Quantification of network dissimilarities and application to modeling the Power Grid network

Cristina Masoller

Universitat Politecnica de Catalunya, Barcelona

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

Cristina.masoller@upc.edu

www.fisica.edu.uy/~cris

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

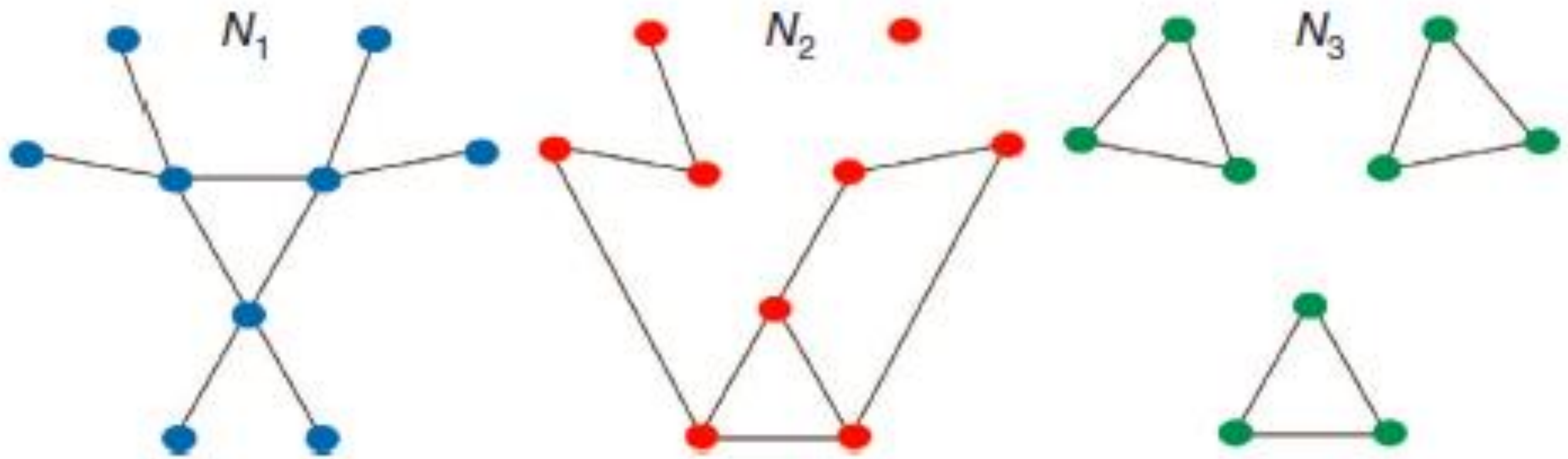
Campus d'Excel·lència Internacional

Complex Networks 2017
Lyon, France, November 2017

ICREA

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA E INNOVACIÓN

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.

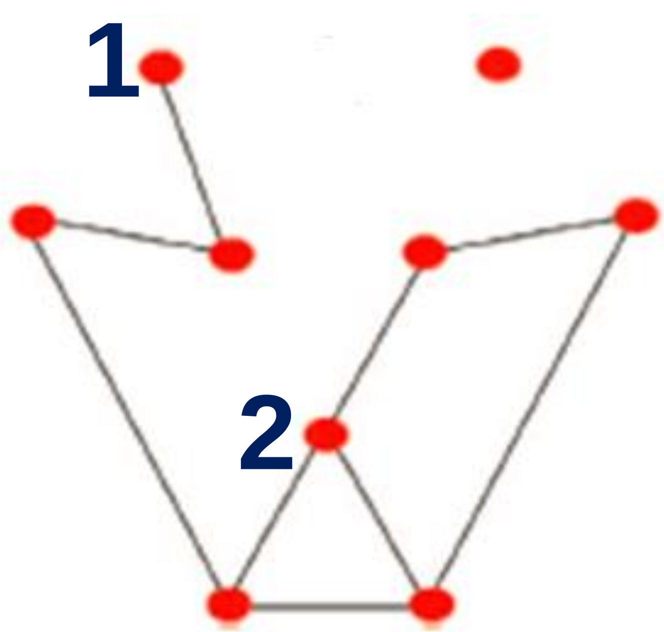
- Provide *partial* information.

- 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)

- $p_i(j)$ of node i = fraction of nodes connected to i at distance j

Same number of nodes and links



Node 1:

j	# nodes at distance j
1	1
2	1
3	1
4	2
5	2
∞	1

Node 2:

j	# nodes at distance j
1	3
2	2
3	1
4	1
∞	1

- With *N* nodes, the Node Distance Distributions is a vector of N pdfs $\{p_1, p_2, \ldots, p_N\}$
- If two networks have the same set of NDDs $\Rightarrow$ they have the same diameter, average path length, etc.

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

- The *Network Node Dispersion (NND*) measures the heterogeneity of the N pdfs $\{p_1, p_2, \ldots, p_N\}$
- Quantifies the heterogeneity of connectivity distances.

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

d = diameter

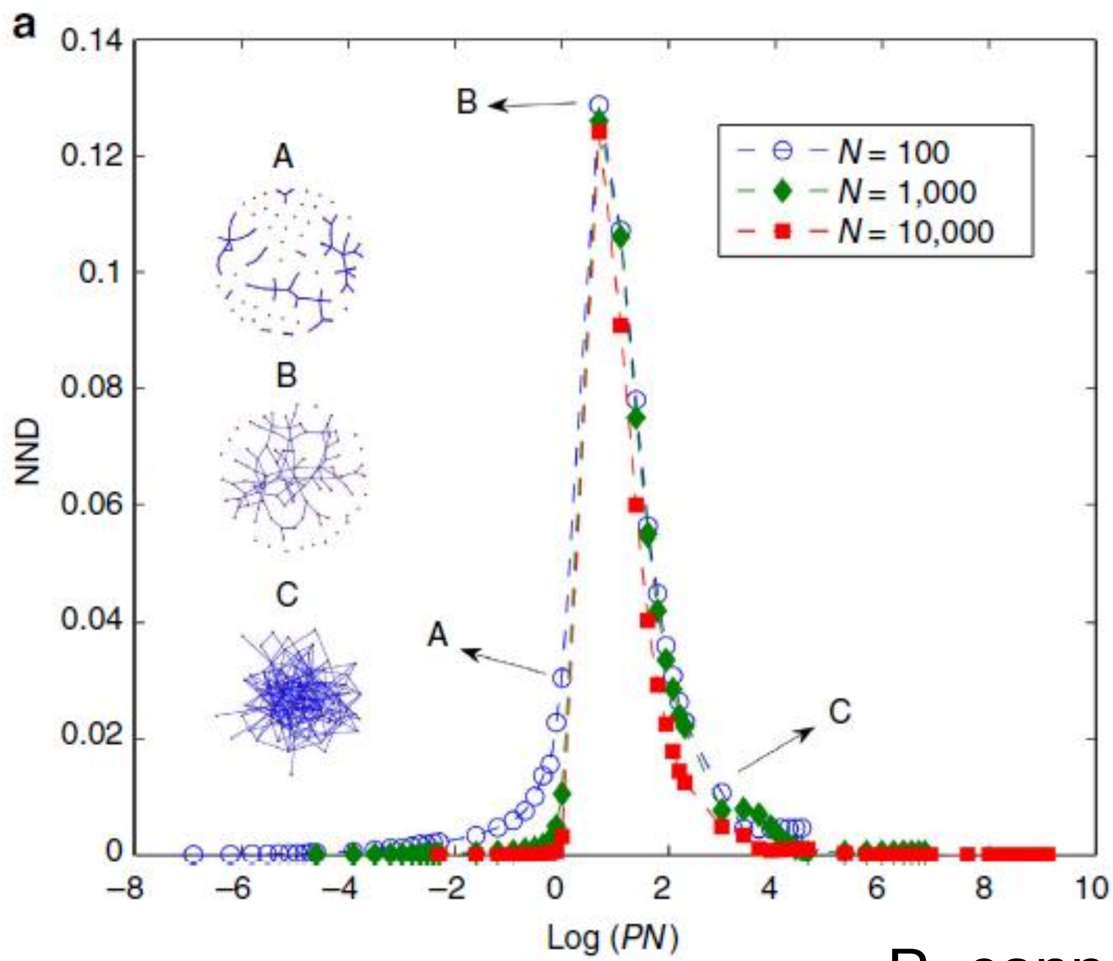
$$\mathcal{J}(\mathbf{P}_1, \ldots, \mathbf{P}_\mathrm{N}) = \frac{1}{N}\sum_{\mathrm{i,j}} p_\mathrm{i}(j)\log\left(\frac{p_\mathrm{i}(j)}{\mu_\mathrm{j}}\right)$$

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

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Example of application: percolation transition in a random network



⇒ the Network Node Dispersion 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

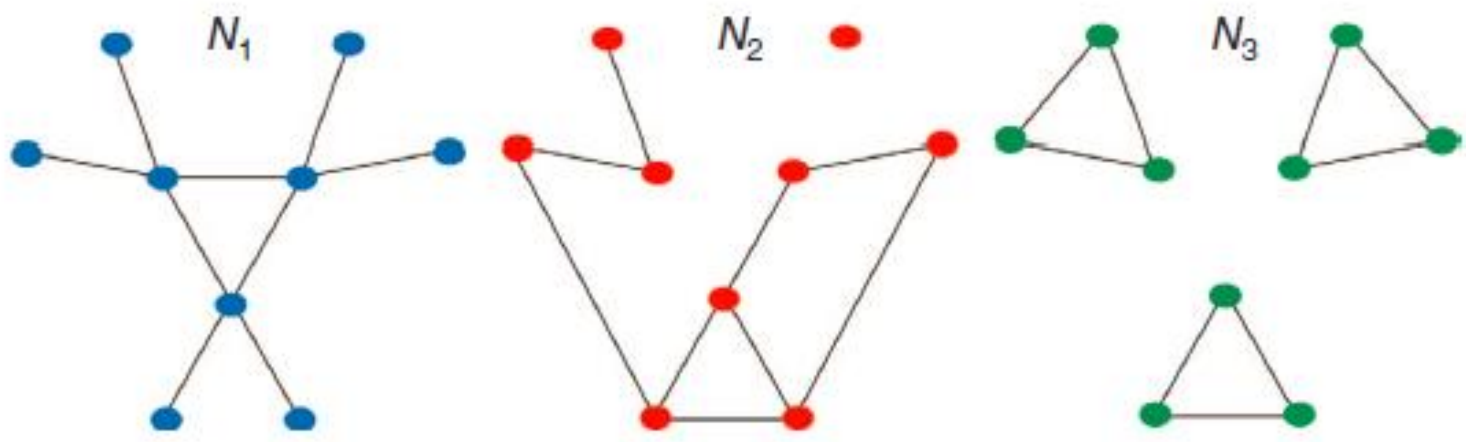
compares the heterogeneity of the connectivity distances

- Extensive numerical experiments demonstrate that isomorphic graphs return ***D=0***
- Can be applied to networks of different sizes.
- *D* time complexity is polynomial because it relies on the computation of all shortest paths length which is a polynomial problem.

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

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

Percolation on real networks

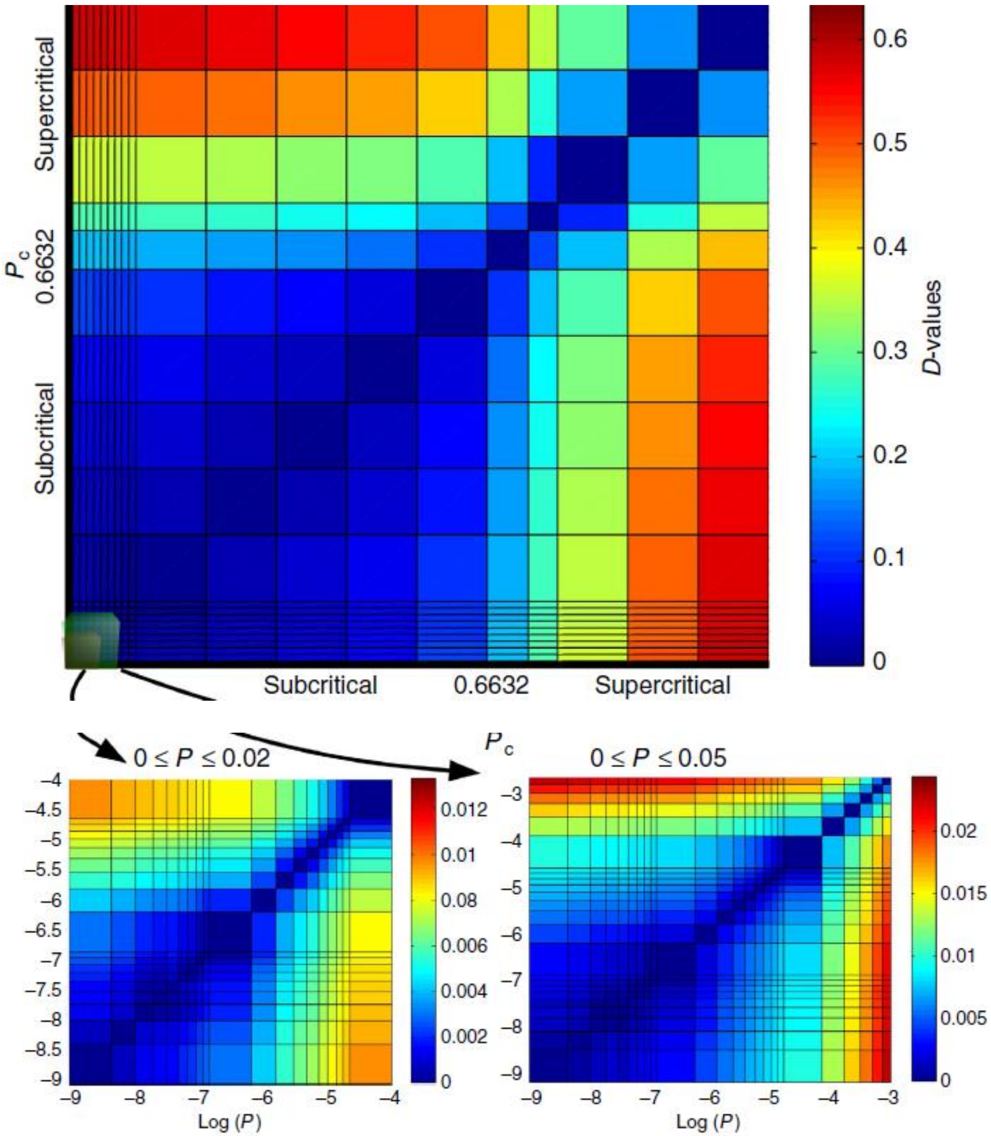
- Main idea: when looking for the percolation transition, two graphs in the same phase (subcritical or supercritical) present smaller D-values than a pair of graphs in different phases.

- Start with two probabilities, β and α, on the supercritical and subcritical phases. $Pm=(\alpha + \beta)/2$

- If $D(Gm, G\alpha)>D(Gm, G\beta)$ then β =Pm else α =Pm

- Stop when $|\beta - \alpha| <$ precision ε

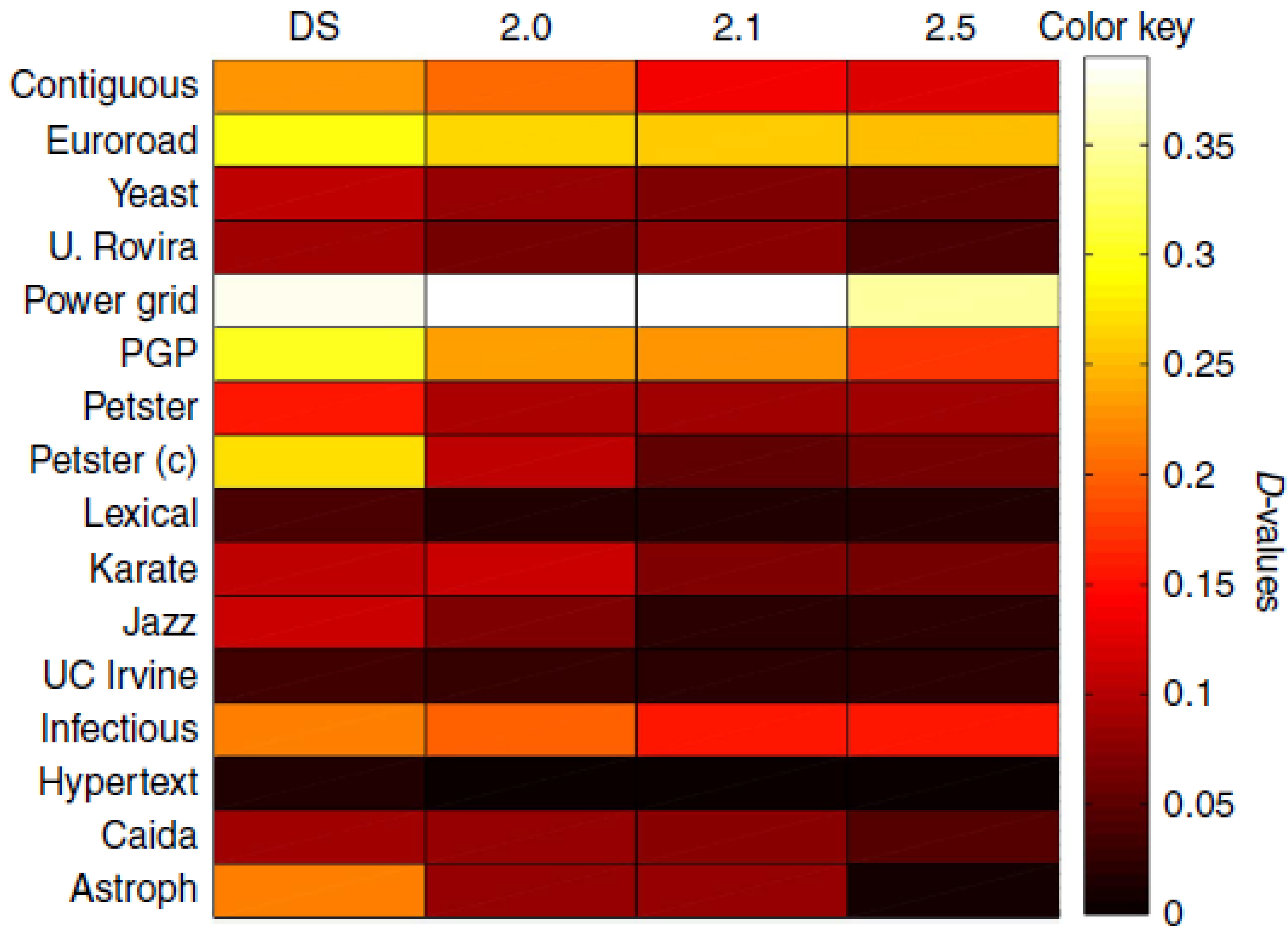
Percolation on the Power Grid network



	Pc	Comp. time(s)
D	0.6826	5,500
MC	0.6632	35,000

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



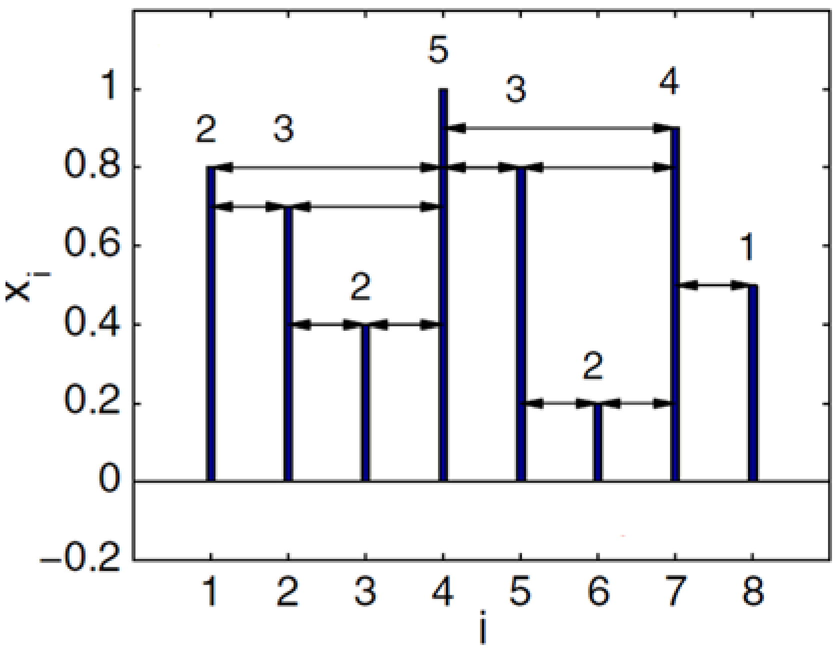
Each model is run 30 times, ⟨D⟩ is plotted in color code

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

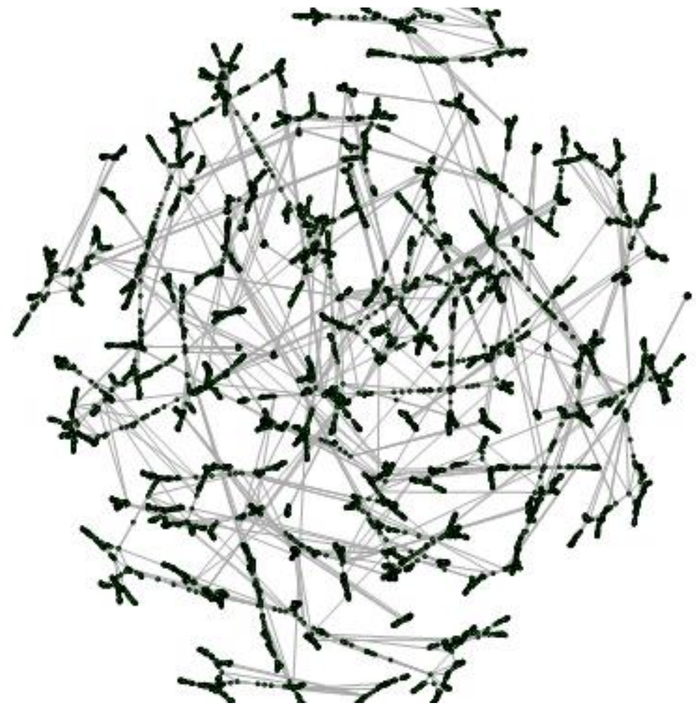
- Synthetic model for Power Grid Network?

Horizontal Visibility Graph: graph representation of a time series

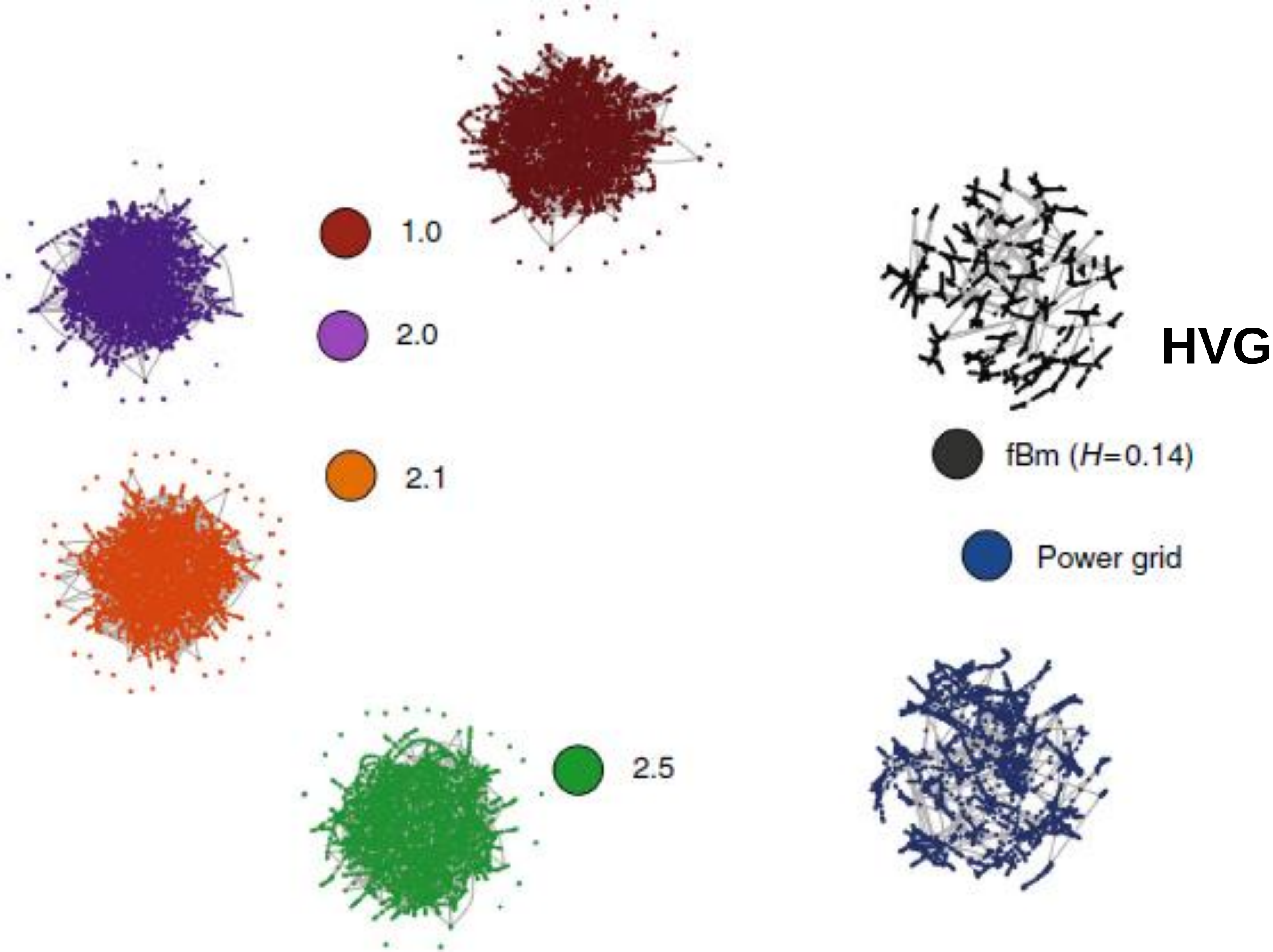
Synthetic time series: fractional Brownian Motion (fBm) with controllable Hurst exponent



⇒



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



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

Summary

- New measure to quantify the heterogeneity of the connectivity paths of a single network.
  - It 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 networks, classification of biological networks, etc.

THANK YOU FOR YOUR ATTENTION !

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

Advertising

- School on "*Nonlinear Time Series Analysis and Complex Networks in the Big Data Era*", co-organized with Jesus Gomez-Gardenes and Hilda Cerdeira

  ICTP-SAIFR (Sao Paulo): February 19 – March 2, 2018

- Workshop on "*Predicting transitions in complex systems*", co-organized with K. Lehnertz and J. Hlinka

  Max Planck Institute for Physics of Complex Systems (Dresden): 23 – 27 April 2018