Tutorial: Topological Data Analysis on Dynamic Brain Networks

Moo K. Chung
University of Wisconsin-Madison
www.stat.wisc.edu/~mchung

Abstract

The tutorial introduces a data-driven topological data analysis (TDA) framework, designed to elucidate the state spaces in dynamically changing functional brain networks. This educational session will guide participants through fundamental concepts of TDA, moving towards a comprehensive understanding of how topological distance can be leveraged to cluster brain networks into distinct states without models. Special attention will be given to the incorporation of the temporal dimension of brain network data, utilizing the scalability of Wasserstein distance to provide a more nuanced analysis of network changes over time. Participants will gain in-depth experience with this method, learning why it is advantageous over traditional methods such as k-means clustering for estimating state spaces. The tutorial will delve into the intriguing investigation of if TDA is sensitive and flexible enough to determine the heritability of state changes. The tutorial is based on arXiv:2201.00087 (PLOS Computational Biology).

Grants: NIH U01NS093650, NS117568, EB022856, EB028753, MH133614, MH101504, P30HD003352, U54HD09025, UL1TR002373, NSF DMS-2010778, 2112455

Acknowledgement

Felipe Branco De Paiva, Camille Garcia Ramos, Zijian Chen, Tahmineh Azizi, D. Vijay Anand, Soumya Das, Tananun Songdechakraiwut, Vivek Prabharakaren, Veena A. Nair, Elizabeth Meyerand, Bruce P. Hermann, Aaron F. Struck, Ian C. Carroll, H. Hill Goldsmith, Seth Pollack, Richard Davdison
Univ. of Wisconsin-Madison

Jeffrey R. Binder, Medical College of Wisconsin (MCW)

Vince D. Calhoun TRenDs, Georgia State, Georgia Tech, Emory State University, Georgia

Grants: NIH U01NS093650, NS117568, EB022856, EB028753, MH133614, MH101504, P30HD003352, U54HD09025, UL1TR002373, NSF DMS-2010778, 2112455

Acknowledgement

James Gee, Li Shen University of Pennsylvania

Jamie Hanson University of Pittsburgh

Shih-Gu Huang, Anqi Qiu National University of Singapore
Anass El Yaagoubi Bourakna, Hernando Ombao KAUST, Saudi Arabia

Sunah Choi, Minah Kim, Hyekyoung Lee, Dong Soo Lee, Jun Soo Kwon *Seoul National University, Korea*
Jong Chul Ye, *KAIST, Korea*
Ilwoo Lyu, Jae-Hun Jung *POSTECH, Korea*

Grants: NIH U01NS093650, NS117568, EB022856, EB028753, MH133614, MH101504, P30HD003352, U54HD09025, UL1TR002373, NSF DMS-2010778, 2112455

Satellite meeting of OHBM

Workshop: NeuroImaging Statistics

SNU, Seoul, Korea, June 21-22, 2024

https://sites.google.com/view/neuroimaging2024

Organizers :

Moo K. Chung, University of Wisconsin-Madison

Inha Lee, Seoul National University

Tom Nichols, Oxford University

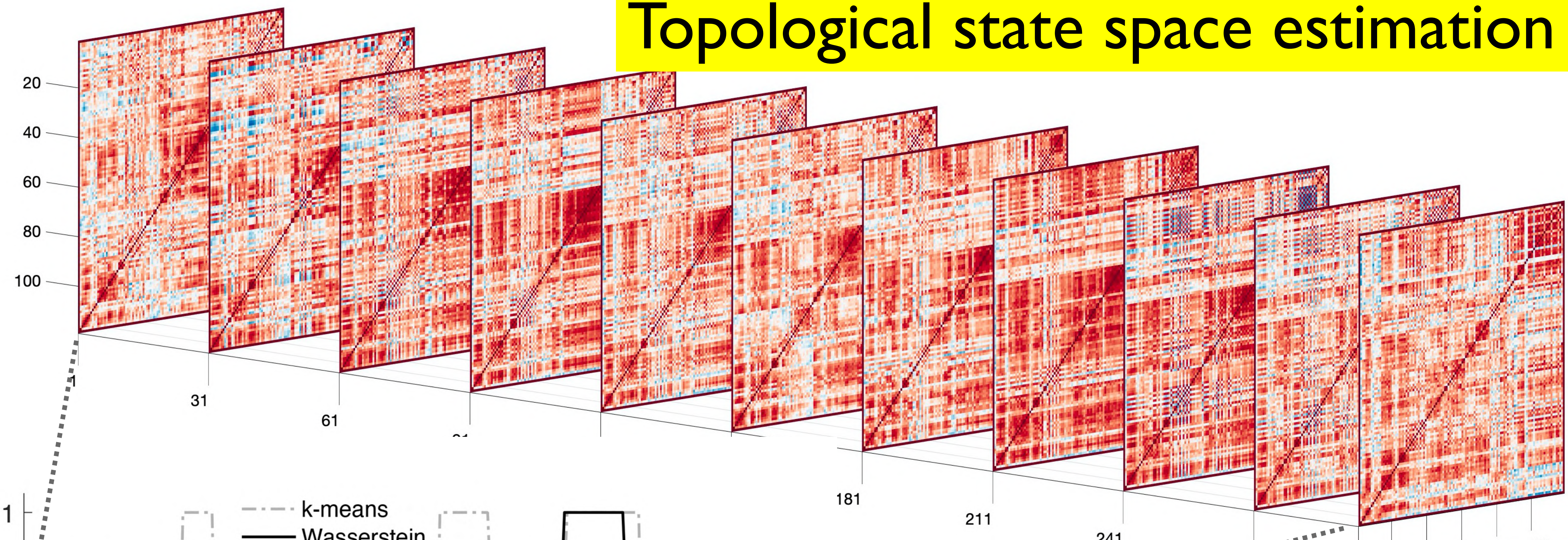
Hernando Ombao, KAUST

Jean-Baptiste Poline, McGill University

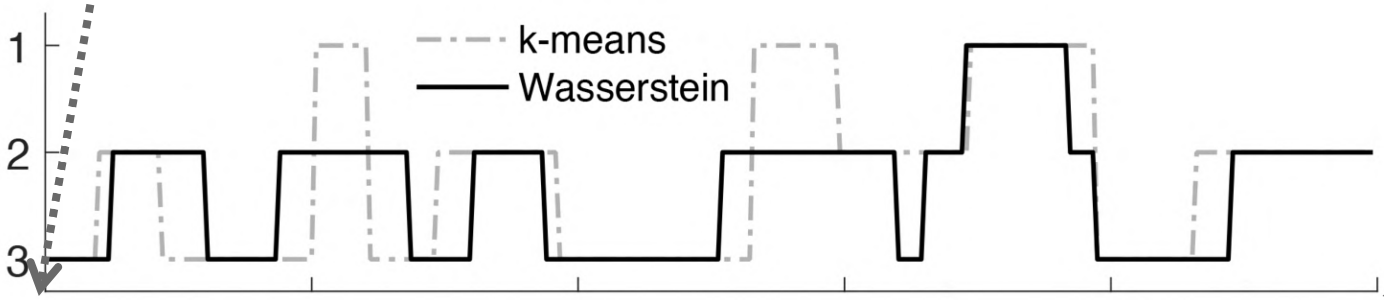
Anqi Qiu, Hong Kong Polytechnic University

Problem statement

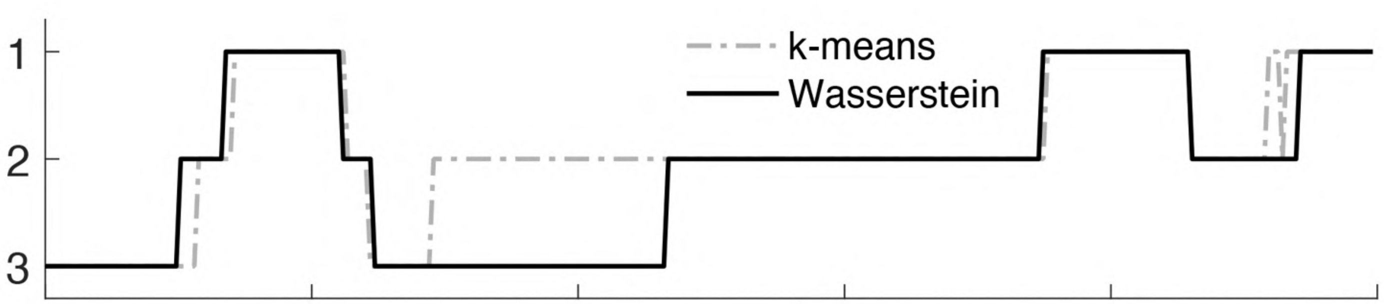
Topological state space estimation



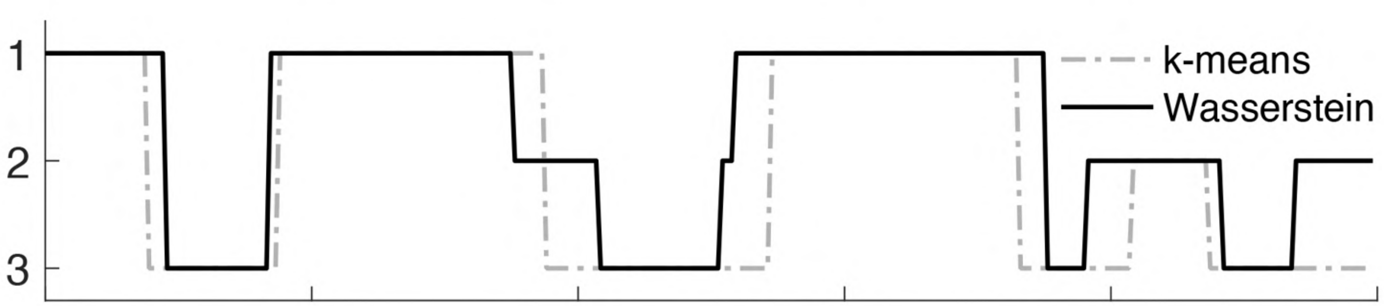
479 subjects



Subject 1



Subject 2



Subject 3

Is this make sense?

How many states?

Brain Imaging Data

T1-MRI

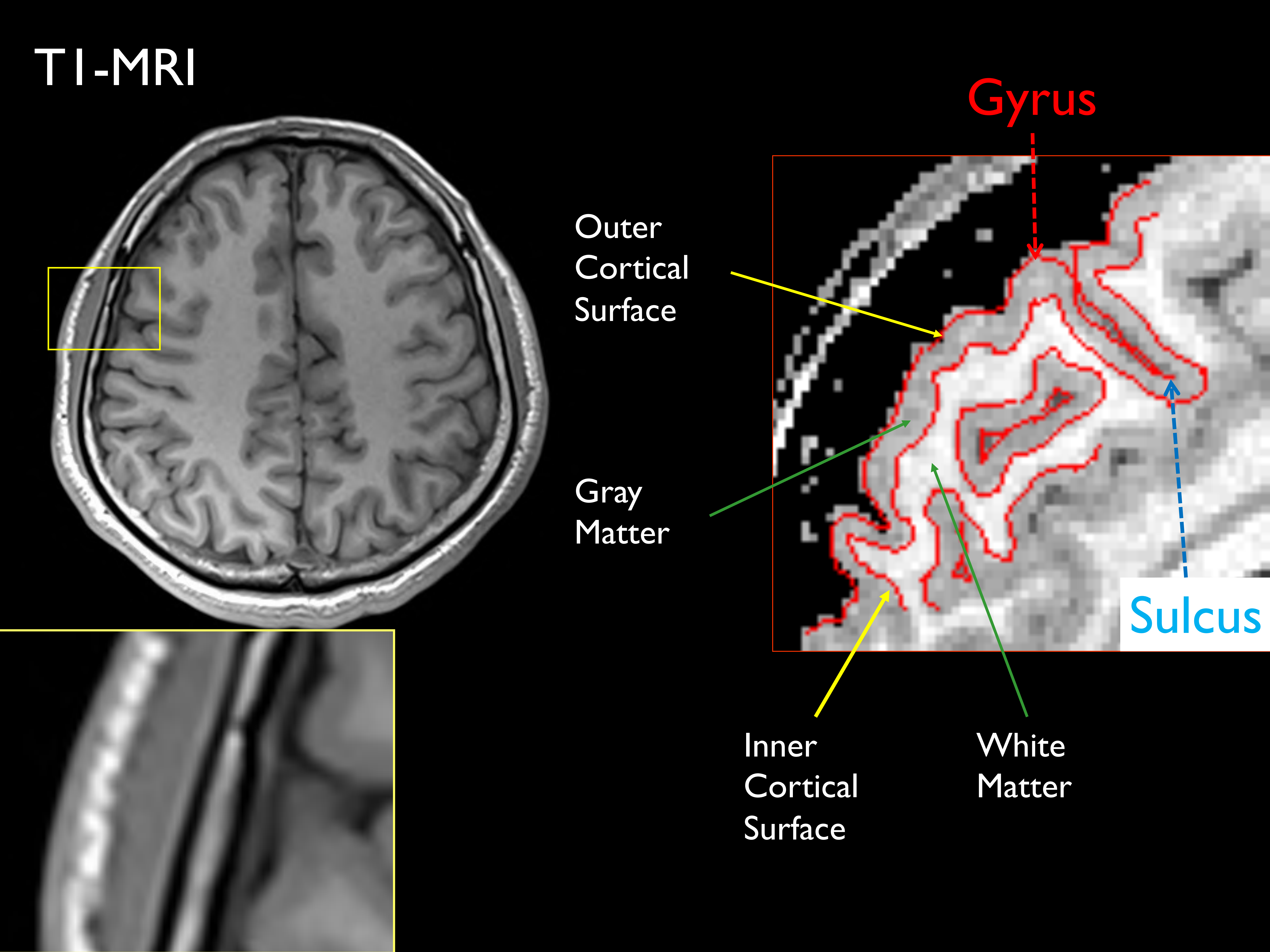
functional MRI

diffusion MRI

Magnetic resonance imaging (MRI)

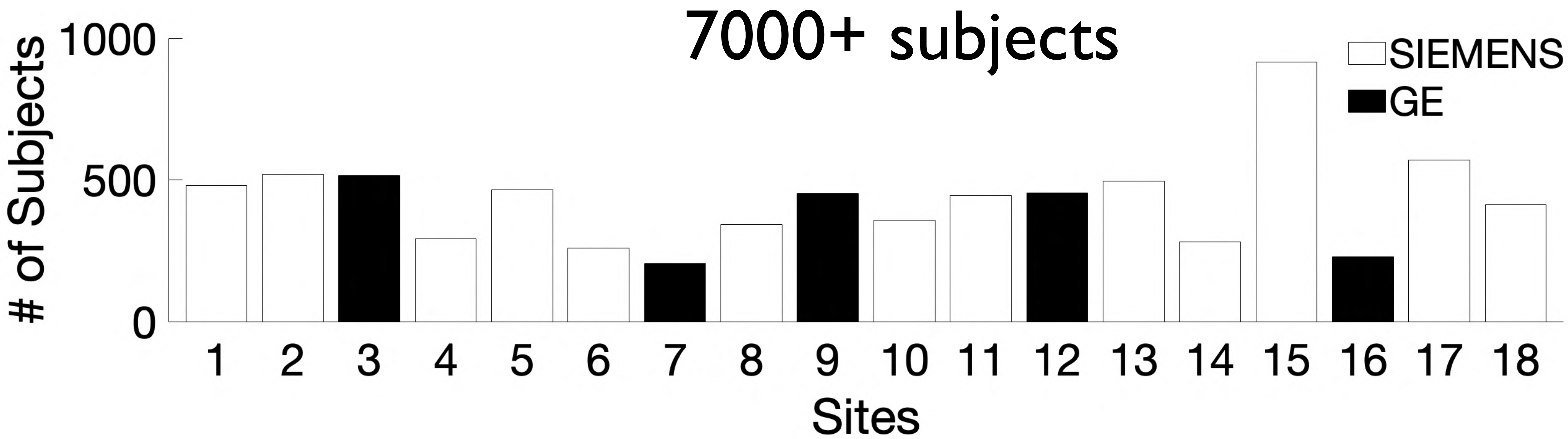
3T GE Discovery X750
Waisman Brain Imaging Laboratory
University of Wisconsin-Madison

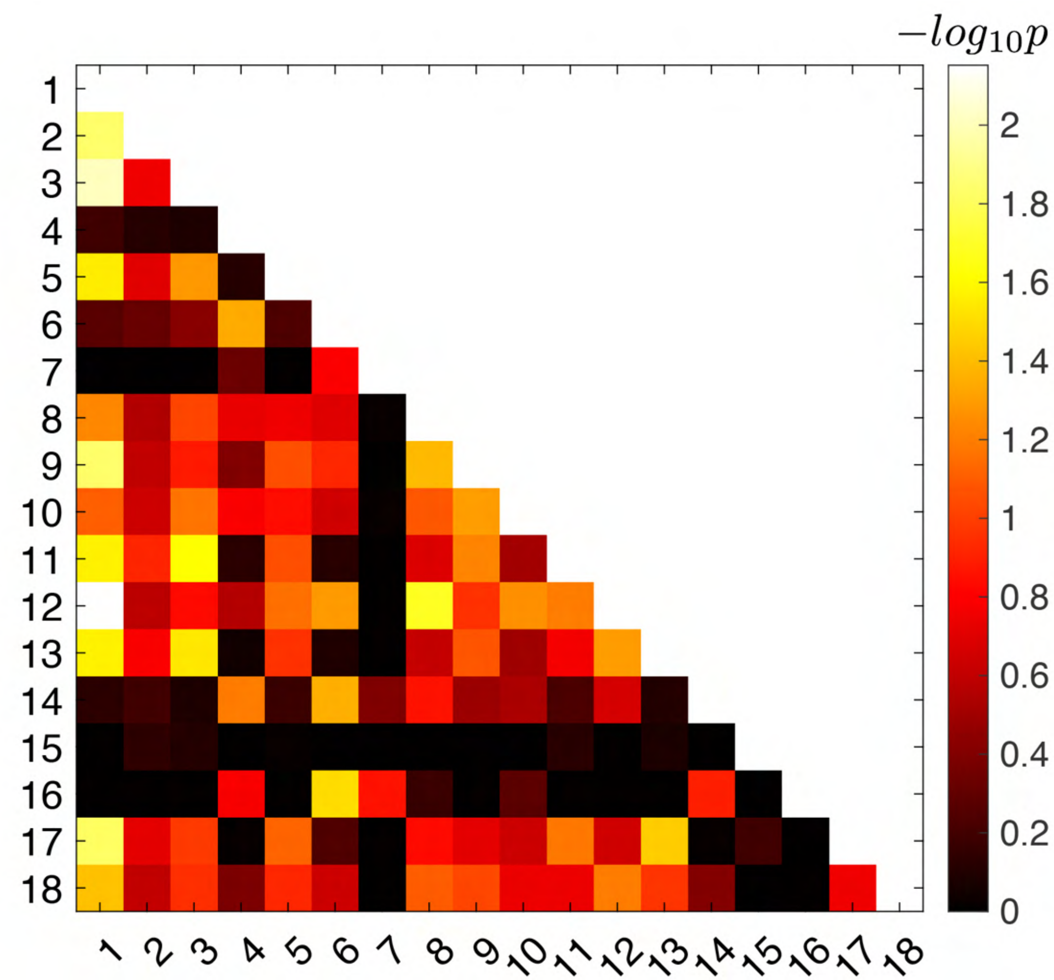
3T GE Discovery MR750
Center for Imaging Research
Medical College of Wisconsin, Milwaukee, WI

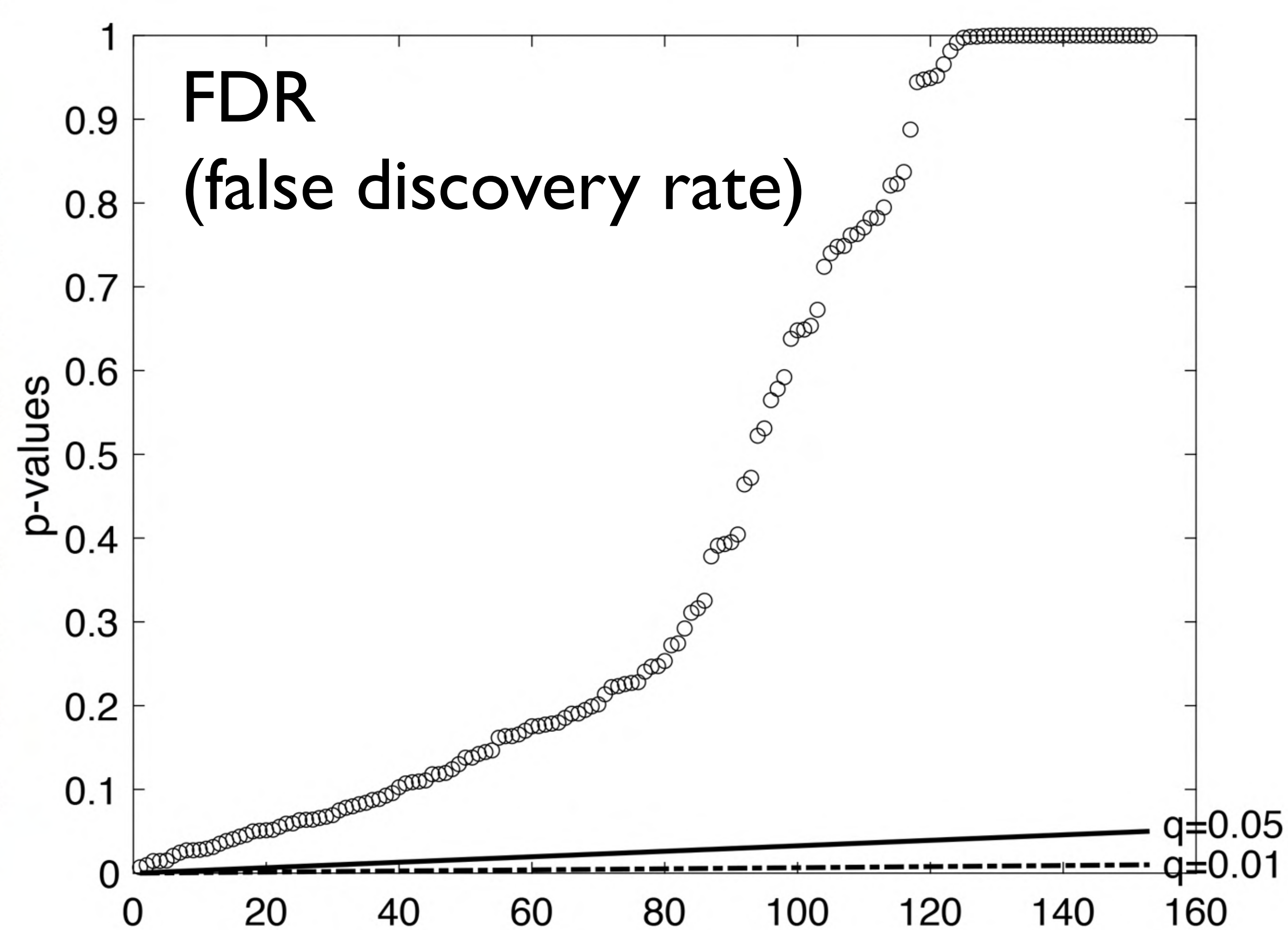


T1-MRI
Gyrus
Outer Cortical Surface
Gray Matter
Sulcus
Inner Cortical Surface
White Matter

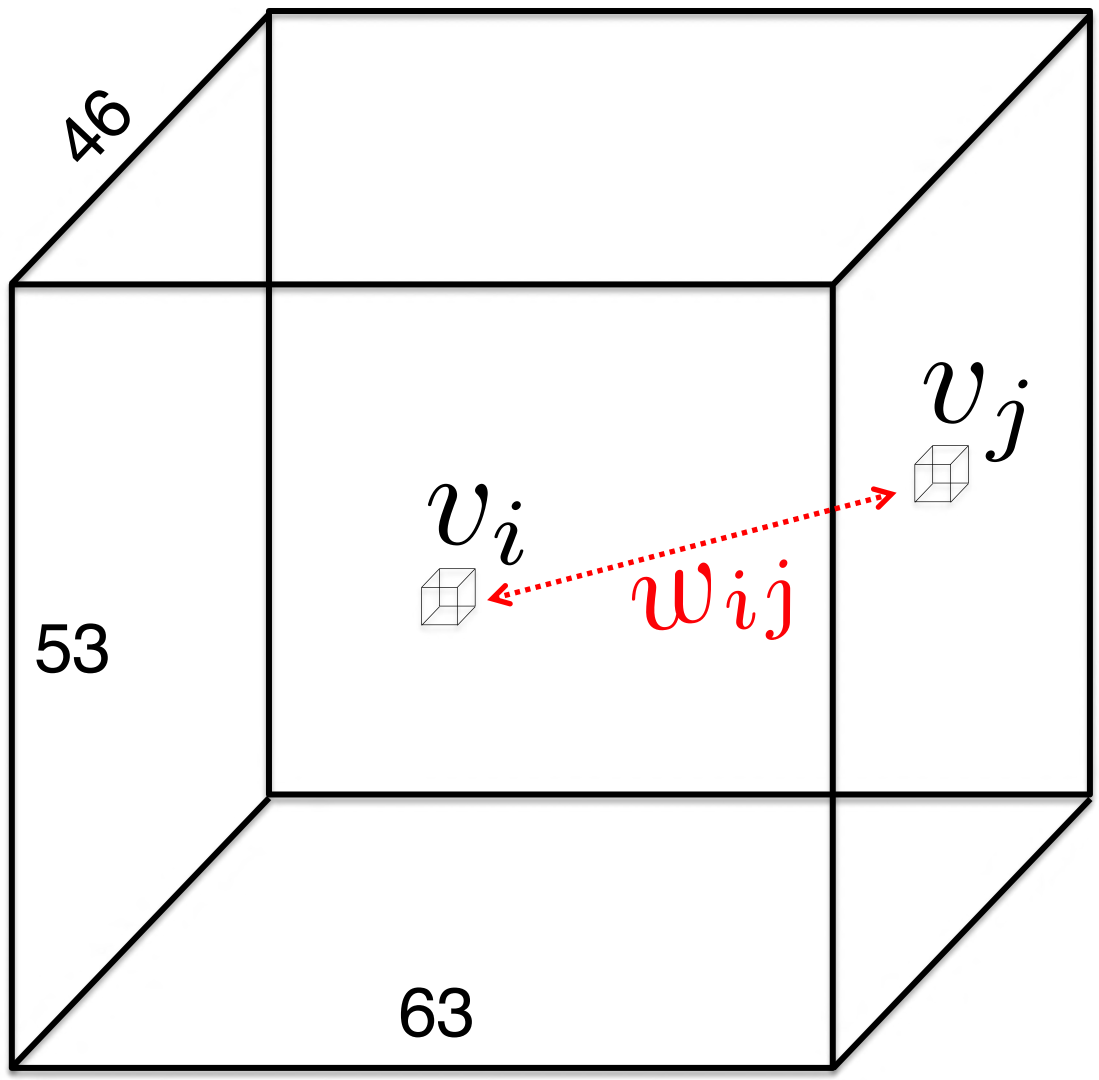
Topological methods will not detect site and sex effects – ABCD study



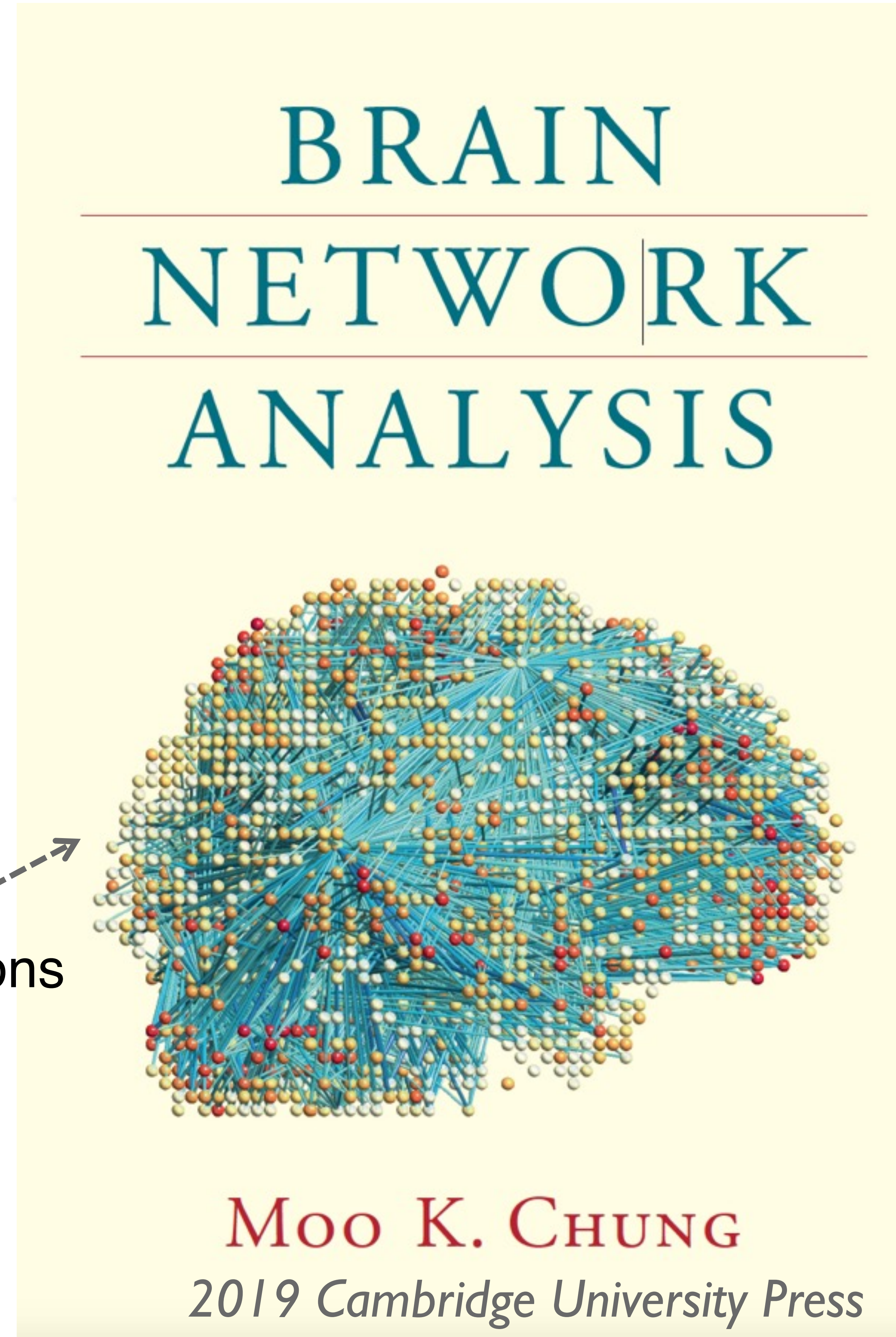




How big brain network data is?



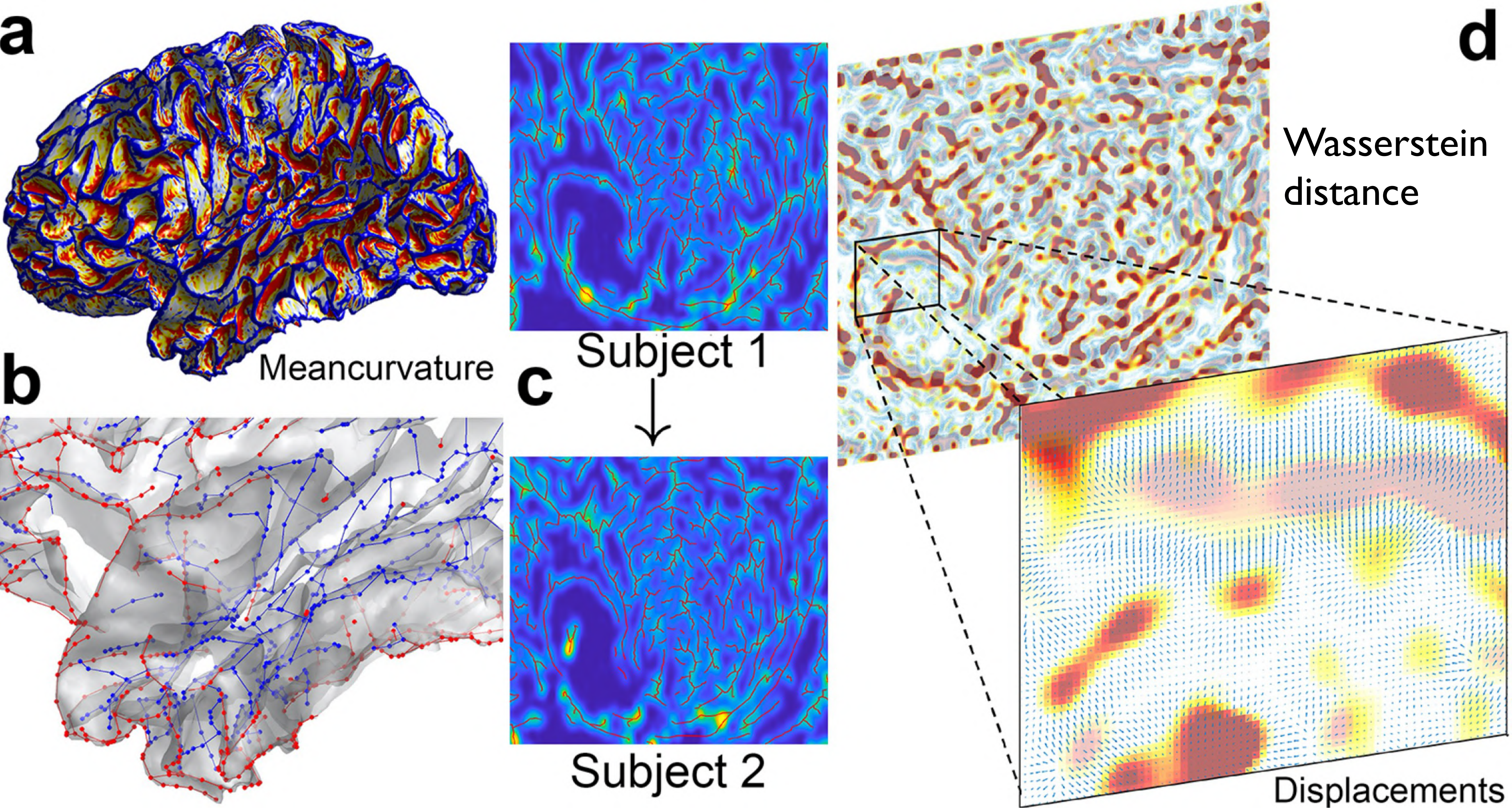
v_i



p=25972 voxels (3mm) in the brain
→ 25972 x 25972 = 0.67 billion connections
5.2GB memory

300000 voxels (1mm)
→ 90 billion connections
→ 700 GB memory

T1-MRI → Sulcal and gyral trees on cortical manifolds

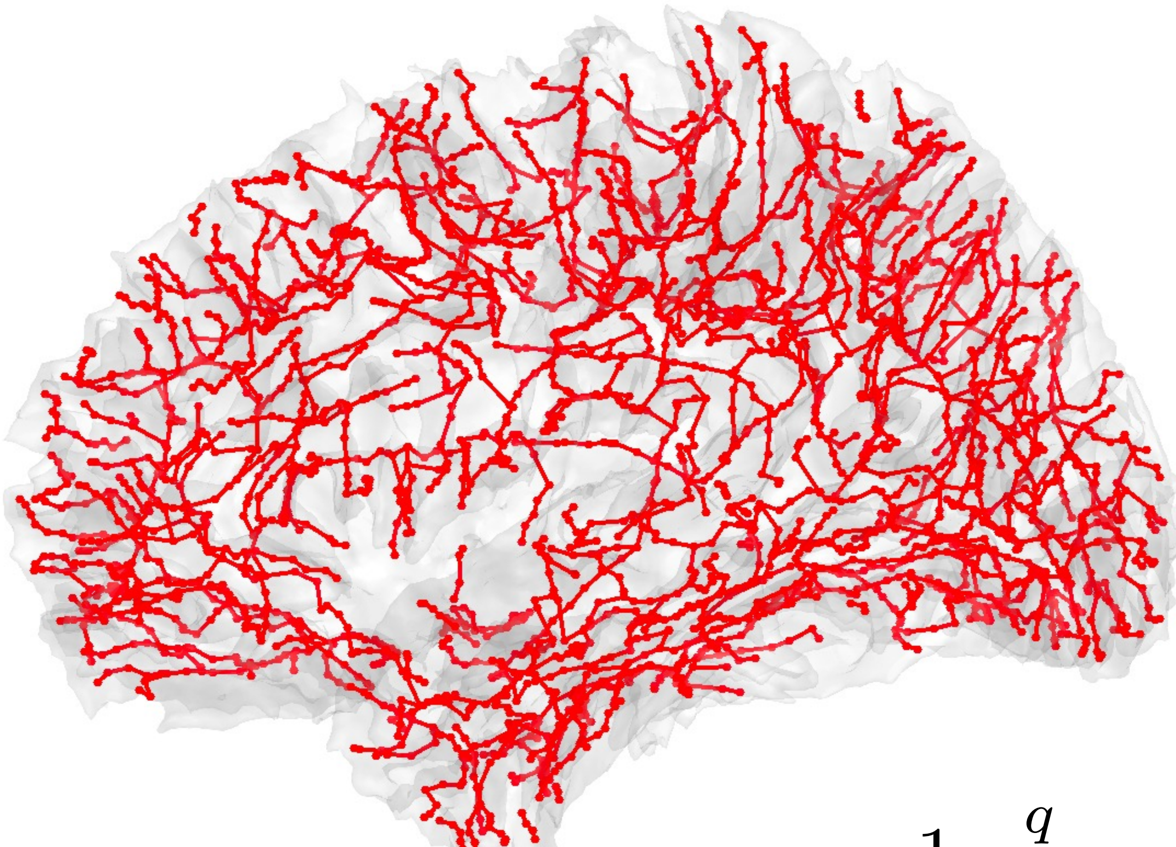


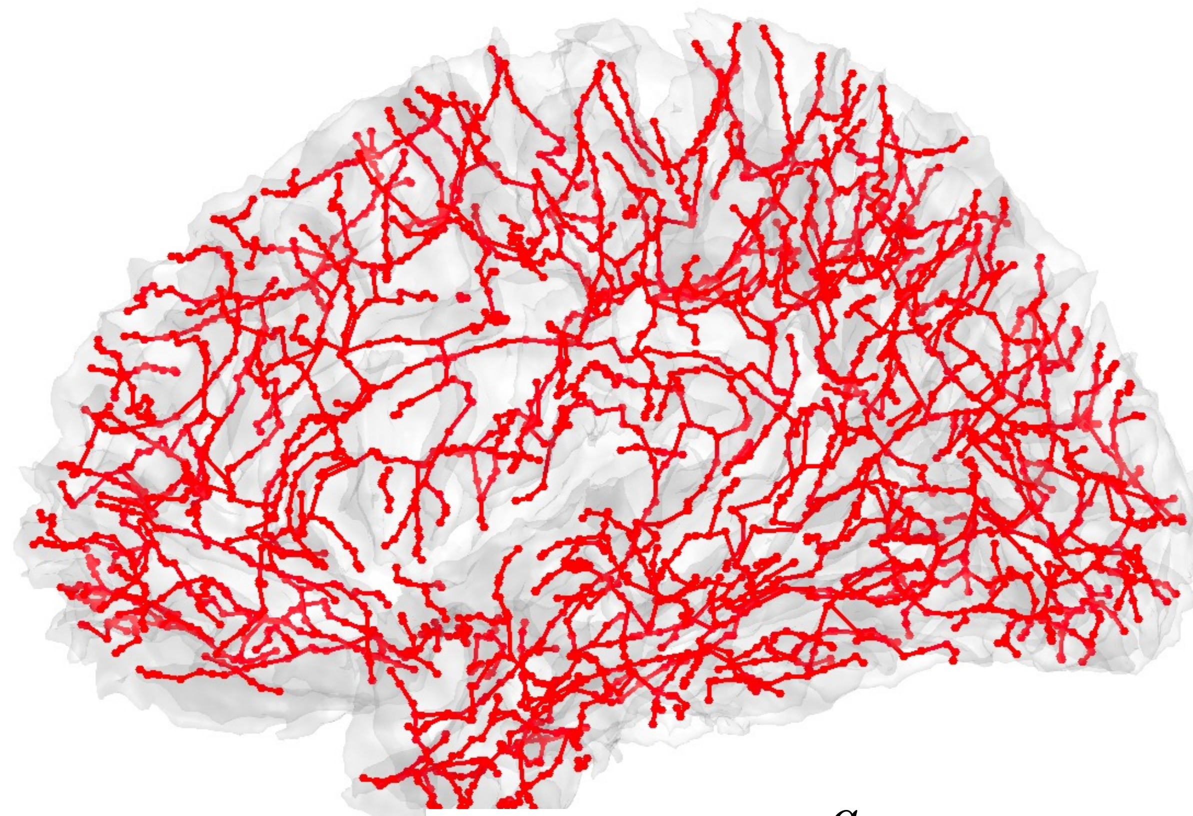
Huang et al. 2020 IEEE Transactions on Medical Imaging
Chen et al. 2023 arXiv:2307.00385

2-Wasserstein distance between vertices of sucal graphs

Random variables: $X \sim f_1 \quad Y \sim f_2$

2-Wasserstein distance: $\mathcal{D}(X,Y) = \big(\inf \mathbb{E}\|X - Y\|^2\big)^{1/2}$





$$f_1(x) = \frac{1}{q}\sum_{i=1}^{q}\delta(x - x_i)$$

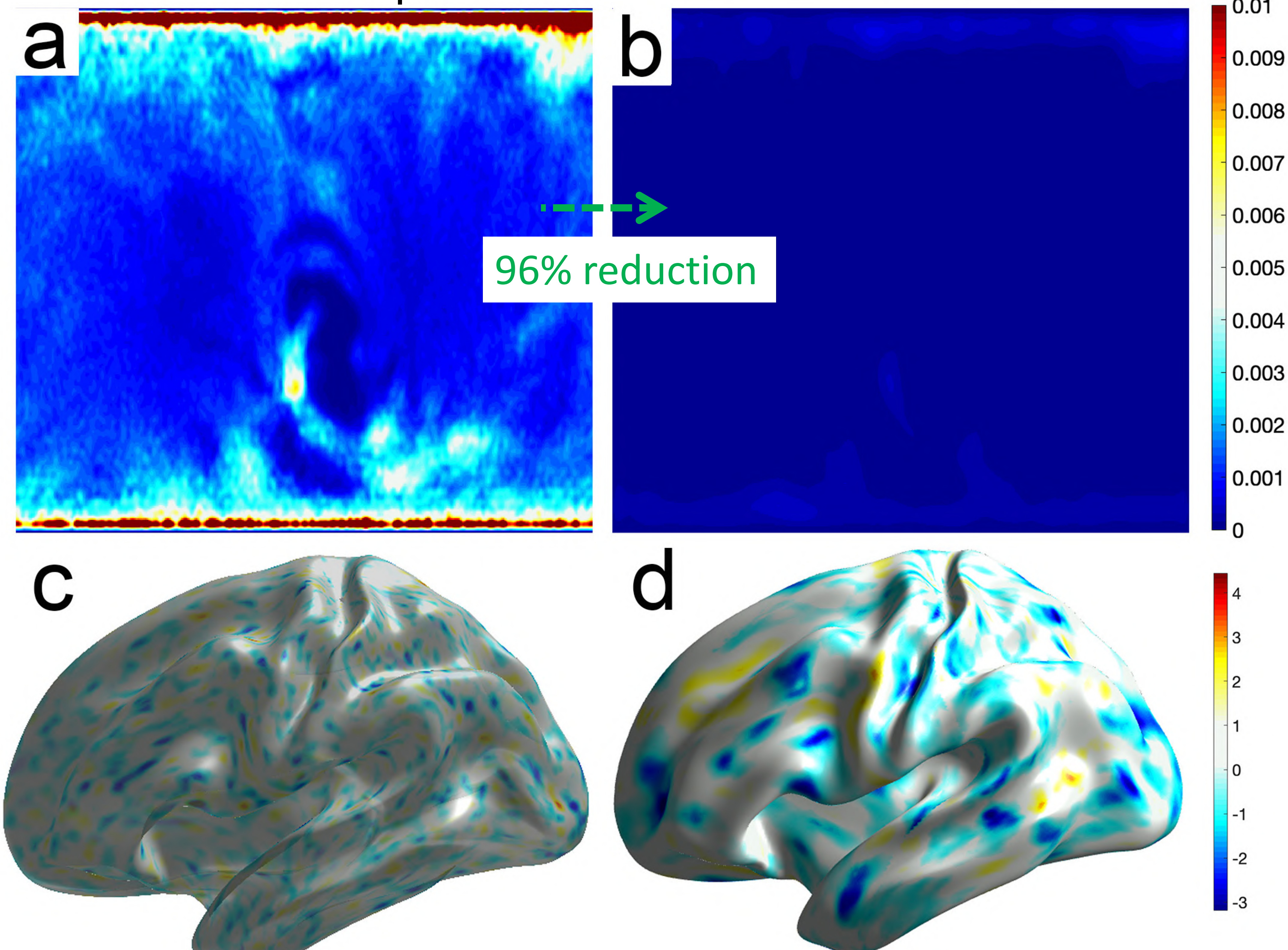
$$f_2(y) = \frac{1}{q}\sum_{i=1}^{q}\delta(y - y_i)$$

$$\mathcal{L}(P_1, P_2) = \inf_{\psi: P_1 \to P_2} \Big(\sum_{x \in P_1} \|x - \psi(x)\|^2\Big)^{1/2}$$

Hungarian algorithm in $\mathcal{O}(q^3)$

Intersubject variability in FreeSurfer output

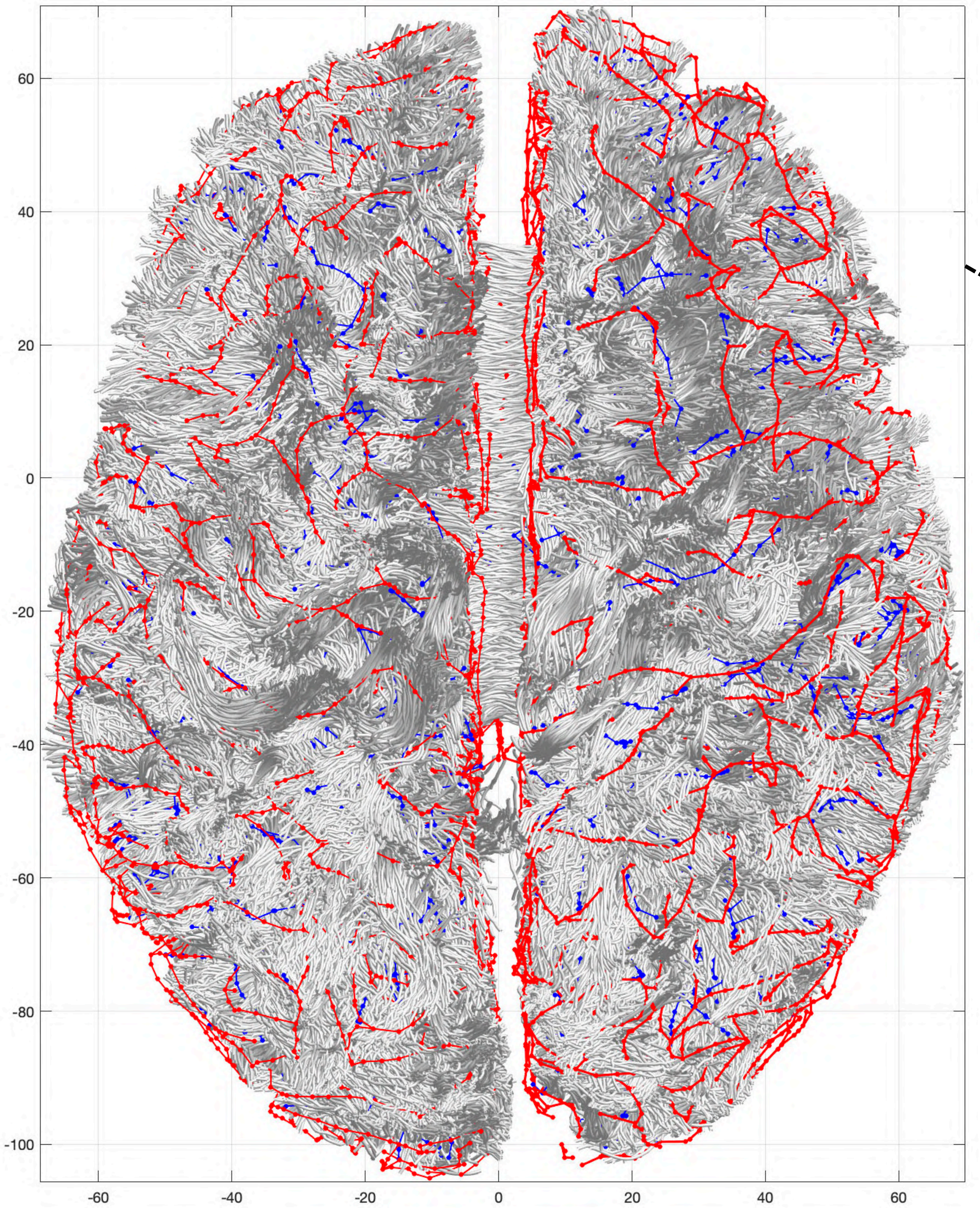
Intersubject variability after Wasserstein distance



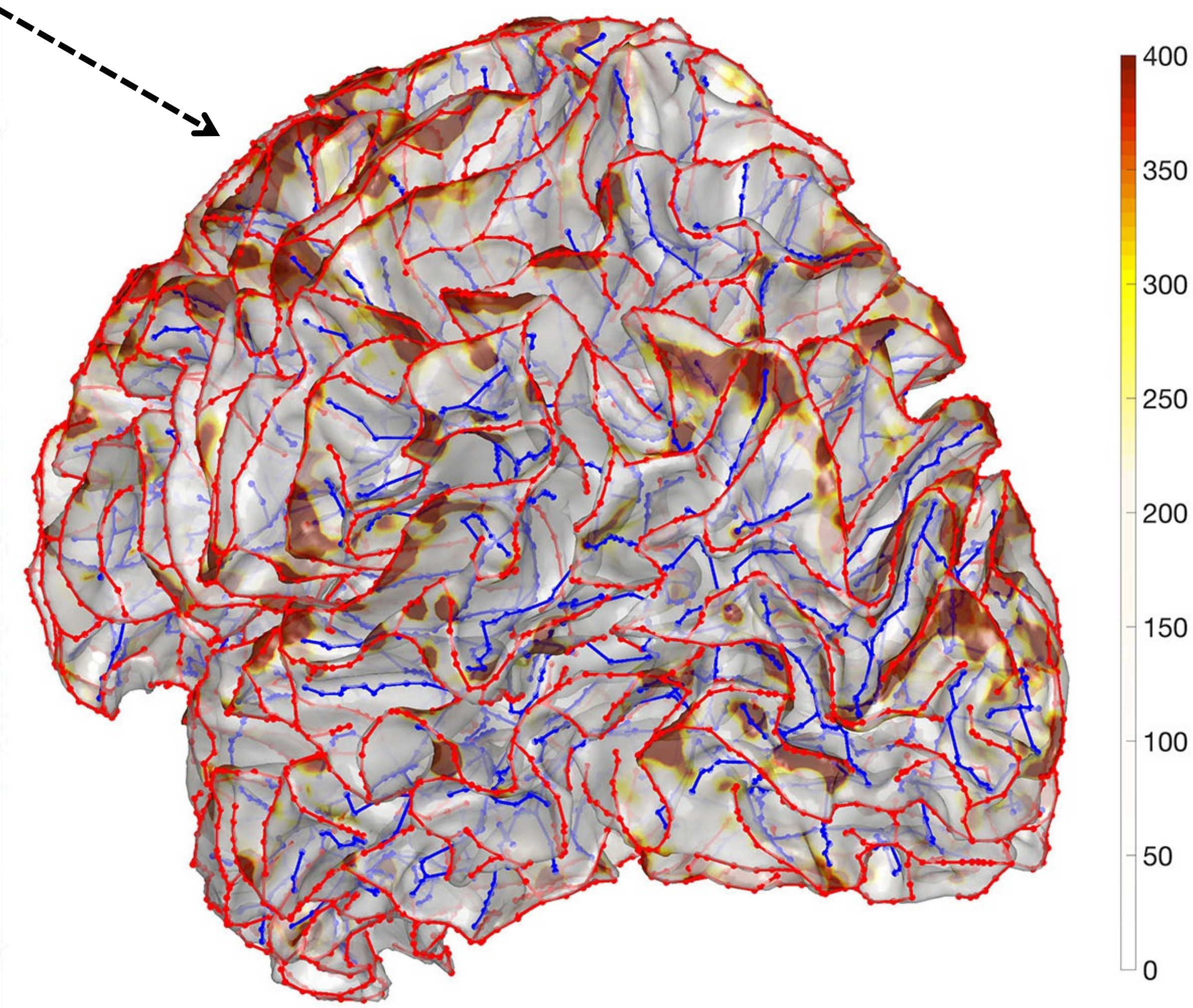
T-stat (274 females – 182 males)

Chen et al. 2023 *arXiv:2307.00385*

dMRI → 1 million white matter fiber tracts per subject

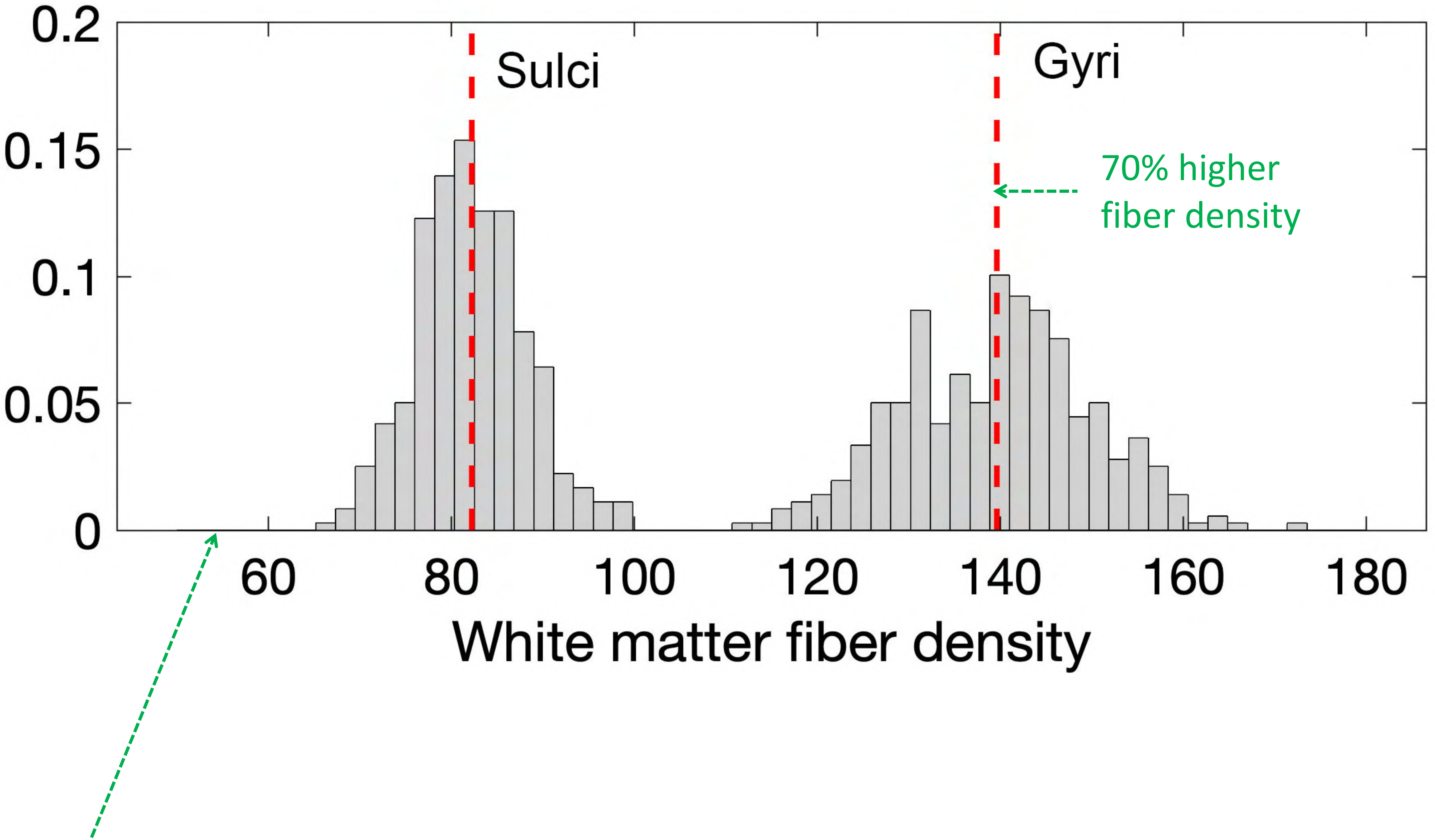


Fiber count within 2mm radius around nodes on trees



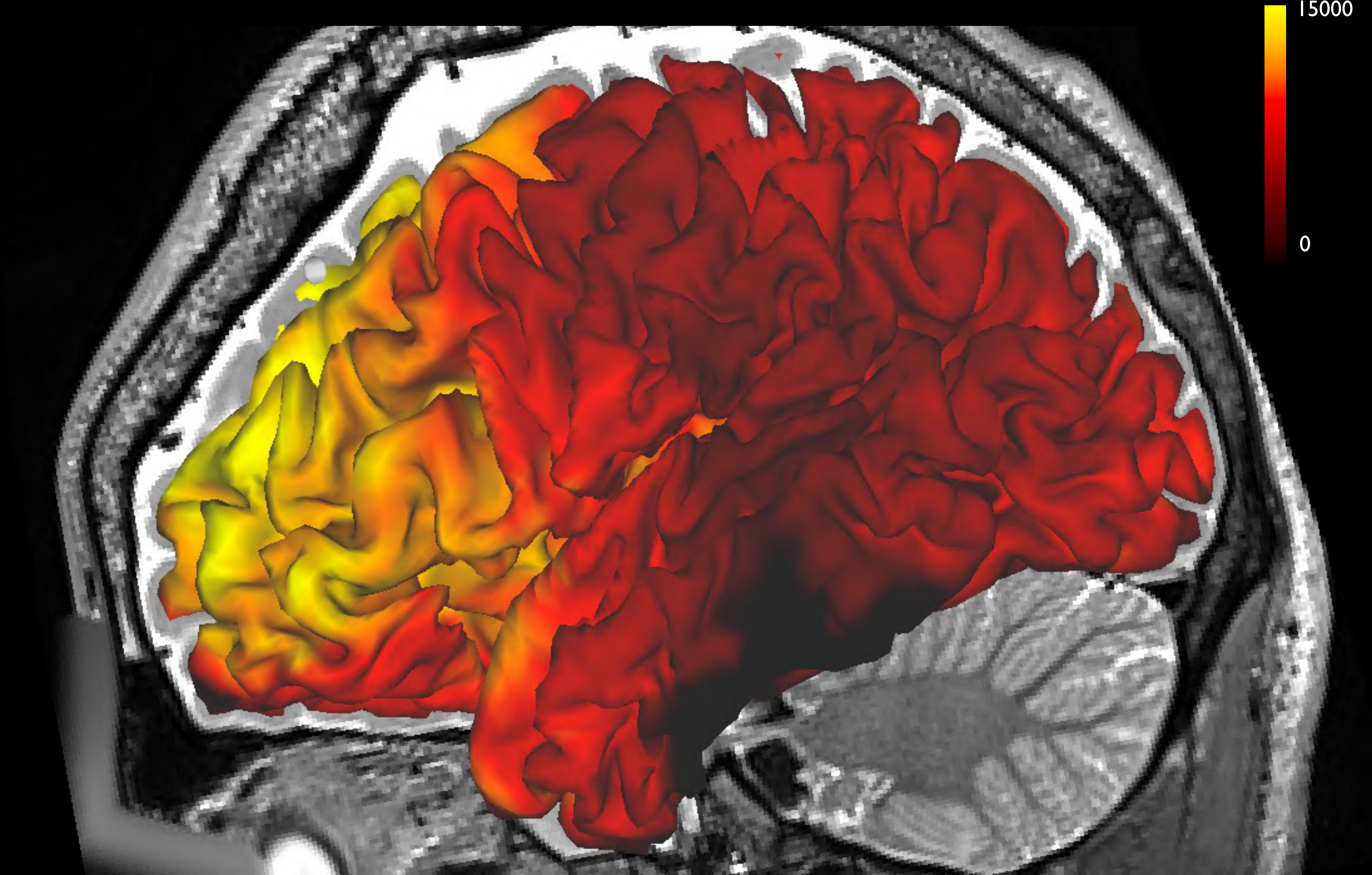
Chung et al. 2019 LNCS 11848:42-53

Differential structural connectivity between sulci/gyri

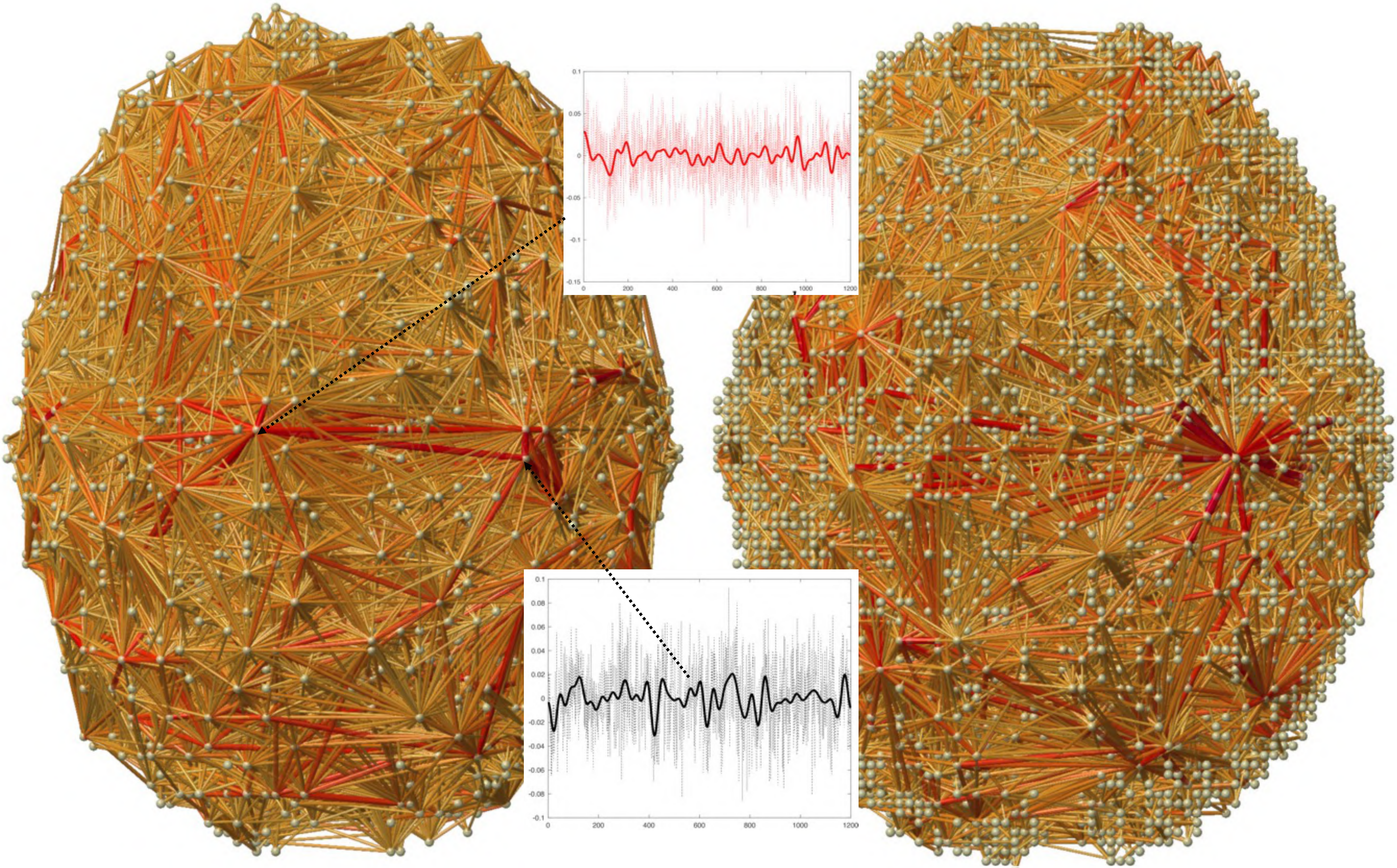


Distribution out of 358 subjects

rs-fMRI (every 30 second)

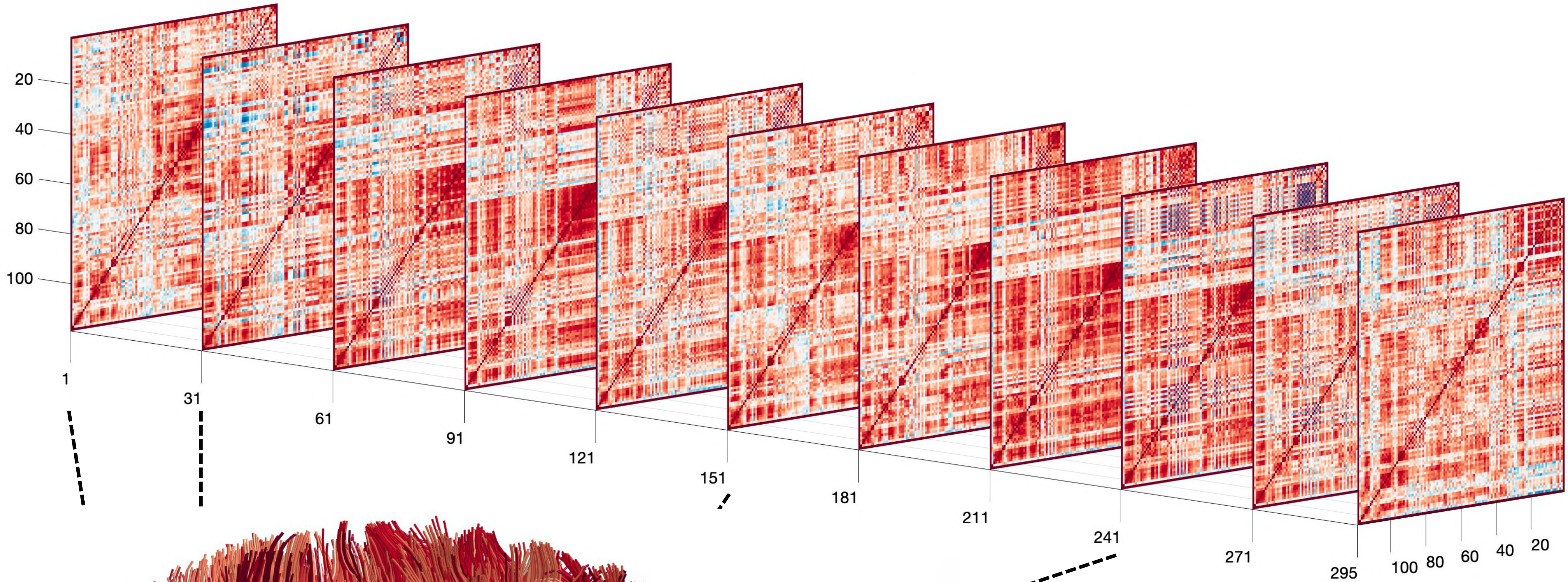


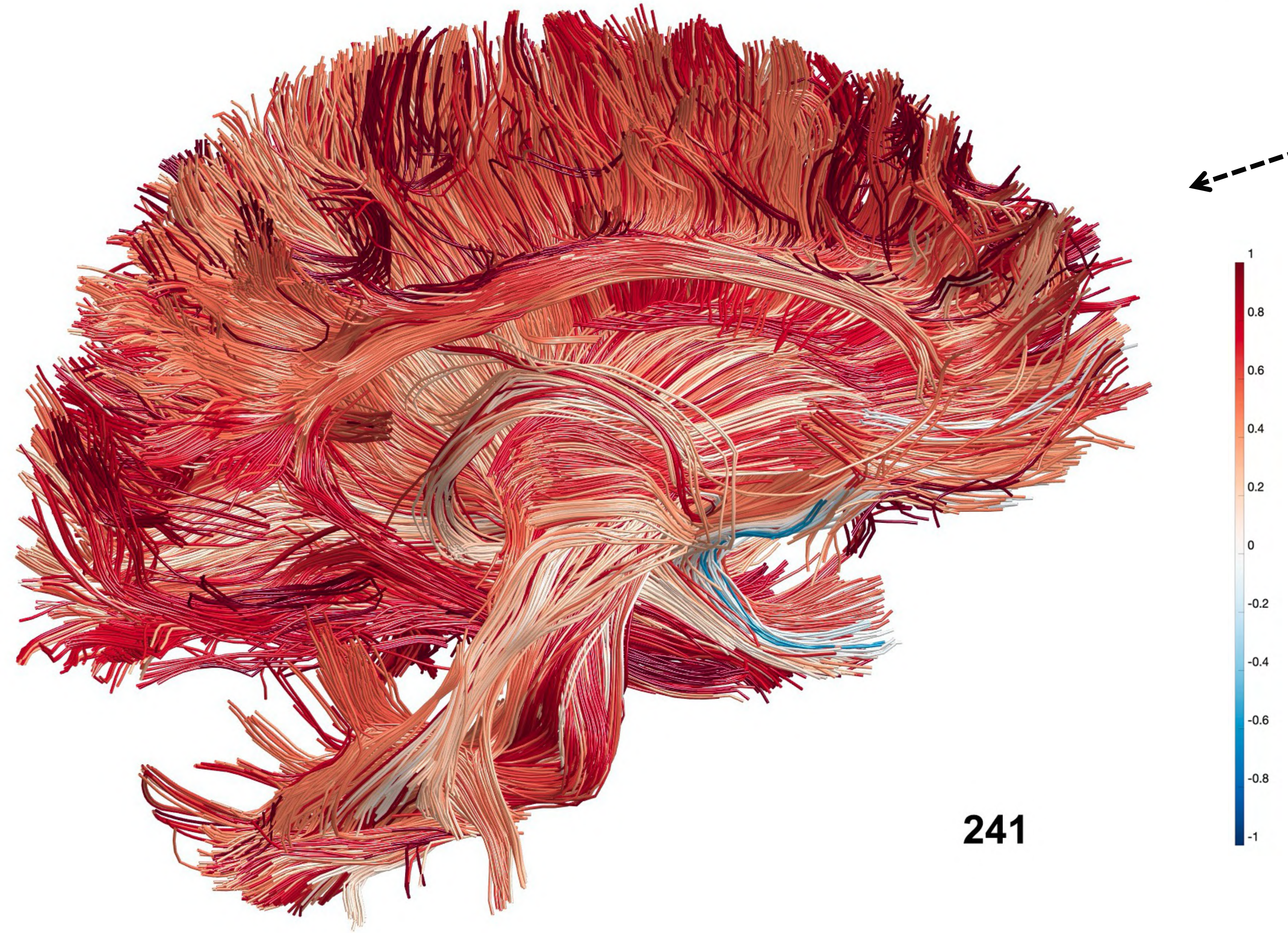
Dynamically changing correlation brain network at voxel level



Correlation network of 300000 time series

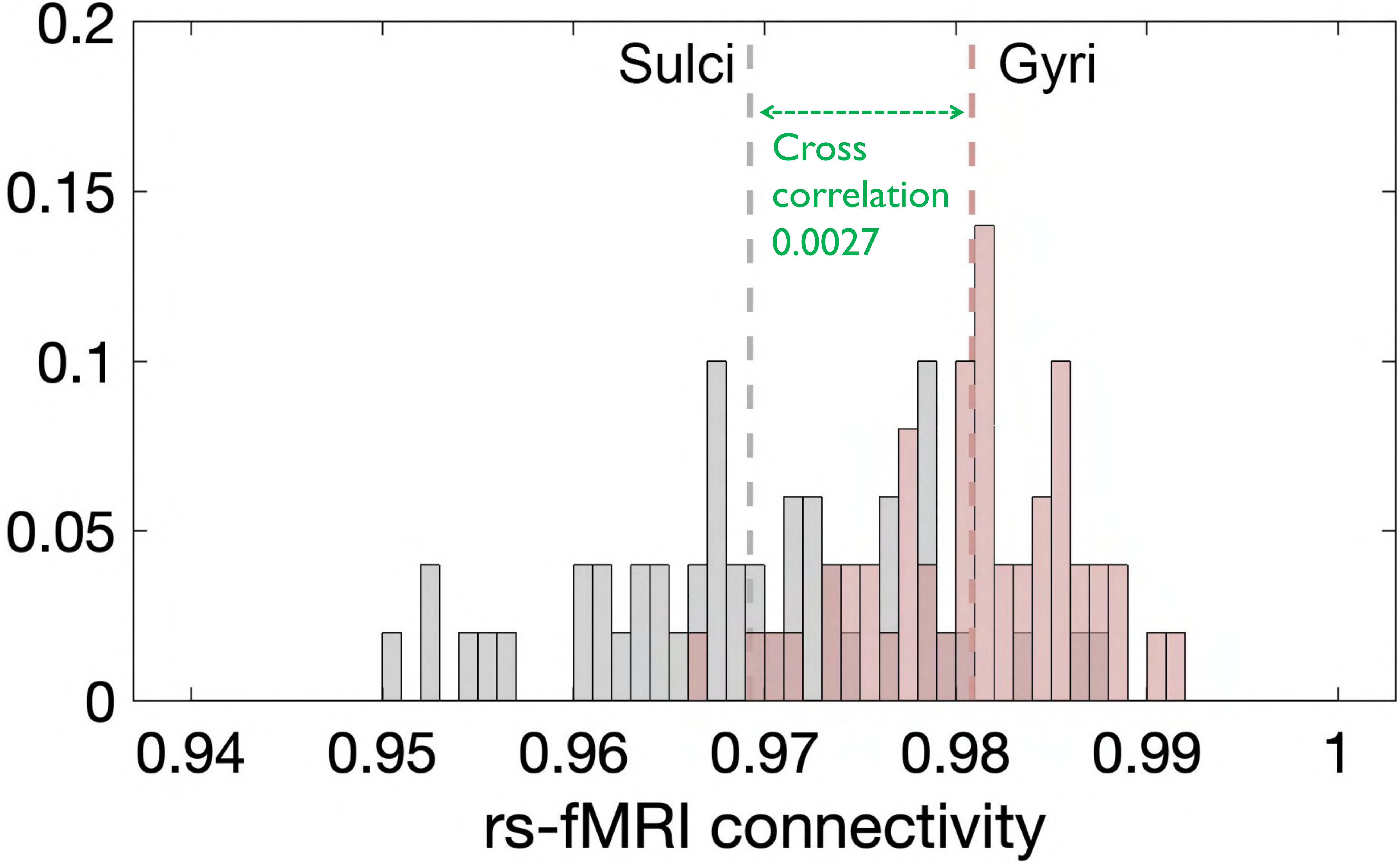
Dynamically changing omplete graph with about $300000^2/2$ cycles.





Dynamically changing functional connectivity superimposed on top of white matter fibers

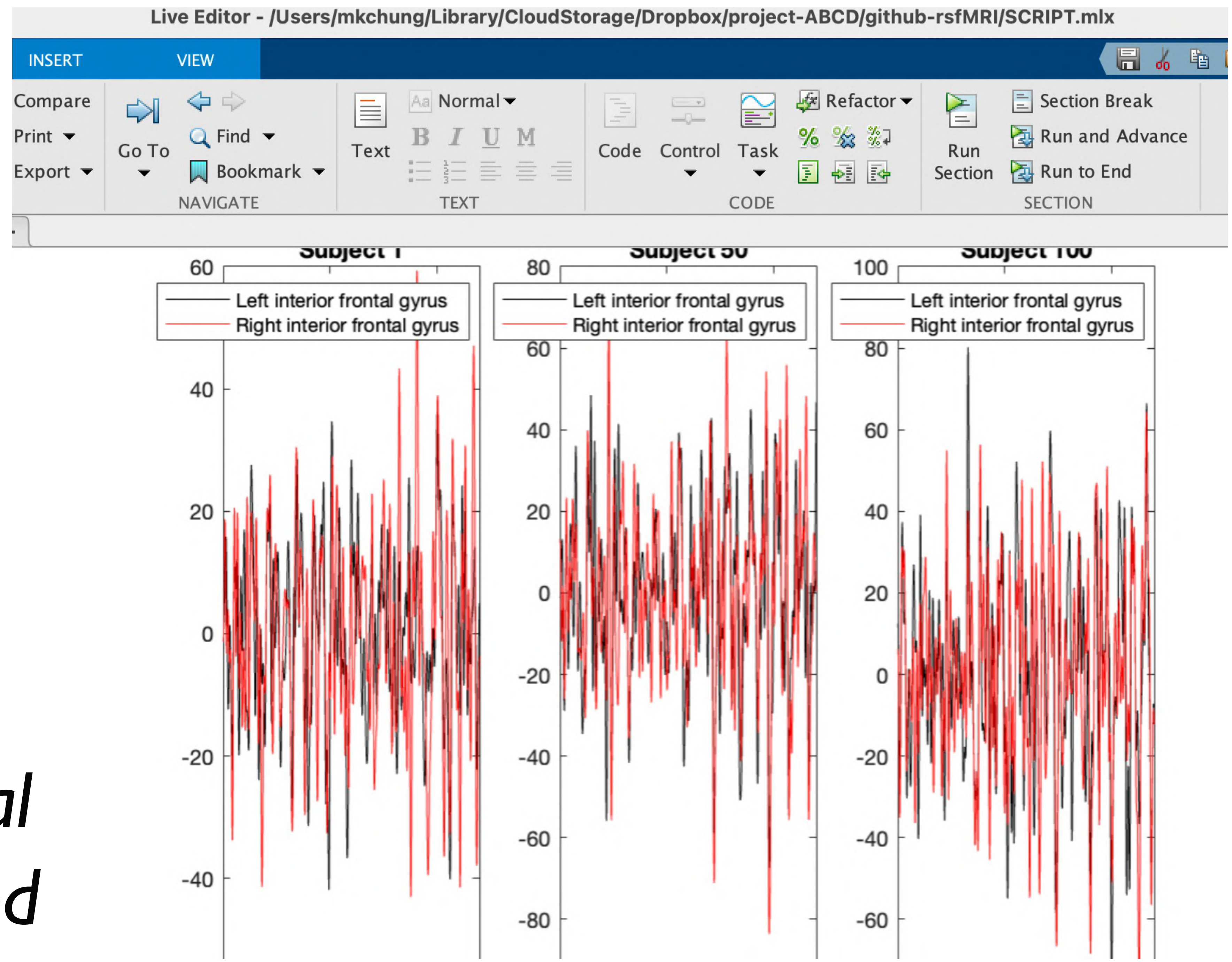
Differential functional connectivity across sulci and gyri



rs-fMRI time series data

https://github.com/laplcebeltrami/rsfMRI

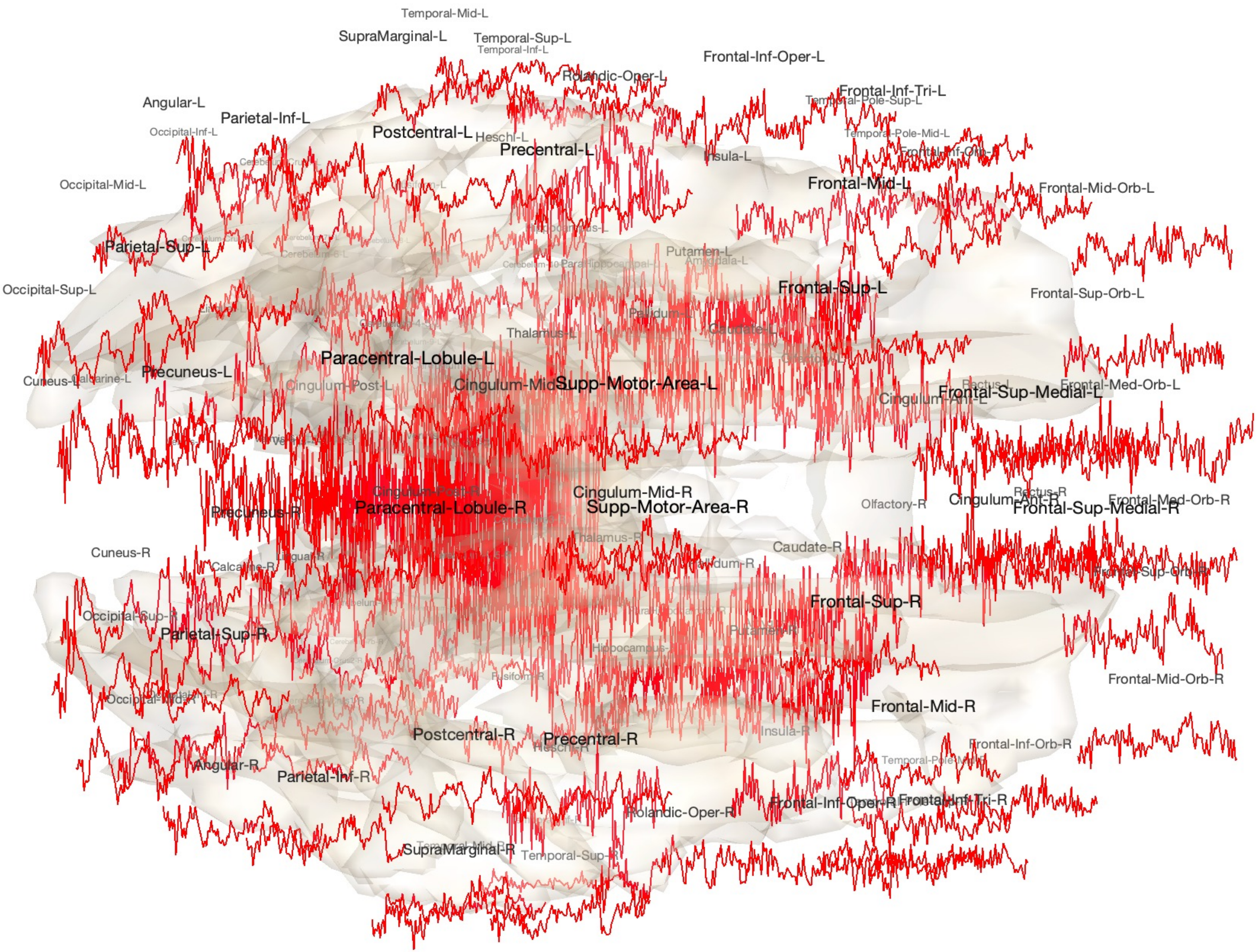
MATLAB



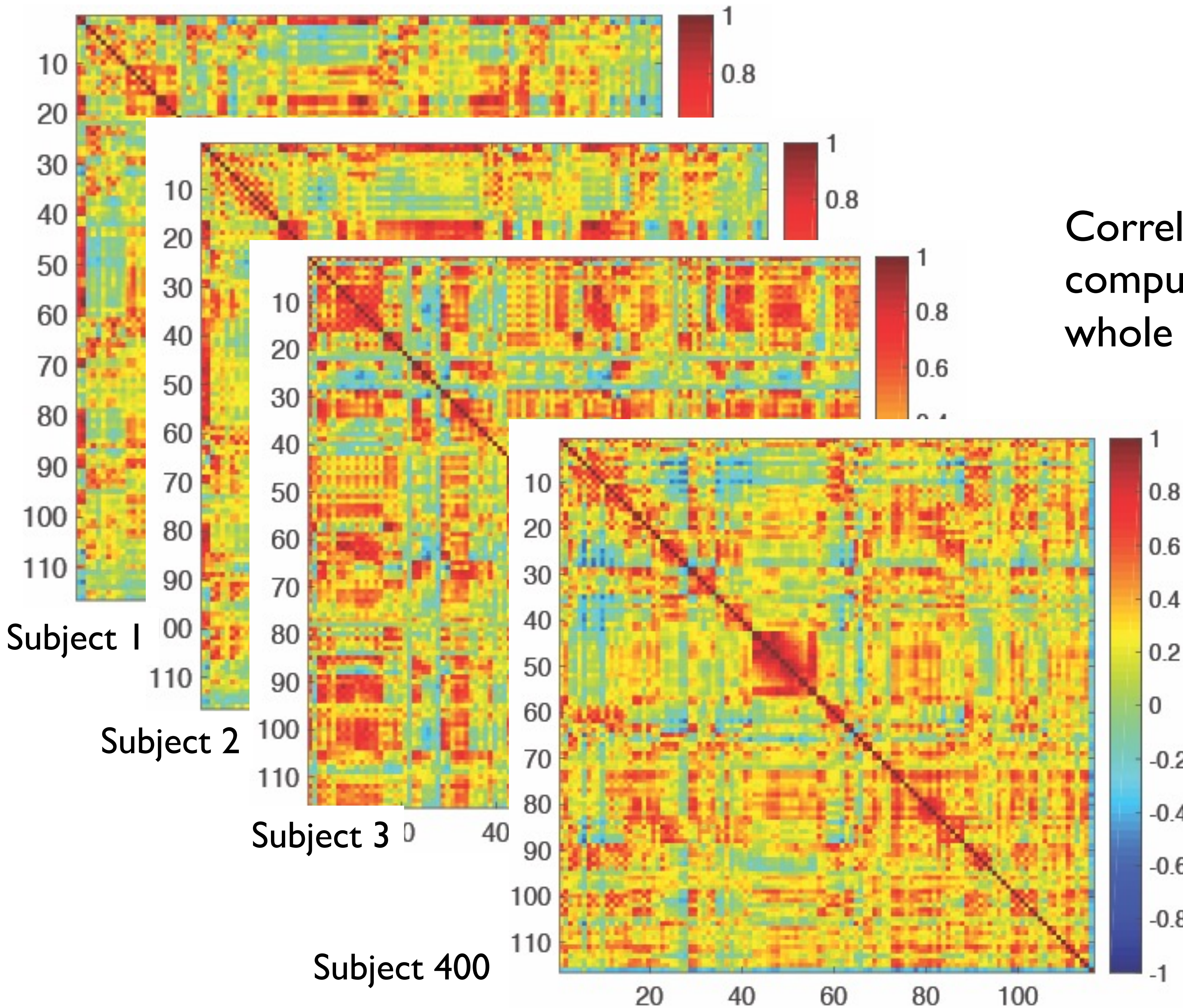
Important biological questions are added

[Huang et al. 2020 Neuroscience Methods 331:108480]()

Time series averaged into 116 brain regions

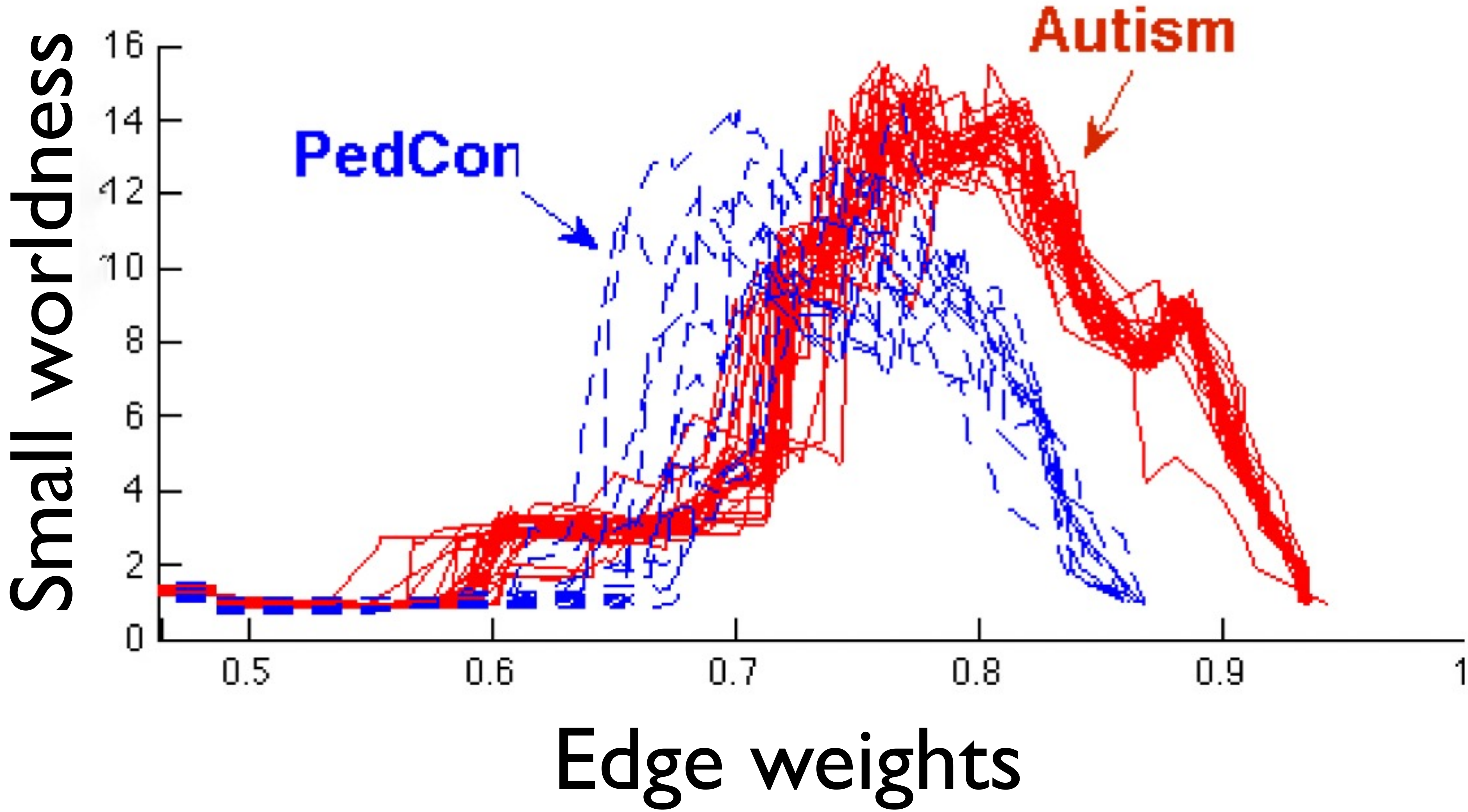


Subject level brain connectivity matrix



Correlation computed over whole time points

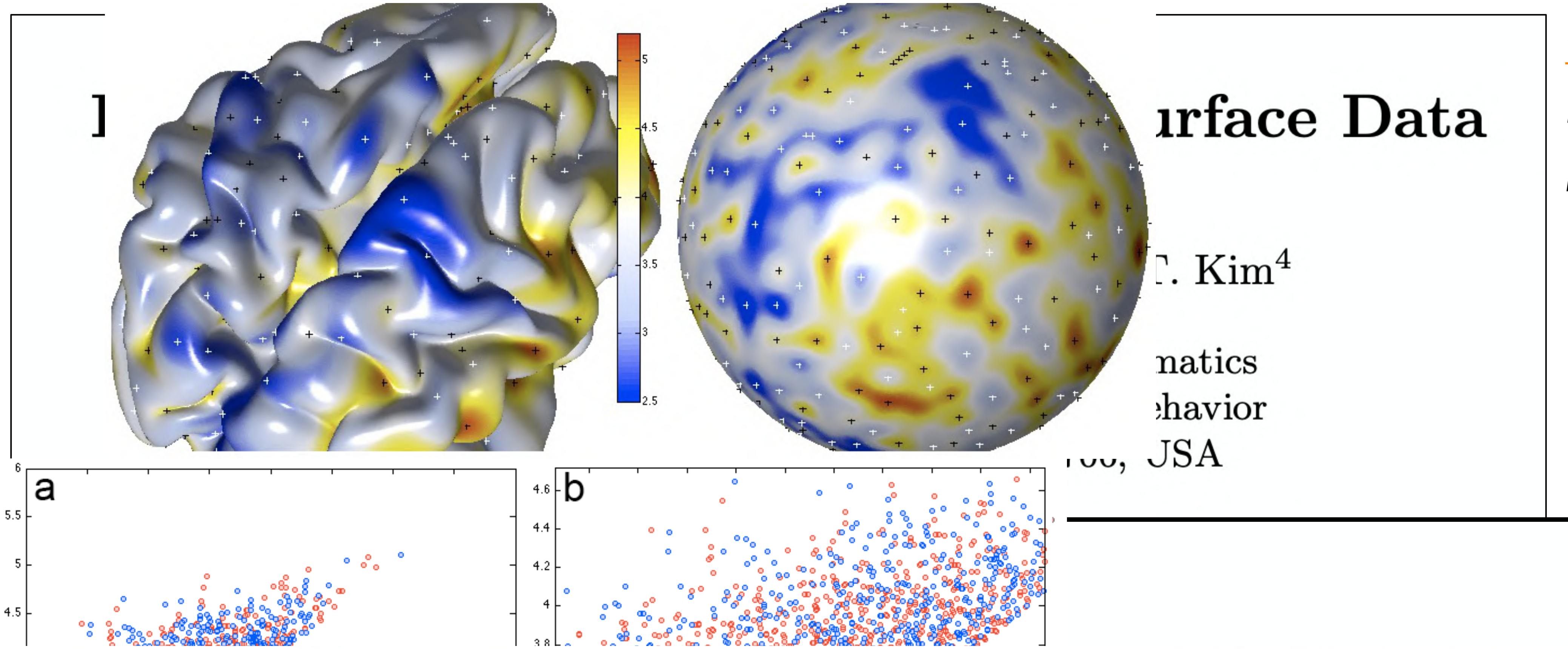
Why we need to avoid graph theory features?



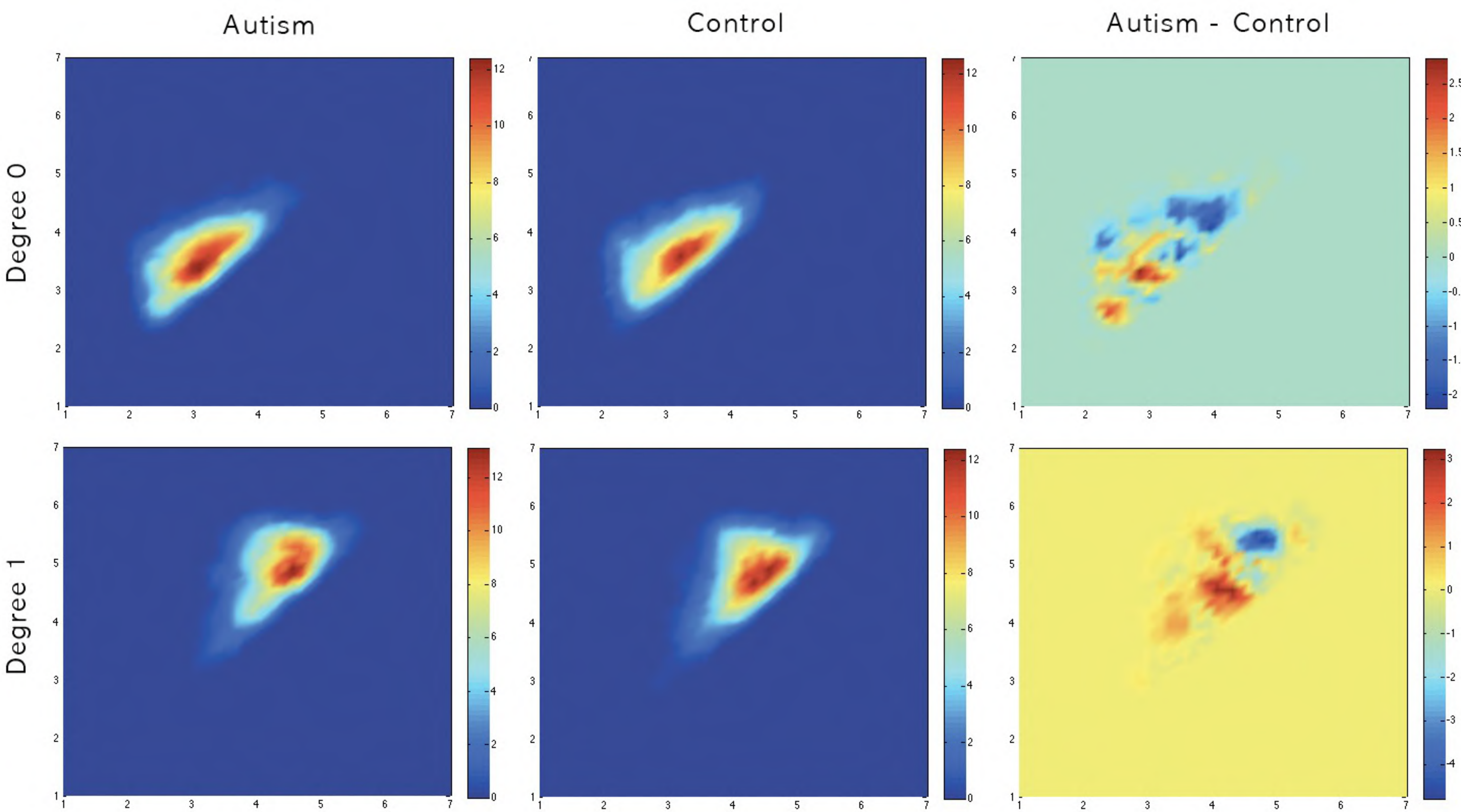
Lee et al. 2011 IEEE Transactions on Medical Imaging 30:1154-1165

Topological data analysis (TDA)

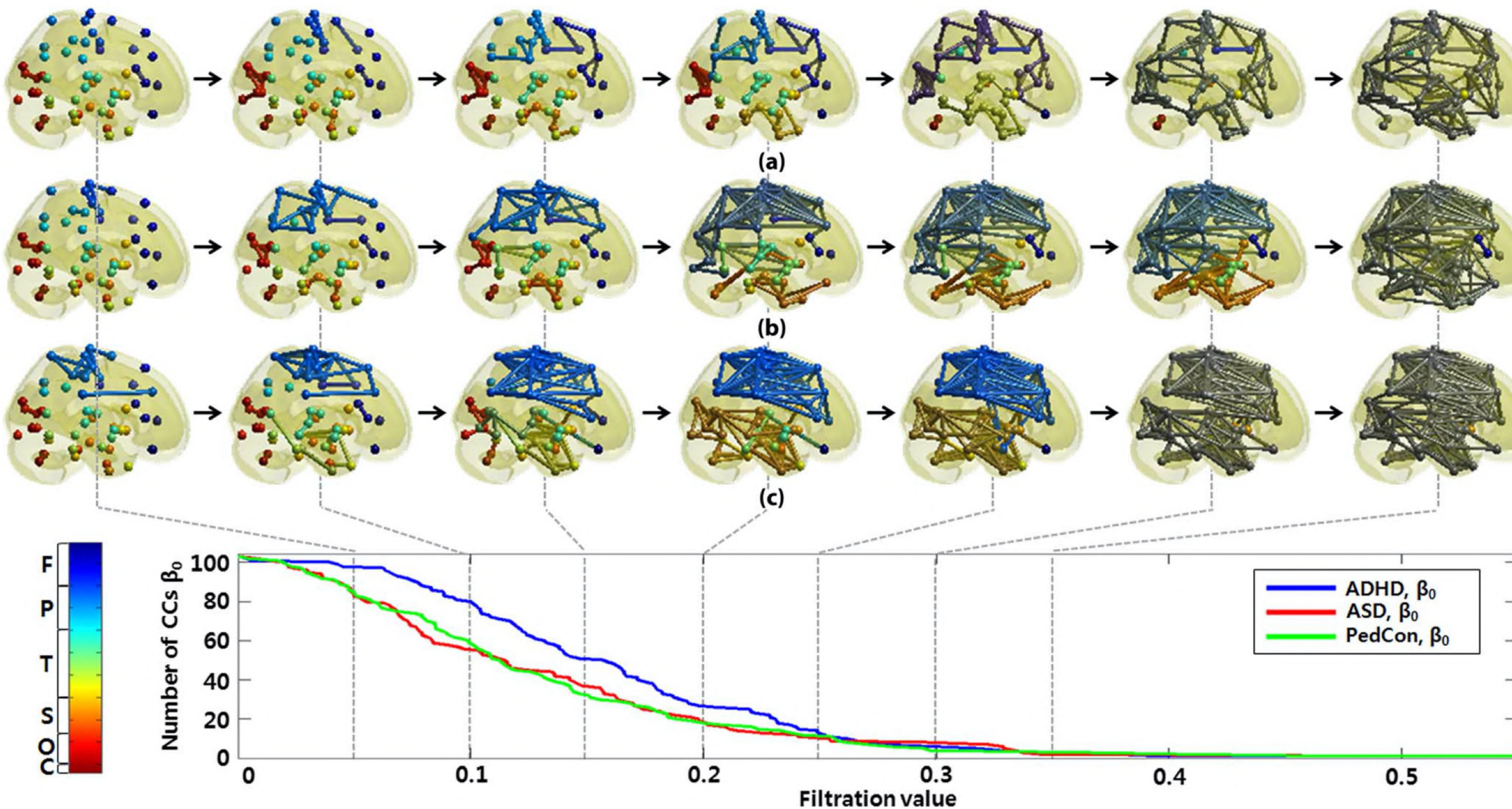
Completely data driven!
No explicit model!
No distributional assumption!



Chung et al., 2009
Information Processing in Medical Imaging (IPMI) 5636:386-397.



First persistent homology paper in brain imaging



First persistent homology paper in brain network analysis

Lee et al. (2011) ISBI

Lee et al. 2012 IEEE Transactions on Medical Imaging 31:2267-2277

Matlab toolbox PH-STAT

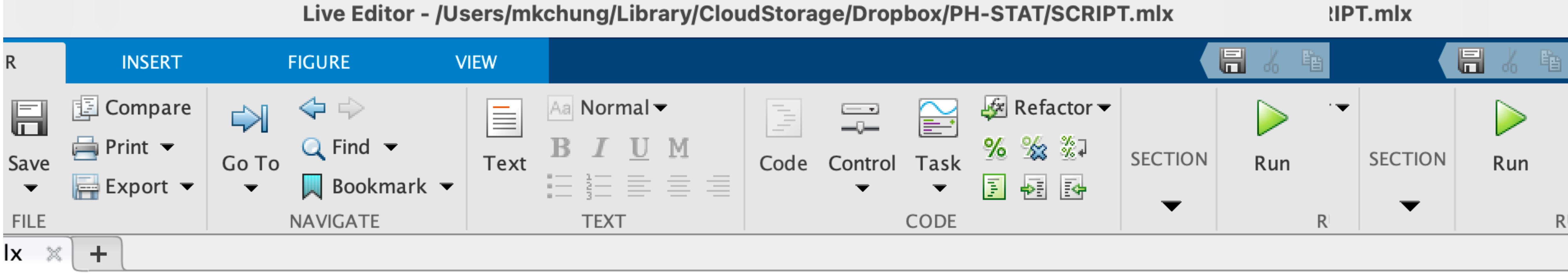
Statistical Inference on Persistent Homology

https://github.com/laplcebeltrami/PH-STAT

Manual:

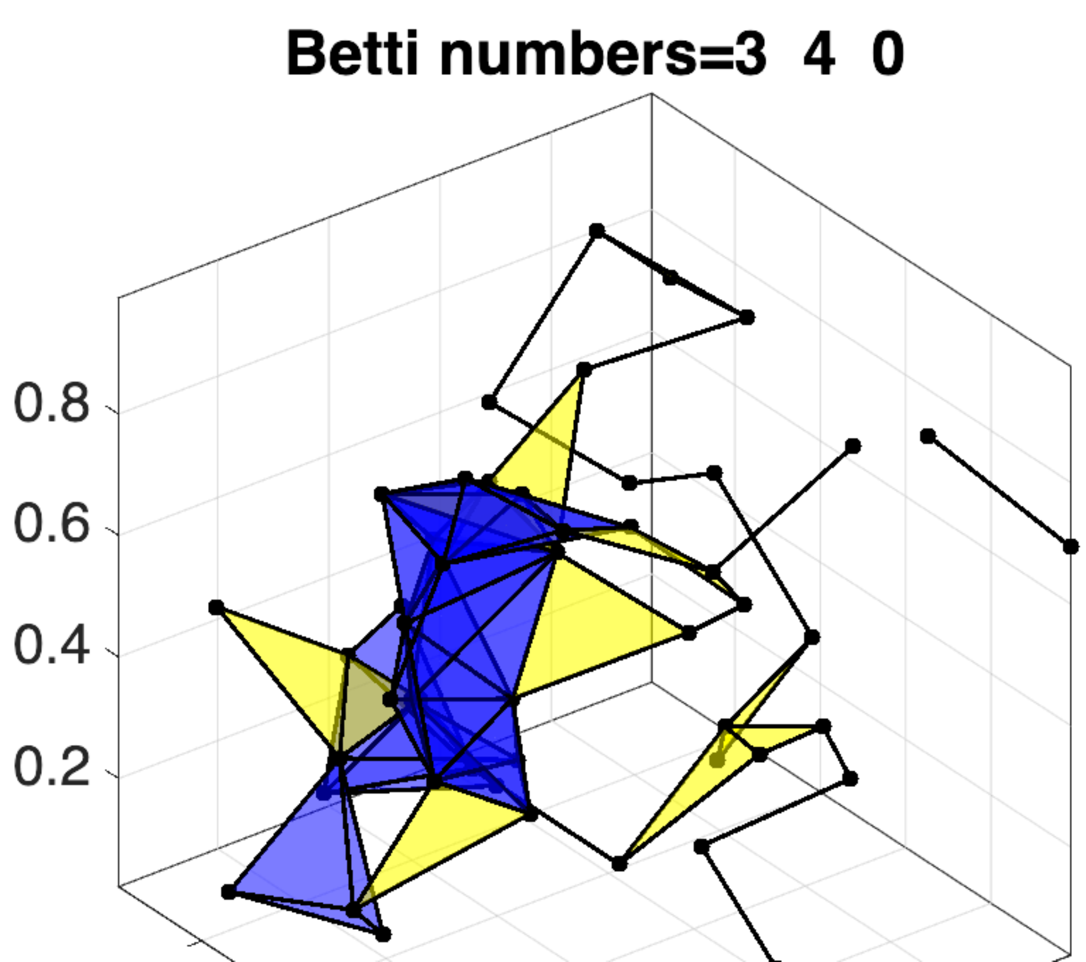
Chung 2023, PH-STAT arXIv:2304.05912

The self-contained package can do topological clustering and inference explained in this talk




```
%Display Ris complex
PH_rips_display(X,S);
%labels = cellstr(num2str((1:p)', '% d'));
%text(X(:,1)+0.01, X(:,2)+0.01, X(:,3)+0.01, labels, 'Color', 'r', 'FontSize',16)

% Boundary matrices
B = PH_boundary(S);
betti = PH_boundary_betti(B);
title(['Betti numbers=' num2str(betti)])
```



Will be built on top of 7000+ custom functions

Goal: scalable computation in laptop

Graph Filtrations

Weighted complete graph

$$\mathcal{X} = (V, w) \quad w = (w_{ij})$$

Node set

Edge weight

Binary graph

$$\mathcal{X}_\epsilon = (V, w_\epsilon)$$

$$w_{\epsilon,ij} = \begin{cases} 1 & \text{if } w_{ij} > \epsilon; \\ 0 & \text{otherwise.} \end{cases}$$

Graph filtration

$$\mathcal{X}_{\epsilon_0} \supset \mathcal{X}_{\epsilon_1} \supset \mathcal{X}_{\epsilon_2} \supset \cdots$$

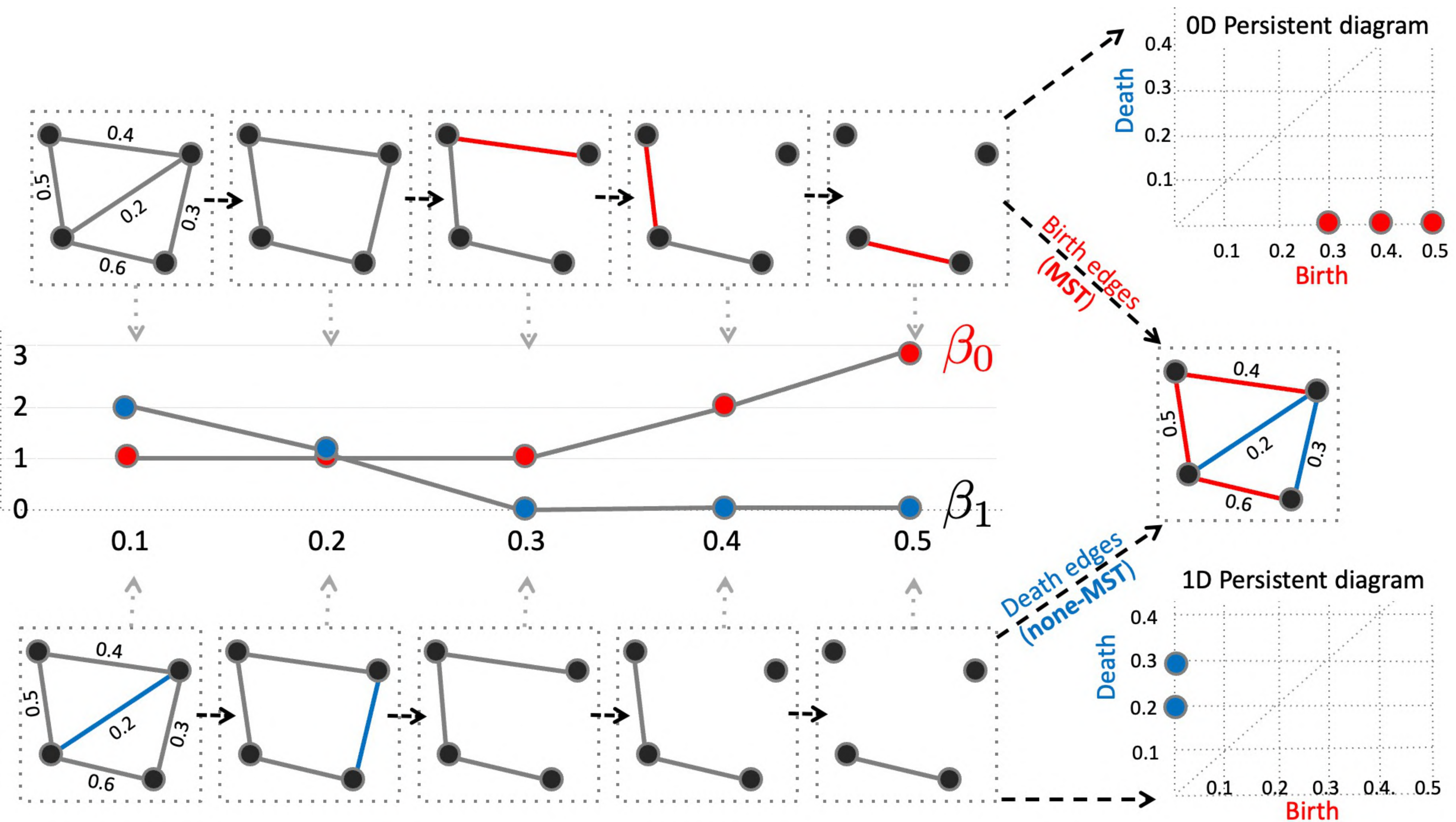
for increased edge weights

$$\epsilon_0 < \epsilon_1 < \epsilon_2 < \cdots$$

Lee et al. 2011 MICCAI 6892:302-309

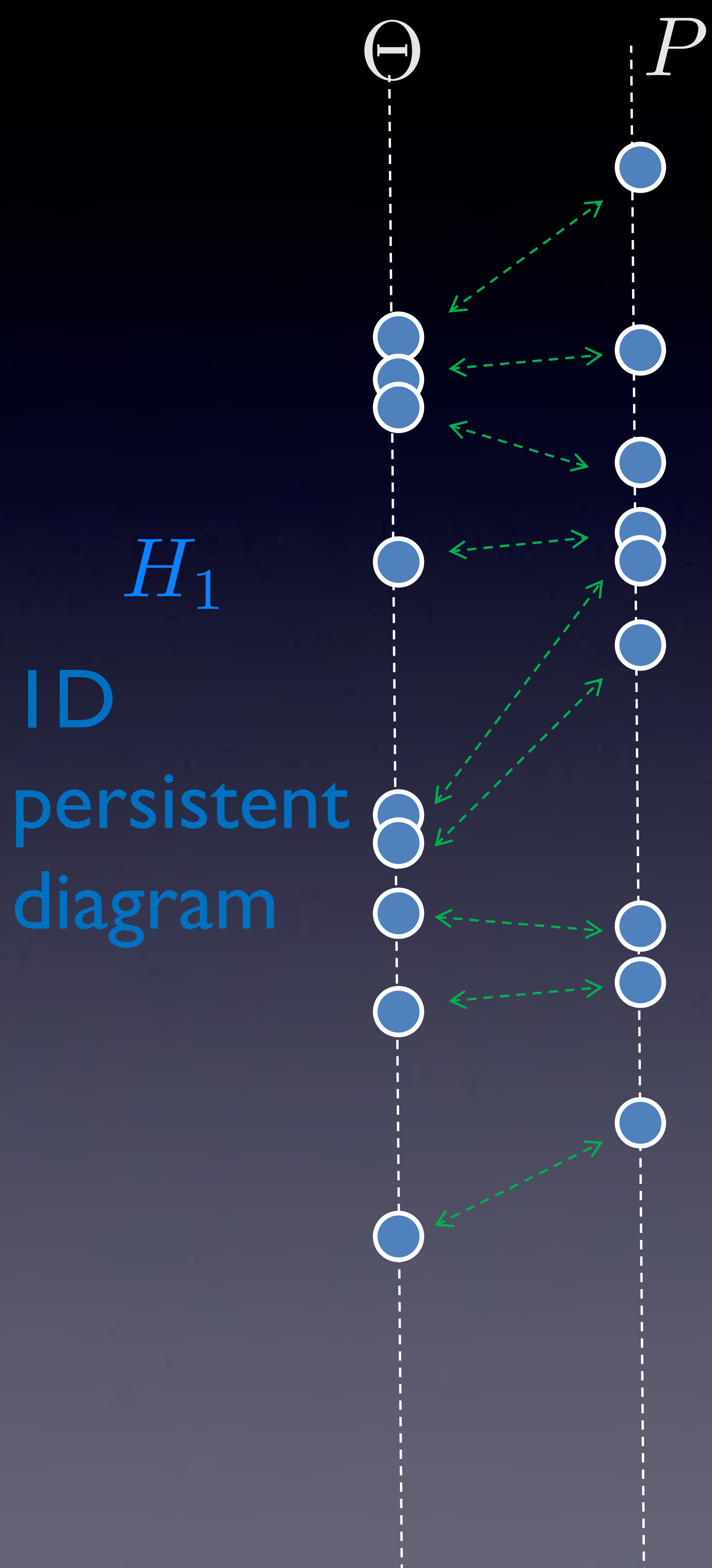
Chung et al. 2019 Network Neuroscience 3:674-694

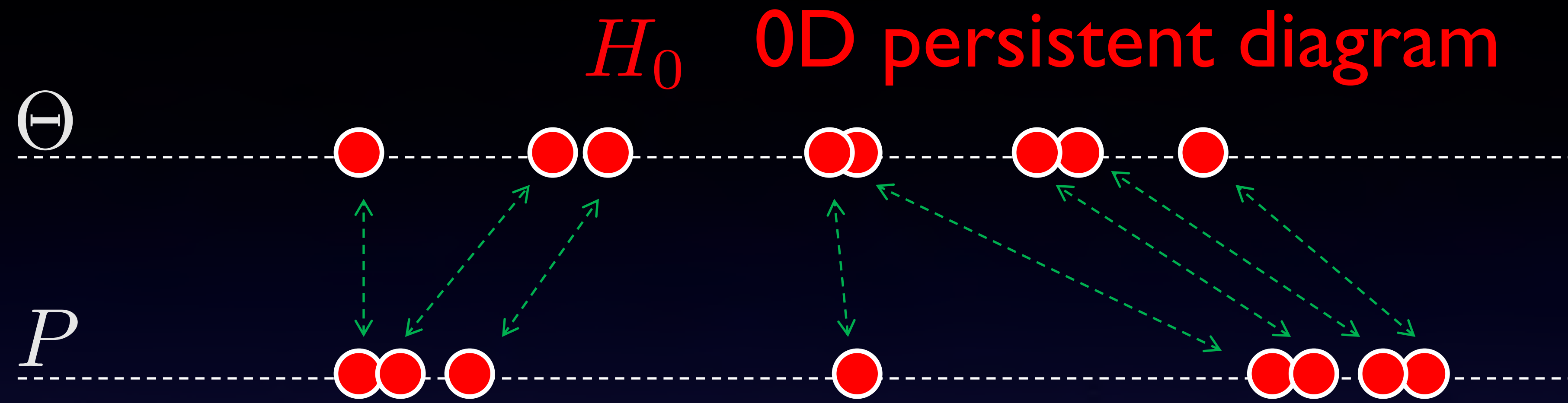
Theorem: Birth & death decomposition



Songdechakraiwut and Chung. 2023, Annals of Appled Statistics 17:403-433

Wasserstein distance for graph filtrations





1D assignment problem has known analytic solution

$\mathcal{O}(q \log q)$

Theorem: Wasserstein distance on graph filtrations

$$\mathcal{L}_{0D}(\Theta, P) = \min_{\tau} \sum_{b \in E_0} \left[b - \tau(b)\right]^2$$

Birth set

$$= \sum_{b \in E_0} \left[b - \tau_0^*(b)\right]^2$$

τ_0^* *:The i-th smallest birth value to the i-th smallest birth value*

$$\mathcal{L}_{1D}(\Theta, P) = \min_{\tau} \sum_{d \in E_1} \left[d - \tau(d)\right]^2$$

Death set

$$= \sum_{d \in E_1} \left[d - \tau_1^*(d)\right]^2$$

τ_1^* *:The i-th smallest death value to the i-th smallest death value*

Songdechakraiwut and Chung. 2023, Annals of Appled Statistics

Wasserstein distance between networks

$$C_1 \cup C_2 = \{\mathcal{X}_1, \cdots, \mathcal{X}_n\}, \quad C_1 \cap C_2 = \emptyset$$

Between-group distance

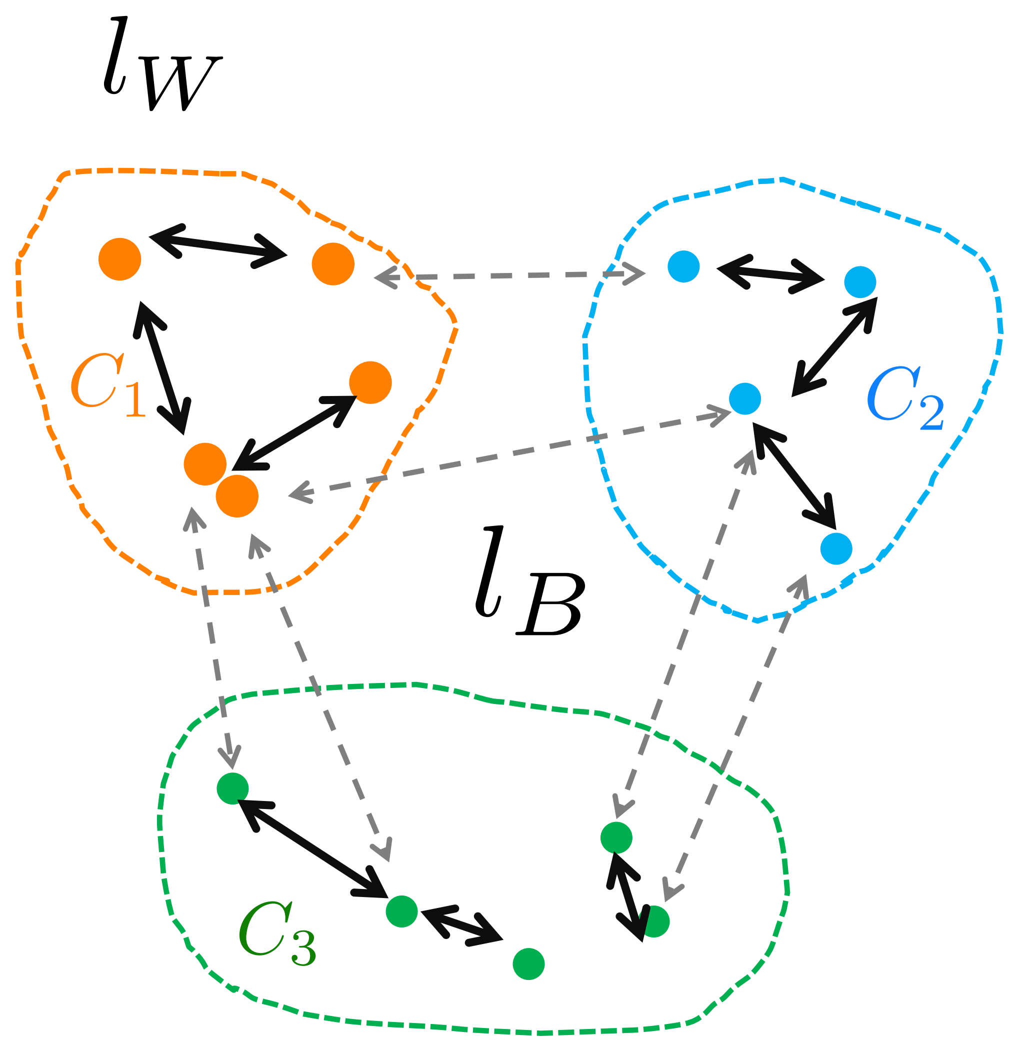
$$l_B \propto \sum_{i \in C_1, j \in C_2} \mathcal{L}(\mathcal{X}_i, \mathcal{X}_j)$$

0D and 1D combined distances

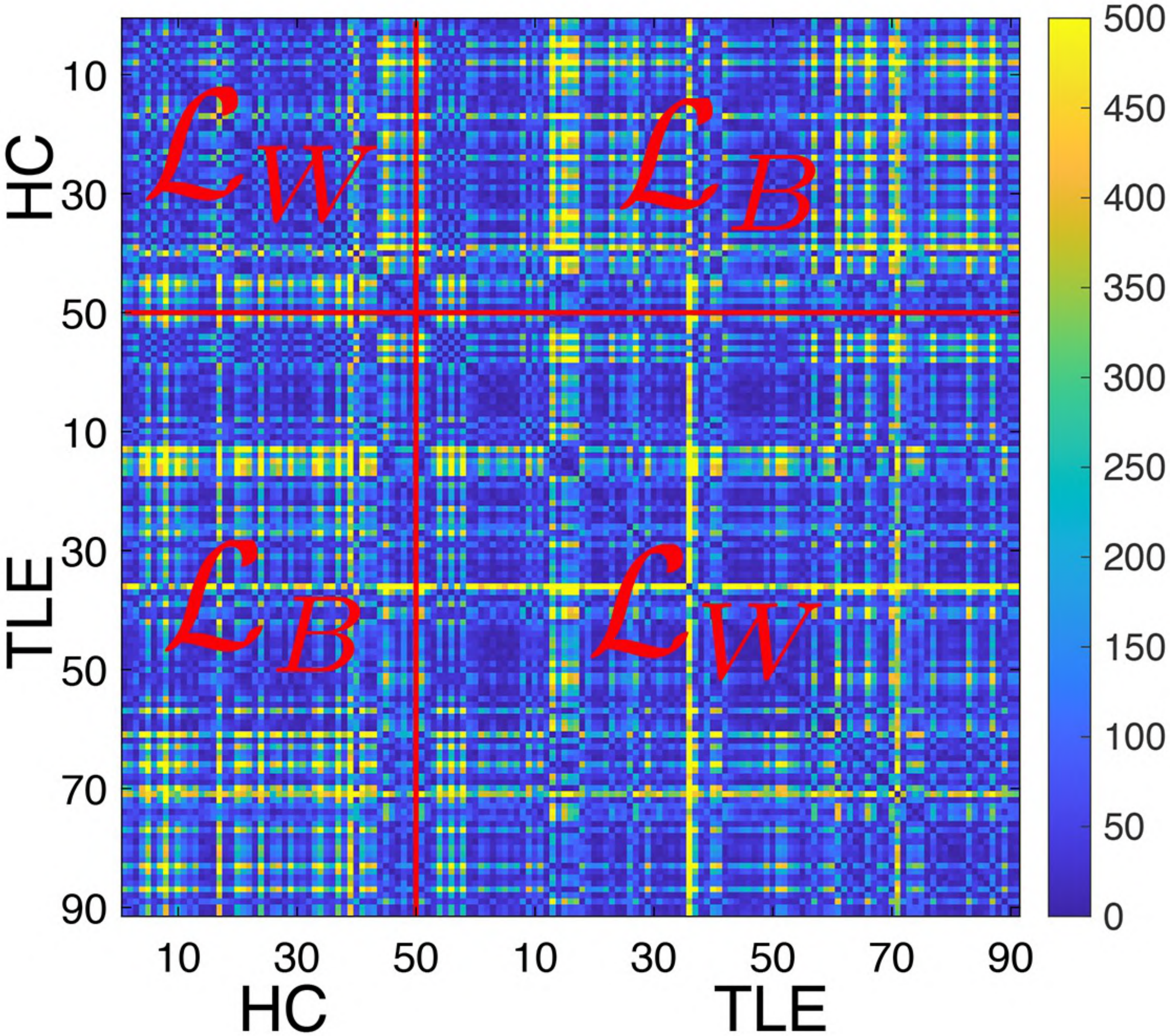
Within-group distance

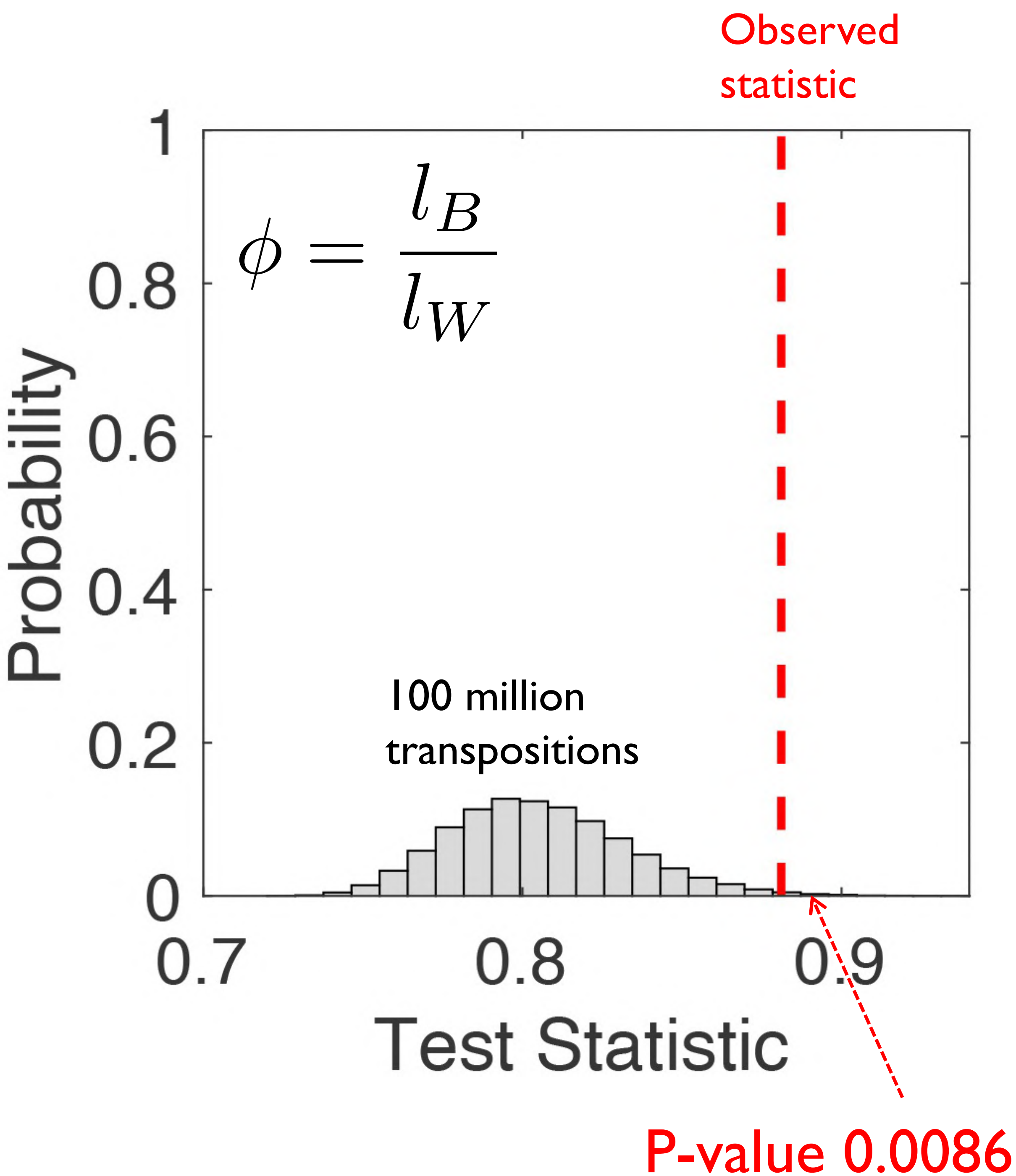
$$l_W \propto \sum_k \sum_{i,j \in C_k} \mathcal{L}(\mathcal{X}_i, \mathcal{X}_j)$$

$$l_B + l_W = \sum_{i,j} \mathcal{L}(\mathcal{X}_i, \mathcal{X}_j)$$



Topological inference on the ratio statistic





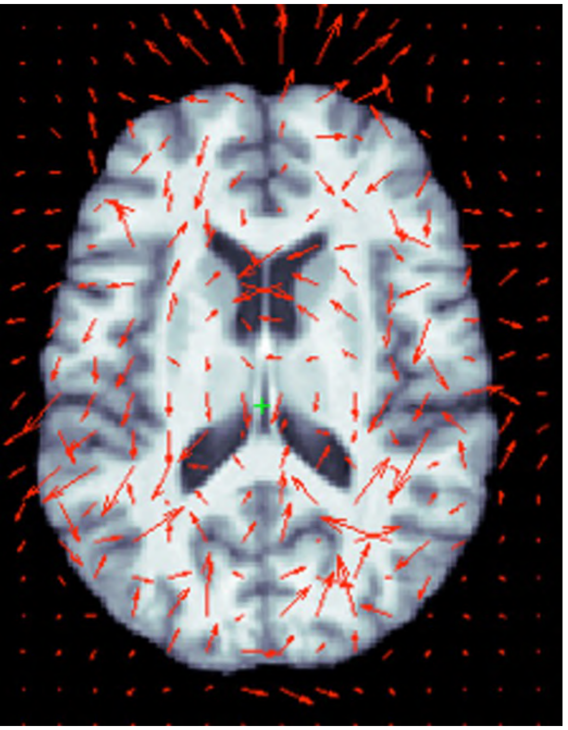
$$l_W \rightarrow l_W + \Delta(tranposition)$$

$$l_B \rightarrow l_B + \Delta(tranposition)$$

Songdechakraiwut and Chung 2023
Annals of Appled Statistics

Structural covariance network data

https://github.com/laplcebeltrami/maltreated



Jacobian matrix

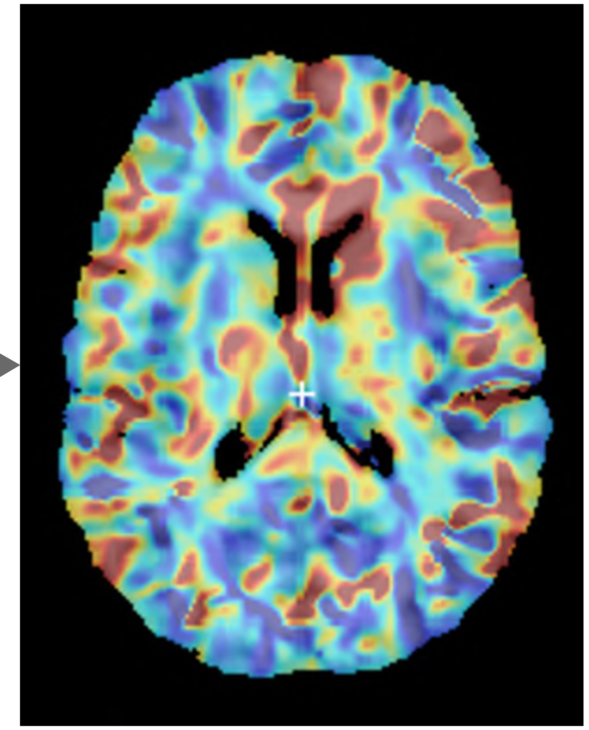
$$J = \frac{\partial d}{\partial x}$$

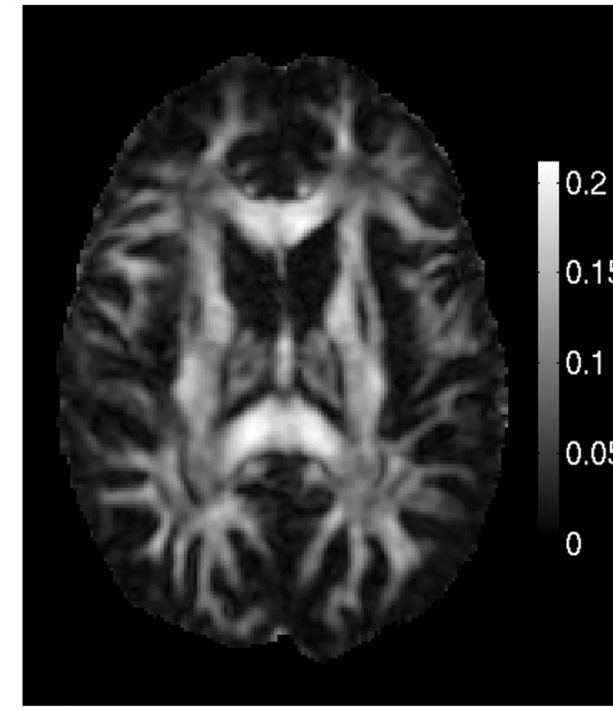
Riemannian metric tensor

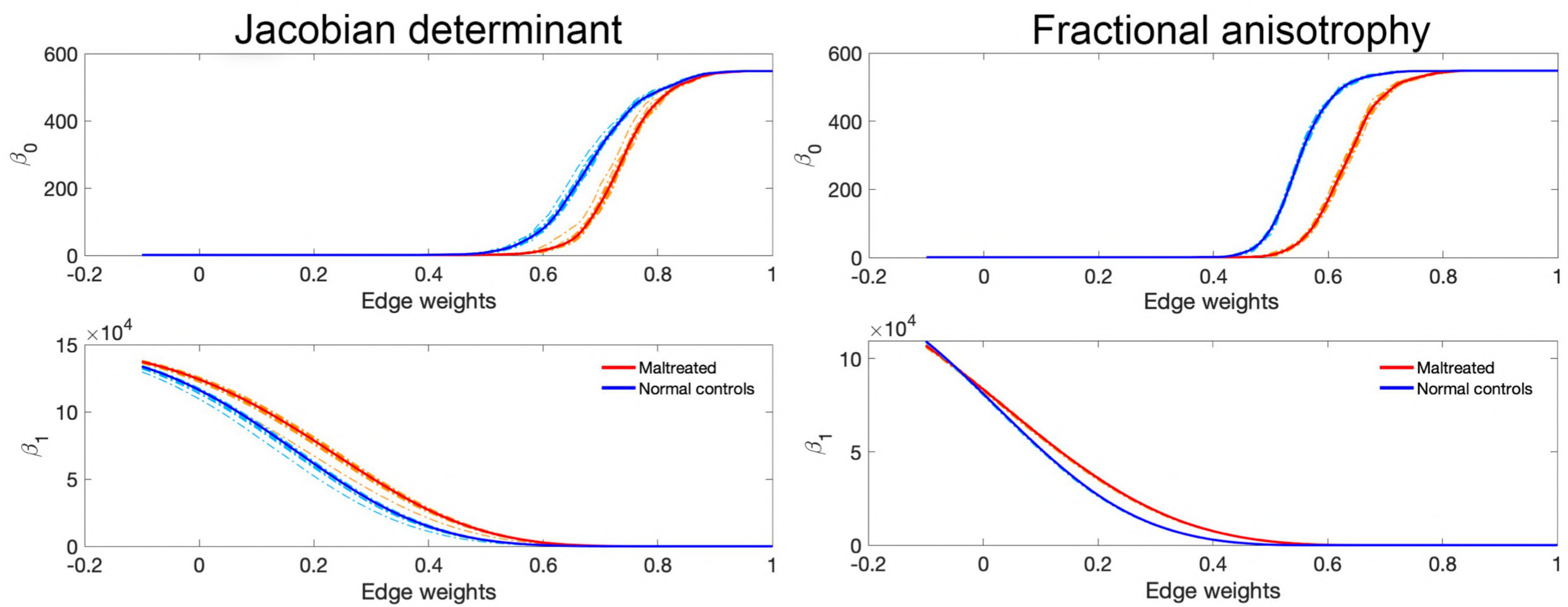
$$g = J^\top J$$

Volume element

$$\sqrt{\det g}$$







Chung et al. 2024 Network Neuroscience arXiv:2304.05908

Topological clustering

Minimize the within cluster distance

$$l_W \propto \sum_k \sum_{i,j \in C_k} \mathcal{L}(\mathcal{X}_i, \mathcal{X}_j)$$

Theorem: Topological clustering converges locally.

Algebraic proof:
Chung et al. 2023 NeuroImage

Geometric proof:
Chung et al. 2024 PLOS Computational Biology

The Wasserstein distance is equivalent to the Euclidean distance in the convex set $\mathcal{T}_0 \otimes \mathcal{T}_1$

Mathematical equivalence of topological clustering and topological inference

There exists a monotonically decreasing function f satisfying

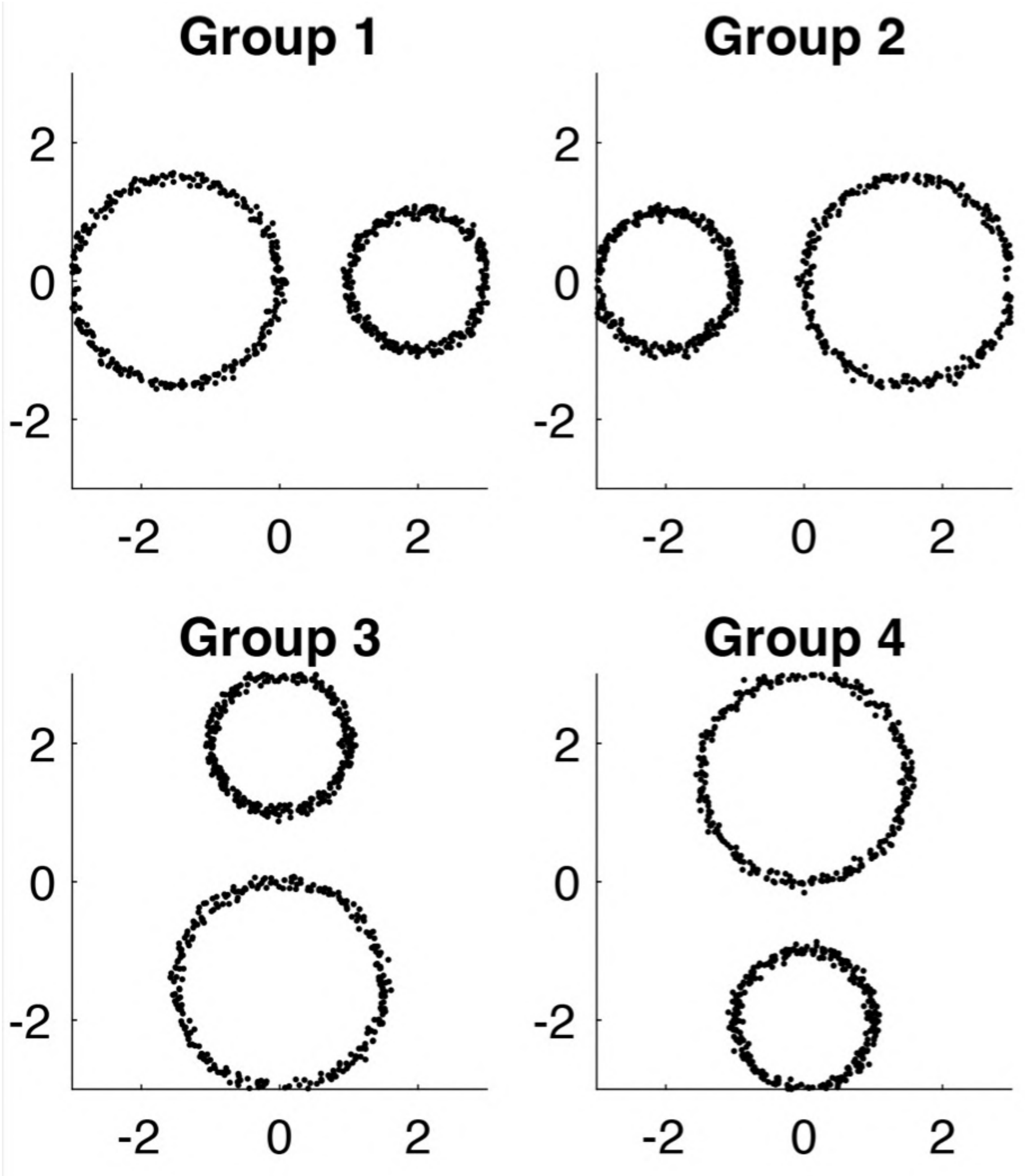
p-value = f (clustering accuracy)

Inference

Clustering

Proof in Chung et al. 2023 NeuroImage

Geometric methods fail topological clustering task



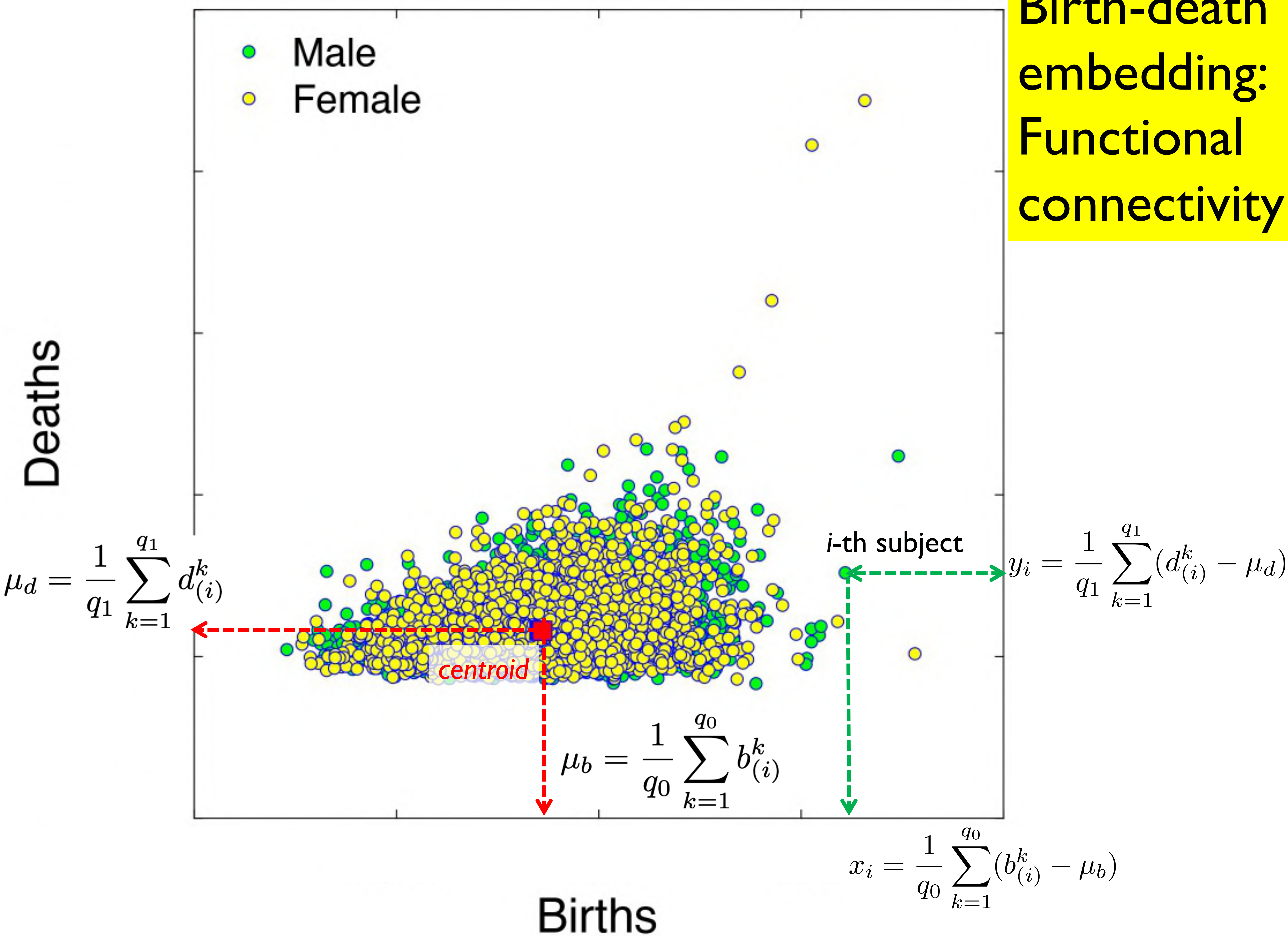
All false positives

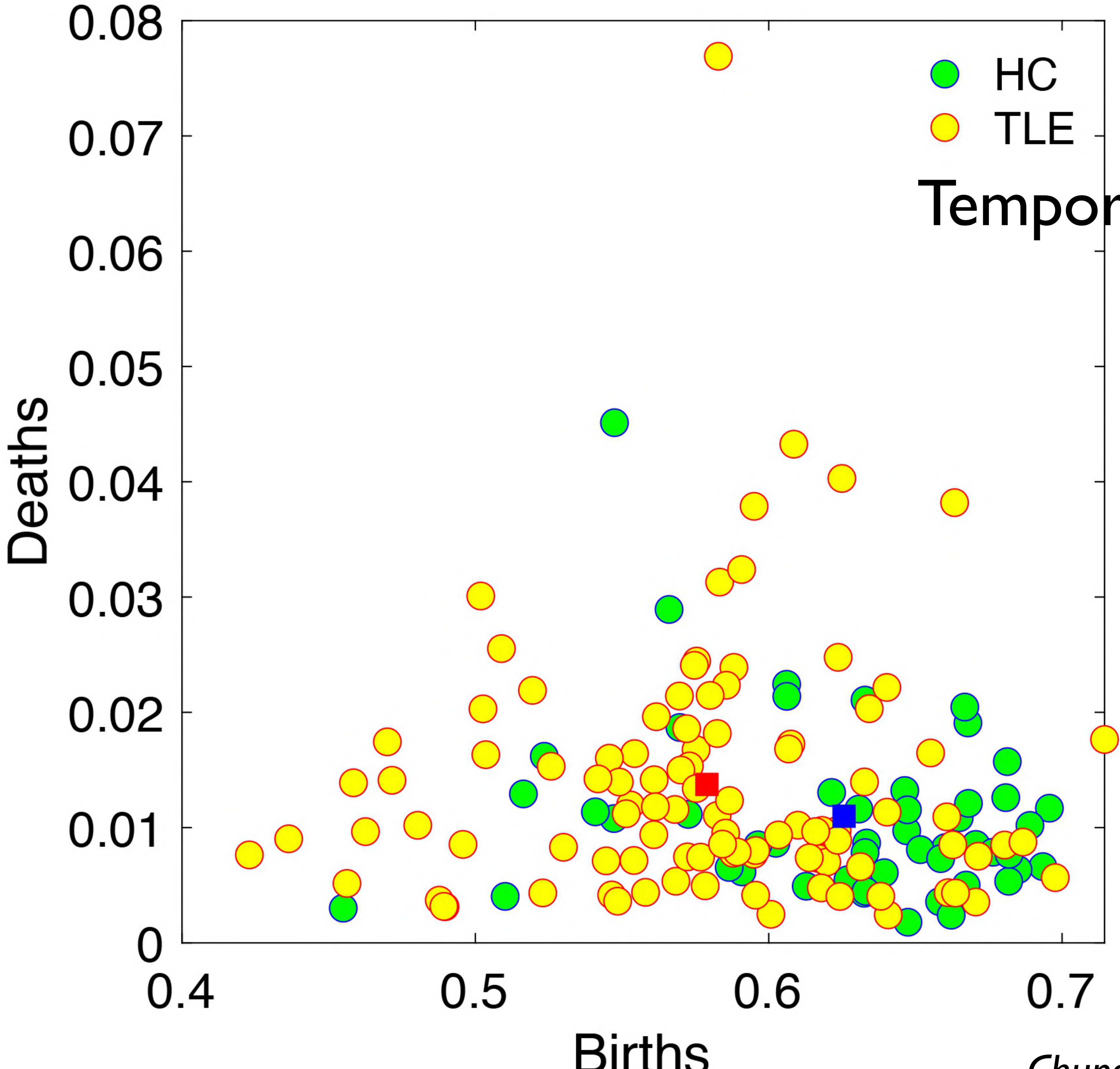
K-means clustering
0.98 +- 0.01

Hierarchical clustering
1.00 +- 0.00

Topological clustering
0.63 +- 0.04

Birth-death embedding: Functional connectivity

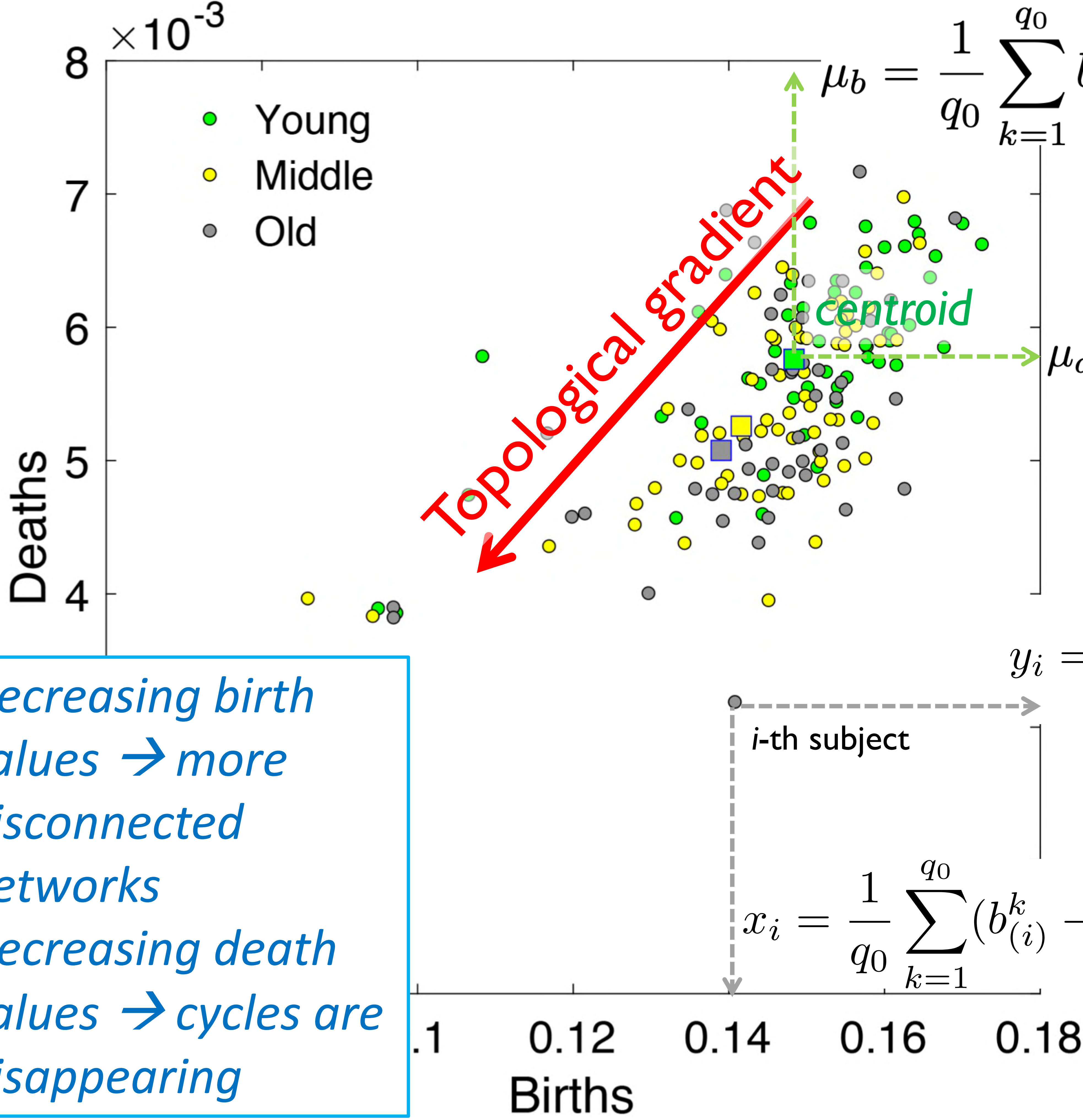




Temporal lobe epilepsy

Birth-death embedding: Functional connectivity

Chung et al. 2023 NeuroImage



$$\mu_b = \frac{1}{q_0}\sum_{k=1}^{q_0} b_{(i)}^k$$

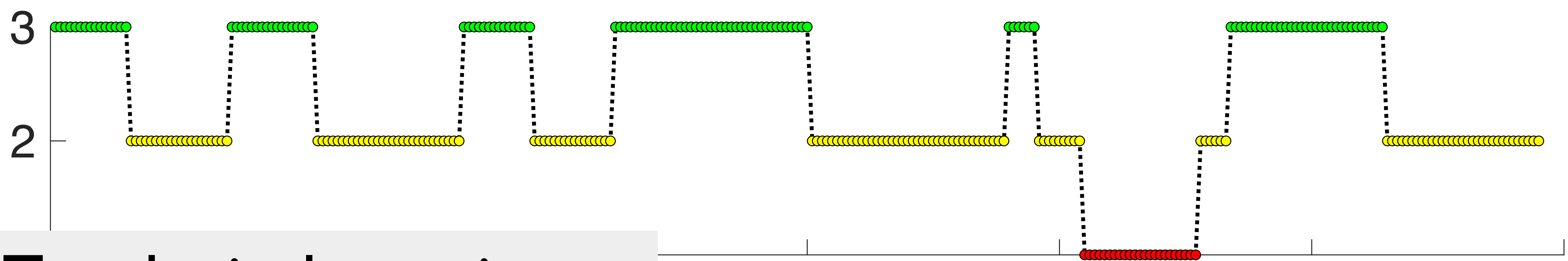
$$\mu_d = \frac{1}{q_1}\sum_{k=1}^{q_1} d_{(i)}^k$$

$$y_i = \frac{1}{q_1}\sum_{k=1}^{q_1}(d_{(i)}^k - \mu_d)$$

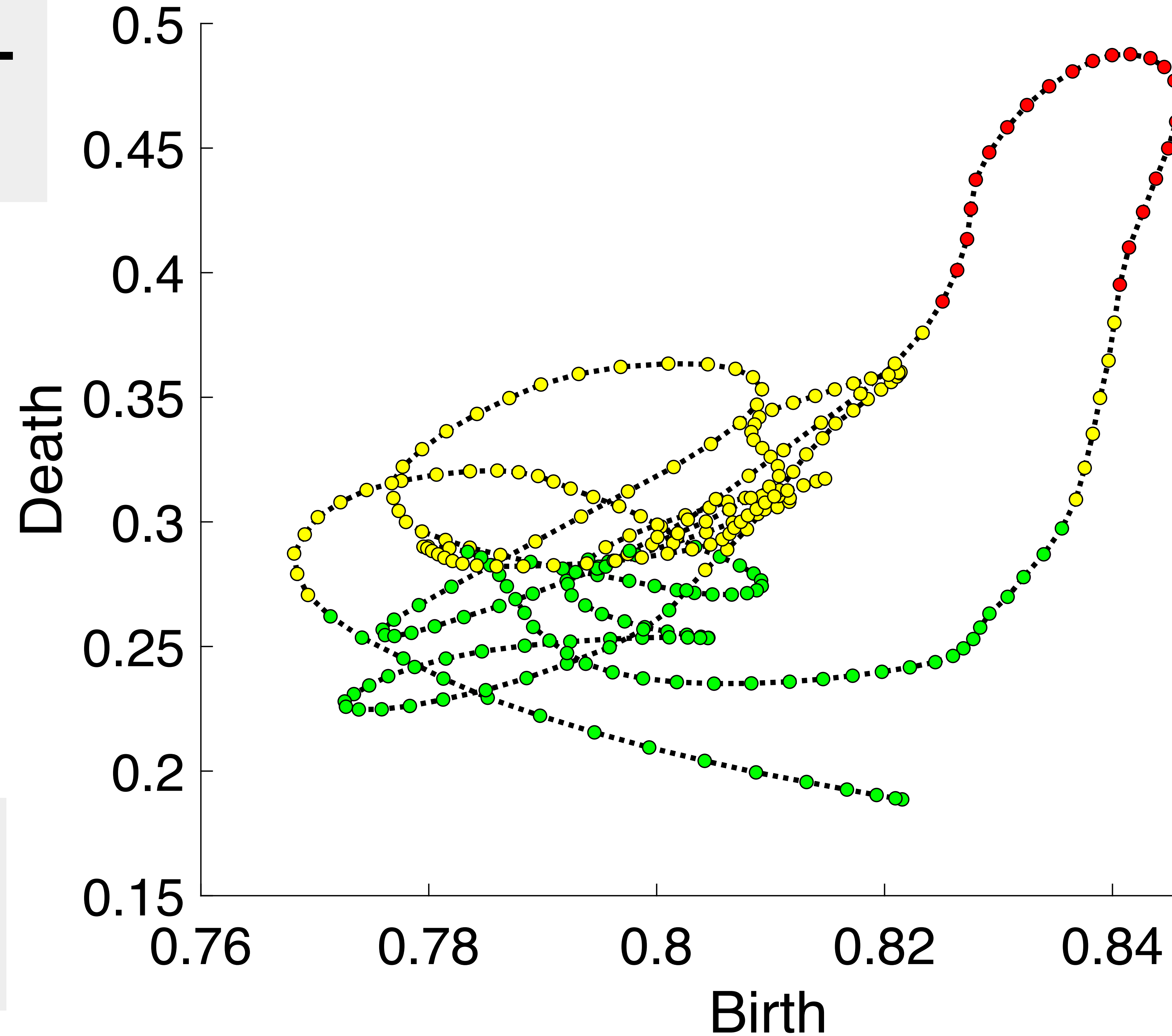
$$x_i = \frac{1}{q_0}\sum_{k=1}^{q_0}(b_{(i)}^k - \mu_b)$$

Decreasing birth values → more disconnected networks
Decreasing death values → cycles are disappearing

Structural connectivity

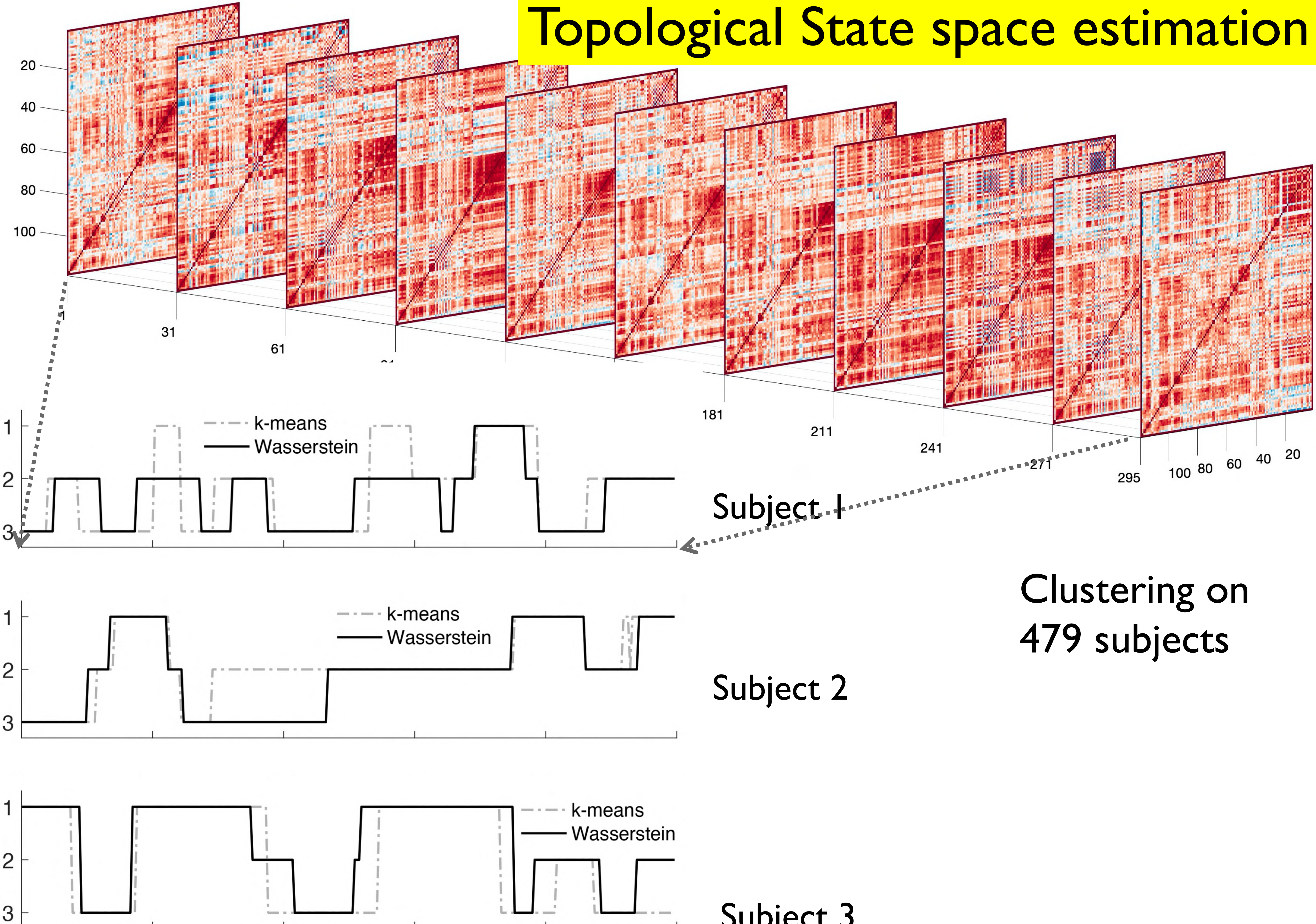


Topological transient connectivity patterns - state space estimation



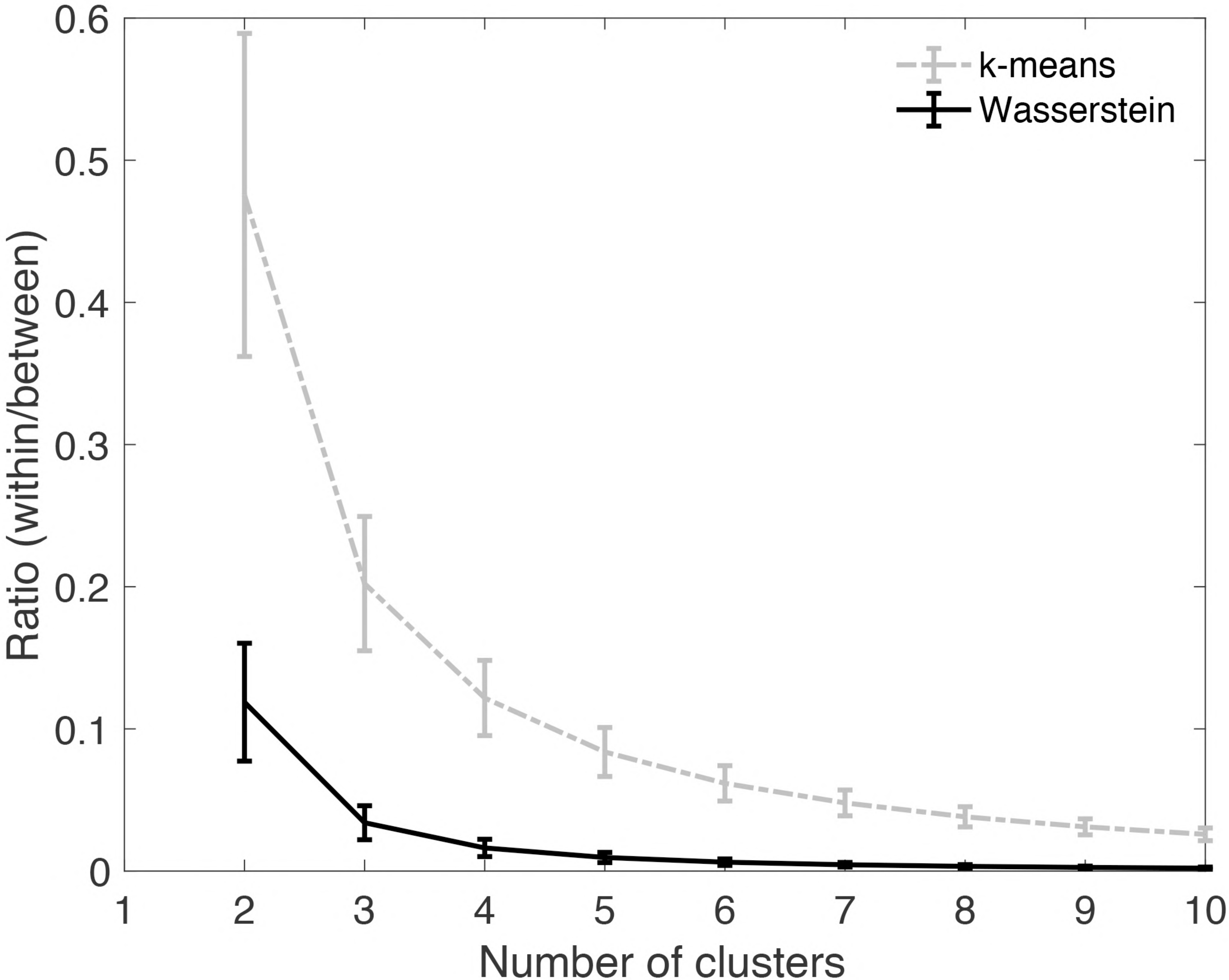
Birth-death embedding

Topological State space estimation

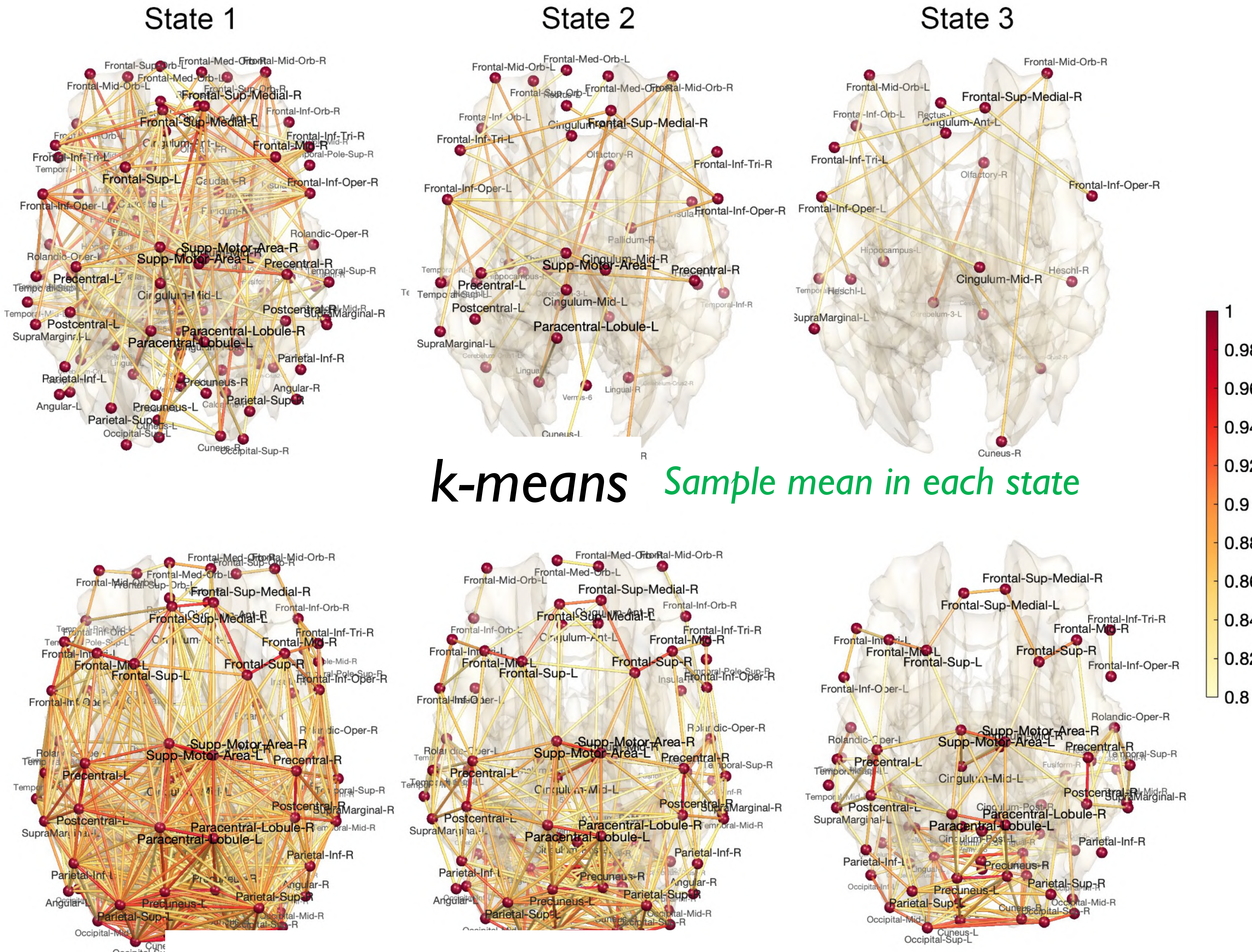


Clustering on 479 subjects

$$\frac{1}{\phi} = \frac{l_W}{l_B}$$



The within cluster variance *6 times* smaller



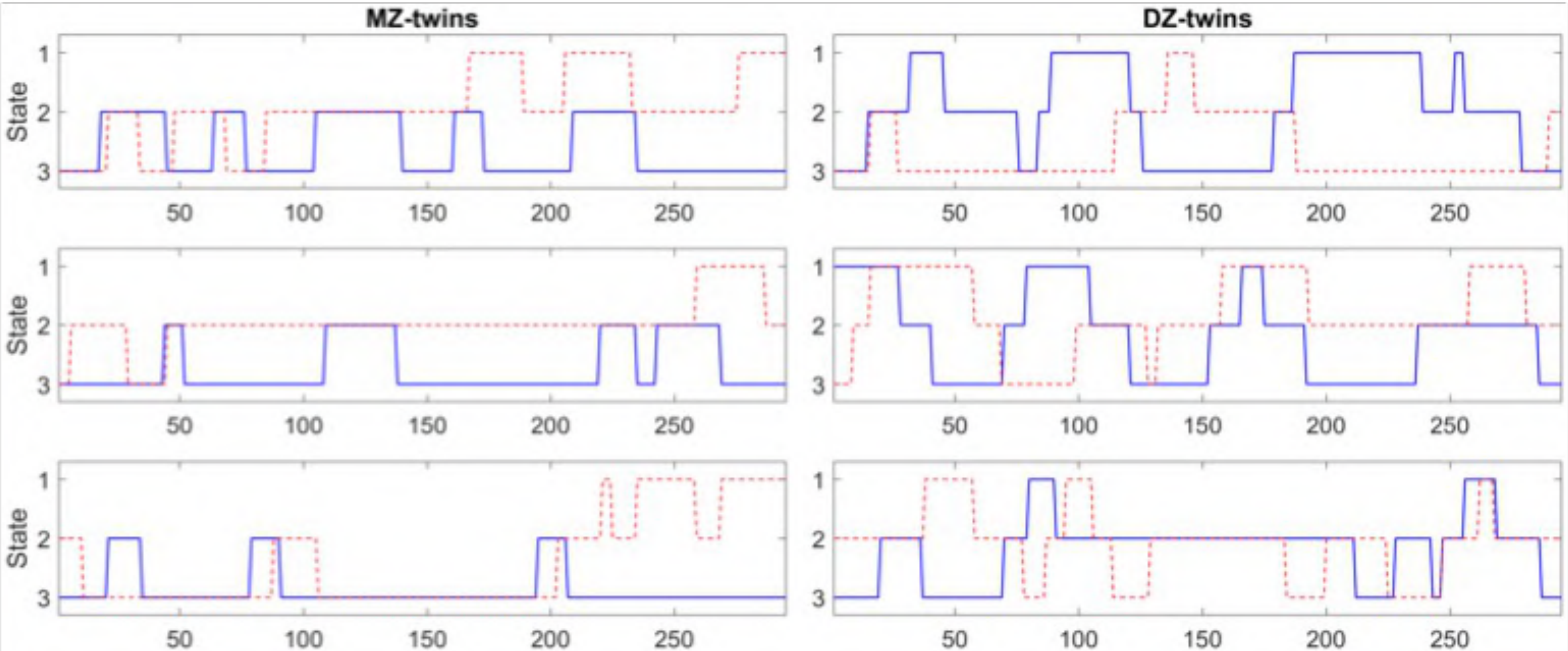
k-means *Sample mean in each state*

Topological clustering *Topological mean in each state*

State space estimation on 479 subjects

Chung at al. 2023 arXiv:2201.00087

Is the state-change heritable?



UW-Madison twin study (200 twin pairs)

ACE genetic model for twins

MZ-twins share 100% of genes
DZ-twins share 50% of genes

$$\underbrace{\rho_{\mathrm{MZ}}}_{\text{Twin correlation}} = \underbrace{A}_{\text{Additive genetics}} + \underbrace{C}_{\text{Common environment}}$$

$$\rho_{\mathrm{DZ}} = A/2 + C$$

Falconer's formula for heritability index (HI)

$$\mathrm{HI} = A = 2(\rho_{\mathrm{MZ}} - \rho_{\mathrm{DZ}})$$

Genetic control over the resting brain

D. C. Glahn[a,b,1], A. M. Winkler[a,b], P. Kochunov[c], L. Almasy[d], R. Duggirala[d], M. A. Carless[d], J. C. Curran[d], R. L. Olvera[e], A. R. Laird[c], S. M. Smith[f], C. F. Beckmann[f,g], P. T. Fox[c], and J. Blangero[d]

[a]Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, Hartford, CT 06106; [b]Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511; [c]Research Imaging Institute, University of Texas Health Science Center, San Antonio, TX 78229; [d]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78245; [e]Department of Psychiatry, University of Texas Health Science Center, San Antonio, TX 78229; [f]Functional Magnetic Resonance Imaging of the Brain (FMRIB) Centre, University of Oxford, Oxford OX3 9DU, United Kingdom; and [g]Department of Clinical Neuroscience, Imperial College, Hammersmith Campus, London W12 0NN, United Kingdom

Edited by Marcus E. Raichle, Washington University, St. Louis, MO, and approved December 10, 2009 (received for review August 31, 2009)

Table 2. Heritability estimates for regions within the default mode

	Functional connectivity		Gray-matter density	
Region*	Heritability†	*P* value‡	Heritability†	*P* value‡
Posterior cingulate/precuneus	0.423 (0.17)	4.4×10^{-3}	0.623 (0.16)	6.8×10^{-5}
Medial prefrontal cortex	0.376 (0.15)	3.8×10^{-3}	0.631 (0.15)	5.3×10^{-6}
Left temporal–parietal region	0.331 (0.19)	3.1×10^{-2}	0.387 (0.21)	3.1×10^{-2}
Right temporal–parietal region	0.420 (0.16)	3.5×10^{-3}	0.365 (0.21)	3.4×10^{-2}
Left cerebellum	0.104 (0.13)	2.0×10^{-1}	0.493 (0.15)	4.9×10^{-4}
Right cerebellum	0.304 (0.16)	1.6×10^{-2}	0.596 (0.14)	1.6×10^{-5}
Cerebellar tonsil	0.219 (0.19)	1.1×10^{-1}	0.271 (0.16)	3.2×10^{-2}
Left parahippocampal gyrus	0.273 (0.14)	1.7×10^{-2}	0.420 (0.18)	7.5×10^{-3}

*Bolded figures are significant at 5% FDR.
†Estimated heritability, h2 (SE).
‡*P* value for the heritability estimate.

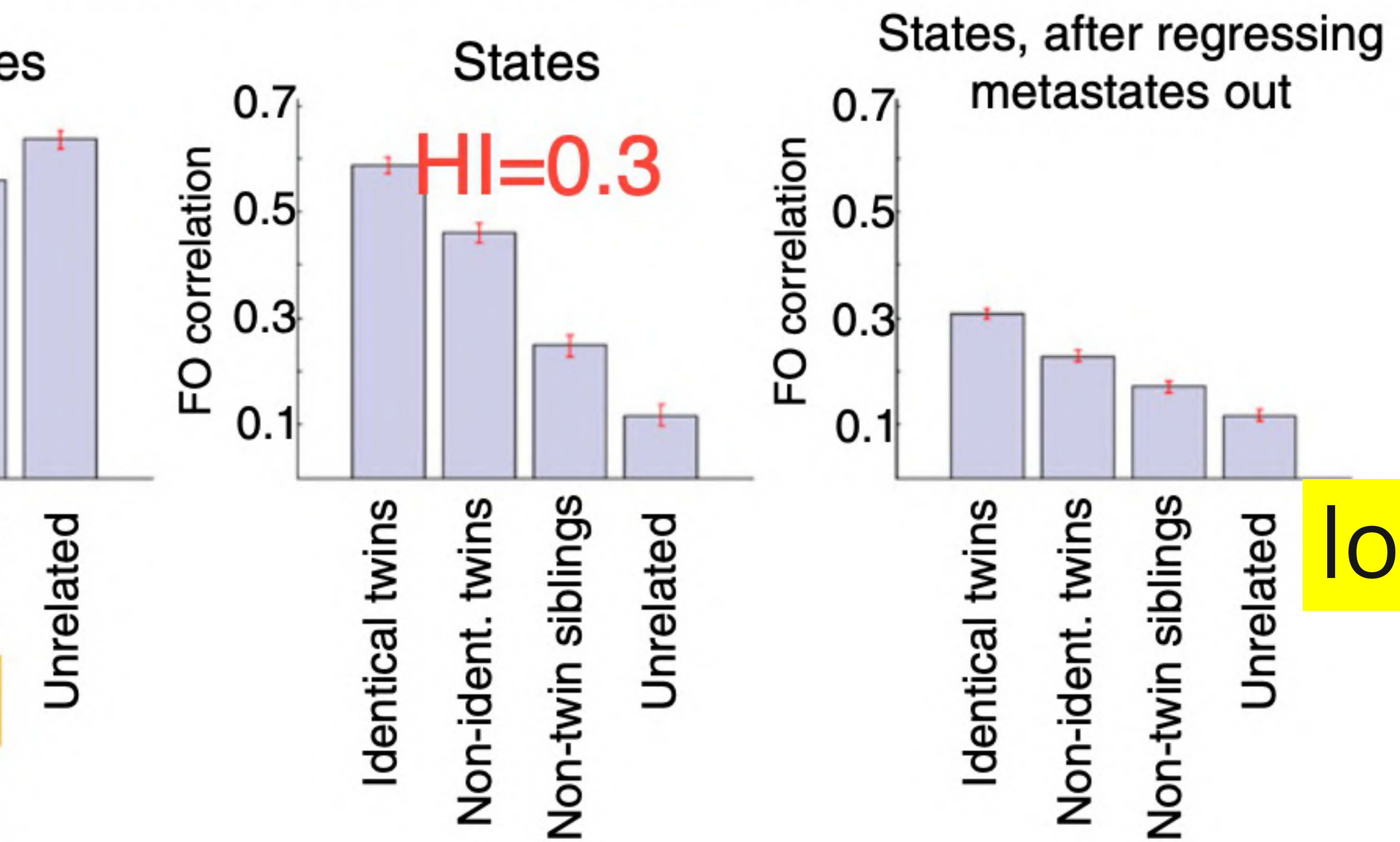
low heritability index

Brain network dynamics are hierarchically organized in time

Diego Vidaurre[a,1], Stephen M. Smith[b], and Mark W. Woolrich[a,b]

[a]Oxford Centre for Human Brain Activity (OHBA), Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxfor OX3 7JX, United Kingdom; and [b]Oxford Centre for Functional MRI of the Brain (FMRIB), Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, United Kingdom

Metastates and states are heritable



Hidden Markov model (HMM)

low heritability index

PNAS 2017 114:12827-12832

OXFORD

Cerebral Cortex, December 2017;27: 5626–5634

doi: 10
Advan
Origin

ORIGINAL ARTICLE

Heritability of the Effective Con Resting-State Default Mode Net

Junhai Xu[1,2,3], Xuntao Yin[2], Haitao Ge[2], Yar
Baolin Liu[1], Shuwei Liu[2] and Karl Friston[3]

[1]Tianjin Key Laboratory of Cognitive Computing and Applicati
Technology, Tianjin University, Tianjin 300350, P.R. China, [2]Re
Anatomy, Shandong University School of Medicine, Jinan, Sha
Neuroimaging, Institute of Neurology, University College Lond
Affiliated Hospital of Medical College, Qingdao University, Qin
Epidemiology, Qingdao Municipal Central for Disease Control

Address correspondence to Prof. Baolin Liu, Tianjin Key Laboratory of Cognitive Comp
Technology, Tianjin University, Tianjin, 300350, P. R. China. Email: liubaolin@tsinghua
Imaging Anatomy, Shandong University School of Medicine, Jinan, Shandong, 250012,

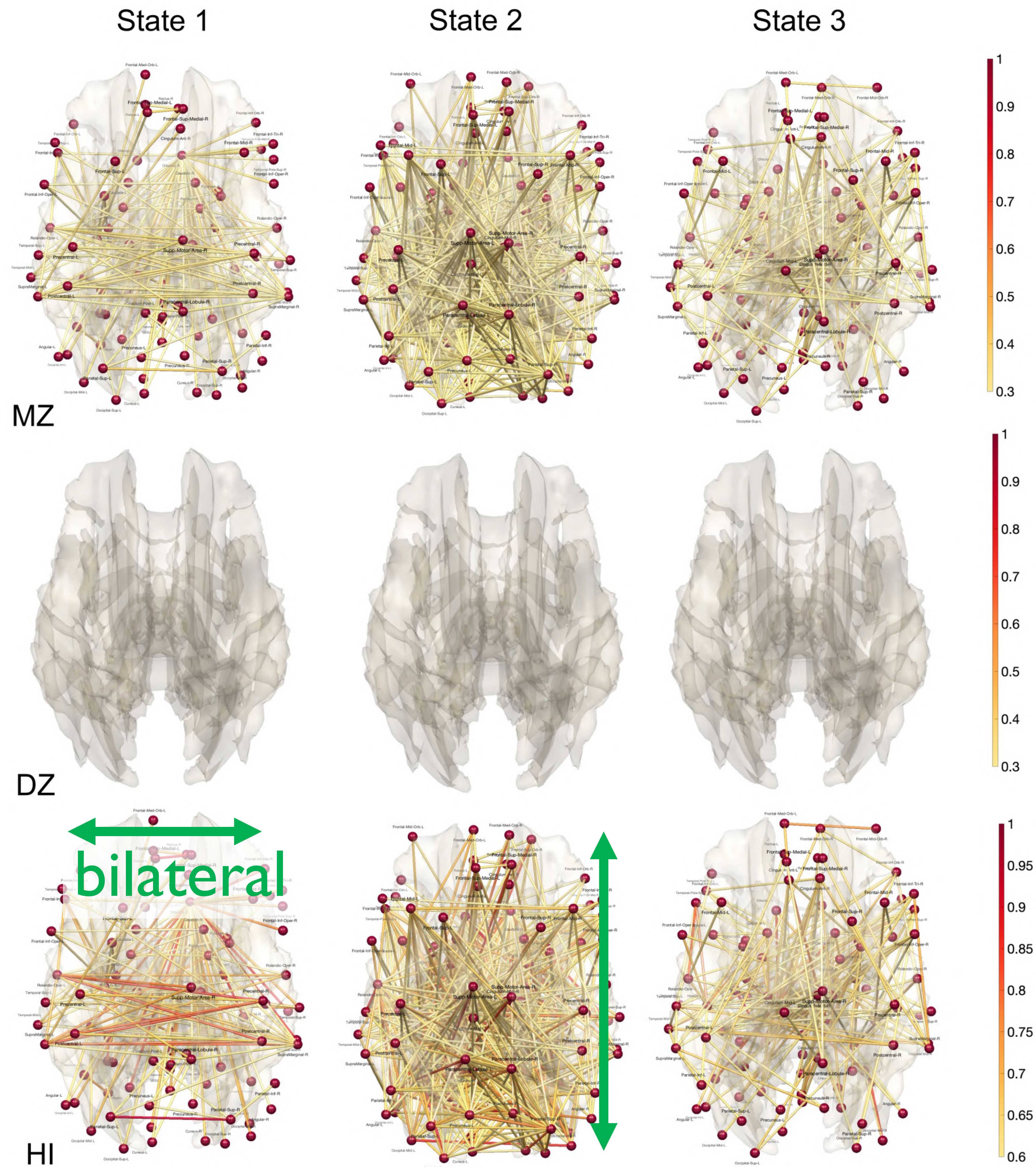
Tables 3 Additive genetics/common environment/unique environment (A/C/E) model estimates and significant value for each effective connectivity

DMN effective connectivity	r_{MZ}	r_{DZ}	V_A	V_C	V_E	P values
PCC-> mPFC	0.79	0.58	0.44	0.32	0.24	0.37
PCC->LPC	0.76	0.63	0.40	0.33	0.28	<0.001*
PCC->RPC	0.94	0.53	0.56	0.36	0.08	<0.001*
PCC->PCC	0.78	0.55	0.18	0.49	0.33	0.025*
mPFC->PCC	0.86	0.55	0.76	0.09	0.15	0.068
mPFC->LPC	0.62	0.49	0.15	0.44	0.42	0.92
mPFC->RPC	0.34	0.72	0.00	0.54	0.46	0.28
mPFC->mPFC	0.56	0.14	0.51	0.00	0.49	0.43
LPC->PCC	0.89	0.6	0.44	0.42	0.14	<0.001*
LPC->mPFC	0.82	0.61	0.39	0.40	0.21	0.33
LPC->RPC	0.75	0.51	0.38	0.30	0.32	0.077
LPC->LPC	0.84	0.71	0.24	0.58	0.18	0.43
RPC->PCC	0.68	0.72	0.36	0.44	0.20	<0.001*
RPC->mPFC	0.61	0.63	0.33	0.39	0.28	0.24
RPC->LPC	0.47	0.39	0.38	0.15	0.47	0.83
RPC->RPC	0.75	0.55	0.37	0.35	0.28	0.65

Structural equation models

dditive genetics variance estimate; V_C, common environment ate; V_E, unique environment variance estimate.

low heritability index



MZ-twin correlation

DZ-twin correlation

Heritability index

Chung et al. 2024
arXiv:2201.00087 (PLOS Computational Biology)

Thank you.

We are located here. Join us for
postdoc and *graduate research*
if you are better than chat-GPT