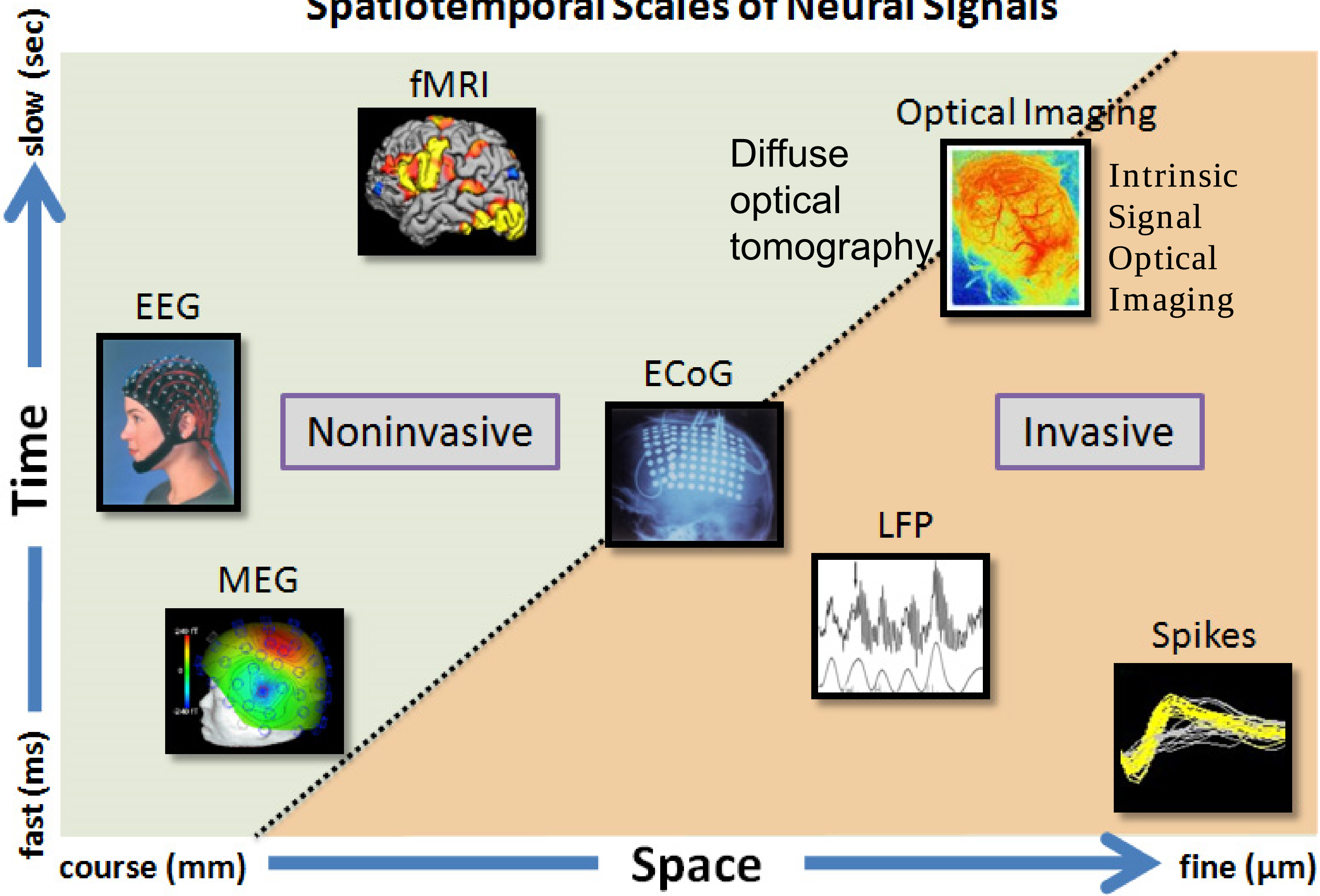
25th Symmer School "Dynamical Systems and Complexity"

RECONSTRUCTION OF FUNCTIONAL CONNECTIVITY NETWORKS FROM FMRI DATA USING MANIFOLD LEARNING ALGORITHMS

IOANNIS GALLOS, CONSTANTINOS SIETTOS

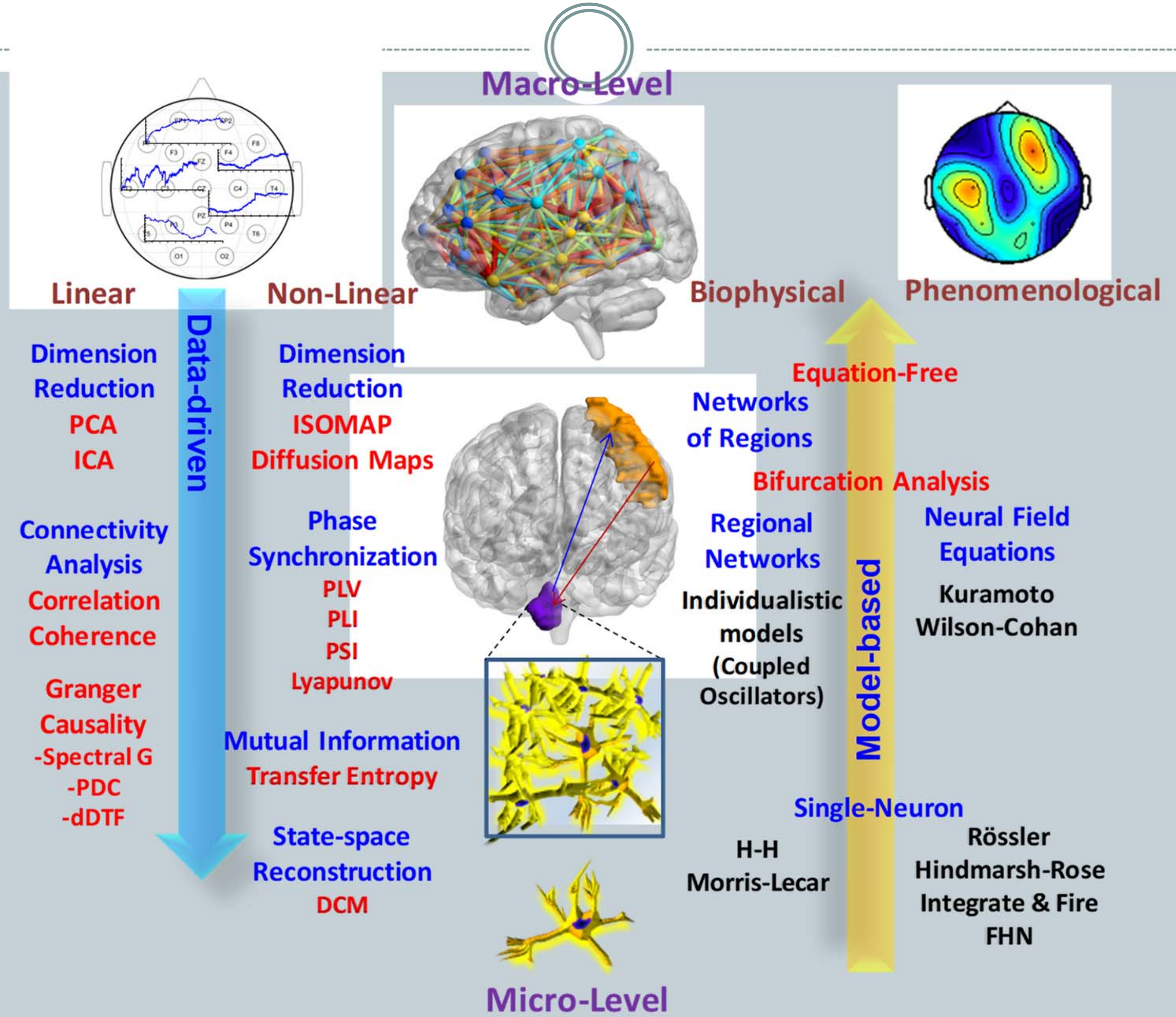
SCHOOL OF APPLIED MATHEMATICS AND PHYSICAL SCIENCES
NATIONAL TECHNICAL UNIVERSITY OF ATHENS, GREECE

Spatiotemporal Scales of Neural Signals



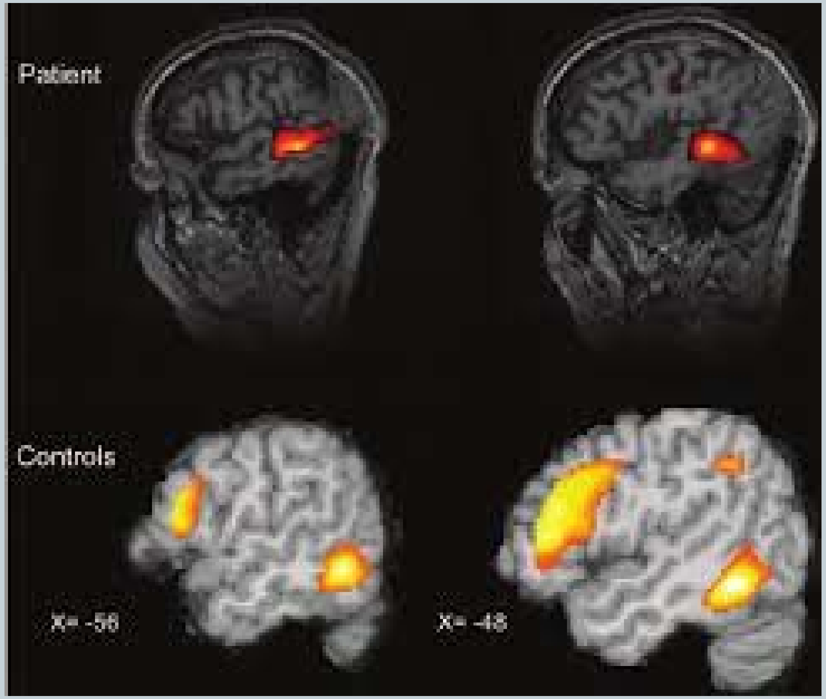
Kim et al. 2015,*BioMedical Engineering OnLine*

Modeling Brain Dynamics- Top Down and Bottom Up



Siettos, C.I., Starke, J., 2016, Multiscale Modeling of Brain Dynamics: from Single Neurons and Networks to Mathematical Tools, WIREs Systems Biology and Medicine, ***8****(5), 438-458.*

FMRI NeuroImaging and Data Mining

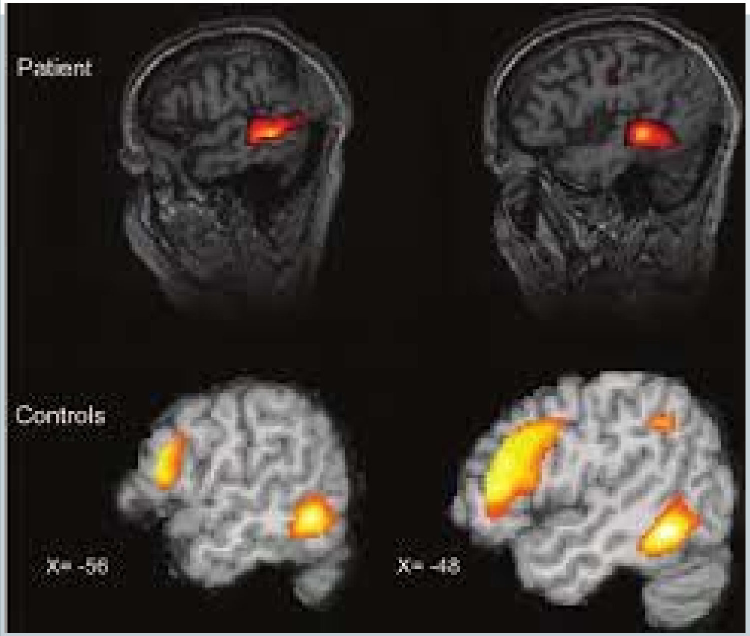


• The primary form of fMRI uses the **blood-oxygen-level dependent (BOLD) contrast**.

•This is a type of specialized brain and body scan used to map neural activity in the brain or spinal cord of humans or other animals by **imaging the change in blood flow (hemodynamic response) related to the energy used by brain cells**.

• fMRI provide signals that are contained in a high 4D dimensional space.

FMRI NeuroImaging and Data Mining



- In a typical 4D functional connectivity fMRI scan more than 100,000 voxels (which contain the activity information of millions of neurons) are recorded at each time instance.

- thus in the four dimensions of space and time, one gets **several millions of data measurements**.

Hence, one is confronted with two challenges:

A. that of dimensionality reduction and the problem of data mining, i.e. extracting features (biomarkers) that pertain to the investigated brain function within each subject and

B. that of classifying differences among subjects or groups of subjects (e.g Controls vs. Patients)

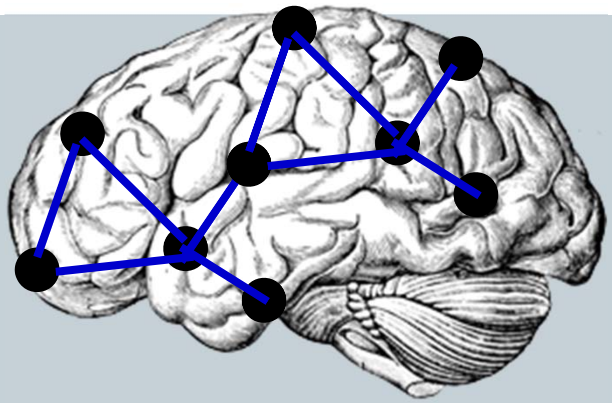
MAIN GOAL

- Our main goal is to find differences between Controls and Patients of Neurological Disorders (such as Epilepsy, Schizophrenia) on

- (functional**)** **connectivity** of brain dynamics from FMRI data

- **embedded on a low-dimensional space.**

Brain Connectivity



Functional Connectivity:
estimates whether there is any functional connection between functional regions, even if indirect. (Smith et al., 2013)
It is defined as statistical dependencies among remote neurophysiological events.

Effective Connectivity:
refers explicitly to the influence that one neural system exerts over another, either at a synaptic or population level. (Friston, 2011)

Causal Connectivity:

estimates whether the response of a region causes that of another.

Granger causality

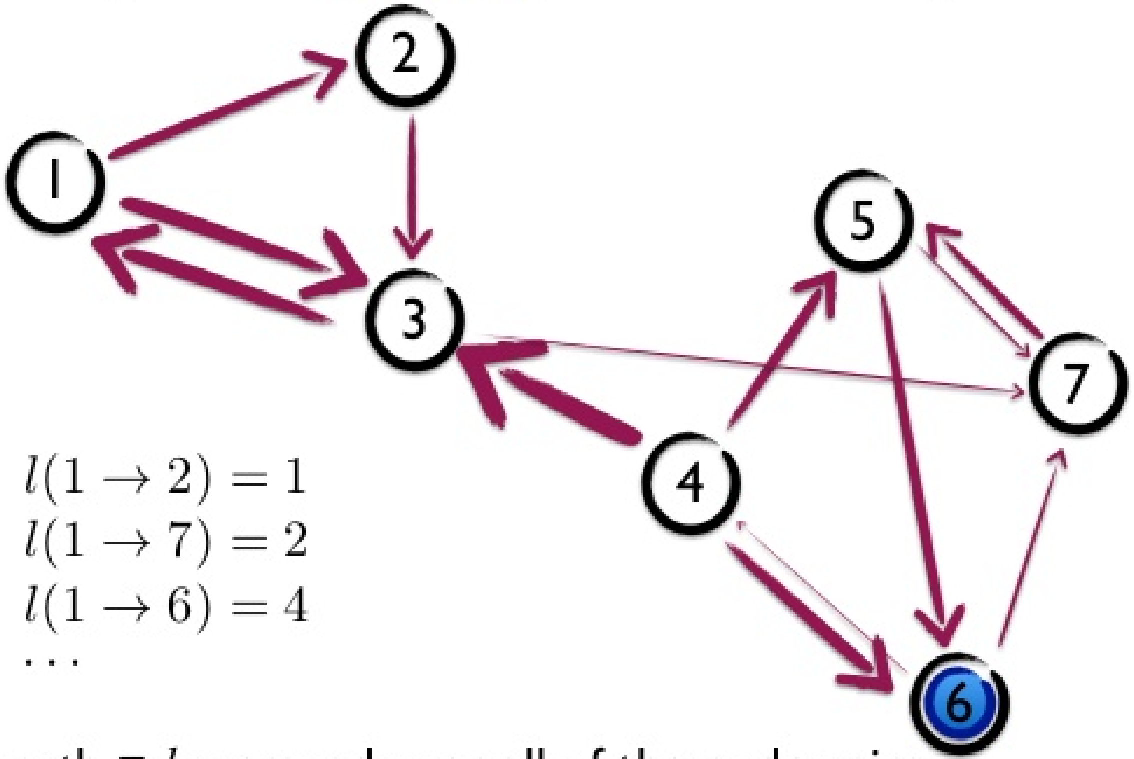
Brain Connectivity

The average inverse shortest path length is a related measure known as the **global efficiency**

average path length of a network

path length l between i and j:

the number of edges in the **shortest** path between i and j

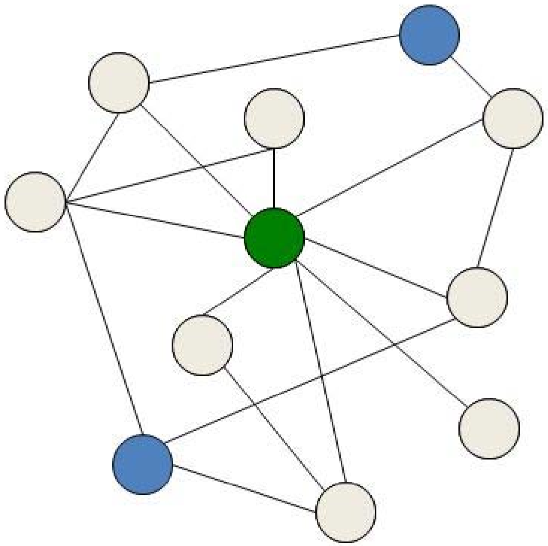


$$l(1 \to 2) = 1$$
$$l(1 \to 7) = 2$$
$$l(1 \to 6) = 4$$
$$\dots$$

average path length = l averaged over all of the node pairs

real networks: much smaller average path length than regular networks!

Clustering coefficient



The density of the network surrounding node I, characterized as the number of triangles through I. Related to network modularity

$$C_I = \frac{n_I}{\binom{k}{2}} = \frac{2n_I}{k \cdot (k-1)}$$

k: neighbors of *I*

n_I: edges between node *I*'s neighbors

The center node has 8 (grey) neighbors

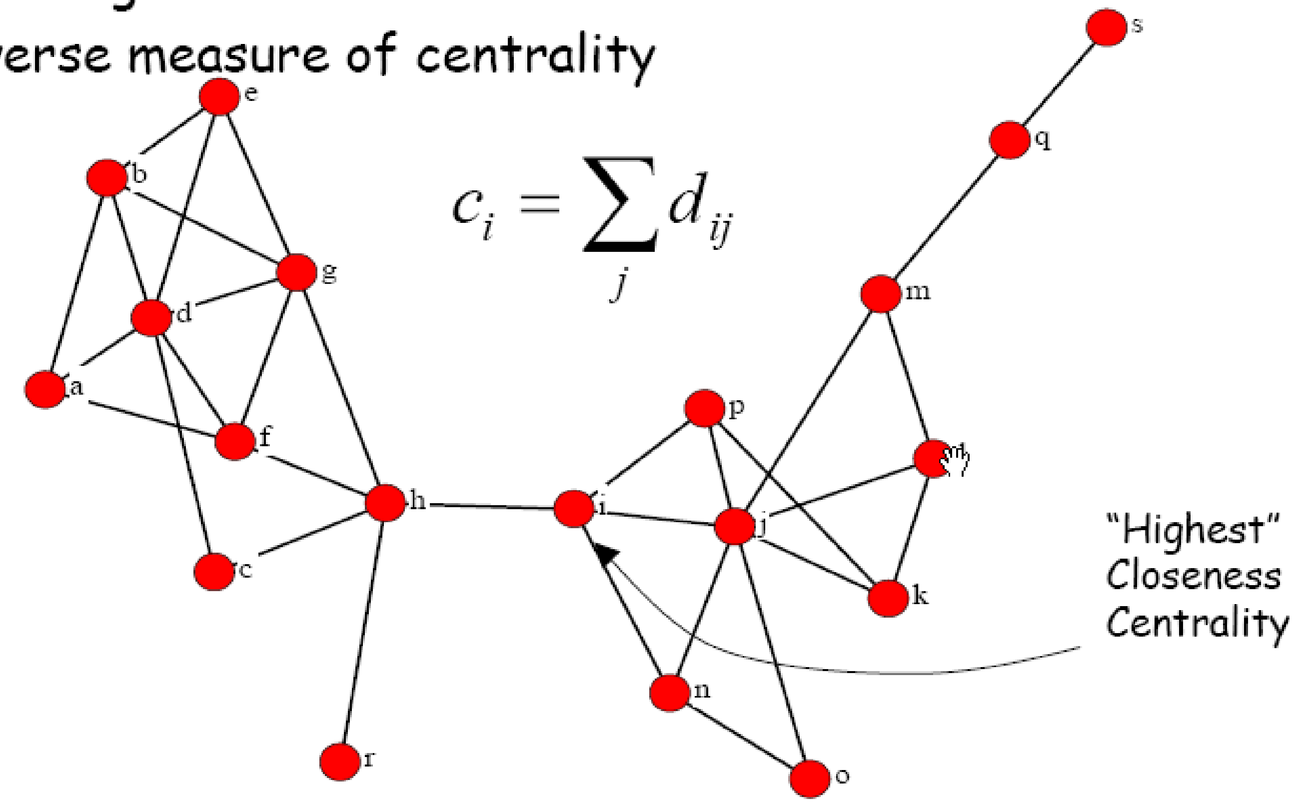
There are 4 edges between the neighbors

C = 2*4 /(8*(8-1)) = 8/56 = 1/7

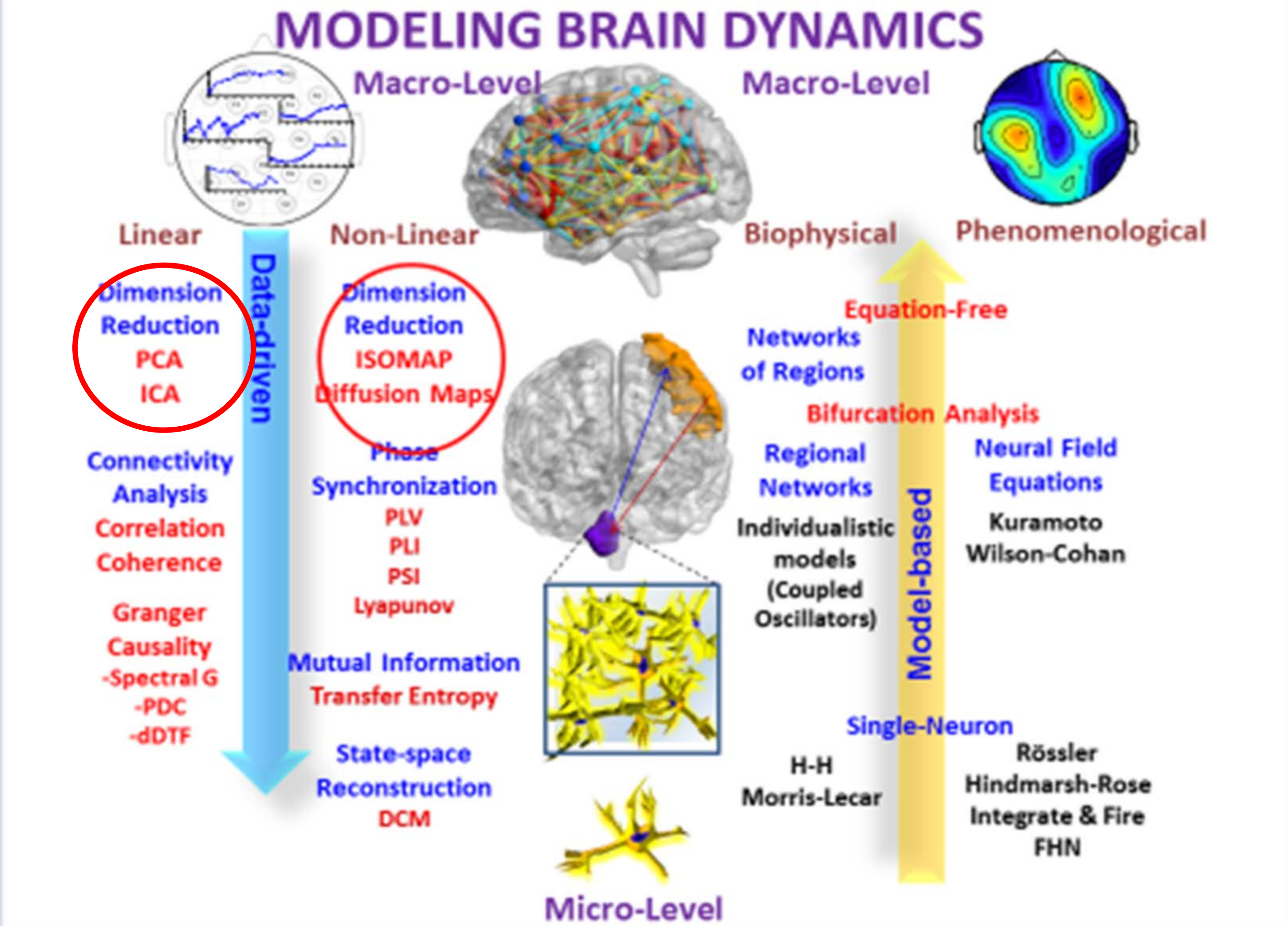
Closeness Centrality

- Sum of geodesic distances to all other nodes
- Inverse measure of centrality

$$c_i = \sum_j d_{ij}$$

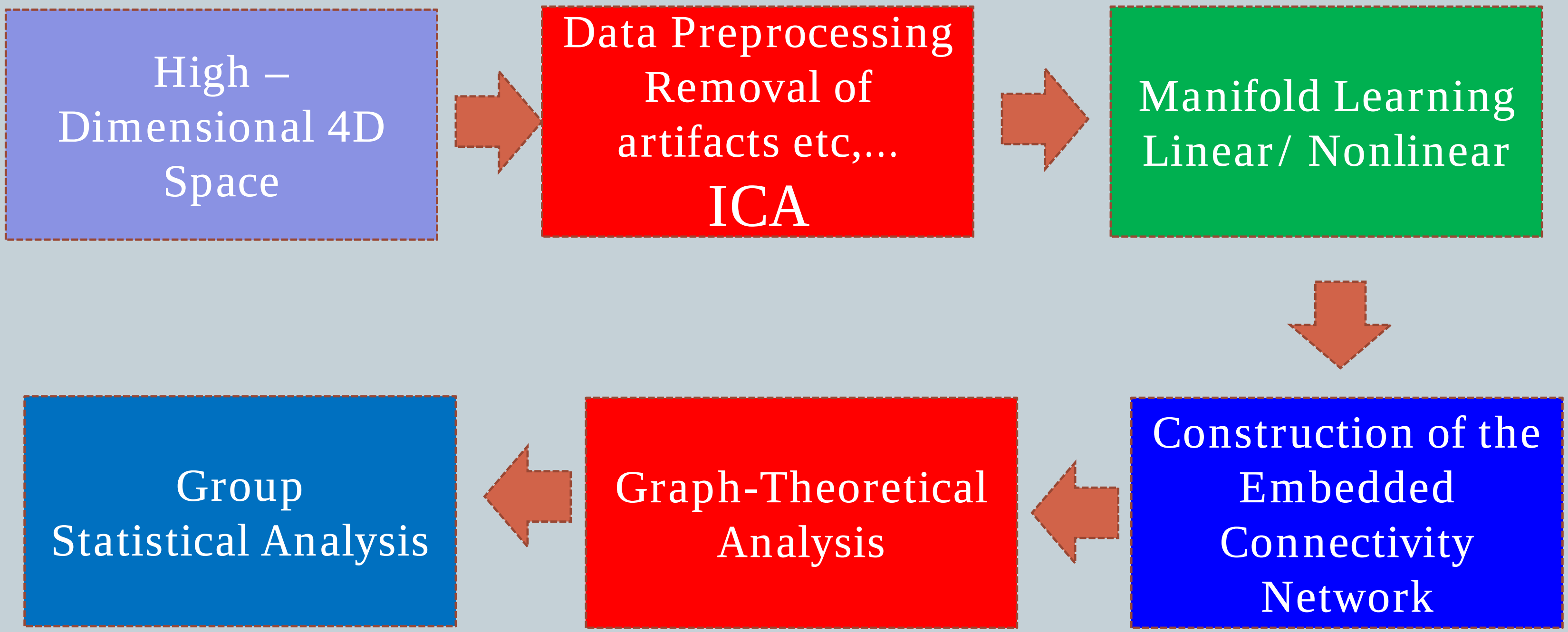


Construction of Connectivity Networks Based on Manifold Learning Techniques



*Siettos, C.I., Starke, J., 2016, Multiscale Modeling of Brain Dynamics: from Single Neurons and Networks to Mathematical Tools, WIREs Systems Biology and Medicine, **8**(5), 438-458.*

Construction of Connectivity Networks Based on Manifold Learning Techniques



Siettos, C.I., Starke, J., 2016, Multiscale Modeling of Brain Dynamics: from Single Neurons and Networks to Mathematical Tools, WIREs Systems Biology and Medicine, ***8****(5), 438-458.*

Preprocessing of the raw FMRI

FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following pre-statistics processing was applied:

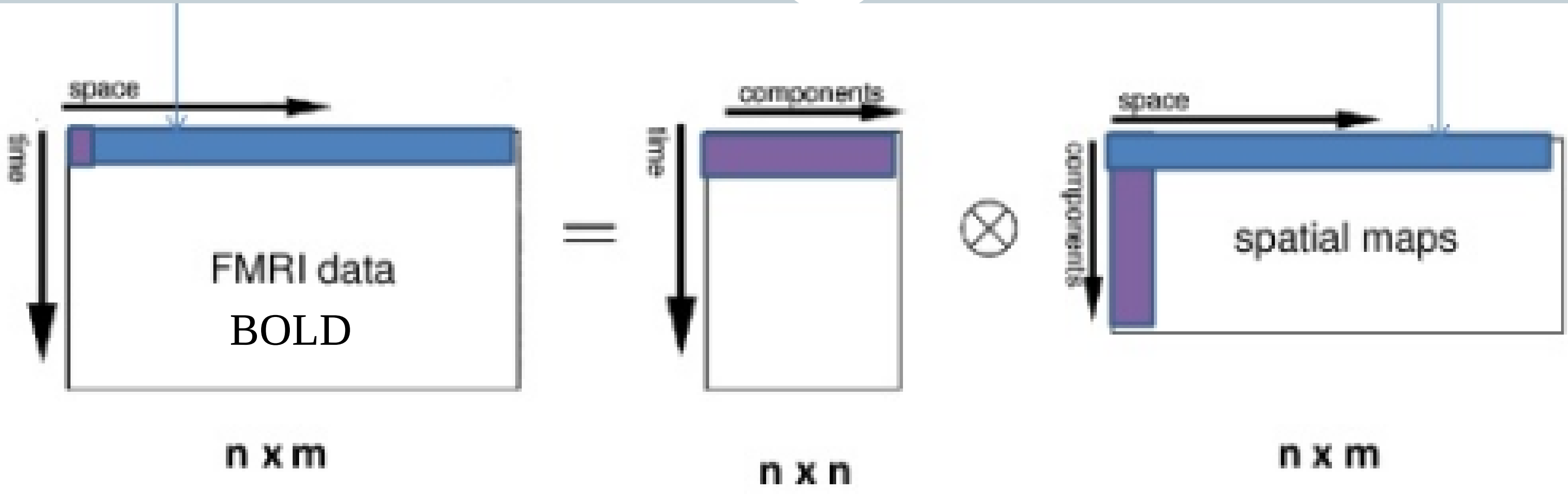
- motion correction using MCFLIRT [Jenkinson 2002];
- slice-timing correction using Fourier-space time-series phase-shifting;
- non-brain removal using BET [Smith 2002];
- spatial smoothing using a Gaussian kernel of FWHM 5mm;
- grand-mean intensity normalization of the entire 4D dataset.

Then we used ICA-AROMA* for removing motion related ICA components of the FMRI data and at last we applied:

- High-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s).

*Pruim, Raimon HR, et al. "ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data." Neuroimage 112 (2015): 267-277.

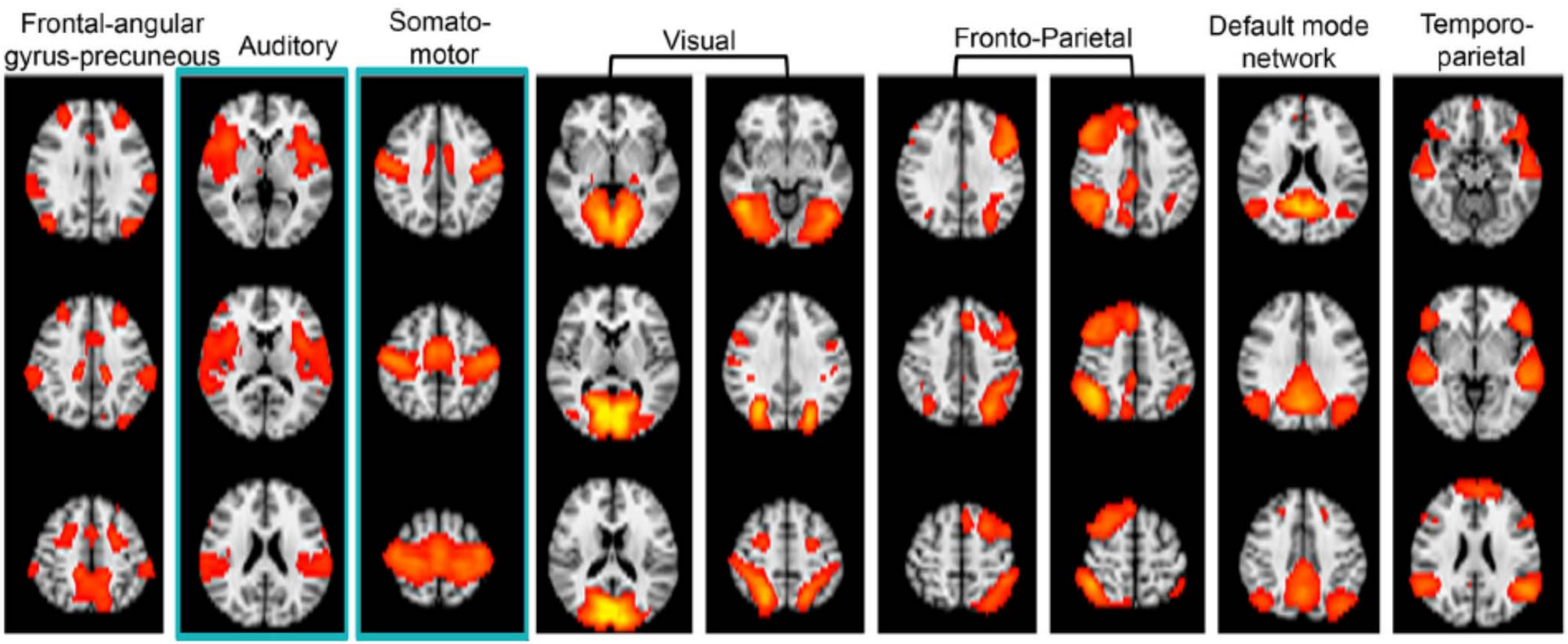
Independent Component Analysis



n time points
m voxels

$$s(t, x, y, z) = \sum_{i=1}^{M} A_i(t) C_i(x, y, z)$$

Independent Component Analysis



Manifold Learning

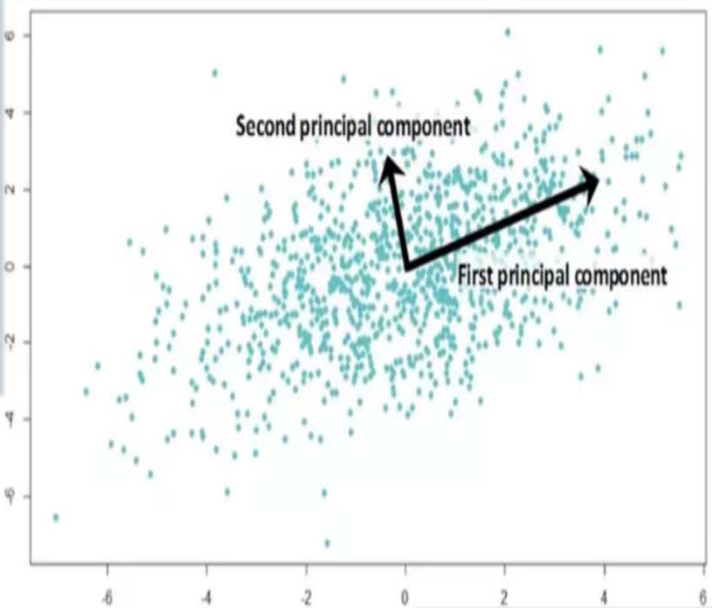
- **PCA**
- **Multidimensional Scaling (MDS)**
- **Local Linear Embedding**
- **Isomap (Isoparametric mapping)***
- **Diffusion Maps ****

Tenenbaum, Joshua B., Vin De Silva, and John C. Langford. "A global geometric framework for nonlinear dimensionality reduction." *science* 290.5500 (2000): 2319-2323.

**Coifman, R.R.; Lafon, S; Lee, A B; Maggioni, M; Nadler, B; Warner, F; Zucker, S W (2005). "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps". PNAS. 102: 7426–7431.

Principal Component Analysis

- Orthogonal linear transformation of data to a new coordinate system so that the greatest variance by some projection of data is captured on the first coordinate (1st principal component), the second greatest variance on the second, and so on.



Step 1: $\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_i$

Step 2: subtract the mean: $\Phi_i = x_i - \bar{x}$

Step 3: form the matrix $A = [\Phi_1 \; \Phi_2 \cdots \Phi_M]$ (NxM matrix), then compute:

$$C = \frac{1}{M}\sum_{n=1}^{M} \Phi_n \Phi_n^T = AA^T$$

(sample **covariance** matrix, NxN, characterizes the *scatter* of the data)

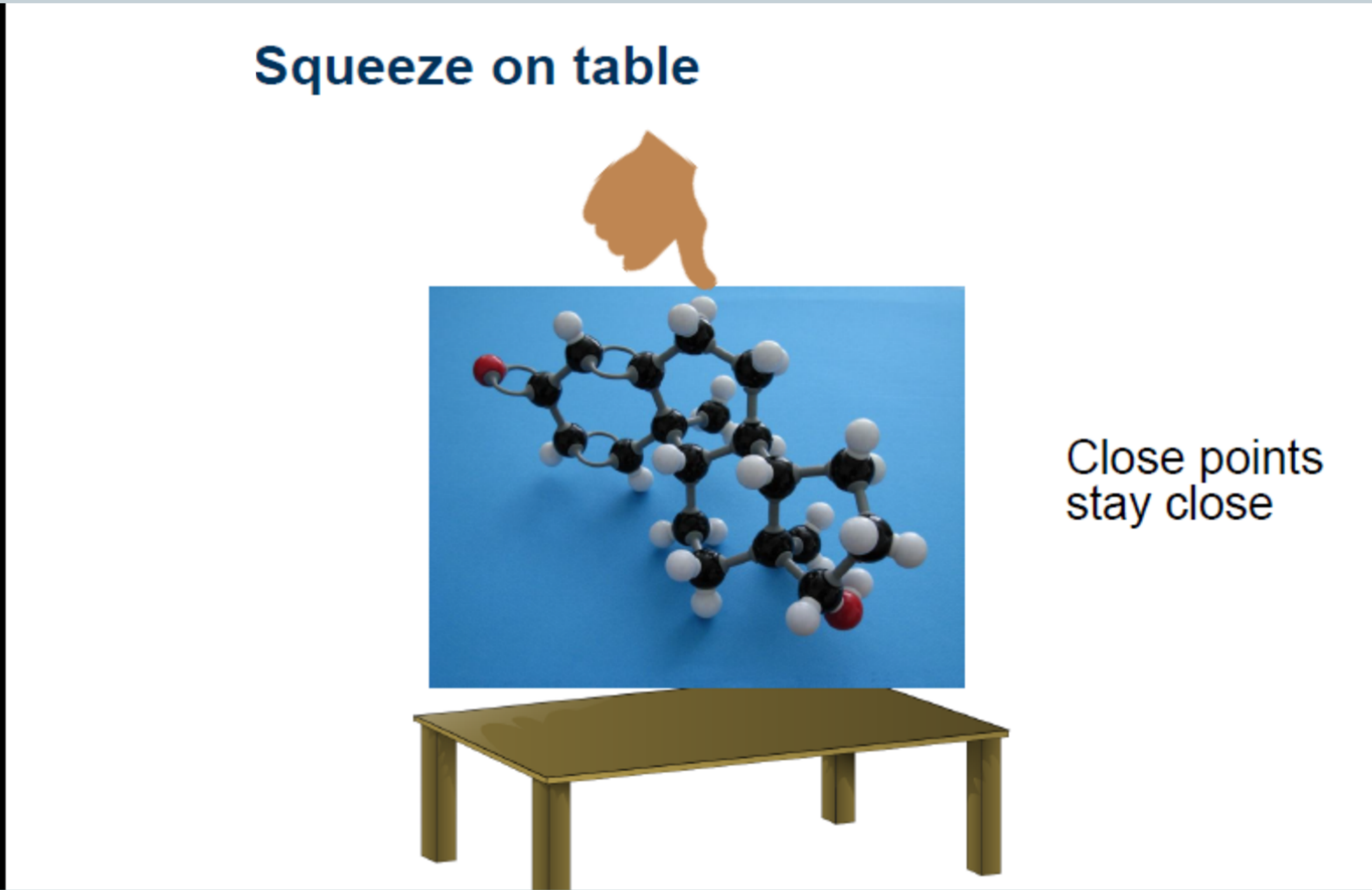
Step 4: compute the eigenvalues of C: $\lambda_1 > \lambda_2 > \cdots > \lambda_N$

Step 5: compute the eigenvectors of C: $u_1, u_2, \ldots, u_N$

MDS

Idea: Represent high dimensional points in a few dimensions keeping distances between points similar

Given distances between points recover the positions of the points!



MDS

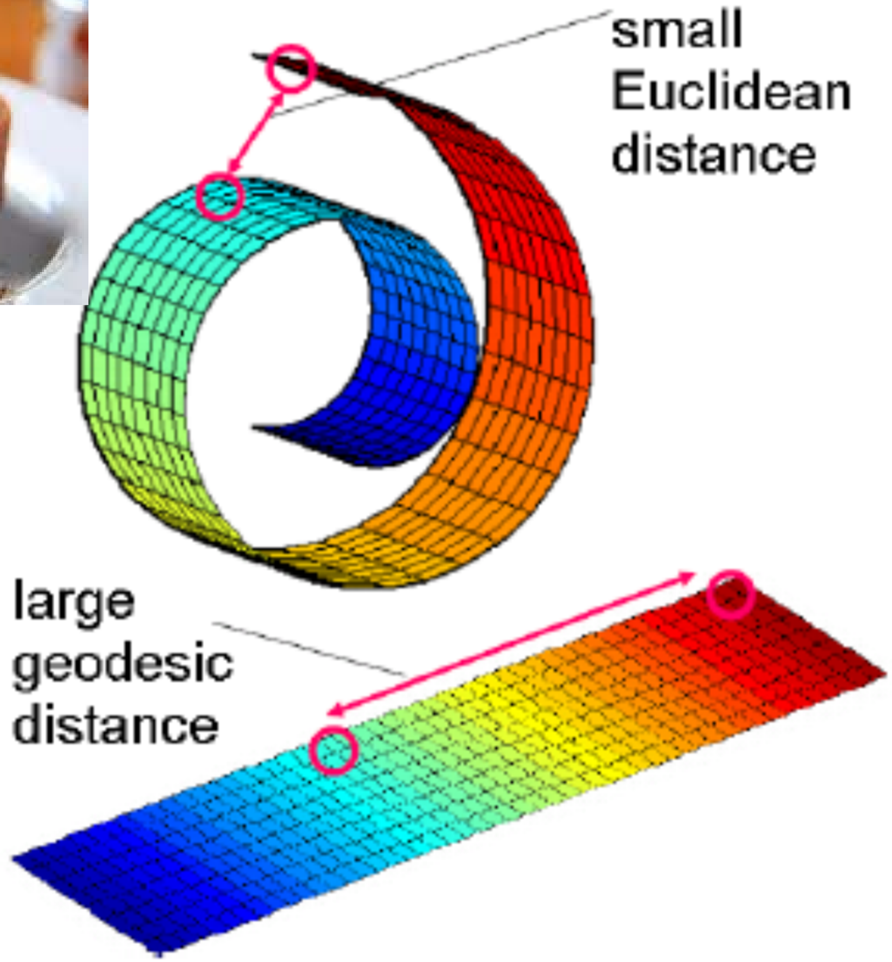
1. Set up the matrix of squared proximities $\mathbf{P^{(2)}} = [p^2]$.

2. Apply the double centering: $\mathbf{B} = -\frac{1}{2}\mathbf{J}\mathbf{P^{(2)}}\mathbf{J}$ using the matrix $\mathbf{J} = \mathbf{I} - n^{-1}\mathbf{1}\mathbf{1}'$, where n is the number of objects.

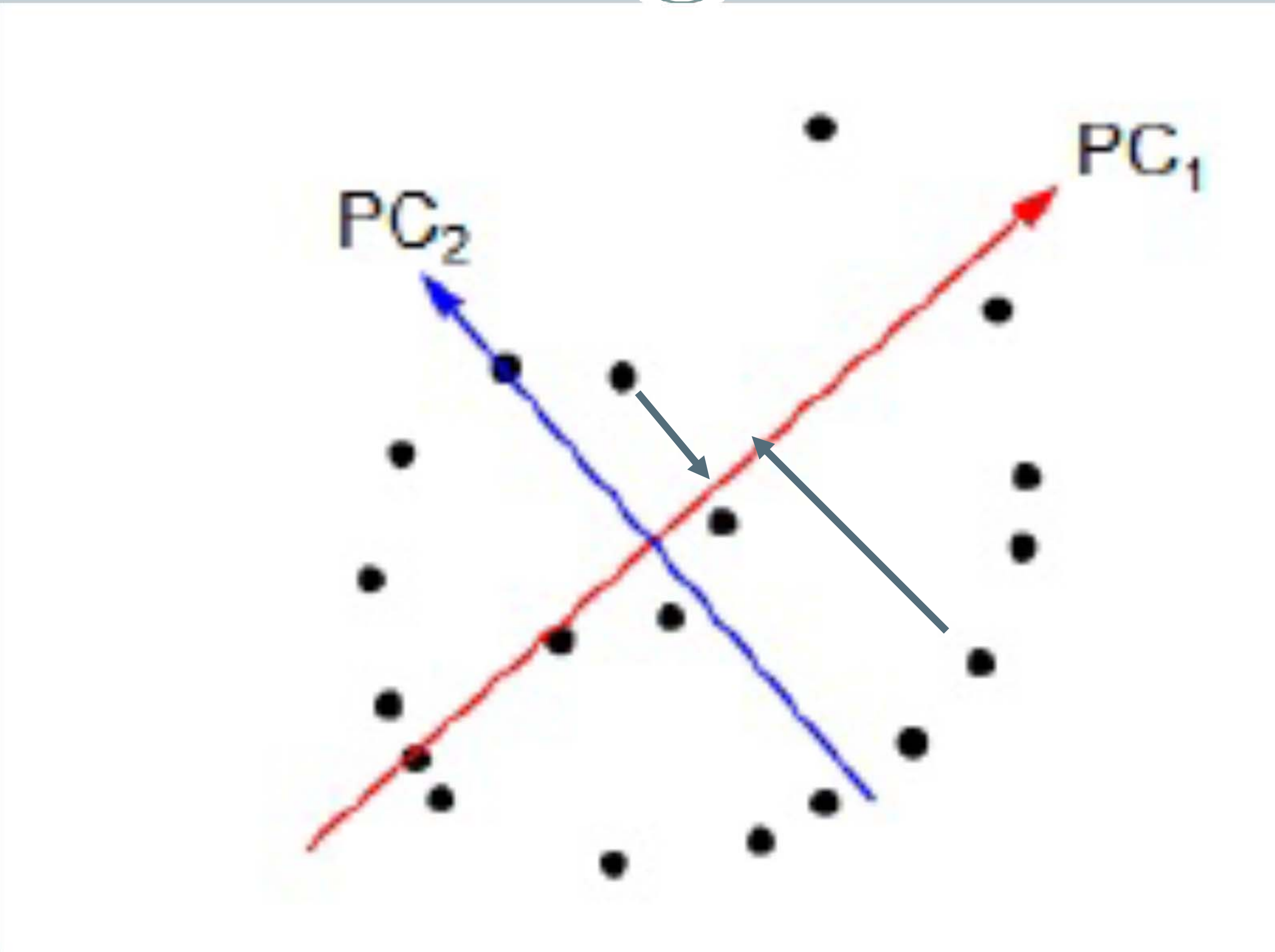
3. Extract the m largest positive eigenvalues $\lambda_1 \ldots \lambda_m$ of $\mathbf{B}$ and the corresponding m eigenvectors $\mathbf{e_1} \ldots \mathbf{e_m}$.

4. A m-dimensional spatial configuration of the n objects is derived from the coordinate matrix $\mathbf{X} = \mathbf{E_m}\mathbf{\Lambda}_m^{\mathbf{1/2}}$, where $\mathbf{E_m}$ is the matrix of m eigenvectors and $\mathbf{\Lambda}_m$ is the diagonal matrix of m eigenvalues of $\mathbf{B}$, respectively.

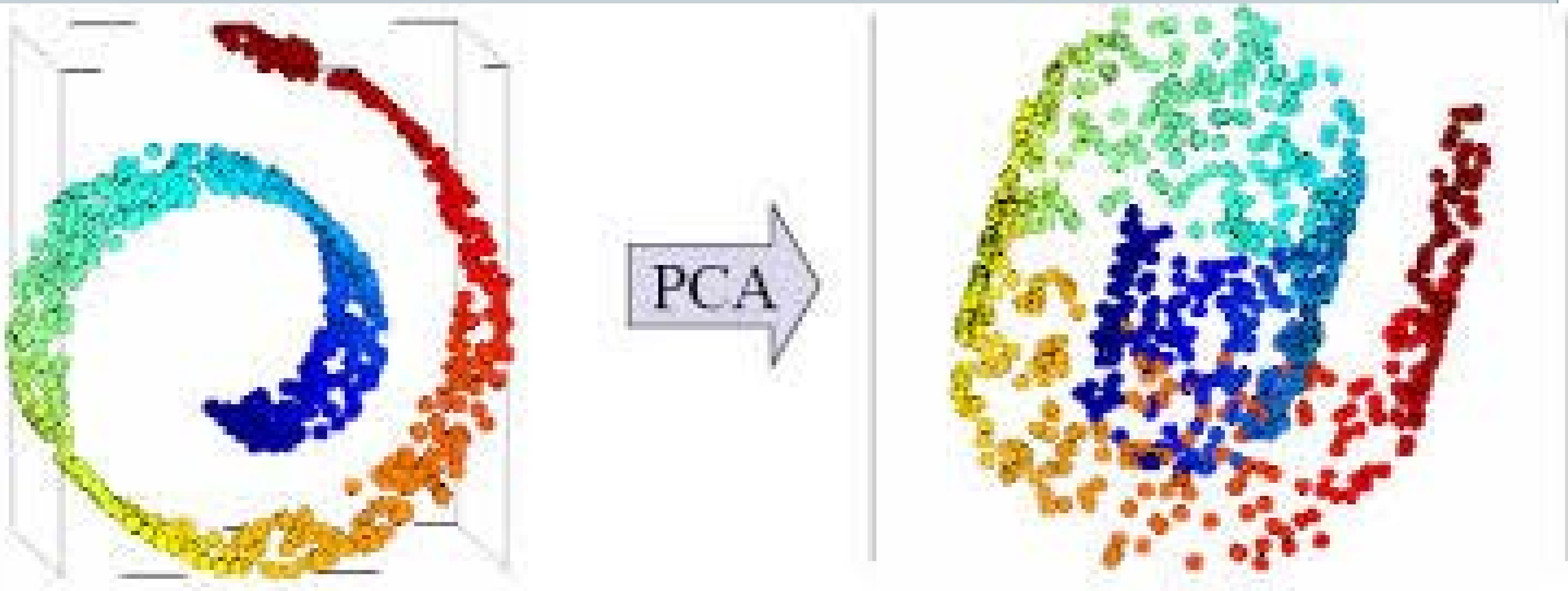
Manifold Learning Algorithms



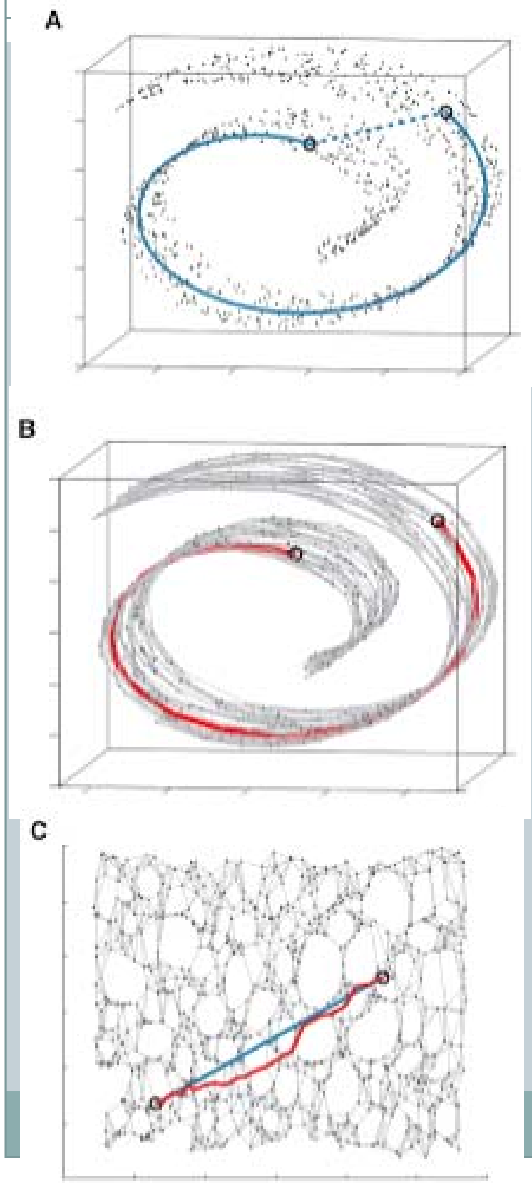
Manifold Learning Algorithms
PCA and MDS would fail



Manifold Learning Algorithms



Manifold Learning Algorithms



Detect the Manifold!

•For two arbitrary points in a high dimensional space the Euclidian distance may not reflect the intrinsic similarity of features

• Hence, the goal is to measure the distance as the blue line in (A).

• This for example can be done by computing the geodesic distances (ISOMAP) on the manifold (B).

To construct a low dimensional space (C)

Construction of a Distance matrix: Metrics

- *Cross/ Pearson Correlation-Based*

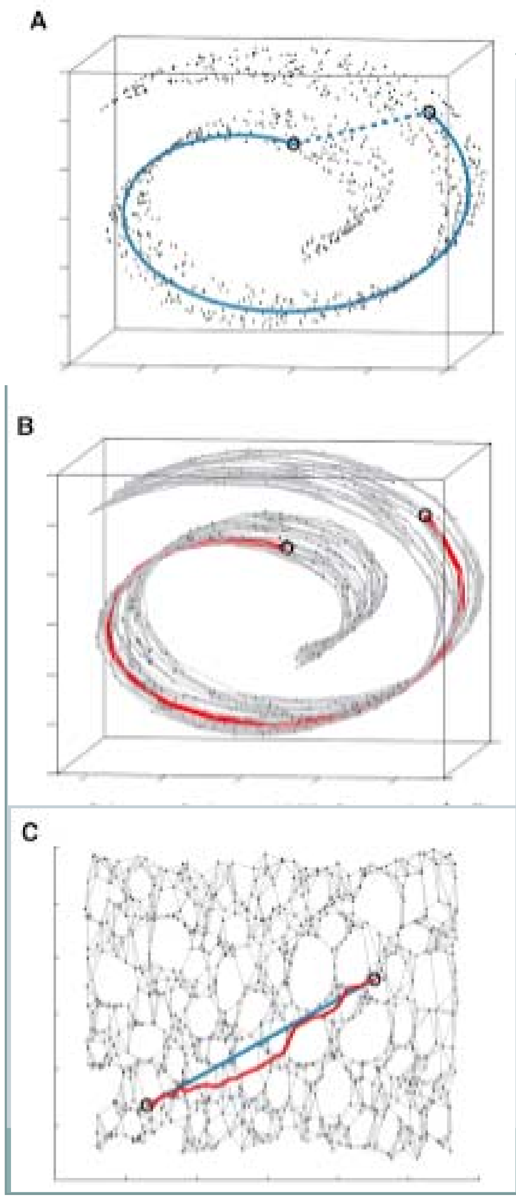
$$d\left(A_{\mu_i}, A_{\mu_j}\right) = 1 - max[\,|CCF\left(A_{\mu_i}, A_{\mu_j}, l\right)|\,]$$

- *Euclidian norm*

- *Gaussian Kernel* $\exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\varepsilon}\right)$

- …….

ISOMAP

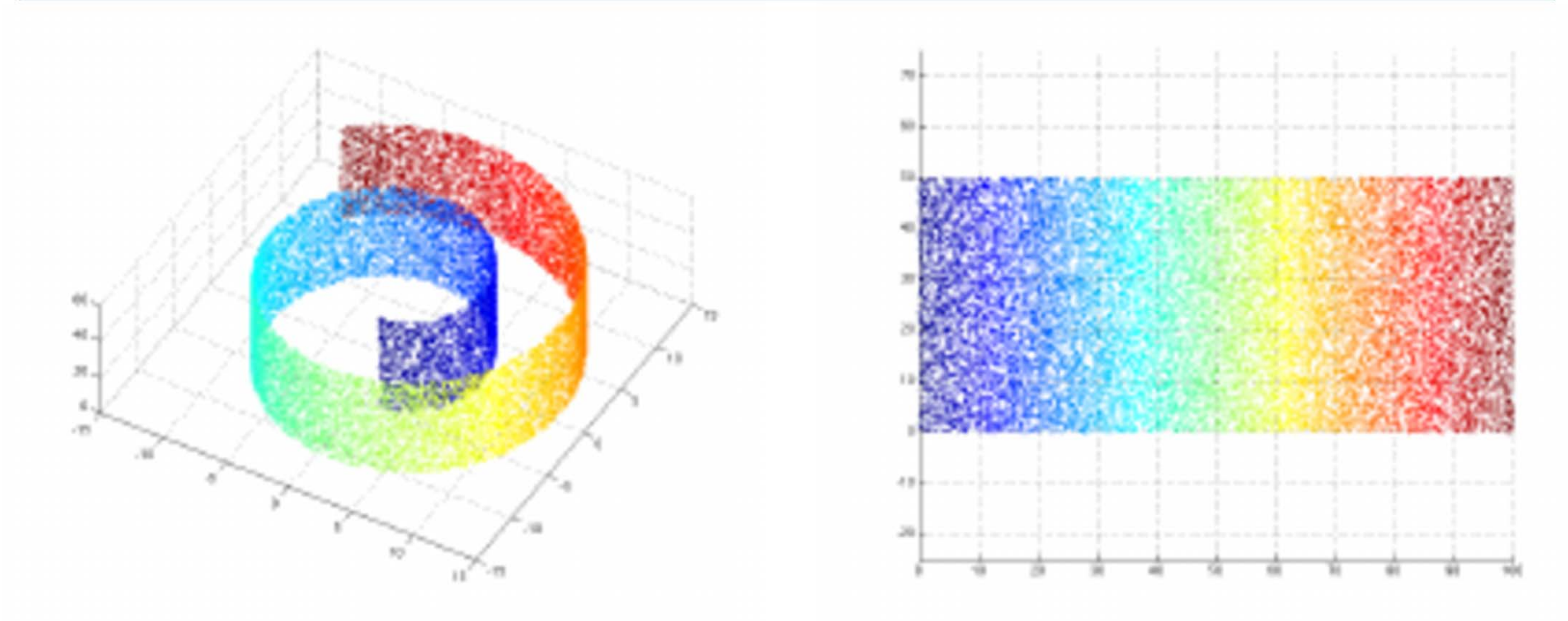


•Construction of the connectivity graph based on the k-nearest neighbors of the components based on a metric (Euclidian, Correlation, etc..).

•Computation of the shortest path between each pair of nodes in the graph.
This will result to a square matrix *D* with elements representing the length of the shortest paths **(called geodesic distances)** among nodes.

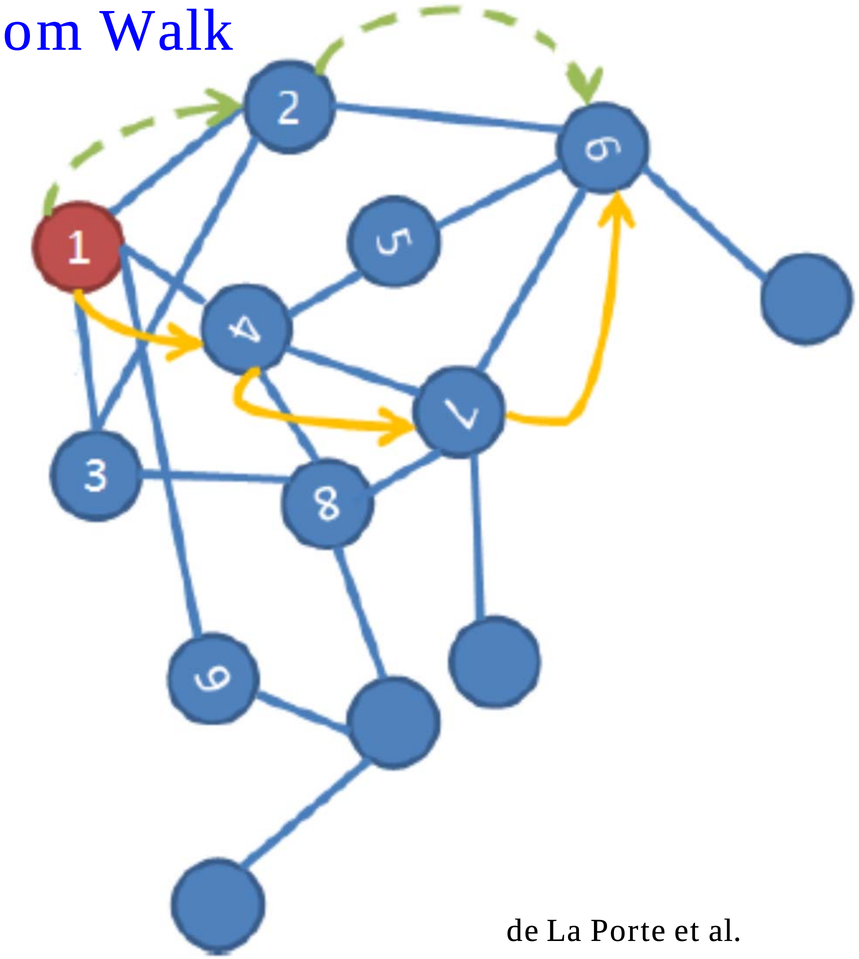
•Construct the d-dimensional embedding space.
The embedded space is spanned by the first *d* eigenvectors corresponding to the largest eigenvalues of *D*.

ISOMAP



Diffusion Maps

Random Walk



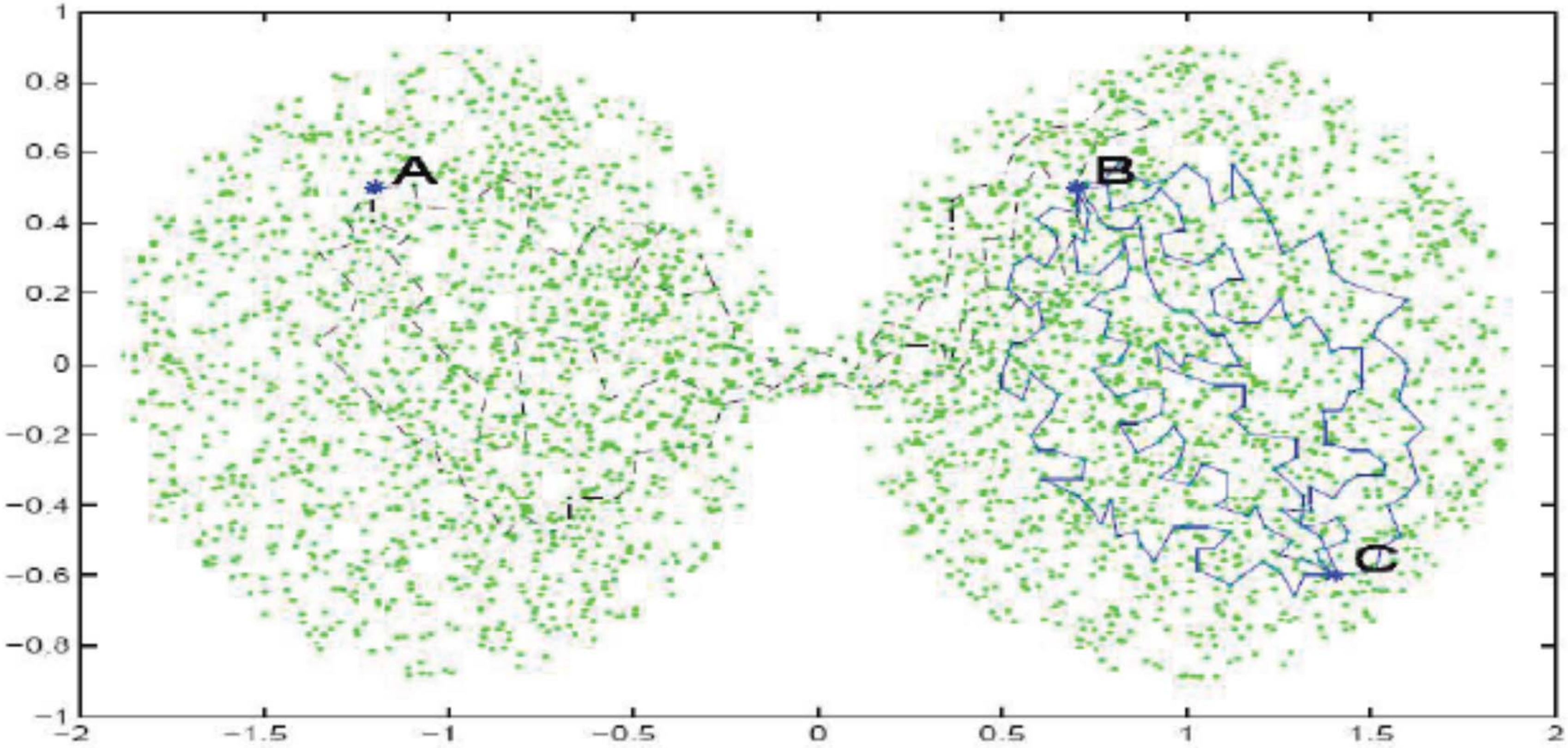
de La Porte et al.

Diffusion Maps

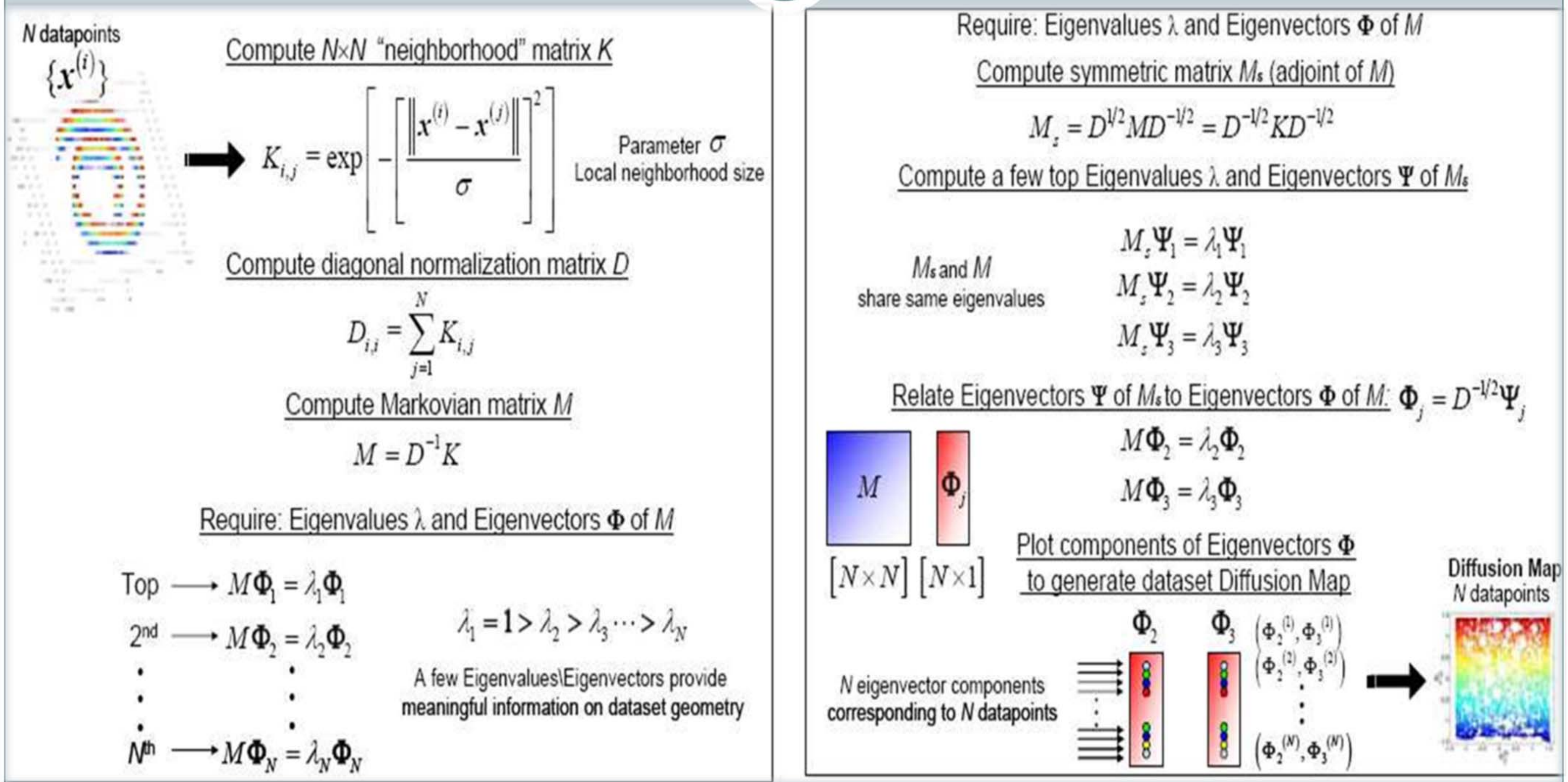
$d_{geod.}(A,B) \sim d_{geod.}(C,B)$, however $d^{(t)}(A,B) >> d^{(t)}(C,B)$.

Coifman & Lafon 2006

$$Dis_t^2(\boldsymbol{x}_0,\boldsymbol{x}_1) = \sum_{\boldsymbol{y}=\boldsymbol{x}_1}^{\boldsymbol{x}_N} \left(p(t,\boldsymbol{y}\mid\boldsymbol{x}_0) - p(t,\boldsymbol{y}\mid\boldsymbol{x}_1)\right)^2 w(y)$$



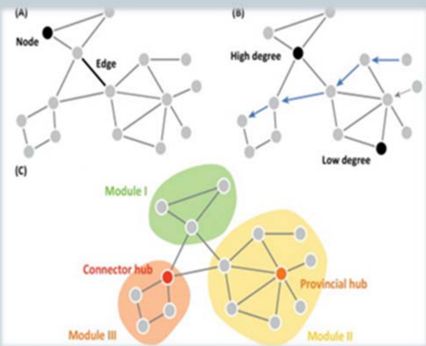
Diffusion Maps

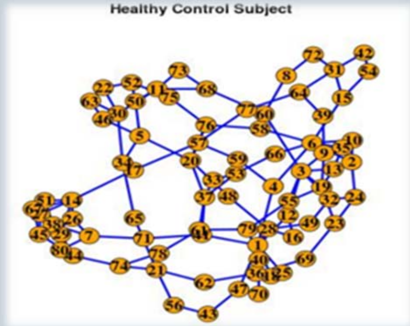


Diffusion distance = Euclidean distance in diffusion map space

Construction of networks: Thresholding

- ISOMAP constructs embedded graphs using k-NN algorithm, thus no further thresholding is needed.





- PCA and Diffusion Maps result to FULL CONNECTED GRAPHS. So we construct (sparse) networks using the 5% to 70% of the strongest connections.

Graph metrics

Most used brain graph metrics

	metric	features			requirements
		weighted	directed	negative	connected
large	characteristic path length	yes	yes	no	yes
	global-efficiency	yes	yes	no	no
	clustering coefficient	yes	yes	no	no
	local-efficiency	yes	yes	no	no
	modularity	yes	yes	yes	no
intermediate	communities	yes	yes	no	no
	motifs	yes	yes	no	no
	edge betweenness	yes	yes	no	no
	redundancy	no	yes	no	no
small	degree	yes	yes	yes	no
	node betweenness	yes	yes	no	no
	eigenvector centrality	yes	yes	yes	yes
	accessibility	yes	yes	no	yes

We make sure that the graphs between the two groups are comparable!

We check if graphs have no significant differences in the number of edges, vertices, graph density and median degree of nodes*.Differences in any of these measures could bias the analysis*.

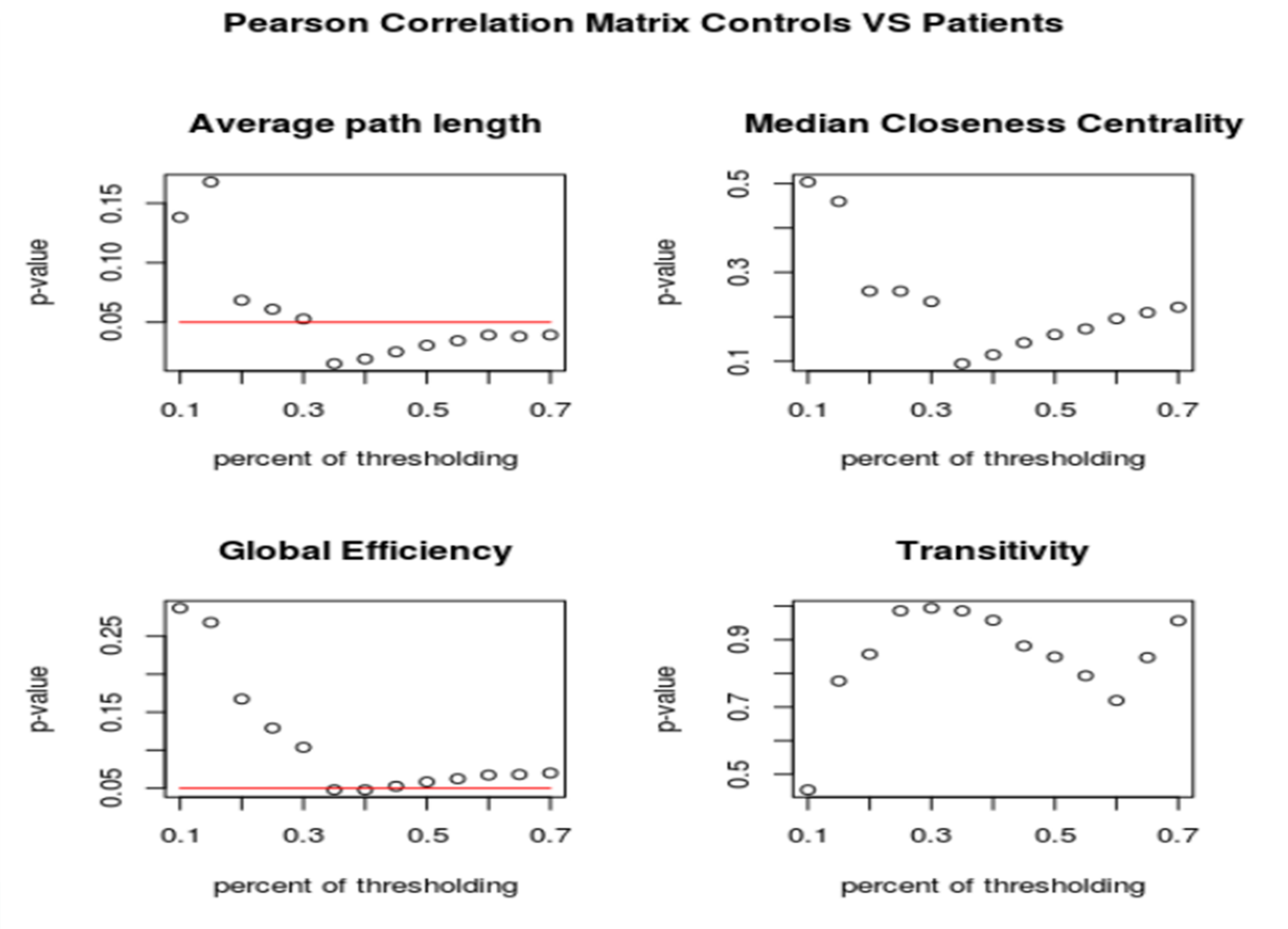
*Van Wijk, B. C., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. PloS one, 5(10), e13701.

The application: Clinical Data

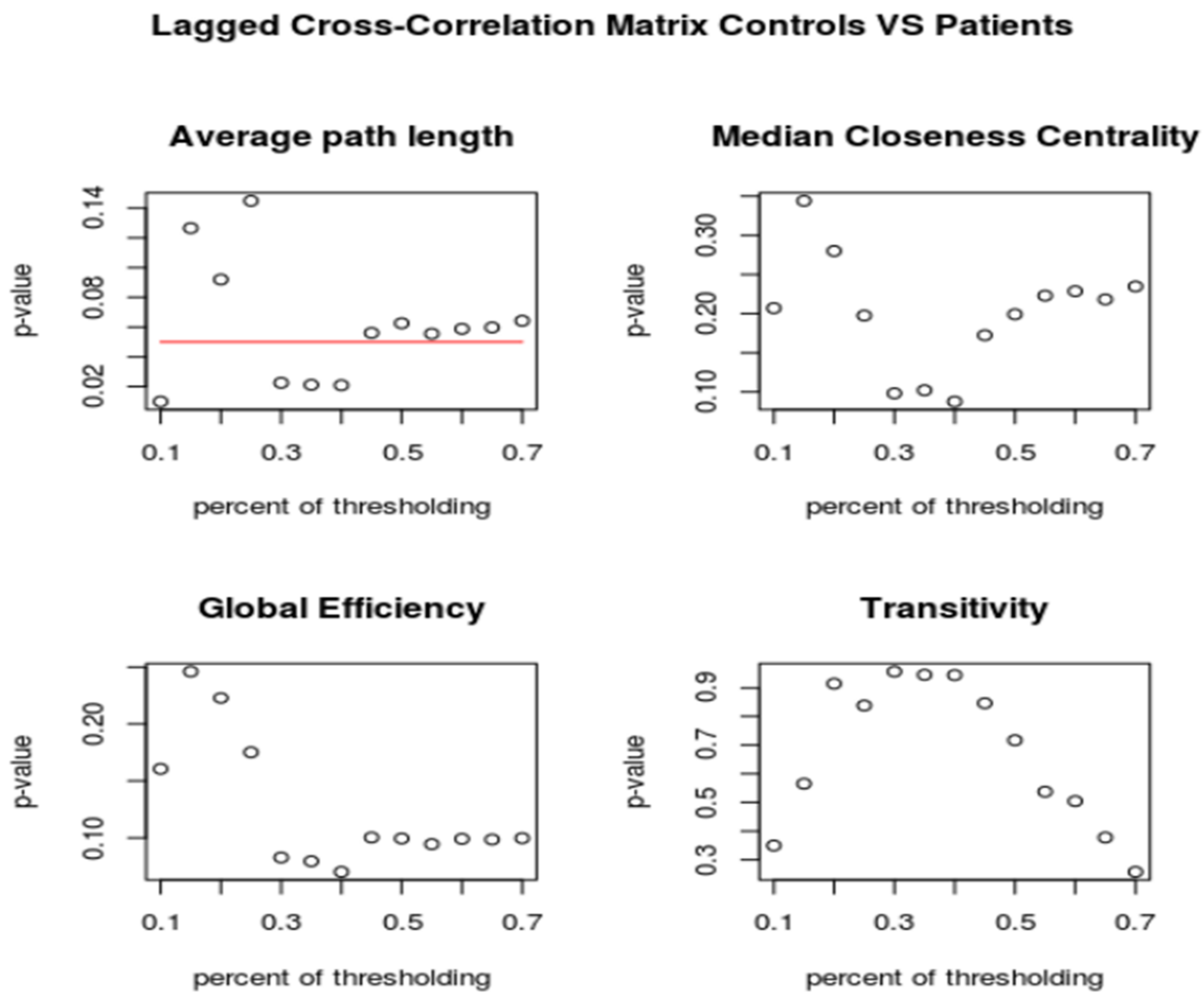
- Our clinical data consists of 74 healthy controls and 72 patients with Schizophrenia.

- The Fmri Resting-State data and phenotypic information were collected and shared by the Mind Research Network and the University of New Mexico funded by a National Institute of Health Center of Biomedical Research Excellence (COBRE)

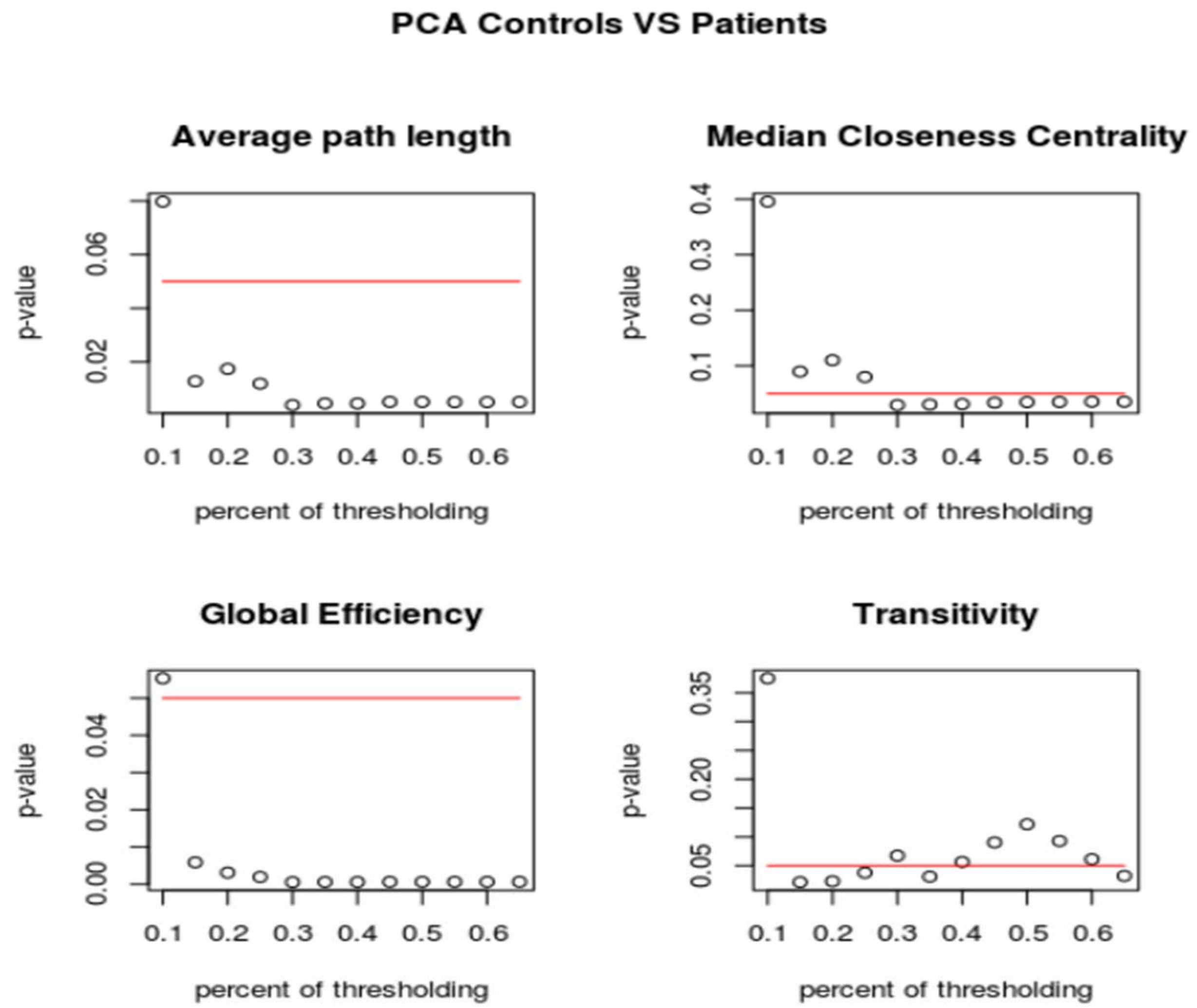
Global Network Metrics: Pearson Correlation Based



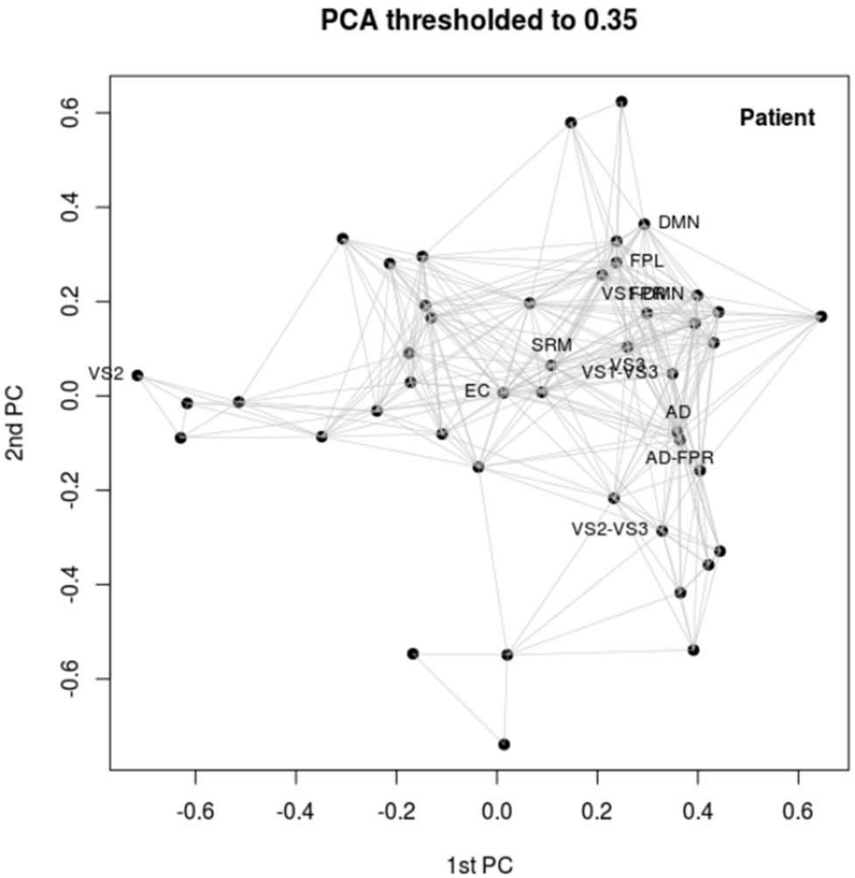
Global Network Metrics: Cross-Correlation Based

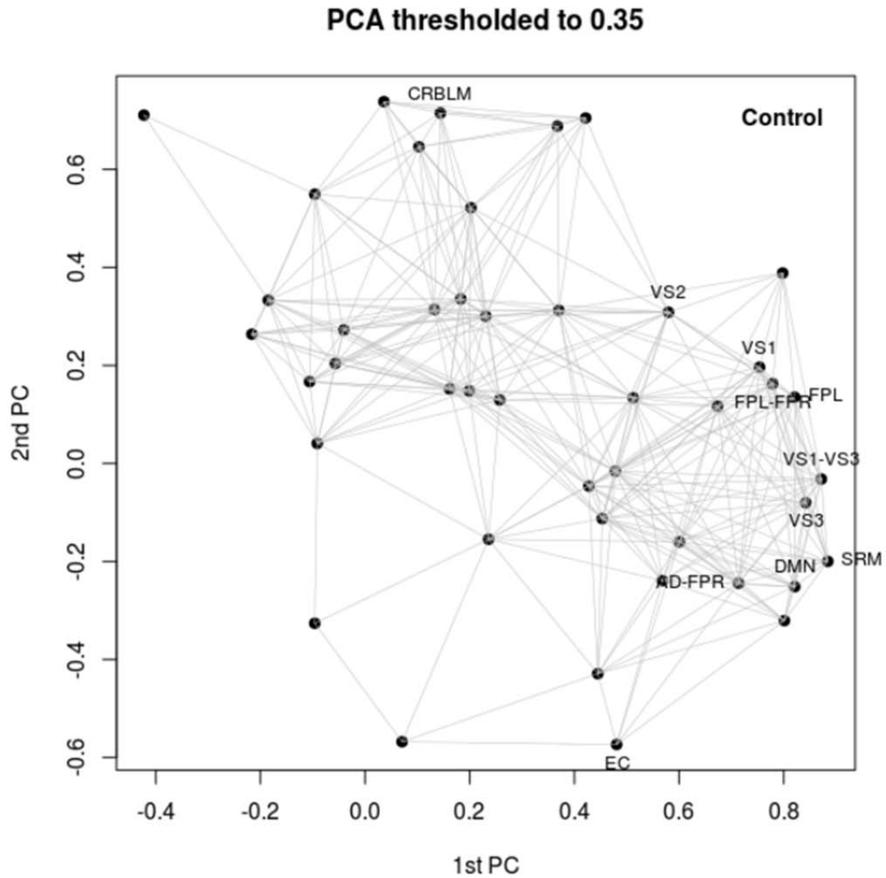


Global Network Metrics: PCA

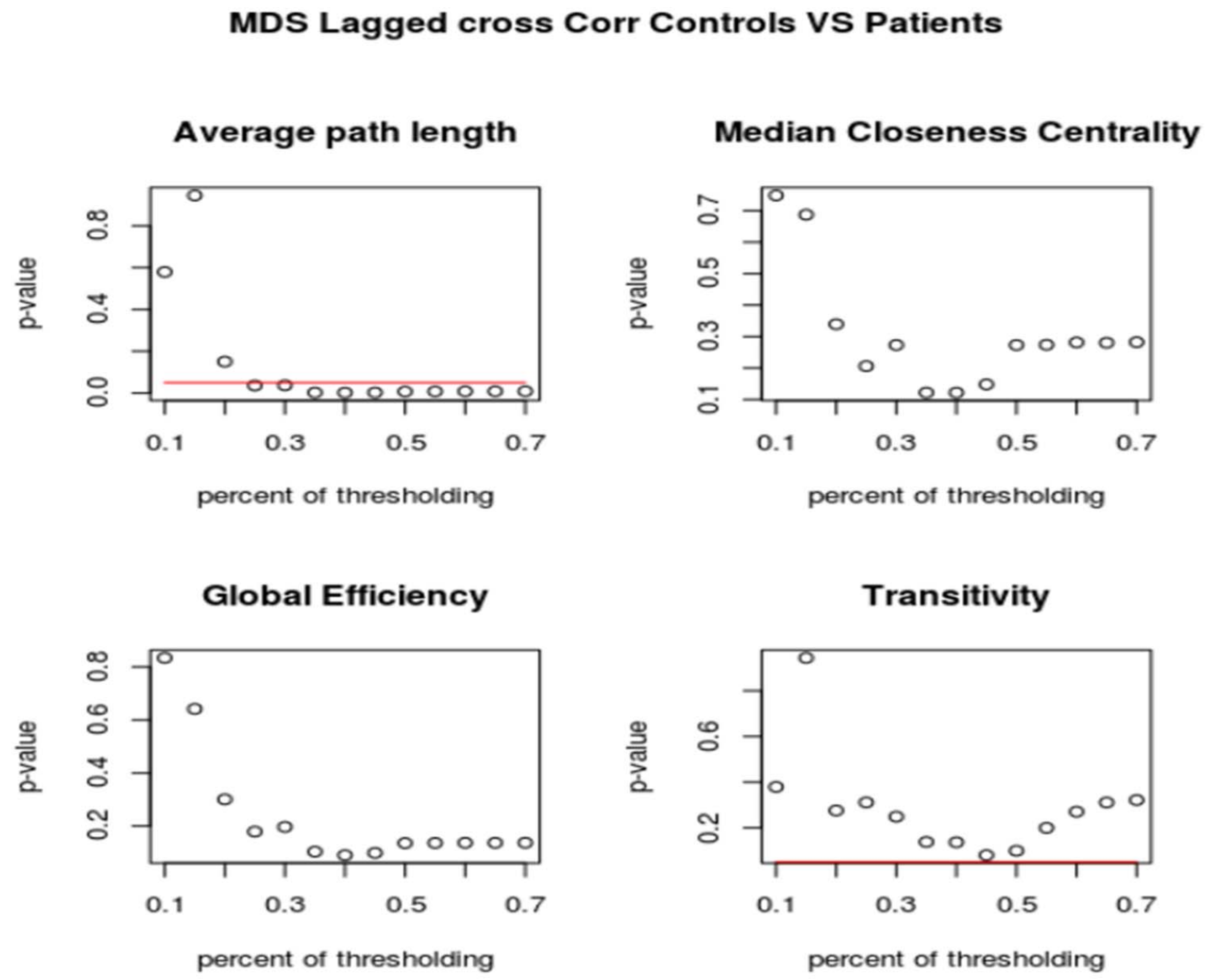


Embedded Networks-PCA

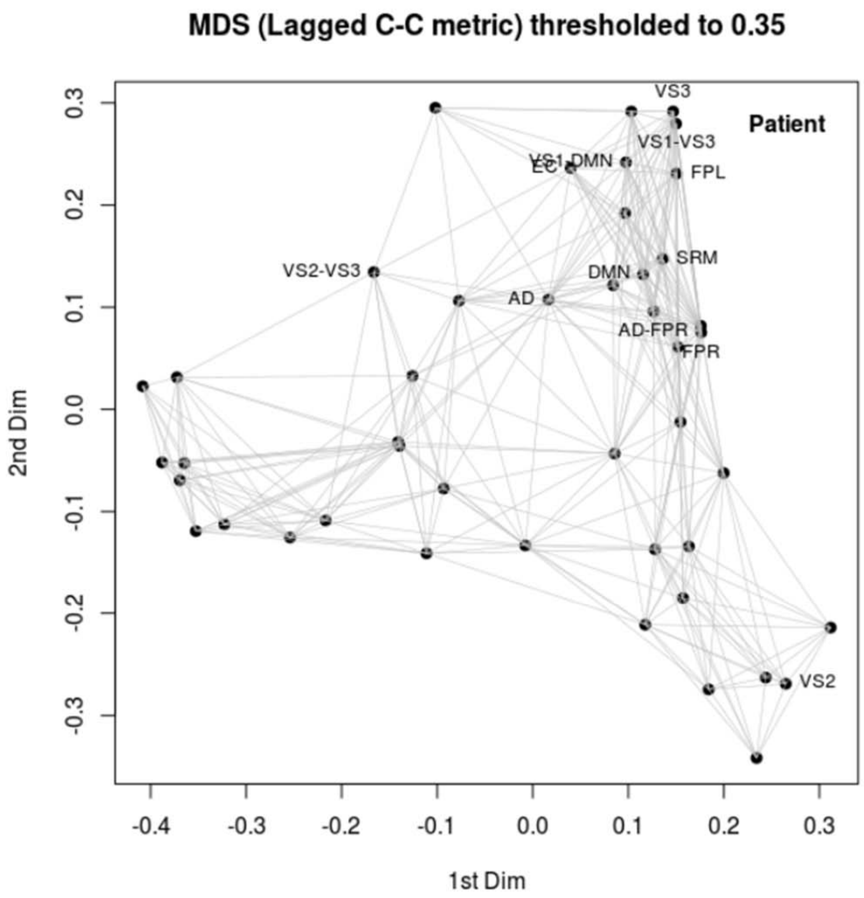


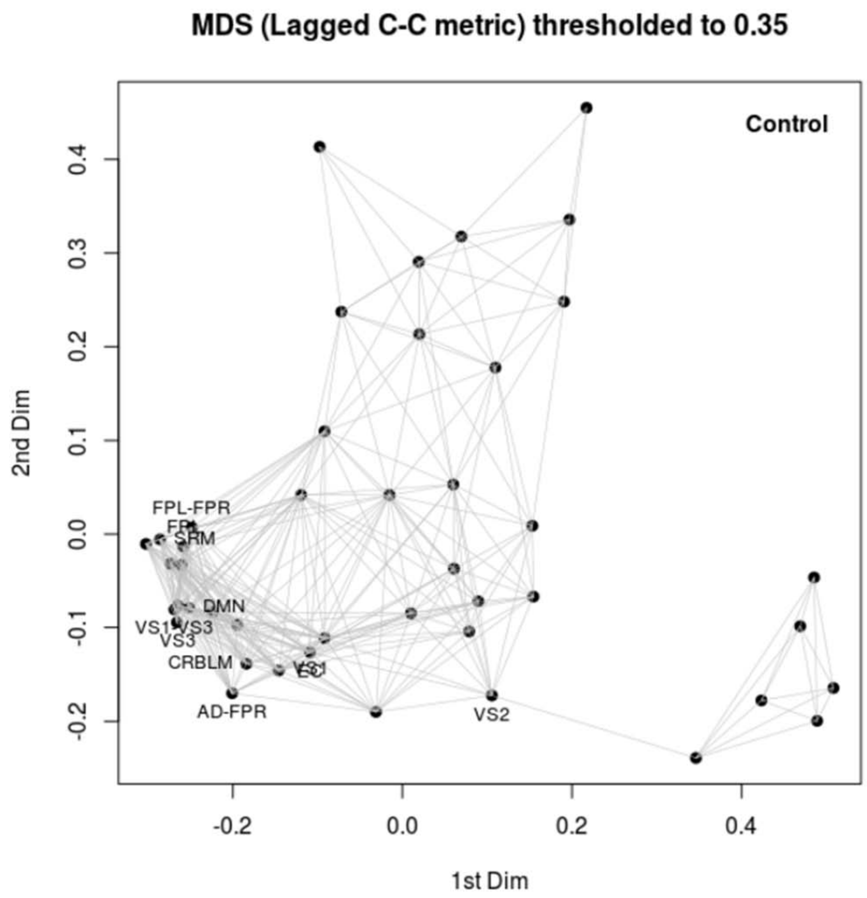


Global Network Metrics: MDS-Correlation Based

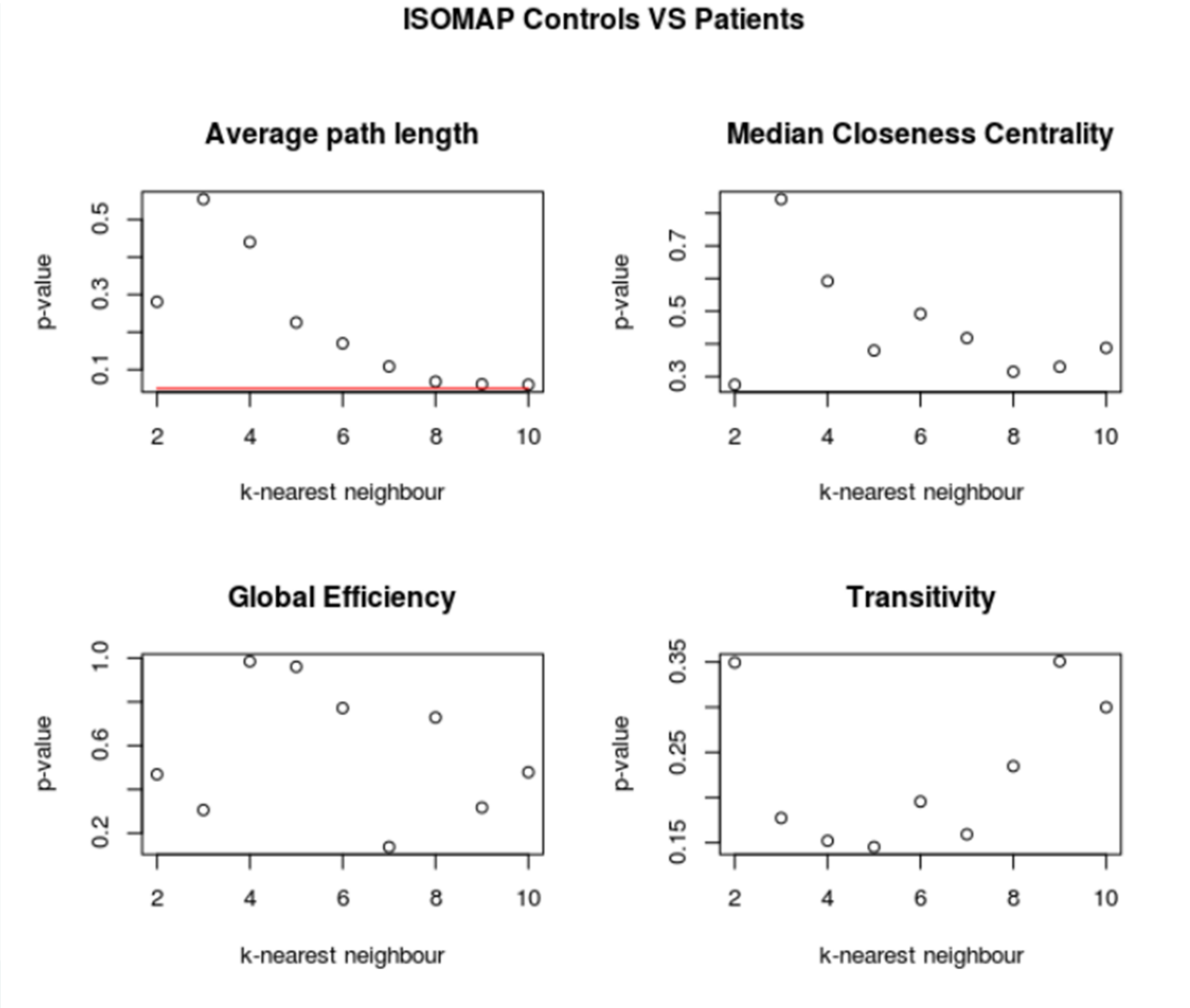


Embedded Networks: MDS-Correlation Based

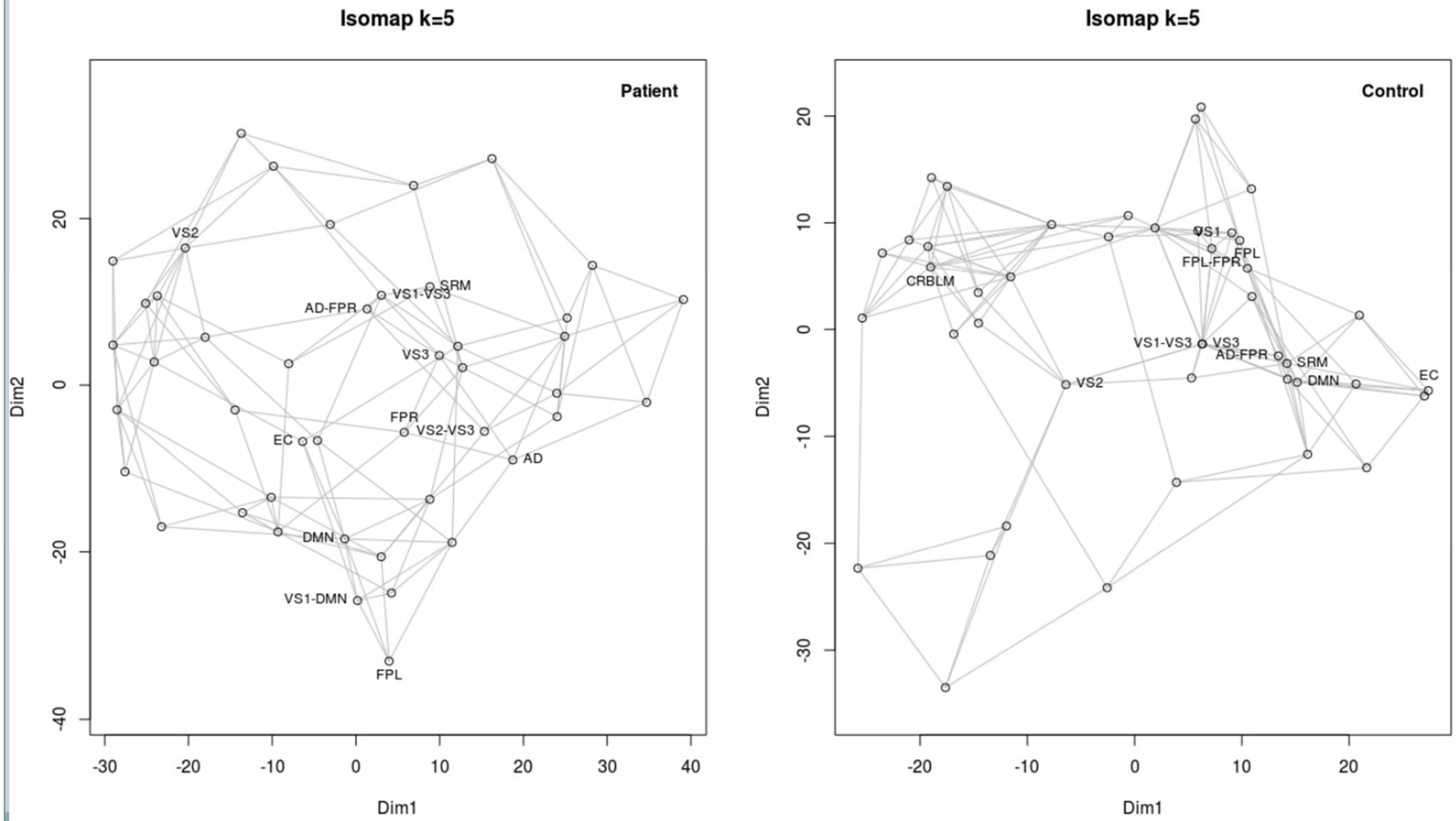




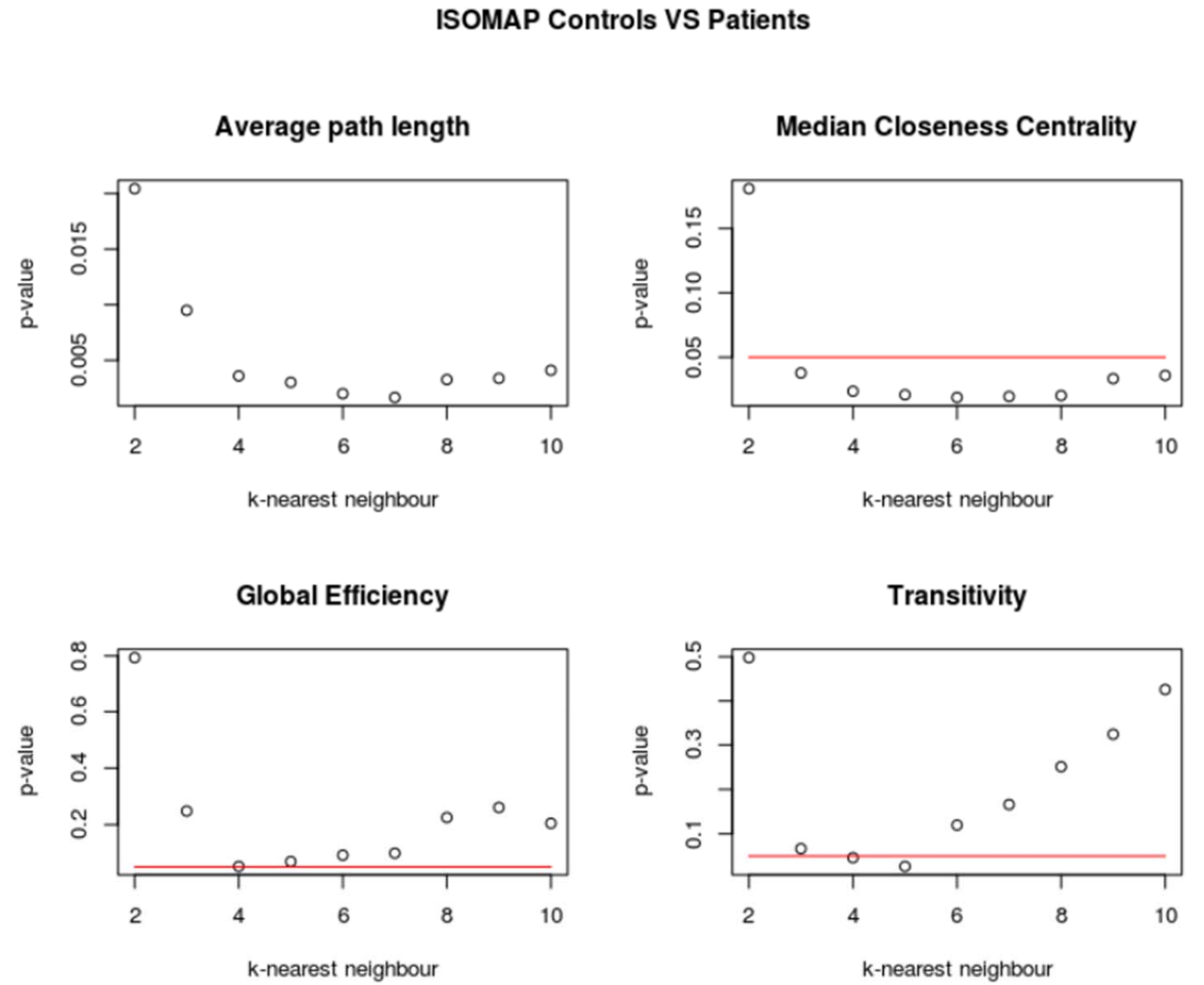
Isomap-Euclidian Metric



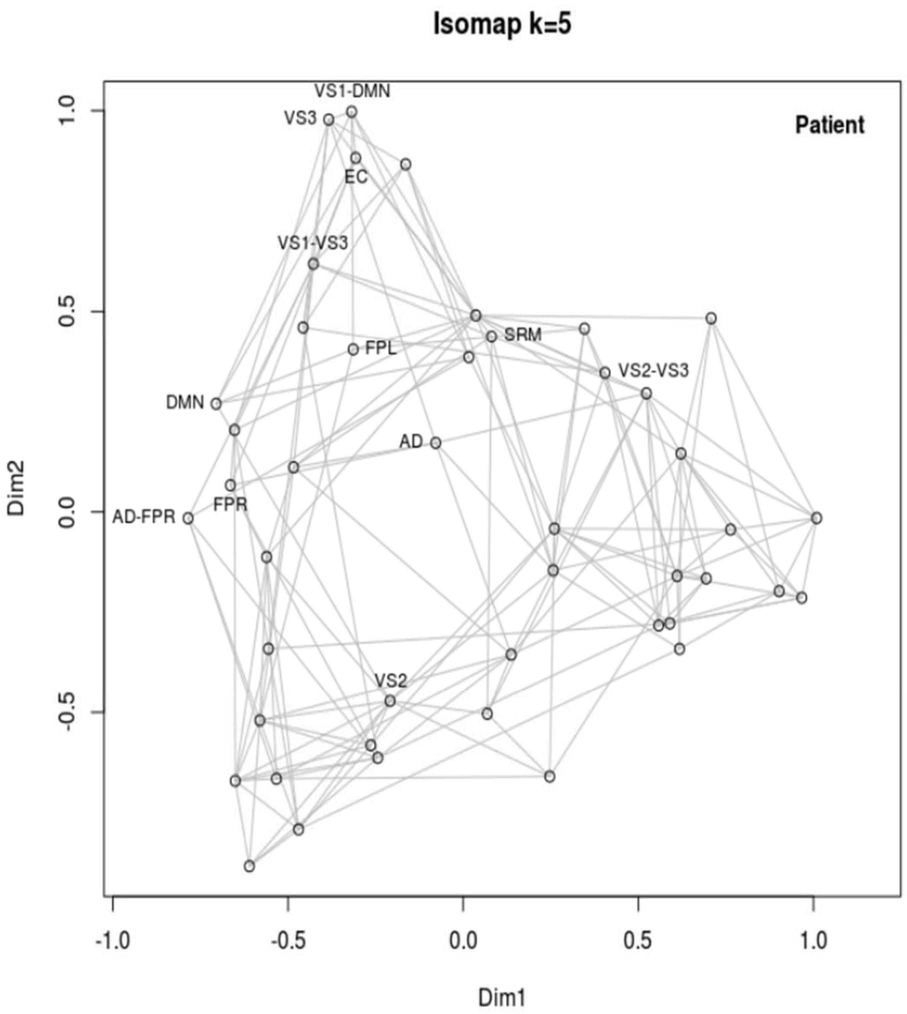
Embedded Networks – Isomap-Euclidian

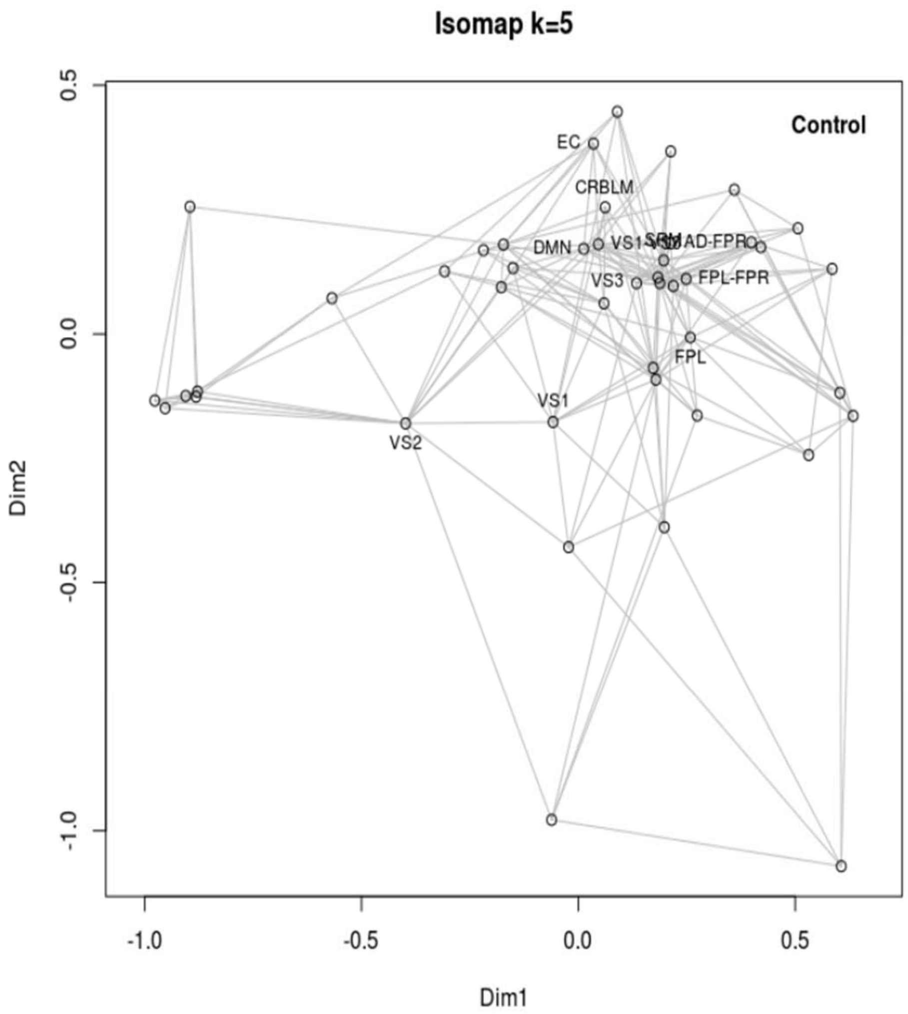


Isomap-Cross Correlation-based

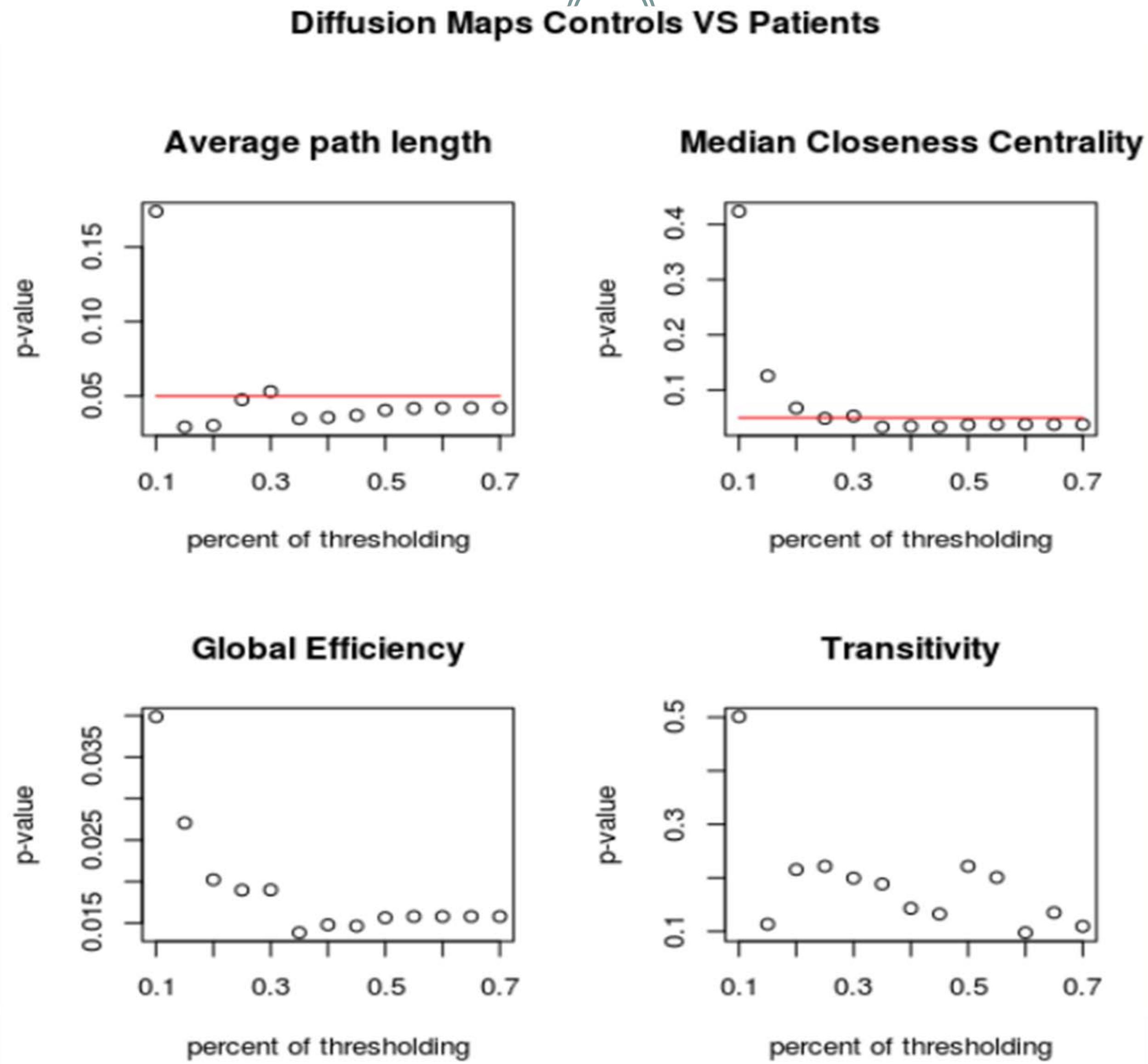


Embedded Networks
Isomap – Correlation Based

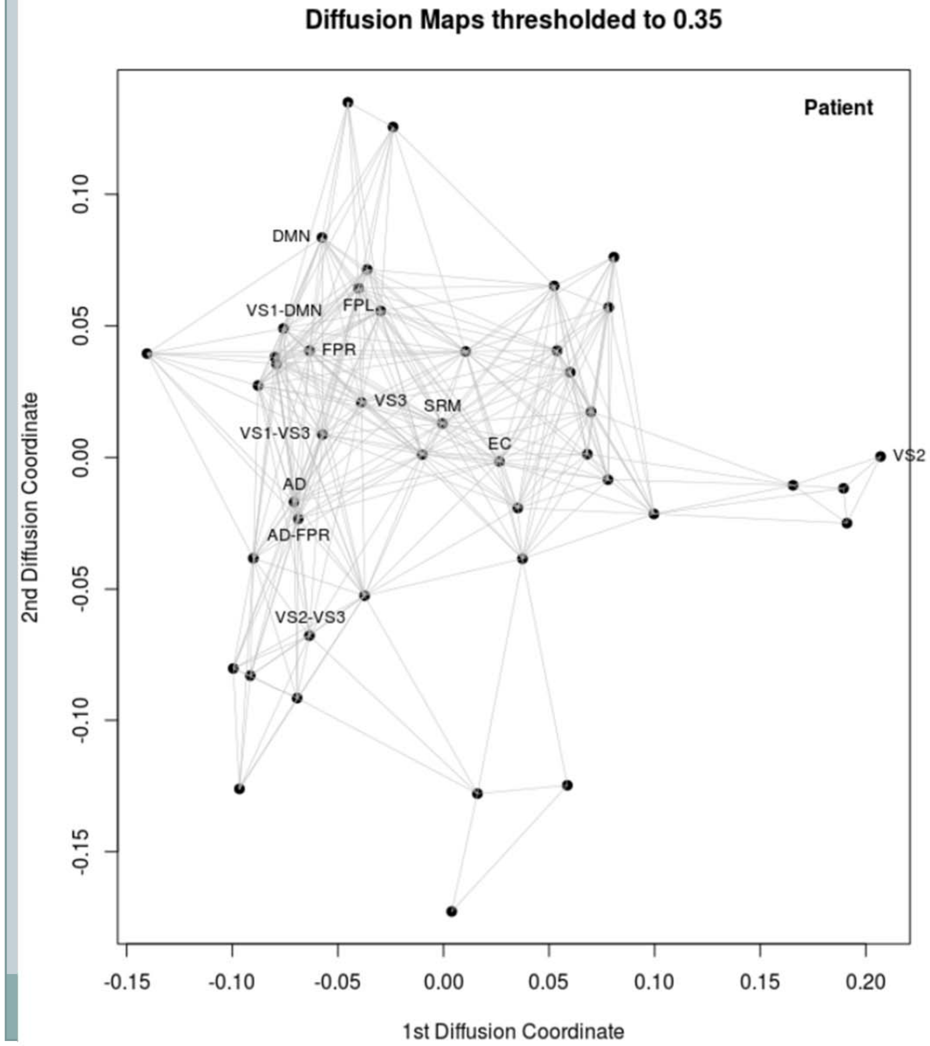


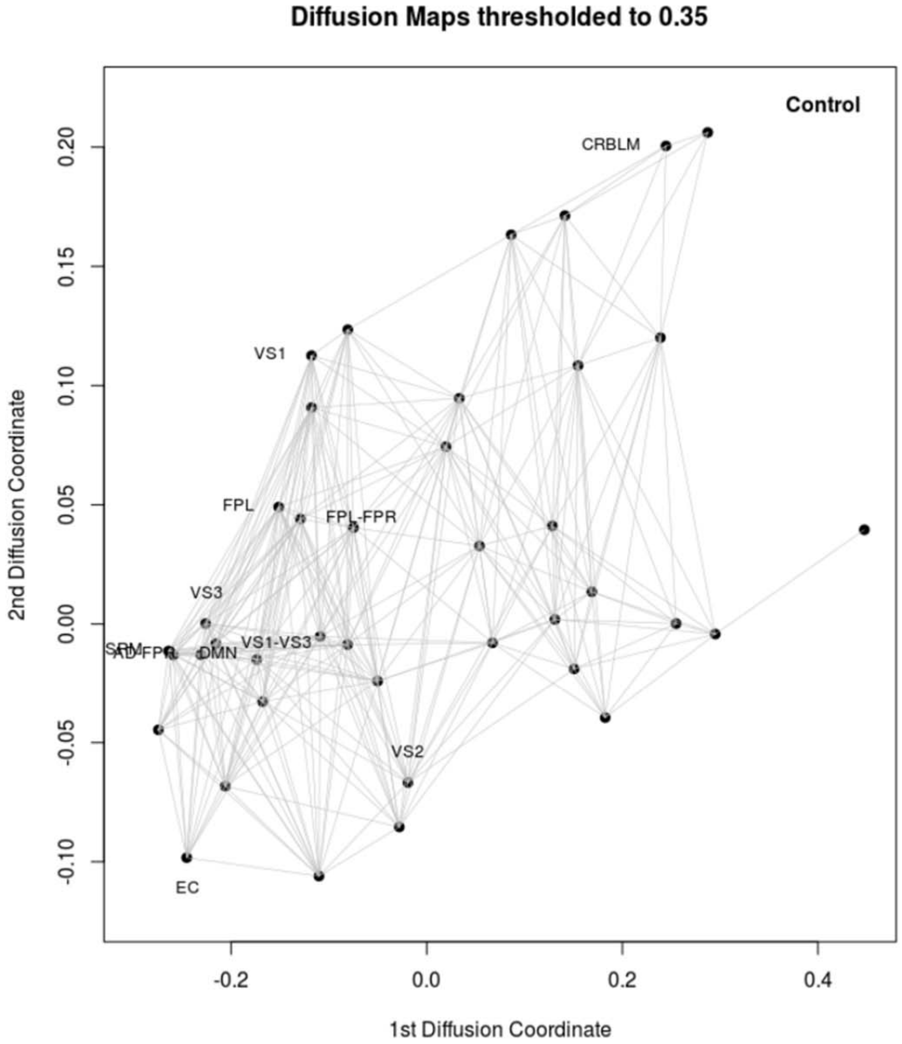


Global Network Metrics: Diffusion Maps-Heat-Euclidian

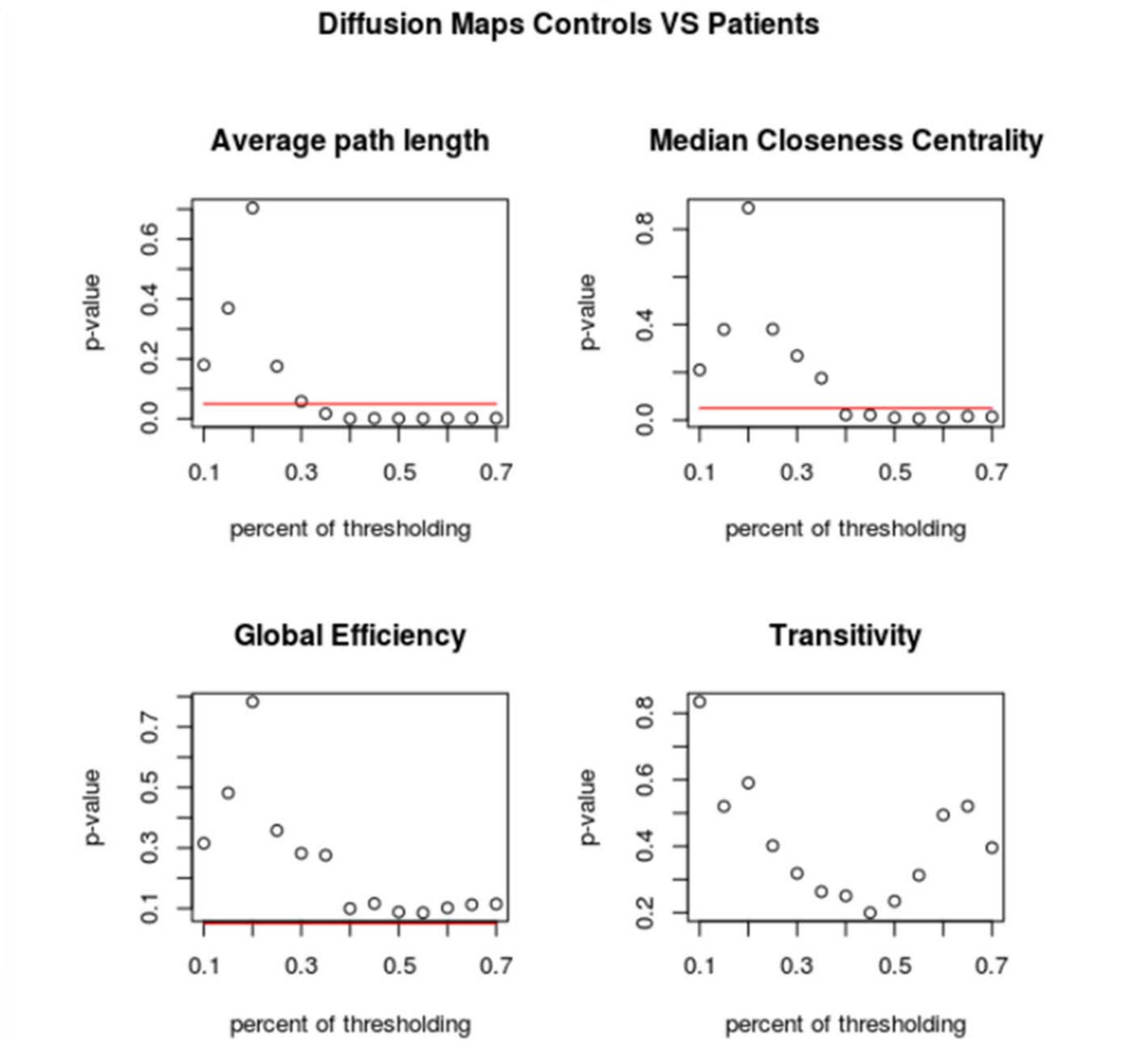


Embedded Networks: Diffusion Maps-Heat-Euclidian

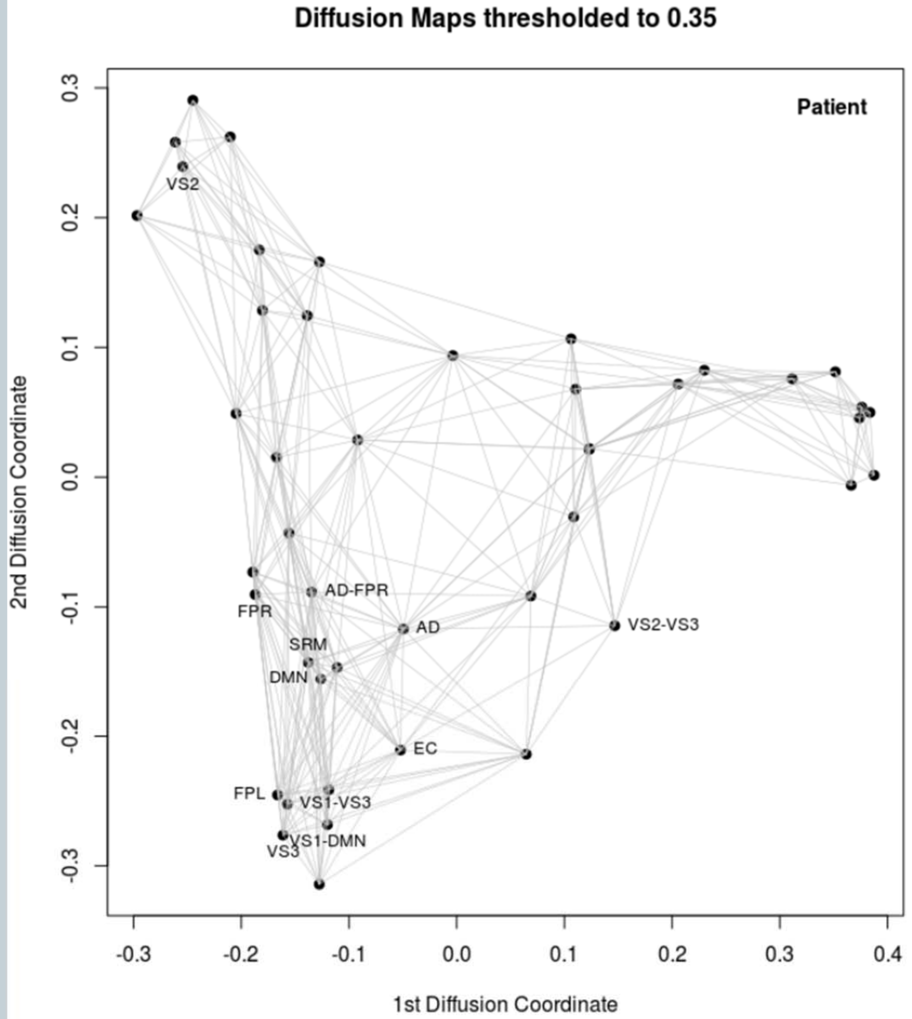


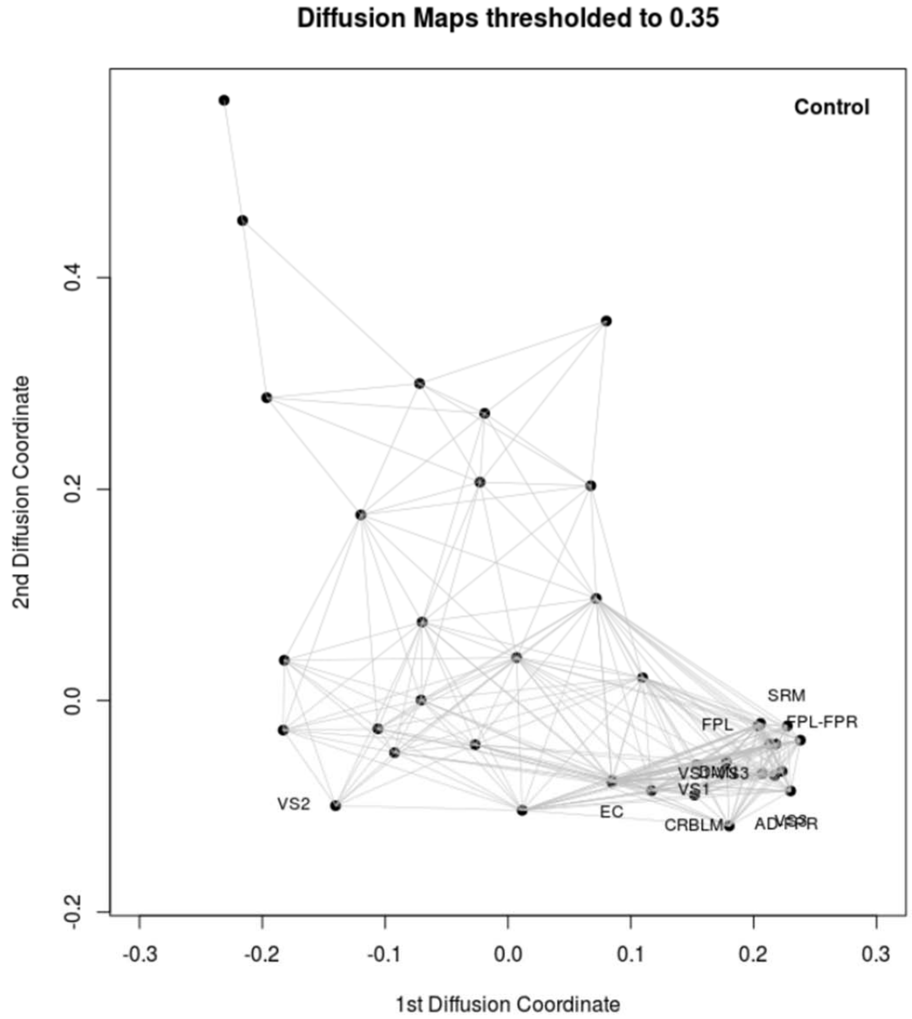


Global Network Metrics: Diffusion Maps-Gaussian-CCorrelation Based



Embedded Networks:
Diffusion Maps-Heat-Correlation Based





Conclusions

- A data-driven methodology based on nonlinear manifold learning techniques to construct low-dimensional embedded functional connectivity networks for fMRI data
- ICA, PCA, MDS, ISOMAP and Diffusion MAPS were employed towards this direction
- The approach allowed the identification of significant differences between groups (controls, and patients treatment).
- Cross Correlation, PCA, MDS
- ISOMAP seems sensitive to metrics. Fails with the Euclidian.
- Diffusion Maps seems to be slightly more robust w.r.t thresholding

Further work

- Task-Based fMRI Experiments– Nonlinear Dynamics!
- Compare and implement non-convex Machine Learning techniques such as SOM

Thank you for your attention