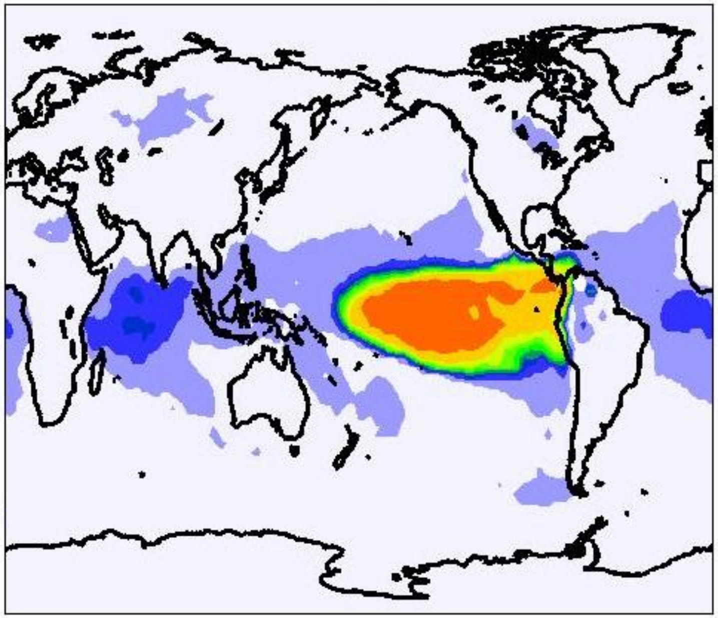


Introduction to symbolic time series analysis applied to climatological data

Cristina Masoller
Universitat Politècnica de Catalunya
cristina.masoller@upc.edu, www.fisica.edu.uy/~cris

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

"LINC"
Learning about Interacting Networks in Climate

Outline

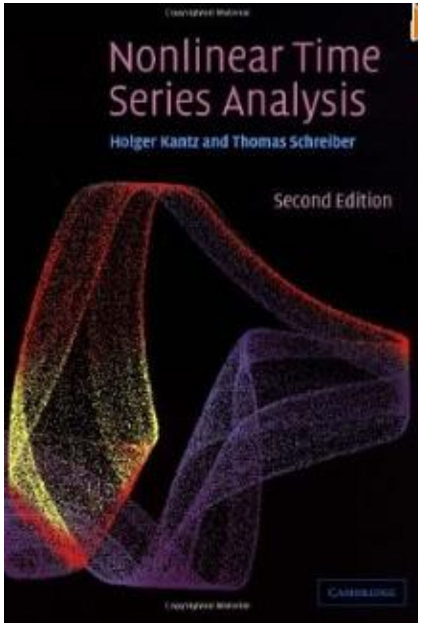
- Symbolic time series analysis
- Ordinal analysis
- Information theory measures
- Examples
- Application to climate data analysis

Nonlinear time-series analysis

- Many methods have been developed to test for determinism, nonlinearity and correlations in data generated from complex systems (climate, brain EEGs, financial data, social systems, etc).

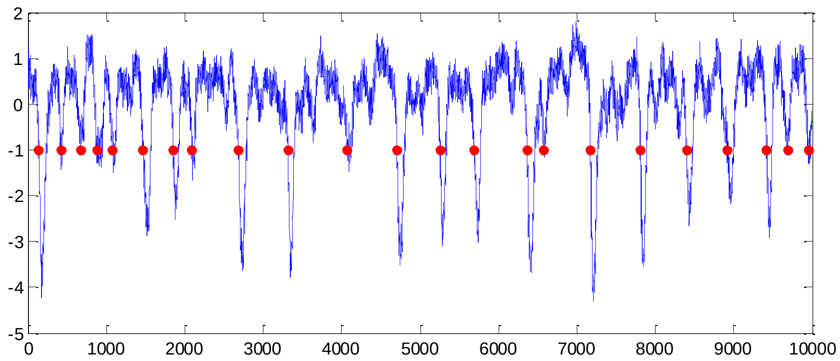
- The appropriateness of the method depends on the characteristics of the time series
  - short or long;
  - stationary or not;
  - more or less noisy;
  - multi or single channel measurements,
  - discrete or continuous values, etc.

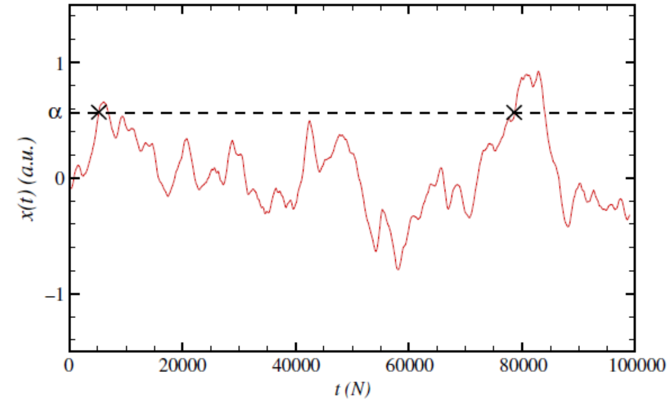
- Different methods provide complementary information.



Event-like description of a signal

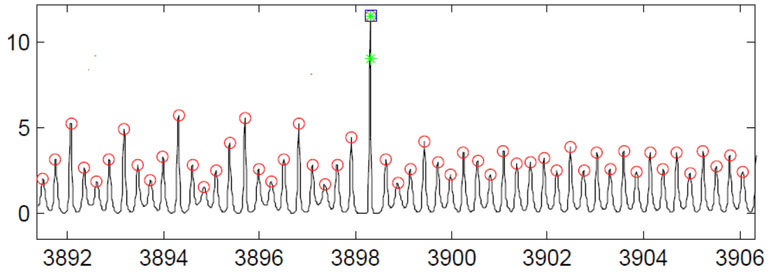
- Threshold crossings



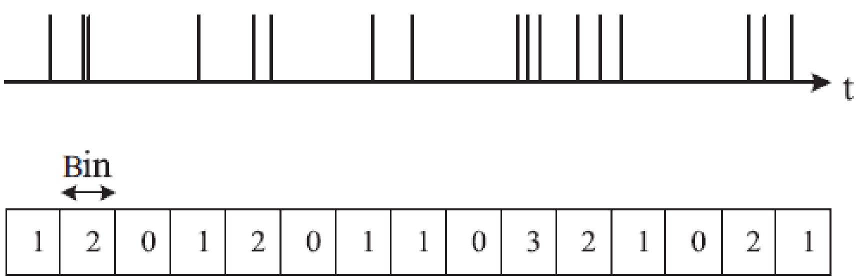


R. Friedrich et al. / Physics Reports 506 (2011) 87–162

- Extreme values



- Bin counting

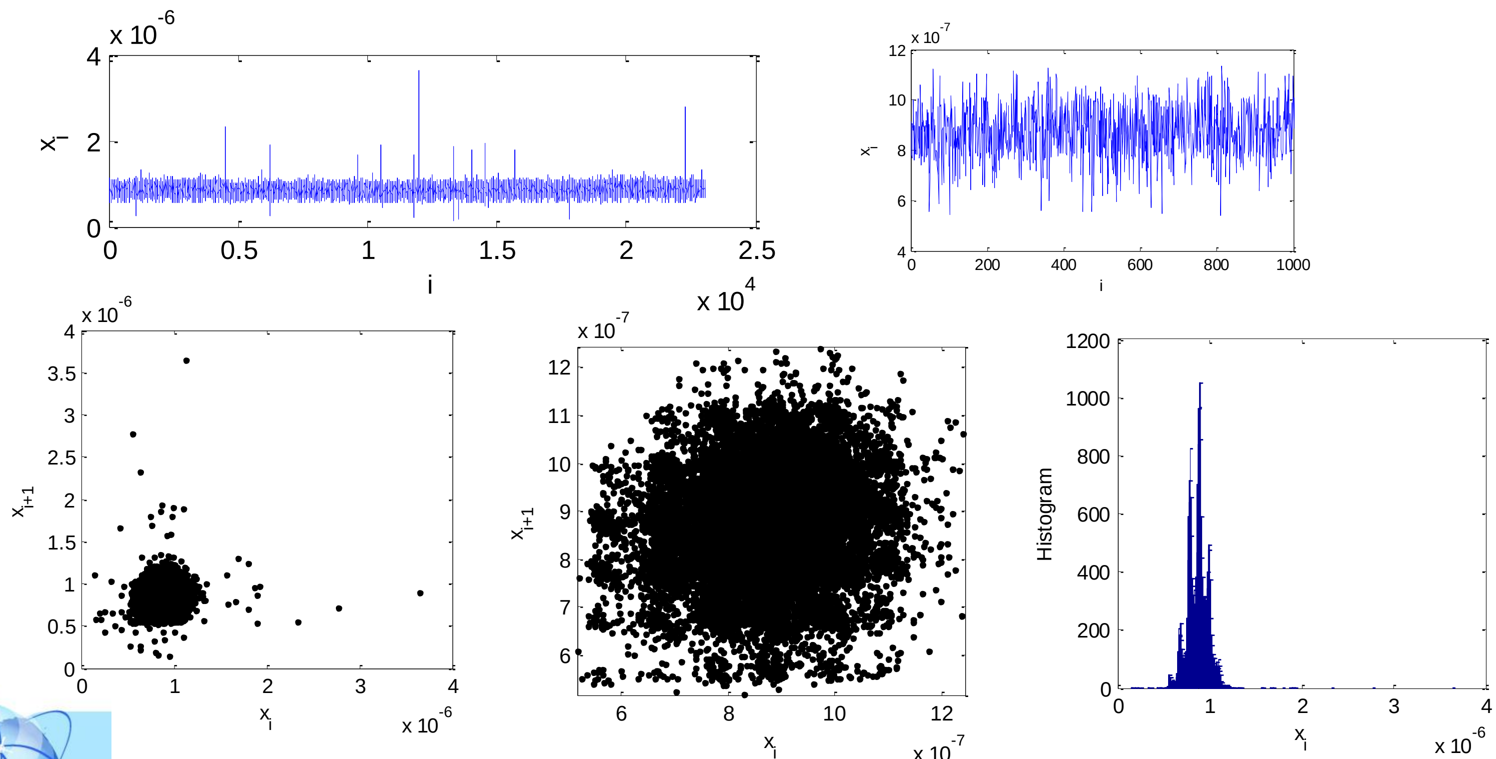


… and many others.

Visual inspection of time-series

- Consider a time series $\{x_1, x_2, \ldots, x_N\}$ generated from a complex system.

- First step: Look at the time series. Examine simple properties: auto/cross correlation, Fourier spectrum, return map (x_i vs $x_{i+\tau}$), histogram, etc.



Next: capture hidden structures

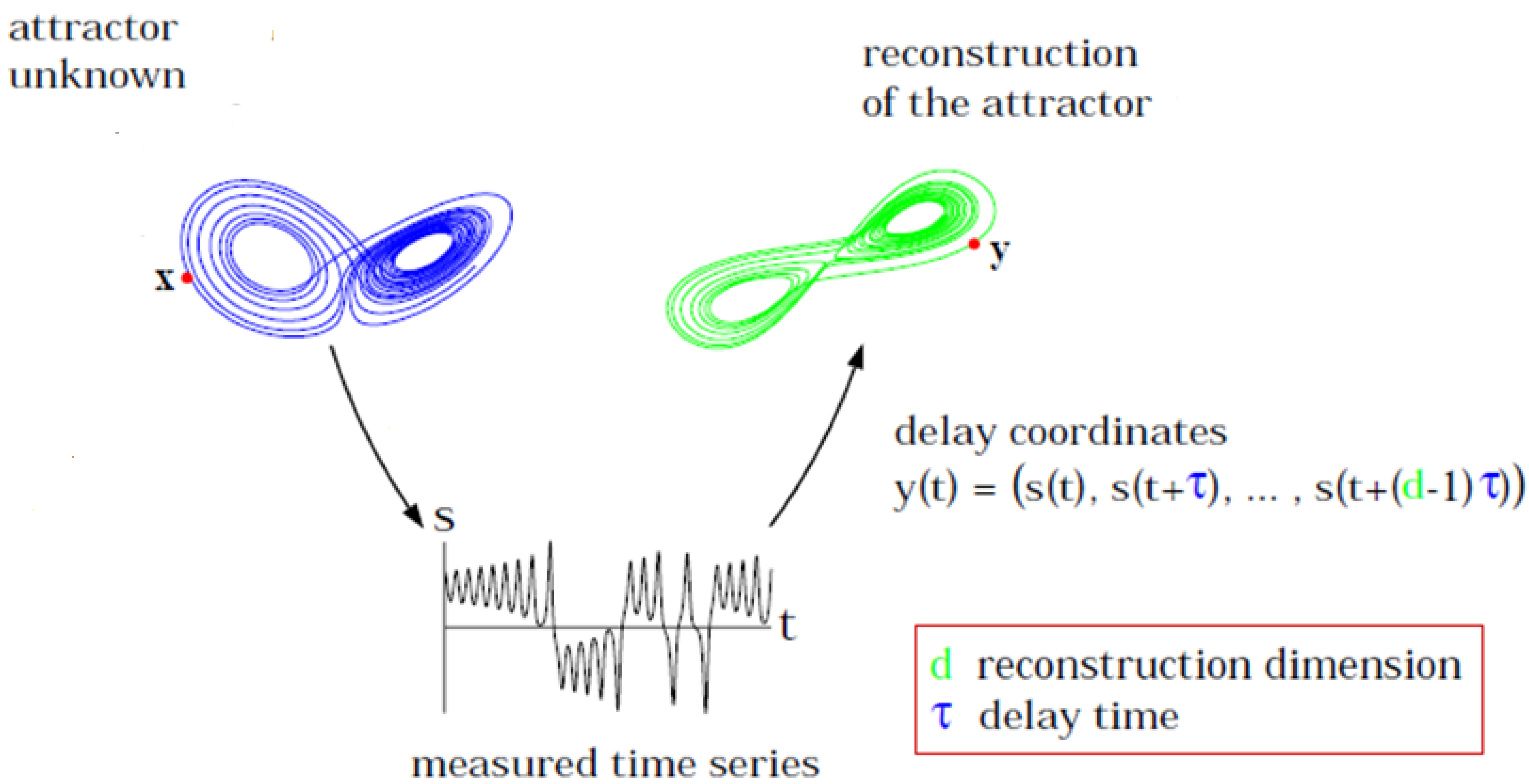
Two main approaches to identify patterns and ordering in the sequence

- Phase-space reconstruction methods
  - Time-delay coordinates
  - Derivative coordinates

- Symbolic methods

They allow for model verification, forecasting, classification of different types of behaviors, noise reduction, etc.

Reconstruction using delay coordinates



A problem: finding good embedding

Adapted from U. Parlitz, MPI for Complex Systems, Germany

Symbolic analysis

- The time series $\{x_1, x_2, x_3, \ldots\}$ is transformed (using an appropriated rule) into a sequence of symbols $\{s_1, s_2, \ldots\}$
- taken from an "alphabet" of possible symbols $\{a_1, a_2, \ldots\}$.
- Then consider "blocks" of D symbols ("patterns" or "words").
- All the possible words form the "dictionary".
- Then analyze the "language" of the sequence of words
  - the probabilities of the words,
  - missing/forbidden words,
  - transition probabilities,
  - symbolic information measures (entropy, mutual information, etc).

Example

- Binary transformation rule

  if $x_i > x_{th} \Rightarrow s_i = 0$; else $s_i = 1$

  transforms a time series into a sequence of 0s and 1s, e.g., {011100001011111…}

- Considering "blocks" of D=3 letters gives the sequence of words:

  {011 100 001 011 111 …}

Another symbolic transformation

- Ordinal transformation:

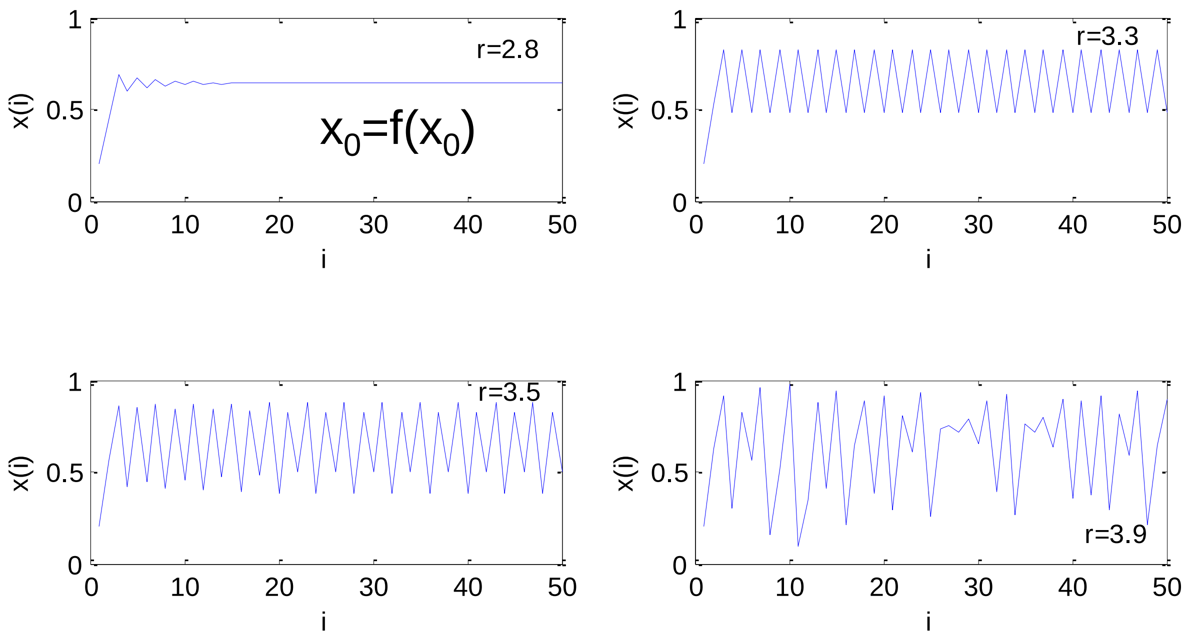
  if $x_i > x_{i-1} \Rightarrow s_i = 0$; else $s_i = 1$

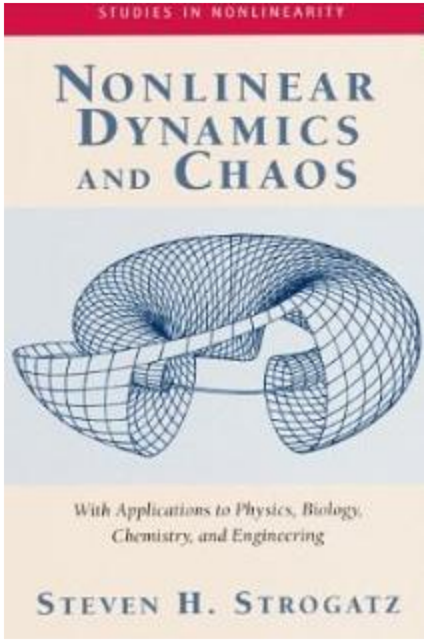
  also transforms a time-series into a sequence of 0s and 1s.

- "words" of D letters are formed by considering the order relation between sets of D values $\{\ldots x_i, x_{i+1}, x_{i+2}, \ldots\}$.

Example: the logistic map

- $x(i+1)=r\ x(i)[1-x(i)]$ $\qquad$ $f(x)=r\ x\ (1-x)$

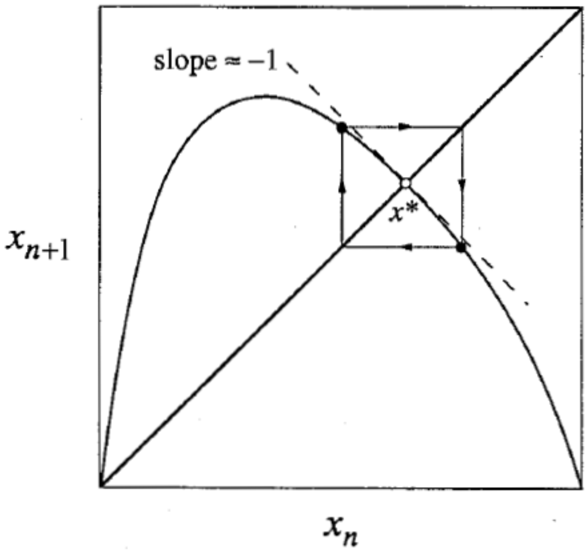




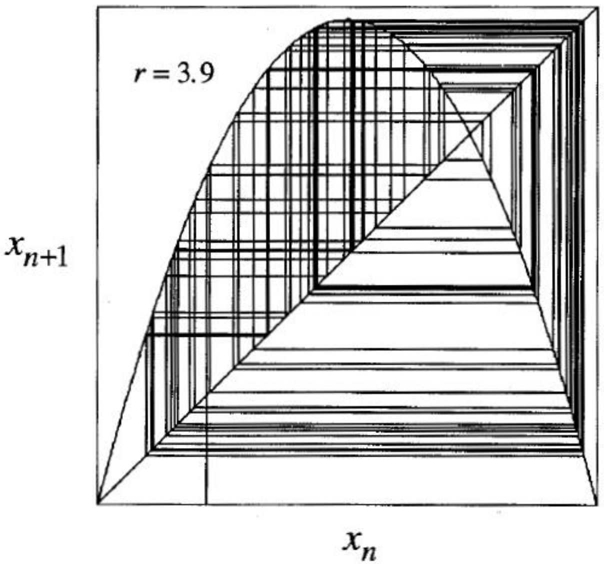
Example: the logistic map

- x(i+1)=r x(i)[1-x(i)]

Period 2 (r=3.5)

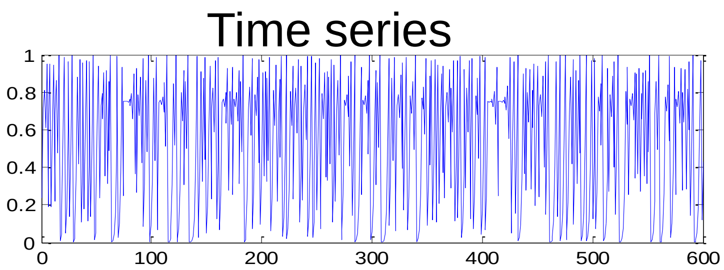


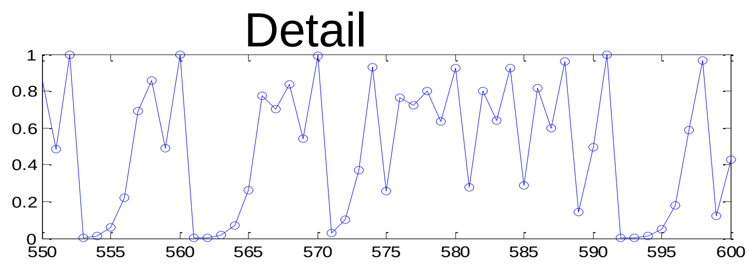
Chaos



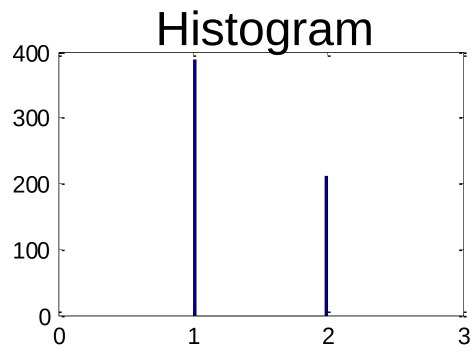
Ordinal analysis of the dynamics of the logistic map

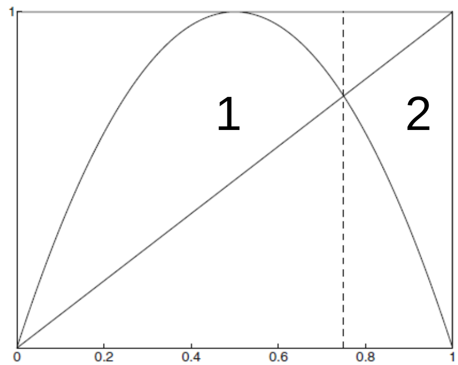
- $x(i+1)=4x(i)[1-x(i)]$





- Rule: if $x_i > x_{i-1} \Rightarrow s_i = 1$; else $s_i = 2$

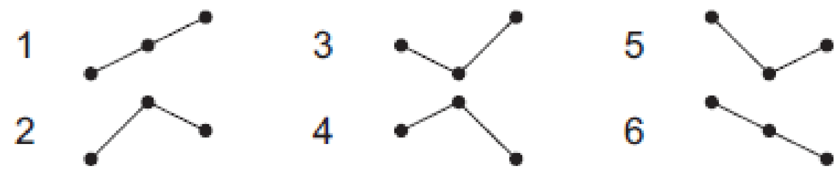




Words of length D

- Words of length D are determined by the order the values appear in the time series: each element of a "block" of length D is replaced by a number from 0 to D − 1 (0: the smallest element; D − 1: the longest element in each "block").

- Example: D=3

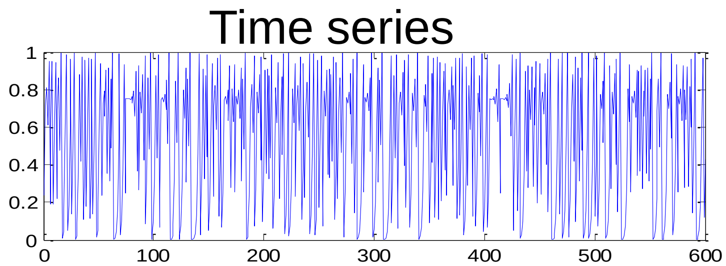


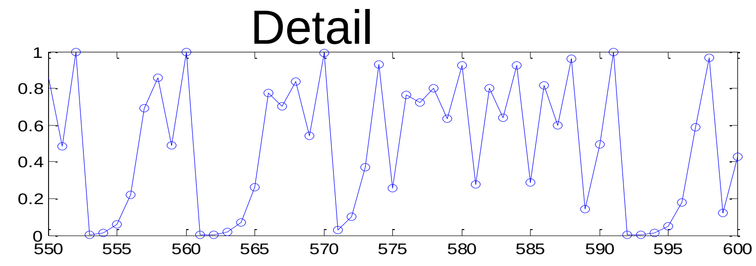
$\{\ldots\ x(t),\ x(t+1),\ x(t+2)\ldots\} = \{\ldots 5, -1,\ 10\ldots\}$

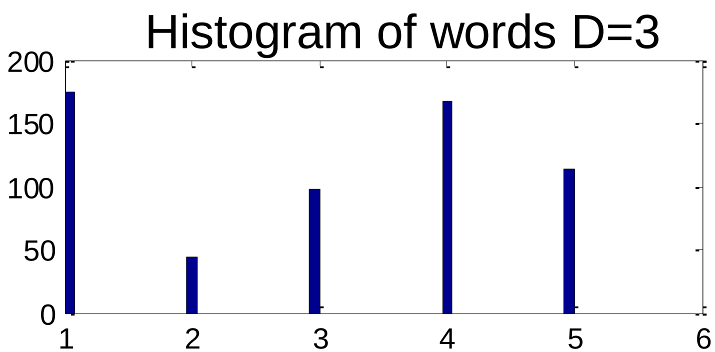
the set (5, -1, 10) gives word 102 because $x(t+1)<x(t)<x(t+2)$

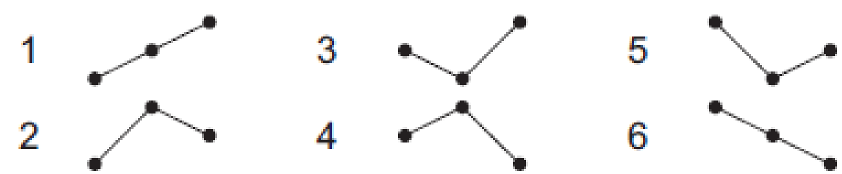
In the list 102 is word number 3

Logistic map: symbolic dynamics characterized with D=3 words









Word 6 forbidden

Ordinal analysis

- Was proposed by Bandt and Pompe in 2002 (Phys. Rev. Lett. 88, 174102).

- It has been successfully applied to the analysis of complex signals
  - Financial
  - Biological, life sciences
  - Geosciences, climate
  - Physics, chemistry, etc

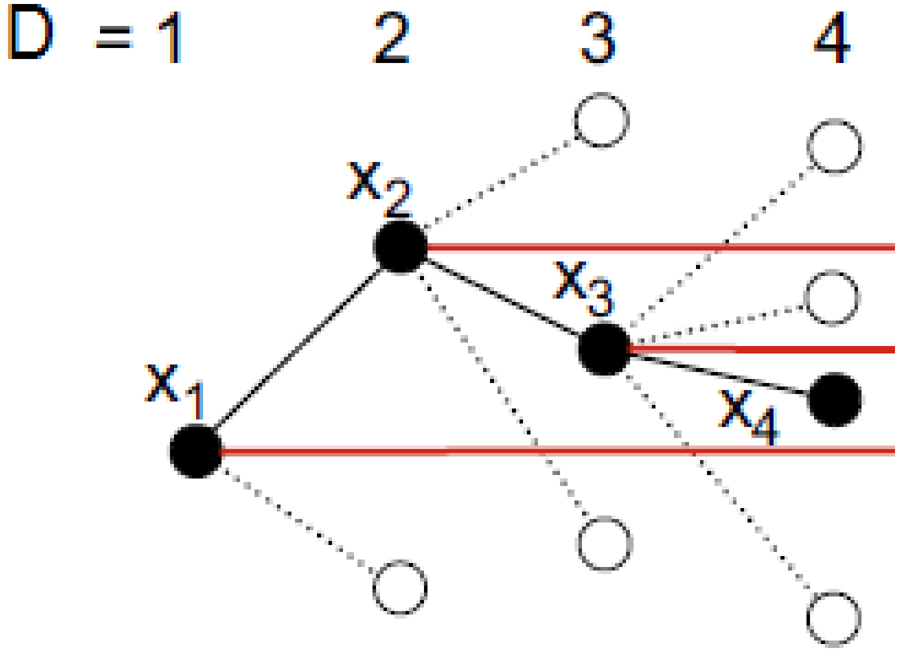
- It has been used to:
  - Distinguish stochasticity and determinism in high-dimensional systems
  - Classify different types of dynamical behaviors (pathological, healthy)
  - Quantify complexity
  - Identify coupling and directionality, etc.

Construction principle of ordinal patterns (OPs) of length D

Only 2 possible directions from x_1 to x_2: up or down.

From x_3: 3 possible directions.

From x_4: 4 possible directions.

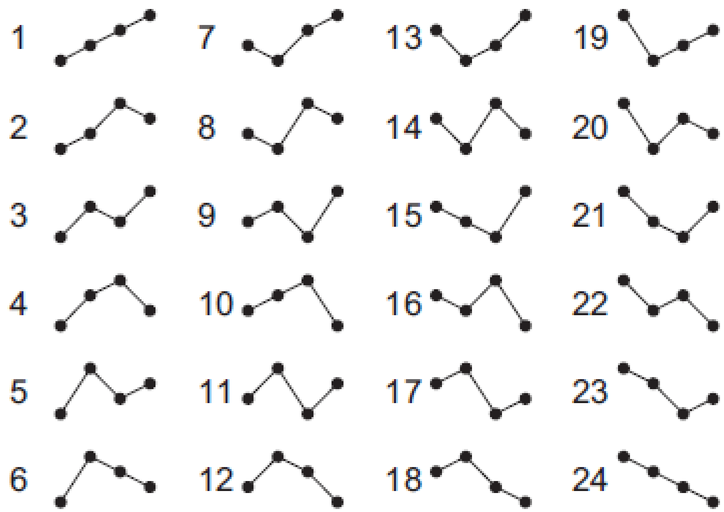


U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

Ordinal patterns

- For words of length D there are D! possible words in the dictionary.

D=4



U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

$$(2.22, 4.21, 1.30, 3.76)$$

$$\downarrow$$

$$(3,\ 1,\ 4,\ 2)$$

$$\downarrow$$

$$I = 11$$

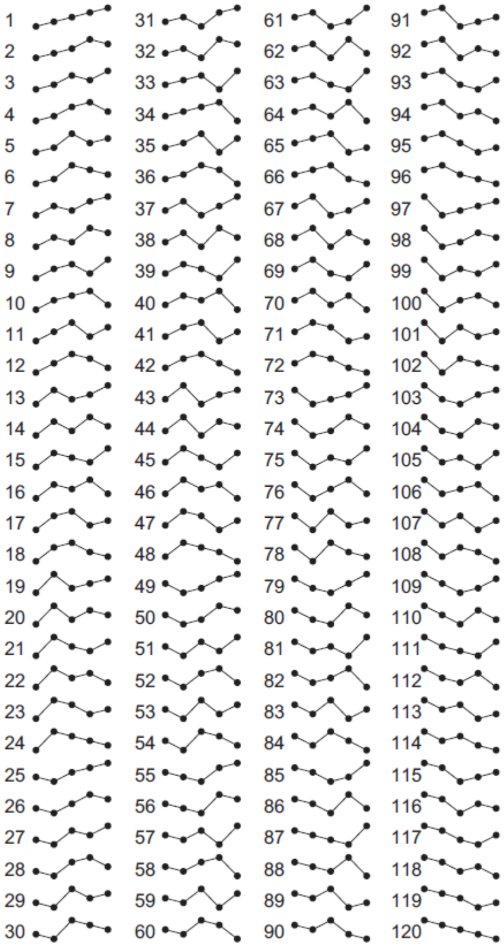
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Ordinal patterns

D=5

- The number of words in the "dictionary" grows fast with D



U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

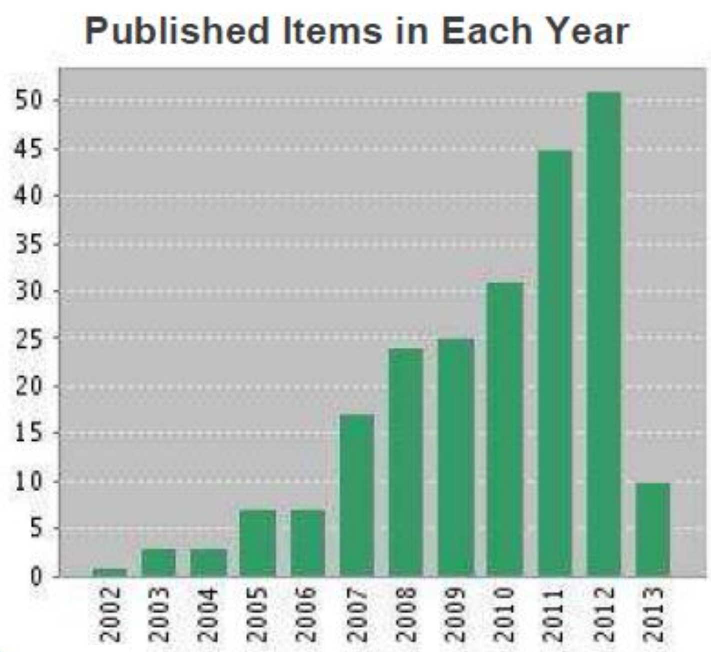
Ordinal analysis is becoming very popular

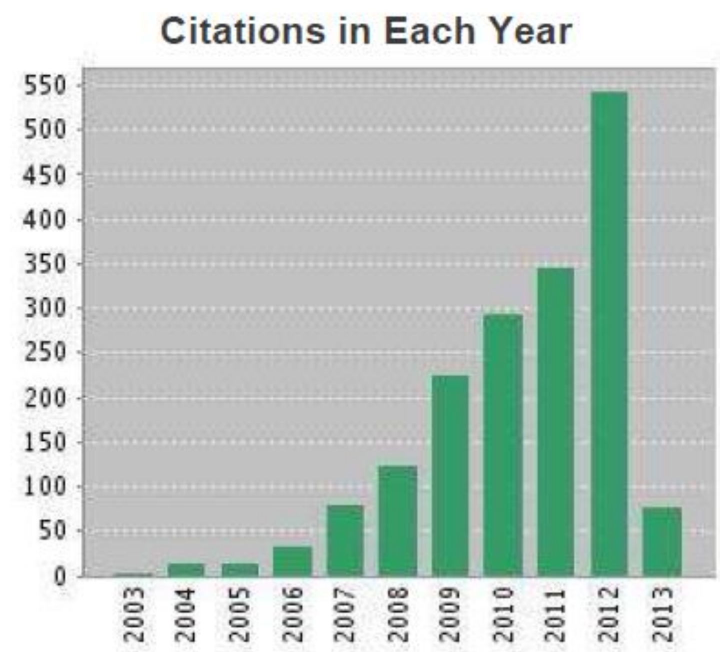
Citation Report

Title: **Permutation entropy: A natural complexity measure for time series**
Author(s): **Bandt, C ; Pompe, B**
Source: **PHYSICAL REVIEW LETTERS** Volume: **88** Issue: **17** Article Number: **174102**
10.1103/PhysRevLett.88.174102 Published: **APR 29 2002**

Timespan=All Years. Databases=SCI-EXPANDED, A&HCI, SSCI, CPCI-SSH, CPCI-S.

This report reflects citations to source items indexed within Web of Science. Perform a Cited Reference Search to include c
indexed within Web of Science.





Advantages and drawbacks of fix threshold method

- Binary transformation: if $x_i > x_{th} \Rightarrow s_i = 0$; else $s_i = 1$

  Number of words of length D in the "dictionary": 2^D

  D=3 000 001 010 011 100 101 110 111

- More thresholds allow for more letters in the "alphabet" (and more words in the dictionary). Example:
  if $x_i > x_{th1} \Rightarrow s_i = 0$;
  else if $x_i < x_{th2} \Rightarrow s_i = 2$;
  else $(x_{th2} < x_i < x_{th1}) \Rightarrow s_i = 1$.

- Advantage: keeps information about the magnitude of the values.

- Drawback: requires the existence of one or more adequate thresholds.

Advantages and drawbacks of ordinal symbolic method

- Ordinal transformation: words defined by order relations

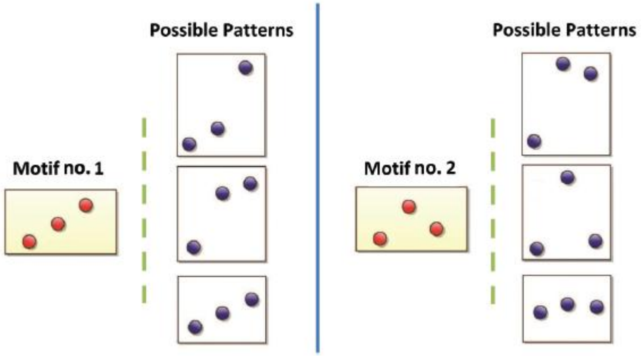
  Number of words of length D in the “dictionary”: D!

  D=3 012 021 102 120 201 210

  – Advantage: keeps information about the order in which the values appear in the sequence; does not need threshold

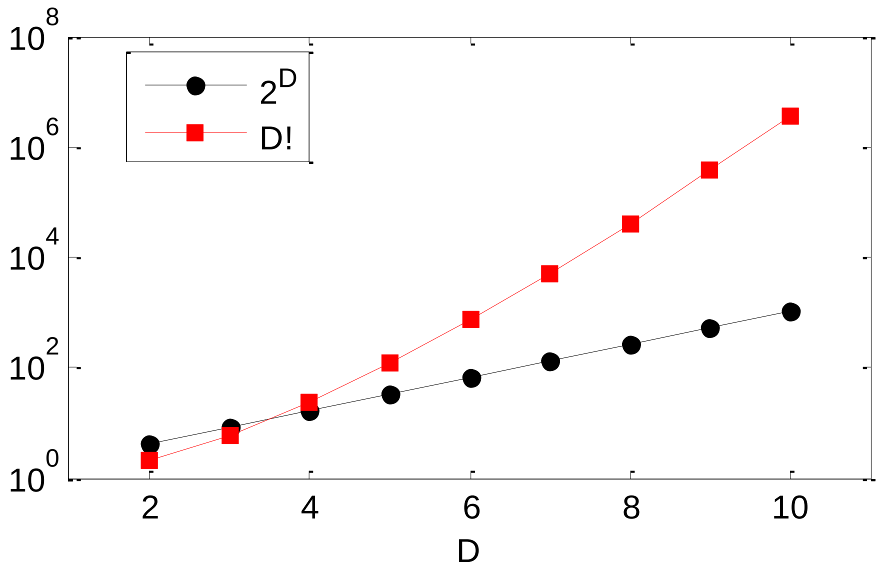
  – Drawback: the information about the absolute magnitudes is lost

Not surprisingly, extensions are being proposed to overcome this problem.



Fadlallah et al, PRE 2013

Number of words of D letters in the dictionary



- With OP method, to consider long time scales we need either a very long time series (to reliable compute probabilities) or a lag (more latter).

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

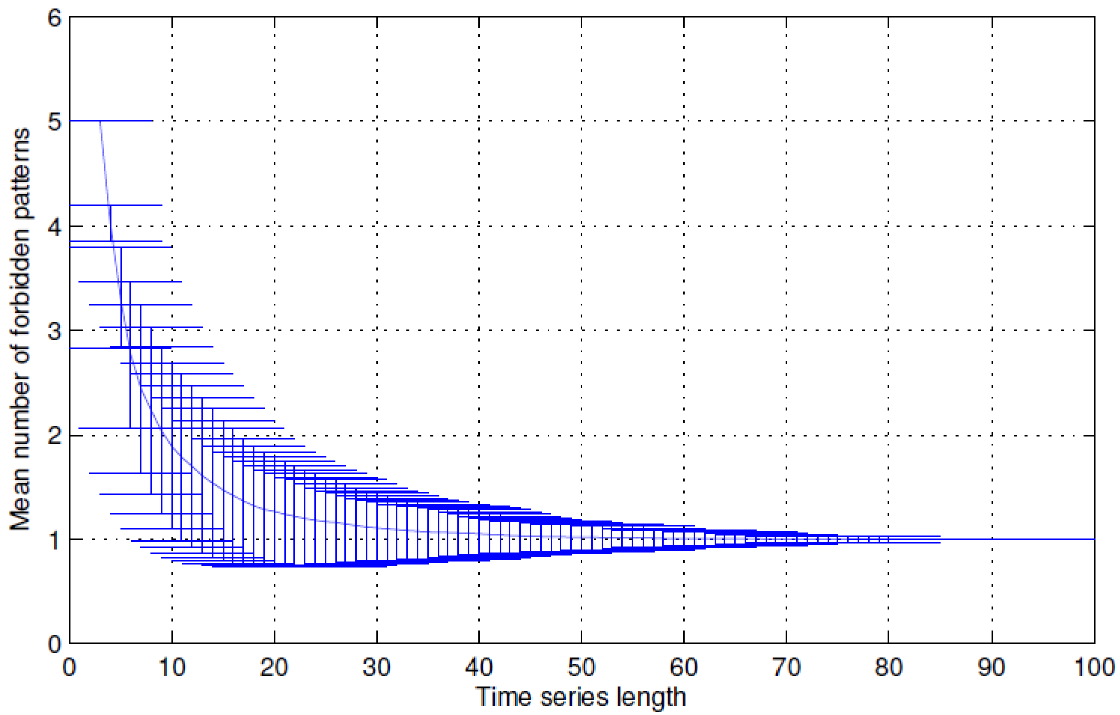
Campus d'Excel·lència Internacional

Next?

- Assuming that we have a suitable symbolic description of the time series.
- What information can we obtain from the sequence of "words"?
- Analogy with deciphering a foreign text.

Forbidden patterns

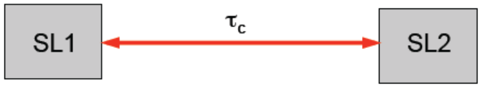
- Number of forbidden patterns (D = 3), found in 1000 time series generated with the logistic map (r=4), as a function of the length of the series.

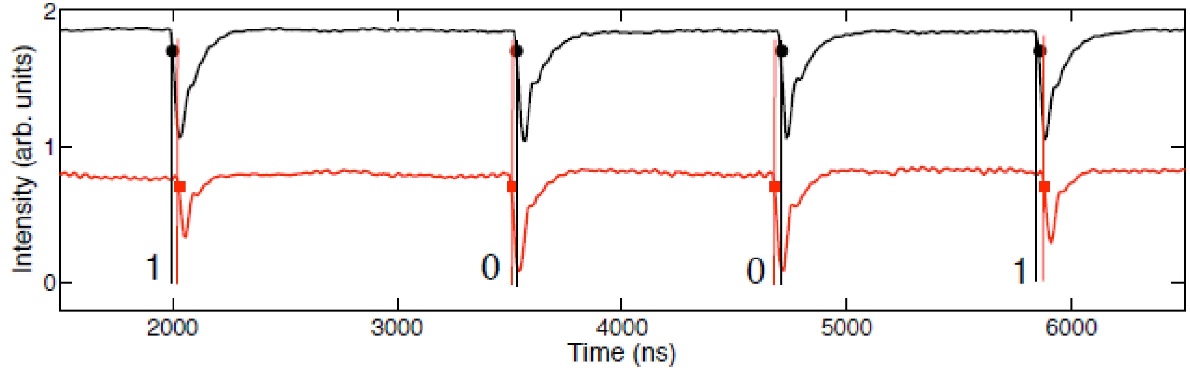


M. Zanin et al, Entropy 14, 1553 (2012)

Application: missing patterns as a signature of stochasticity

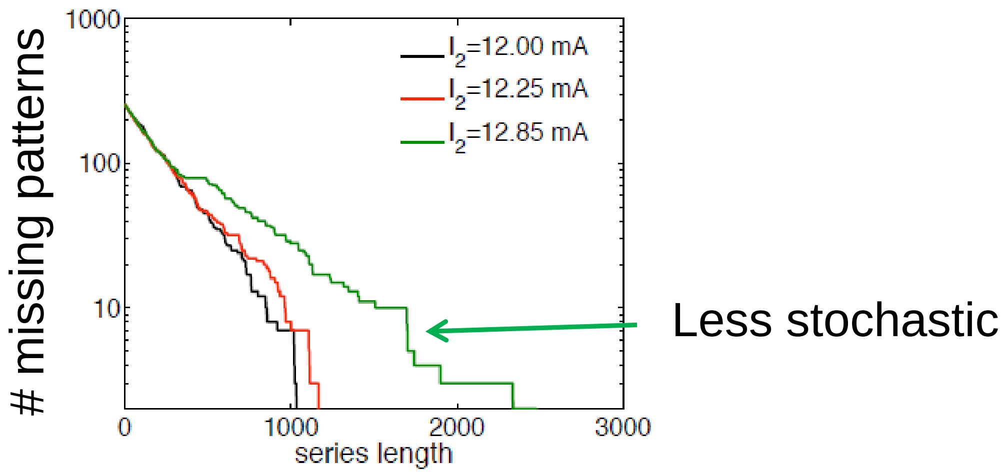
- Intensities of two coupled lasers





Pairs of synchronized dropouts are labeled as 1 if the dropout of SL1 occurs earlier than the one of SL2, else are labeled 0.

- Words are formed with 8 letters; the number of words in the dictionary is $2^8 = 256$.



J. Tiana-Alsina et al, Phil. Trans. Royal Soc. A 368, 367 (2010)

Information theory measure: Shannon entropy

- How much information is in a time-series?
- Consider the probabilities associated to a discrete variable

$$\sum_{i=1}^{N} p_i = 1$$

- Shannon Entropy: $H = -\sum_i p_i \log_2 p_i$
- Interpretation: "quantity of surprise one should feel upon reading the result of a measurement" [K. Hlavackova-Schindler et al, Physics Reports 441 (2007)]

.

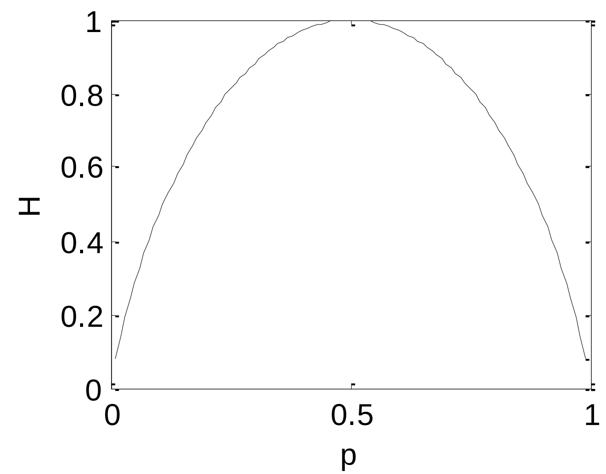
Shannon entropy

$$H = -\sum_i p_i \log_2 p_i$$

- Simple example: suppose that a random variable takes values 0 or 1 with probabilities:

  p(0) = p, p(1) = 1 − p.

- $H = -p \log_2(p) - (1-p)\log_2(1-p)$.

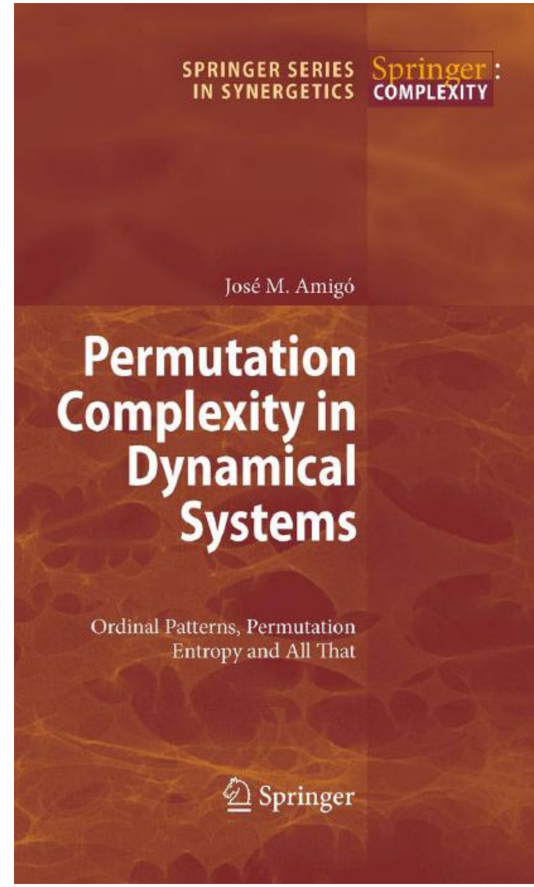


.

Permutation Entropy

- Consider a time series and its ordinal representation in terms of "words" of length D.

  The entropy computed from the probabilities of the words is the Permutation Entropy.



Recent reviews

Entropy **2012**, *14*, 1553-1577; doi:10.3390/e14081553

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

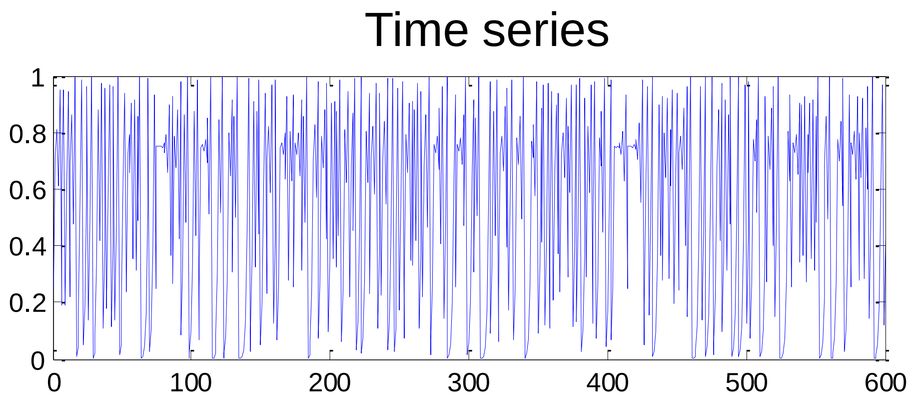
Article

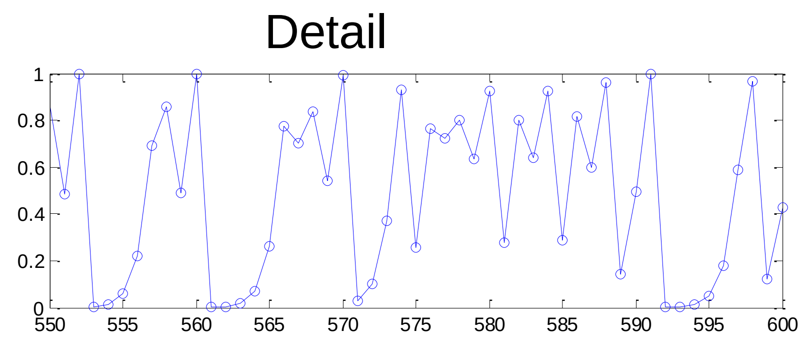
Permutation Entropy and Its Main Biomedical and Econophysics Applications: A Review

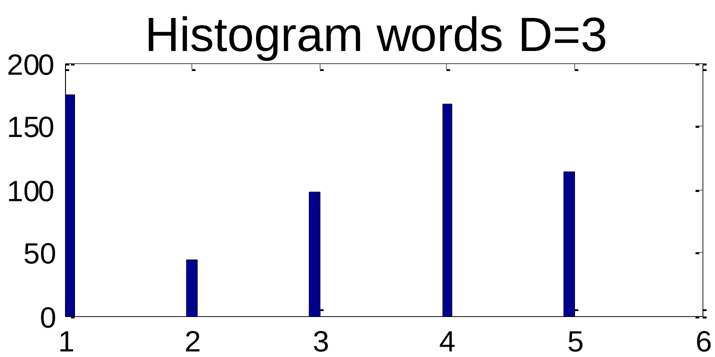
Massimiliano Zanin [1,2,3,] *, Luciano Zunino [4,5], Osvaldo A. Rosso [6,7] and David Papo [1]

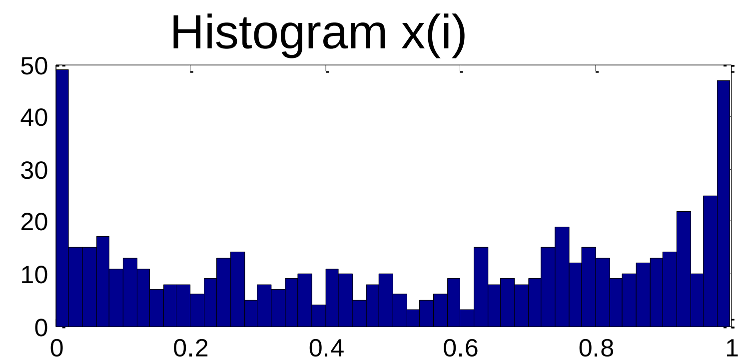
- See also EPJST 2013, special issue on PE.

Example: the logistic map $x(i+1)=4x(i)[1-x(i)]$





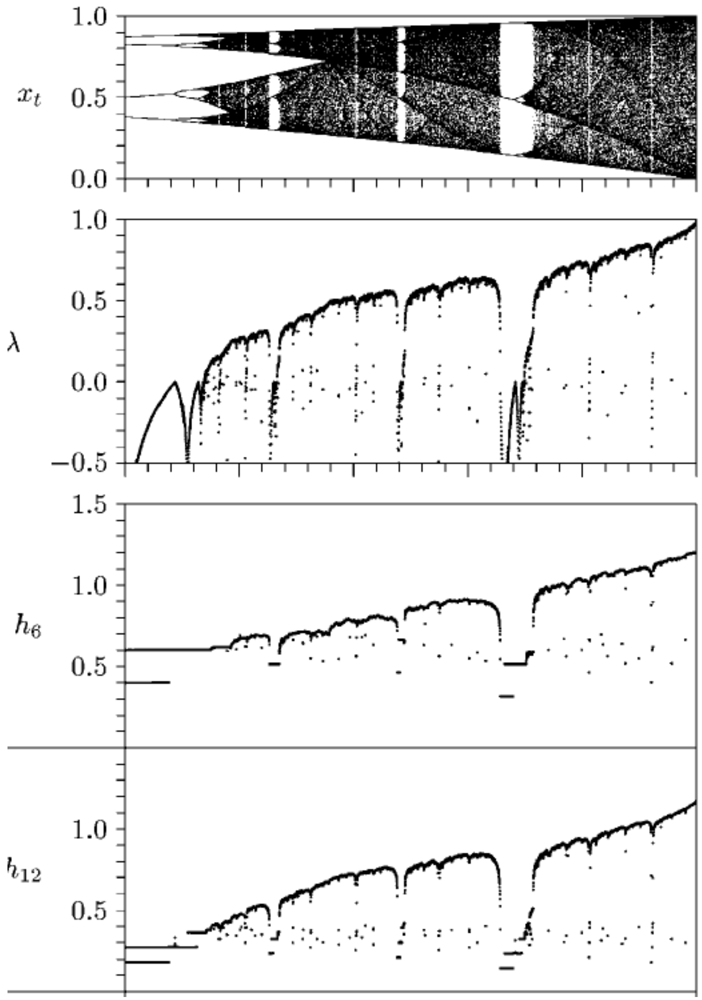




Permutation entropy is computed from the word probabilities

Shannon entropy is computed from x(i) probability distribution function (PDF).

Permutation entropy and Lyapunov exponent



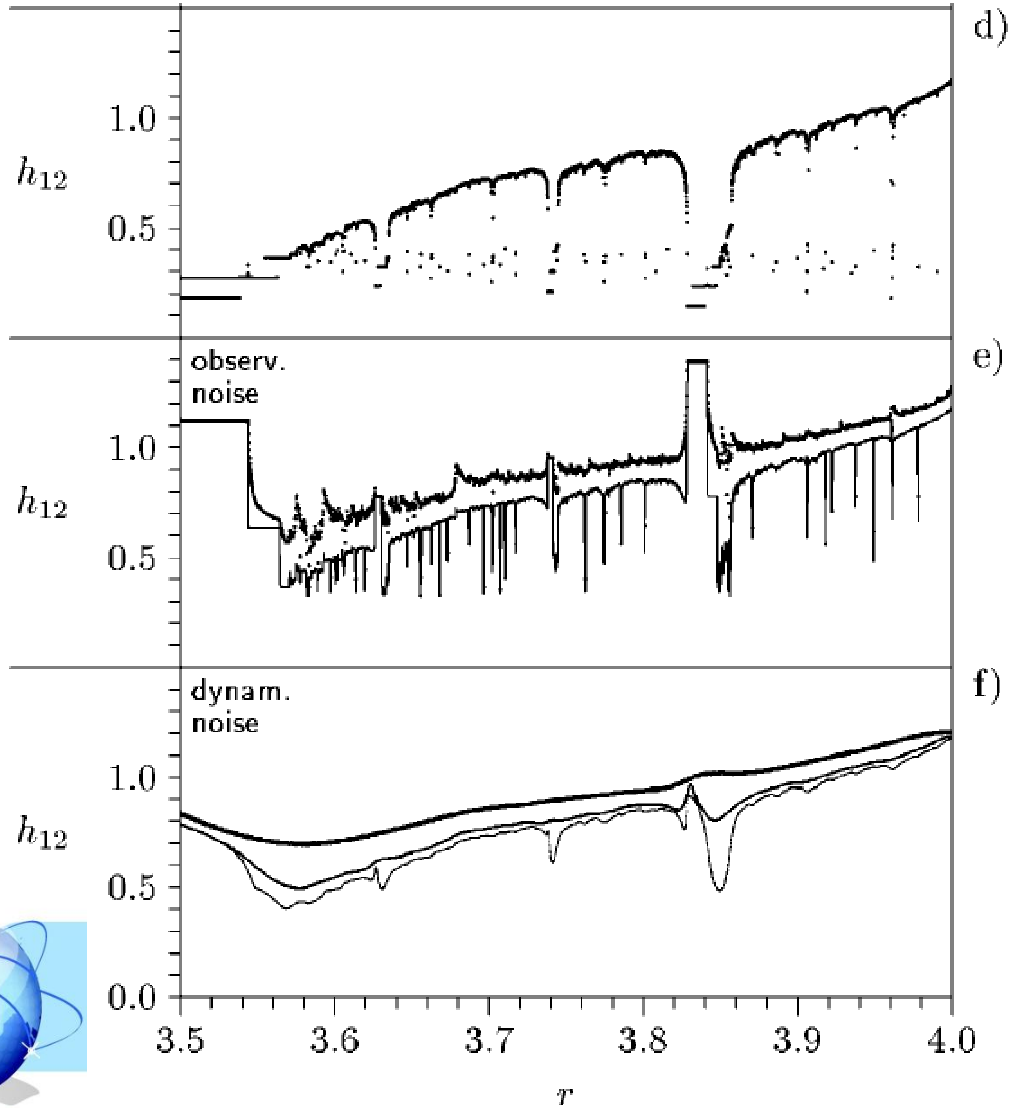
Parameter r

Entropy per symbol:

$$h_n = H(n)/(n-1)$$

Bandt and Pompe PRL 2002

Permutation entropy and noise



No noise

Weak noise (lower line), stronger noise (upper line)

Weak noise (lower line) to stronger noise (upper line)

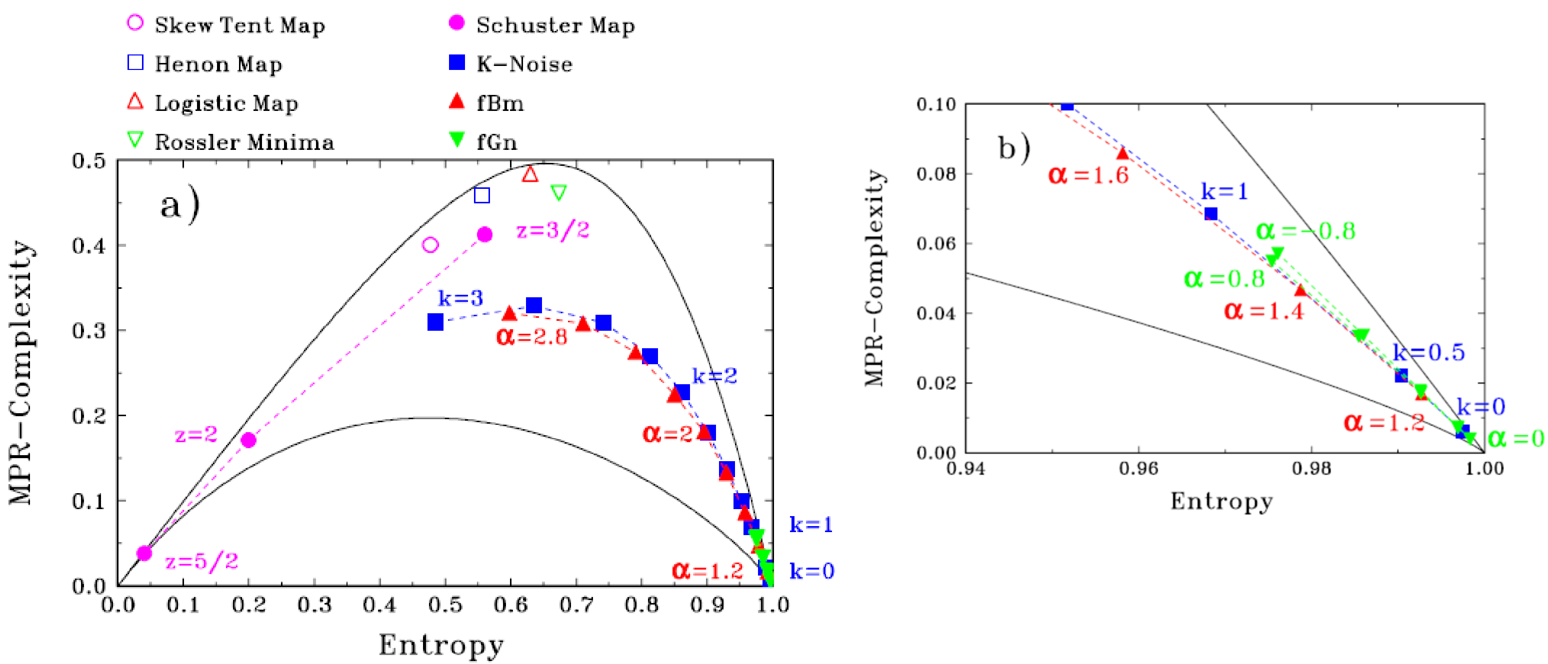
Bandt and Pompe PRL 2002

Application

PRL 99, 154102 (2007) PHYSICAL REVIEW LETTERS week ending 12 OCTOBER 2007

Distinguishing Noise from Chaos

O. A. Rosso,[1,2] H. A. Larrondo,[3] M. T. Martin,[4] A. Plastino,[4] and M. A. Fuentes[5,6]



Constructing longer words

$$[\ldots\ x(t),\ x(t+1),\ x(t+2),\ x(t+3),\ x(t+4),\ x(t+5)\ldots]$$

- But long time series will be required to estimate the probabilities of the fast growing number of words in the dictionary (D!).

- Solution: a lag allows considering long time-scales without having to use words of many letters

$$[\ldots\ x(t),\ x(t+2),\ x(t+4),\ldots]$$

- For climatological data (assuming monthly data):
  - Consecutive months: $[\ldots x_i(t),\ x_i(t+1),\ x_i(t+2)\ldots]$
  - One year: $[\ldots x_i(t),\ldots\ x_i(t+4),\ldots\ x_i(t+8)\ldots]$
  - Consecutive years: $[\ldots x_i(t),\ldots\ x_i(t+12),\ldots\ x_i(t+24)\ldots]$
  - etc

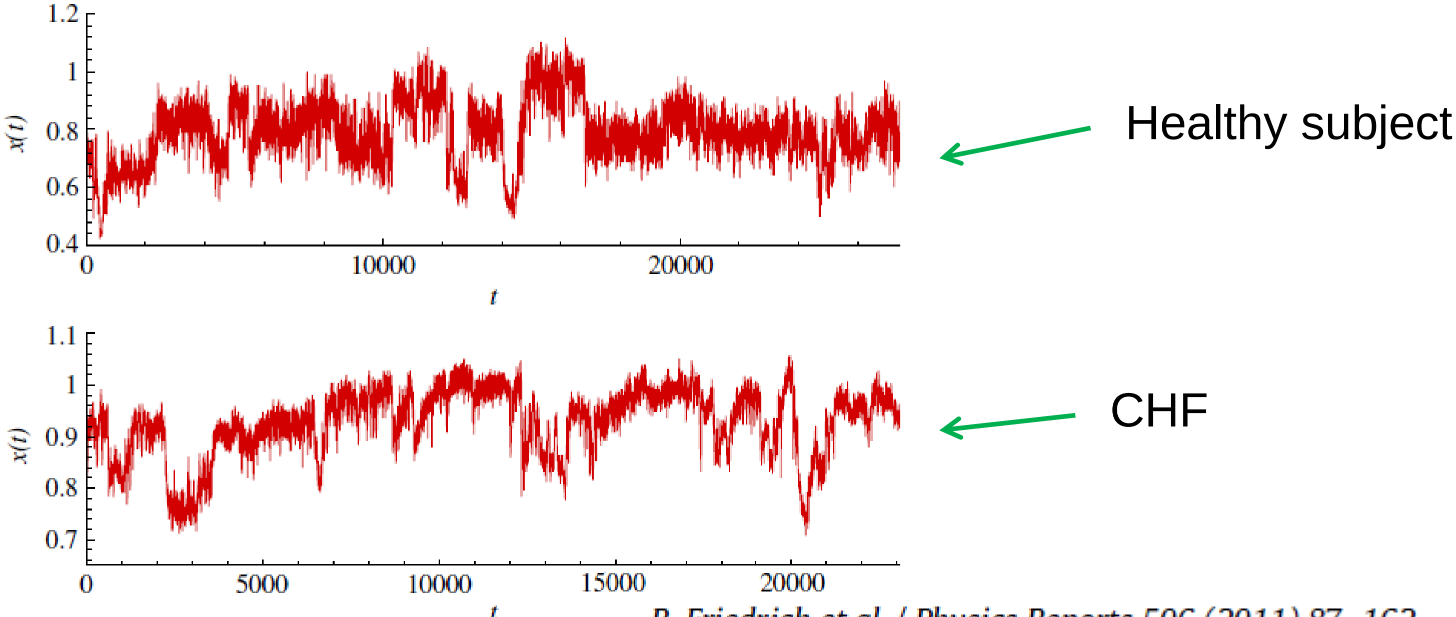
Application

Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics

U. Parlitz [a,b,*], S. Berg [c], S. Luther [a,b,d], A. Schirdewan [e], J. Kurths [f,g], N. Wessel [f]

Computers in Biology and Medicine 42 (2012) 319–327

- Distinguishing patients suffering from congestive heart failure (CHF) from a (healthy) control group using beat-to-beat (inter-beat intervals) time series



R. Friedrich et al. / Physics Reports 506 (2011) 87–162

Application

- After pre-processing the signals, classification is done in terms of the probability of occurrence of a word "i" with "D" letters, constructed with lag "l"

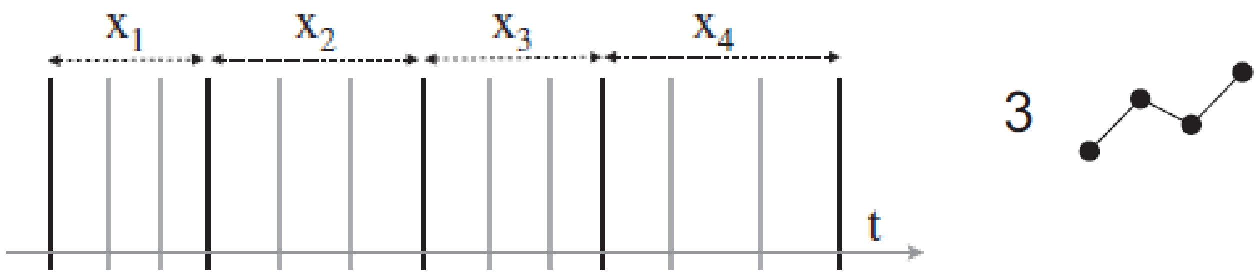
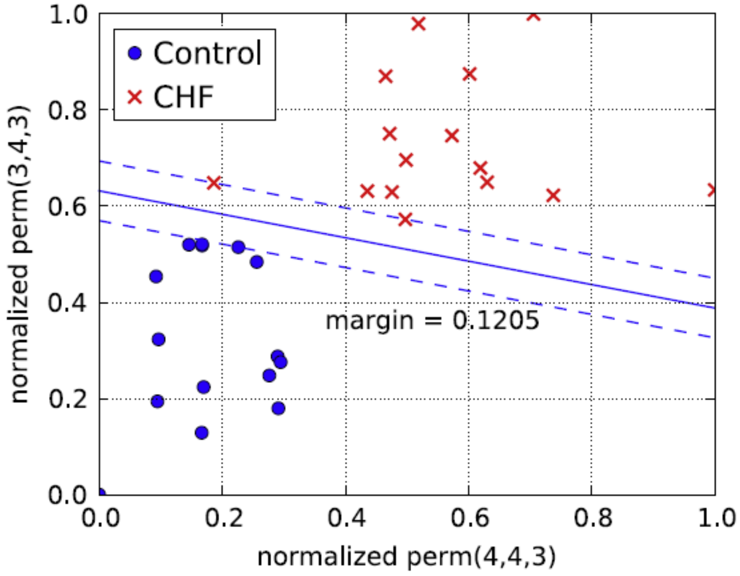
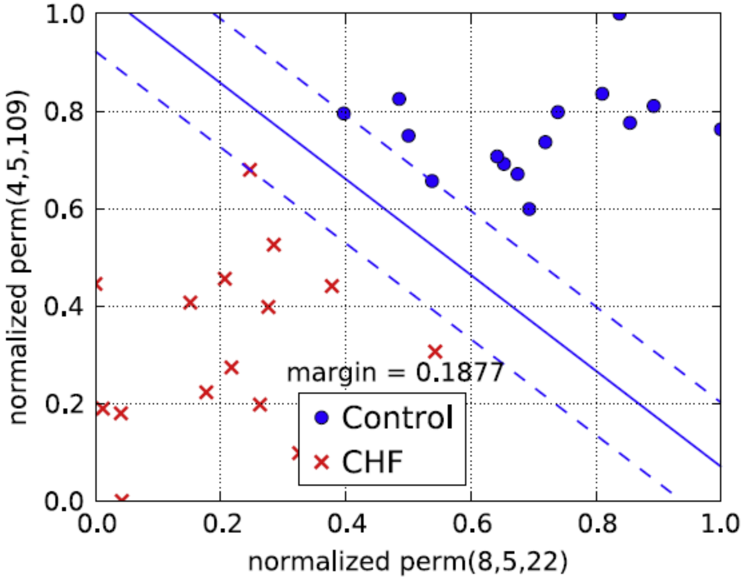


Fig. 4. Sketch of an ECG-signal (sequence of *R*-peaks) corresponding to an ordinal pattern perm(3,4,3). The time intervals x_k between every third beat ($T=3$) are ordered as $x_1 < x_3 < x_2 < x_4$.

Lag: l=3 (skip 3 peaks)
Letters: D=4
Word: i=3

Classifying ECG-signals according to the appearance of words





Perm (i,D,lag)

(the probabilities are normalized with respect to the smallest and the largest value occurring in the data set)

Matlab code + Exercise

Code for generating the words:

```
function indcs = perm_indices(ts, wl, lag) ;
m = length(ts)-(wl-1)*lag;
 indcs = zeros(m,1) ;
for i = 1:wl-1 ;
   st = ts(1+(i-1)*lag : m+(i-1)*lag) ;
   for j = i:wl-1 ;
      indcs = indcs+(st > ts(1+j*lag : m+j*lag));
   end
   indcs = indcs*(wl-i) ;
end
indcs=indcs + 1 ;
```

- ts: Time series
- wl: Length of the word (D)
- lag

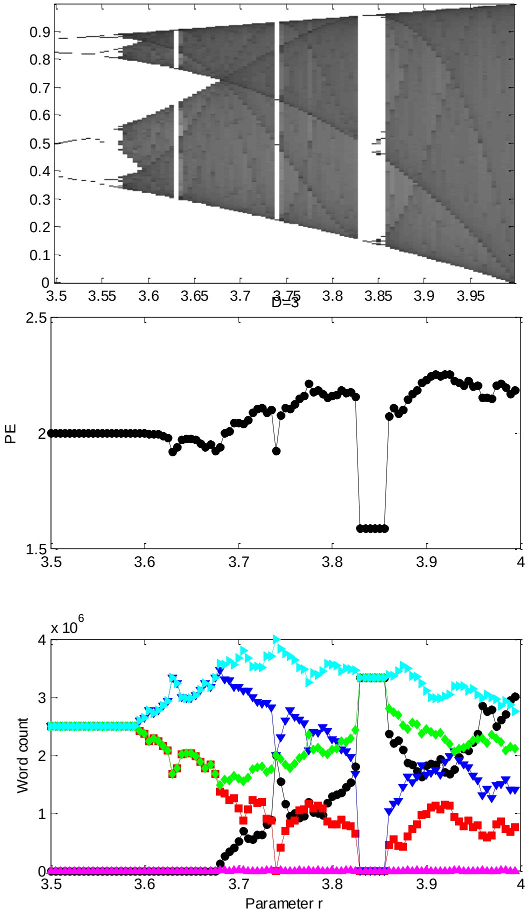
Also for Pyton:

U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

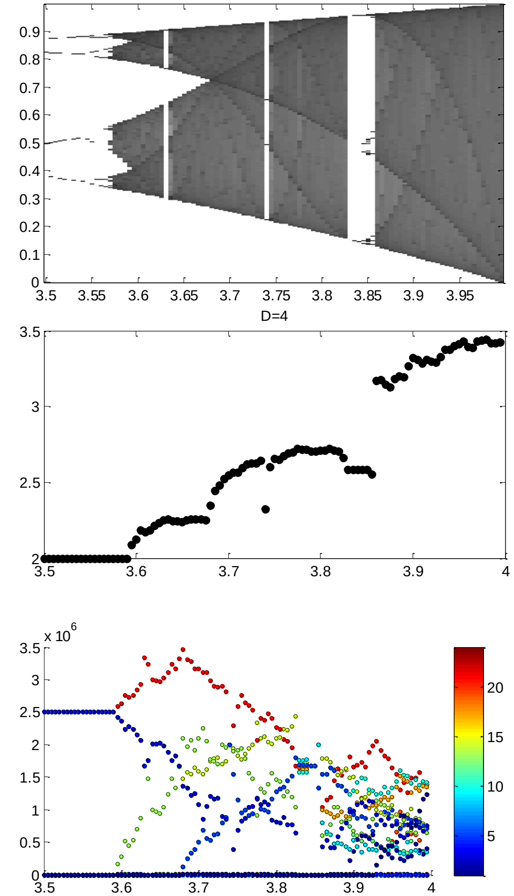
- Exercise: compute the "ordinal" bifurcation diagram of the logistic map (word probabilities vs r) for various D values, compute the PE and discuss the effect of "observational" noise.

Expect something like this

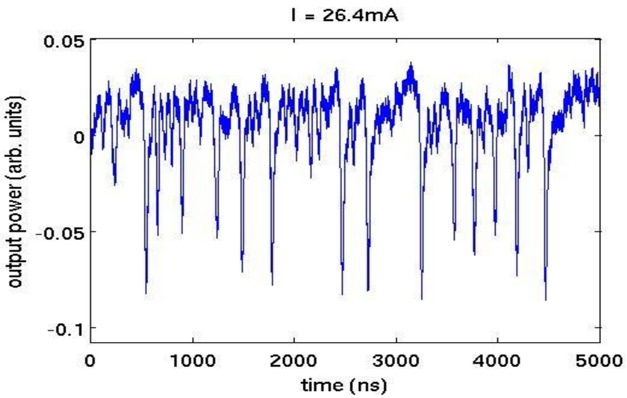
D=3

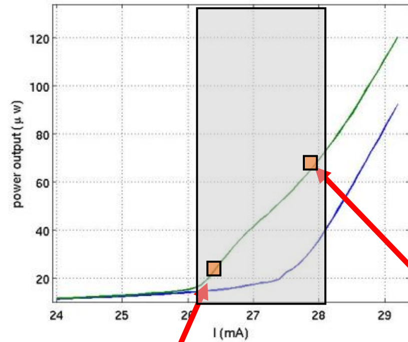


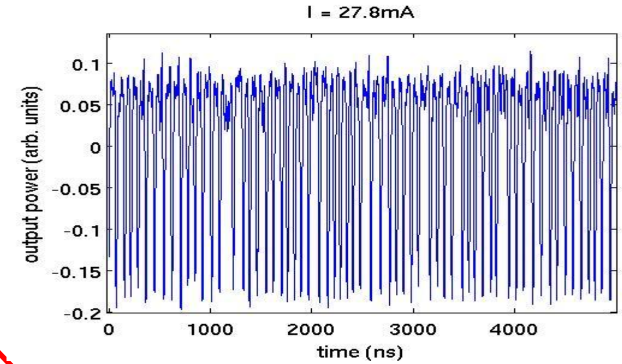
D=4

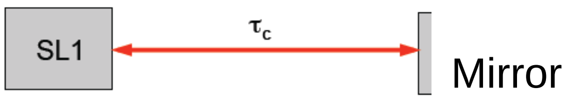


Application: decoding the spike code of a diode laser with feedback

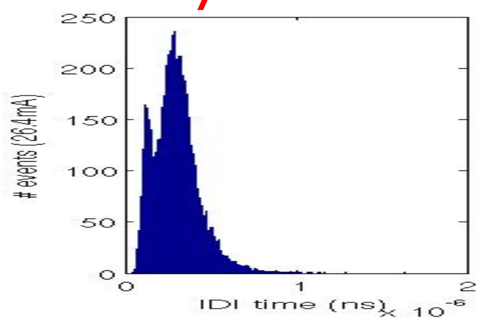


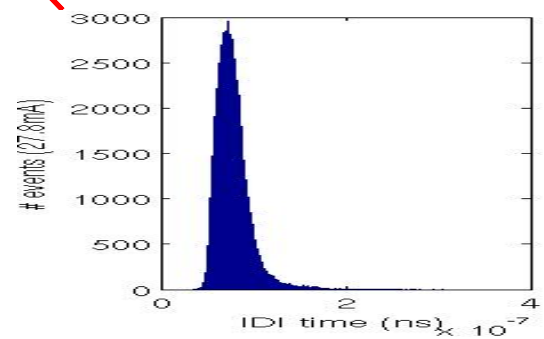






High dimensional and stochastic system





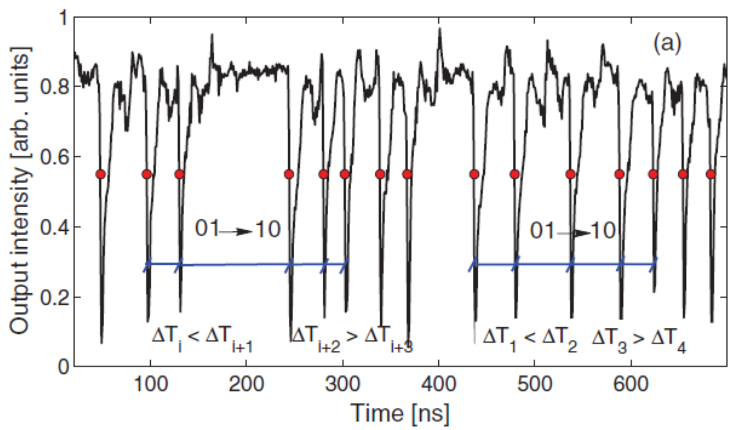
Is there any information in the inter-dropout-interval (IDI) sequence?

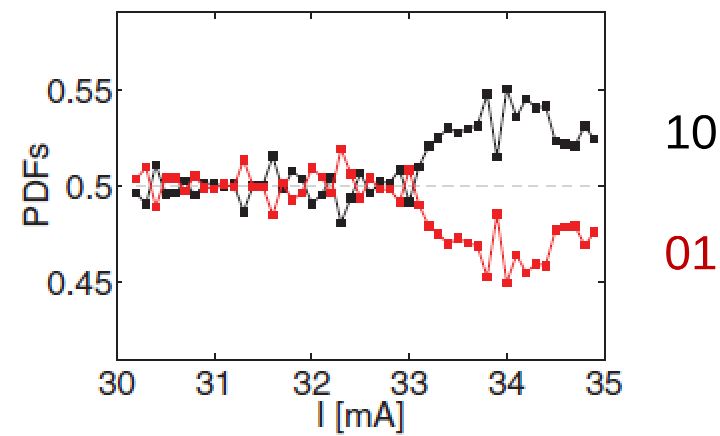
"language" analysis: word probabilities

PHYSICAL REVIEW E **84**, 026202 (2011)

Language organization and temporal correlations in the spiking activity of an excitable laser: Experiments and model comparison

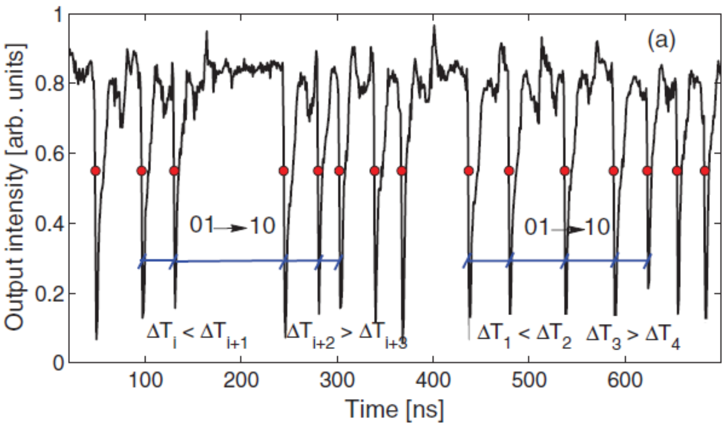
Nicolas Rubido,[1] Jordi Tiana-Alsina,[2] M. C. Torrent,[2] Jordi Garcia-Ojalvo,[2] and Cristina Masoller[2]

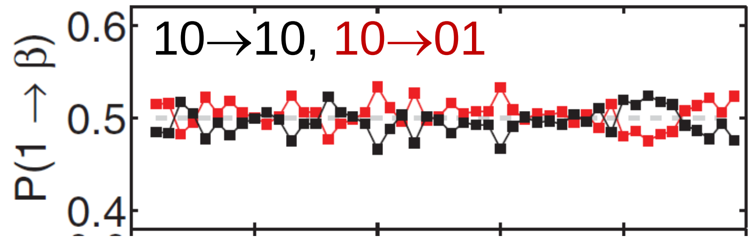


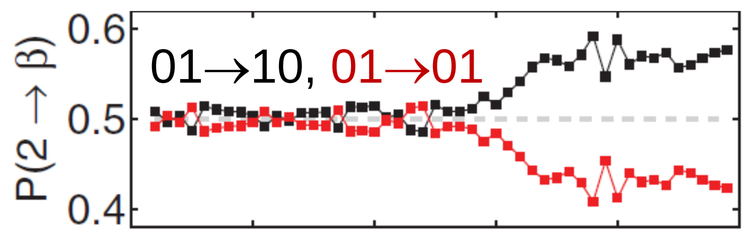


Consistent with stochastic dynamics at low pumps, but signatures of determinism at higher pump currents

"language" analysis: transition probabilities



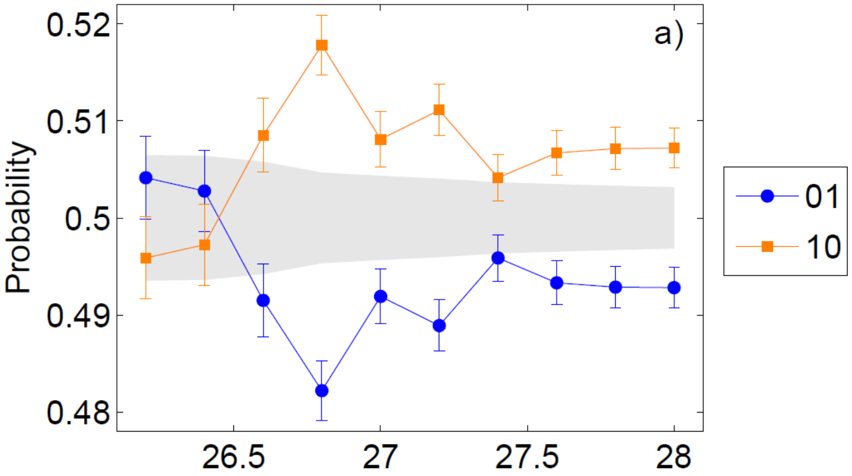




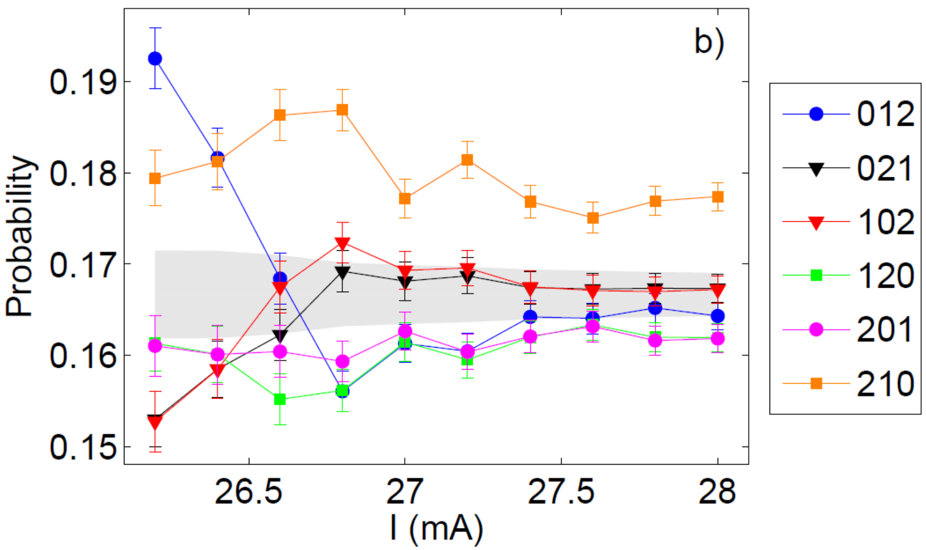
Consistent with stochastic dynamics at low pumps, but signatures of determinism at higher pump currents.

But at low pump currents: inter-dropout-intervals not fully random

D=2



D=3



A. Aragoneses et al, to appear in Scientific Reports (2013)

Two or more time series X, Y

- Ordinal analysis can be used to quantify similarity, infer coupling and directionality from time series.

- A key concept: the mutual information.

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Physics Reports 441 (2007) 1–46
PHYSICS REPORTS
www.elsevier.com/locate/physrep

Causality detection based on information-theoretic approaches in time series analysis

Katerina Hlaváčková-Schindler[a,*], Milan Paluš[b], Martin Vejmelka[b], Joydeep Bhattacharya[a,c]

BIOINFORMATICS
Vol. 18 Suppl. 2 2002
Pages S231–S240

The mutual information: Detecting and evaluating dependencies between variables

R. Steuer[1], J. Kurths[1], C. O. Daub[2], J. Weise[2] and J. Selbig[2]

Similarity measure: mutual information

- Joint entropy: $H(X,Y) = -\sum_{i=1}^{m_X}\sum_{j=1}^{m_Y} p(x_i, y_j)\log p(x_i, y_j)$

- If X and Y are independent: H(X,Y) = H(X) + H(Y)

- Mutual Information: MI(X,Y) = H(X) + H(Y) – H(X,Y)

$$MI(X,Y) = \sum_{i=1}^{m_X}\sum_{j=1}^{m_Y} p(x_i, y_j)\log\frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

- It reflects the reduction in uncertainty of one variable by knowing the other one.

- X and Y are independent ⇔ MI(X,Y) = 0.

- This does not hold for the cross-correlation.

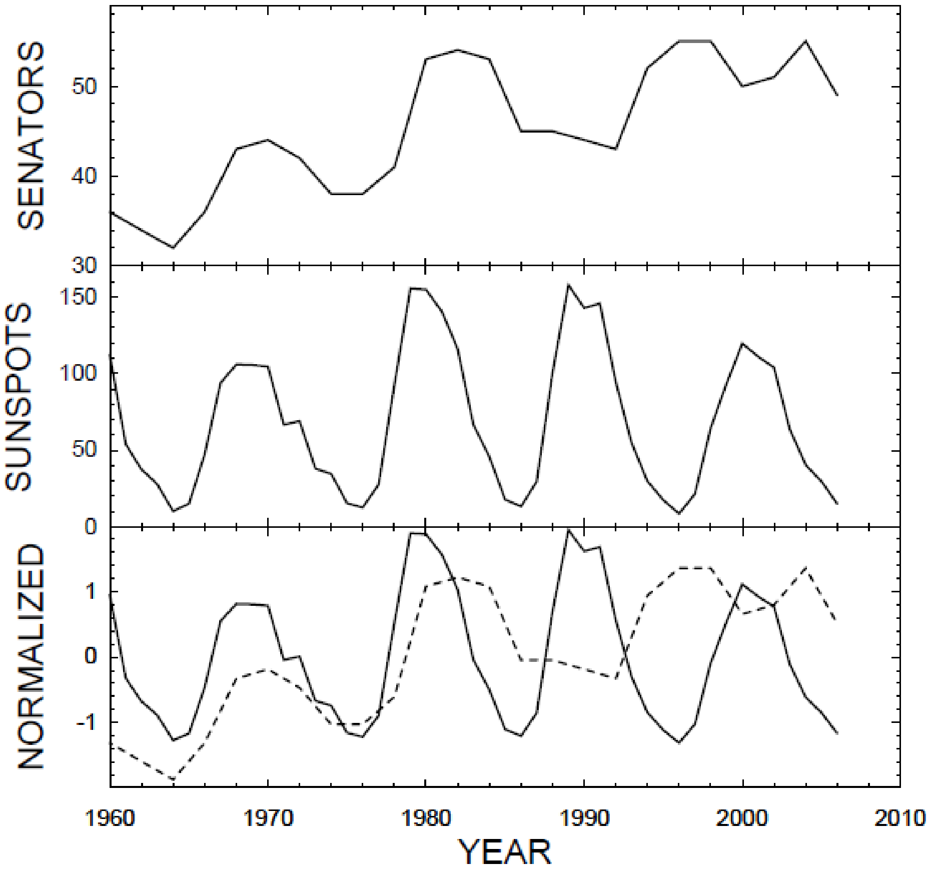
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Problem with cross-correlation

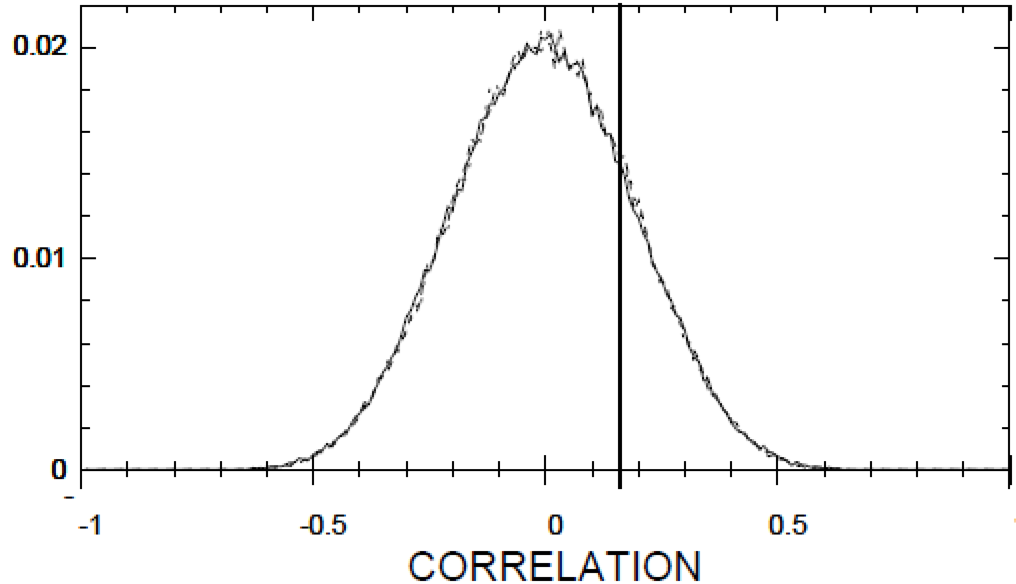
$$C_{ij} = \sum_{t=1}^{N} x_i(t)\, x_j(t)$$ (x_i, x_j normalized to zero mean and unit standard deviation)

- Illustrative example: The number of Republicans in the U.S. Senate and the sunspot number in the period 1960-2006.



An illustrative example: number of sunspots and senators

- Distribution of CC values
- Between the number of the Republican senators in the period 1960-2006 (24 samples) with 24-sample sets randomly drawn from the Gaussian distribution (dashed);
- between the number of the Republican senators in the period 1960-2006 (24 samples) with the 24-sample segment of the sunspot numbers randomly permutated in the temporal order (IID surrogate, dash-and-dotted)
- two 24-sample sets randomly drawn from a Gaussian distribution (solid).



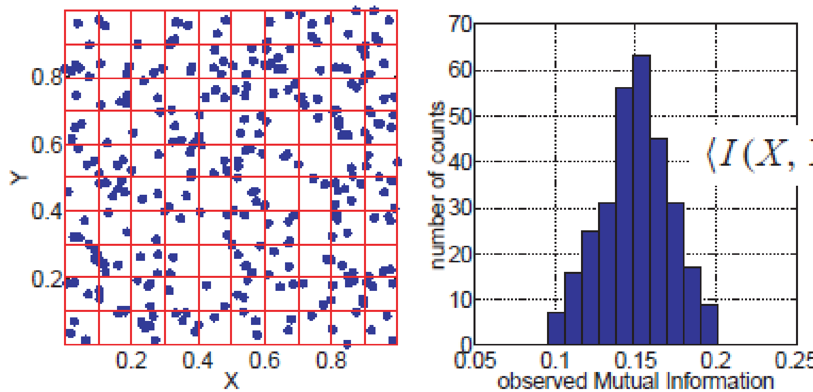
Vertical line: correlation between the number of the Republican senators and the sunspot numbers for the period 1960-2006.

M. Palus, Contemporary Physics 48, 307 (2007)

Problem with mutual information

- X and Y are independent ⇔ MI(X,Y) = 0.

- OK. But: Computing probabilities from histograms give MI values that fluctuate or are systematically overestimated.



$$\langle I(X,Y)^{estimated} \rangle \approx 0.15 \pm 0.02$$

Fig. 1. Naive estimation of the mutual information for finite data. Left: The dataset consists of $N = 300$ artificially generated independent and equidistributed random numbers. The probabilities are estimated using a histogram which divides each axis into $M_x = M_y = 10$ bins. Right: The histogram of the estimated mutual information $I(X,Y)$ obtained from 300 independent realizations.

R. Steuer et al, Bioinformatics 18, suppl 2, S231 (2002).

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

An example from neural coding

Invited Review

J Neurophysiol 98: 1064–1072, 2007.
First published July 5, 2007; doi:10.1152/jn.00559.2007.

Correcting for the Sampling Bias Problem in Spike Train Information Measures

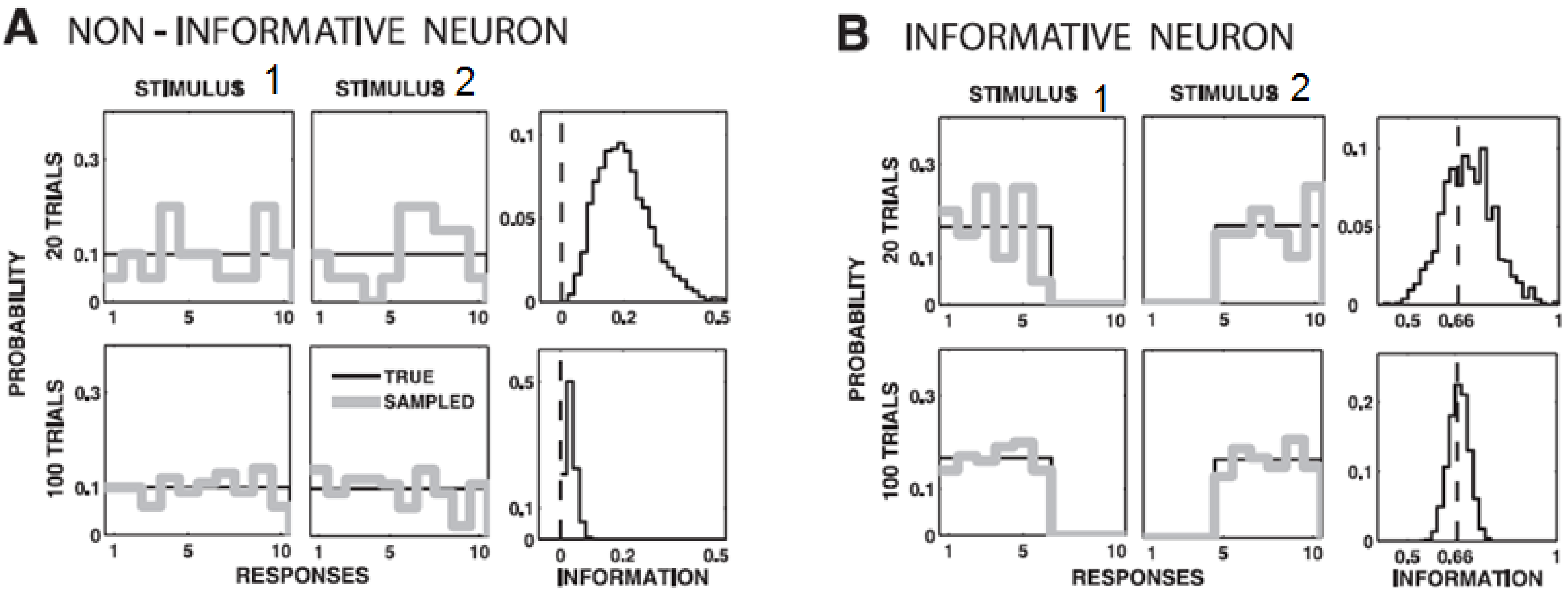
Stefano Panzeri, Riccardo Senatore, Marcelo A. Montemurro, and Rasmus S. Petersen

Numerical simulation of neuronal response to 2 different stimulus:

- Non-informative neuron: fires (with uniform probability) 1-10 regardless of the stimulus

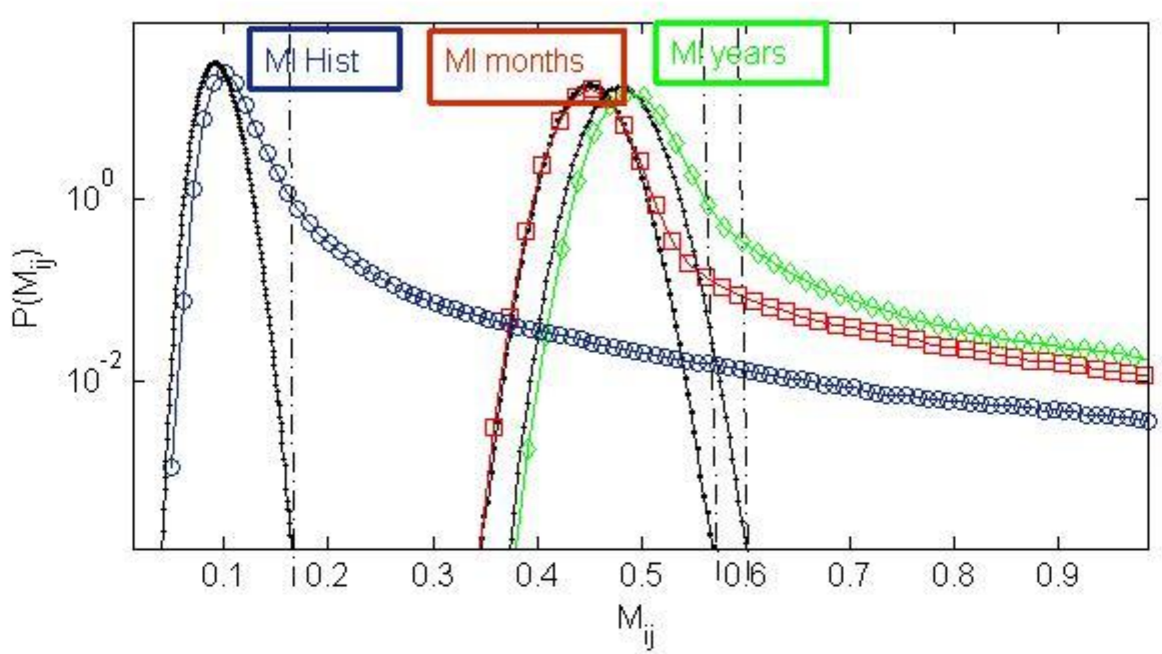
- Informative neuron: fires (with uniform probability)
  - 1–6 spikes to stimulus 1
  - 5–10 spike to stimulus 2.

An example from neural coding



- Distribution (5,000 simulations) of the MI values obtained with 20 (top) and 100 (bottom) trials per stimulus respectively. As the number of trials increases, both the information bias and the dispersion decrease. The dashed line indicates the true MI value.

Another example: monthly-averaged SAT anomalies



- MI values for real and surrogated and data.
- problem for identifying weak significant links.

Directionality: conditional mutual information

$$MI(X,Y) = H(X) + H(Y) - H(X,Y)$$

- Conditional mutual information: $CMI(X,Y|Z) = H(X|Z)+H(Y|Z)-H(X,Y|Z)$
- If X and Y are independent of Z: $CMI(X,Y|Z) = MI(X,Y)$
- We want to estimate the net information concerning the future of the process X_1 that is contained within the process X_2

$$\Delta X_1 = [\ldots x_1(t) - x_1(t+\tau), x_i(t+1) - x_i(t+1+\tau), \ldots]$$

- $I_{21} = CMI(X_2, \Delta X_1 | X_1)$
- $I_{21}=0$: there is no information in X_2 about the future of X_1

PCMI = Permutation conditional mutual information

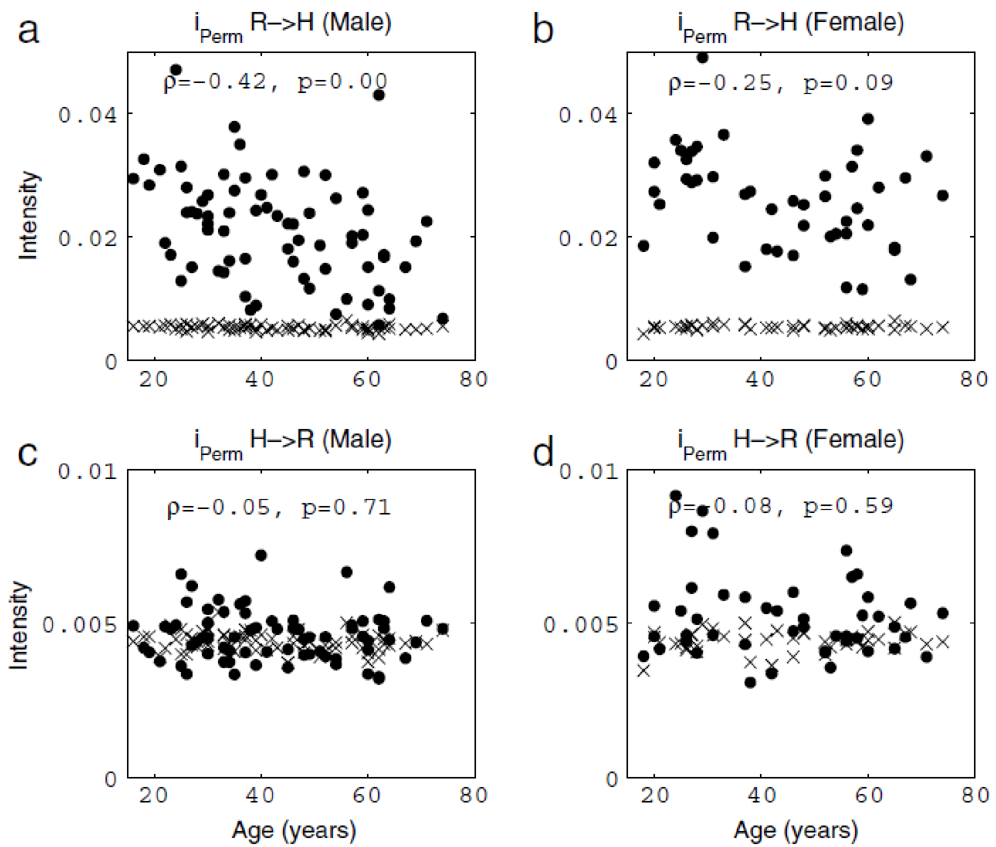
Example: the cardio-respiratory interaction

- X_1 = phase of the heart
- X_2= phase of the respiration

$$i(X_2 \to X_1) = \frac{1}{N} \sum_{\tau=\tau_{min}}^{\tau_{max}} I_{12}(\tau)$$

$$N = 46, \tau_{min} = 5 \text{ and } \tau_{max} = 50$$

- dots: real data (averaged for each patient, using two methods for computing pdfs: 8-bin histograms and D=4 OPs)
- x: the same, surrogate data



Y. Shiogai et al. / Physics Reports 488 (2010) 51–110

(Permutation) Directionality index

PRL **100**, 084101 (2008) PHYSICAL REVIEW LETTERS week ending 29 FEBRUARY 2008

Direction of Coupling from Phases of Interacting Oscillators: A Permutation Information Approach

A. Bahraminasab,[1] F. Ghasemi,[2] A. Stefanovska,[1] P. V. E. McClintock,[1] and H. Kantz[3]

PHYSICAL REVIEW E **84**, 021929 (2011)

Characterization of the causality between spike trains with permutation conditional mutual information

Zhaohui Li,[1] Gaoxiang Ouyang,[2] Duan Li,[1] and Xiaoli Li[2,3,*]

$$I^{\delta}_{X\to Y} = I(X;Y_\delta|Y) = H(X|Y) + H(Y_\delta|Y) - H(X,Y_\delta|Y),$$

$$I^{\delta}_{Y\to X} = I(Y;X_\delta|X) = H(Y|X) + H(X_\delta|X) - H(Y,X_\delta|X)$$

$$D^{P}_{X\to Y} = \left(\frac{I^{\eta}_{X\to Y} - I^{\eta}_{Y\to X}}{I^{\eta}_{X\to Y} + I^{\eta}_{Y\to X}}\right)$$

$D^{P}_{X\to Y} > 0$ means that S_X drives S_Y

$D^{P}_{X\to Y} < 0$ means that S_Y drives S_X

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Coupled chaotic systems: symbolic analysis to characterize the synchronization transition

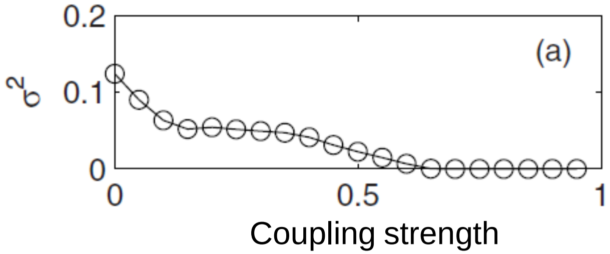
$$x_i(t+1) = f(x_i(t)) + \frac{\varepsilon}{b_i}\sum_{j=1}^{N} A_{ij}\left[f(x_j(t-\tau_{ij})) - f(x_i(t))\right]$$

- With sufficiently distributed delays: steady-state synchronization.

$$x_i(t) = x_j(t) = x_0 = f(x_0) \quad \forall i, j, t$$

- A popular synchronization quantifier:

$$\sigma^2 = \left\langle \sum_{i=1}^{N}\left(x_i(t) - \langle x\rangle_s\right)^2 \right\rangle_t$$



- is a “global” indicator: it provides no information about the microscopic local dynamics in the nodes.

- Alternative: analyze the transition to synchronization in terms of the diversity of the symbolic “language” of the nodes.

Symbolic synchronization

- In each network node, compute the transition probability from word α to word β

$$P_i(\alpha,\beta)=\frac{\sum_{t=1}^{N} n(s(t)=\alpha,\, s(t+1)=\beta)}{\sum_{t=1}^{N} n(s(t)=\alpha)}$$

  where n is a count of the number of occurrences in node i.

- Then, the "language" diversity can be quantified in terms of the heterogeneity of the probabilities among the nodes:

$$\xi^2(\alpha,\beta)=\left\langle \sum_{i=1}^{N}\left(P_i(\alpha,\beta)-\left\langle P_i(\alpha,\beta)\right\rangle_s\right)^2 \right\rangle_t \qquad \sigma^2=\left\langle \sum_{i=1}^{N}\left(x_i(t)-\langle x\rangle_s\right)^2 \right\rangle_t$$

- This provides a quantifier for each TP ($\alpha\to\beta$).

Two clusters before synchronization when the coupling strength increases

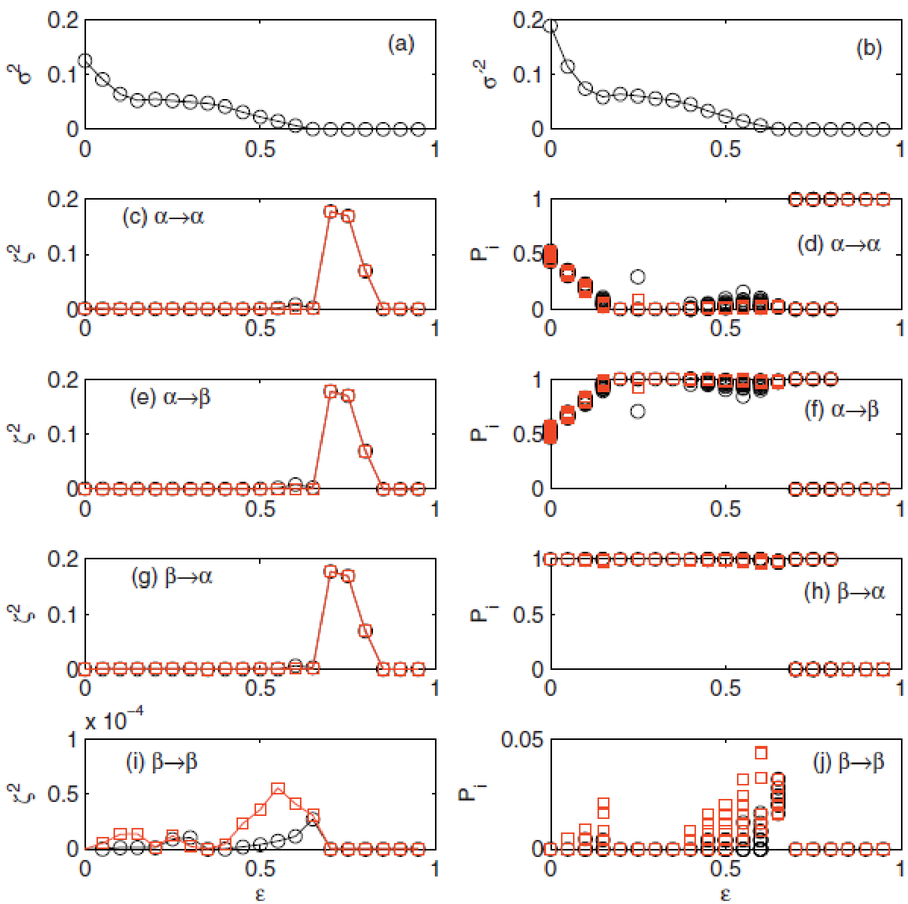
$$\sigma^2 = \frac{1}{N}\left\langle \sum_{i=1}^{N} (x_i(t) - \langle x \rangle_s)^2 \right\rangle_t \qquad \sigma'^2 = \frac{1}{N}\left\langle \sum_i (x_i - x_0)^2 \right\rangle_t$$

N=200

D=2

o: bin

□: OP



Transition probabilities in 20 randomly selected nodes

C. Masoller

Masoller & Atay, EPJD 62, 119 (2011)

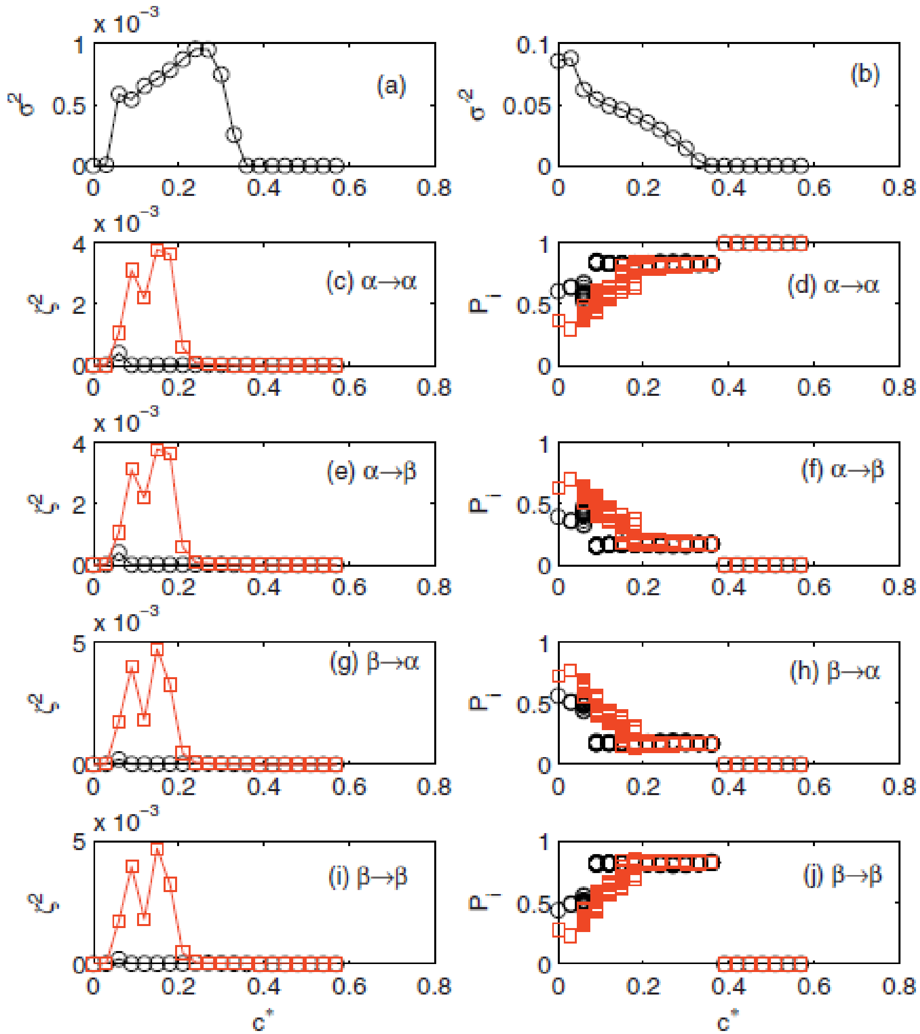
No clustering before synchronization when the delay heterogeneity increases

N=200

D=2

o: bin

□: OP



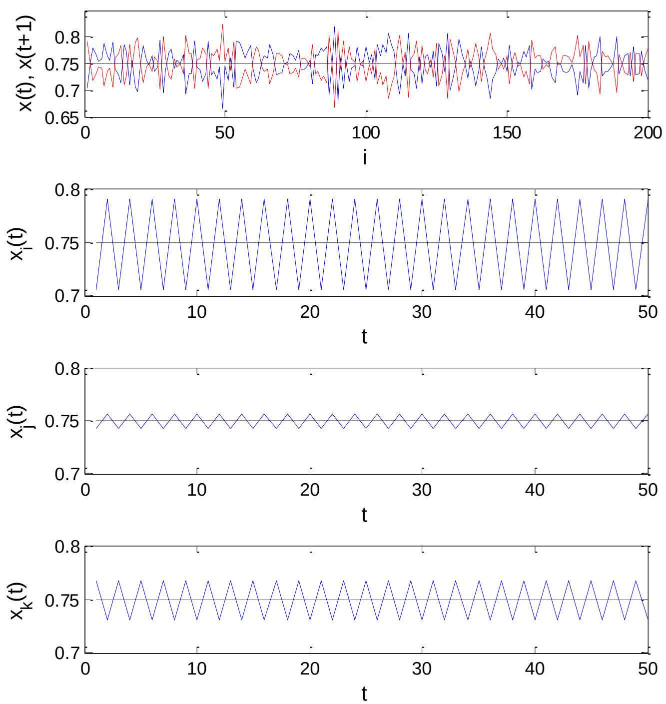
Transition probabilities in 20 randomly selected nodes

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

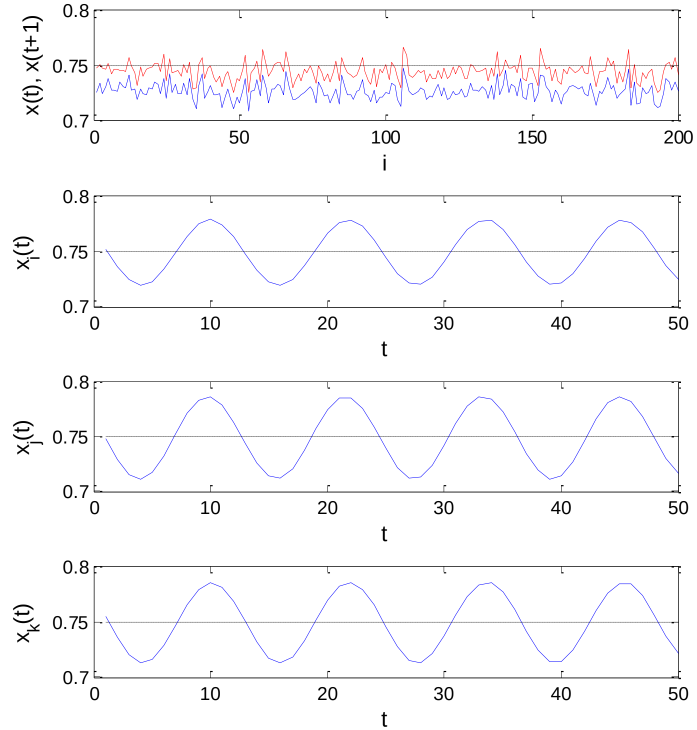
Campus d'Excel·lència Internacional

Network visualization before the onset of synchronization

The coupling strength is close to synchronization



The delay heterogeneity is close to synchronization



Network configuration at two consecutive times and time-evolution of three randomly selected nodes

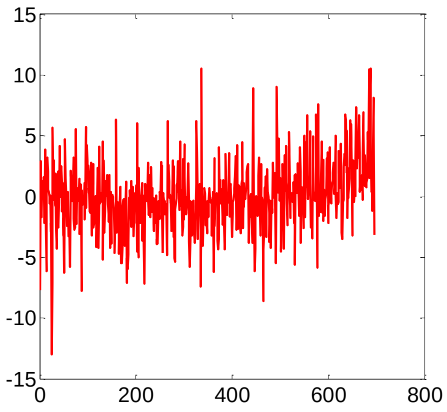
Outline

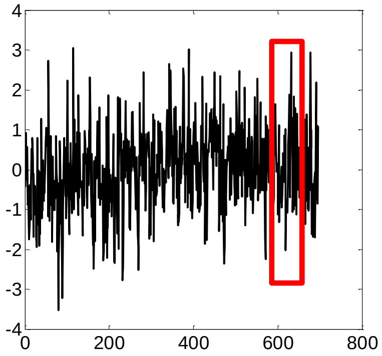
- Symbolic time series analysis
- Ordinal analysis
- Information theory measures
- Examples
- Application to climate data analysis

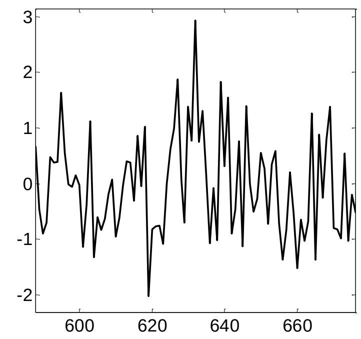
Ordinal pattern analysis of climatological data

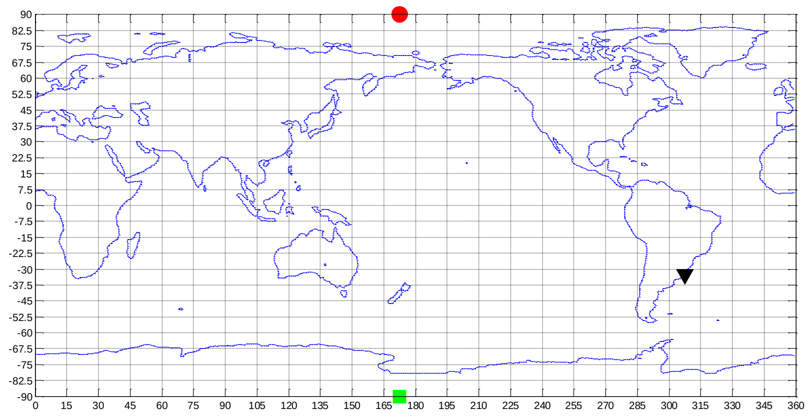
- **The data**: monthly-averaged surface air temperature anomalies (SATA).
  - Reanalysis data from National Center for Environmental Prediction, National Center for Atmospheric Research (NCEP-NCAR, USA).
  - Regular grid of nodes covering the Earth's surface with resolution 2.5 x 2.5 (about 250 kms by 250 kms in the equator): **10,226 nodes**.
  - January 1949 -- December 2006: in each node we have **696 data points** (58 years x 12 months).
- Network **representation**:
  - Area-weighted connectivity: plot of the % of the Earth each node is connected to (no information about the connections).
  - Connections of one node with all the other nodes: plot of cross-correlation (CC) or MI values.

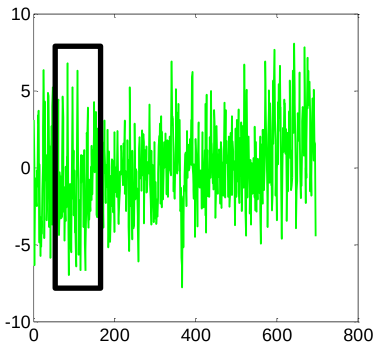
Visual inspection of time series (monthly averaged SAT anomalies)

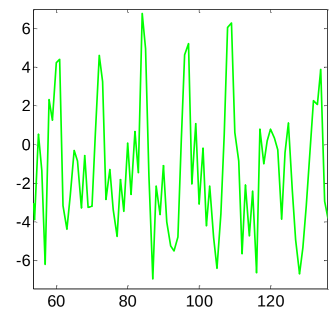












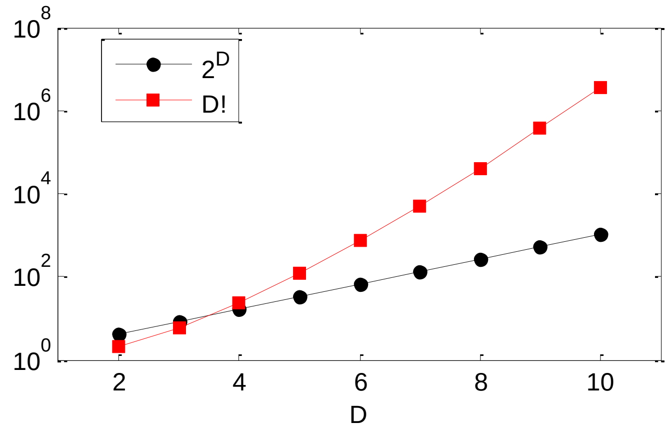
Similarity measures used to construct climate networks

- CC: computed from SATA values

- MI: computed from histogram of SATA values.

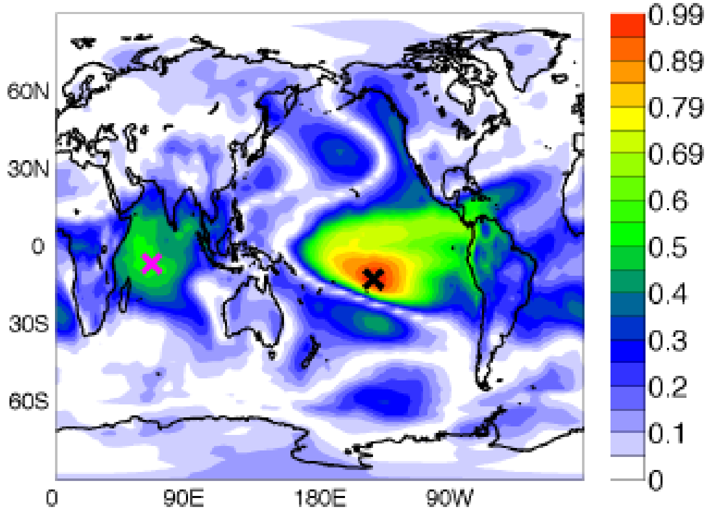
- MI: computed from ordinal patterns (D=4, D=5 lag=1 allows to consider consecutive months; lag=4 allows to consider 1 year period, lag=12 allows to consider consecutive years).

- MI: also computed from binary representations (SATA, one threshold at x_{th} =0), allows to consider words with more letters.

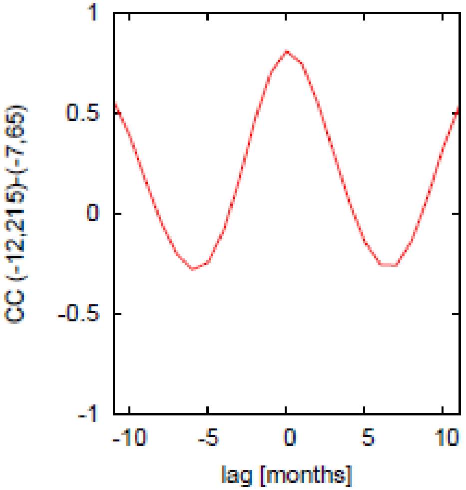


An example of cross correlation plot

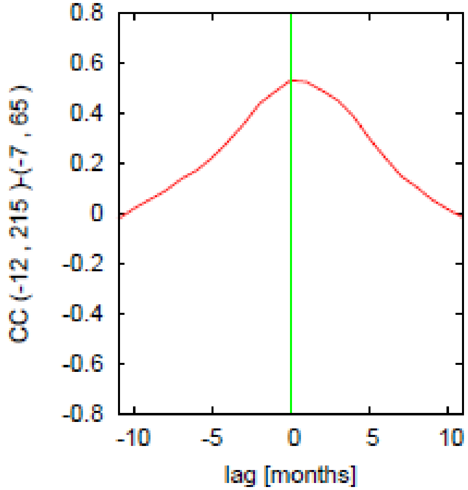
$$C_{i\,j}(\tau) = \sum_{t=1}^{N} x_i(t+\tau)\, x_j(t)$$



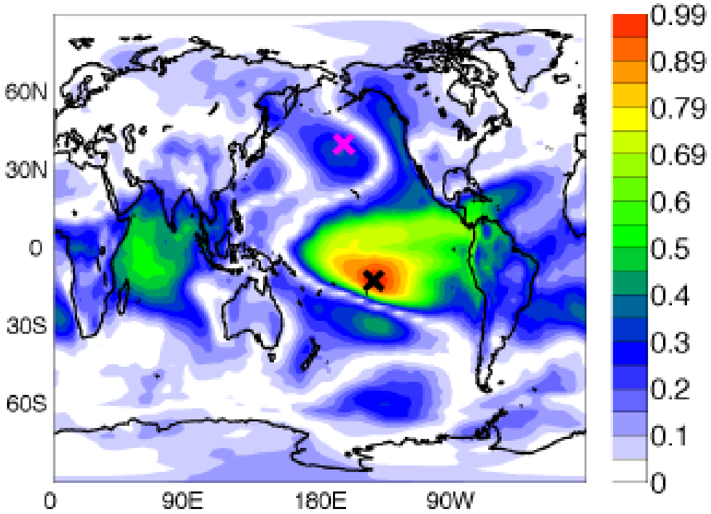
SAT

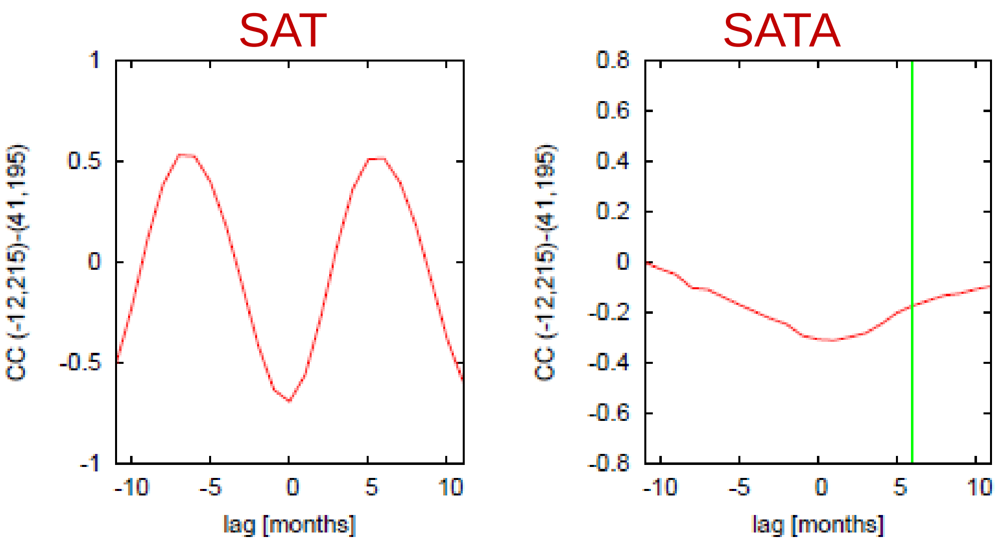


SATA



Cross correlation plots

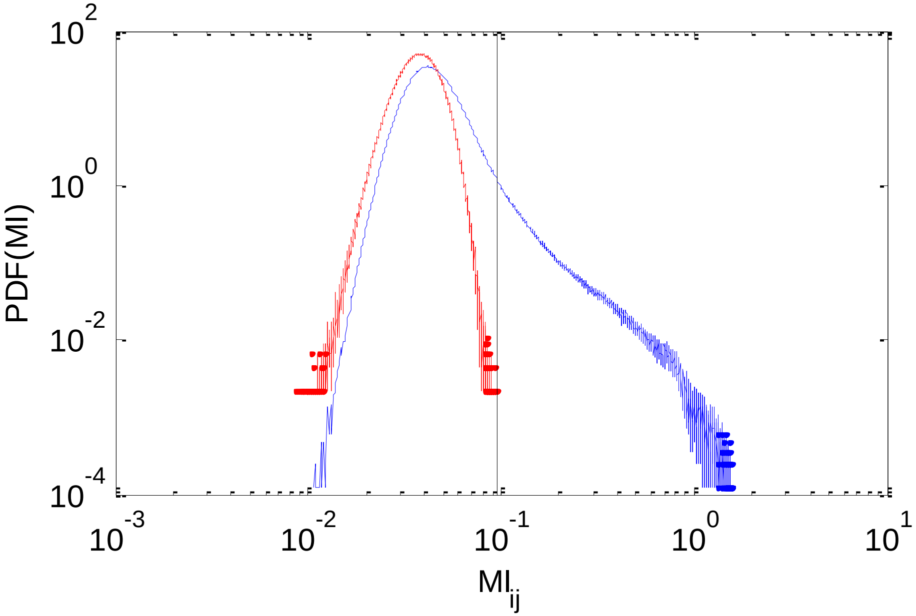




In what follows: CC/MI computed with zero-lag. For lag-times effects: Giulio's presentation

Thresholding

- Simplest option: consider that links are statistically significant if similarity values (CC/MI) are larger than those obtained with surrogated shuffled data.



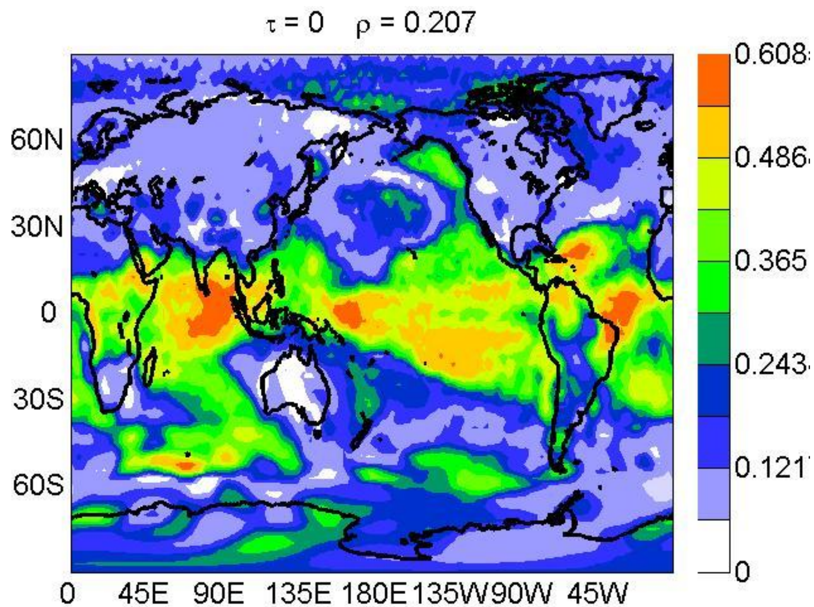
PDF computed with 10,226x10,226 values. Similar results when using a local threshold for each node (from PDFs computed with 10,226 values).

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

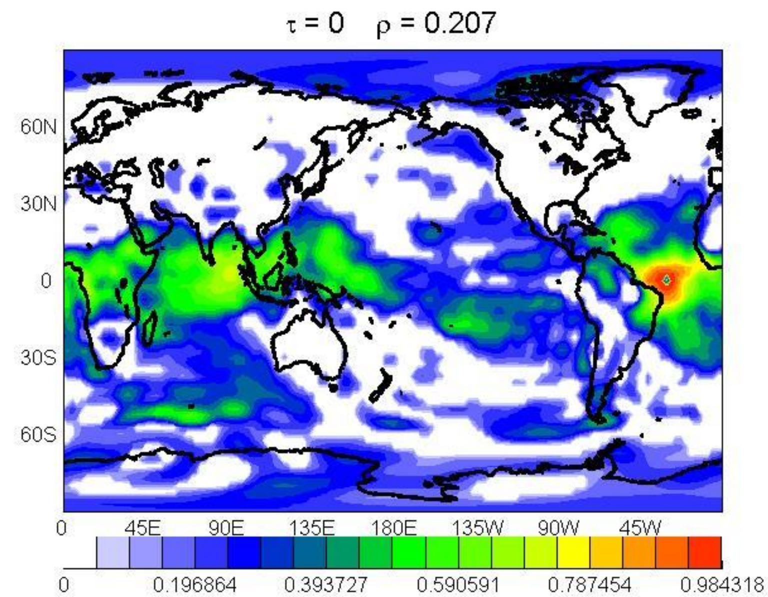
Campus d'Excel·lència Internacional

Climate network constructed with CC and keeping all the significant links

AWC

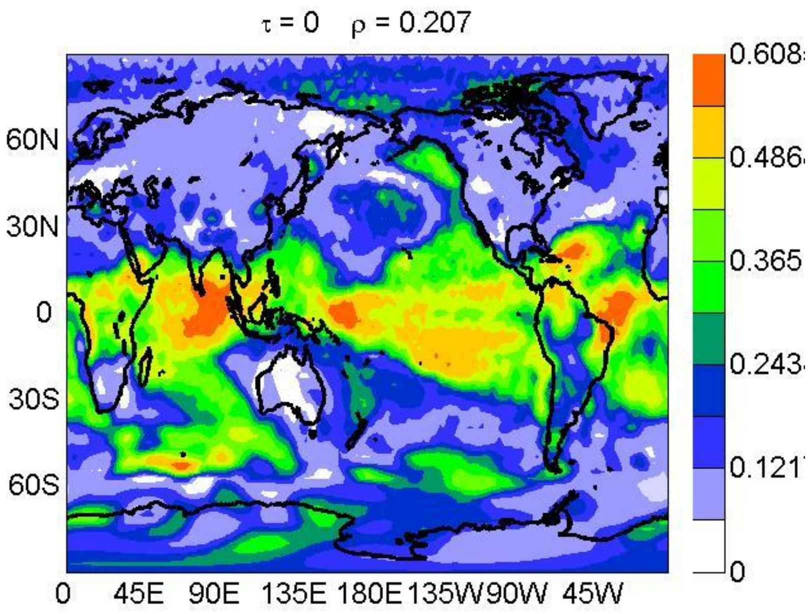


CC values

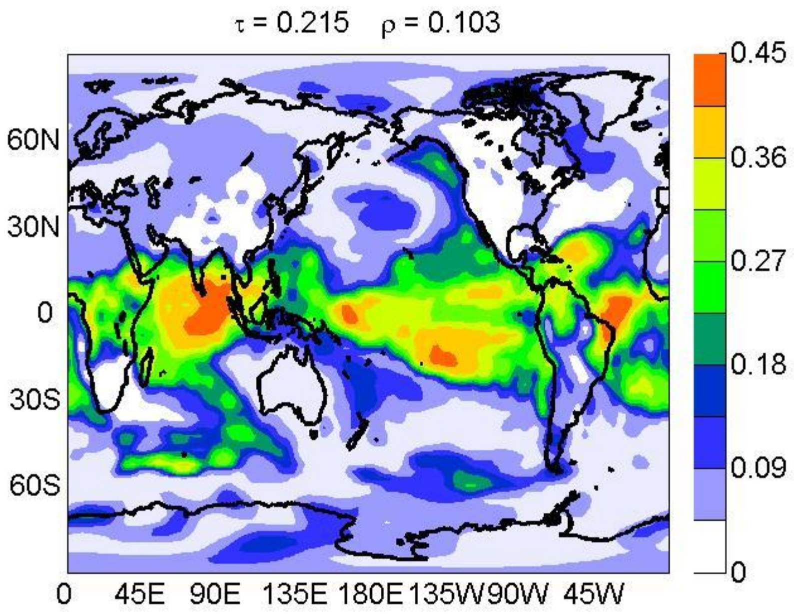


Climate network constructed with CC

AWC (all significant)

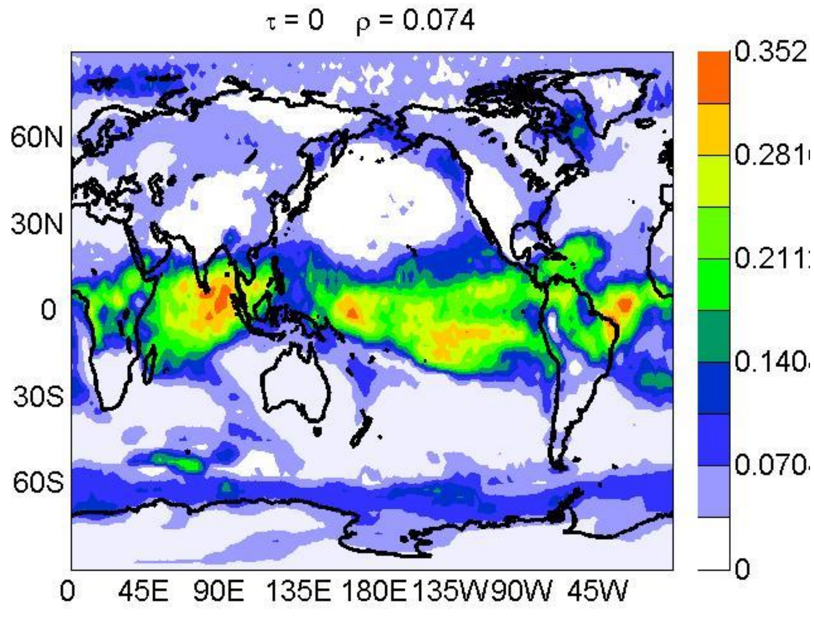


10% strongest links

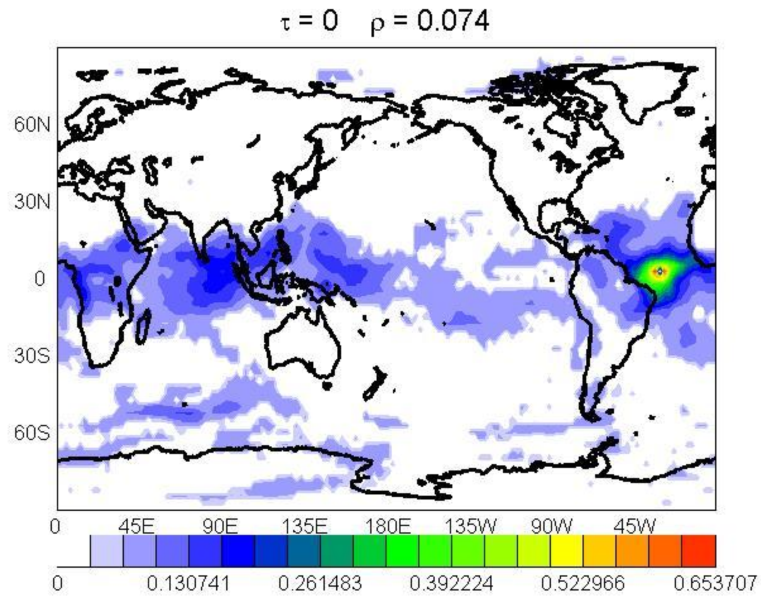


MI (histogram of anomaly values) and keeping all the significant links

AWC

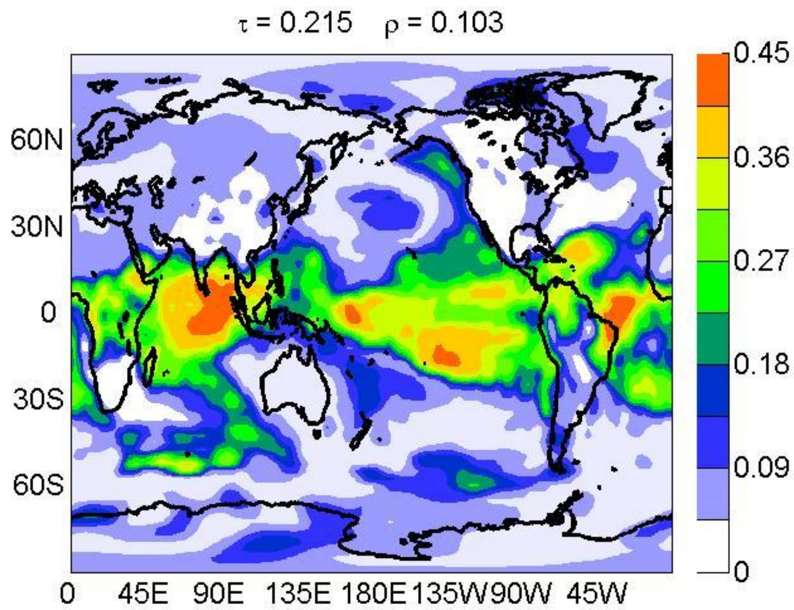


MI values

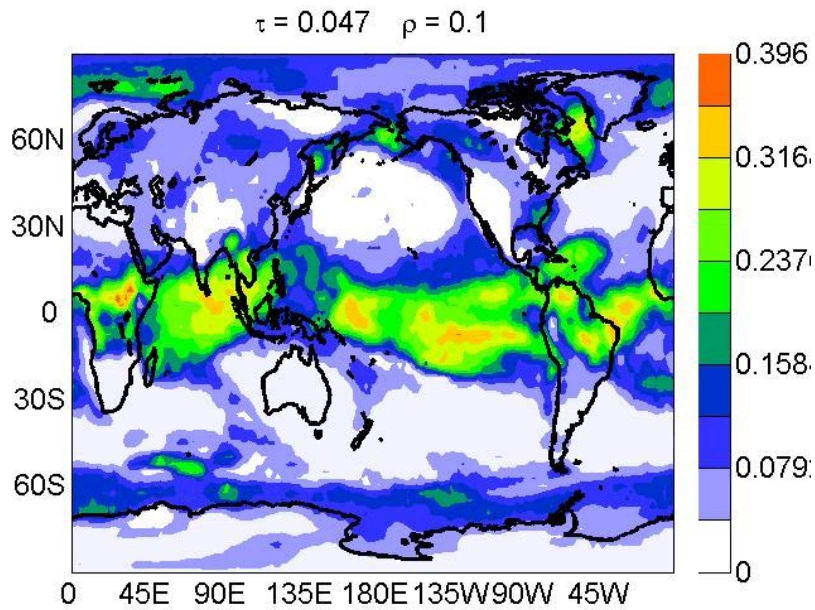


Comparison CC – MI at 10% density

AWC (CC)



AWC (MI)

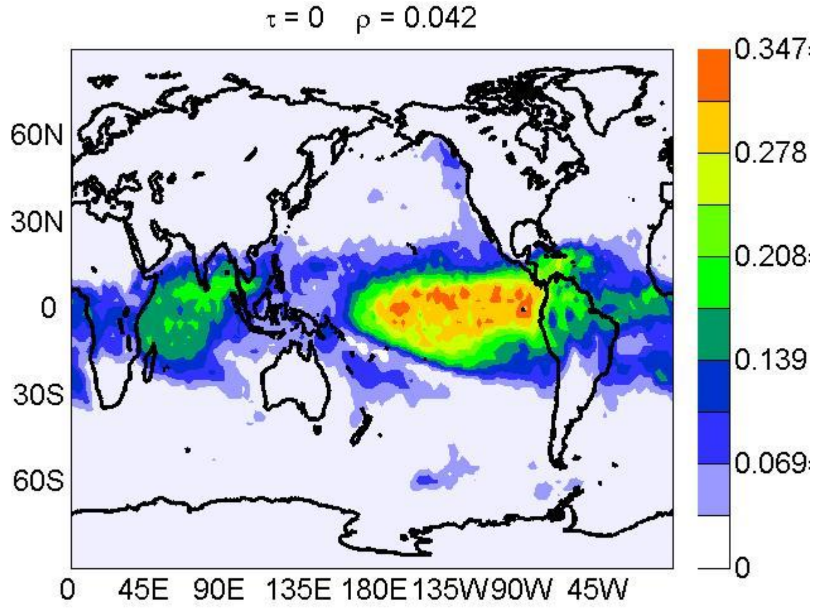


MI (OPs 4 years) and keeping all the significant links

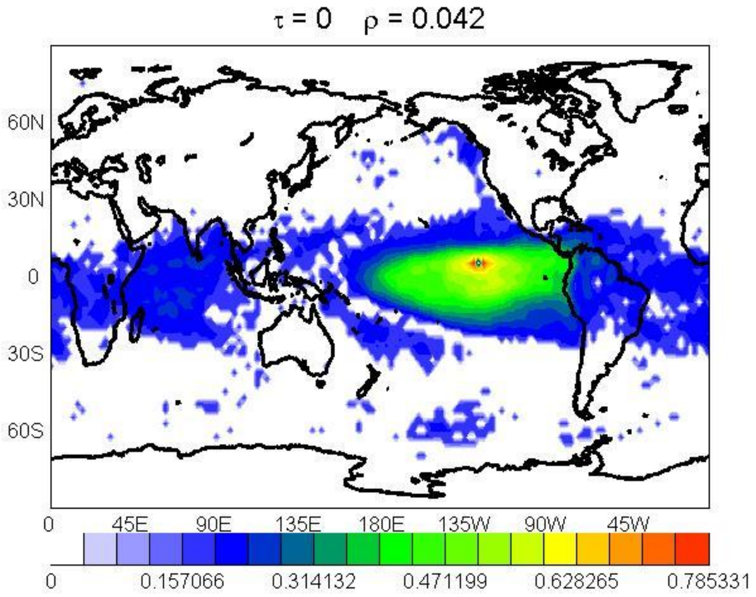
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

AWC



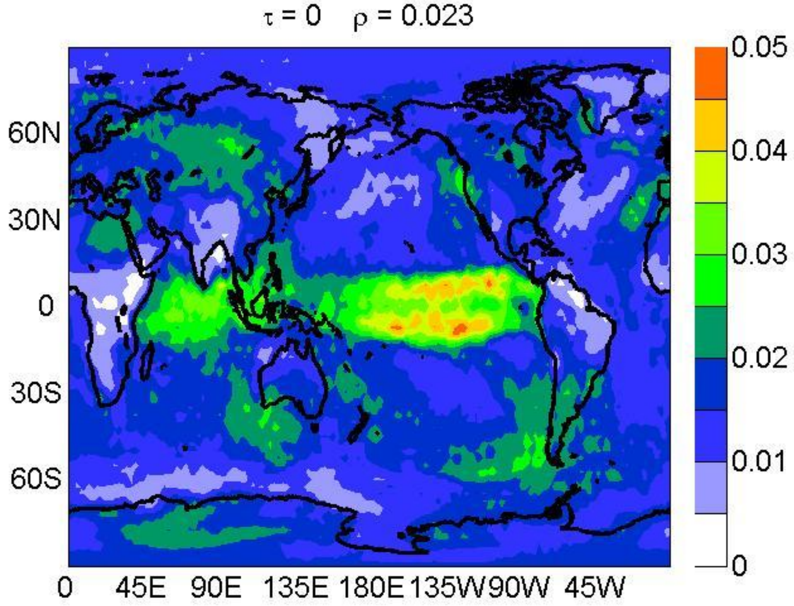
MI values



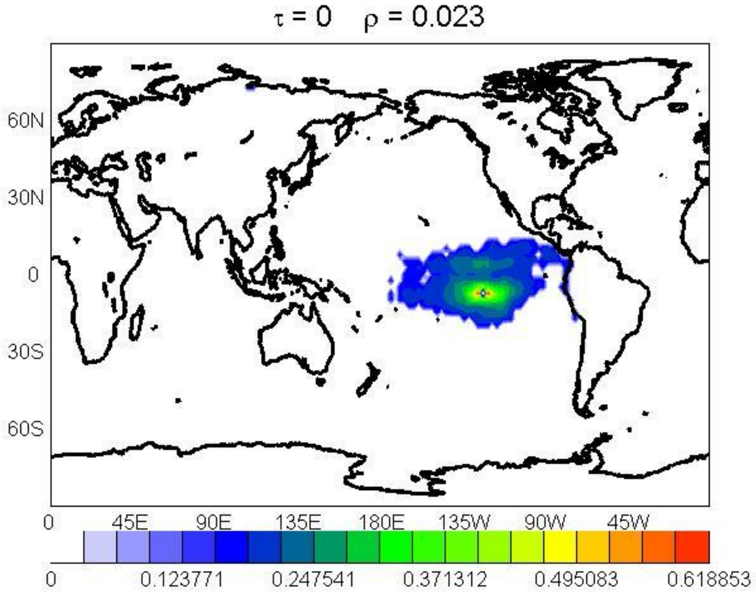
Barreiro, Martí and Masoller, Chaos 21, 013101 (2011)

MI (OPs 4 months) and keeping all the significant links

AWC

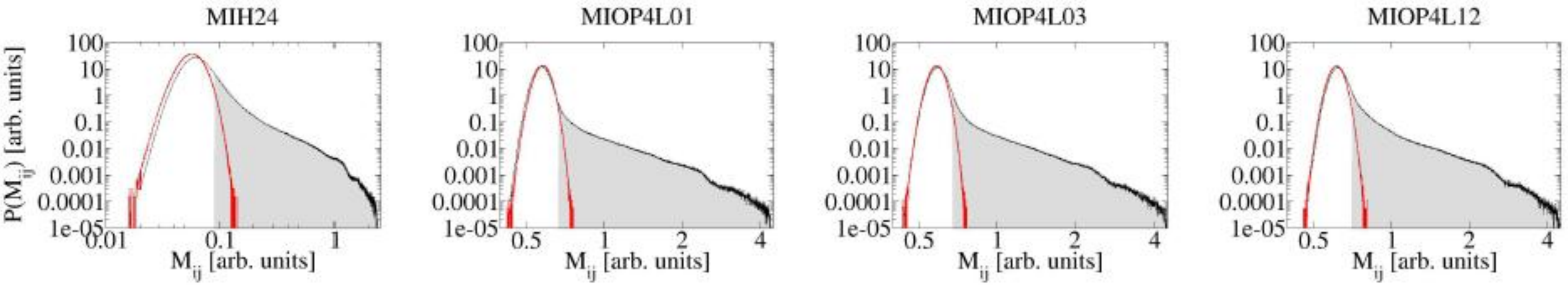


MI values



The results are robust when a more tolerant threshold is used

$$\tau = \mu + 3\sigma$$



Networks with higher link density are obtained, which display a richer pattern of teleconections: Ignacio's presentation.

The link significance should be carefully examined in order to avoid disregarding weak but significant links.

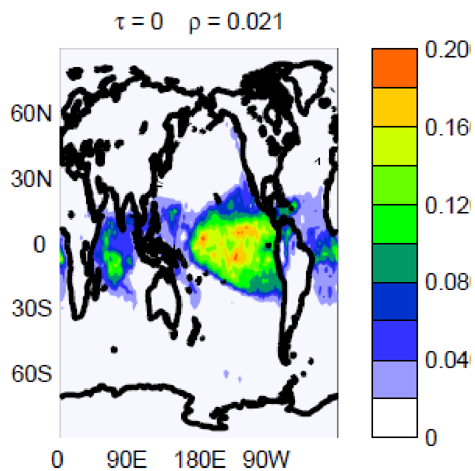
MI (BINARY consecutive years)

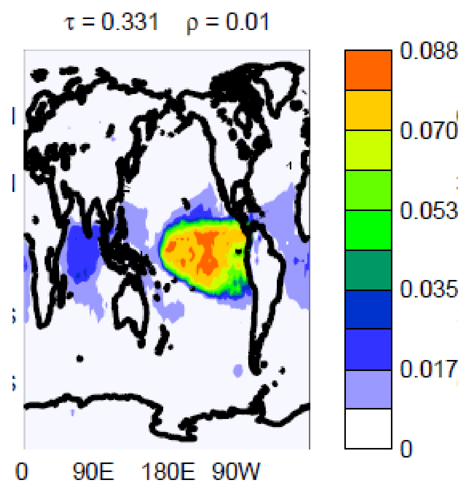
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

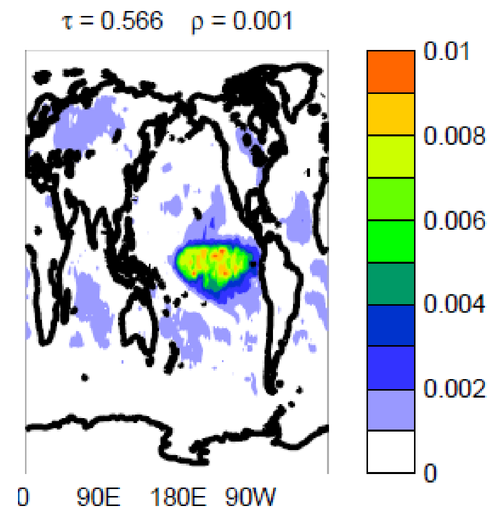
Campus d'Excel·lència Internacional

D=5

2^5=32

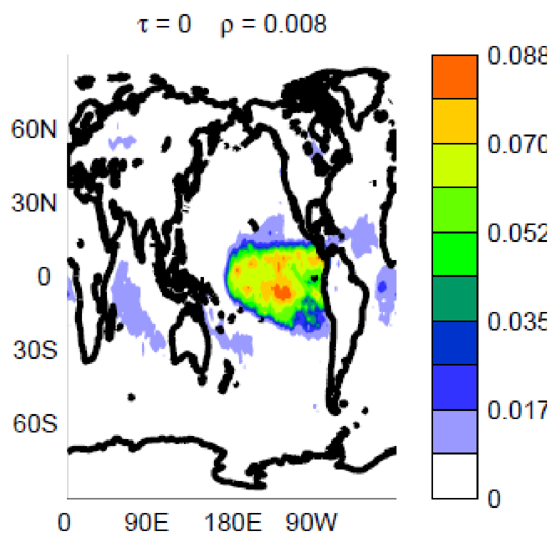


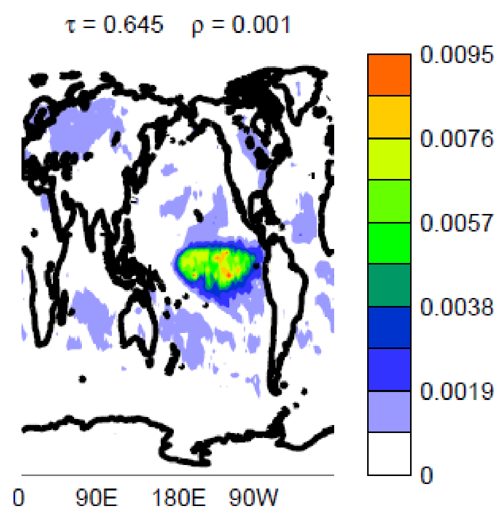




D=6

2^6=64



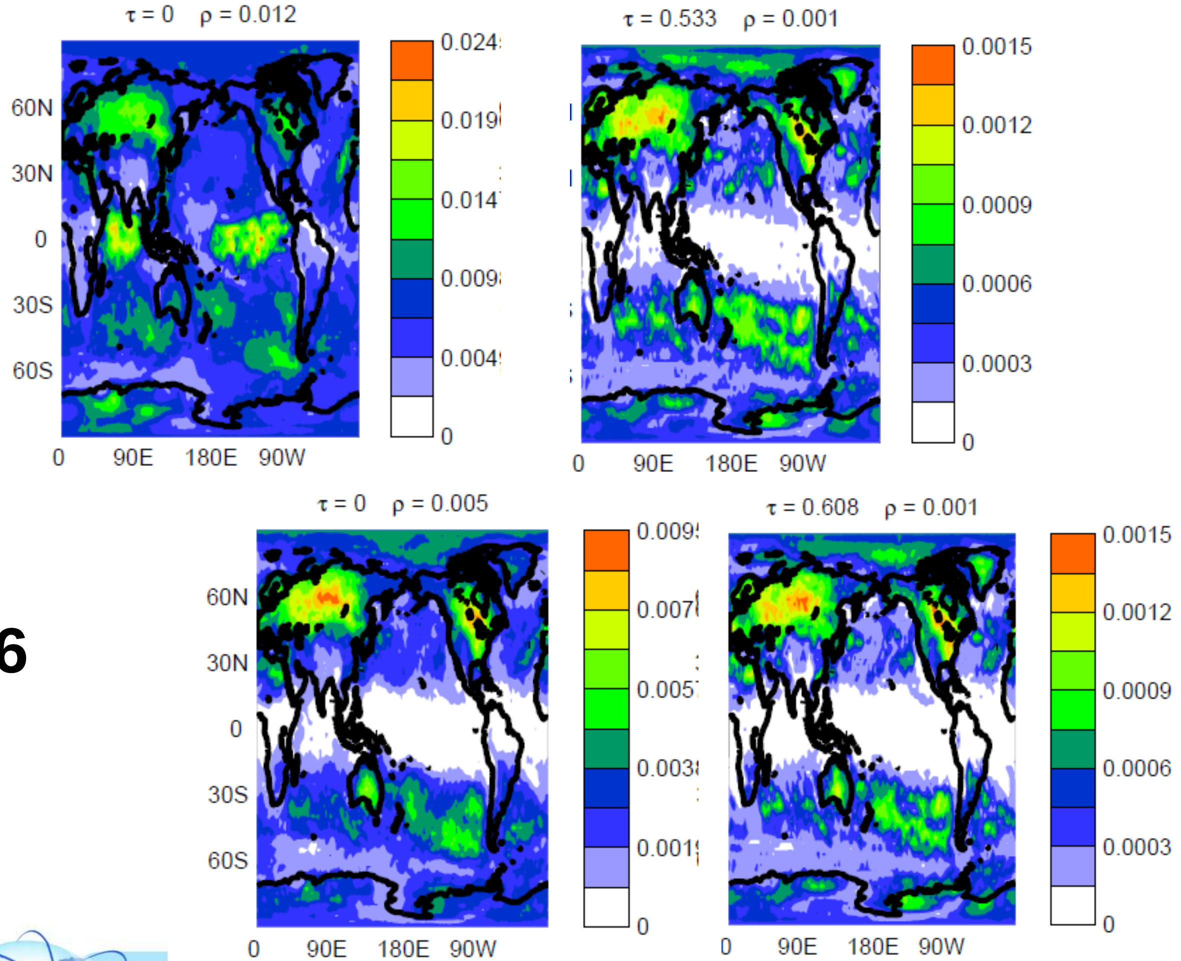


MI (BINARY consecutive months)

Campus d'Excel·lència Internacional

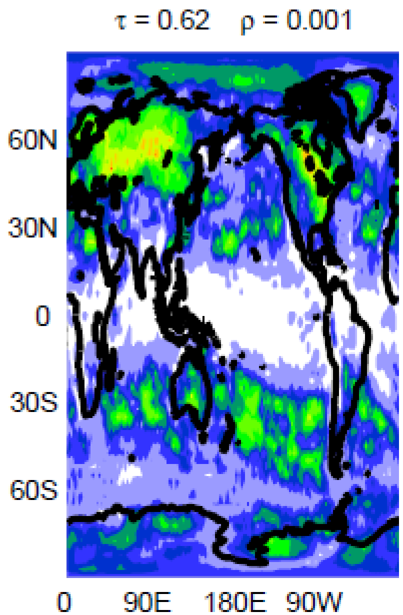
D=5

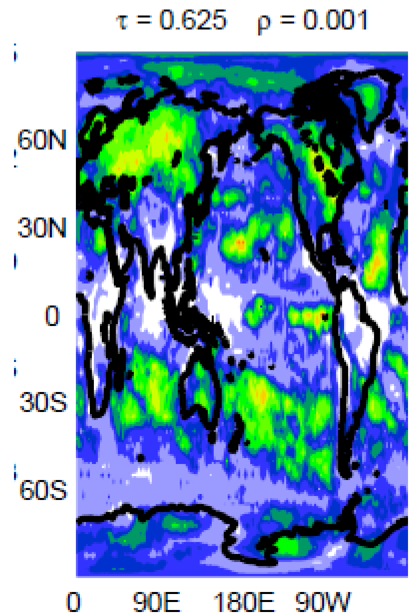
D=6

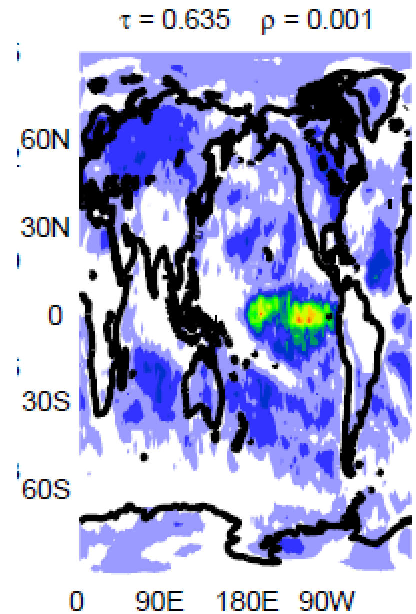


- 1% and 0.1% connectivity: very different networks.
- Stronger links (0.1%): the network is almost the same for D=5 and D=6.

MI (BINARY): influence of the pattern time-interval for fixed the length (D=6) and network density (0.1%)







6 months

$$[x_i(t), x_i(t+1), x_i(t+2),\ x_i(t+3), x_i(t+4), x_i(t+5)]$$

1 year

$$[x_i(t), x_i(t+2), x_i(t+4),\ x_i(t+6), x_i(t+8), x_i(t+10)]$$

2 years

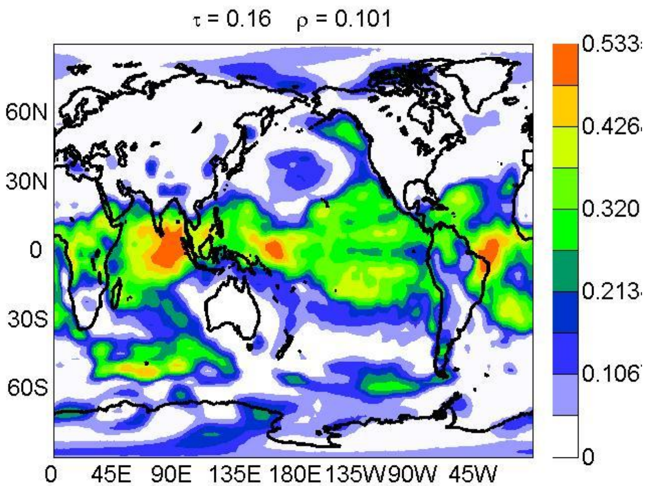
$$[x_i(t), x_i(t+4), x_i(t+8),\ x_i(t+12), x_i(t+16), x_i(t+22)]$$

Problem: visualization of the network via the AWC

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

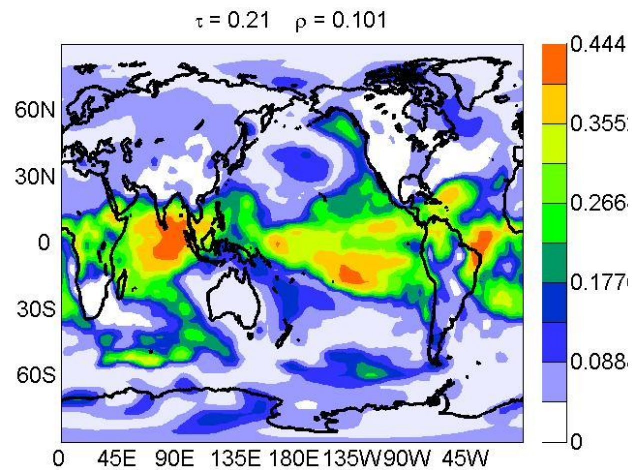
10%

τ_{ij} random in [0,11]



$$C_{i\,j}(\tau) = \sum_{t=1}^{N} x_i(t+\tau_{ij})\, x_j(t)$$

Zero-lag



$$C_{i\,j}(\tau) = \sum_{t=1}^{N} x_i(t)\, x_j(t)$$

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

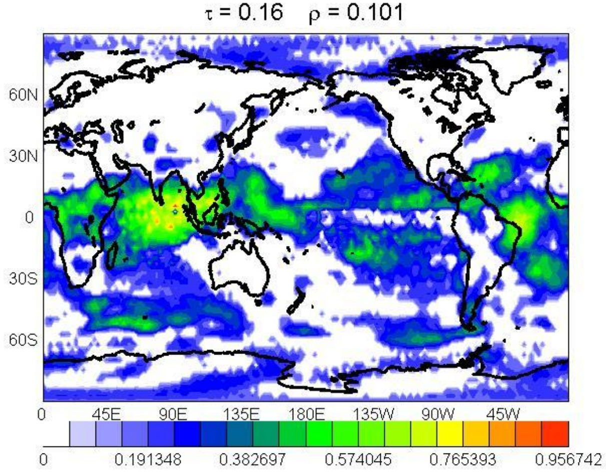
Campus d'Excel·lència Internacional

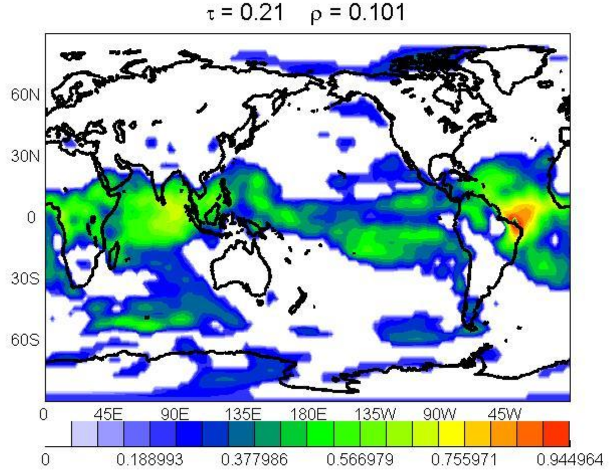
CC of the node with largest AWC ("hub")

τ_{ij} random in [0,11]

Zero lag

10%



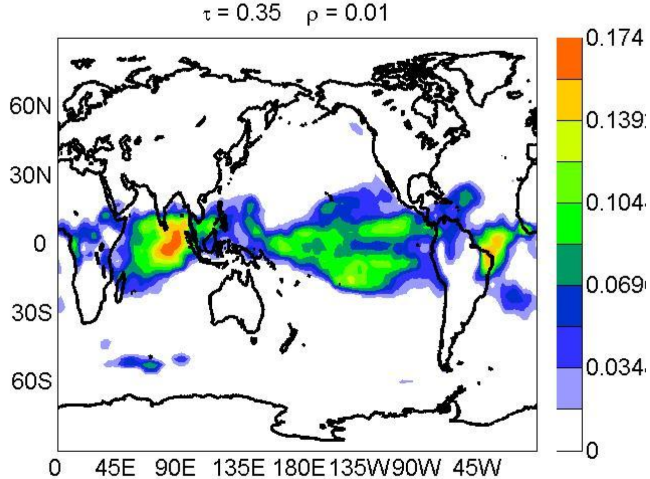


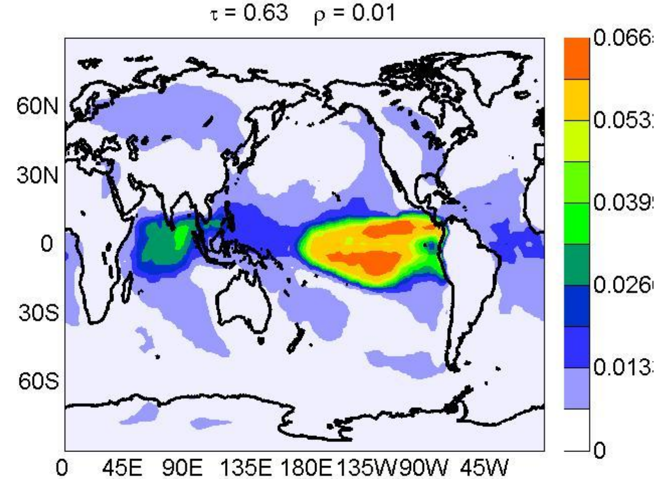
Considering the strongest links (1%)

τ_{ij} random in [0,11]

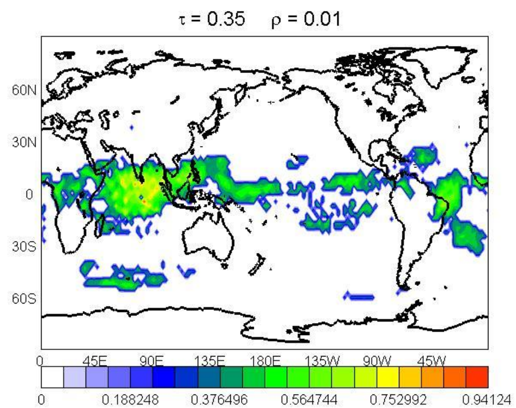
Zero-lags

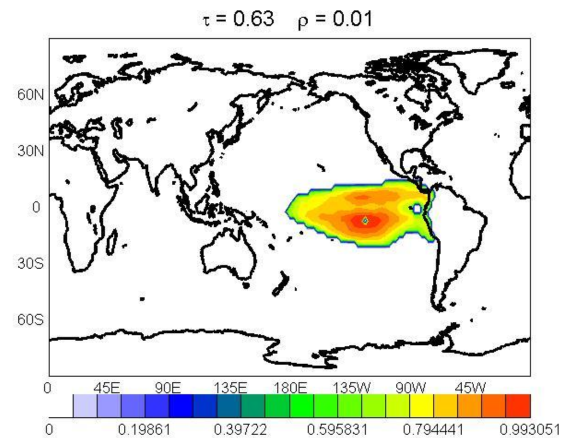
AWC





HUB





UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

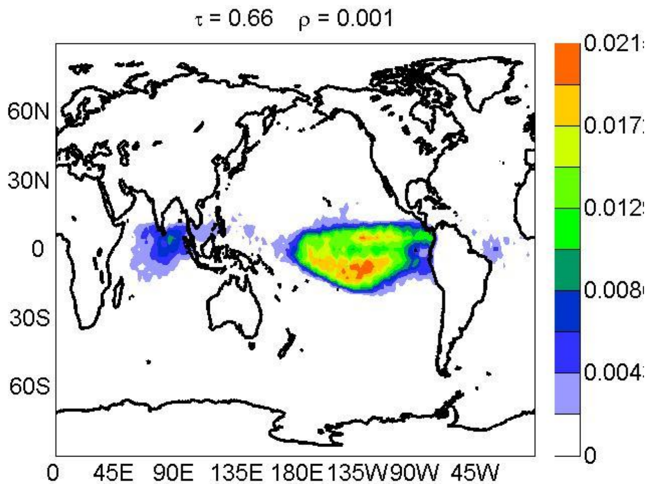
Campus d'Excel·lència Internacional

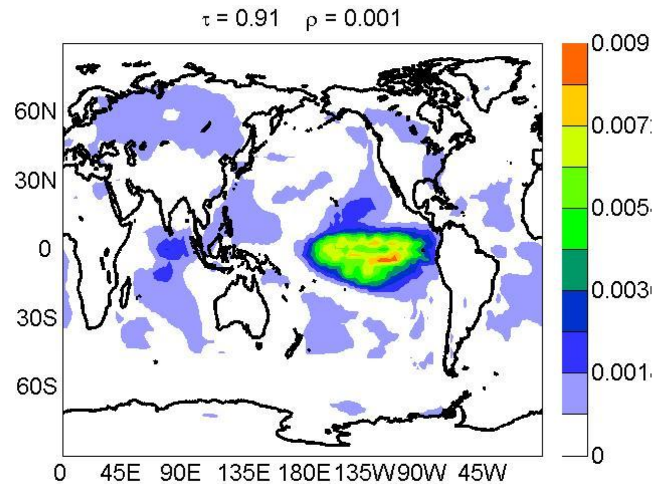
Even stronger (0.1%)

τ_{ij} random in [0,11]

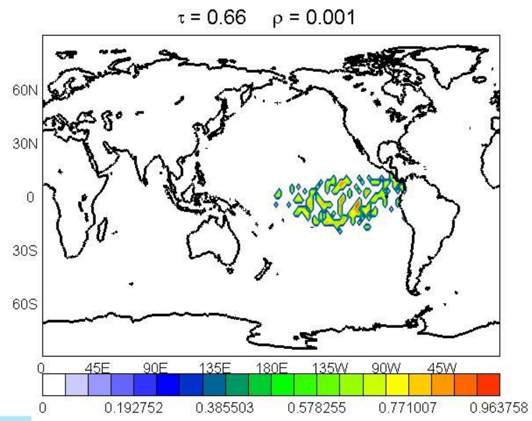
Zero-lag

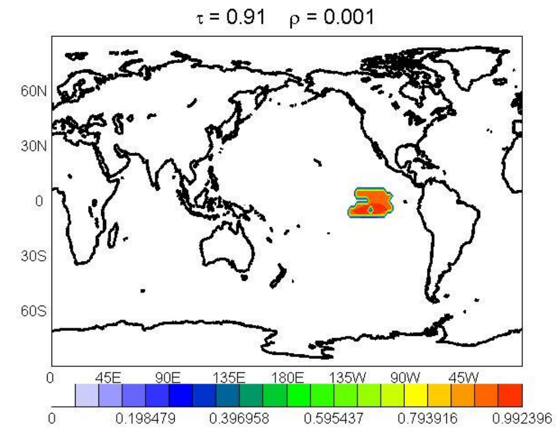
AWC





HUB





UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

Summary and future work

- Symbolic analysis is a powerful tools for the analysis of data from complex systems such as the climate.
- The success of the method is based on an appropriate symbolic representation that fully characterizes the diversity of patterns present in the time-series.
- Allows to study processes with different time scales.
- Problems identified: i) significant weak links might be hidden by noise; ii) because the network is embedded in a regular grid, the stronger links are mainly local connections and iii) the AWC does not reveal the rich underlying pattern of weak non-local connections.
- Future work: detection of link directionality and relations among different variables (construction of interacting networks).

THANK YOU FOR YOUR ATTENTION