Network dissimilarity measure and application to brain network differentiation

Cristina Masoller
Universitat Politecnica de Catalunya, Barcelona

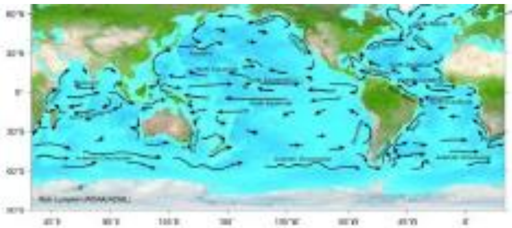
T. A. Schieber, L. Carpi, M. G. Ravetti (Bello Horizonte), A. Diaz-Guilera (Barcelona), P. Pardalos (Florida)

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Campus d'Excel·lència Internacional

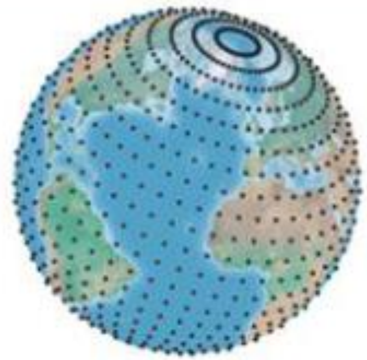
BIFI 2018: **"Complexity, Networks & Collective Behavior"**
Zaragoza, February 2018

ICREA
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

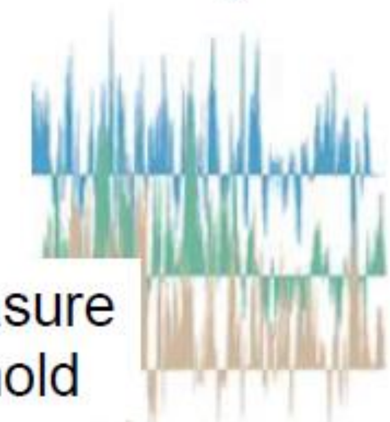
Motivation: how can we compare climate networks?



Back to the climate system: interpretation (currents, winds, etc.)



More than 10000 nodes.



Daily resolution: more than 13000 data points in each TS

Sim. measure + threshold

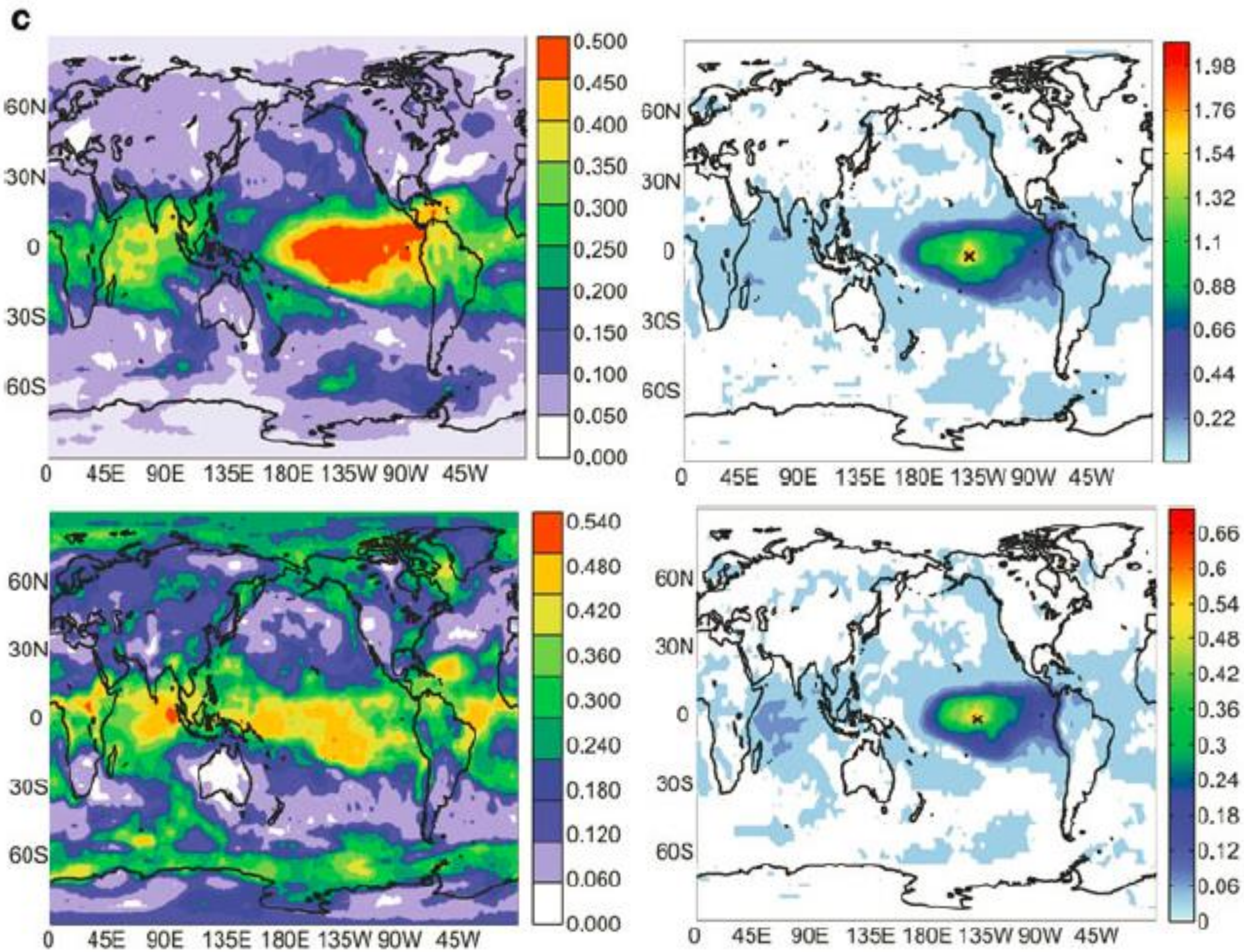
Surface Air Temperature <u>Anomalies</u> (solar cycle removed)

Donges et al, Chaos 2015

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

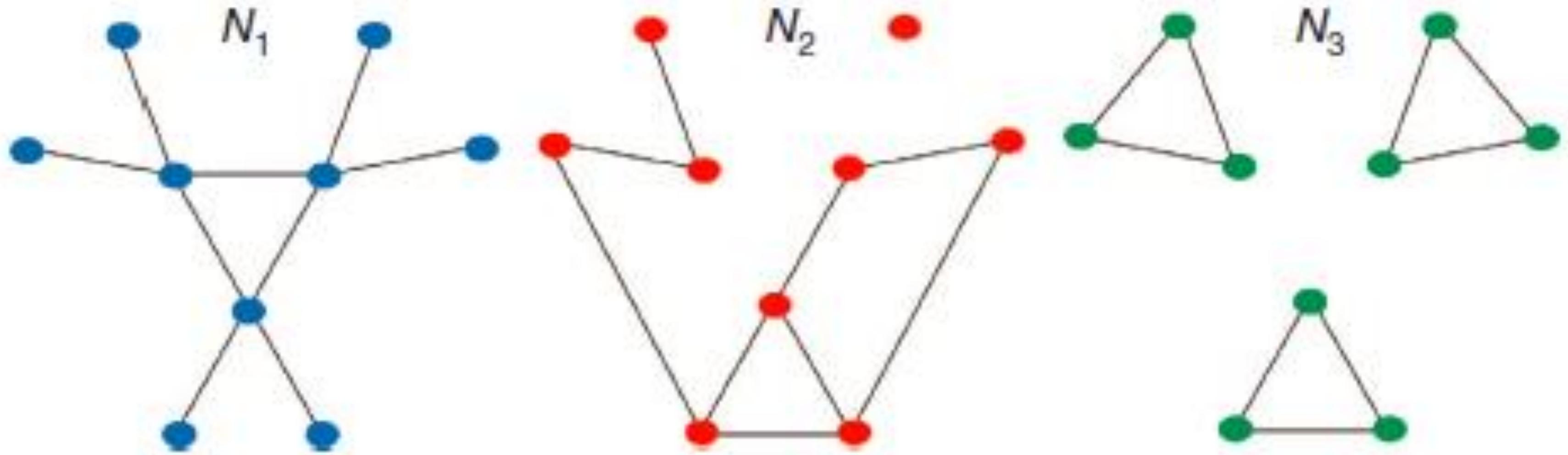
Motivation: how can we compare climate networks?



Main Goal: to develop a measure that allows a precise comparison of complex networks (including different sizes)

A.A. Tsonis (ed.), *Advances in Nonlinear Geosciences*,
DOI 10.1007/978-3-319-58895-7_4

Same number of nodes and links



How to measure distances between networks?

Complex network measures

- Degree distribution, closeness centrality, betweenness centrality, average path length, etc.
- Provide *partial* information.
- How to define a measure that contains detailed information about the *global topology* of a network, in a *compact* way?

$\Rightarrow$ Node Distance Distributions (NDDs)

- $\boldsymbol{P_i(j)}$ of node $\boldsymbol{i}$ = fraction of nodes connected to $\boldsymbol{i}$ at distance $\boldsymbol{j}$ (shortest path)
- N nodes: have a vector of N pdfs $\{P_1, P_2, \ldots, P_N\}$

- If two networks have the same set of node distance distributions $\Rightarrow$ they have the same diameter, average path length, etc.

- How to condense the information contained in the N node-distance distributions?

Network Node Dispersion (NND)

- Measures the heterogeneity of the N pdfs $\{P_1, P_2, \ldots, P_N\}$

$$\mathrm{NND}(G) = \frac{\mathcal{J}(\mathbf{P}_1, \ldots, \mathbf{P}_\mathrm{N})}{\log(d+1)}$$

d = diameter

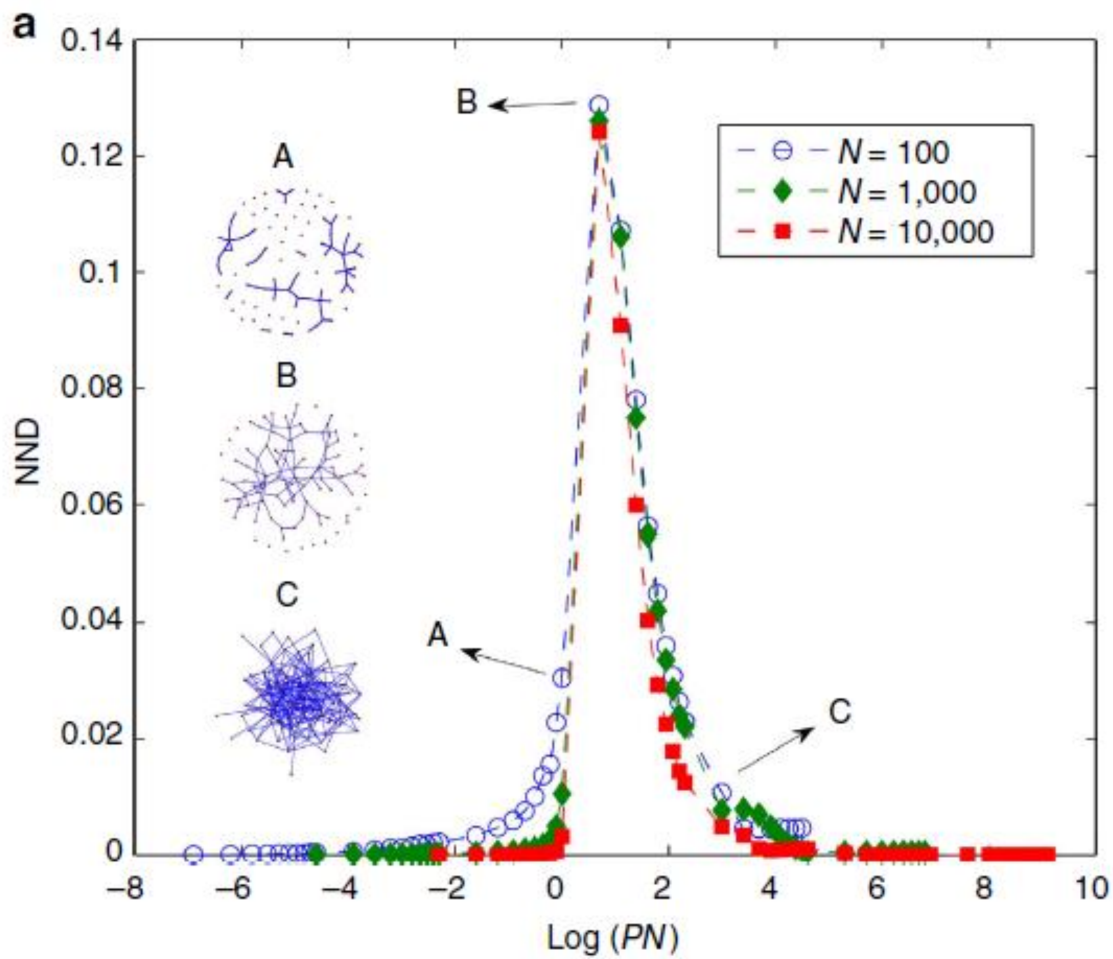
$$\mathcal{J}(\mathbf{P}_1, \ldots, \mathbf{P}_\mathrm{N}) = \frac{1}{N} \sum_{\mathrm{i,j}} p_\mathrm{i}(j) \log\left(\frac{p_\mathrm{i}(j)}{\mu_\mathrm{j}}\right)$$

$$\mu_\mathrm{j} = \left(\sum_{\mathrm{i}=1}^{N} p_\mathrm{i}(j)\right) / N$$

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Example of application: percolation transition in a random network



⇒ NND detects the percolation transition

P=connection probability

T. A. Schieber, L. Carpi, A. Diaz-Guilera, P. M. Pardalos, C. Masoller and M. G. Ravetti, Nat. Comm. 8:13928 (2017).

Dissimilarity between two networks

$$D(G,G') = w_1 \sqrt{\frac{\mathcal{J}(\mu_G, \mu_{G'})}{\log 2}} + w_2 \left| \sqrt{\mathrm{NND}(G)} - \sqrt{\mathrm{NND}(G')} \right|$$

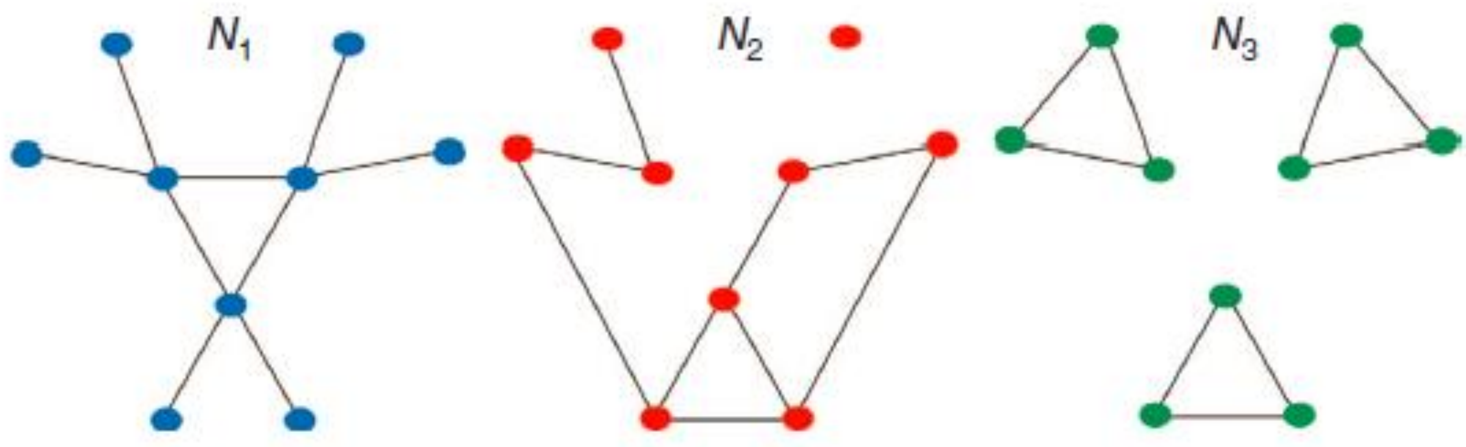
$w_1 = w_2 = 0.5$

compares the averaged connectivity

compares the heterogeneity of the connectivity distances

- Extensive numerical experiments demonstrate that isomorphic graphs return ***D=0***
- Can be applied to networks of different sizes
- *Computationally efficient*.

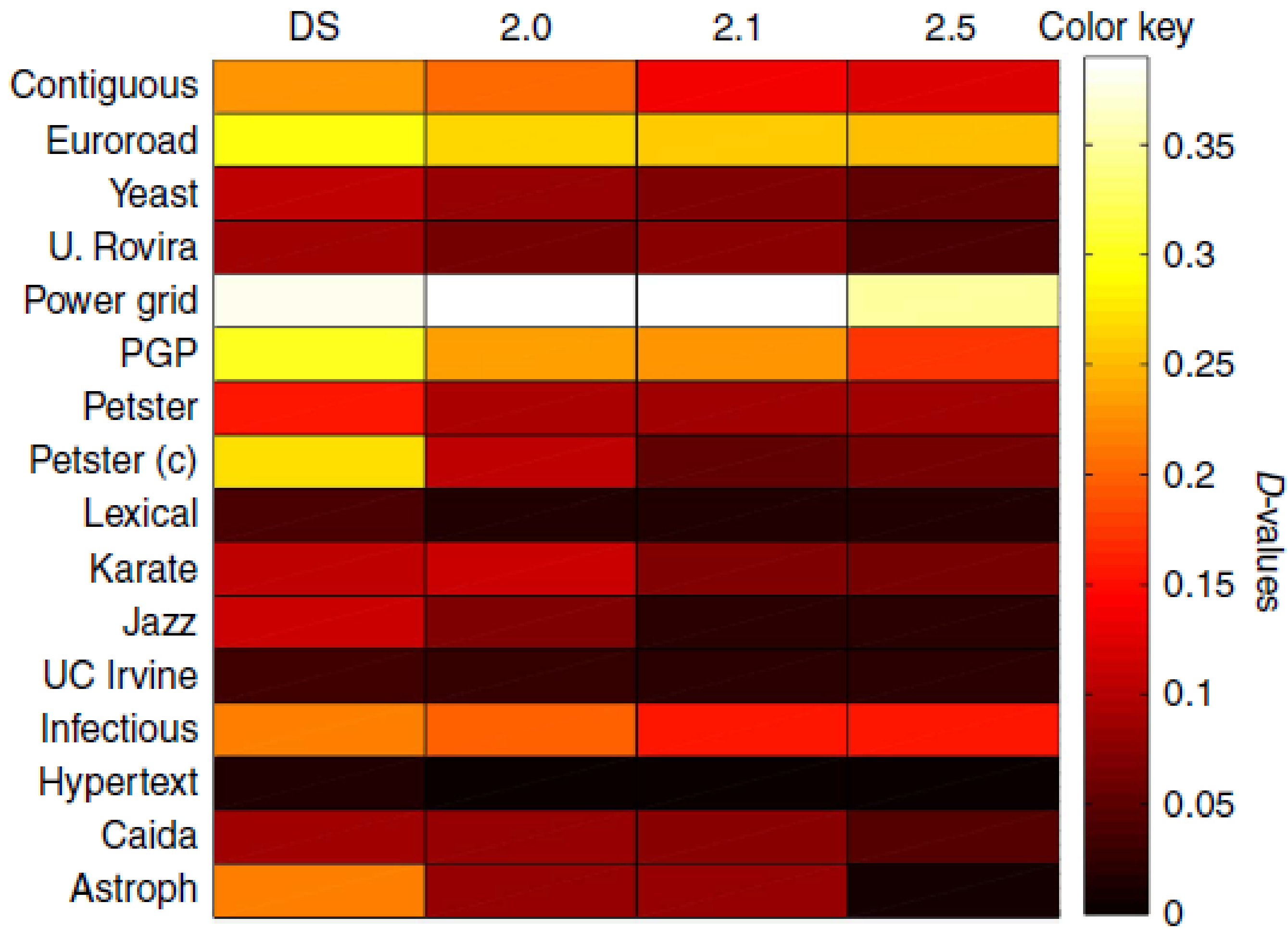
Comparing three networks with the same number of nodes and links



	D	Hamming	Graph Edit Distance
N_1,N_2	0.25	12	6
N_1,N_3	0.56	12	6
N_2,N_3	0.47	12	6

Comparing real networks to null models

DS preserves the degree sequence; 2.0 also preserves the degree correlation; 2.1 also the clustering coefficient; 2.5 also the clustering spectrum



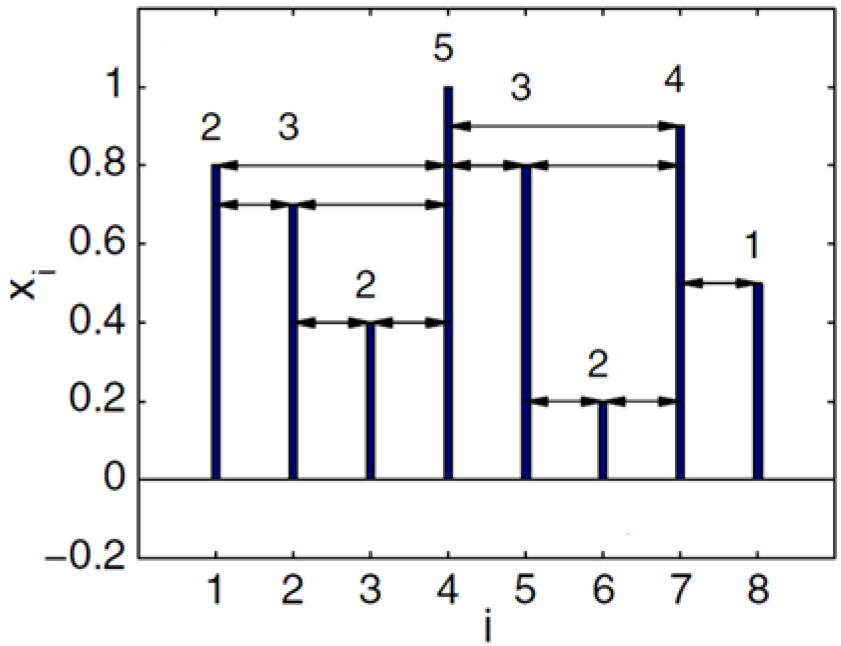
Each model is run 30 times, ⟨D⟩ is plotted in color code

dk model: *Orsini, C. et al. Nat. Commun. 6, 8627 (2015)*

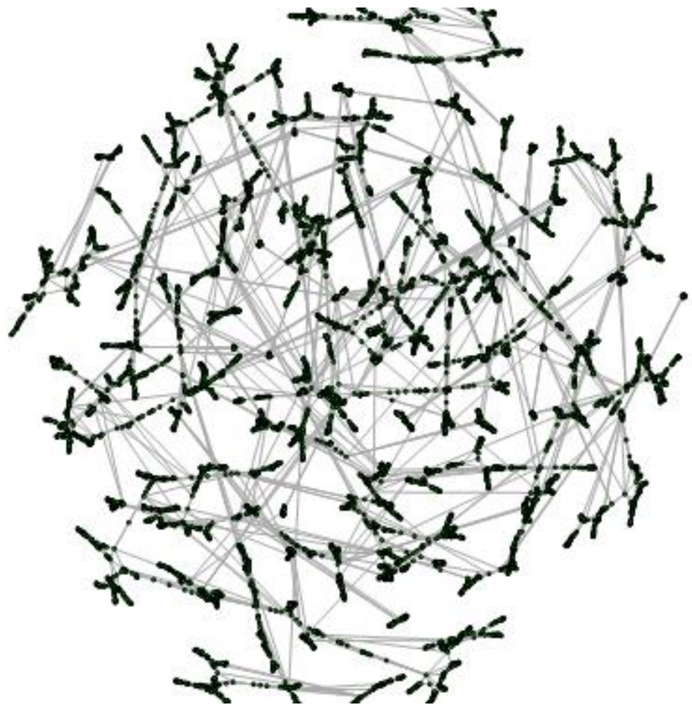
- Synthetic model for Power Grid Network?

Horizontal Visibility Graph: graph representation of a time series

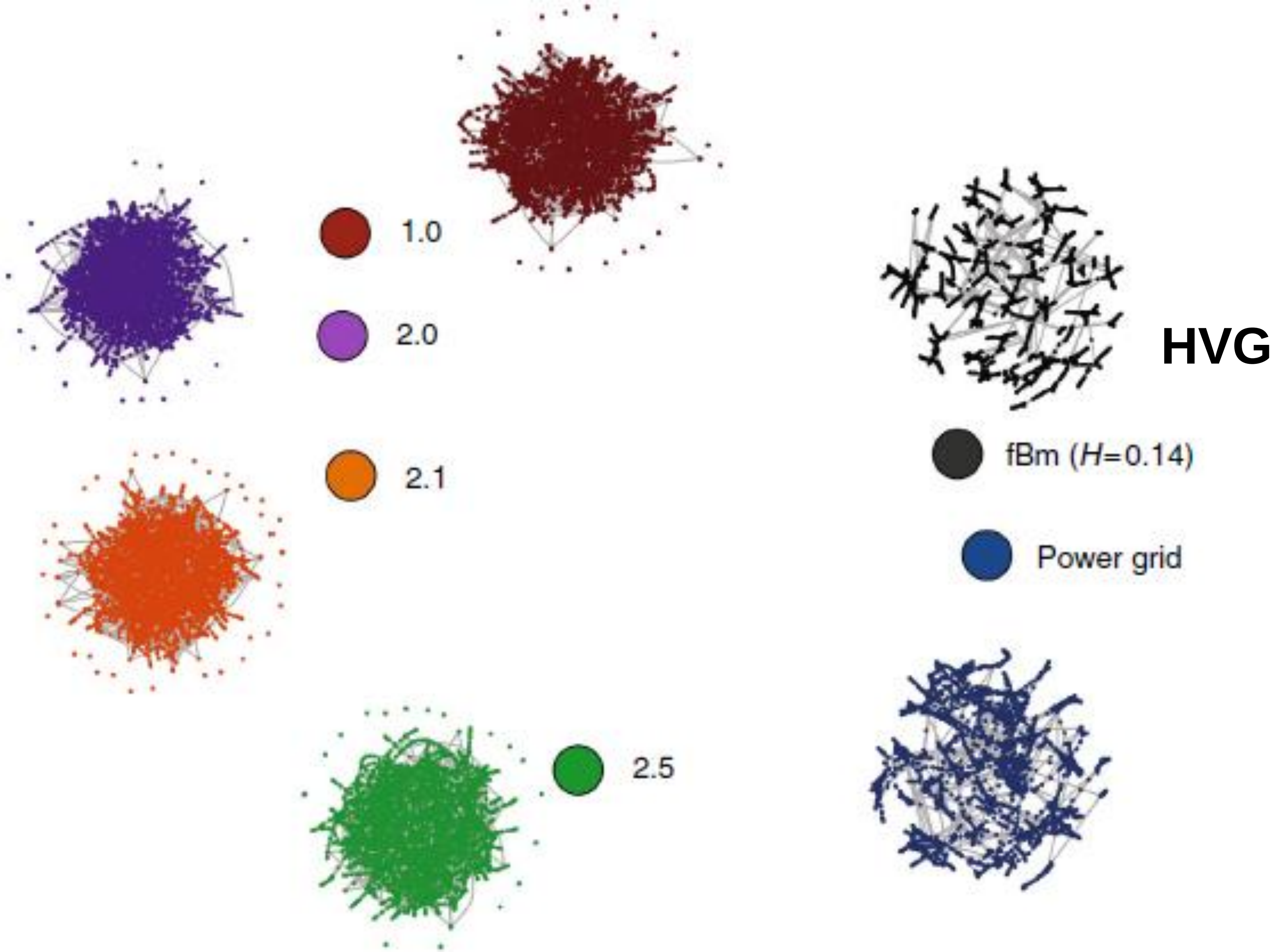
Synthetic time series: fractional Brownian Motion (fBm) with controllable Hurst exponent



⇒



HVG method: Luque et al, Phys. Rev. E 80, 046103 (2009).



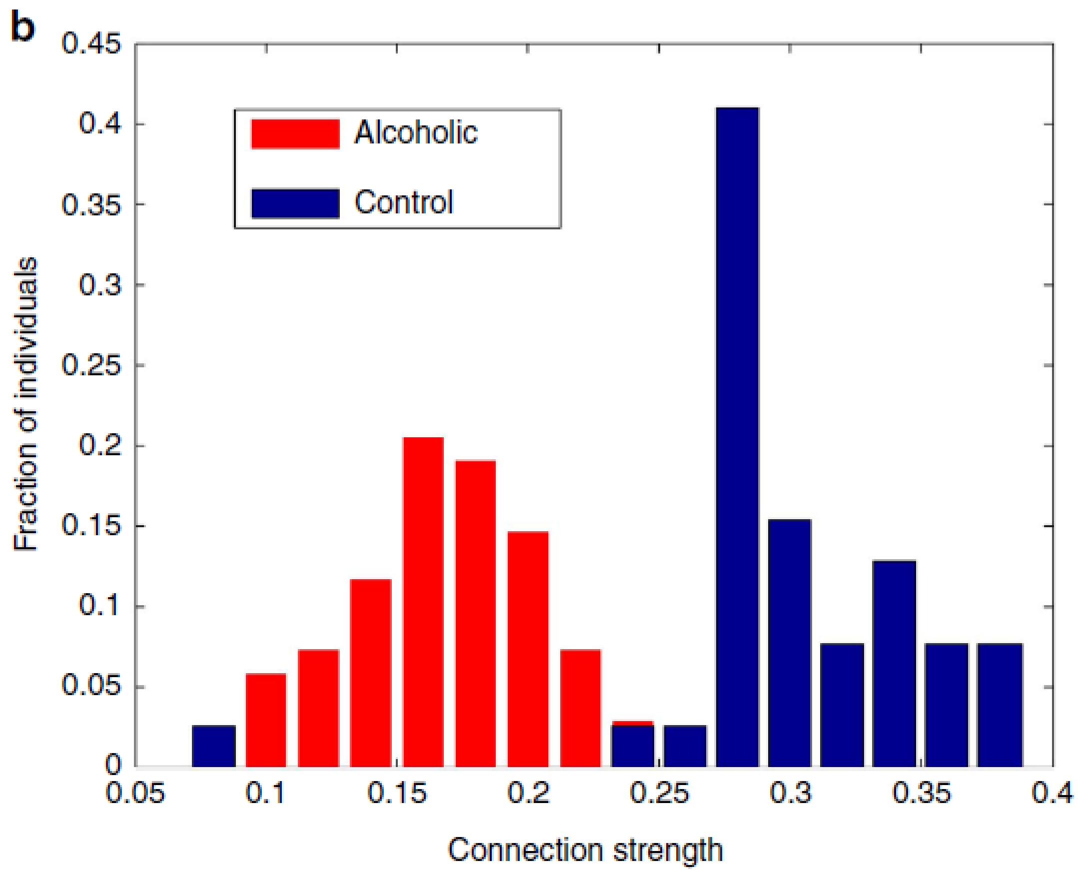
T. A. Schieber et al, Nat. Comm. 2017

Comparing brain networks

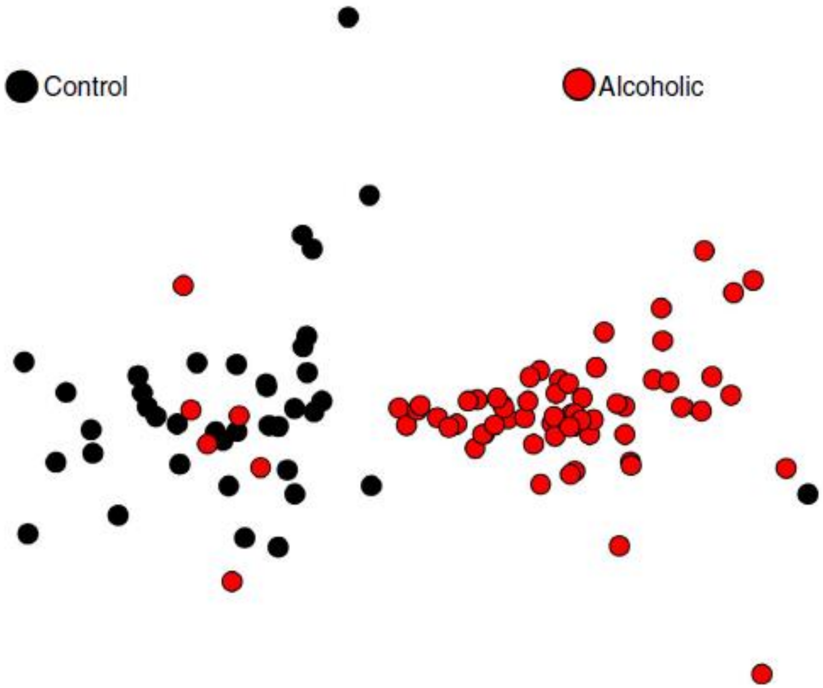
- EEG data
  - https://archive.ics.uci.edu/ml/datasets/eeg+database
  - 64 electrodes placed on the subject's scalp sampled at 256 Hz during 1s
  - 107 subjects: 39 control and 68 alcoholic
- Use HVG to transform each EEG TS into a network G.
- Weight between two brain regions: **1-*D*(G,G')**
- The resulting network represents the weighted similarity between the brain regions of an individual.

  $\Rightarrow$ We can compare the different individuals.

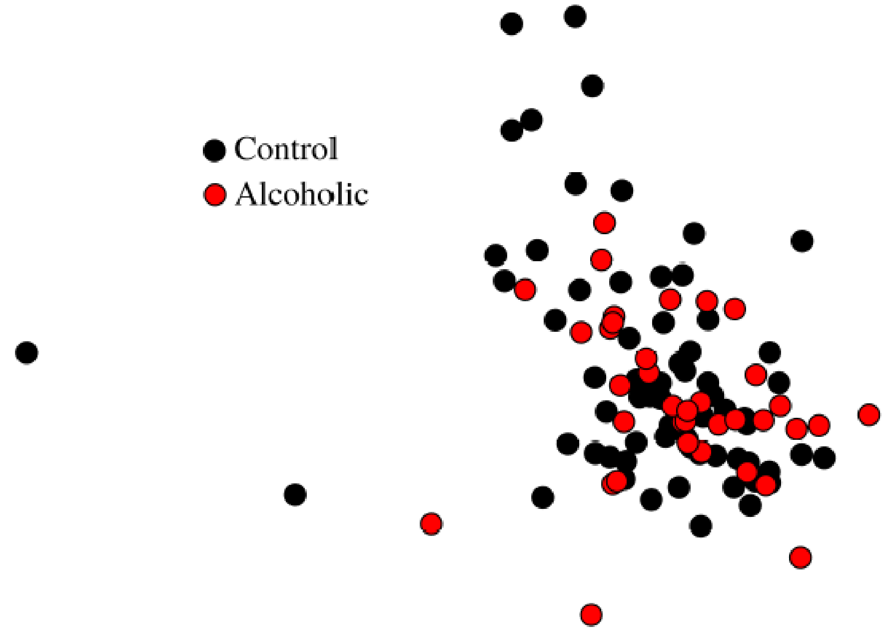
We identified two regions of the brain (called 'nd' and 'y'), where the weight of the connections between these regions is higher in control than in alcoholic networks.



Dissimilarity measure



Hamming distance



T. A. Schieber et al, Nat. Comm. 2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Summary

- New measure to quantify the heterogeneity of the connectivity paths of a single network.
  - detects the percolation transition in a random network.
- New measure to calculate the distance between two networks
  - Can be applied to networks of different sizes.
  - Returns $D=0$ only if the two networks are isomorphic.
- Many possible applications: characterizing time-evolving climate networks, classification of networks generated from biomedical data, etc.

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

<cristina.masoller@upc.edu>

Universitat Politecnica de Catalunya

http://www.fisica.edu.uy/~cris/

THANK YOU FOR YOUR ATTENTION !

T. A. Schieber et al, "*Quantification of network structural dissimilarities*", Nat. Comm. 8:13928 (2017).