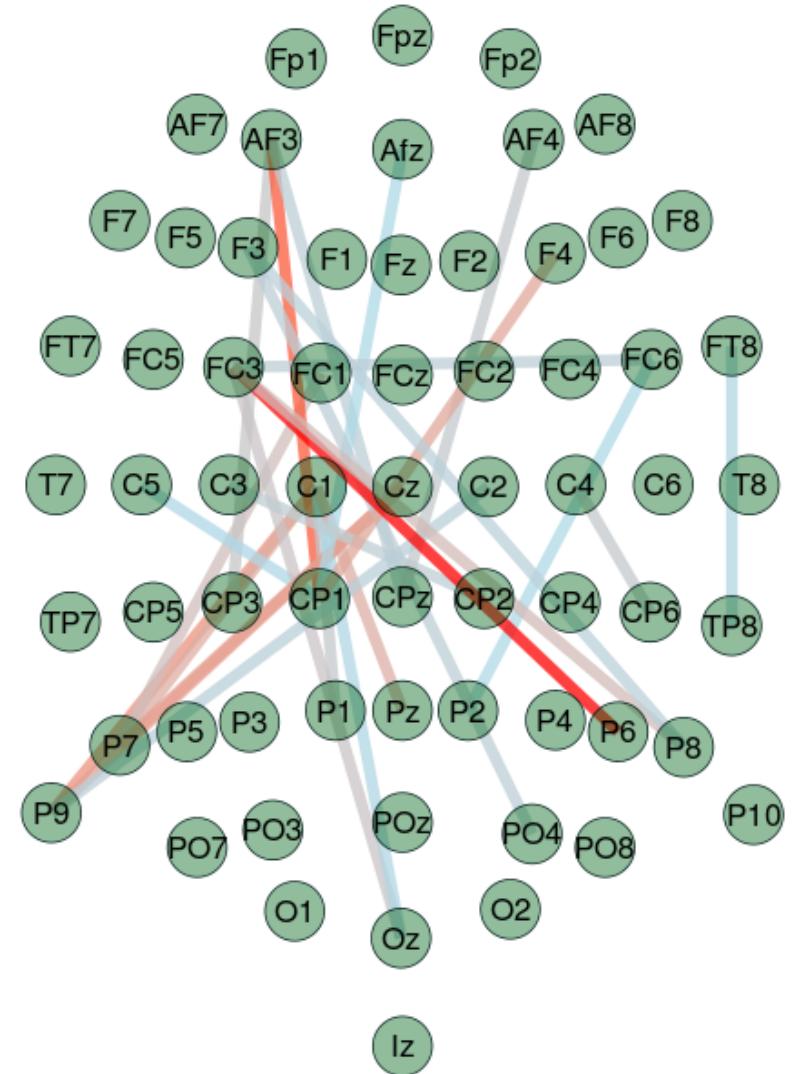
3.5
3.0
2.5
2.0



h^2

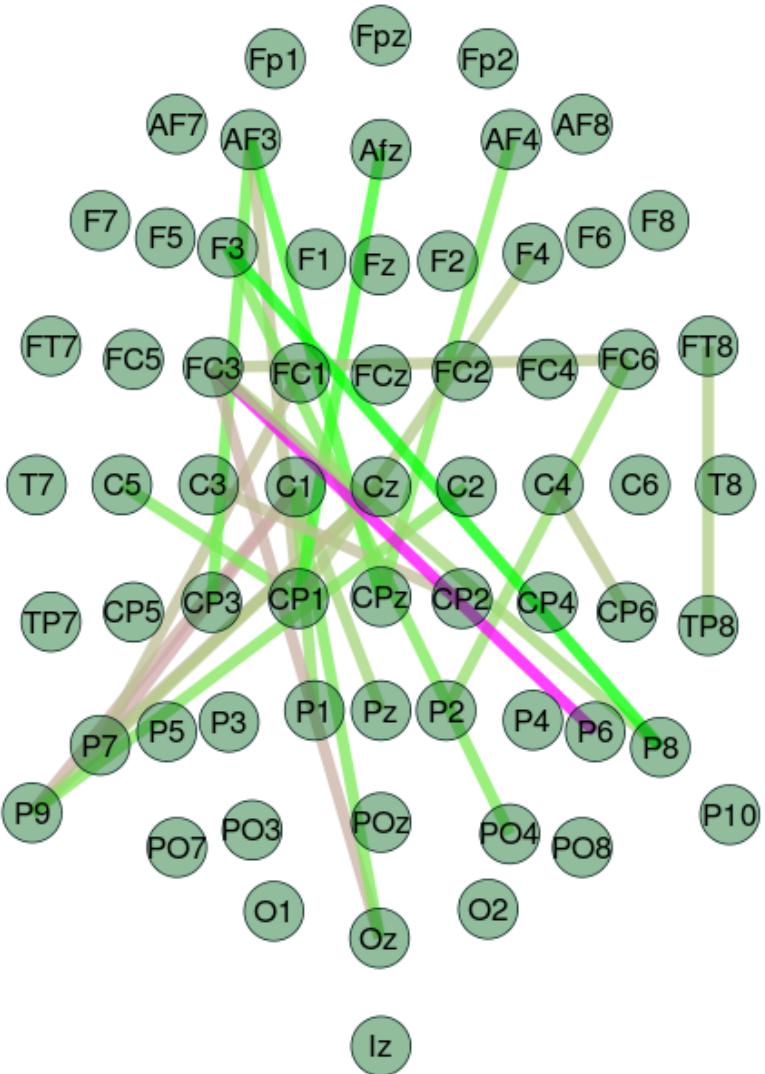
0.10
0.08
0.06
0.04

h2 Theta Coherence



$-\log(P)$

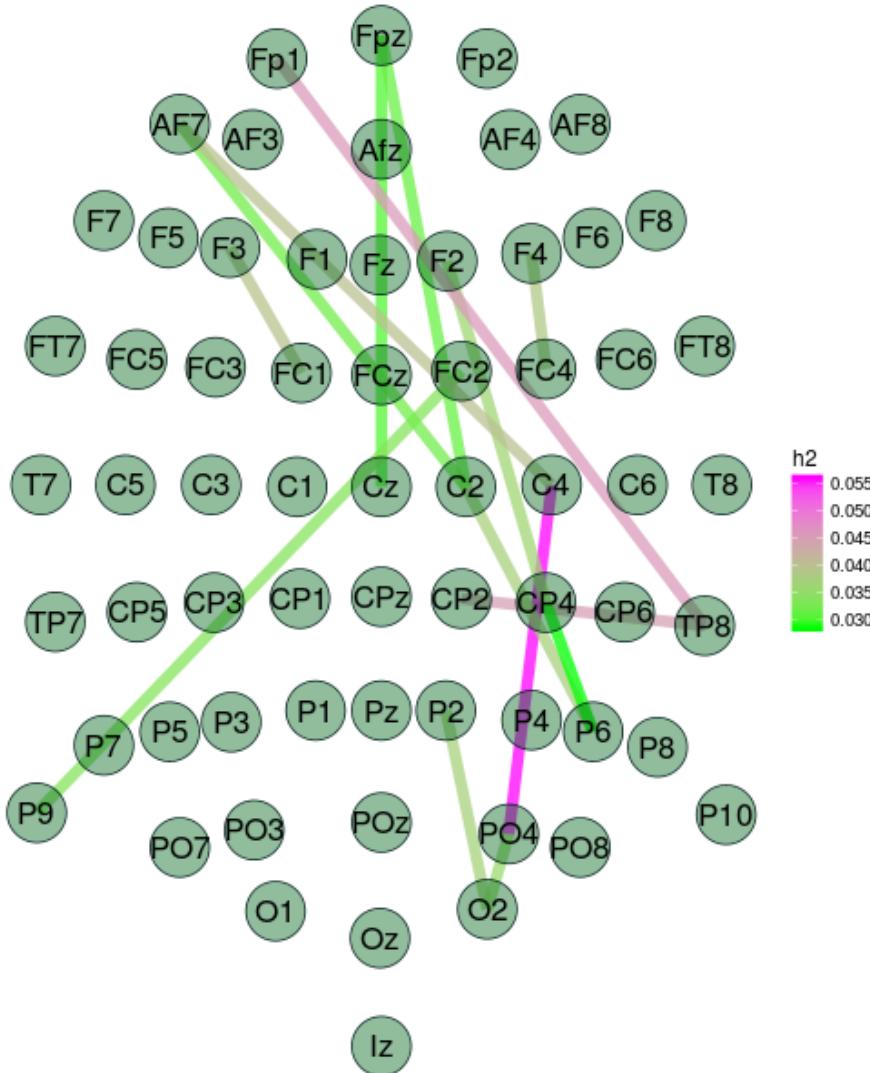
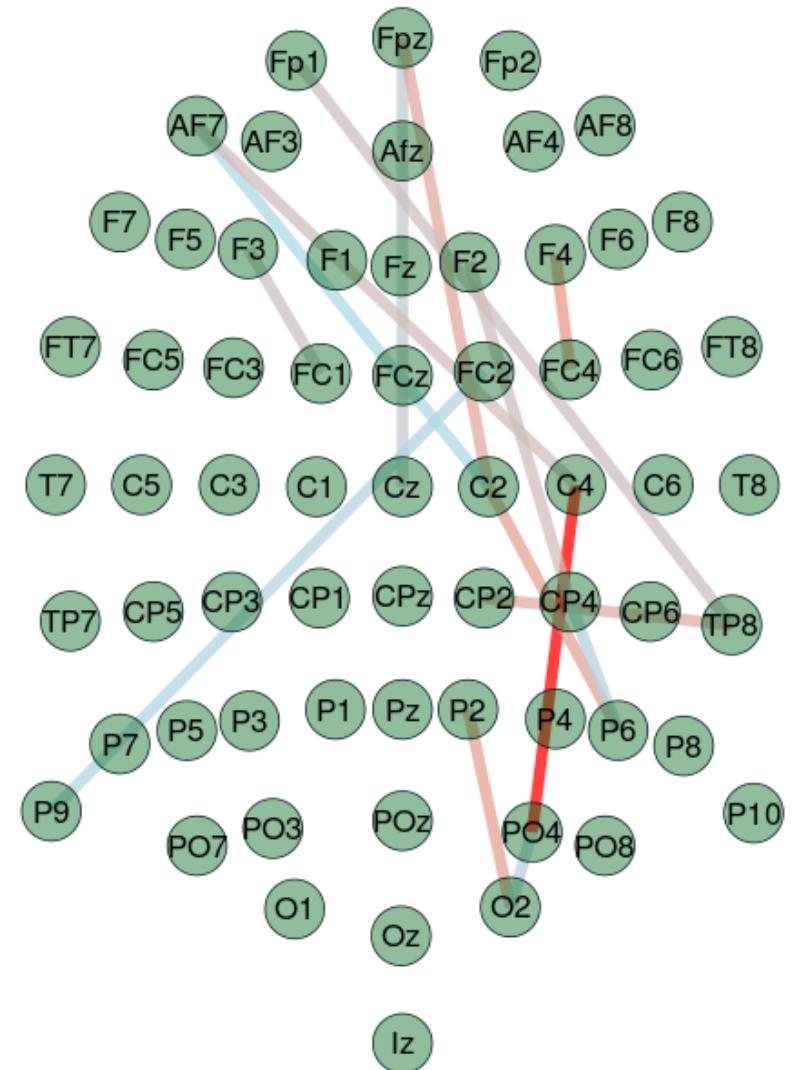
2.8
2.4
2.0



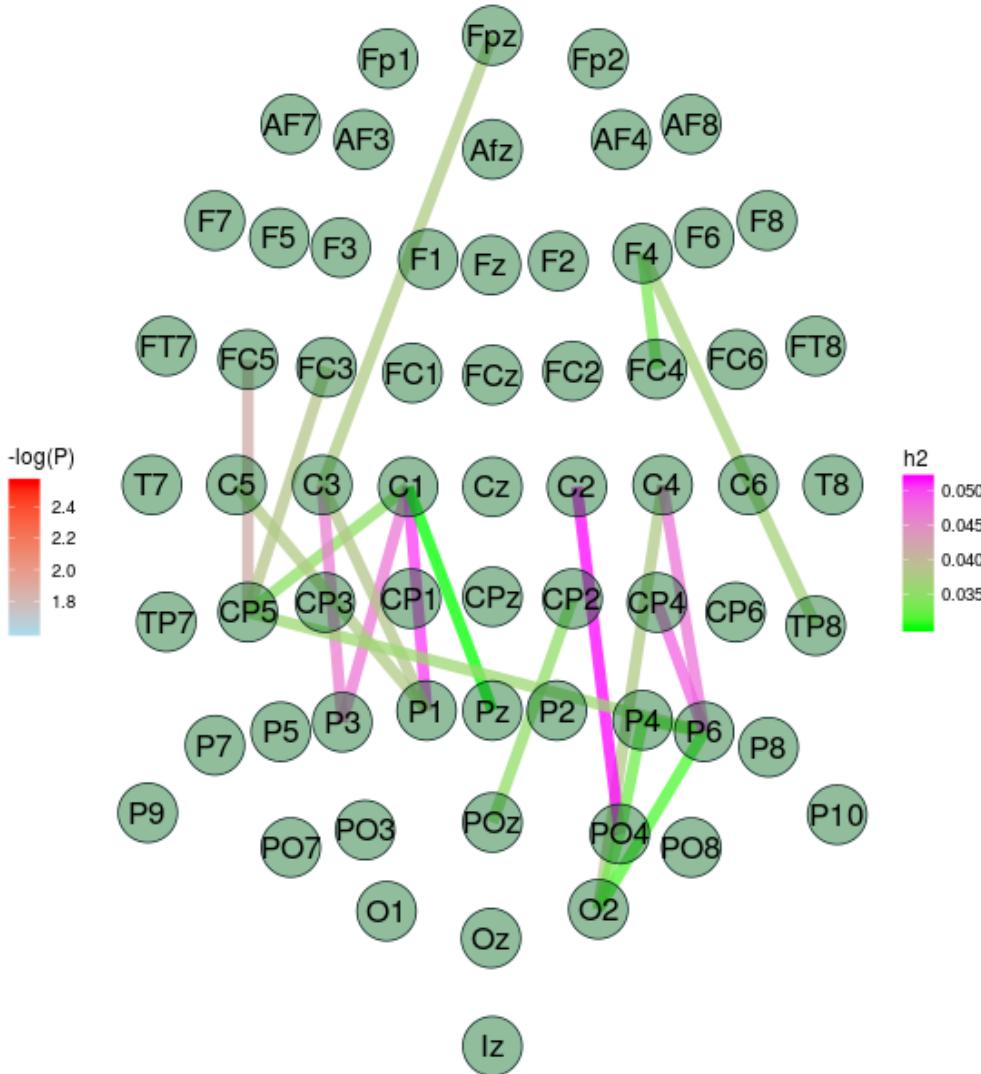
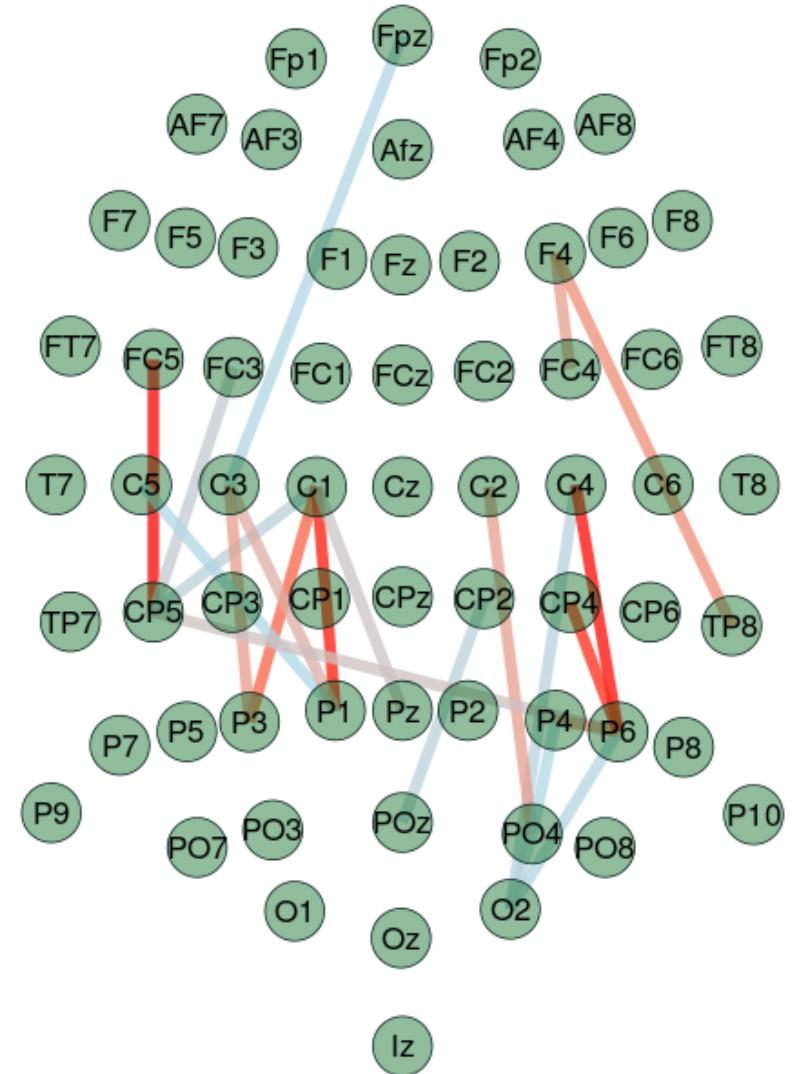
$h2$

0.06
0.05
0.04
0.03

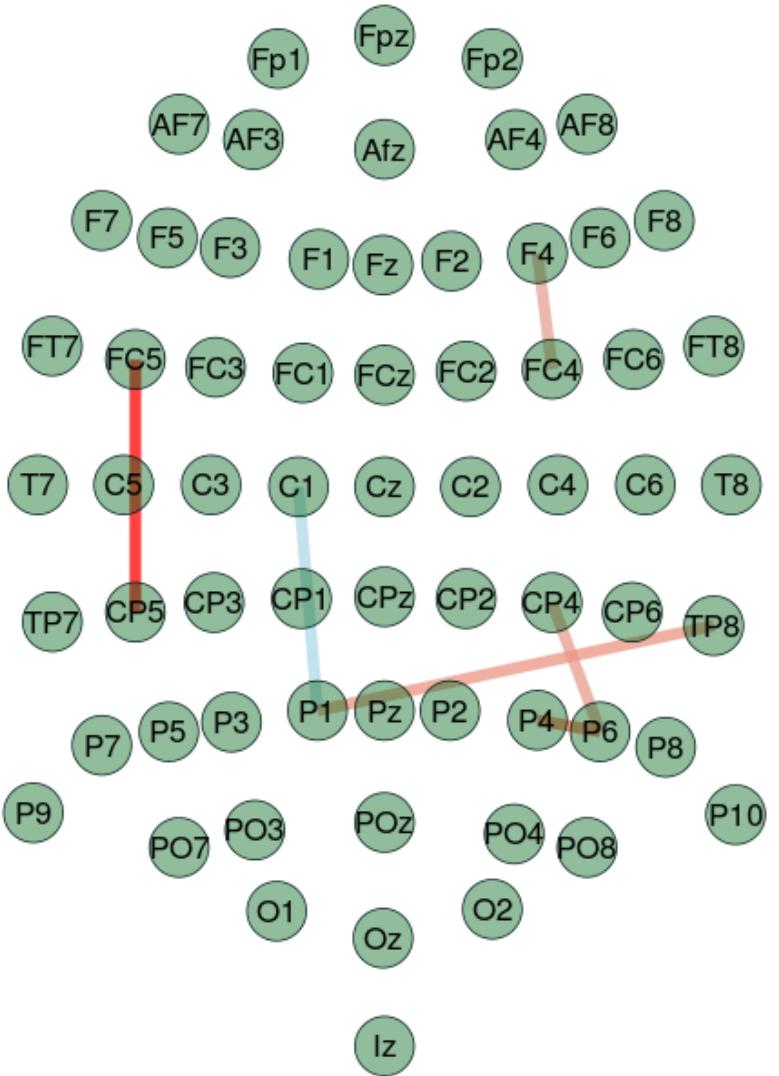
h2 Alpha Coherence



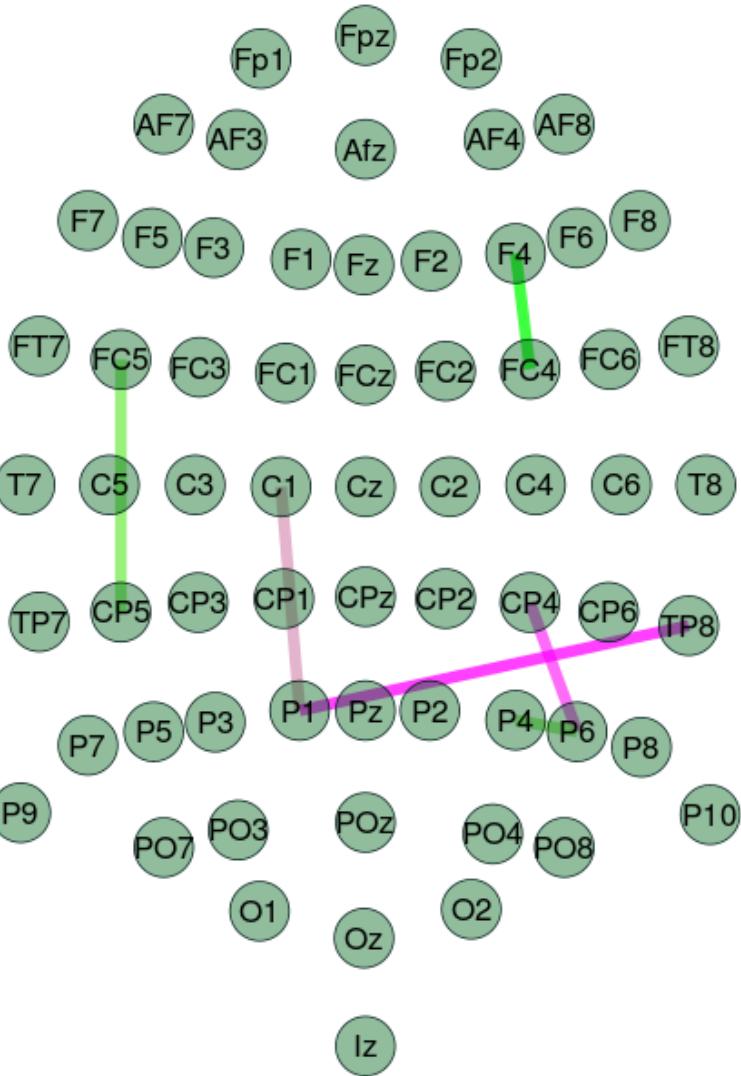
h2 Beta Coherence



h2 Gamma Coherence



-log(P)
1.9
1.8
1.7



h2
0.038
0.036
0.034
0.032

Current and Future Work

Further Analysis of VWM Coherence

Goals

- **Testing and comparing coherence and genetic effects during different phases of the VWM experiment**, e.g. encoding vs comparison
- **Test other important SNP sets and connectivity measures:**
 - SNPs: Dopamine pathways, neurotransmitters, evolutionarily recently selected SNPs
 - Connectivity: Partial coherence, PDC, total variation distance
- **Modeling genetic effects on EEG dynamic connectivity**
 - Extending Chee-Ming's switching factor VAR approach
 - Bayesian approaches, e.g. Multi-output Gaussian Process model

Switching Factor VAR for Dynamic Connectivity

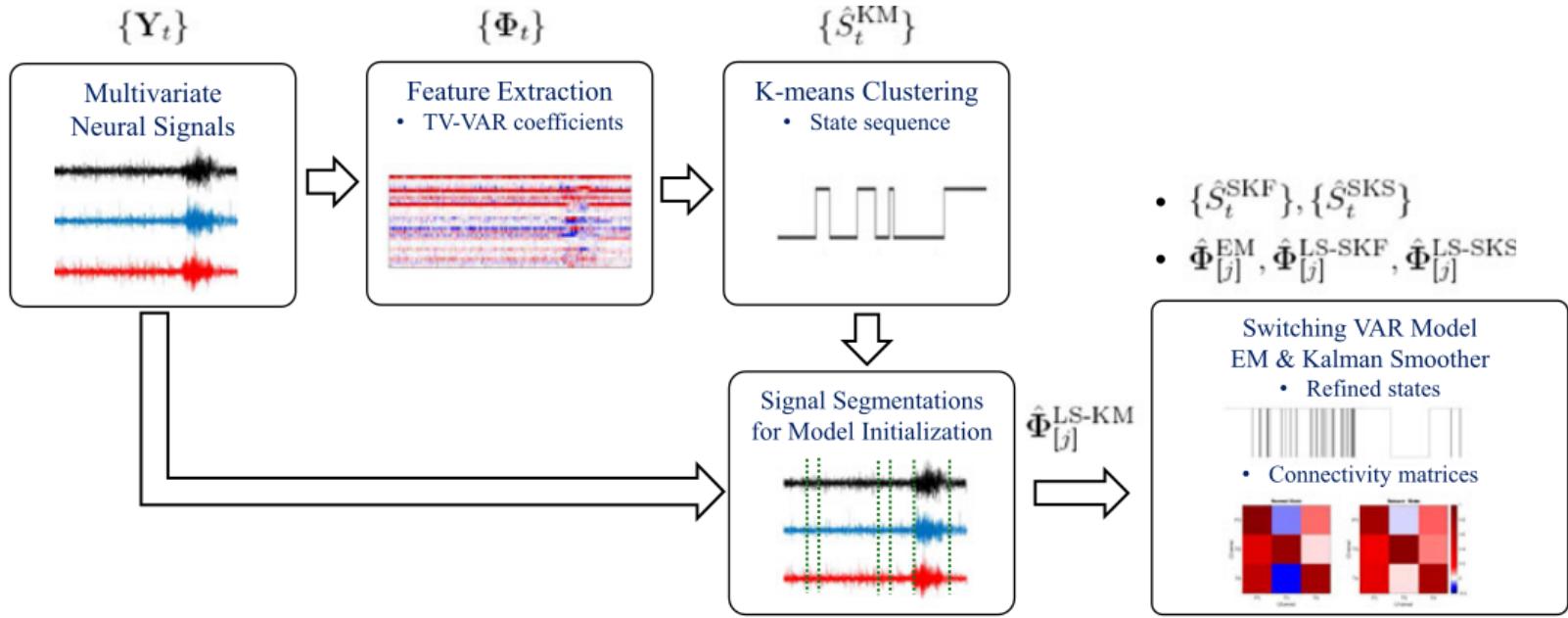
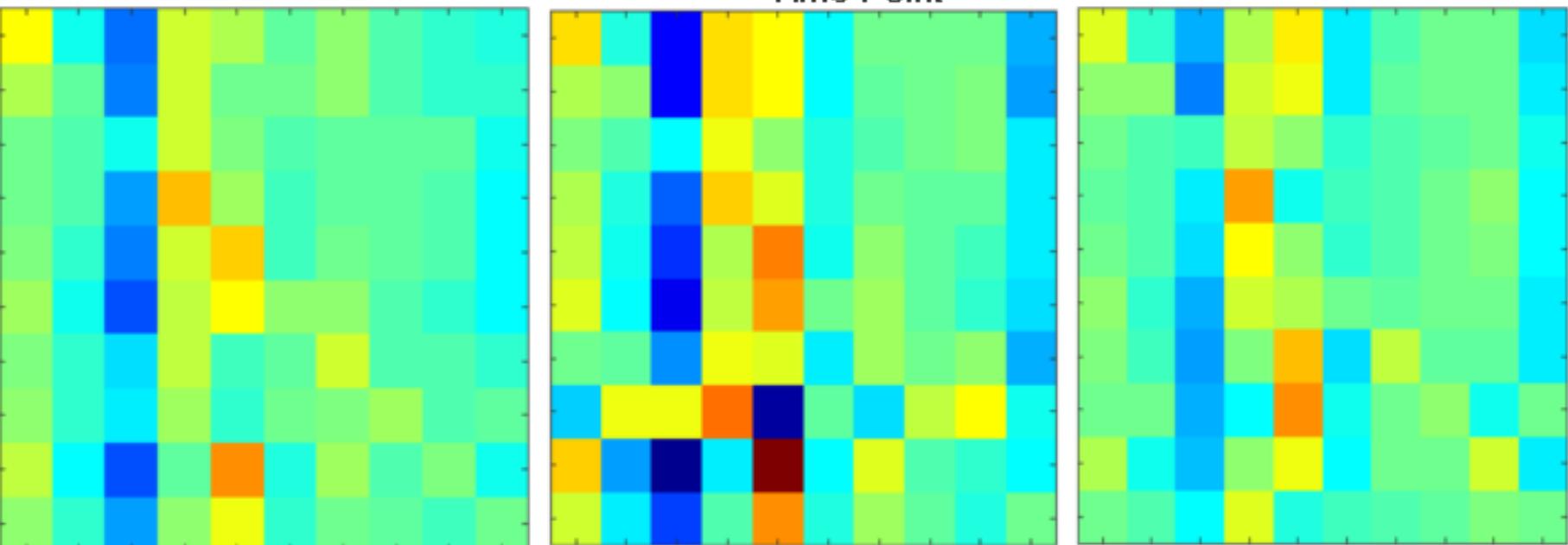
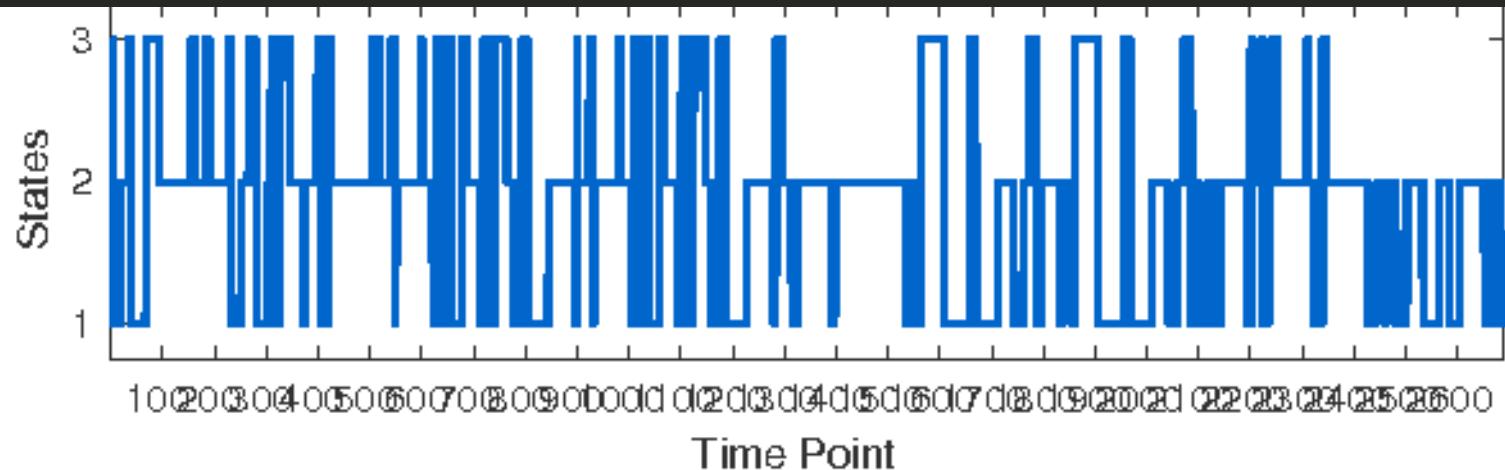


Fig. 1. Unified framework for estimating dynamic effective connectivity states in brain signals. The connectivity state estimation is initialized by K-means clustering of time-evolving connectivity features extracted using a TV-VAR process, and then refined by the SKF and EM algorithm based on a SVAR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

Switching Factor VAR for Dynamic Connectivity

Statistical Challenges

Choosing model parameters:

- **Lag Order**
 - AIC or Hannan-Quinn Criterion
- **Number of Clusters**
 - Scientifically informed (how many distinct states might we reasonably expect?)
 - Exploratory analysis
 - DP Mixture Model
- **Window size and Minimum Regime Length**
 - Scientifically informed (lower bound given by analysis targets, upper bound by experimental setting)
 - Exploratory Analysis
 - Selection via goodness of fit methods?

Switching Factor VAR for Dynamic Connectivity

Statistical Challenges

Incorporating Covariates like Genetics

- Requires common states across subjects
- How genetics may influence different features of connectivity is largely unknown, making it difficult to know which models will be scientifically sensible *a priori*.
- Model interpretation may be challenging

Multi-subject Dynamic Connectivity Modeling

Other Considerations

- DP Clustering instead of K -means
- Penalized fitting for TV-VAR to encourage smoothness over time
- Fixed design experiment and temporal alignment
- Fixed choice of basis states for scientific relevance and greater interpretability
- Can we learn the states? Inspiration from Tensor Regression...

References

- Wu M., et al. Kernel Machine SNP-set Testing under Multiple Candidate Kernels. *Genetic Epidemiology*. 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 $\text{rank}(X) = r \leq n$, 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 :

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

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

- 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