



University of Wisconsin  
**SCHOOL OF MEDICINE  
AND PUBLIC HEALTH**

# Geometric & Topological Distances

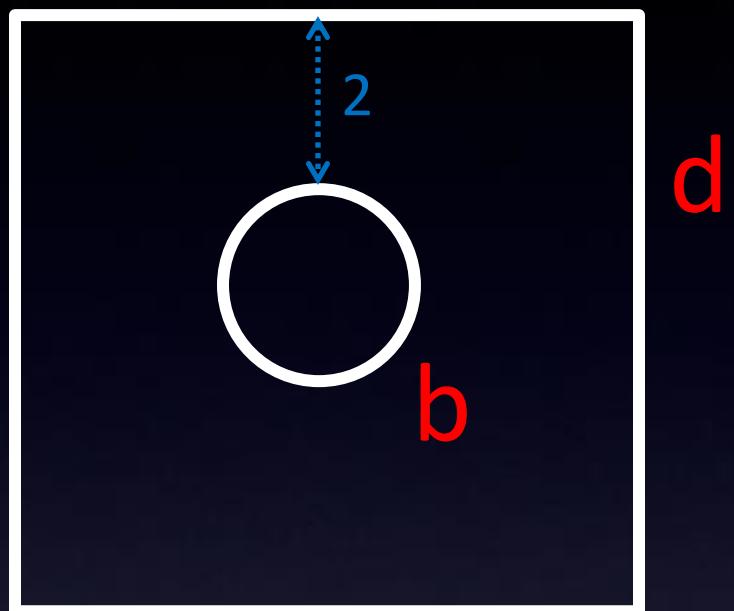
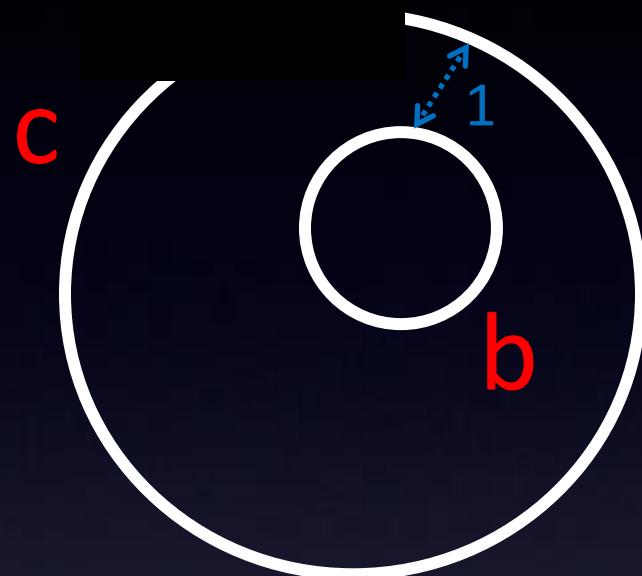
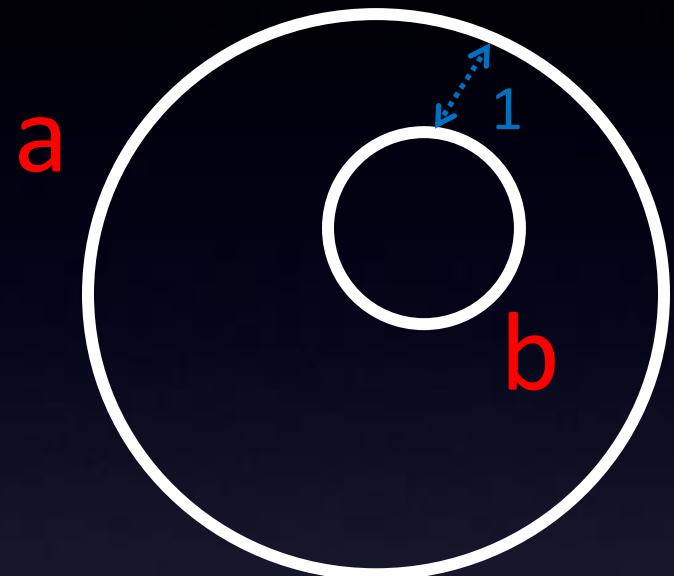
Moo K. Chung  
University of Wisconsin-Madison

[www.stat.wisc.edu/~mchung](http://www.stat.wisc.edu/~mchung) →  
[github.com/laplcebeltrami](https://github.com/laplcebeltrami)

# Geometric distance vs. Topological distances

$d_{geo}$

$d_{top}$



$$d_{geo}(a, b) = 1$$

$$d_{geo}(c, b) = 1$$

$$d_{geo}(d, b) = 2$$

$$d_{top}(a, b) = 0$$

$$d_{top}(c, b) = 1$$

$$d_{top}(d, b) = 0$$

Wasserstein  
distance  
(optimal  
transport)

# 2-Wasserstein distance between scatter points

Random variables:

$$X \sim f_1 \quad Y \sim f_2$$

2-Wasserstein distance:

$$\mathcal{D}(X, Y) = \left( \inf \mathbb{E} \|X - Y\|^2 \right)^{1/2}$$

Scatter points

$$P_1 = \{x_1, \dots, x_q\} \subset \mathbb{R}^2 \qquad P_2 = \{y_1, \dots, y_q\} \in \mathbb{R}^2$$

Empirical distributions

$$f_1(x) = \frac{1}{q} \sum_{i=1}^q \delta(x - x_i) \qquad f_2(y) = \frac{1}{q} \sum_{i=1}^q \delta(y - y_i)$$

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

Optimal  
bijection

Assignment problem: Hungarian algorithm  $\mathcal{O}(q^3)$

# Registration of suctal patterns via Wasserstein distance

Random variables:

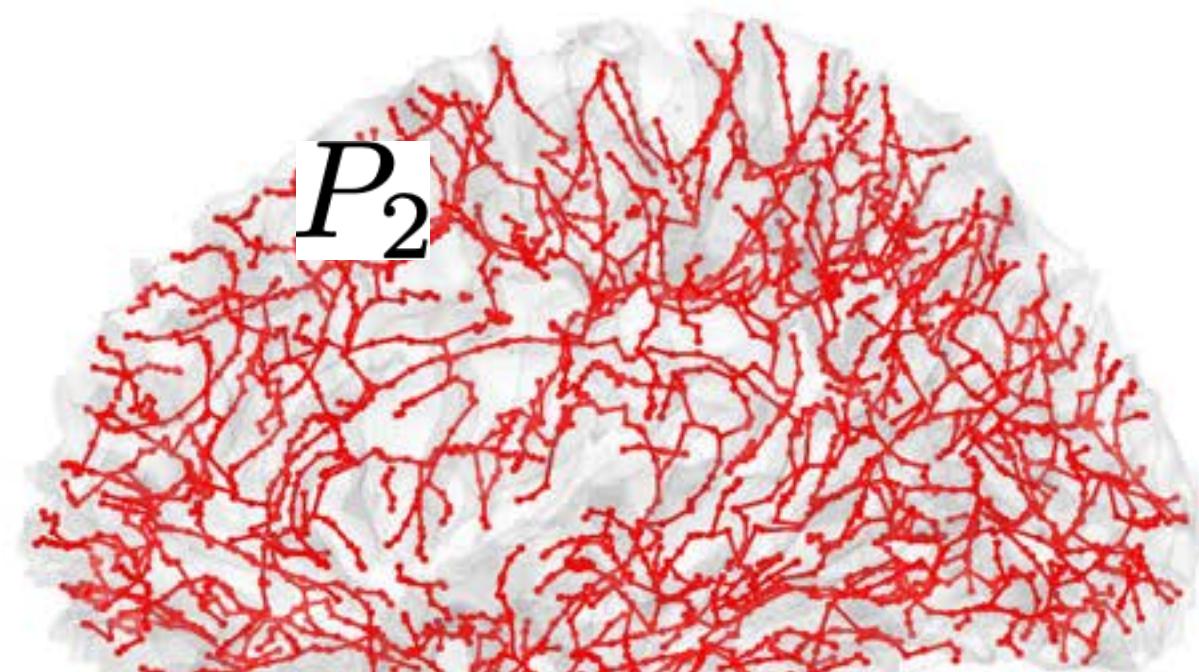
$$X \sim f_1 \quad Y \sim f_2$$

2-Wasserstein distance:

$$d(X, Y) = \left( \inf_{f(X, Y)} \mathbb{E} \|X - Y\|^2 \right)^{1/2}$$



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

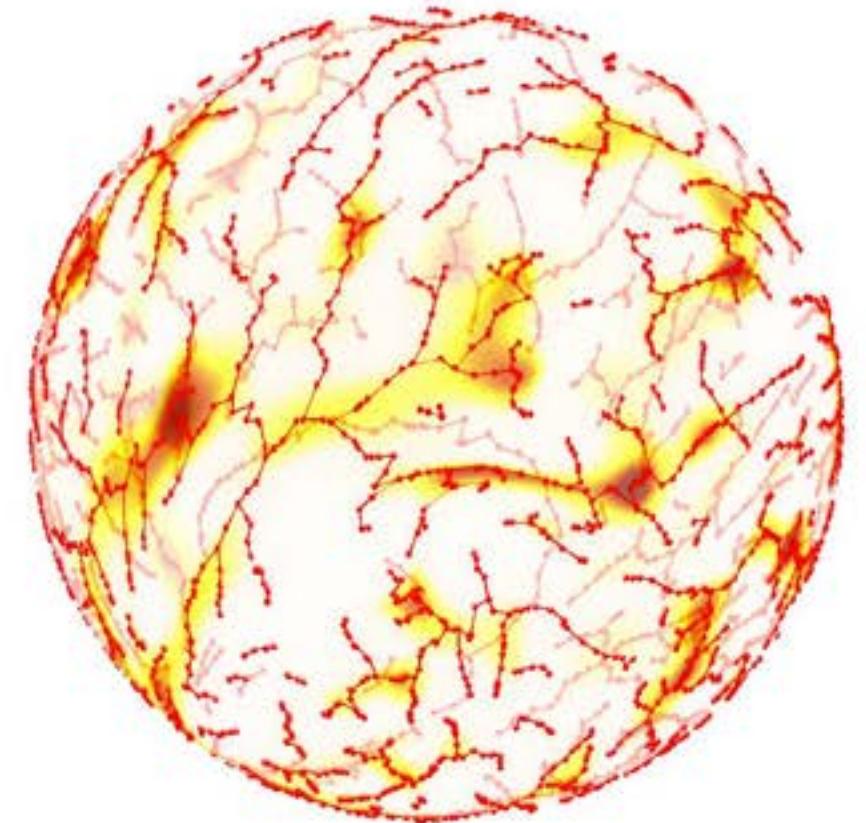
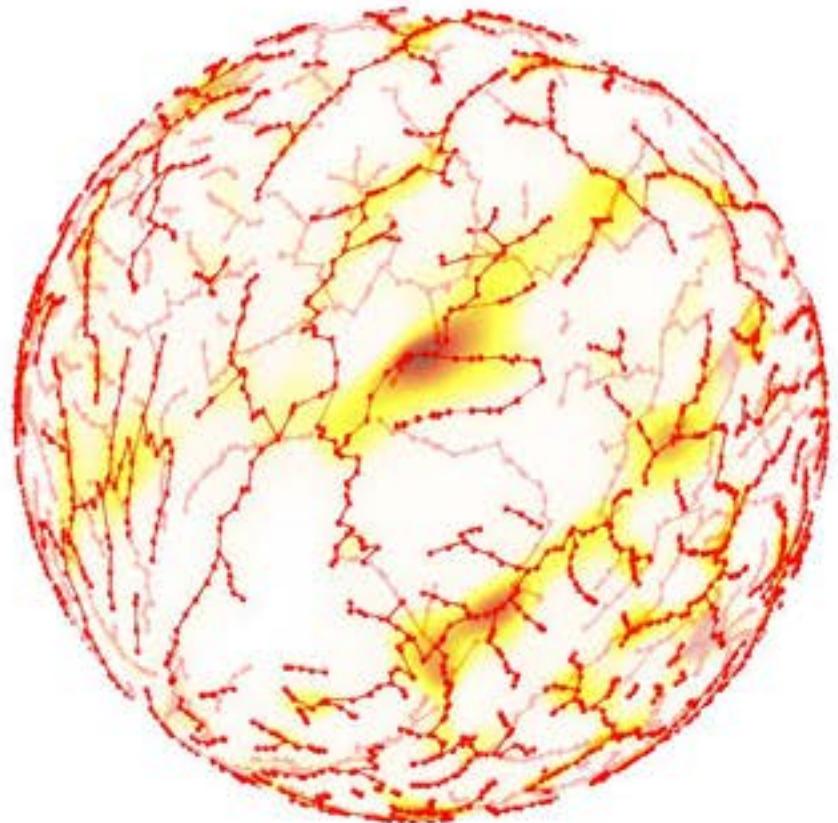


$$f_2(y) = \frac{1}{q} \sum_{i=1}^q \delta(y - y_i)$$

Finding optimal permutation  $\psi$ :

$$d(P_1, P_2) = \inf_{\psi: P_1 \rightarrow P_2} \left( \sum_{x \in P_1} \|x - \psi(x)\|^2 \right)^{1/2}$$

# Diffusion before applying Wasserstein distance



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



$$K_\sigma * f_1(x) = \frac{1}{q} \sum_{i=1}^q K_\sigma(x, x_i)$$

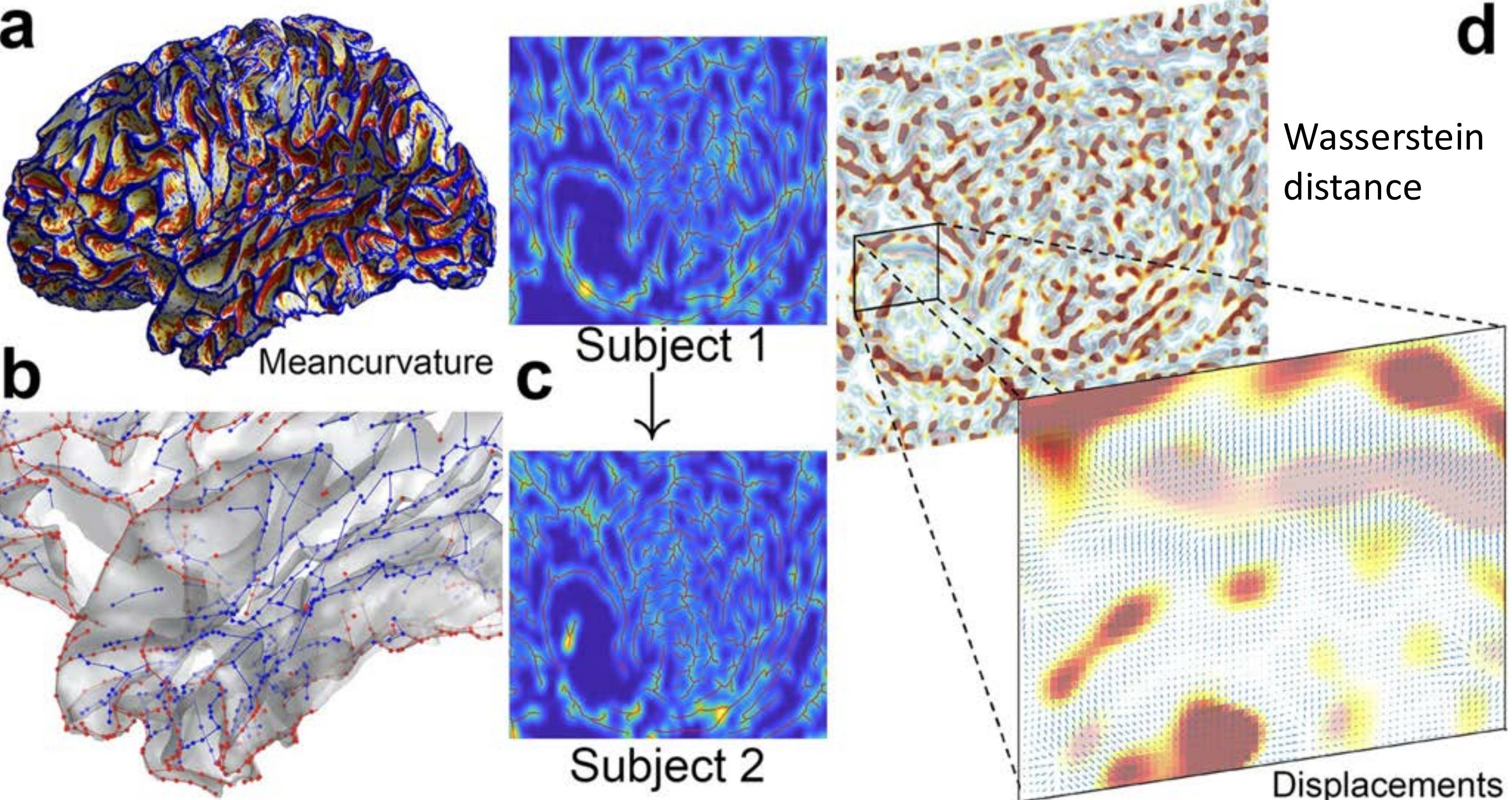
$$f_2(y) = \frac{1}{q} \sum_{i=1}^q \delta(y - y_i)$$



$$K_\sigma * f_2(y) = \frac{1}{q} \sum_{i=1}^q K_\sigma(y, y_i)$$

$$\mathcal{D}(f_1, f_2) = \mathcal{D}(K_\sigma * f_1, K_\sigma * f_2)$$

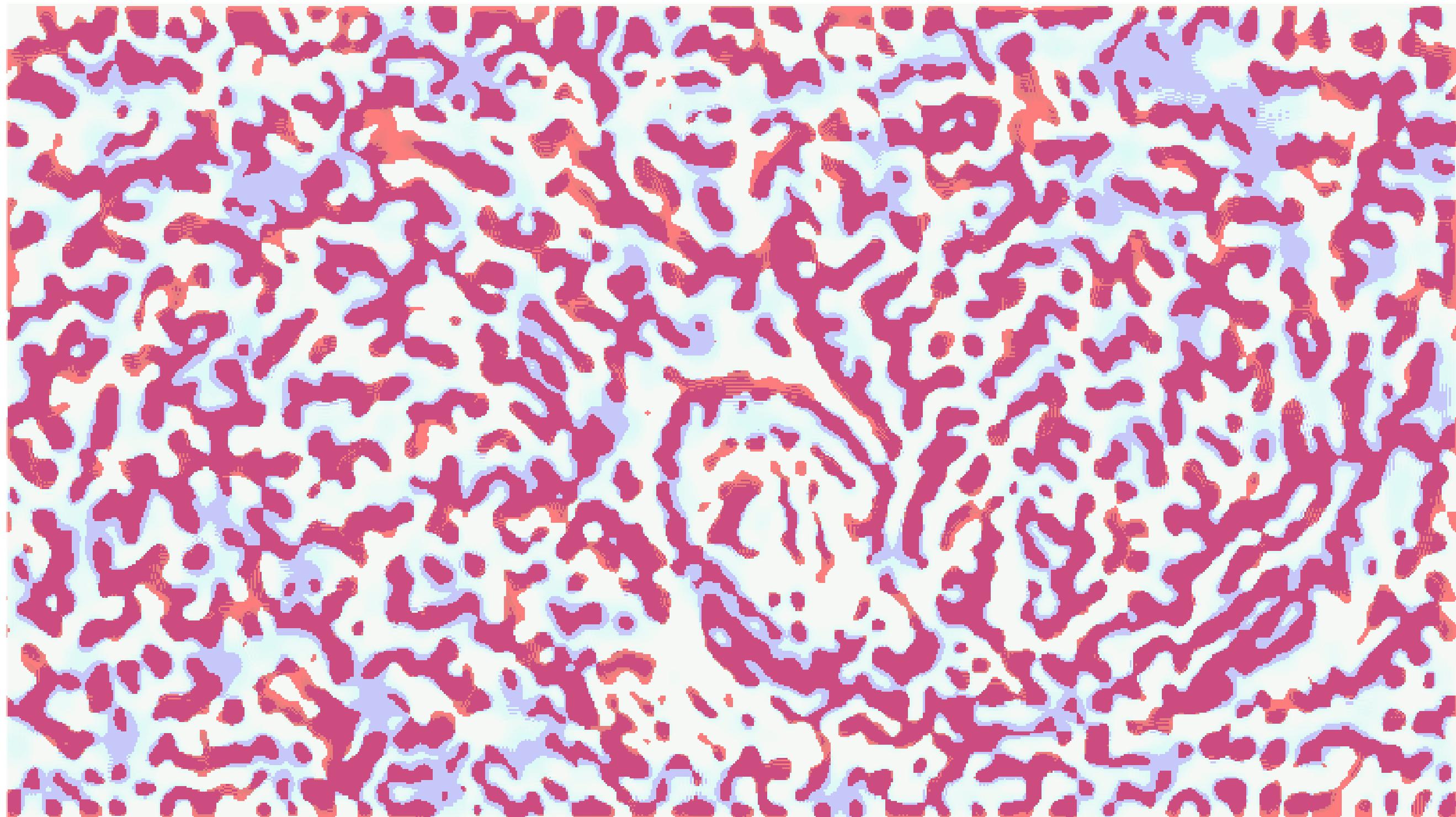
# Registration of sulcal/gyral patterns



Huang et al. 2020 IEEE Transactions on Medical Imaging

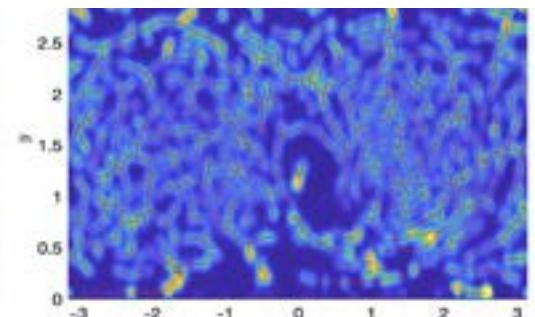
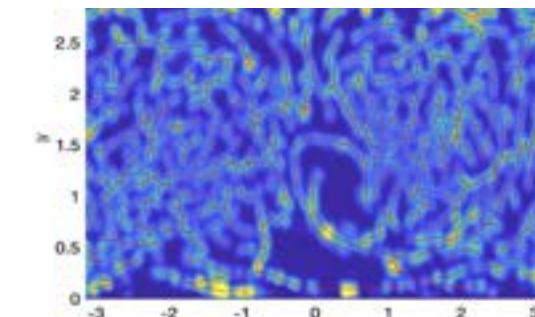
Chen et al. 2023 arXiv:2307.00385

# Wasserstein distance = Optimal transport on probability



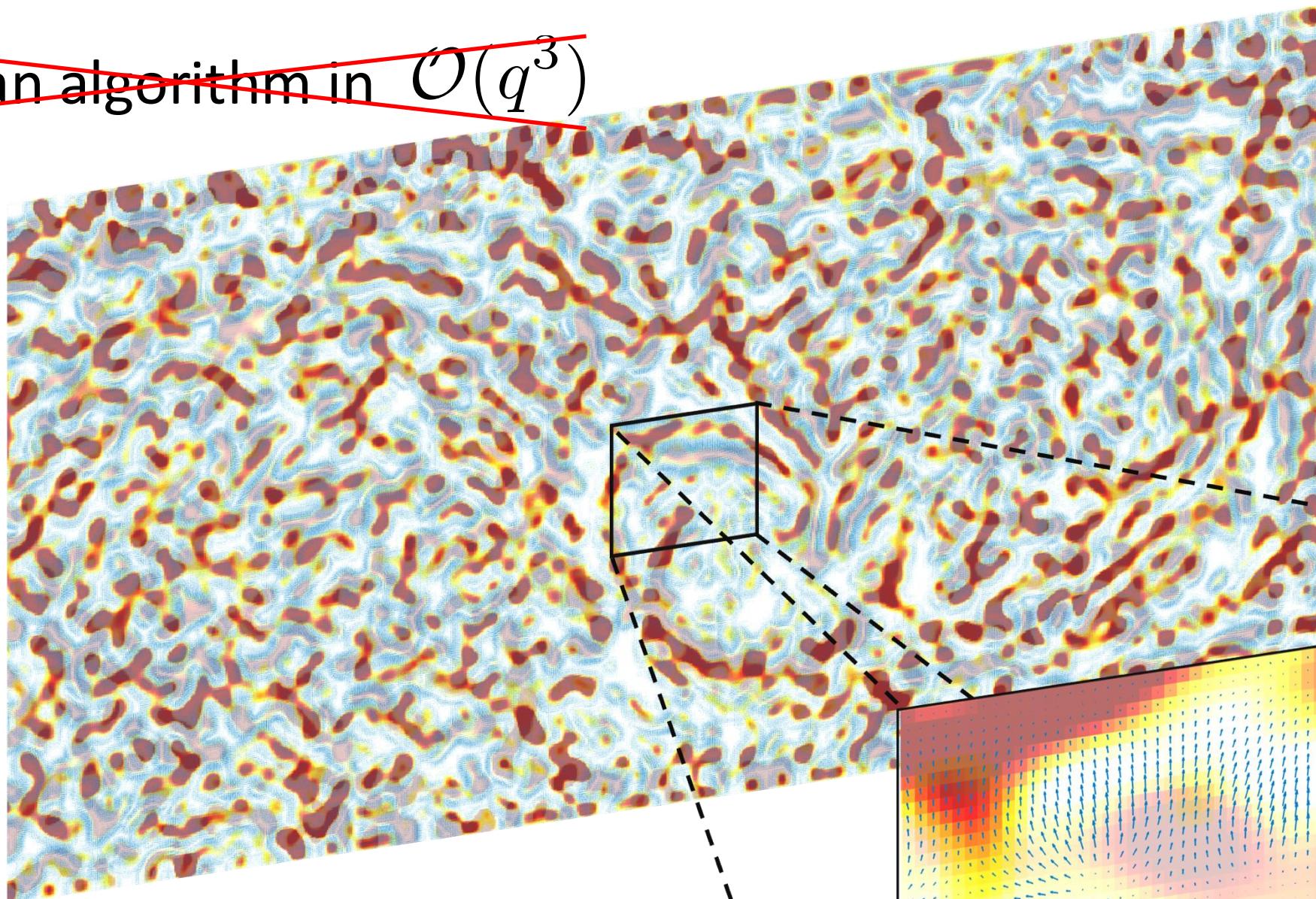
develop the gradient descent algorithms for estimating the deformation field. We further quantify the image registration performance. This method is applied in identifying the differences between male and female sulcal patterns.

Smooth projection



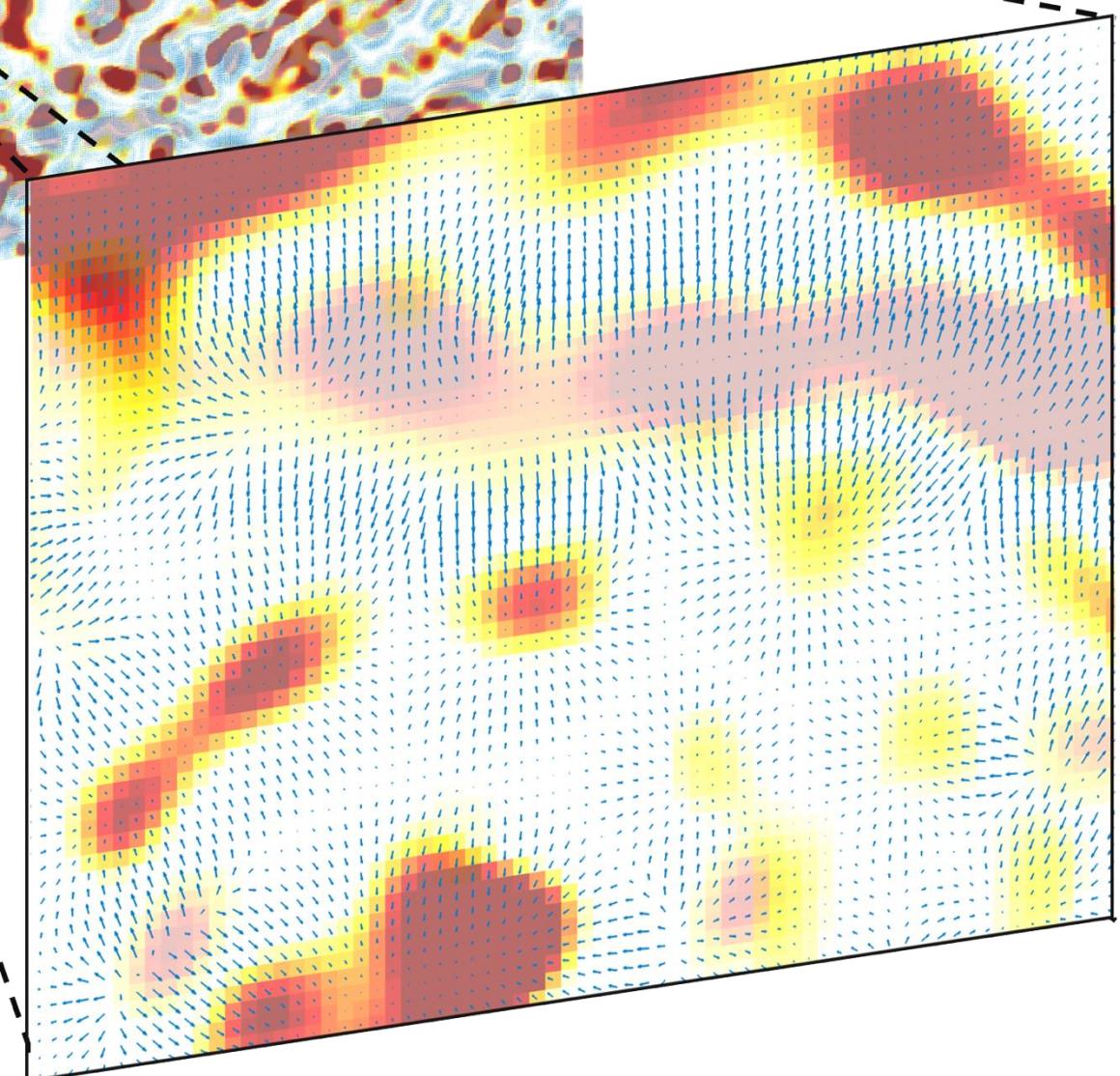
# Numerical implementation

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

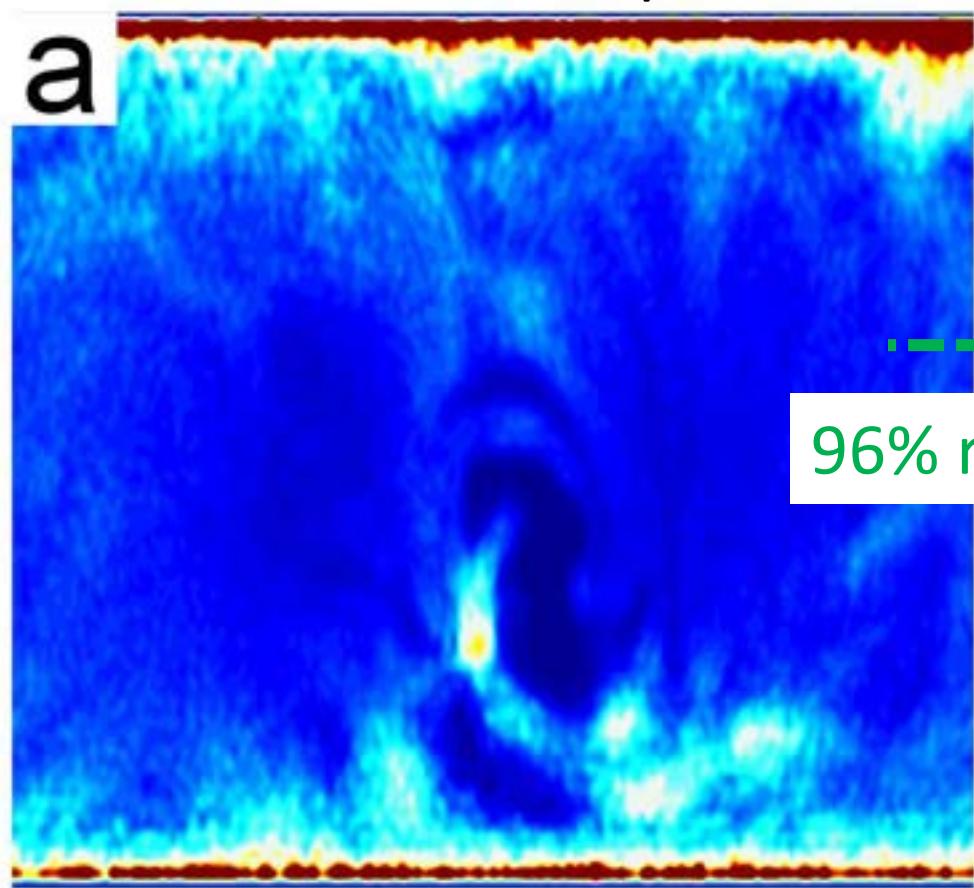


Dual convex formulation  $\mathcal{O}(1)$

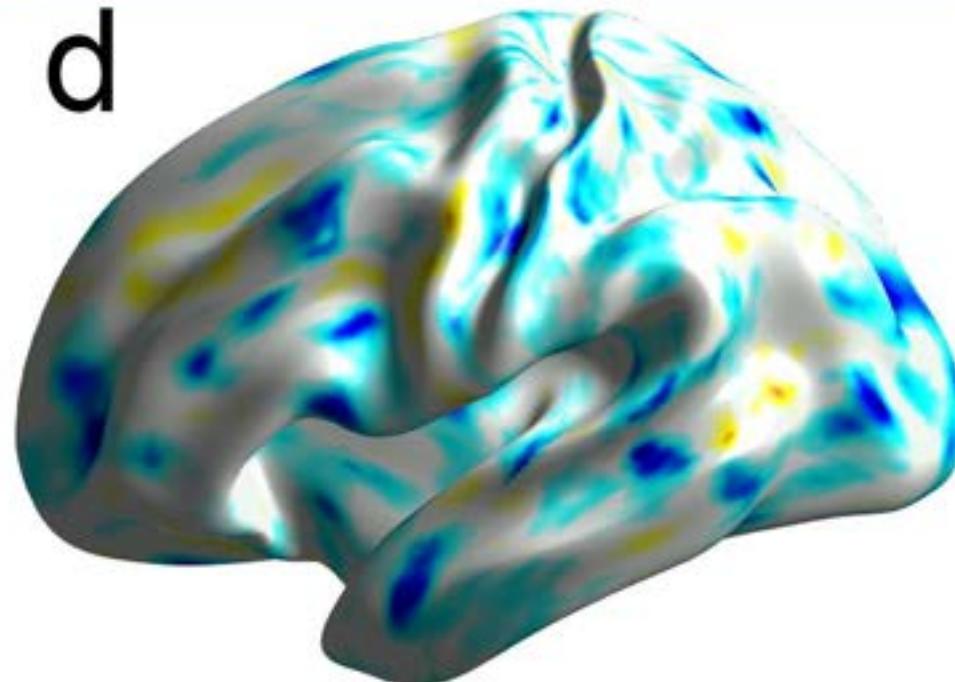
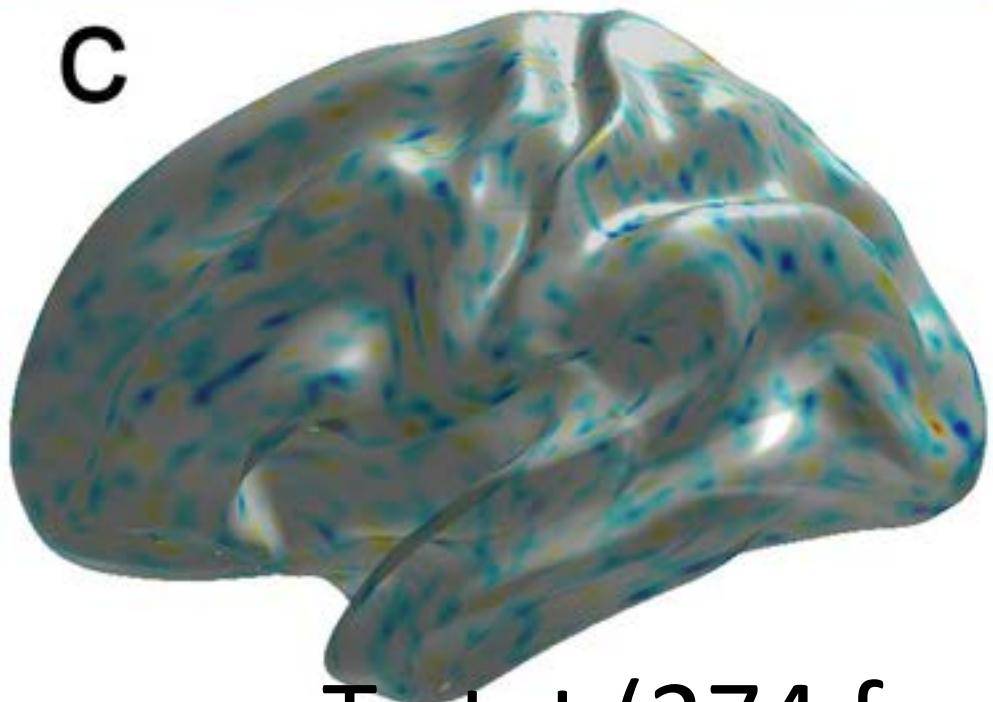
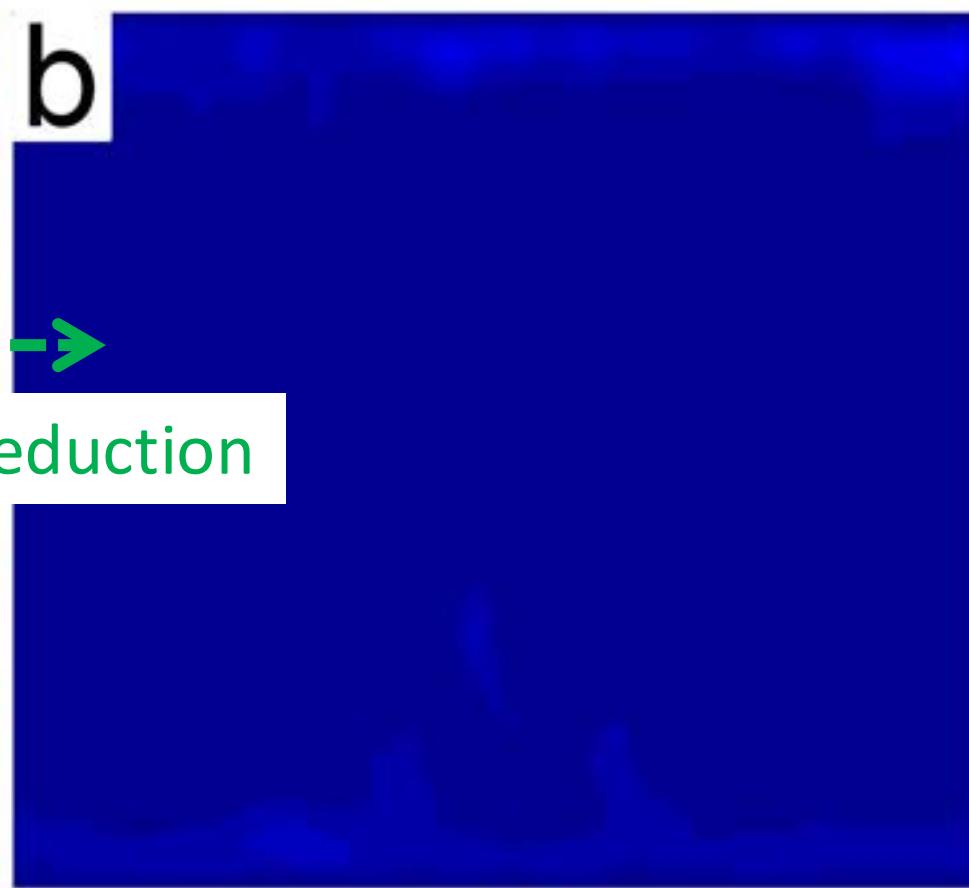
*96% reduction in variability  
compared to FreeSurfer*



Intersubject variability  
in FreeSurfer output



Intersubject variability  
after Wasserstein distance

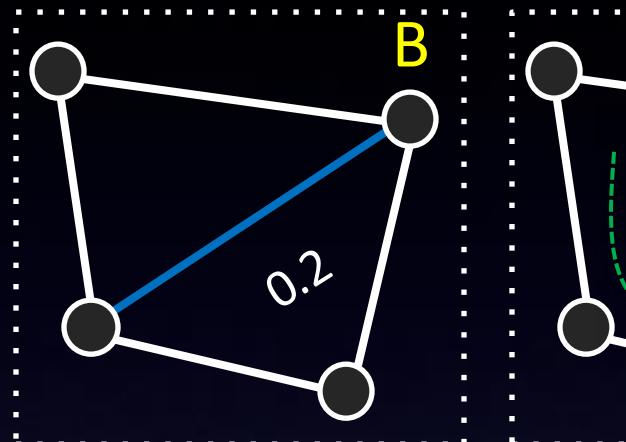


T-stat (274 females – 182 males)

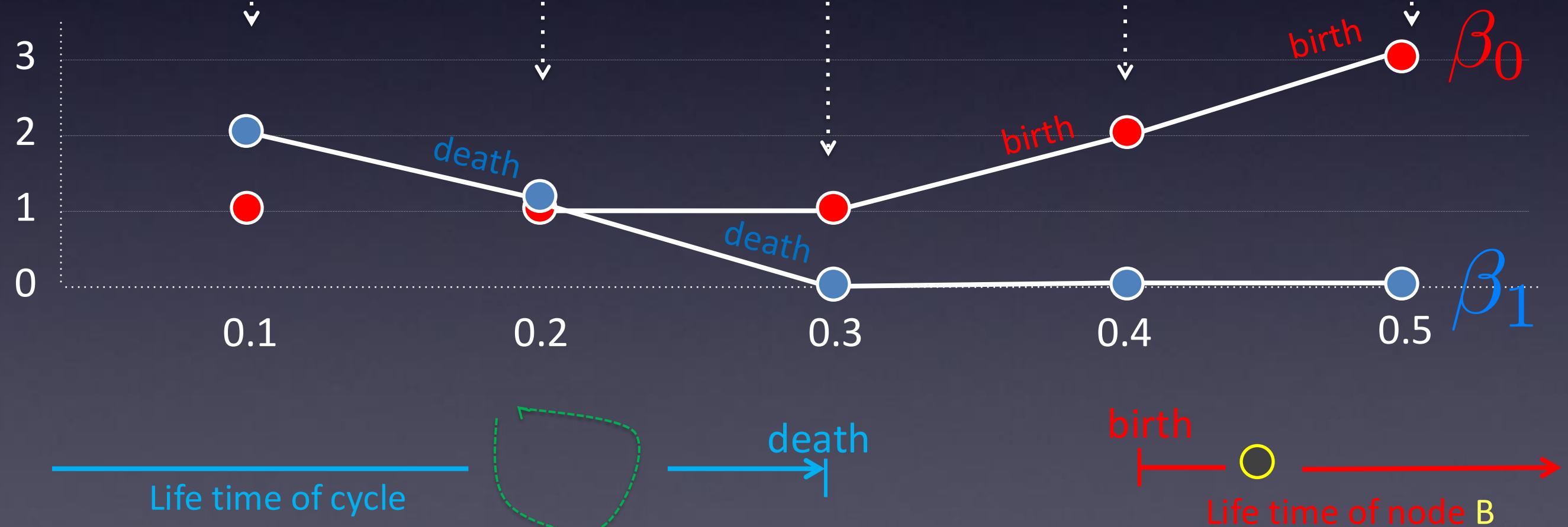
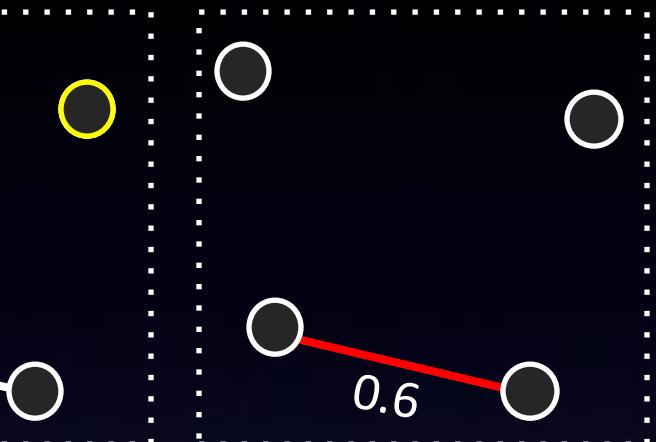
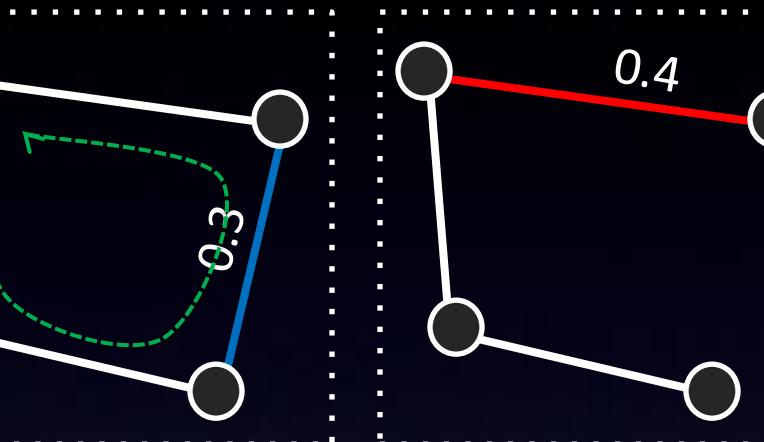
# Topological Distance

Persistence = Life time (death – birth) of a feature

Edges destroy cycles



Edges create components



# Topological registration via $r$ -Wasserstein distance



$$d(X, Y) = \left( \inf_{f(X, Y)} \mathbb{E} \|X - Y\|^r \right)^{1/r}$$



$$d_0(\Theta, P) = \min_{\tau} \sum_b |b - \tau(b)|^r$$

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

$$d_1(\Theta, P) = \min_{\tau} \sum_d |d - \tau(d)|^r$$

Proof in  
Chung et al. 2024,  
Foundations

Theorem: Wasserstein distance (optimal transport) on graph filtrations

$$\begin{aligned}\mathcal{L}_{0D}(\Theta, P) &= \min_{\tau} \sum_{b \in E_0} [b - \tau(b)]^2 \\ &= \sum_{b \in E_0} [b - \tau_0^*(b)]^2\end{aligned}$$

$\tau_0^*$  :The  $i$ -th smallest birth value to the  $i$ -th smallest birth value

$$\begin{aligned}\mathcal{L}_{1D}(\Theta, P) &= \min_{\tau} \sum_{d \in E_1} [d - \tau(d)]^2 \\ &= \sum_{d \in E_1} [d - \tau_1^*(d)]^2\end{aligned}$$

$\tau_1^*$  :The  $i$ -th smallest death value to the  $i$ -th smallest death value

Theorem:  $\infty$ -Wasserstein distance on graph filtrations

0D topology

$$\mathcal{L}_{0D}(\Theta, P) = \max_{\substack{b \in E_0 \\ \text{Birth set}}} |b - \tau_0^*(b)|$$

$\tau_0^*$  : The  $i$ -th smallest birth value to the  $i$ -th smallest birth value

1D topology

$$\mathcal{L}_{1D}(\Theta, P) = \max_{\substack{d \in E_1 \\ \text{Death set}}} |d - \tau_1^*(d)|$$

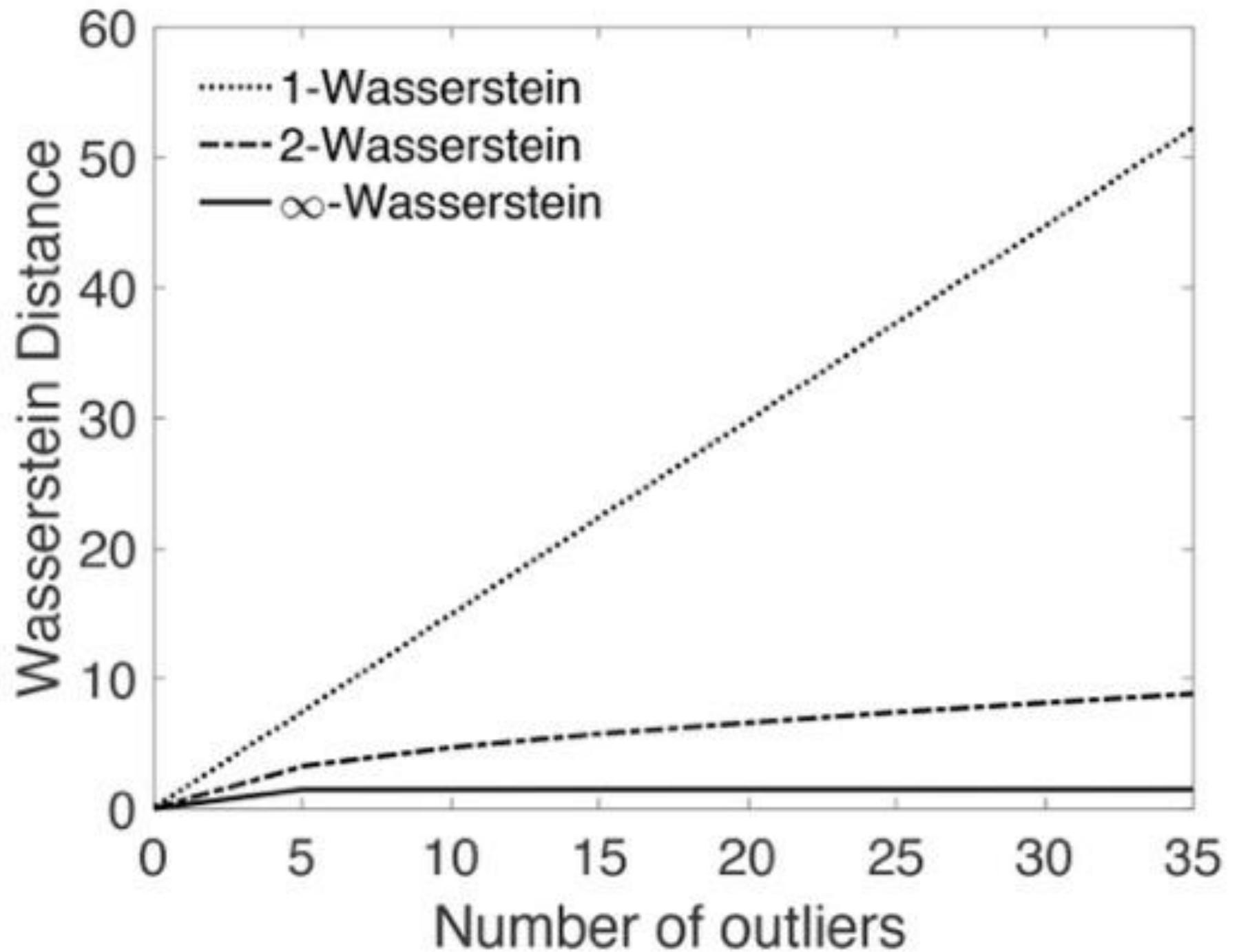
$\tau_1^*$  : The  $i$ -th smallest death value to the  $i$ -th smallest death value

# *Effect of Wasserstein distance on*

**outliers**

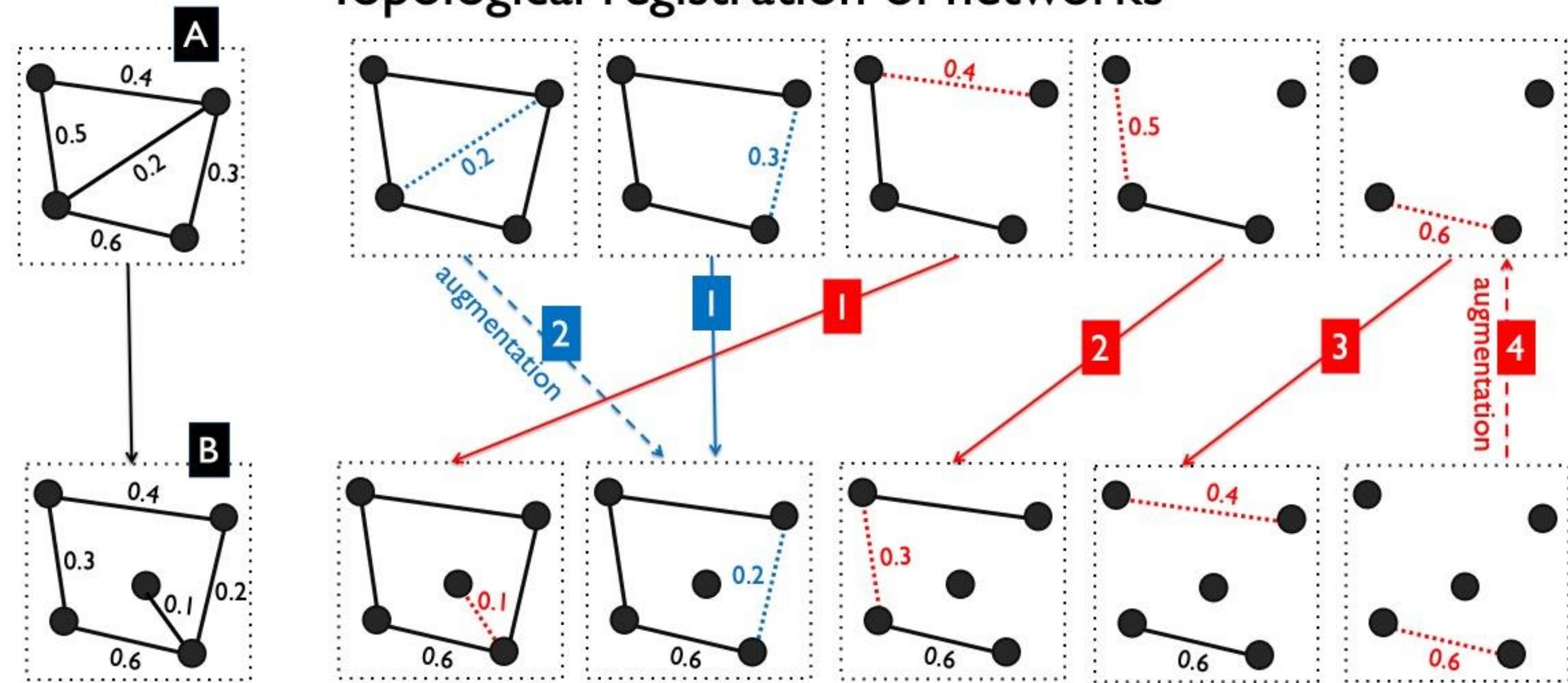
Edge weights  $\sim N(0, 0.01^2)$

Outlying edge weights  $\sim N(1.5, 0.01^2)$



# Topological registration

## Topological registration of networks

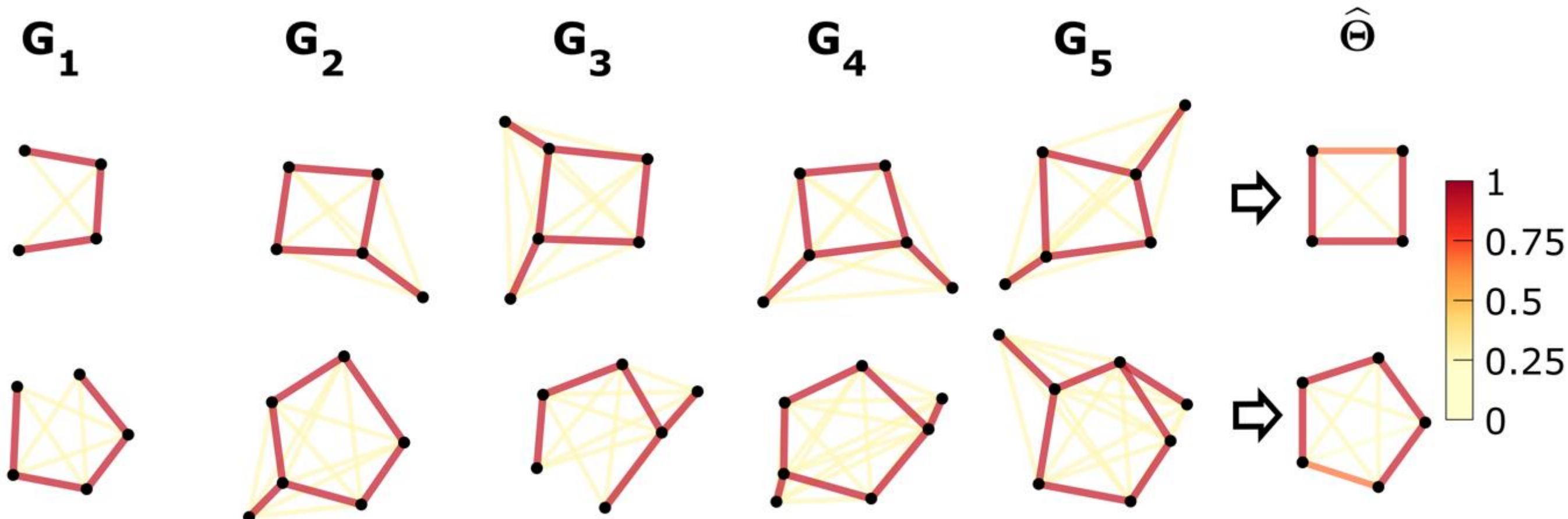


- 1) Match MST to MST, none-MST to none-MST
- 2) Match largest persistence (death – birth) first

Nonlinear registration of topologically different objects

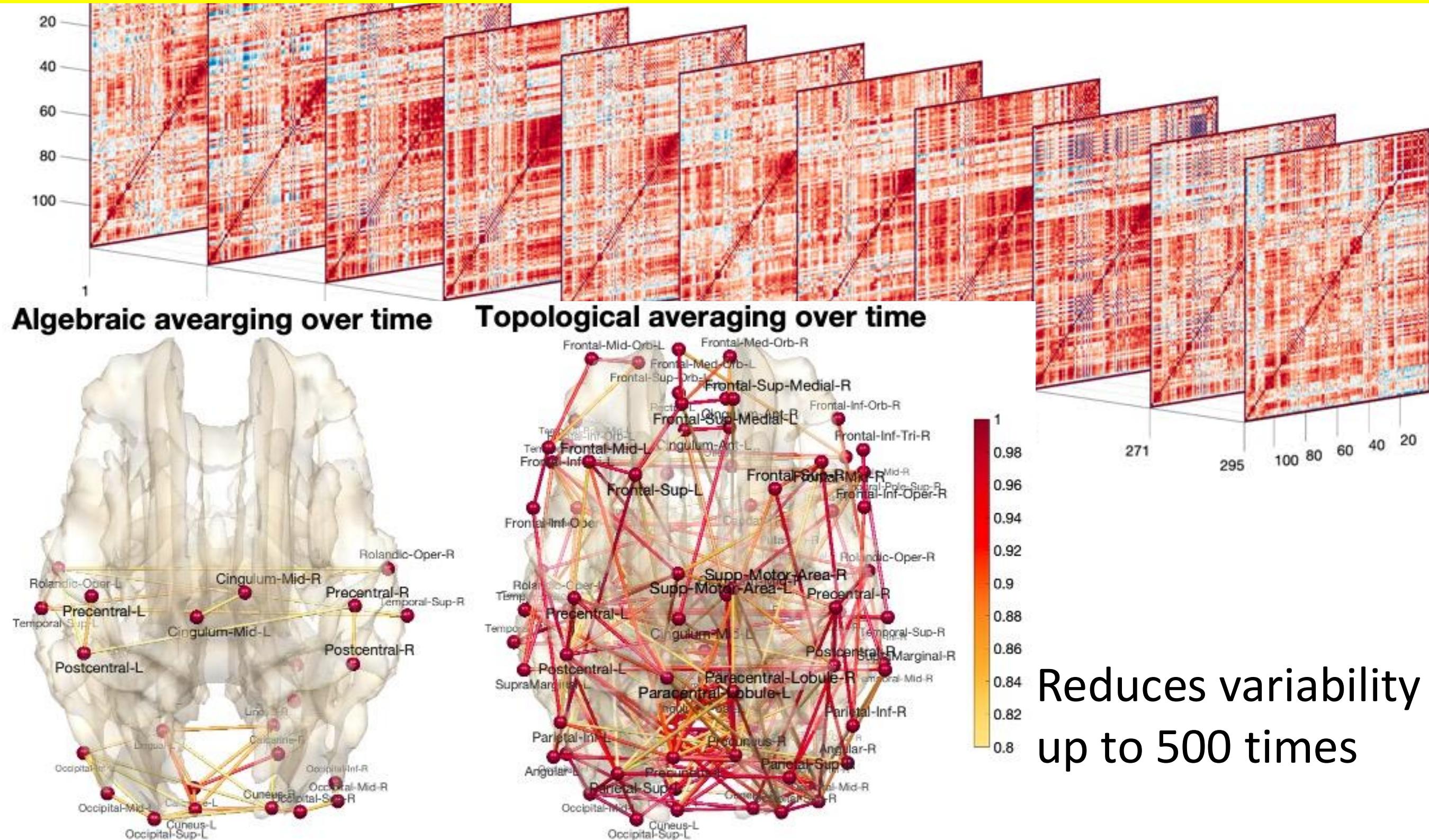
Topological mean

= minimizer of topological distance



Average sorted edge weights over MST and none-MST separately and reconstruct  $\rightarrow$  minimizer of topological distance

# Superiority of topological averaging in rs-fMRI connectivity



Reduces variability  
up to 500 times

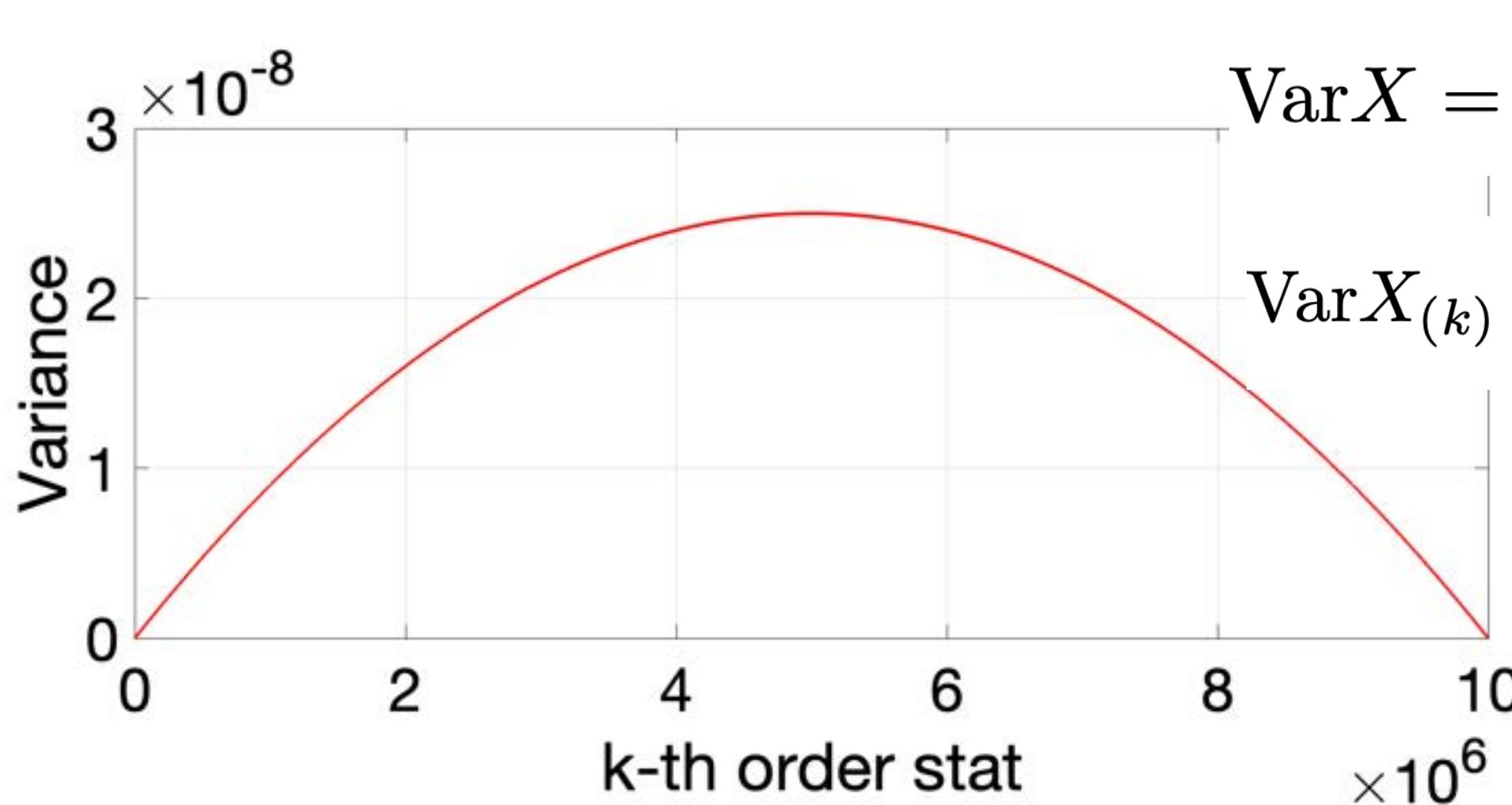
# 400 subjects

*Chung et al. 2024,*

# Why topological methods reduce variability?

For random variables ,  $X_1, \dots, X_n \sim \text{Unif}(0,1)$

Order statistics  $X_{(1)} < X_{(2)} < \dots < X_{(n)}$



$$\text{Var}X = \frac{1}{12}$$

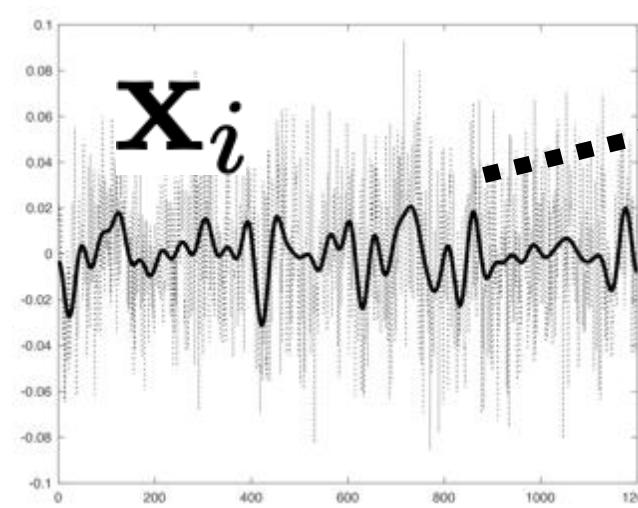
$$\text{Var}X_{(k)} = \frac{k(n - k + 1)}{(n + 1)^2(n + 2)}$$

*Variance  
reduction*

$\sim \mathcal{O}\left(\frac{1}{n}\right)$

Data  
Embedding:  
Multidimensional  
al  
Scaling

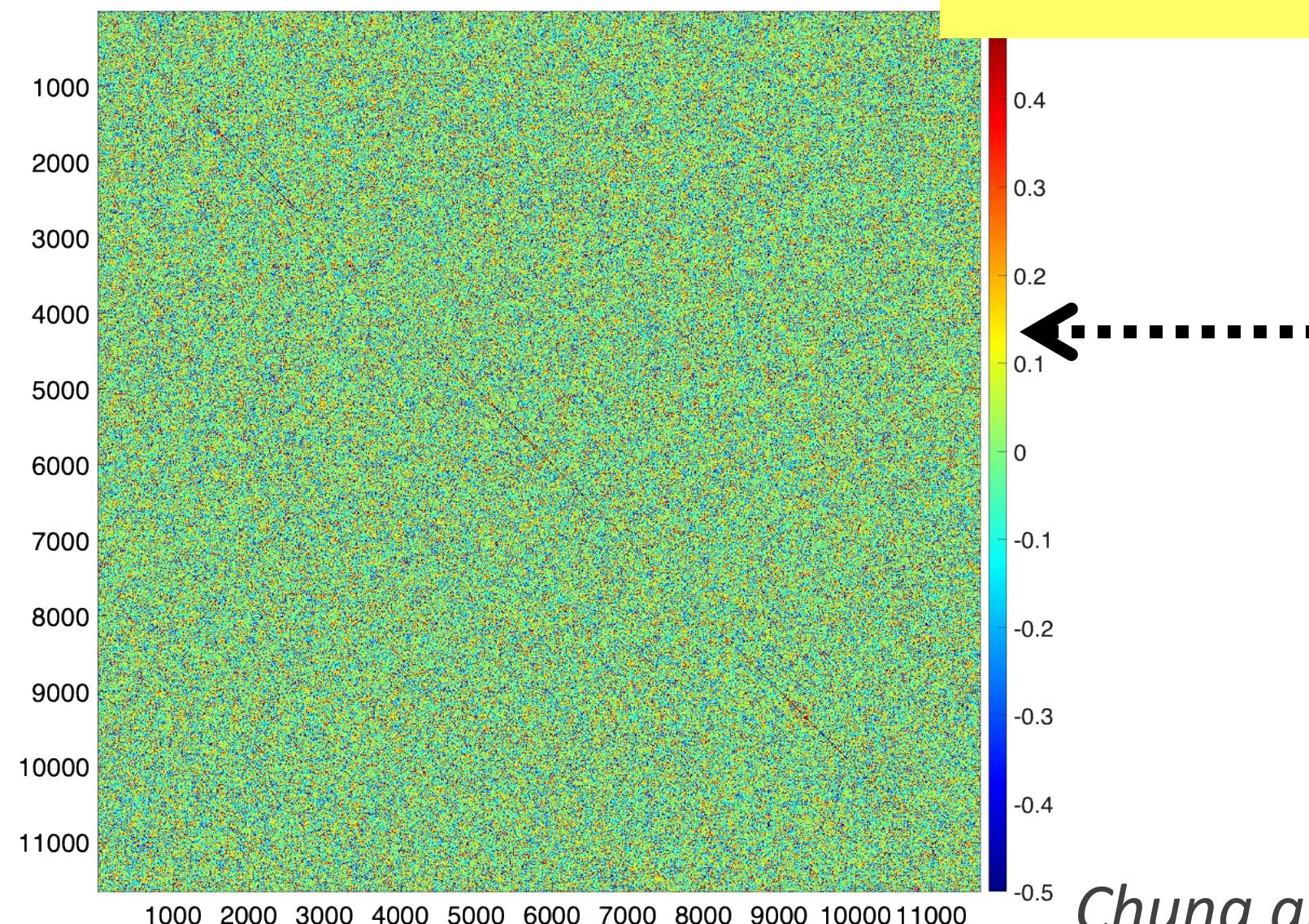
# Geometry of the space of correlation matrices



time series  $\mathbf{x}_i$  at voxel  $i$

Translate and scale:

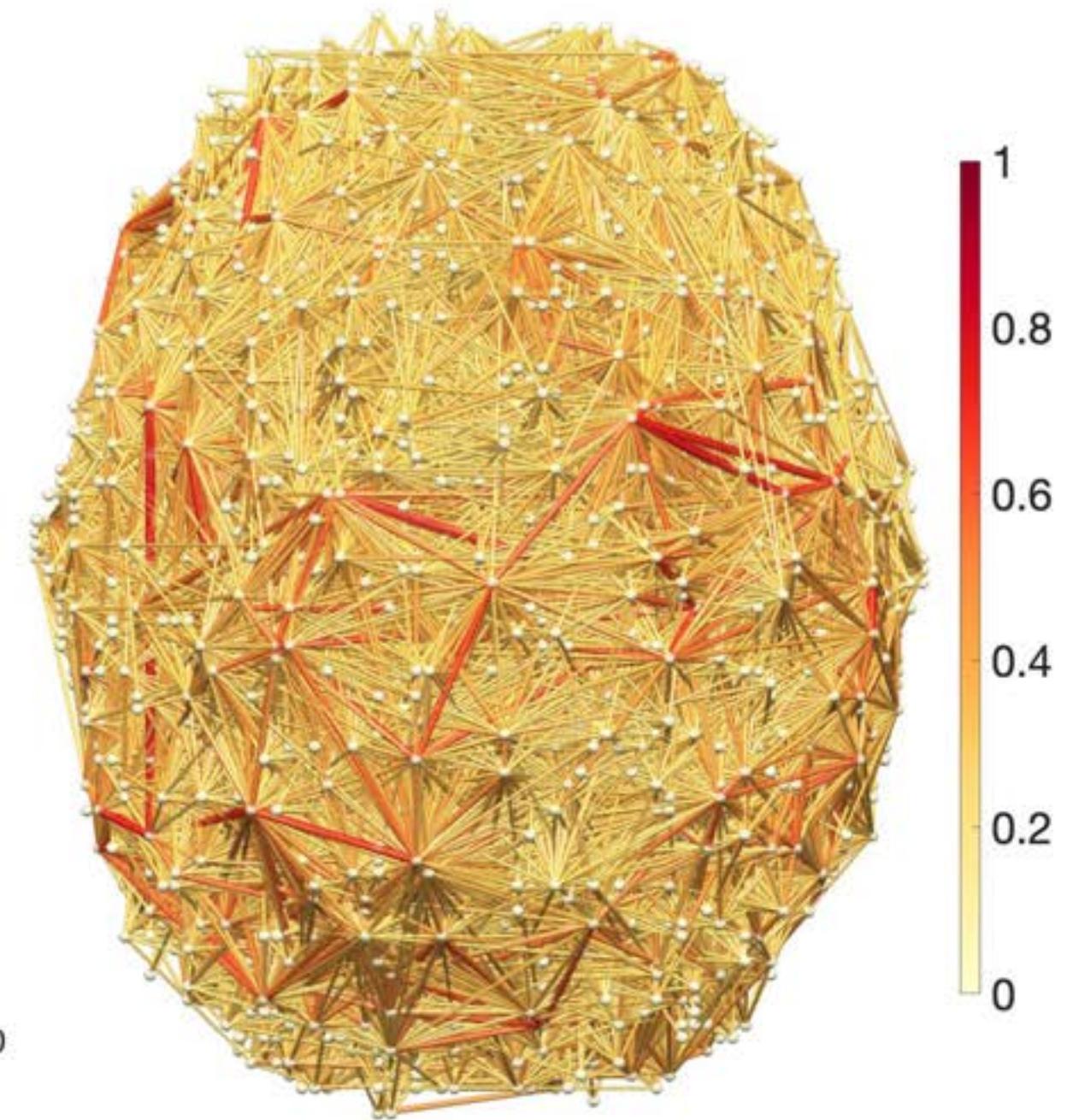
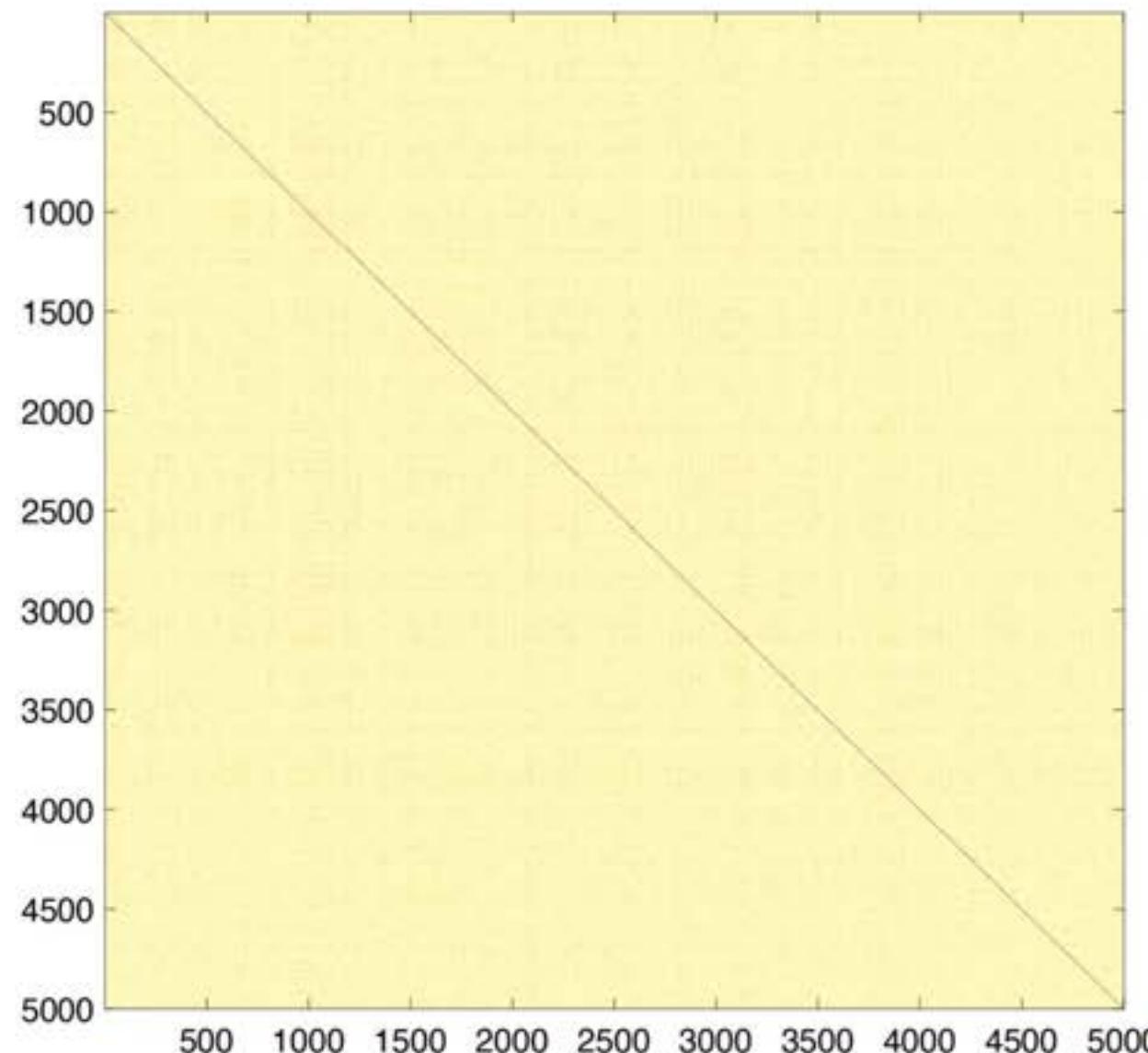
$$\|\mathbf{x}_i\|^2 = \mathbf{x}_i^\top \mathbf{x}_i = 1, \quad \sum_{i=1}^p x_{ij} = 0$$



Pearson correlation

$$\rho_{ij} = \mathbf{x}_i^\top \mathbf{x}_j$$

# Brain network with Pearson correlations



Pearson correlations

# Geometry of correlation matrix

Pearson correlation  $\rho_{ij} = \mathbf{x}_i^\top \mathbf{x}_j$

Nonnegative definite matrix  $\rho = (\rho_{ij})$

What is **dimensionality of data?**  $\mathbb{R}^{n(n-1)/2}$

Claim:  $\mathbb{R}^{n-1}$

Is it even possible?

# Measurement at node $j$

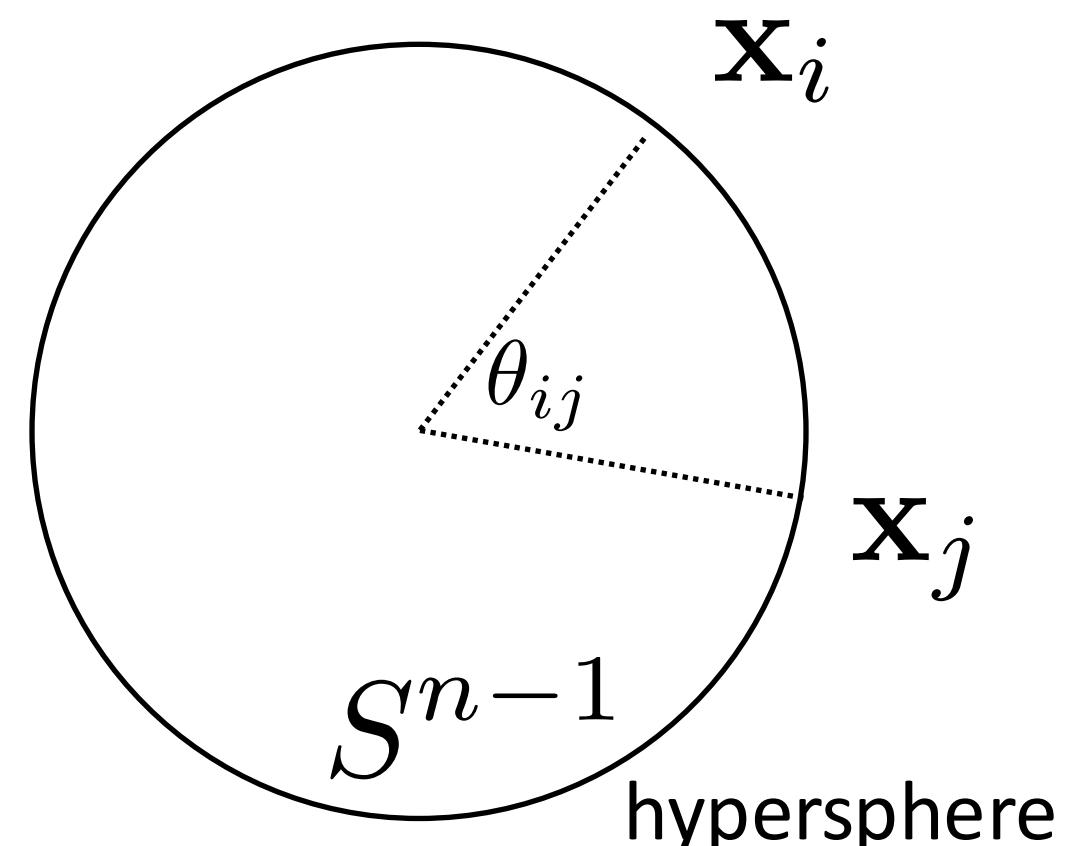
$$\mathbf{x}_j = (x_{1j}, \dots, x_{nj})^\top \in \mathbb{R}^n$$

Scale and translate data such that

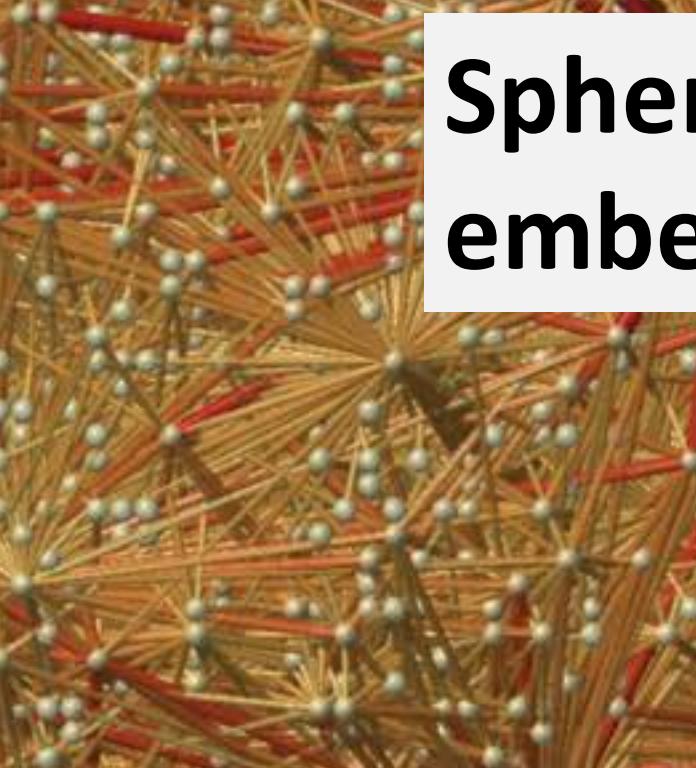
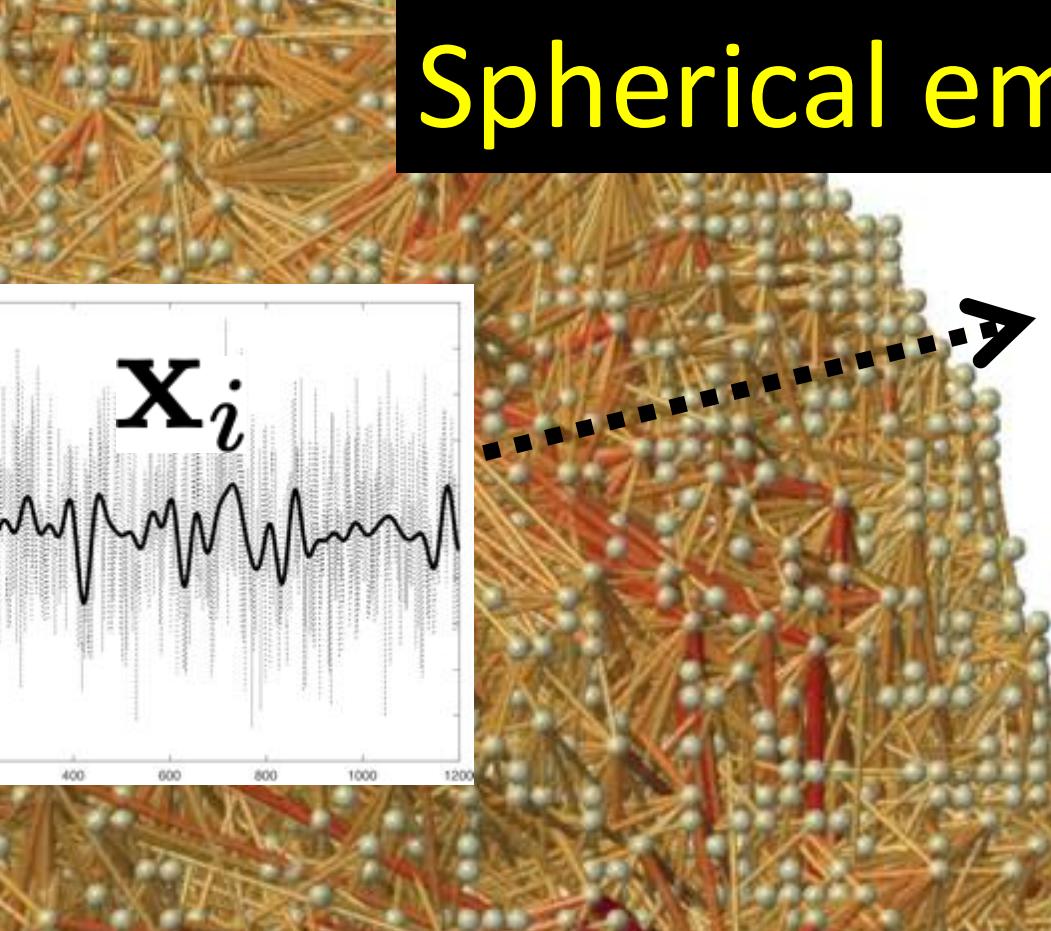
$$\mathbf{x}_j^\top \mathbf{x}_j = \sum_{i=1}^n x_{ij}^2 = 1, \quad \sum_{i=1}^n x_{ij} = 0$$

## Pearson correlation

$$\mathbf{x}_i^\top \mathbf{x}_j = \cos \theta_{ij}$$



# Spherical embedding of correlation network



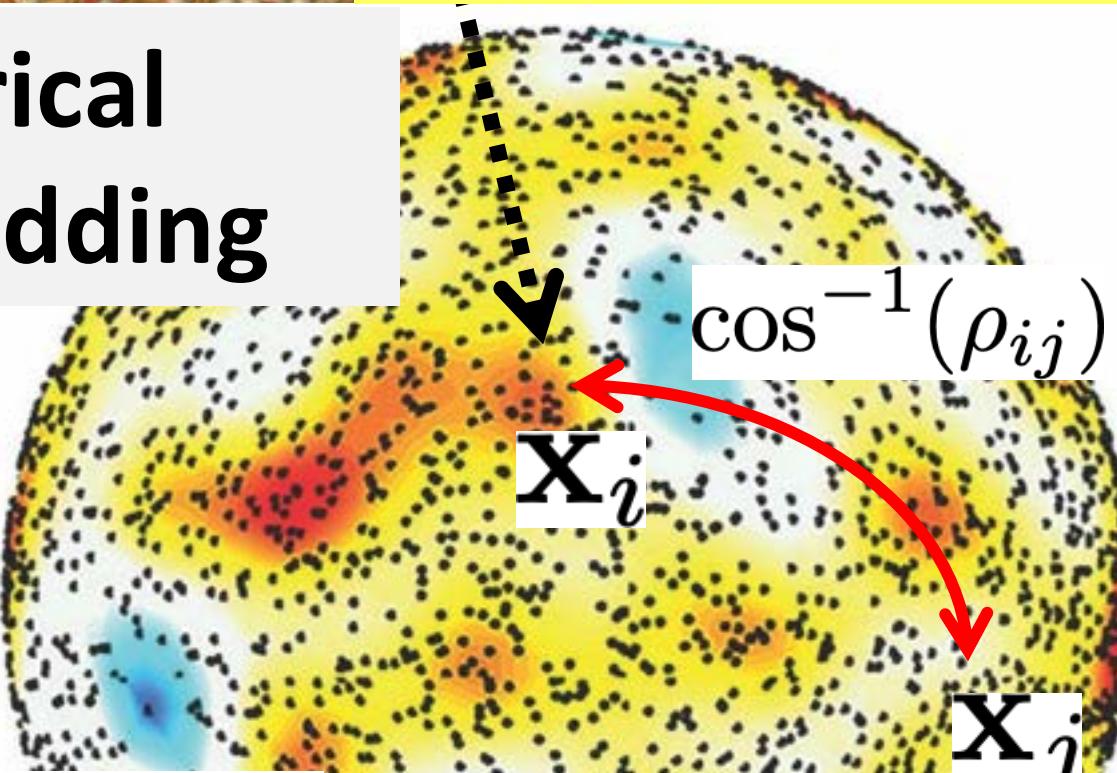
Spherical  
embedding

$n$ -dimensional  
sphere  $S^{n-1}$

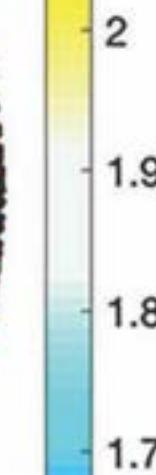
fMRI time series  $\mathbf{x}_i$  at voxel  $i$

Translate and scale:

$$\|\mathbf{x}_i\|^2 = \mathbf{x}_i^\top \mathbf{x}_i = 1, \quad \sum_{i=1}^p x_{ij} = 0$$



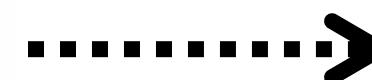
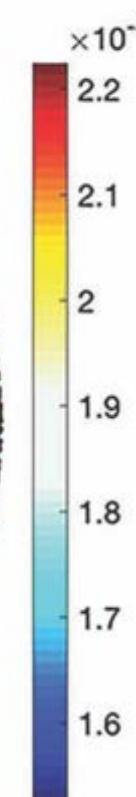
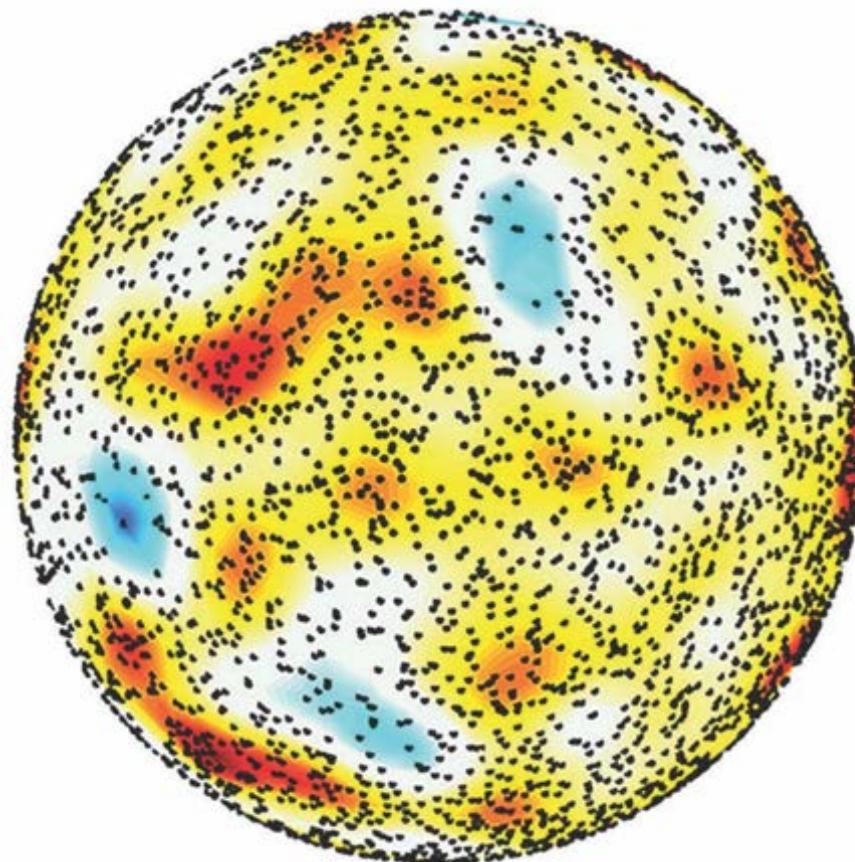
Correlation  
 $\rho_{ij} = \mathbf{x}_i^\top \mathbf{x}_j$



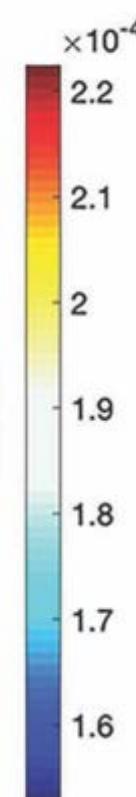
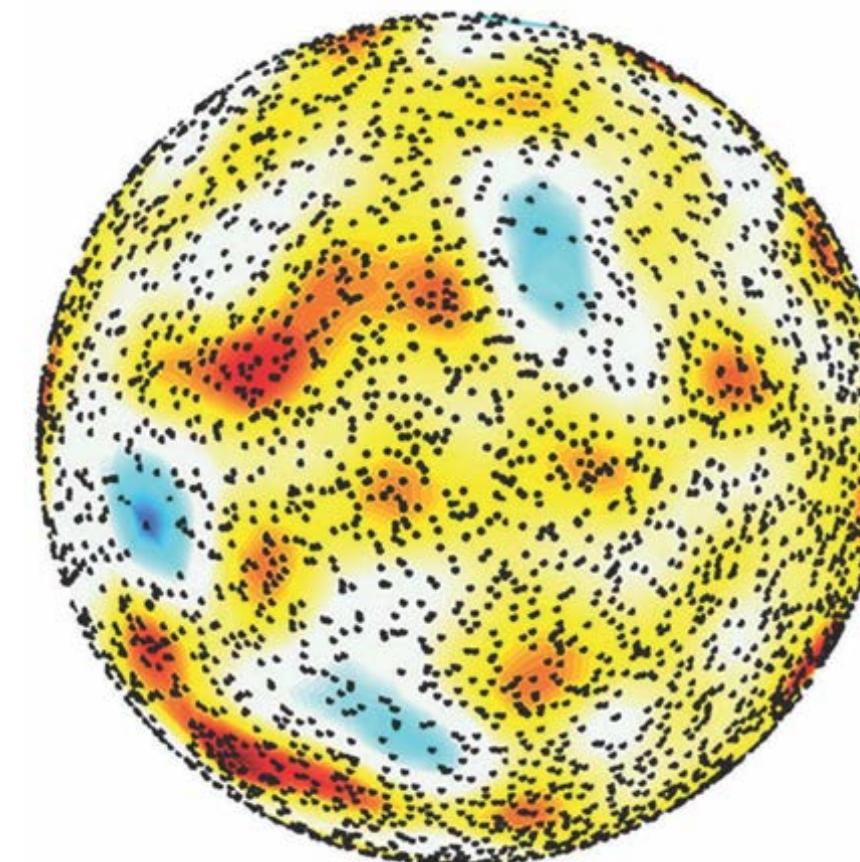
# Multidimensional scaling

*Chung and Chen 2022 arXiv: 2204.03653*

$(n-1)$ -sphere



2-sphere



$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_p]$$



$$\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_p]$$

$$\min_{\mathbf{y}_1, \dots, \mathbf{y}_p \in S^2} \sum_{i,j=1}^p \left[ \cos^{-1} \rho_{ij} - d(\mathbf{y}_i, \mathbf{y}_j) \right]^2$$



Gradient descent

$$\min_{\mathbf{Y}} \mathcal{L}(\mathbf{X}, \mathbf{Y}) = \sum_{i,j=1}^p [\cos^{-1}(\mathbf{x}_i^\top \mathbf{x}_j) - \cos^{-1}(\mathbf{y}_i^\top \mathbf{y}_j)]^2$$



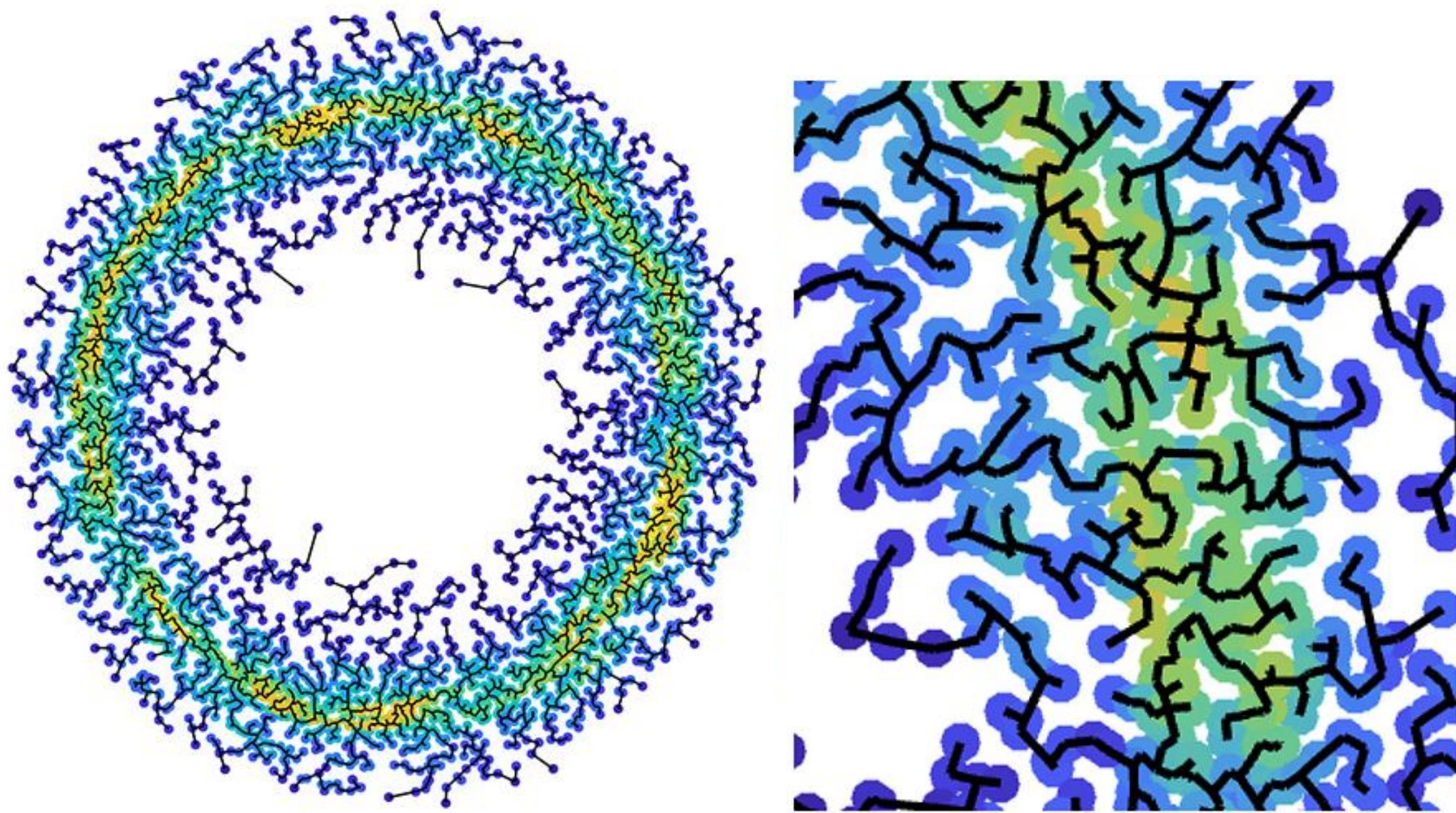
Approximately solve  
by the Taylor expansion

## Low rank approximation problem

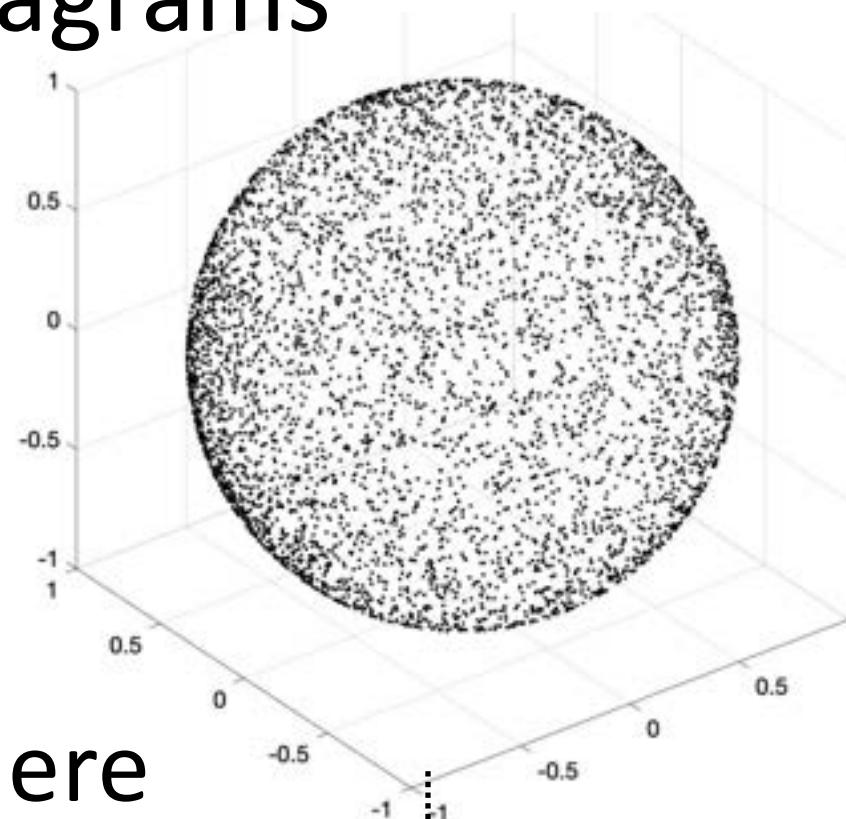
$$\mathcal{L}(\mathbf{X}, \mathbf{Y}) = \text{tr}(\mathbf{X}^\top \mathbf{X} - \mathbf{Y}^\top \mathbf{Y})^2 = \|\mathbf{X}^\top \mathbf{X} - \mathbf{Y}^\top \mathbf{Y}\|_F^2$$

rank n > rank 3

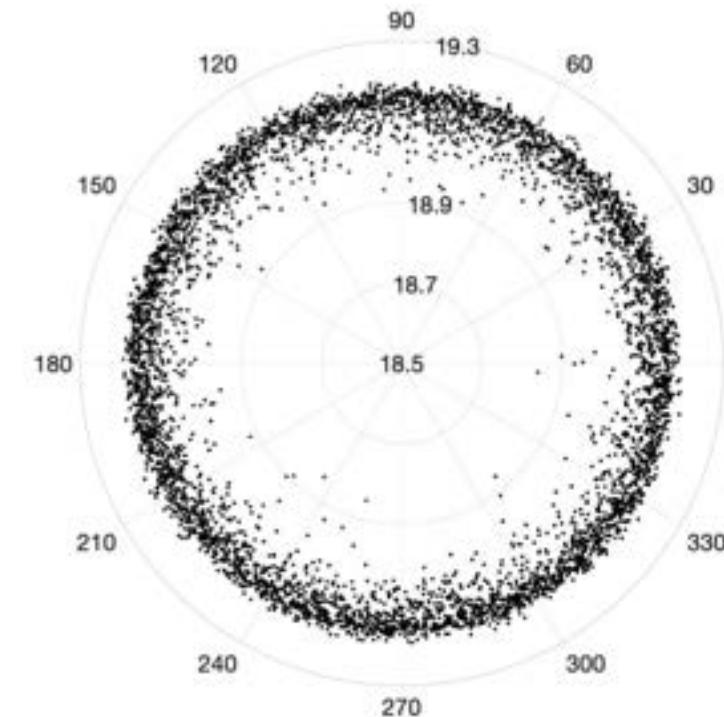
OD topology (MST) of rs-fMRI correlation  
network embedded onto Poincaré disk:  
hyperbolic space with negative curvature



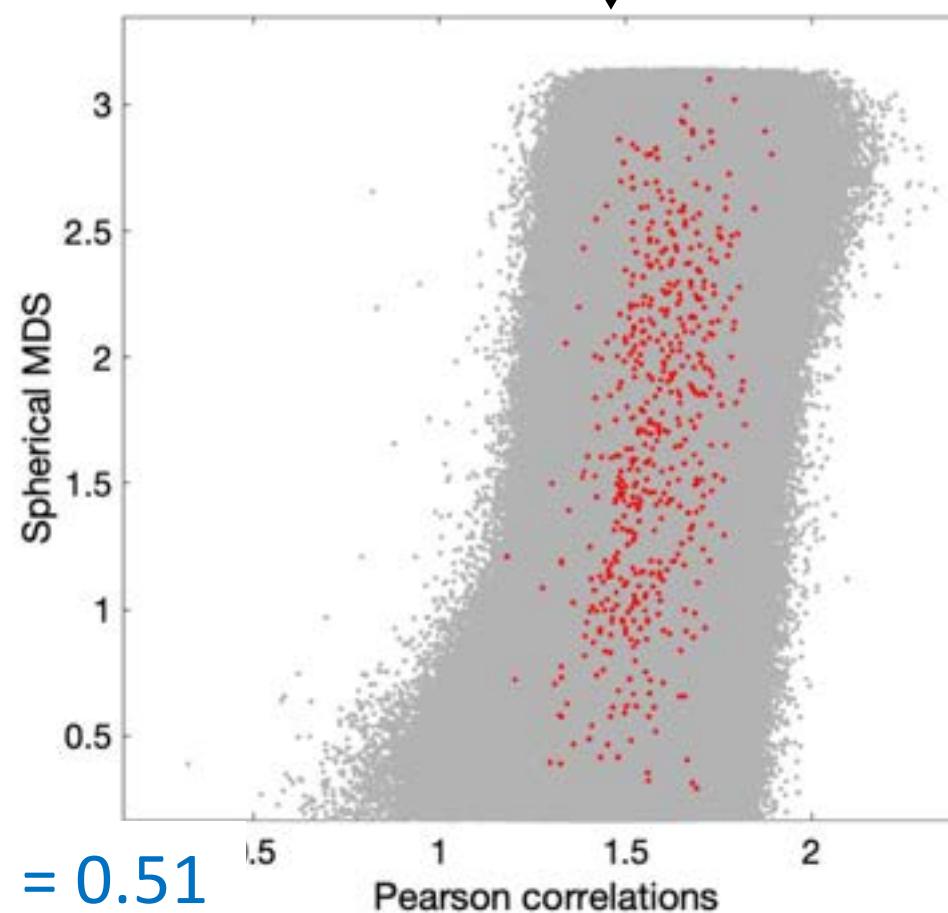
# Shepard diagrams



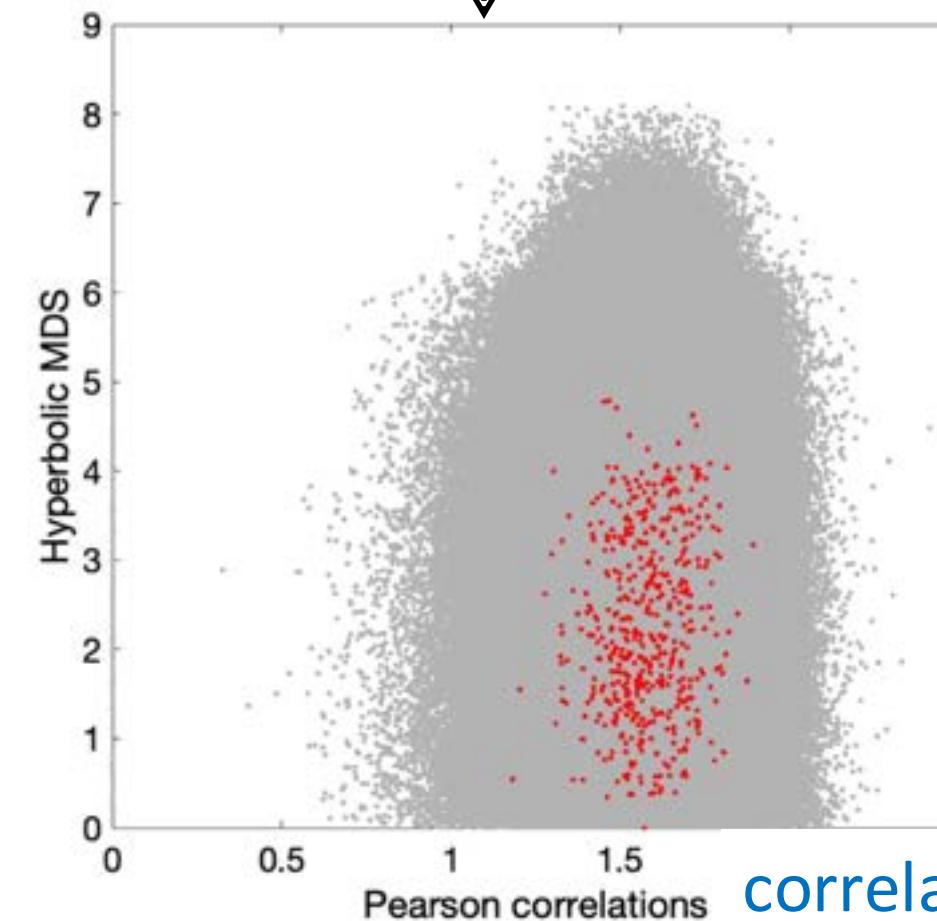
Sphere



Poincare disk



correlation = 0.51



correlation = 0.0018

ARTICLE

<https://doi.org/10.1038/s41467-019-12915-x>

OPEN

# Network curvature as a hallmark of brain structural connectivity

Hamza Farooq<sup>1\*</sup>, Yongxin Chen<sup>2</sup>, Tryphon T. Georgiou<sup>3</sup>, Allen Tannenbaum<sup>4</sup> & Christophe Lenglet<sup>5</sup>

Although brain functionality is often remarkably robust to lesions and other insults, it may be fragile when these take place in specific locations. Previous attempts to quantify robustness and fragility sought to understand how the functional connectivity of brain networks is affected by structural changes, using either model-based predictions or empirical studies of the effects of lesions. We advance a geometric viewpoint relying on a notion of network curvature, the so-called Ollivier-Ricci curvature. This approach has been proposed to assess financial market robustness and to differentiate biological networks of cancer cells from healthy ones. Here, we apply curvature-based measures to brain structural networks to identify robust and fragile brain regions in healthy subjects. We show that curvature can also be used to track changes in brain connectivity related to age and autism spectrum disorder (ASD), and we obtain results that are in agreement with previous MRI studies.

Several other directions may be worthy of investigation along the same lines. In particular, studying curvature changes between nodes at the edge level may prove particularly effective as, in that case, critical changes in interactions between areas in the brain may be easier to detect. We propose these future directions with the caveat that curvature is sensitive to the way connectivity matrices are generated, i.e., curvature is affected by the choice of parcellation scale, tractography algorithms, as well as the particular type of diffusion data, e.g., DTI, HARDI, DSI, etc. Therefore, care must be exercised to minimize such possible effects. The present work focused mainly on exploring the concept of node curvature as a measure of robustness of brain structural networks, in comparison with existing measures.

## Methods

**Overview.** First, we describe generic notions of distance and curvature on metric spaces (i.e., Riemannian manifolds). These concepts are needed to understand how brain networks (e.g., graphs) curvature and robustness can be characterized. Next, we describe how curvature can be defined and computed on discrete spaces, such as brain networks with finite (and usually low) number of nodes. Finally, we relate curvature to robustness, and explain how it can be efficiently computed and used to assess the ability of a graph to withstand perturbations.

**Wasserstein distance and optimal mass transport.** Let  $p$  and  $q$  be two probability distributions on the discrete metric space  $\mathcal{X}$  equipped with metric  $d(\cdot, \cdot)$ . The transportation cost of a unit mass from point  $x_i \in \mathcal{X}$  to  $x_j \in \mathcal{X}$  is denoted as  $c_{i,j} \geq 0$ . Denote by  $\pi_{i,j} \geq 0$  the transference plan, i.e., the (probability) measure of the amount of mass transferred from  $x_i$  to  $x_j$ .

The optimal mass transportation (OMT) problem seeks an optimal transference plan ( $\pi$ ) that minimizes the total cost of moving  $p$  to  $q$ . This can be formulated as the following optimization problem<sup>55–57</sup>:

$$\min_{\pi} \sum_{i,j} c_{i,j} \pi_{i,j} \quad (1)$$

$$\begin{aligned} \text{subject to } & \sum_j \pi_{i,j} = p_i, \quad \forall i \\ & \sum_i \pi_{i,j} = q_j, \quad \forall j \\ & \pi_{i,j} \geq 0, \quad \forall i, j \end{aligned}$$

When the cost  $c$  is defined as  $c_{i,j} = d(x_i, y_j)^r$ , for any positive integer  $r$ , we can define the  $r$ -Wasserstein distance<sup>56,58</sup> as

$$W_r(p, q) := \left( \min_{\Pi \in \mathcal{M}(p,q)} \text{trace}(\Pi^T \Pi) \right)^{1/r}. \quad (3)$$

When  $r = 1$  this is also known as the earth mover's distance (EMD). We will use this version of OMT in the present work.

**Generalities on curvature.** In this section, we introduce the key notion of curvature from Riemannian geometry. For  $X$  an  $n$ -dimensional Riemannian manifold,  $x \in X$ , let  $T_x$  denote the tangent space at  $x$ , and  $v_1, v_2 \in T_x$  orthonormal tangent vectors. Then, for geodesics (local curves of shortest length)  $\alpha_i(t) := \exp(tv_i)$ ,  $i = 1, 2$ , the sectional curvature  $K(v_1, v_2)$  measures the deviation of geodesics relative to Euclidean geometry, i.e.,

$$d(\alpha_1(t), \alpha_2(t)) = \sqrt{2t} \left( 1 - \frac{K(v_1, v_2)}{12} t^2 + O(t^4) \right). \quad (4)$$

The Ricci curvature is the average sectional curvature. In other words, given a (unit) vector  $v \in T_x$ , we complete an orthonormal basis  $v_1, v_2, \dots, v_n$ , and define  $\text{Ric}(v) := \frac{1}{n-1} \sum_{i=2}^n K(v, v_i)$ . The Ricci curvature may be extended to a quadratic form known as the Ricci curvature tensor<sup>59</sup>. The scalar curvature is then defined to be the trace of this quadratic form.

There are a number of alternative characterizations of Ricci curvature<sup>59</sup>. In this paper, we employ the following definition: referring to Fig. 6, consider two very close points  $x$  and  $y$  in  $X$  and associated tangent vectors  $w$  and  $w'$ , where  $w'$  is obtained by parallel transport of  $w$  along a geodesic (in the direction  $v$ ) connecting the two points. Now, geodesics drawn from  $x, y$  along  $w, w'$  will get closer when the curvature is positive (positively curved space). This is also reflected in the fact that the distance between two small (geodesic balls) is less than the distance of their centers. The Ricci curvature  $\text{Ric}(v)$  along the direction  $v$  connecting  $x, y$  quantifies this contraction. Similar considerations apply to negative and zero curvature.

**Curvature and entropy.** In this section, following the previous studies<sup>16,17</sup>, we establish the relationship between curvature and robustness. We consider  $X$  to be a complete, smooth Riemannian manifold. The relations may then be extrapolated to discrete spaces.

We denote the space of probability densities by  $\mathcal{P}(X)$ . Then, one defines the Boltzmann entropy of  $\rho \in \mathcal{P}(X)$  as

$$S(\rho) := - \int_X \rho \log \rho \, dm, \quad (5)$$



# Situating the default-mode network along a principal gradient of macroscale cortical organization

Daniel S. Margulies<sup>a,1</sup>, Satrajit S. Ghosh<sup>b,c</sup>, Alexandros Goulas<sup>d</sup>, Marcel Falkiewicz<sup>a</sup>, Julia M. Huntenburg<sup>a,e</sup>, Georg Langs<sup>f,g</sup>, Gleb Bezgin<sup>h</sup>, Simon B. Eickhoff<sup>i,j</sup>, F. Xavier Castellanos<sup>k,l</sup>, Michael Petrides<sup>m</sup>, Elizabeth Jefferies<sup>n,o</sup>, and Jonathan Smallwood<sup>n,o</sup>

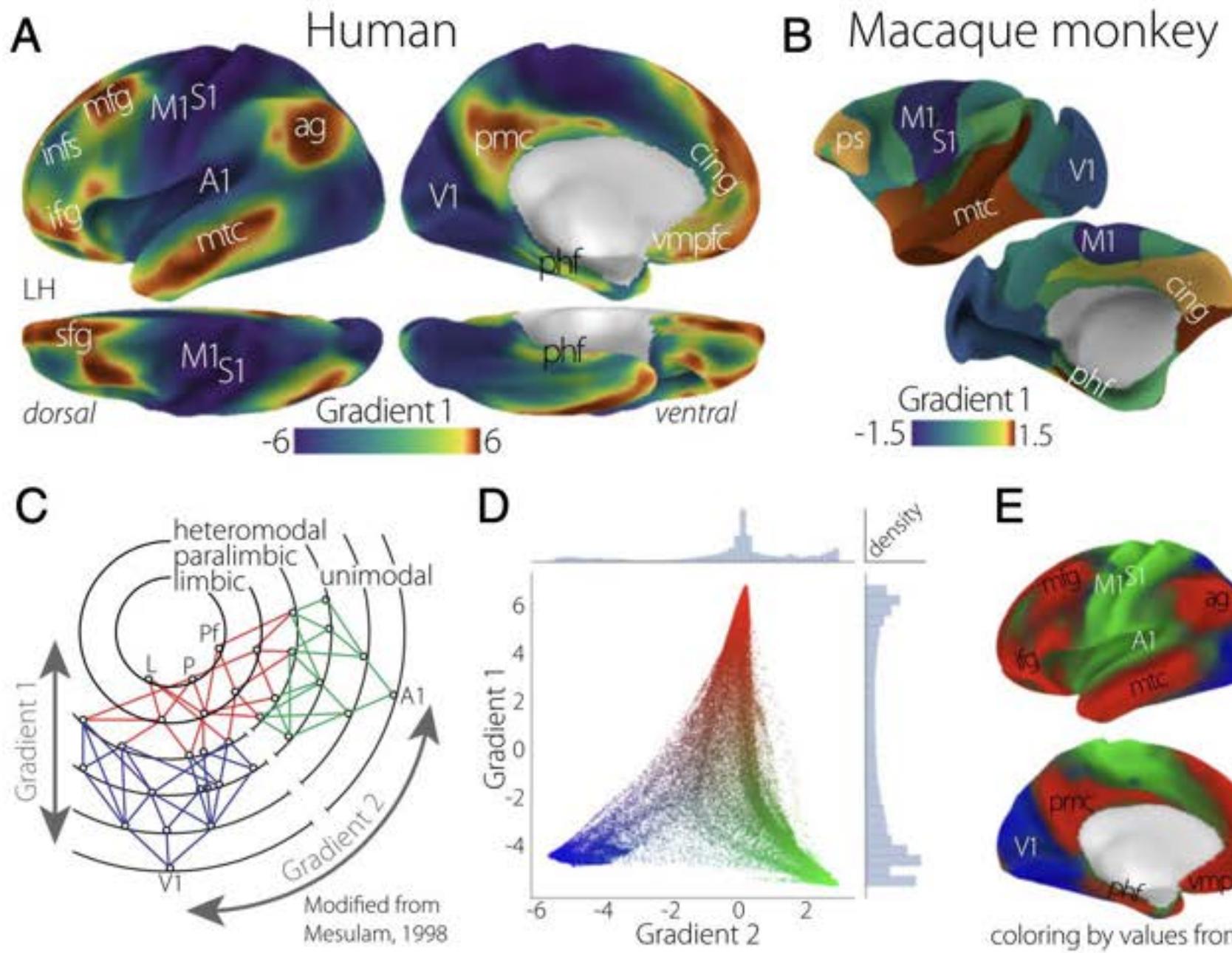
<sup>a</sup>Max Planck Research Group for Neuroanatomy & Connectivity, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany;

<sup>b</sup>McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139; <sup>c</sup>Department of Otolaryngology, Harvard Medical School, Cambridge, MA 02115; <sup>d</sup>Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany;

<sup>e</sup>Neurocomputation and Neuroimaging Unit, Department of Education and Psychology, Free University of Berlin, Berlin 14195, Germany; <sup>f</sup>Department of Biomedical Imaging and Image-Guided Therapy, Computational Imaging Research Laboratory, Medical University of Vienna, Vienna A-1090, Austria;

<sup>g</sup>Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; <sup>h</sup>McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada H3A 2B4; <sup>i</sup>Institute for Neuroscience and Medicine, Research Center Jülich, Juelich 52428, Germany; <sup>j</sup>Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University, Duesseldorf 40225, Germany; <sup>k</sup>Child Study Center, Department of Child and Adolescent Psychiatry, New York University Langone Medical Center, New York, NY 10016; <sup>l</sup>Nathan Kline Institute for Psychiatric Research, Orangeburg, NY 10962; <sup>m</sup>Cognitive Neuroscience Unit, Montreal Neurological Institute, McGill University, Montreal, QC, Canada H3A 2B4; <sup>n</sup>Department of Psychology, University of York, York YO10 5DD, United Kingdom; and <sup>o</sup>York Neuroimaging Centre, University of York, York YO10 5DD, United Kingdom

Edited by Peter L. Strick, University of Pittsburgh, Pittsburgh, PA, and approved September 9, 2016 (received for review May 27, 2016)



**Fig. 1.** The principal gradient of connectivity in both the (A) human and (B) macaque monkey cortices shows a spectrum between unimodal regions (dark blue) and transmodal regions (sienna), which in the human cortex, peaks in regions corresponding to the DMN. The proximity of colors can be interpreted as greater similarity of connectivity patterns. (C) The illustration of connectivity organization suggested by Mesulam (23) proposes a hierarchy of processing from distinct unimodal areas to integrative transmodal areas. Labels Gradient 1 and Gradient 2, which were not included in the original figure, correspond to the results in D. Modified from ref. 23. (D) A scatter plot of the first two connectivity embedding gradients. Gradient 1 extends between primary sensorimotor and transmodal regions (red). Gradient 2 separates somatomotor and auditory cortex (green) from visual cortex (blue). Histograms depicting the distribution of values are presented on the respective axes. (E) Colors from the scatter plot are presented on the cortical surface for anatomical orientation. A1, primary auditory; ag, angular gyrus; cing, anterior cingulate cortex; ifg, inferior frontal gyrus; ifns, intermediate frontal sulcus; L, limbic; M1, primary motor; mfg, middle frontal gyrus; mtc, middle temporal cortex; P, parietal; Pf, prefrontal; phf, para-hippocampal formation; pmc, posteromedial cortex; ps, principal sulcus; S1, primary somatosensory; sfg, superior frontal gyrus; V1, primary visual; vmpfc, ventromedial prefrontal cortex.



# Gradients of structure–function tethering across neocortex

Bertha Vázquez-Rodríguez<sup>a</sup>, Laura E. Suárez<sup>a</sup>, Ross D. Markello<sup>a</sup>, Golia Shafiei<sup>a</sup>, Casey Paquola<sup>a</sup>, Patric Hagmann<sup>b</sup>, Martijn P. van den Heuvel<sup>c</sup>, Boris C. Bernhardt<sup>a</sup>, R. Nathan Spreng<sup>a</sup>, and Bratislav Misic<sup>a,1</sup>

<sup>a</sup>McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, QC H3A 2B4, Canada; <sup>b</sup>Signal Processing Laboratory (LTSS), Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; and <sup>c</sup>Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, 1081 HV, Amsterdam, The Netherlands

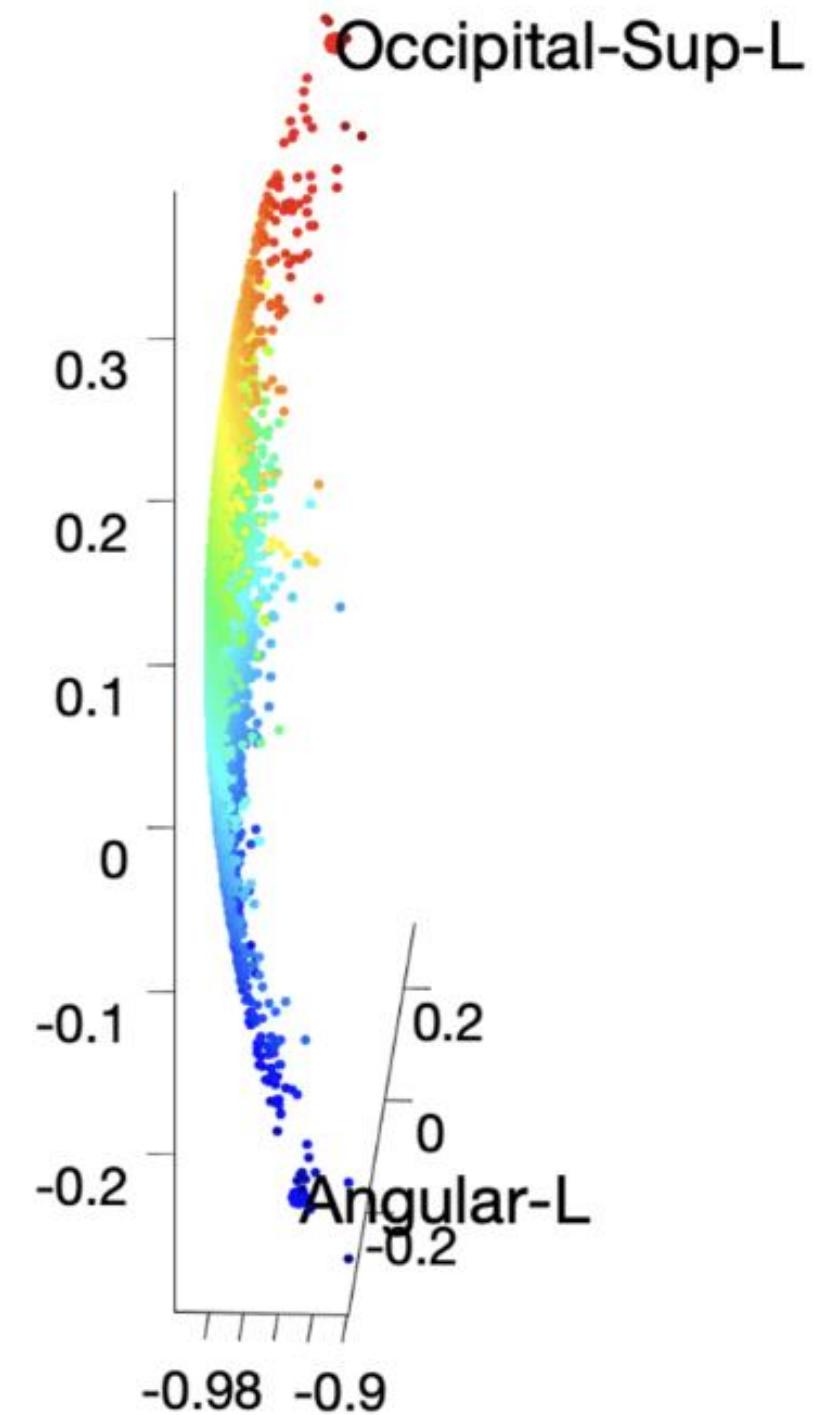
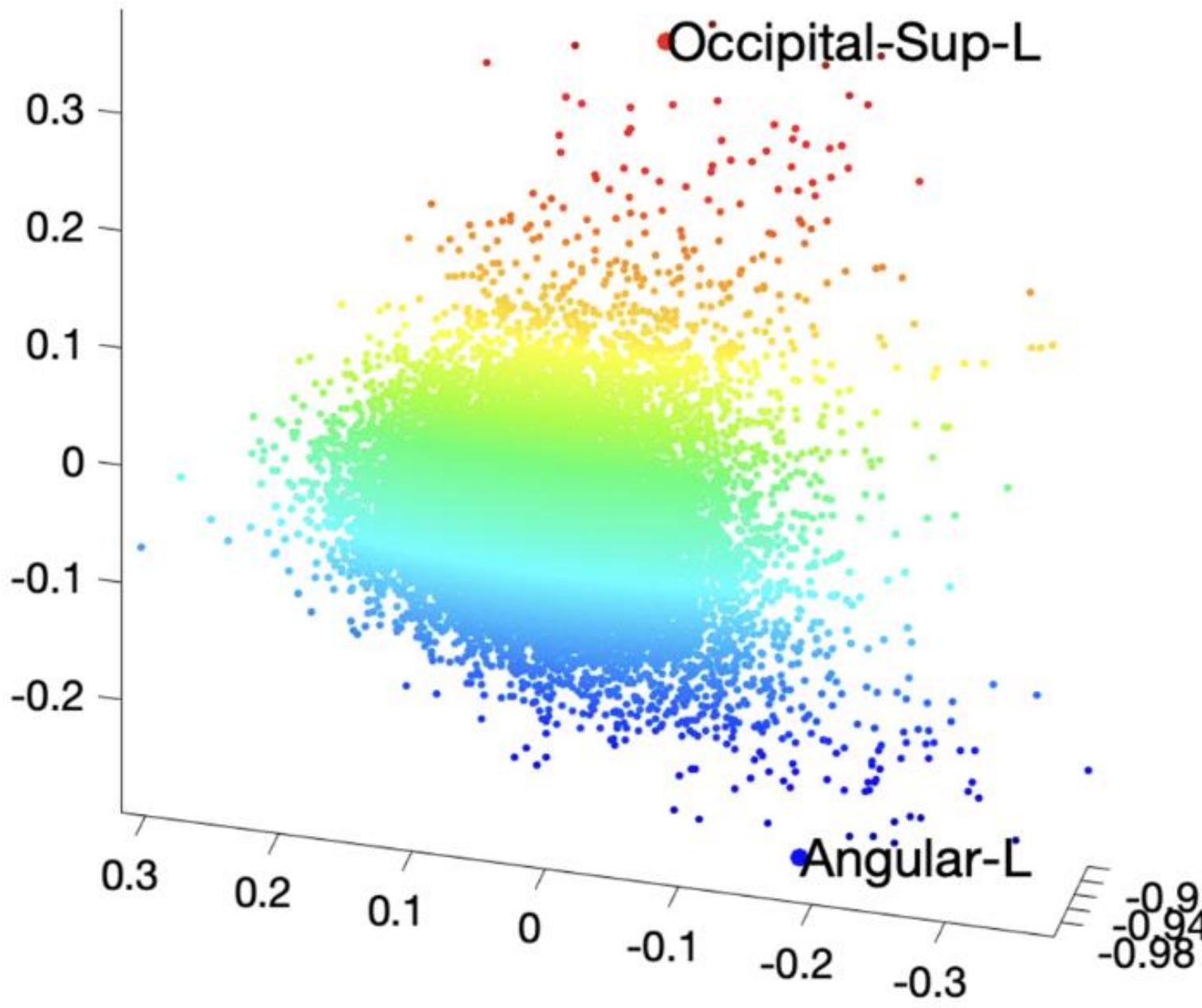
Edited by Marcus E. Raichle, Washington University in St. Louis, St. Louis, MO, and approved September 10, 2019 (received for review February 26, 2019)

The white matter architecture of the brain imparts a distinct signature on neuronal coactivation patterns. Interregional projections promote synchrony among distant neuronal populations, giving rise to richly patterned functional networks. A variety of statistical, communication, and biophysical models have been proposed to study the relationship between brain structure and function, but the link is not yet known. In the present report we seek to relate the structural and functional connection profiles of individual brain areas. We apply a simple multilinear model that incorporates information about spatial proximity, routing, and diffusion between brain regions to predict their functional connectivity. We find that structure–function relationships vary markedly across the neocortex. Structure and function correspond closely in unimodal, primary sensory, and motor regions, but diverge in transmodal cortex, particularly the default mode and salience networks. The divergence between structure and function systematically follows functional and cytoarchitectonic hierarchies. Altogether, the present results demonstrate that structural and functional networks do not align uniformly across the brain, but gradually uncouple in higher-order polysensory areas.

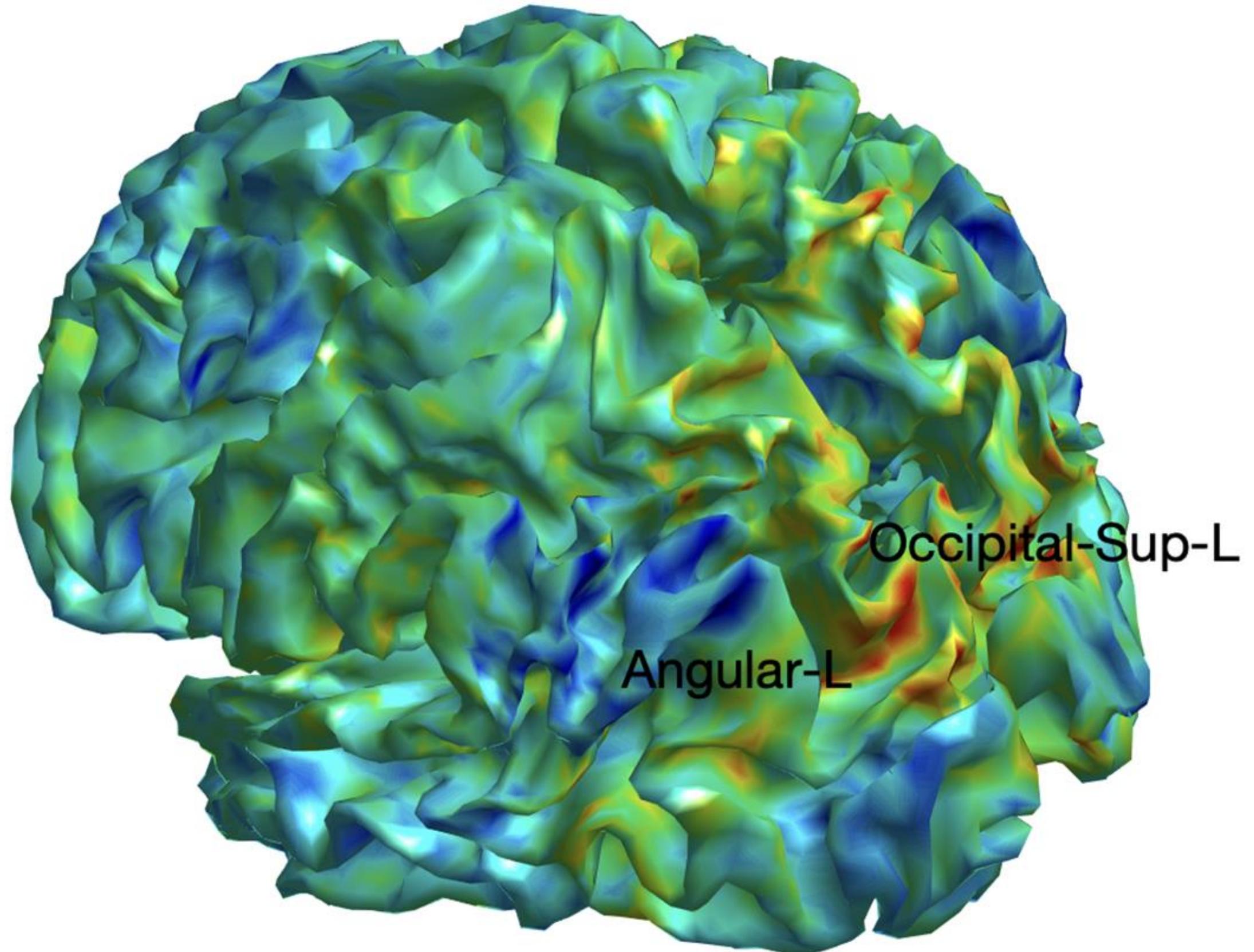
ciations with intracortical myelination (21) and laminar differentiation (22). A prominent account posits that rapid evolutionary expansion of association cortices effectively “untethers” polysensory regions from molecular signaling gradients and canonical sensory-motor activity cascades, resulting in fundamentally different structure–function relationships along the unimodal–transmodal hierarchy (23). Altogether, this work opens the possibility that structure and function may not be related in exactly the same way across the whole brain, but potentially converge or diverge in specific areas.

Here we address the relationship between structure and function by focusing on connection profiles of individual brain regions. We first reconstruct structural and functional networks from diffusion MRI (dMRI) and resting-state functional MRI (fMRI) in a cohort of 40 healthy participants. We then apply a simple multilinear model that uses information about a region’s geometric and structural network embedding to predict its functional network embedding. The method allows us to ask how closely structure and function correspond in individual regions and the extent to which this correspondence reflects affiliation with cognitive systems, cytoarchitecture, and functional hierarchies.

# Largest network gradient from spherical-MDS



# Largest network gradient from spherical-MDS

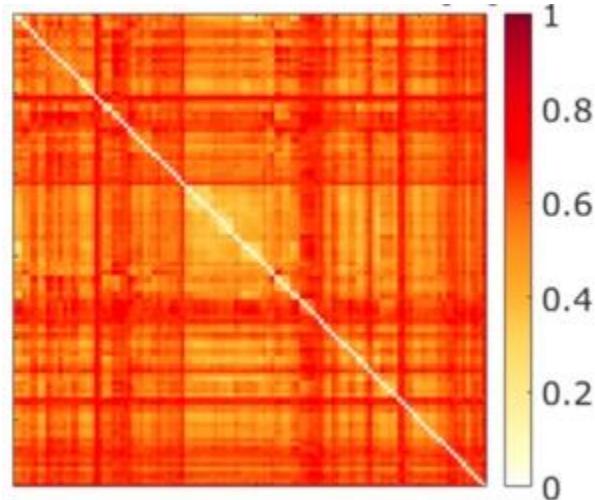


# Distance Regression

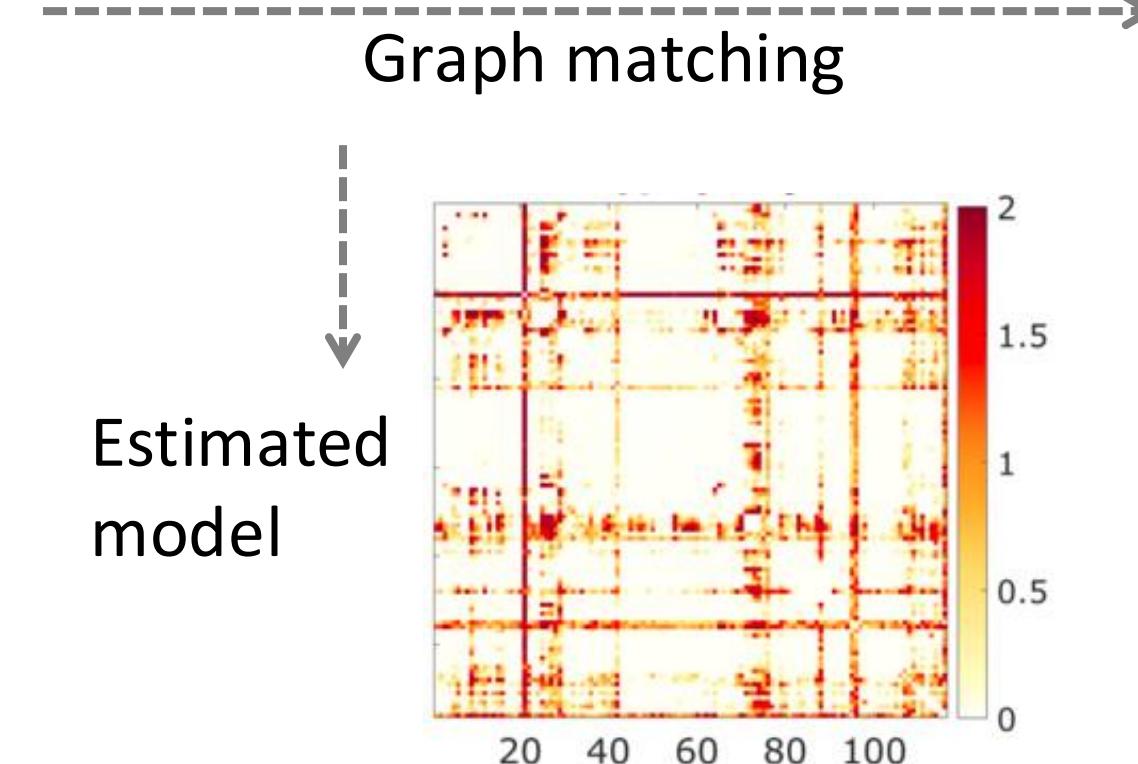
# Subject-level learning

Functional network of subject  $k$

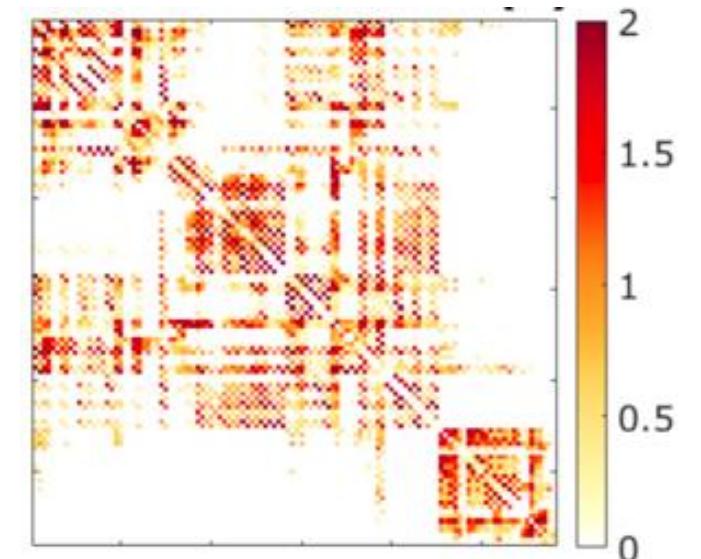
$$G_k = (V, w^k)$$



*Dense cycles*



Estimated model



*Sparse trees*

$$\widehat{\Theta}_k = \arg \min_{\Theta} \mathcal{L}_F(\Theta, G_k) + \lambda_k \mathcal{L}(\Theta, P)$$



Frobenius norm  
Goodness-of-fit



Control amount  
of topology



Topological loss

# Topological gradient descent

$$\hat{\Theta}_k = \arg \min_{\Theta} \mathcal{L}_F(\Theta, G_k) + \lambda_k \mathcal{L}_{top}(\Theta, P)$$

Topological gradient by matching sorted birth and death values

$$\frac{\partial \mathcal{L}_{top}(\Theta, P)}{\partial w_{ij}^{\Theta}} = \begin{cases} 2[w_{ij}^{\Theta} - \tau_{0*}(w_{ij}^{\Theta})] & \text{if } w_{ij}^{\Theta} \in E_0; \\ 2[w_{ij}^{\Theta} - \tau_{1*}(w_{ij}^{\Theta})] & \text{if } w_{ij}^{\Theta} \in E_1 \end{cases}$$

Run time  $O(|E| \log |V|)$