The effect of nonlinearity in computing graph theoretic characteristics of complex networks

D. Hartman J. Hlinka M. Vejmelka M. Paluš

Institute of Computer Science, Academy of Sciences, Prague, Czech Republic

This study was supported by the Czech Science Foundation project No. P103/11/J068.
Special thanks belong to D. Mantini and M. Corbetta

Effect of nonlinearity

Complex systems can be analyzed as networks of mutually interacting subsystems [1]. The critical step in constructing such graphs is the choice of a measure of dependence. Classical choice is linear (Pearson) correlation coefficient. However nonlinear approaches uncovered network phenomena not detectable using linear measures, e.g. in MEG brain studies [2] or in climate systems [3]. Presented results show how the possible nonlinearity [5] in data can influence network analysis in fMRI study [4]. Furthermore some aspects of studying nonlinearities in climate datasets are discussed.

Functional connectivity using mutual information (MI)

- For discrete random variables X_1, X_2 with values $\mathcal{X}_1$ and $\mathcal{X}_2$, the MI is:

$$I(X_1, X_2) = \sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{p(x_1)p(x_2)}$$

- Estimates: box-counting algorithm based on the marginal equiquantization method with $Q = 8$ [6].

Considered network characteristics (unweighted)

- Clustering coefficient

$$c_i = \frac{2|E(\Gamma(i))|}{k_i(k_i-1)} = \frac{\sum_{j,\ell} a_{i,\ell} a_{j,\ell} a_{\ell,i}}{k_i(k_i-1)}$$

- betweenness (BC)

$$C_b(i) = \sum_{j,k \in V, j \neq k} \frac{\sigma_{j,k}(i)}{\sigma_{j,k}}$$

- Efficiency

$$E = \frac{1}{n(n-1)} \sum_{i,j \in V, i \neq j} \frac{1}{d_{i,j}}$$

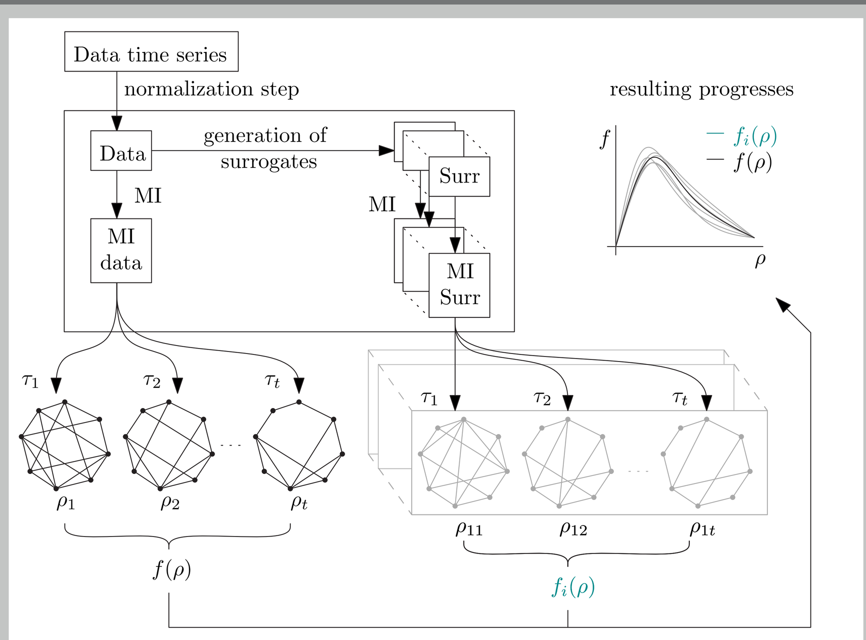
- Assortative coefficient

$$r = \sum_{(i,j)\in E} k_i k_j - \frac{1}{m}\left[\sum_{(i,j)\in E} \frac{1}{2}(k_i + k_j)\right]^2 \Big/ \sum_{(i,j)\in E} \frac{1}{2}(k_i^2 + k_j^2) - \frac{1}{m}\left[\sum_{(i,j)\in E} \frac{1}{2}(k_i + k_j)\right]^2$$

Neurological data

Resting-state fMRI data, 12 healthy volunteers (2 sessions each), 3T Philips Achieva MRI scanner operating at ITAB (Chieti, Italy). After standard preprocessing 90 ROI (by anatomical atlas AAL), 300 timepoints forming 10-min session.

Process description



- Normalization step
  Correct univariate non-Gaussianity
- Data branch
  - Estimate MI for all pairs
  - undir graphs via sequence of thresholds
- Linear surrogate branch
  - Compute muvar FT surrogates
  - Estimate MI for all pairs
  - undir graphs via sequence of thresholds

Systematically compare network characteristics for graph from data and surrogates.

Dominances

Let s denotes session, ρ density, j index of surrogate and i index of vertex. Then for general characteristic f there are $f_j(s, \rho, i)$ for surrogates and $f^D(s, \rho, i)$ for data.

- Vertex i is max (min) dominant if
  $\max_j\{f_j(i)\} < f^D(i)$, $\min_j\{f_j(i)\} < f^D(i)$
- Max dominance indicator function

$$f_M^{dom}(s, \rho, i) = \begin{cases} 1, & i \text{ is max dominant} \\ 0, & \text{otherwise.} \end{cases}$$

Similarly for min dominance indicator function f_m^{dom}

- Dominance function

$$f^{dom}(s, \rho, i) = f_m^{dom}(s, \rho, i) + f_M^{dom}(s, \rho, i)$$

- Graph dominance function

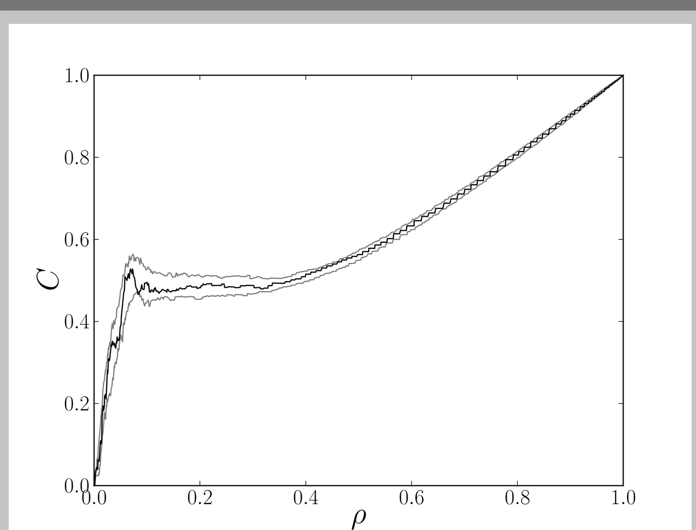
$$f^{G,dom}(s, \rho) = \sum_{i=1}^{n} f^{dom}(s, \rho, i),$$

- Overall dominance function

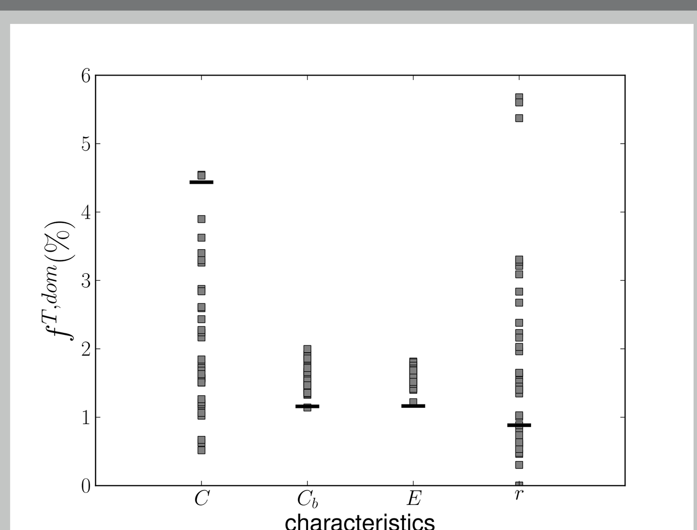
$$f^{T,dom}(\rho) = \frac{1}{N_{sess}} \sum_{s=1}^{N_{sess}} f^{G,dom}(s, \rho).$$

- For global characteristics
  Indices of vertices are excluded

Global characteristics results

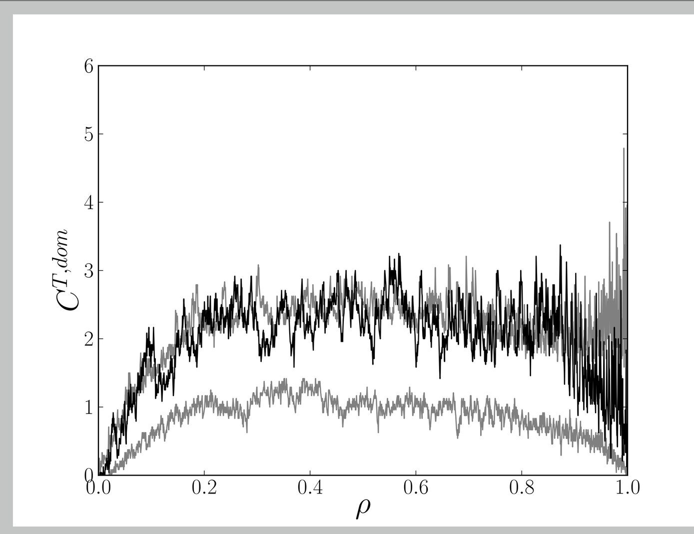


Clustering coefficient for data (black lines) and surrogates (gray lines)

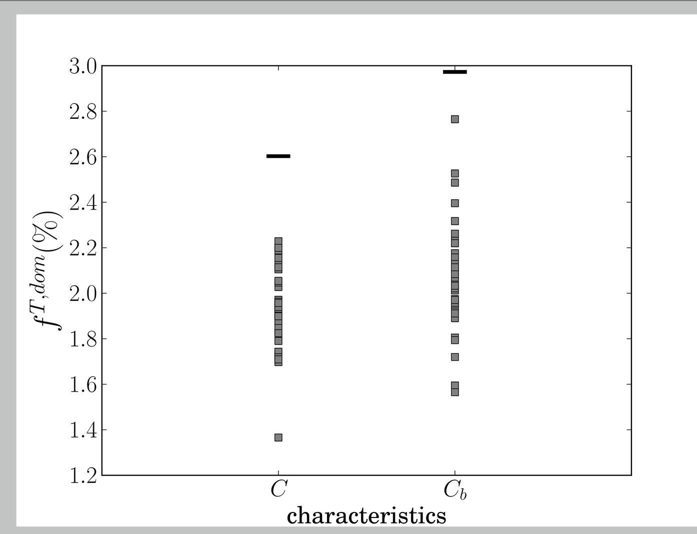


Overall dominances of characteristics for data (black dots) and surrogates (gray dots)

Local characteristics results



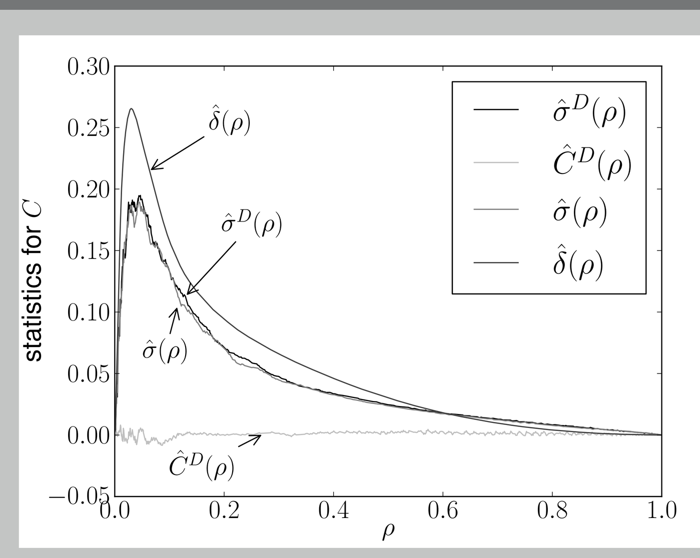
Overall dominances for average clustering coefficient for data (black lines) and surrogates (gray lines)



Overall dominances of characteristics for data (black dots) and surrogates (gray dots)

Comparing nonlinear effect with intra- and inter-session variability

- For clustering coefficient C
  - $\hat{C}^D$ is the average difference between data and average surrogate over all sessions with std. deviation $\hat{\sigma}^D$ (nonlinear effect)
  - $\hat{\delta}$ is the inter-session standard deviation of the session-averaged surrogate values, $\hat{\sigma}$ is the intra-session standard deviation of the surrogate values within a session



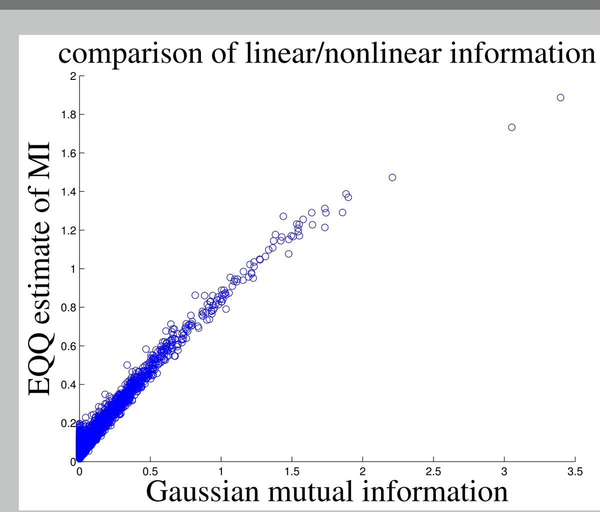
Comparing variabilities for clustering coefficient

Climatic data

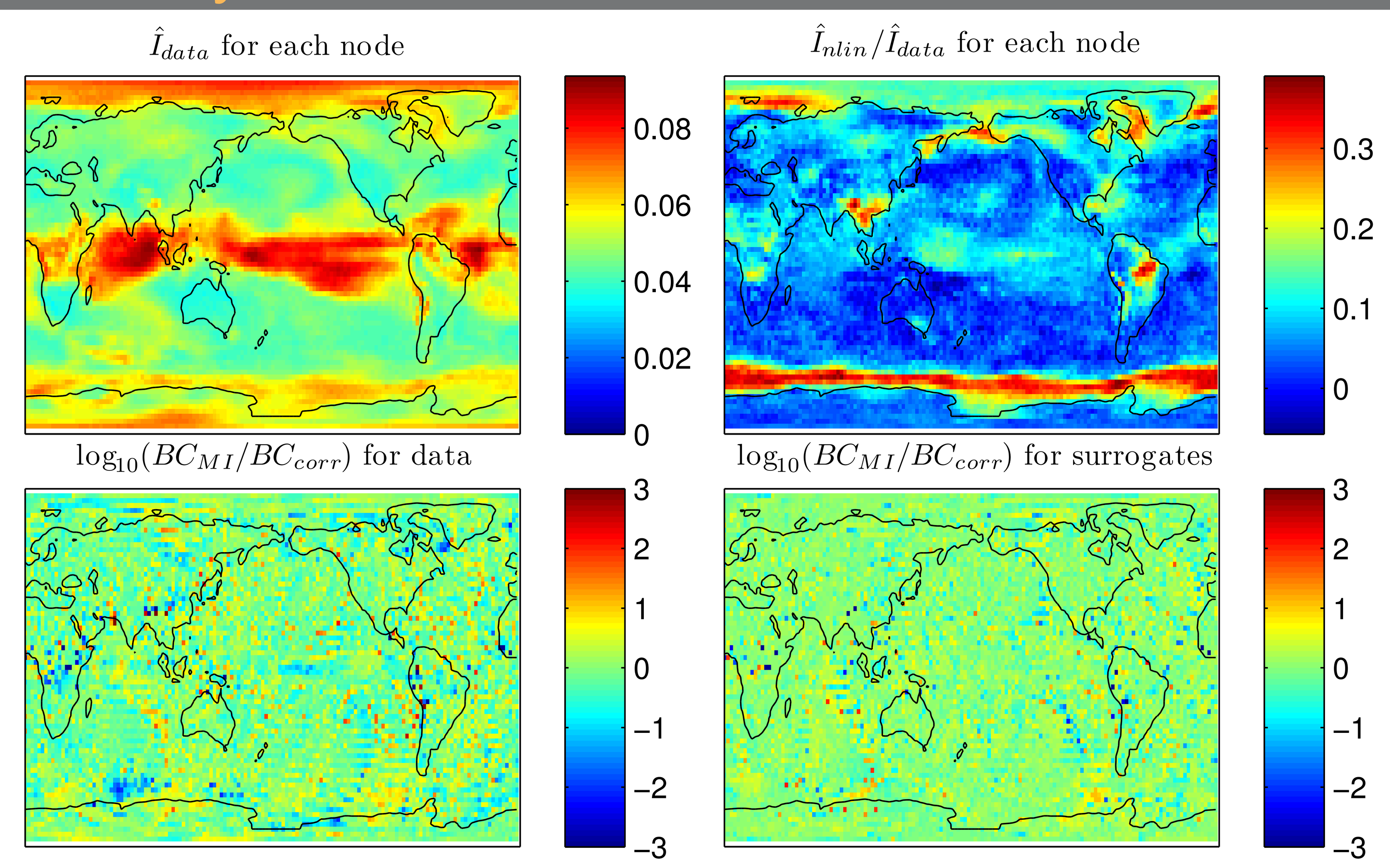
Monthly mean temperatures (756 months, Jan 1948 - Dec 2010) from the NCEP/NCAR reanalysis project [7], both poles removed. Time series were converted into temperature anomalies by subtracting monthly means from each month. Used edge density is 0.005 [3].

Nonlinearity in climate networks

- Nonlinearity role in climate network analysis
  - Observed for betweenness centrality [3]
  - Goal: localize and explain the effect of nonlinear contribution to connectivity and further to network characteristics
  - $I_{nlin} \sim I_{data} - I_{surr}$; $I_{surr} \sim -\frac{1}{2}\log(1 - r^2)$ [5]



Preliminary results for climate networks



References

[1] **E. Bullmore and O. Sporns**. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.*, 10(3):186–198, 2009.

[2] **O. David, D. Cosmelli, and K. J. Friston**. Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage*, 21(2):659–673, 2004.

[3] **J. F. Donges, Y. Zou, N. Marwan, and J. Kurths**. Complex networks in climate dynamics. *Eur. Phys. J.-Spec. Top.*, 174:157–179, 2009.

[4] **D. Hartman, J. Hlinka, M. Paluš, D. Mantini, and M. Corbetta**. The role of nonlinearity in computing graph-theoretical properties of resting-state functional magnetic resonance imaging brain networks. *Chaos*, 21(1):art.num. 013119, 2011.

[5] **J. Hlinka, M. Paluš, M. Vejmelka, D. Mantini, and M. Corbetta**. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? *Neuroimage*, 54(3):2218–2225, 2011.

[6] **M. Paluš**. Testing for Nonlinearity Using Redundancies - Quantitative and Qualitative Aspects. *Physica D*, 80(1-2):186–205, 1995.

[7] **R. Kistler et al**. The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, 82:247–268, 2001.

http://ndw.cs.cas.cz/ {hartman,hlinka,vejmelka,palus}@cs.cas.cz