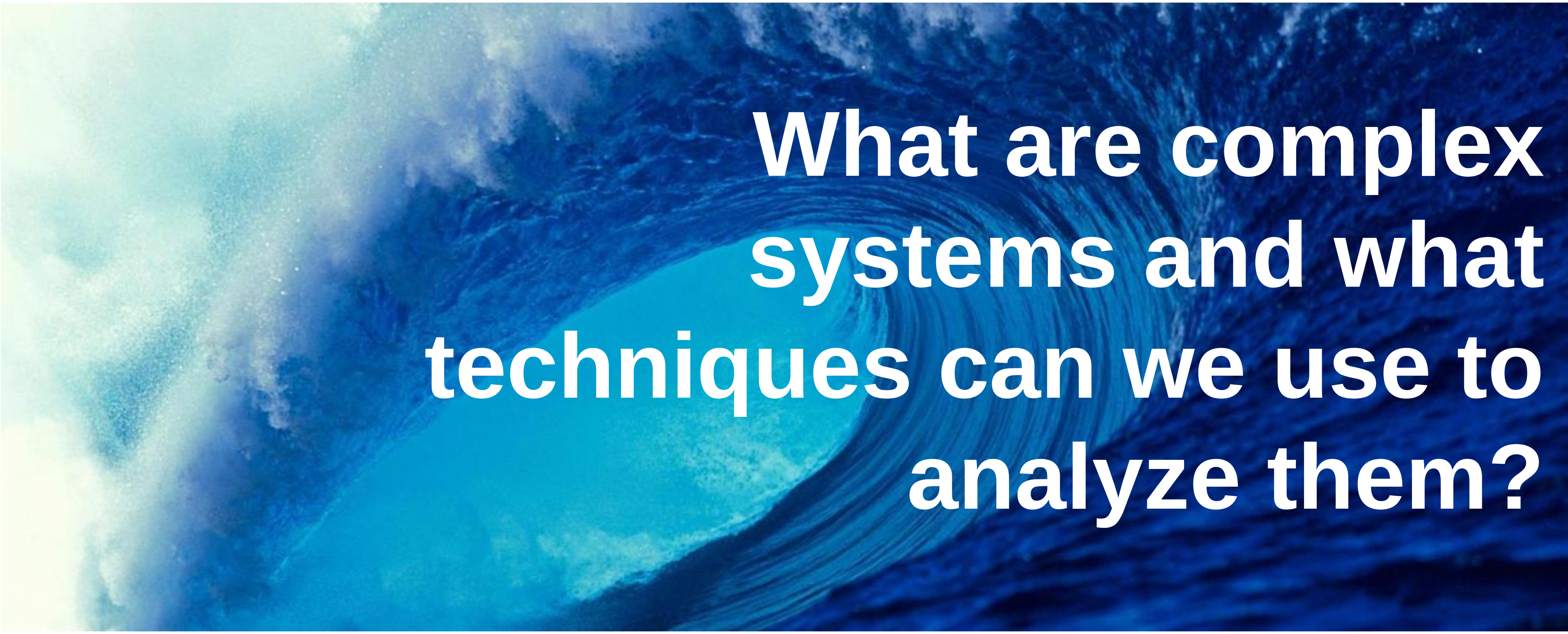


What are complex systems and what techniques can we use to analyze them?

Cristina Masoller,
Departament de Física, UPC

Summer School "Dynamical Systems and Complexity",
Posidi, Halkidiki, Greece
September 5, 2024

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

ICREA

AGAUR

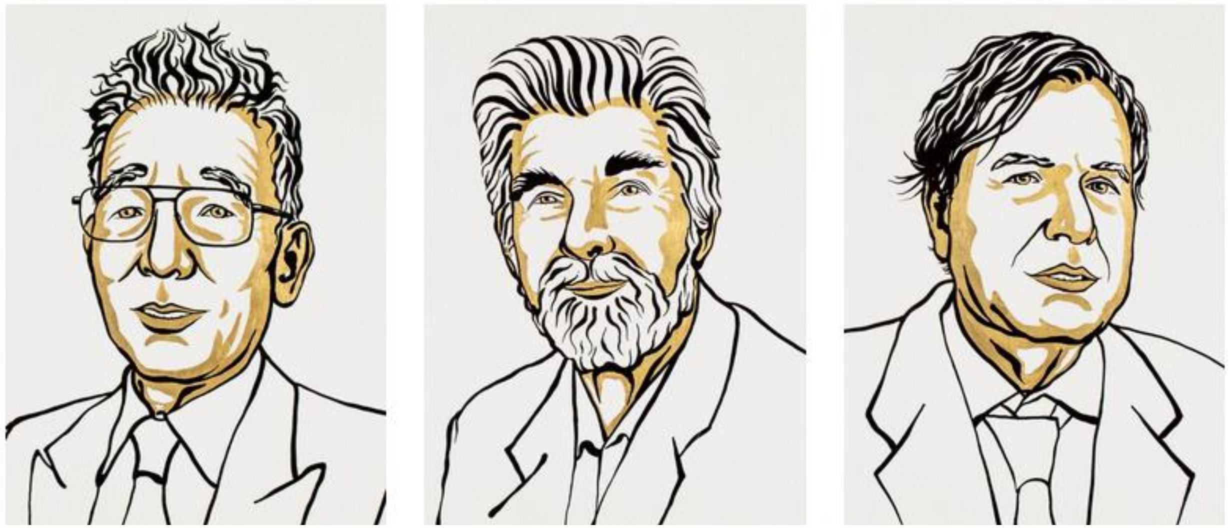
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

Outline

- Complex systems and data analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: laser, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

cristina.masoller@upc.edu @cristinamasoll1

The Nobel Prize in Physics 2021



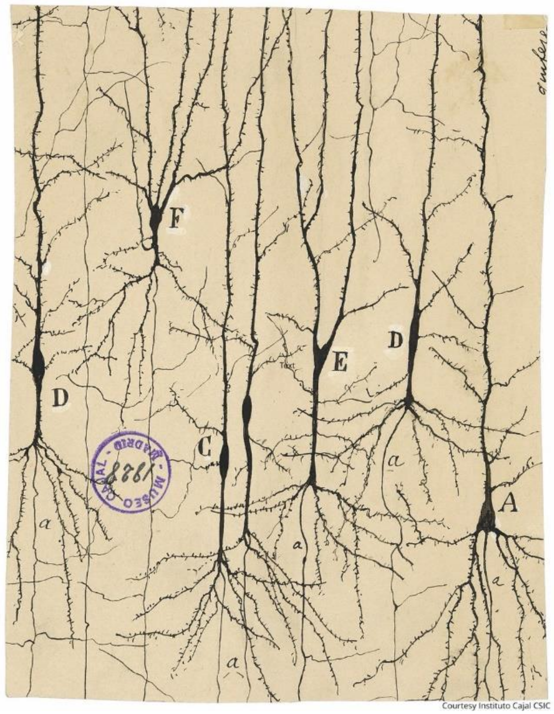
for groundbreaking contributions to our understanding of **complex systems**

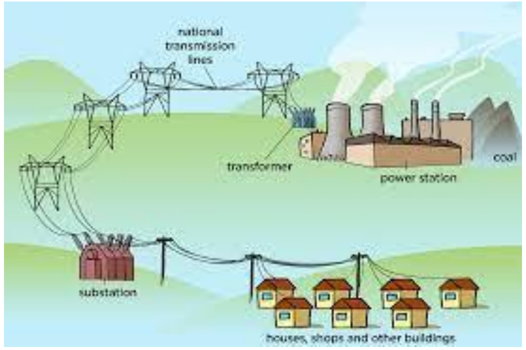
½ Syukuro Manabe and Klaus Hasselmann ½ Giorgio Parisi

cristina.masoller@upc.edu @cristinamasoll1

What is a complex system?

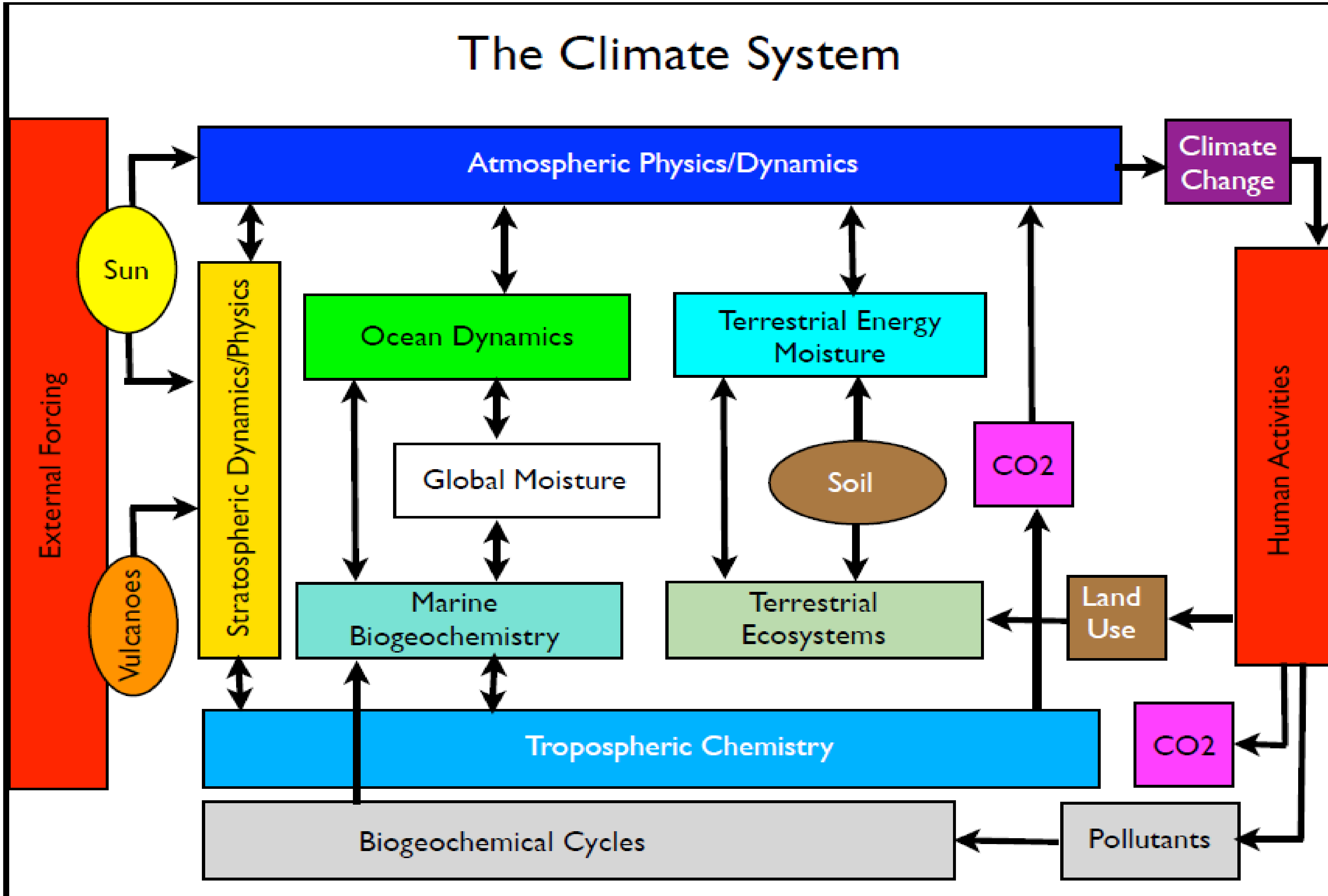
- High-dimensional, large number of interacting elements, heterogeneous structure, multiscale, memory, adaptation.
- The elements and/or the interactions are **nonlinear**.
- Often display abrupt transitions and extreme events.





G. Bianconi et al, *Complex systems in the spotlight: next steps after the 2021 Nobel Prize in Physics*, J. of Phys: Complexity 4, 010201 (2023).

cristina.masoller@upc.edu @cristinamasoll1



Courtesy of Henk Dijkstra (Universidad de Ultrech)

cristina.masoller@upc.edu @cristinamasoll1

Complex systems show "emergent" phenomena such as "synchronization"

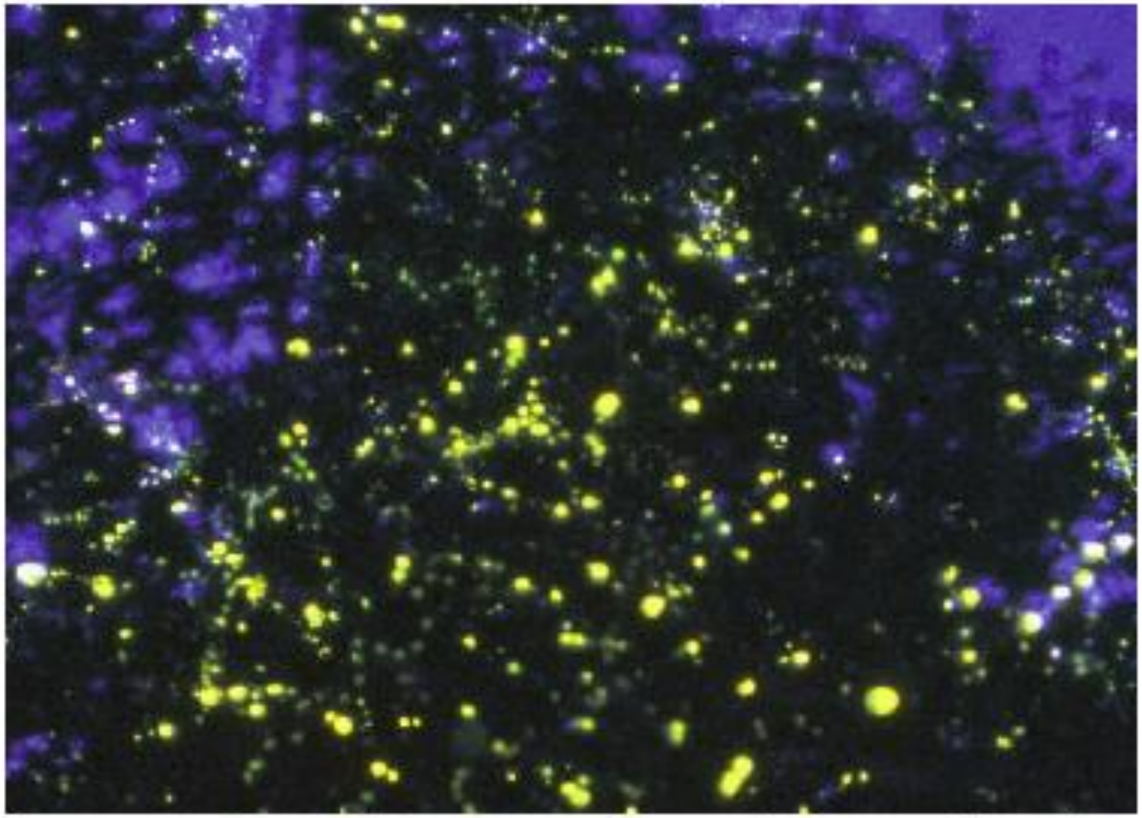


Figure 1 | Fireflies, fireflies burning bright. In the forests of the night, certain species of firefly flash in perfect synchrony — here *Pteroptyx malaccae* in a mangrove apple tree in Malaysia. Kaka *et al.*[2] and Mancoff *et al.*[3] show that the same principle can be applied to oscillators at the nanoscale.

cristina.masoller@upc.edu @cristinamasoll1

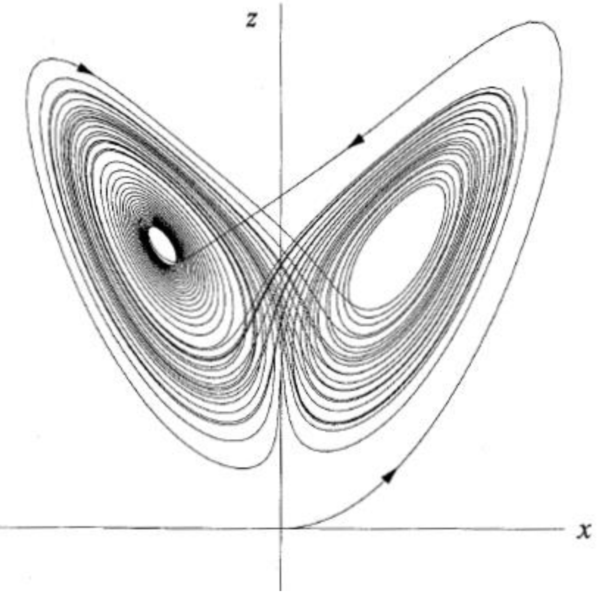
In my opinion, what is NOT a complex system:

- Any system, large or small, that is described by linear equations.
- Low dimensional nonlinear systems.

Example: Lorentz system

$$\frac{dx}{dt} = -\sigma x + \sigma y,$$
$$\frac{dy}{dt} = -xz + rx - y,$$
$$\frac{dz}{dt} = xy - bz.$$

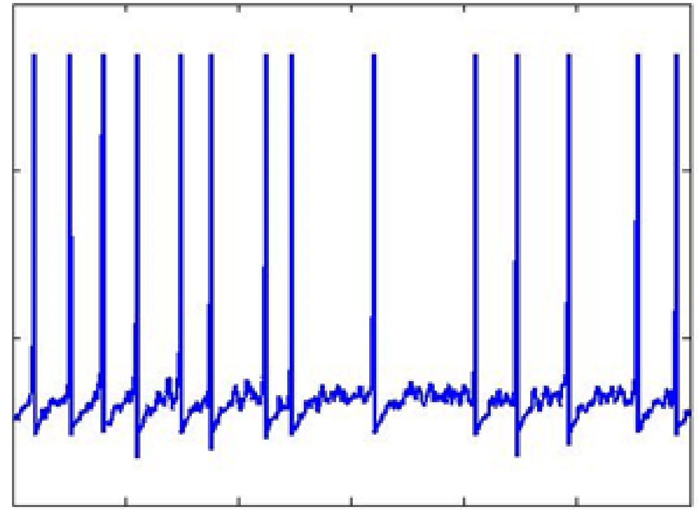
2D projection of 3D attractor



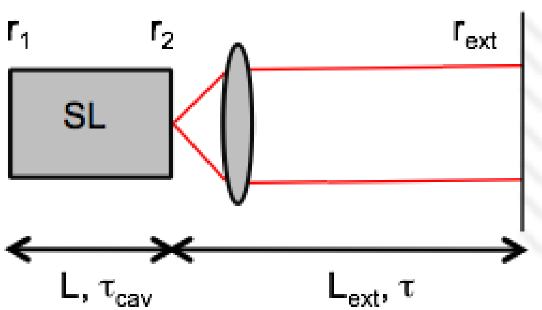
cristina.masoller@upc.edu @cristinamasoll1

Data analysis methods allow to discover statistical similarities in different systems

Time 10^{-9} s



Time 10^{-3} s



$$\varepsilon \frac{dx}{dt} = x - \frac{x^3}{3} - y \,,$$

$$\frac{dy}{dt} = x + a + D\xi(t) \,.$$

(High-dimensional, Memory)

cristina.masoller@upc.edu @cristinamasoll1

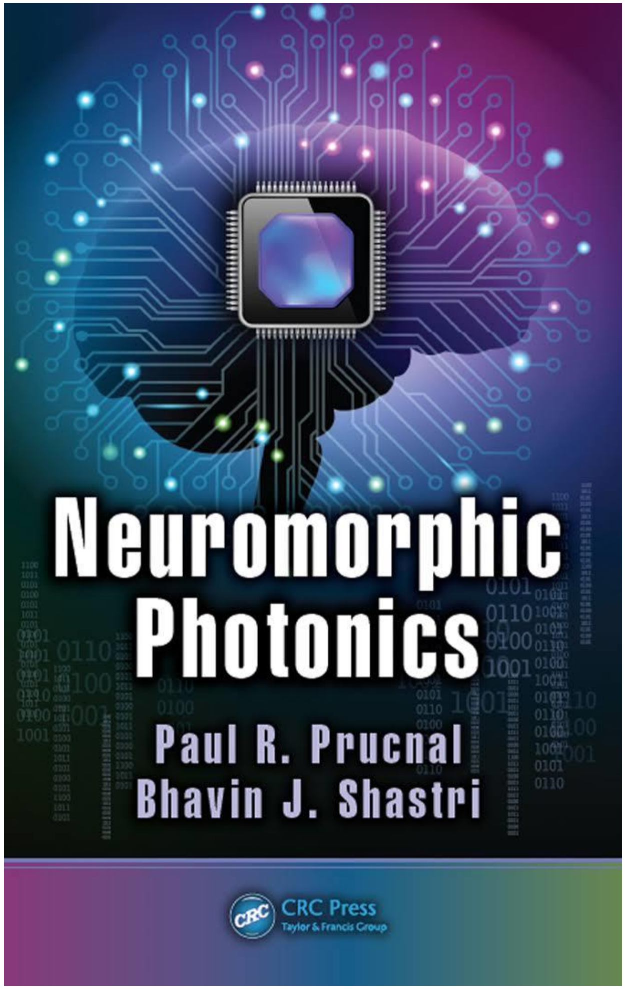
Uncovering similarities between neurons and lasers… Interesting but relevant?

- Data centers, AI systems, HPC consume huge amounts of energy.
- Big concern in the context of climate change.
- The human brain processes huge amounts of information using only 19 Watts.
- Uncovering genuine similarities between neurons and lasers will allow to develop **photonic neurons**, able to process information as real neurons do, but
  - much faster,
  - with much less energy consumption.

European Centre for Medium-Range Weather Forecasts, Reading, UK

cristina.masoller@upc.edu @cristinamasoll1

Photonic neurons

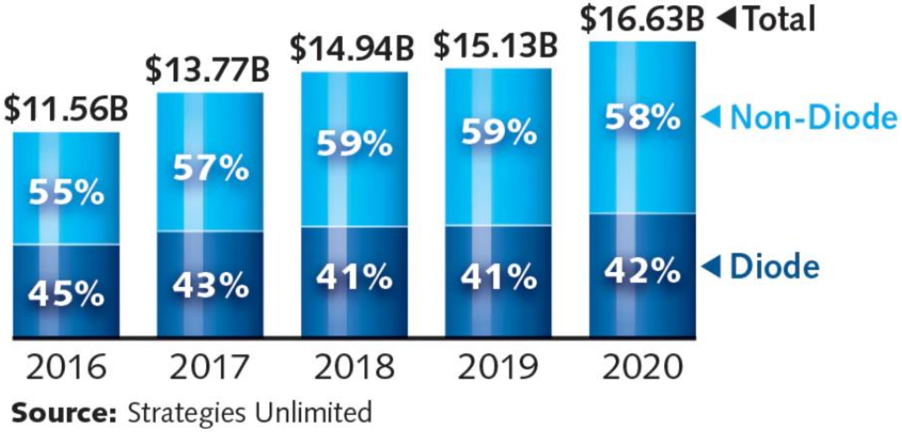


Excitable diode lasers can be artificial neurons in all-optical, ultra-fast, energy-efficient information processing systems.

cristina.masoller@upc.edu @cristinamasoll1

Photonic neurons with diode lasers

- Inexpensive
- Compact, energy-efficient
- Emit wavelengths appropriated for telecom, Datacom and biomedical applications,
- Can be integrated in large arrays,
- Optically perturbed: Rich nonlinear dynamics



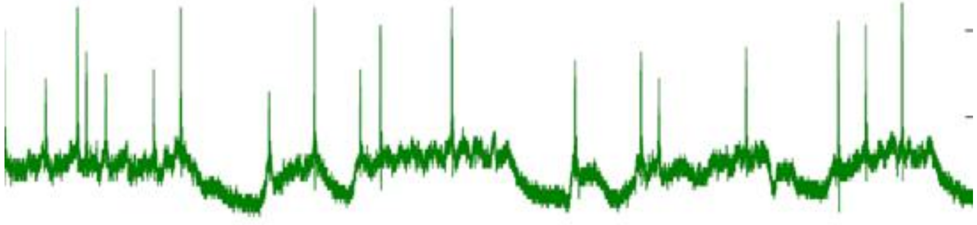
cristina.masoller@upc.edu @cristinamasoll1

Therefore, we want to know:

Can diode lasers mimic real neurons?

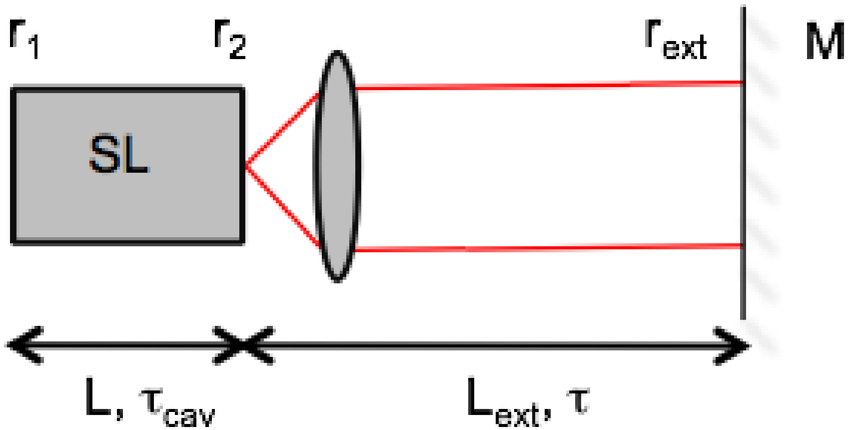
How neurons encode information of weak external inputs in noisy environments?

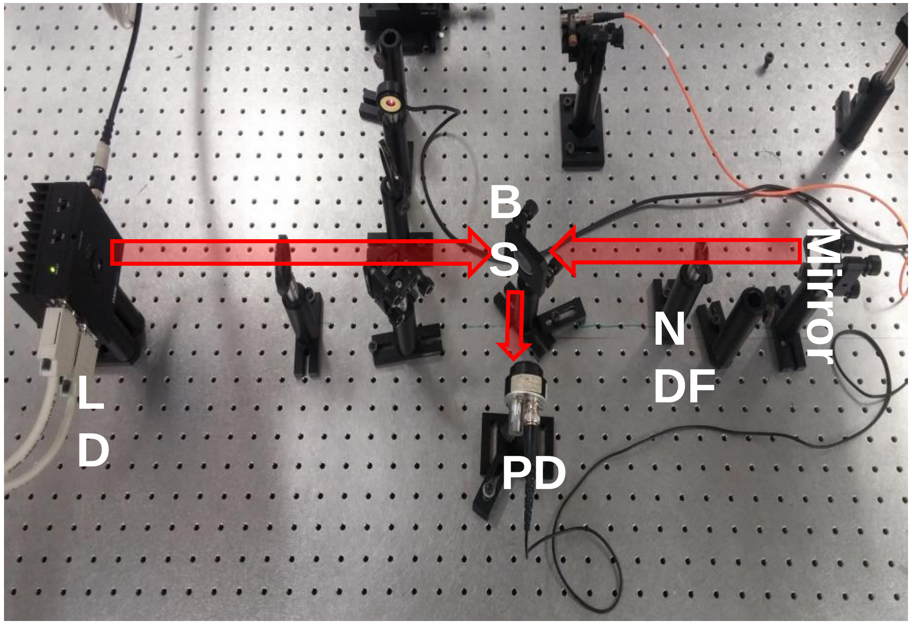
Can the neural code be used by diode lasers?



cristina.masoller@upc.edu @cristinamasoll1

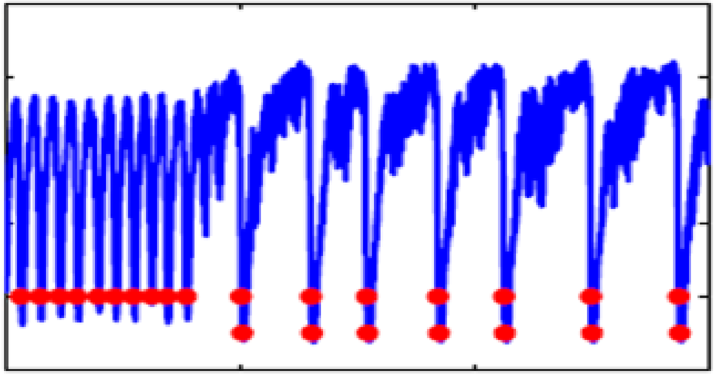
Diode laser with optical feedback

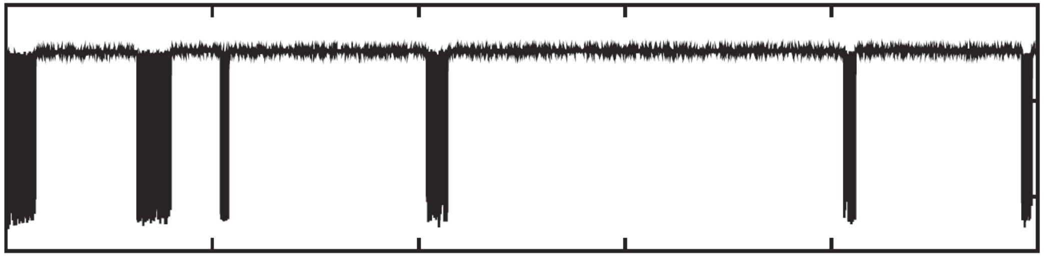


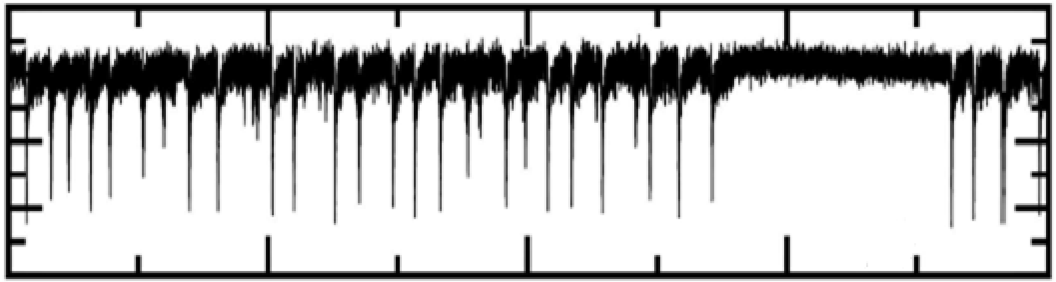


cristina.masoller@upc.edu @cristinamasoll1

The laser dynamics: excitability, tonic spikes and bursting. Similar to real neurons?





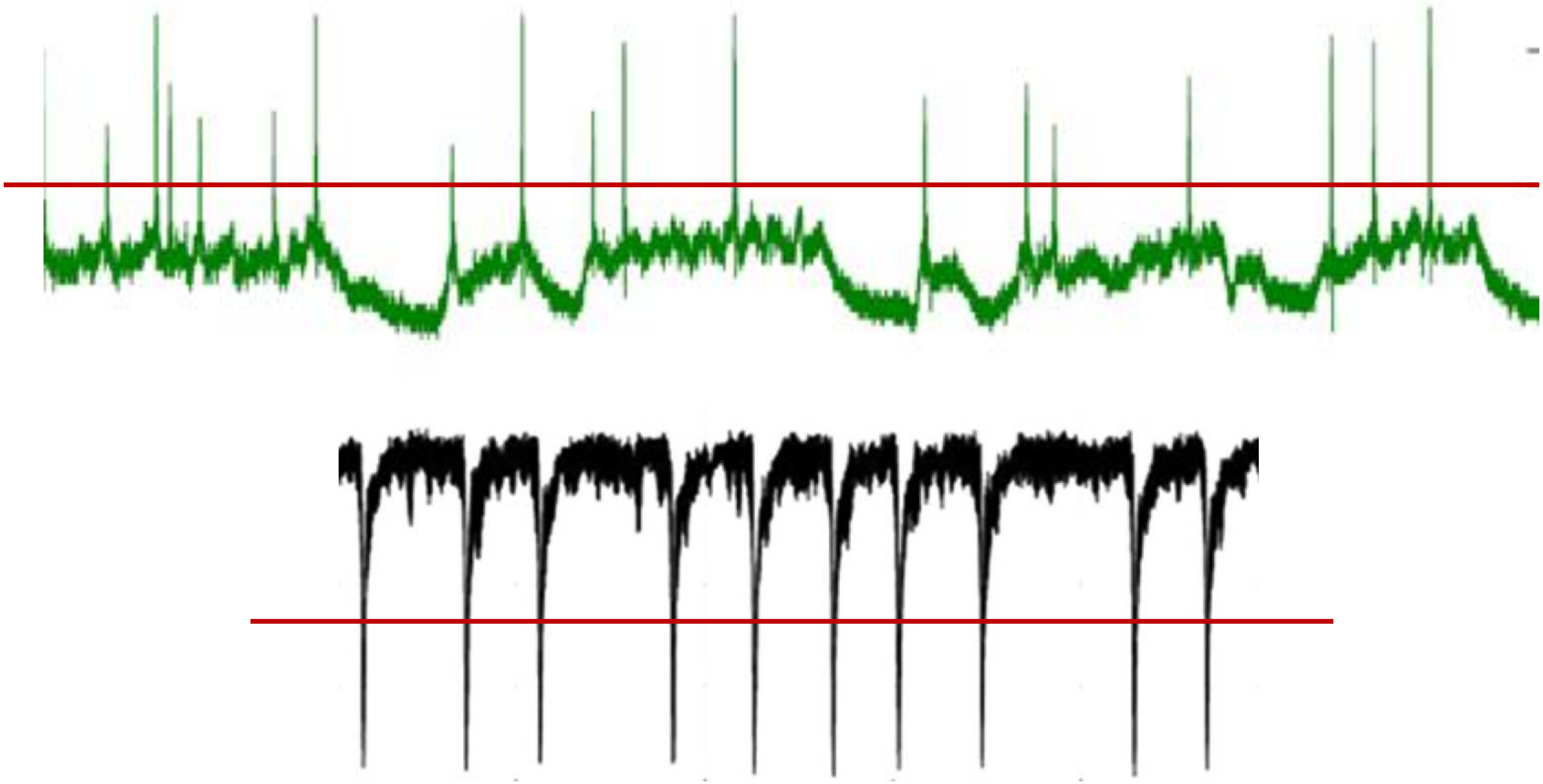


A. Aragoneses, S. Perrone, T. Sorrentino, M. C. Torrent and C. Masoller, "*Unveiling the complex organization of recurrent patterns in spiking dynamical systems*", Sci. Rep. **4**, 4696 (2014).

C. Quintero-Quiroz, J. Tiana-Alsina, J. Roma, M. C. Torrent, and C. Masoller, "*Characterizing how complex optical signals emerge from noisy intensity fluctuations*", Sci. Rep. **6** 37510 (2016).

cristina.masoller@upc.edu @cristinamasoll1

A threshold is used to detect the spike times ⇒ Sequence of inter-spike-intervals (ISIs)



cristina.masoller@upc.edu @cristinamasoll1

With an external signal, are there statistical similarities between neuronal spikes and laser spikes?

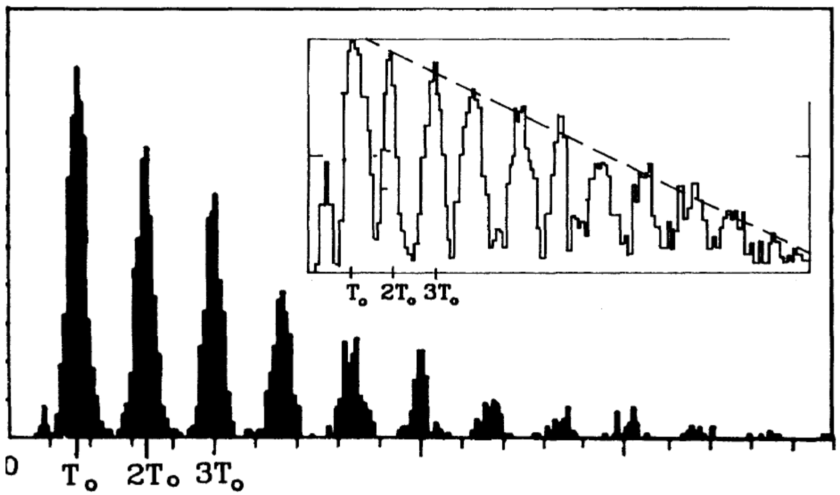
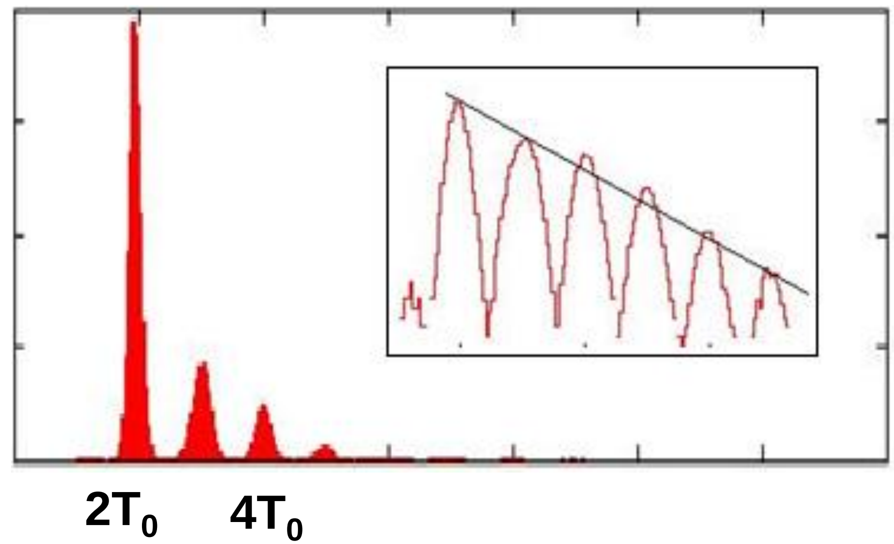


FIG. 1. (a) An experimental ISIH obtained from a single auditory nerve fiber of a squirrel monkey with a sinusoidal 80-dB sound-pressure-level stimulus of period $T_0 = 1.66$ ms applied at the ear. Note the modes at integer multiples of T_0. Inset:

A. Longtin et al. PRL (1991)



Experimental data when the laser current is modulated with a sinusoidal signal of period T_0.

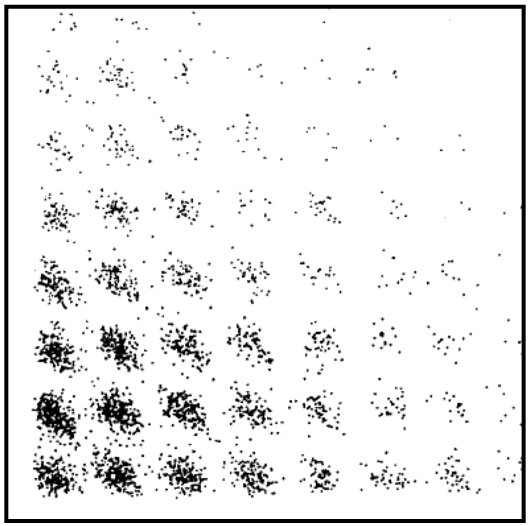
[A. Aragoneses et al. Optics Express (2014)]

cristina.masoller@upc.edu @cristinamasoll1

Return maps of inter-spike-intervals (ISIs)

Neuronal ISIs

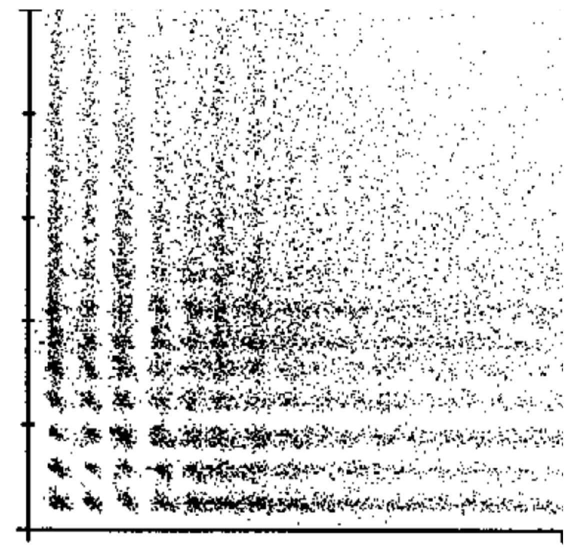
ΔT_{i+1}



ΔT_i

A. Longtin
Int. J. Bif. Chaos (1993)

Laser ISIs



M. Giudici et al PRE (1997)
A. Aragoneses et al
Optics Express (2014)

HOW TO INDENTIFY TEMPORAL ORDER?

cristina.masoller@upc.edu @cristinamasoll1

How to characterize spike sequences?
Analysis of inter-spike-intervals -ISIs

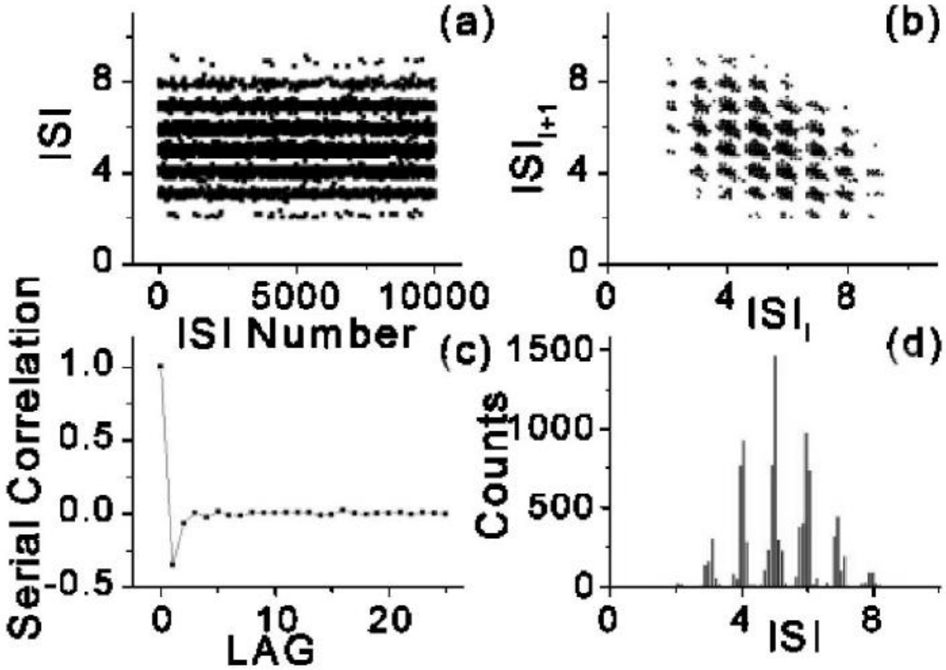
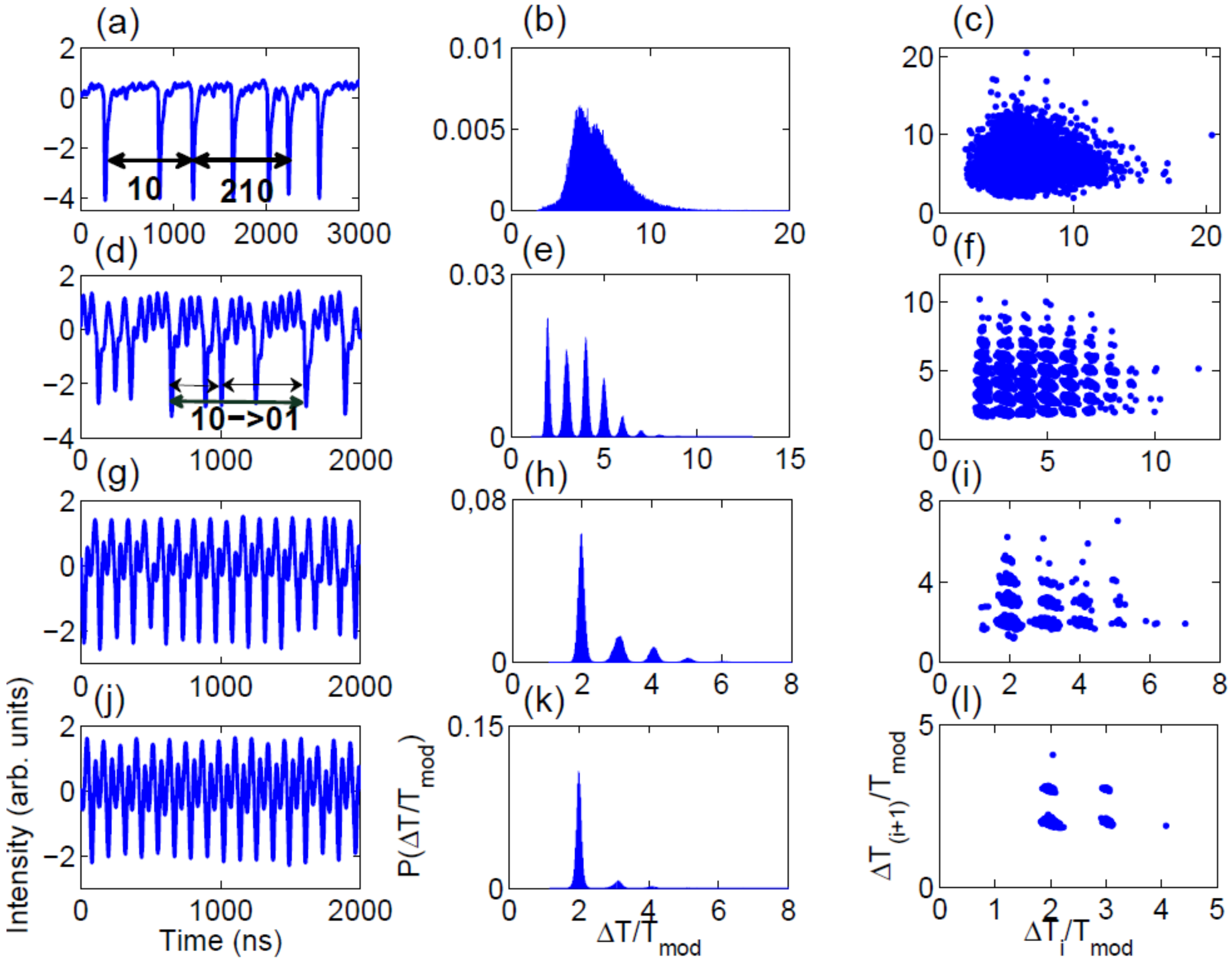


FIG. 1. Analysis of 10 000 consecutive interspike intervals from a P-unit of the weakly electric fish A. Leptorhynchus (data courtesy of Mark Nelson, Beckmann Institute, Illinois, USA; we focus on such "nonbursty" units). Time is in EOD cycles; the EOD frequency is 755 Hz. The firing rate is 145 Hz which corresponds to $P = 0.192$. (a) Raster plot of ISI duration versus ISI number, (b) return map, (c) serial correlation, and (d) histogram.

Chacron, Longtin, et. al, Phys. Rev. Lett. 85, 1576 (2000)

cristina.masoller@upc.edu @cristinamasoll1

Experiments in our lab with a diode laser with feedback



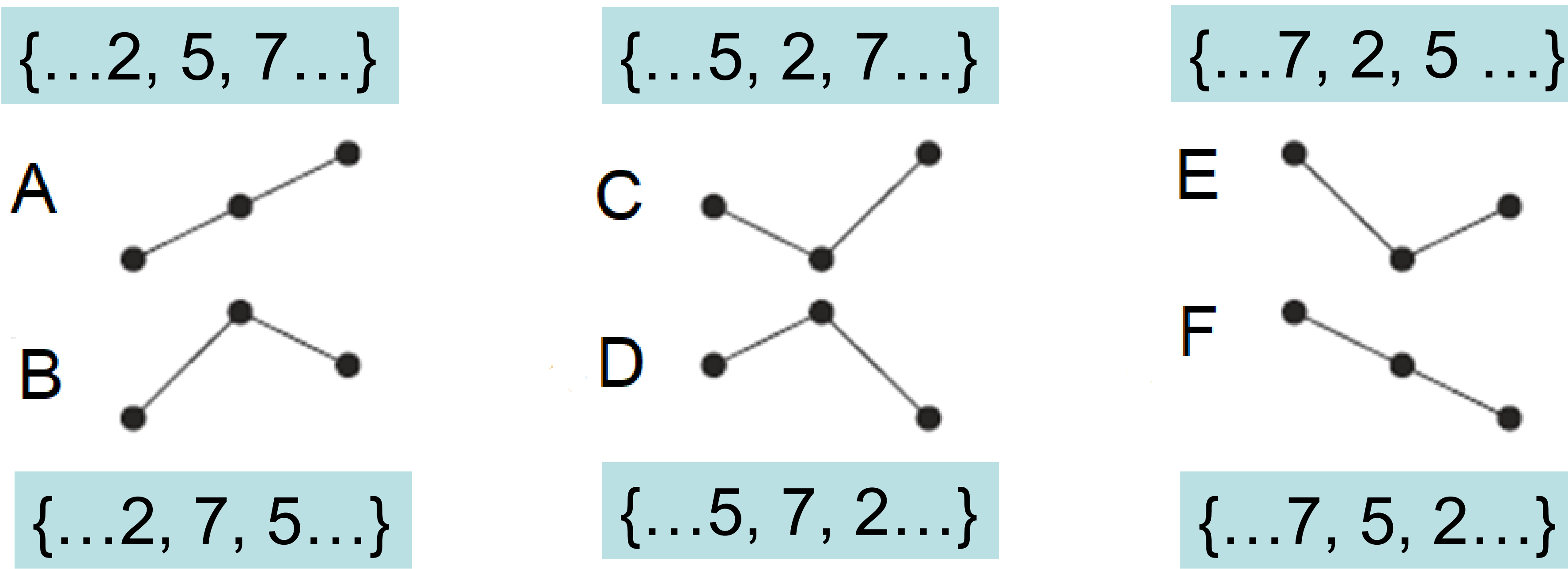
Aragoneses et al, Optics Express 22, 4705 (2014)

cristina.masoller@upc.edu @cristinamasoll1

First data analysis method: ordinal analysis

$$\{\ldots x_i,\ x_{i+1},\ x_{i+2},\ \ldots\}$$

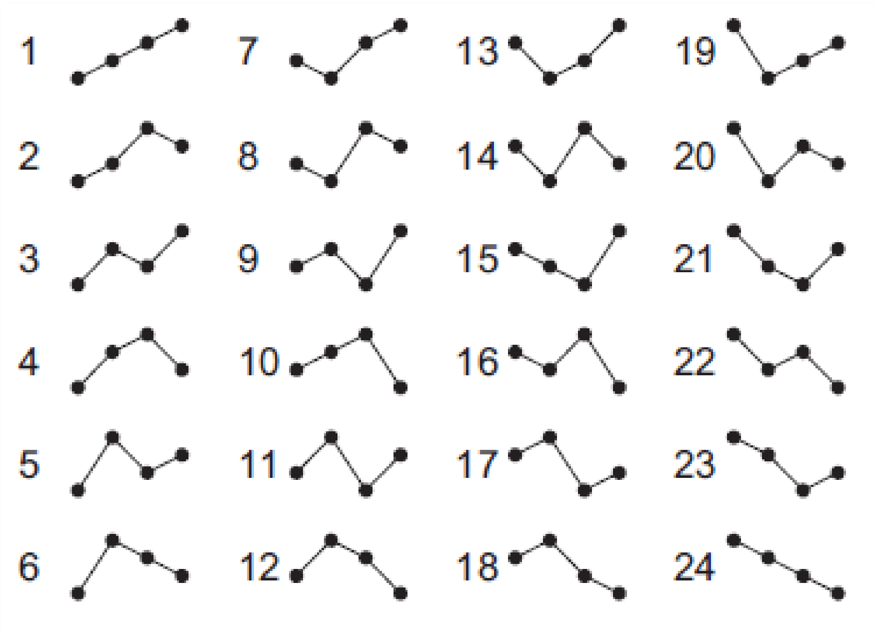
Possible order relations among three numbers (e.g., 2, 5, 7)

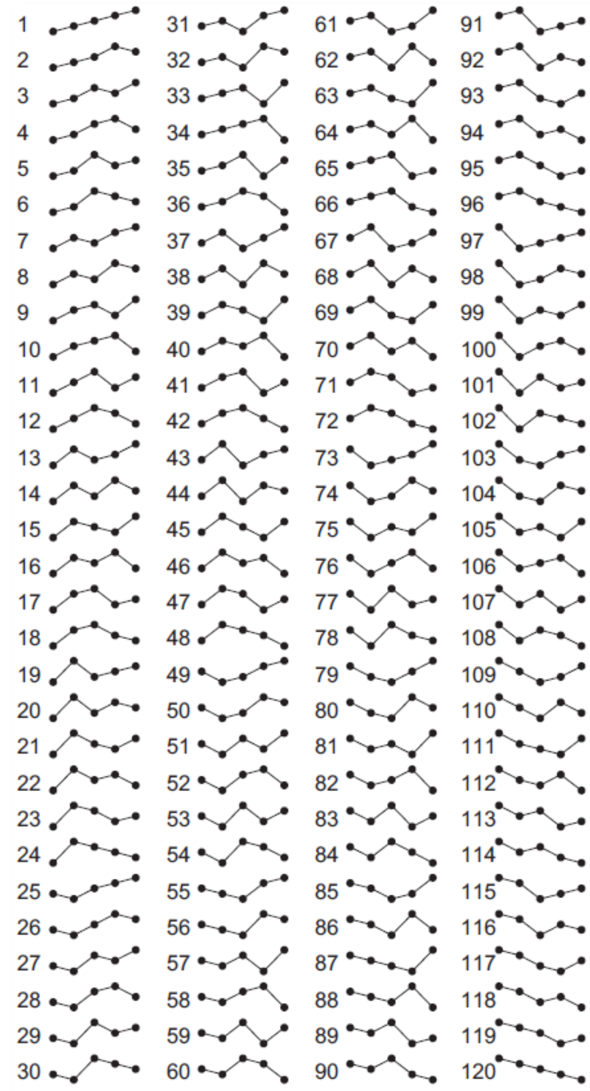


Bandt and Pompe: Phys. Rev. Lett. 2002

cristina.masoller@upc.edu @cristinamasoll1

The number of ordinal patterns increases as D!





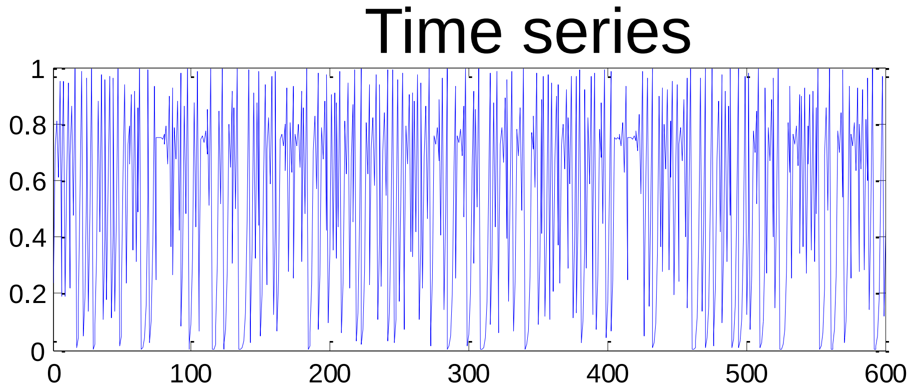
A problem for short datasets.

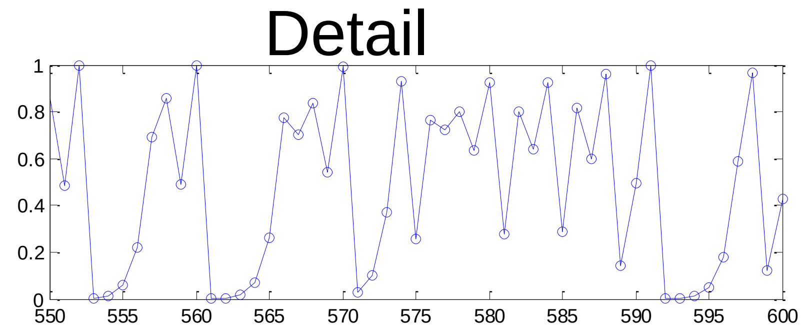
U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

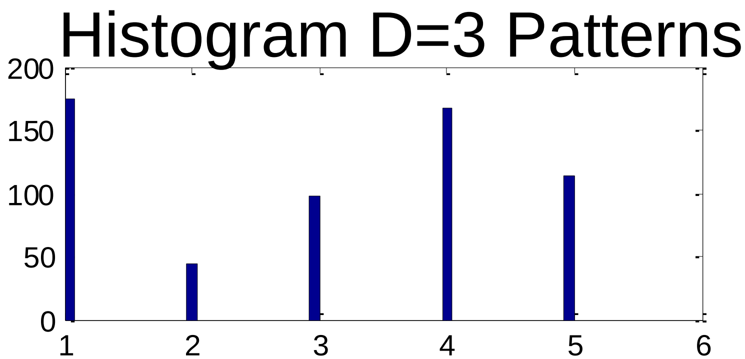
cristina.masoller@upc.edu @cristinamasoll1

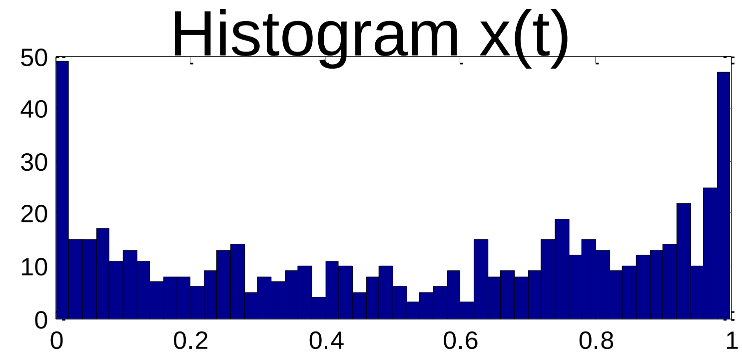
Example: chaotic time series generated with the Logistic map

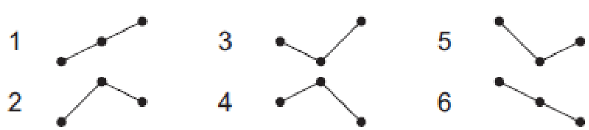
$$x(i+1) = r\,x(i)[1 - x(i)] \quad r=3.99$$









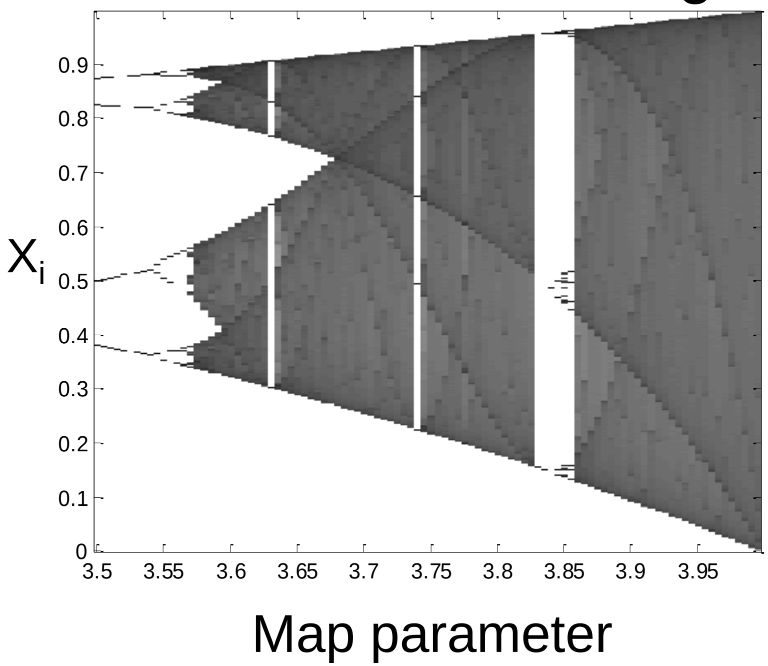


↑ **forbidden**

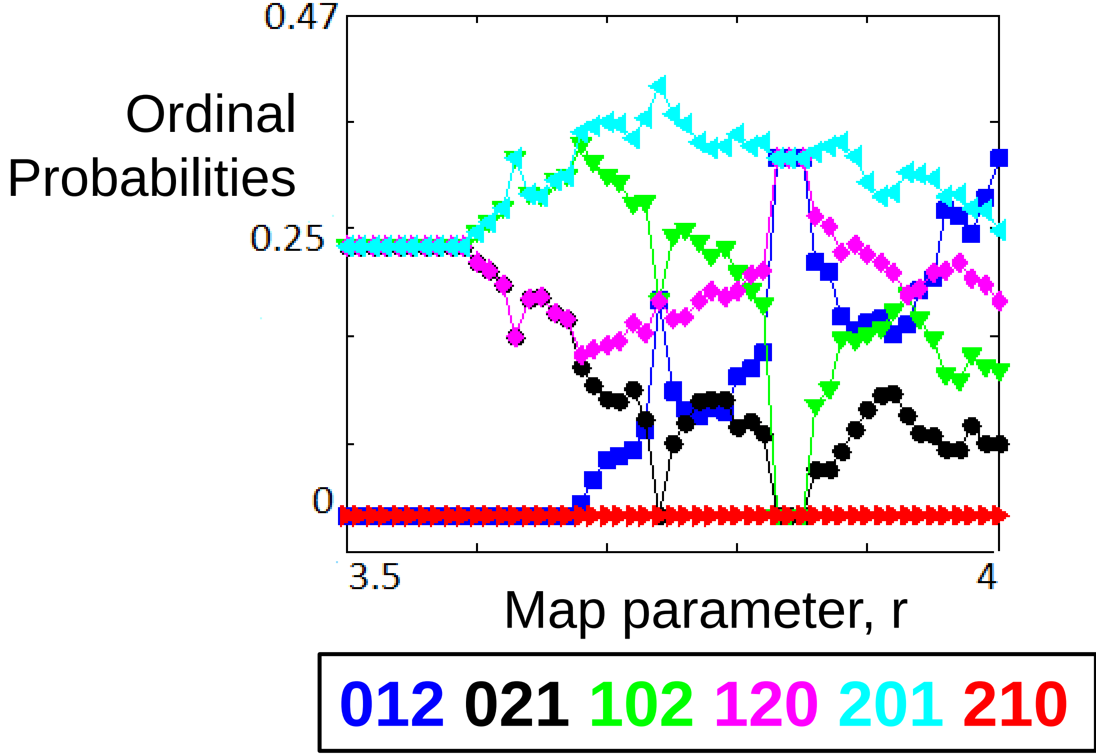
cristina.masoller@upc.edu @cristinamasoll1

“Normal” and “Ordinal” bifurcation diagrams of the Logistic map

Normal bifurcation diagram



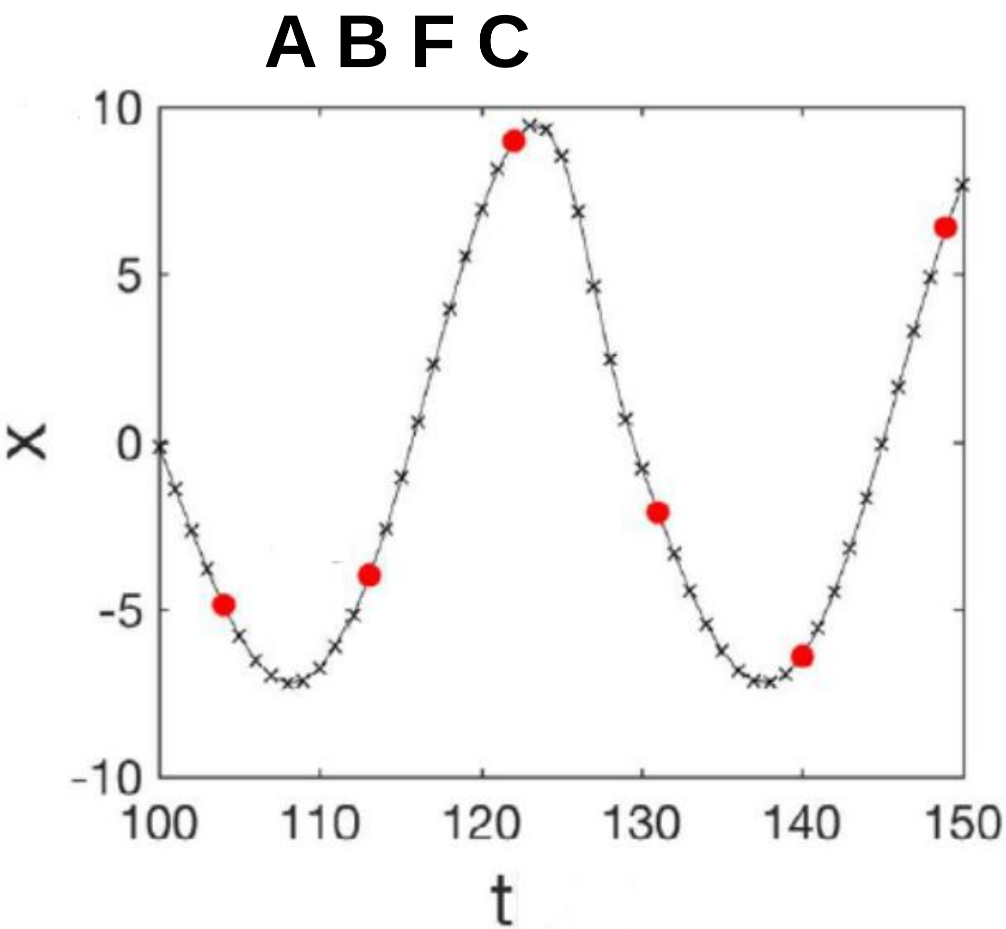
Ordinal diagram with D=3

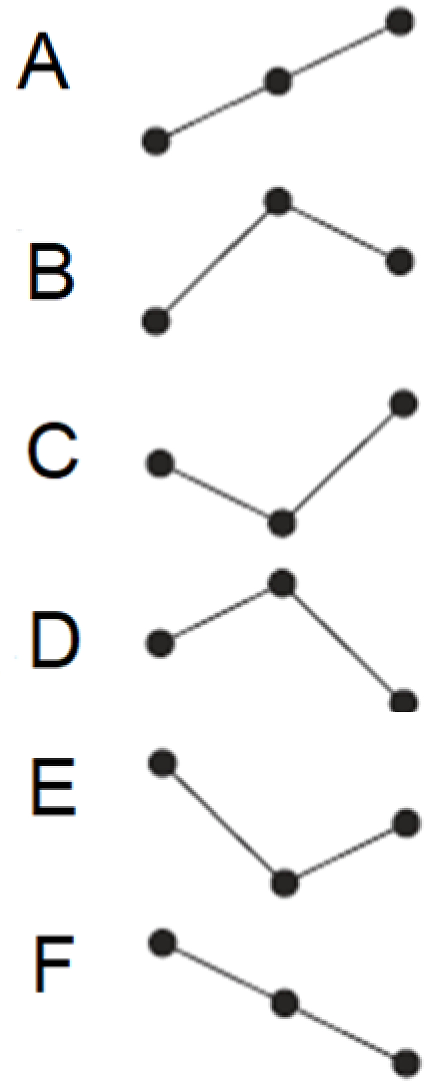


Pattern **210** is always forbidden; pattern **012** is more probable as r increases

cristina.masoller@upc.edu @cristinamasoll1

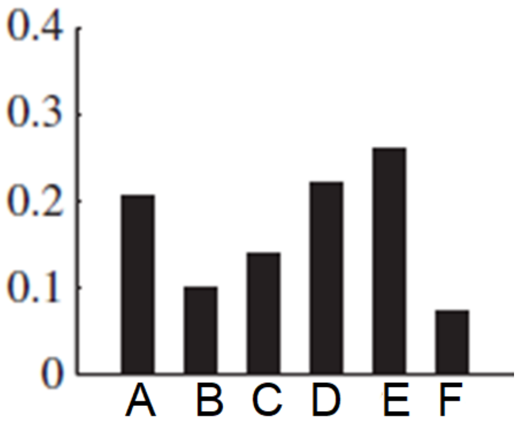
Using the “ordinal code”, which is the message?





cristina.masoller@upc.edu @cristinamasoll1

From a time series, by counting the different patterns, we can calculate the set of "ordinal probabilities"



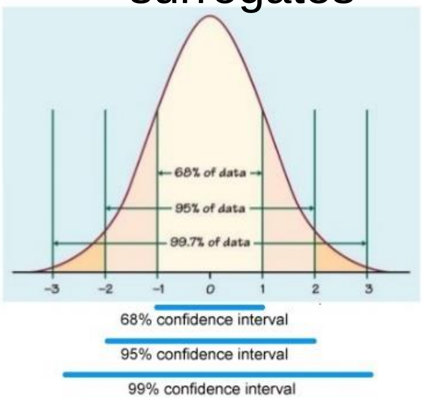
→ ?

A. Analyze the probabilities (are differences statistically significant?)

B. Analyse information theory measures (entropy, complexity)

$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

Compare with surrogates



Ordinal analysis has been extensively used:

- to test if a model is good for the data,
- to fit the model's parameters,
- to classify different types of data based on similarities of probabilities of ordinal patterns.

cristina.masoller@upc.edu @cristinamasoll1

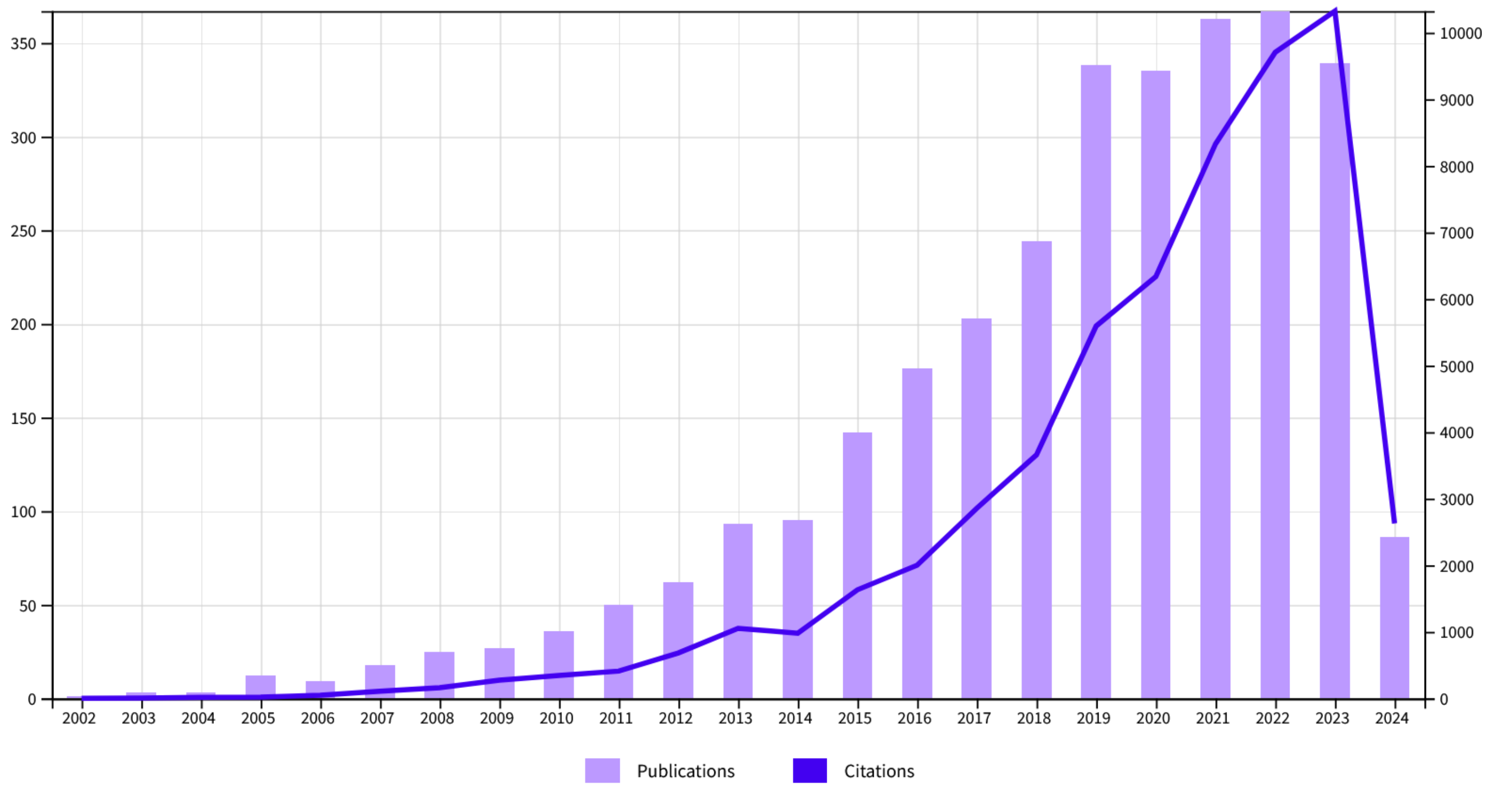
VOLUME 88, NUMBER 17 PHYSICAL REVIEW LETTERS 29 APRIL 2002

Permutation Entropy: A Natural Complexity Measure for Time Series

Christoph Bandt and Bernd Pompe

Institute of Mathematics and Institute of Physics, University of Greifswald, Greifswald, Germany

(Received 19 June 2001; revised manuscript received 20 December 2001; published 11 April 2002)

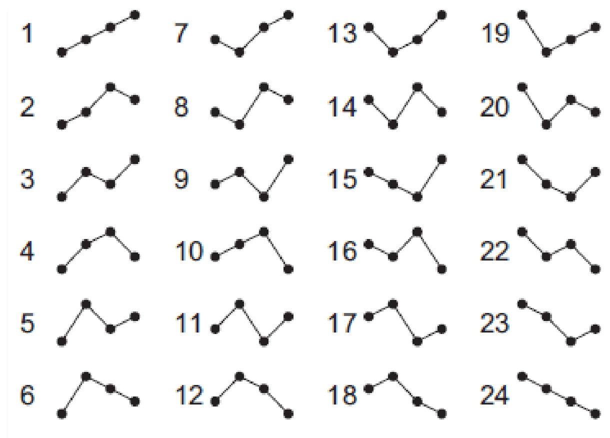


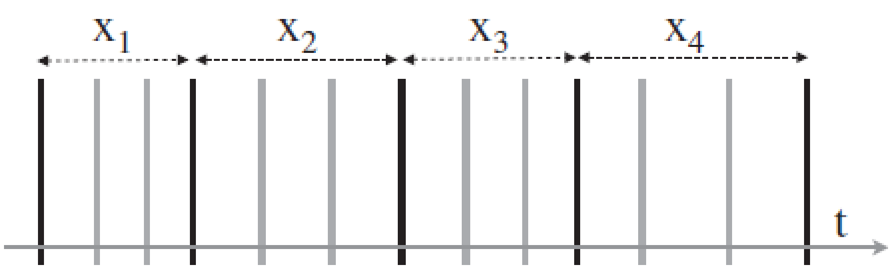
I. Leyva, J. M. Martinez, C. Masoller, O. A. Rosso, M. Zanin, "20 Years of Ordinal Patterns: Perspectives and Challenges", EPL 138, 31001 (2022).

cristina.masoller@upc.edu @cristinamasoll1

Software

Python and Matlab codes for computing the ordinal pattern **index** are available here: U. Parlitz et al. Computers in Biology and Medicine 42, 319 (2012)

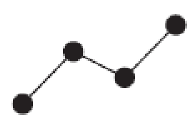




World length (wl): 4

Lag = 3 (skip 2 points)

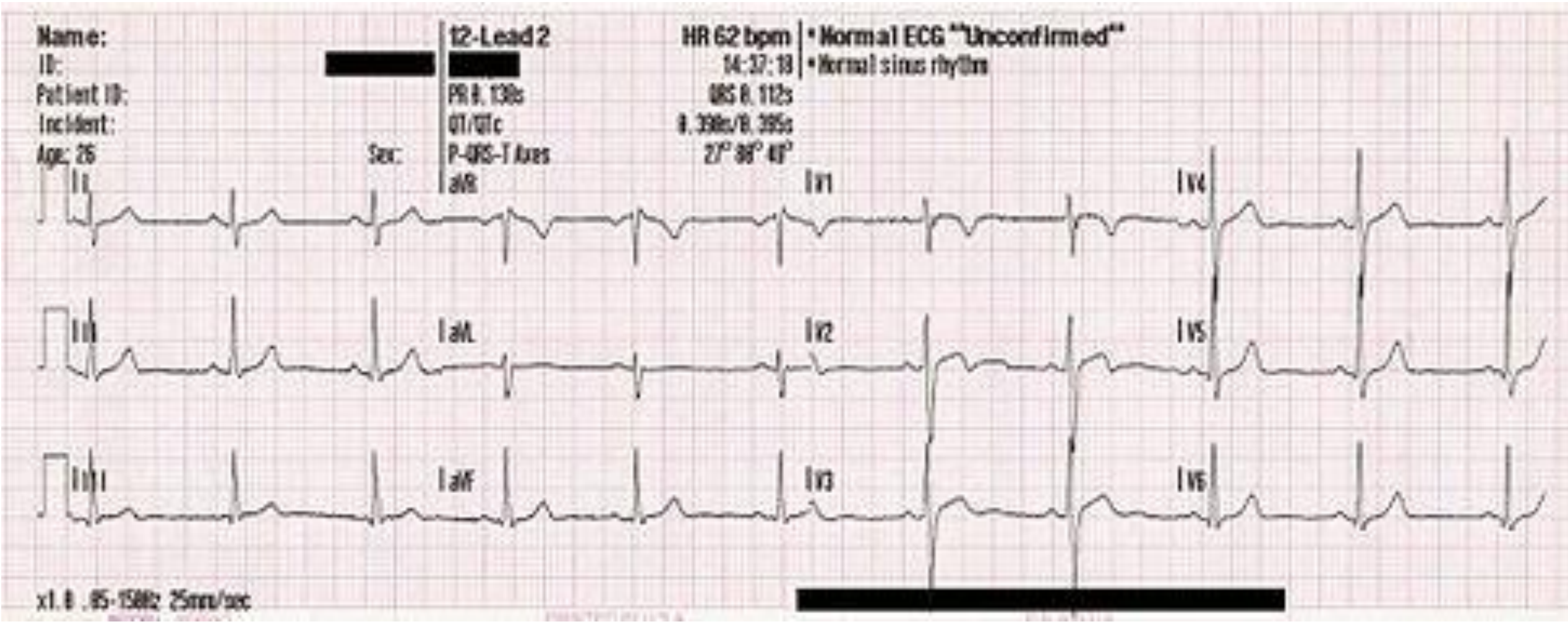
Result:

indcs= 3 

```
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 ;
```

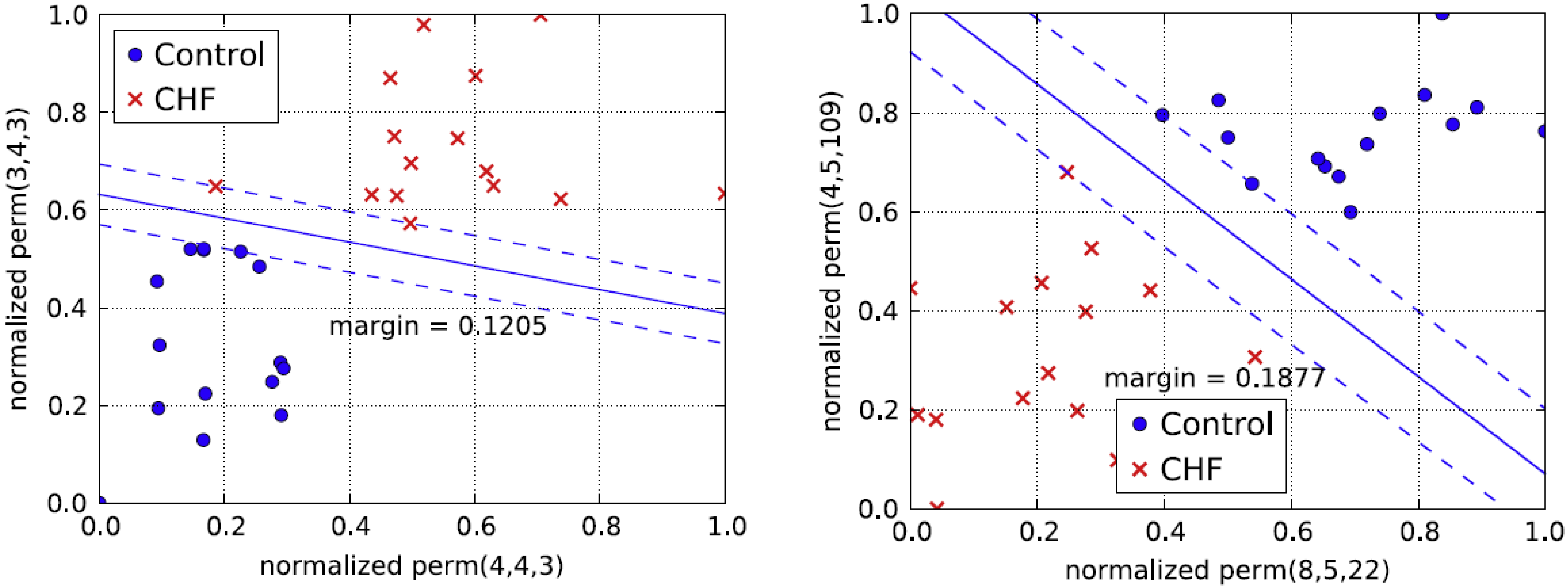

cristina.masoller@upc.edu @cristinamasoll1

Example of application.
ECG signals: analysis of time series of inter-beat intervals



cristina.masoller@upc.edu @cristinamasoll1

Classifying ECG signals according to ordinal probabilities

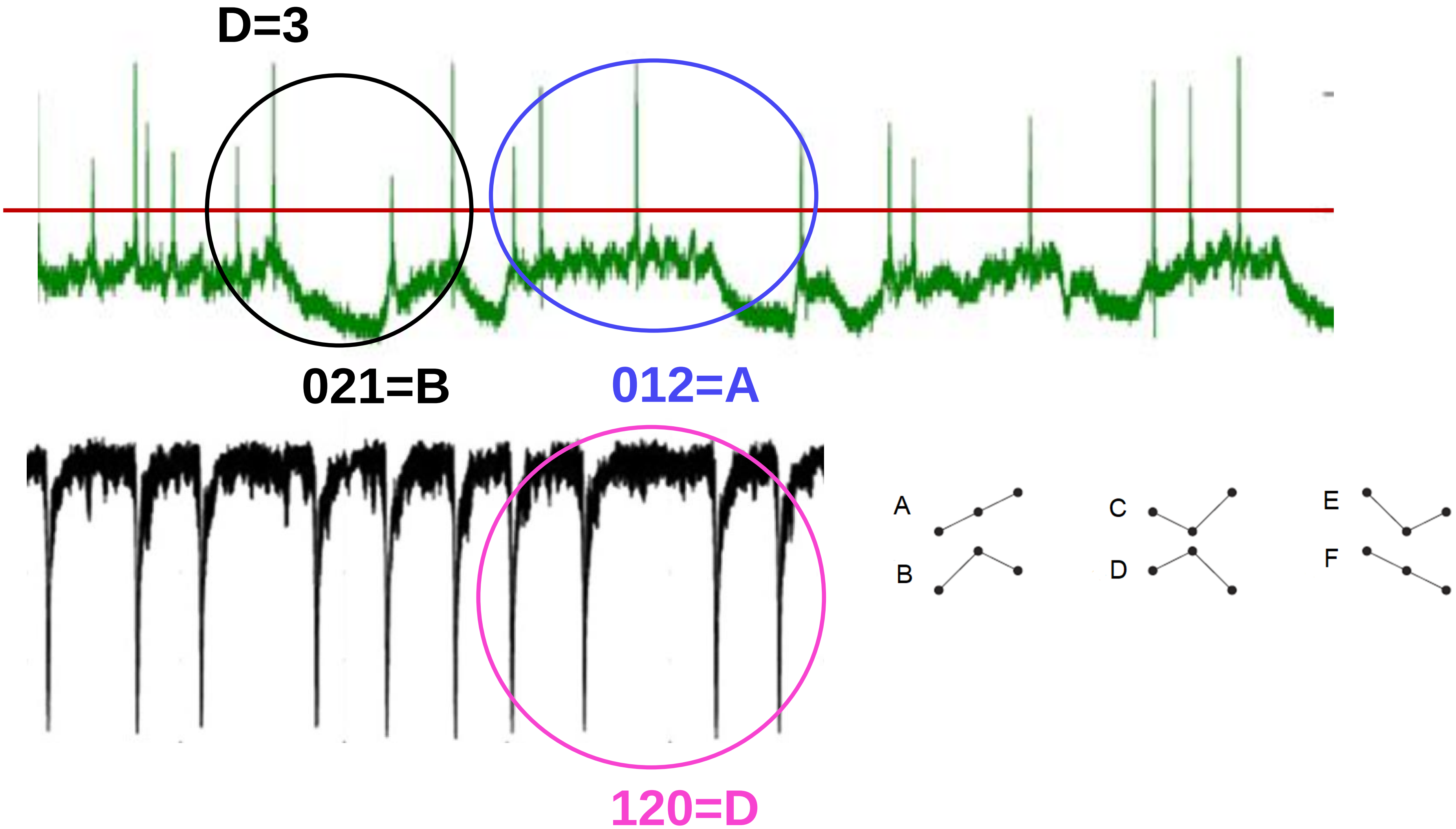


- Analysis of raw data (statistics of ordinal patterns is almost unaffected by a few extreme values)
- The probabilities are normalized with respect to the smallest and the largest value occurring in the data set.

U. Parlitz et al. Computers in Biology and Medicine 42, 319 (2012)

cristina.masoller@upc.edu @cristinamasoll1

Sequence of inter-spike-intervals (ISIs) ⇒ sequence of ordinal patterns



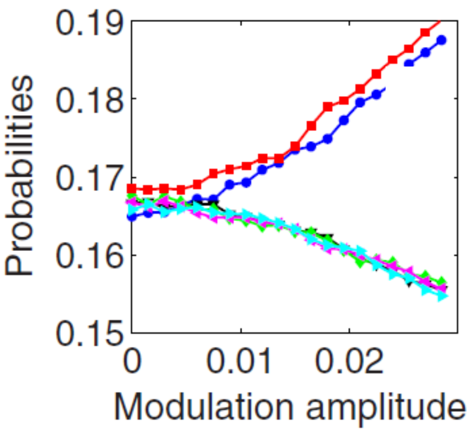
cristina.masoller@upc.edu @cristinamasoll1

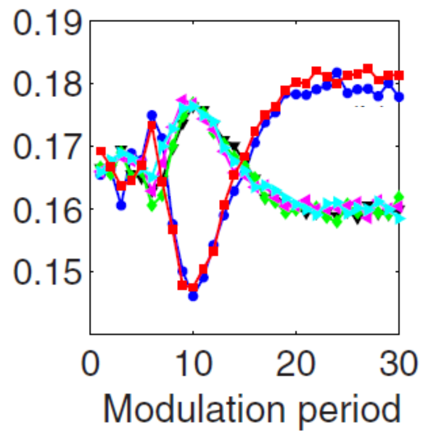
Simulations of a neural model

To try to understand how neurons encode and process weak inputs in noisy environments.

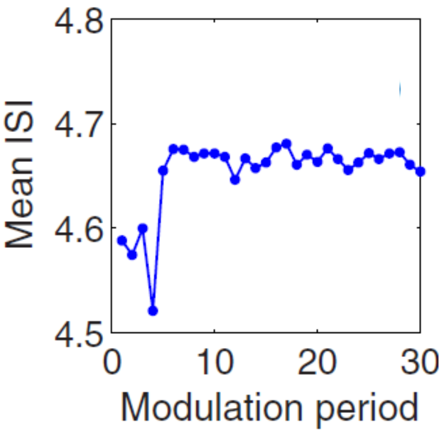
$$\epsilon \frac{dx}{dt} = x - \frac{x^3}{3} - y,$$
$$\frac{dy}{dt} = x + a + a_o \cos(2\pi t/T) + D\xi(t),$$

Weak, subthreshold signal





Rate coding?

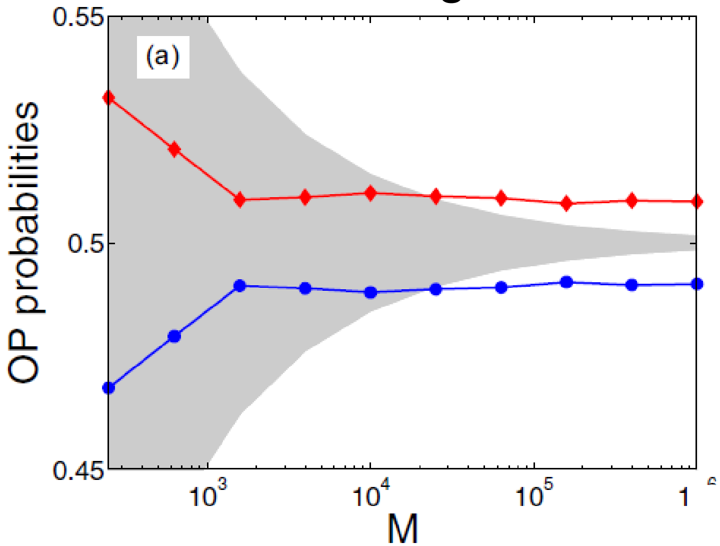


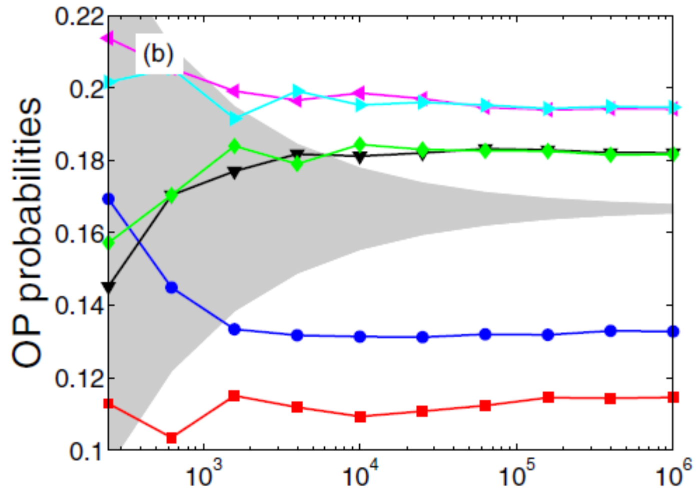
J. A. Reinoso, M. C. Torrent, and C. Masoller, "*Emergence of spike correlations in periodically forced excitable systems*", Phys. Rev. E. 94, 032218 (2016).

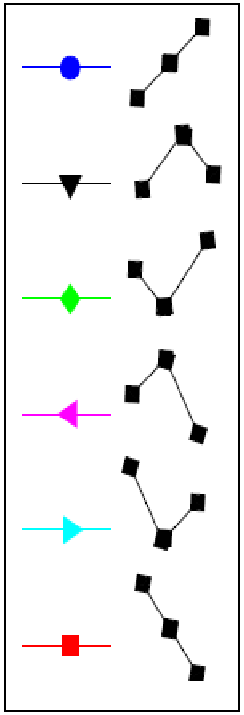
cristina.masoller@upc.edu @cristinamasoll1

How many spikes do we need to estimate the probabilities?

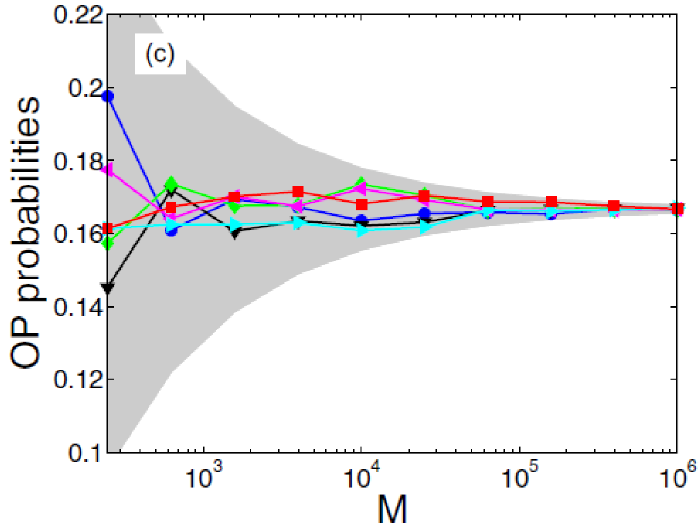
With external signal **01** **10**





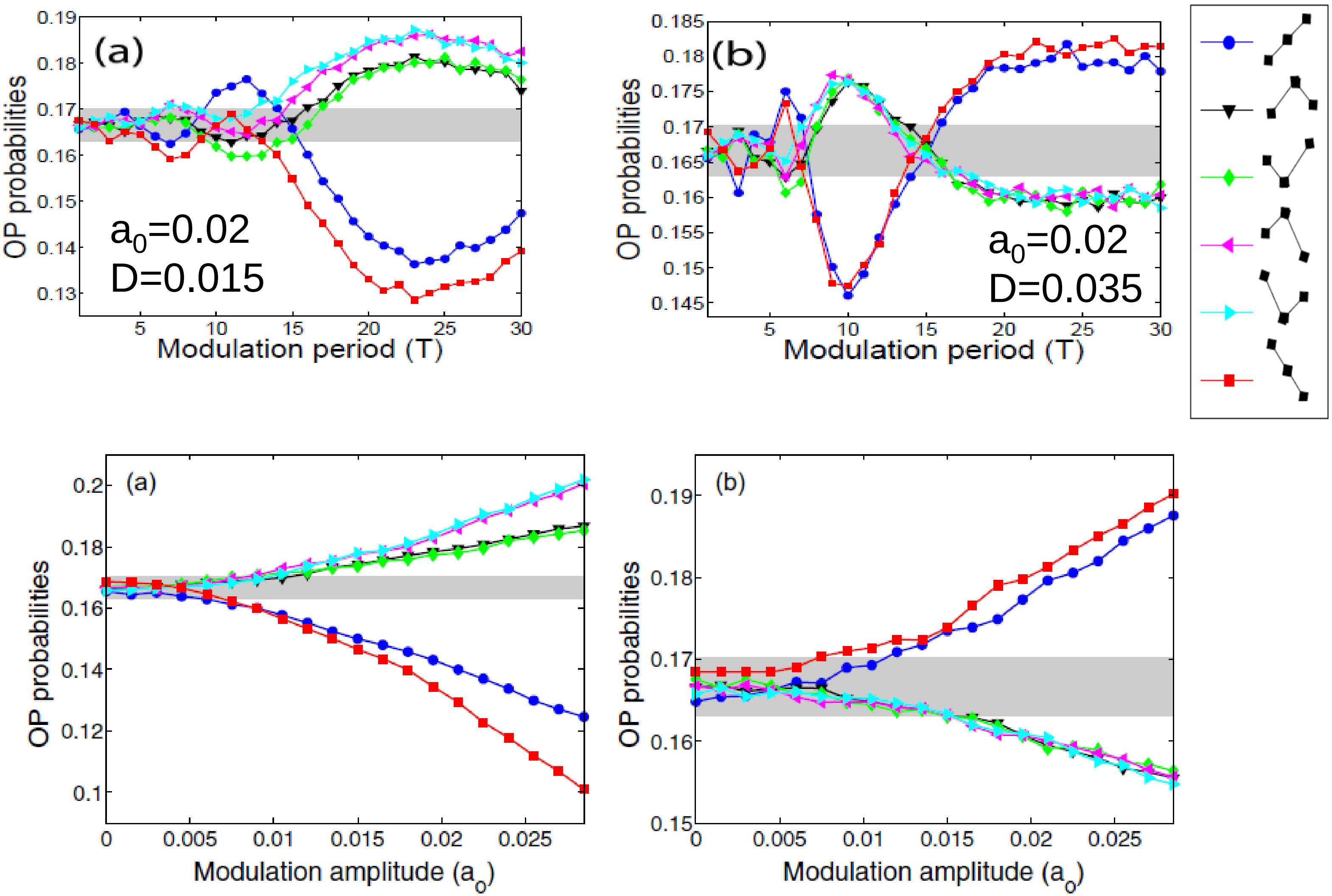


Without external signal



cristina.masoller@upc.edu @cristinamasoll1

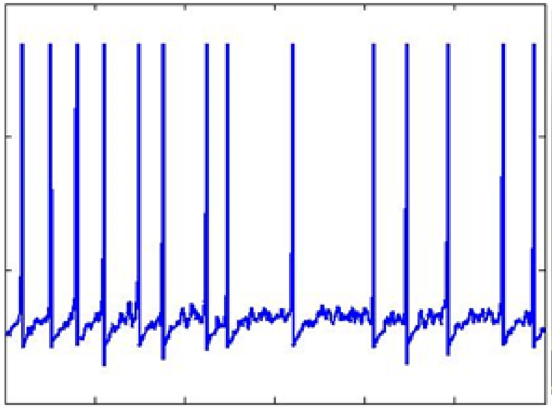
The patterns' probabilities depend not only on the period of the external signal, but also, on the level of noise.

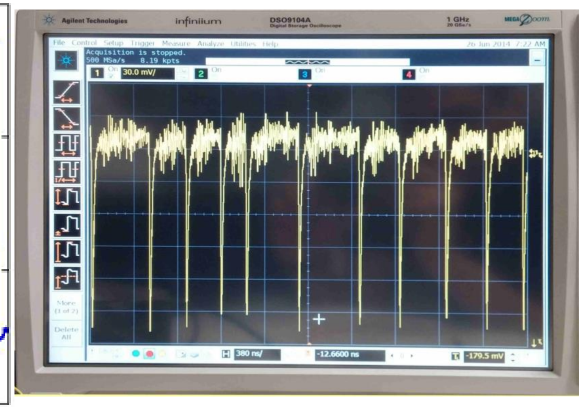


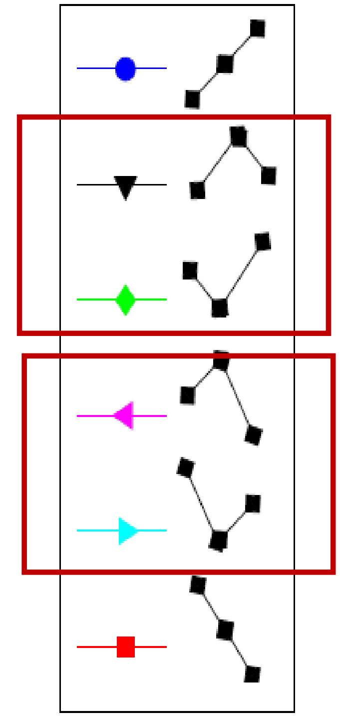
cristina.masoller@upc.edu @cristinamasoll1

The analysis of the ordinal probabilities uncovers similarities in the ISI sequences

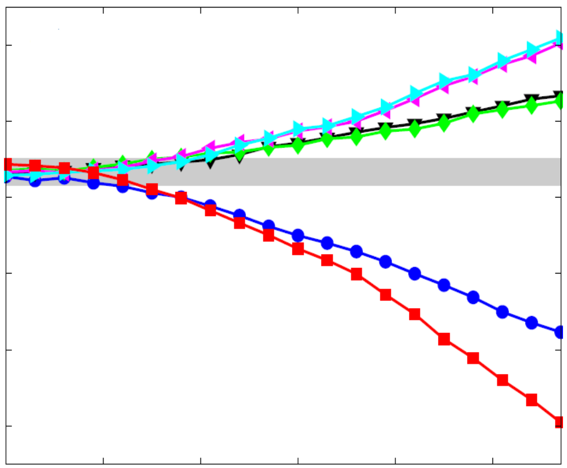
Neuron model **Diode laser with feedback**

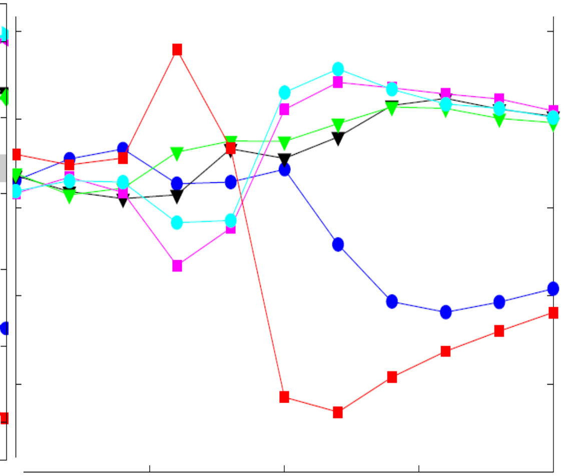






Ordinal probabilities





Modulation amplitude Modulation amplitude

J. M. Aparicio-Reinoso et al PRE 94, 032218 (2016) A. Aragoneses et al, Sci. Rep. **4**, 4696 (2014)

cristina.masoller@upc.edu @cristinamasoll1

Single-neuron vs ensemble encoding

- Single-neuron encoding: **slow** because long spike sequences are needed to estimate the ordinal probabilities.
- Ensemble encoding: can be **fast** because, from the ISI sequences of all the neurons, few spikes per neuron can be enough to accurately estimate the probabilities.

Weak, subthreshold signal

$$\epsilon \dot{u}_i = u_i - \frac{u_i^3}{3} - v_i + a_0 \cos(2\pi t/T) + \frac{\sigma}{k_i} \sum_j^N a_{ij}(u_j - u_i) + \sqrt{2D}\xi_i(t), \qquad i \neq j$$

$$\dot{v}_i = u_i + a.$$

$k_i = \sum_j a_{ij}$

$a_{ij} = a_{ji} = 1$

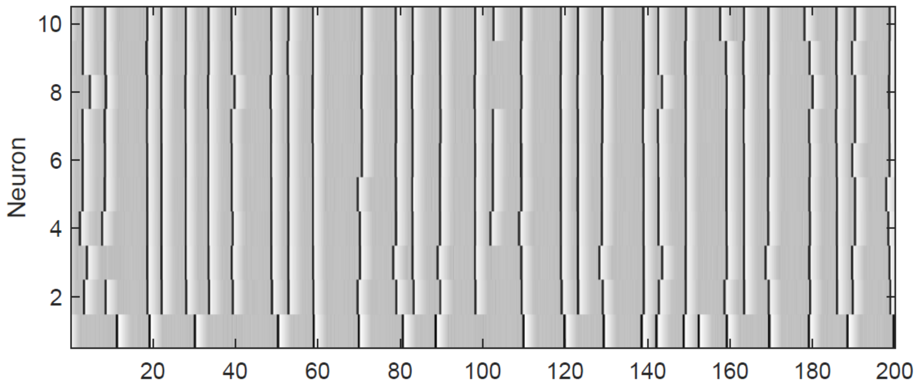
$a_{ij} = a_{ji} = 0$

M. Masoliver and C. Masoller, "*Neuronal coupling benefits the encoding of weak periodic signals in symbolic spike patterns*", Commun. Nonlinear Sci. Numer. Simulat. 88, 105023 (2020).

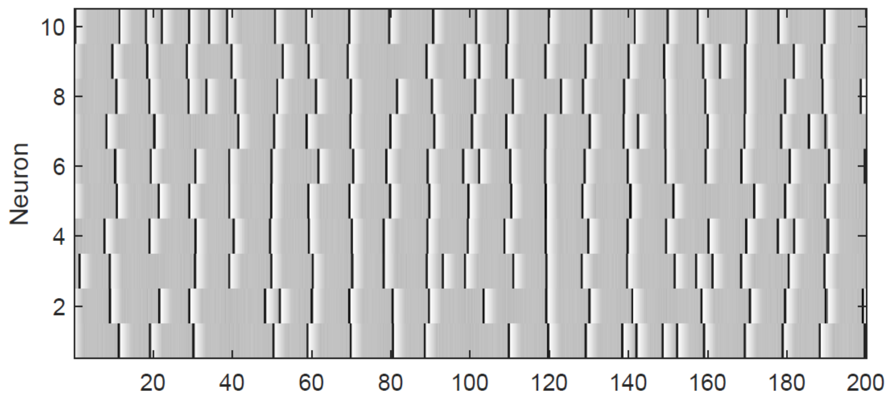
cristina.masoller@upc.edu @cristinamasoll1

Spiking dynamics with/without coupling, with/without external input

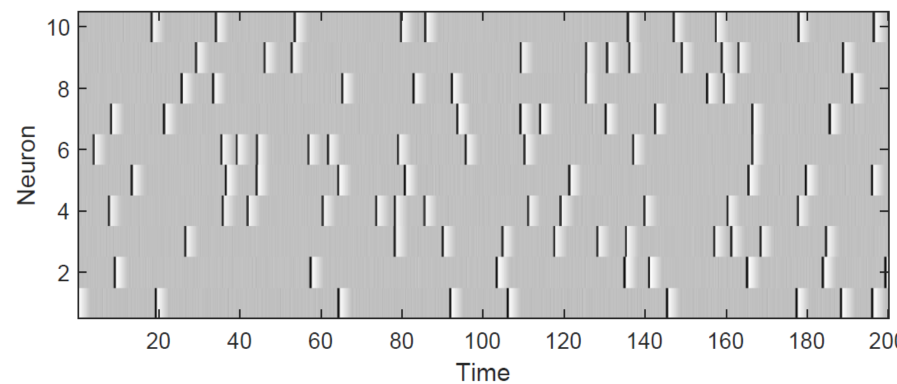
$\sigma \neq 0$
$a \neq 0$



$\sigma = 0$
$a \neq 0$

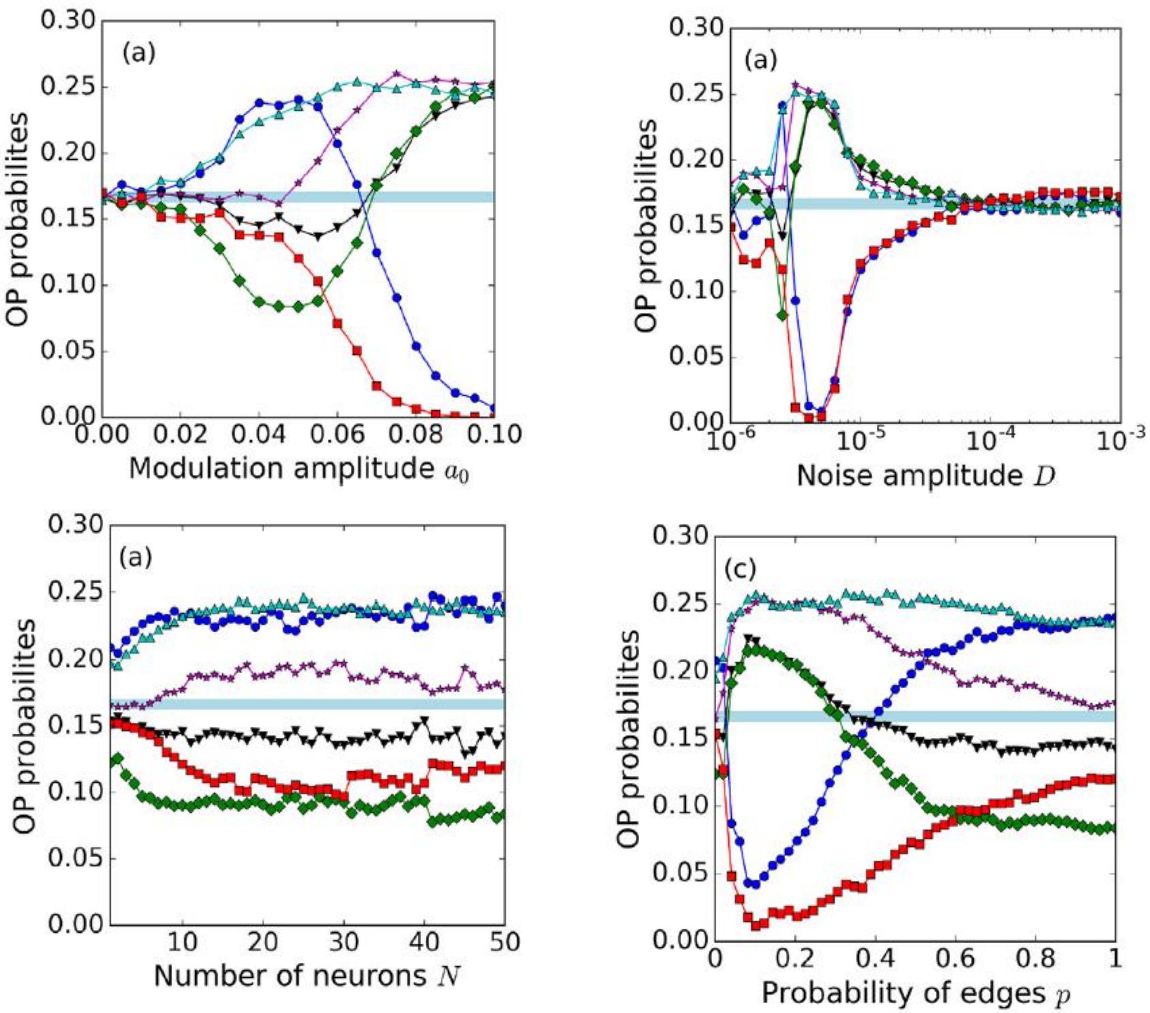


$\sigma = 0$
$a = 0$



cristina.masoller@upc.edu @cristinamasoll1

Ensemble encoding of a weak sinusoidal signal in the frequencies of occurrence of ordinal patterns



M. Masoliver and C. Masoller, Commun. Nonlinear Sci. Numer. Simulat. 88, 105023 (2020).

cristina.masoller@upc.edu @cristinamasoll1

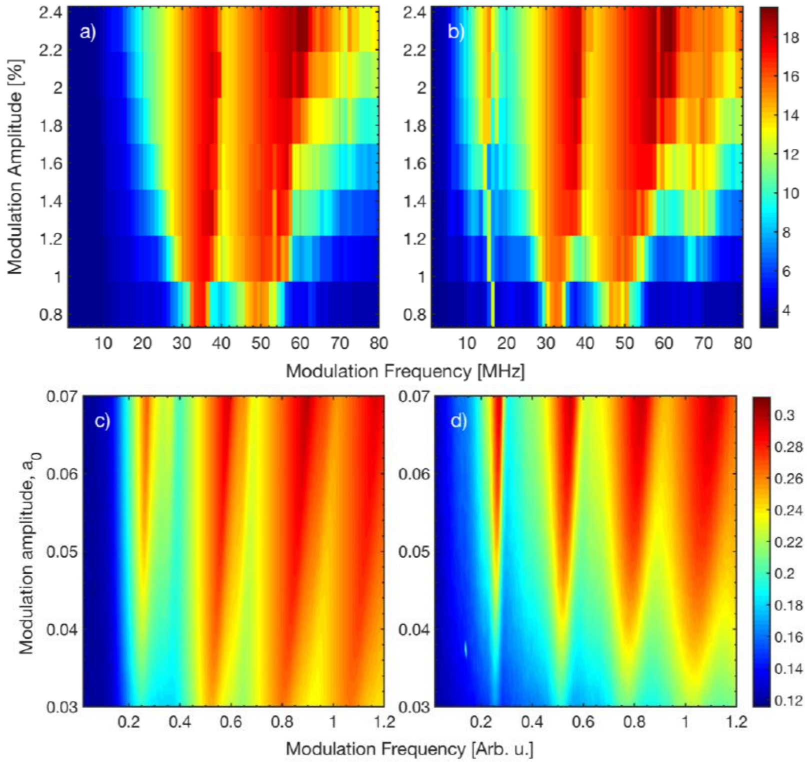
Laser-neuron comparison: encoding a weak periodic signal using spike rate code.

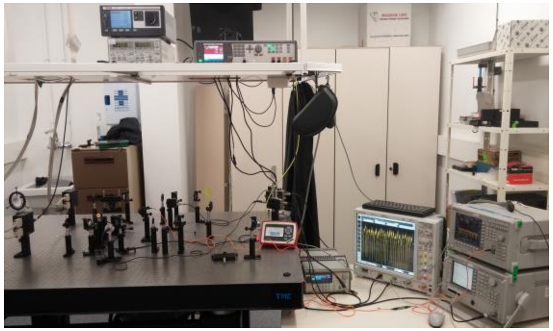
Spike rate in color code

Sinusoidal | Pulsed signal

Experiments modulating the laser current

Neuron model with the same input signal





$$\varepsilon \frac{dx}{dt} = x - \frac{x^3}{3} - y\,,$$

$$\frac{dy}{dt} = x + a + D\xi(t)\,.$$

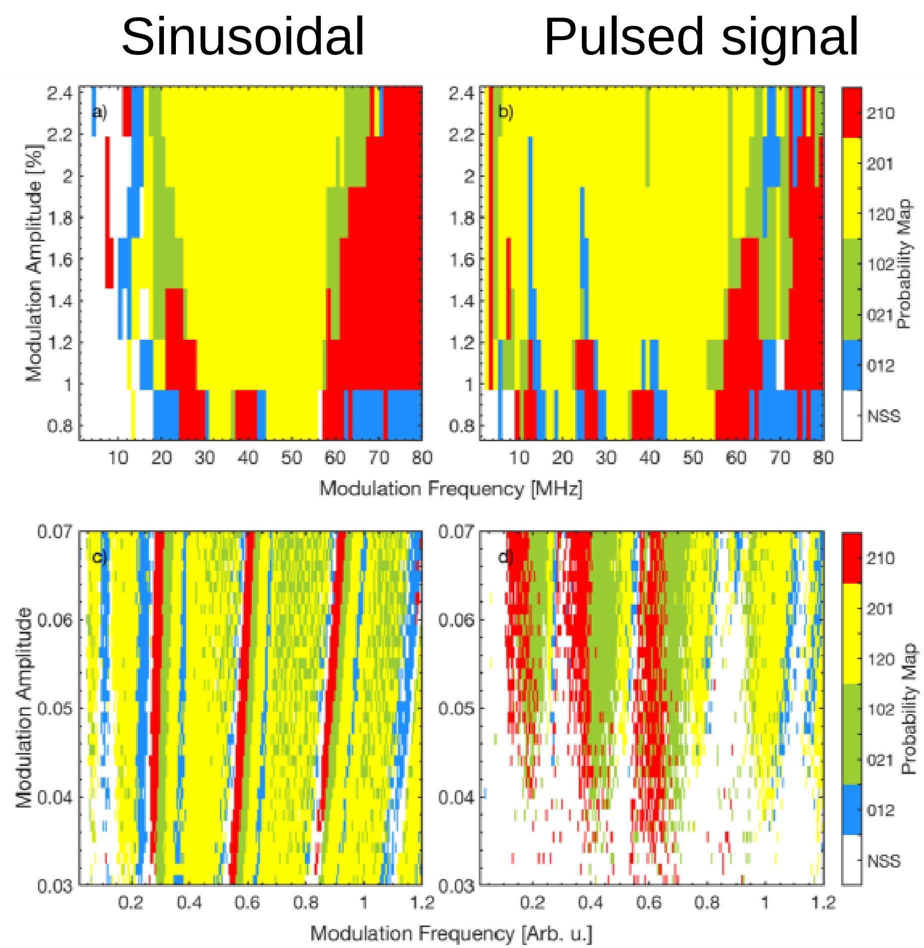
J. Tiana-Alsina, C. Quintero-Quiroz and C. Masoller, "*Comparing the dynamics of periodically forced lasers and neurons*", New J. of Phys. 21, 103039 (2019).

cristina.masoller@upc.edu @cristinamasoll1

How about the temporal code?
Ordinal analysis unveils differences in spike timing.

Diode laser with optical feedback

FitzHugh-Nagumo model

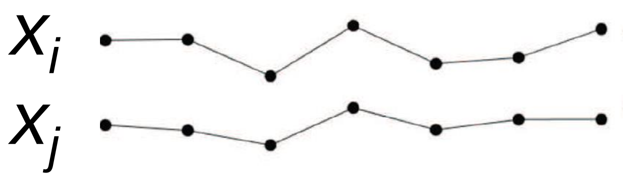


Most probable pattern in color code

J. Tiana-Alsina, C. Quintero-Quiroz and C. Masoller, New J. of Phys. 21, 103039 (2019).

cristina.masoller@upc.edu @cristinamasoll1

Ordinal analysis of bivariate data.
Are two time series statistically independent?



$\longrightarrow$ $p_i(m)$, $p_j(n)$, $p_{ij}(m,n)$ $\longrightarrow$ M_{ij}

Mutual Information: $$M_{ij} = \sum_{m,n} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)p_j(n)}$$

x_i, x_j statistically independent: $p_{ij} = p_i\, p_j \Rightarrow MI=0$

In practice: $MI > 0 \Rightarrow$ surrogate data needed to test significance

MI is not a causal measure: $MI_{ij} = MI_{ji}$

cristina.masoller@upc.edu @cristinamasoll1

A simple example to shown that MI values are overestimated

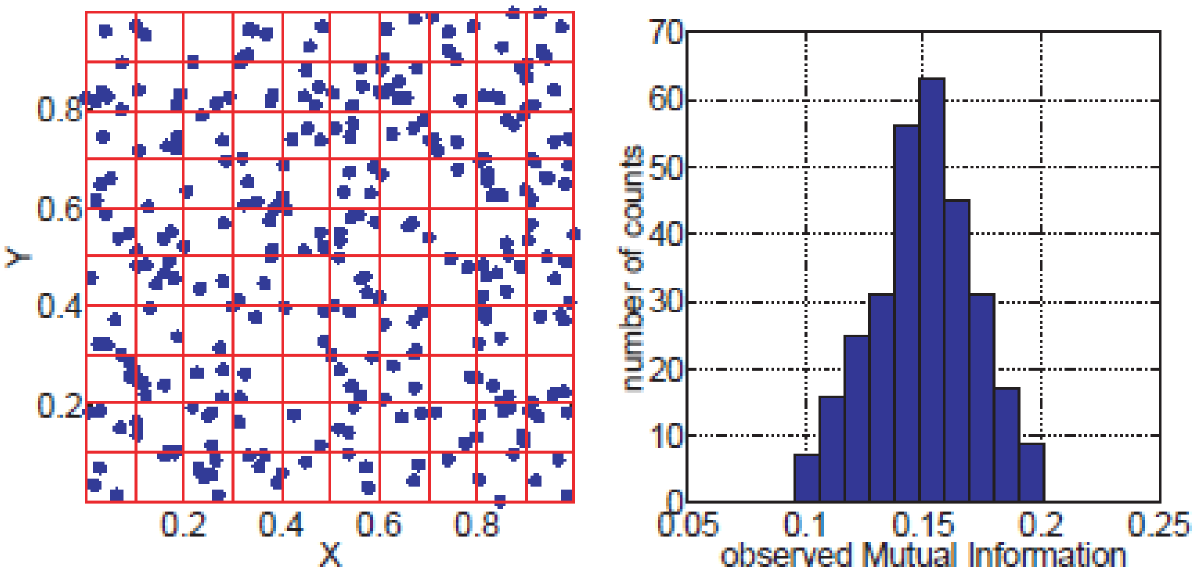


Fig. 1. Naive estimation of the mutual information for finite data. Left: The dataset consists of $N = 300$ artificially generated independent and equidistributed random numbers. The probabilities are estimated using a histogram which divides each axis into $M_X = M_Y = 10$ bins. Right: The histogram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

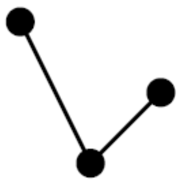
R. Steuer et al, Bioinformatics 18, suppl 2, S231 (2002).

Problem: a reliable estimation of MI requires a large amount of data.

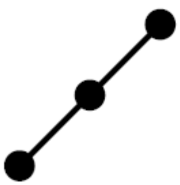
Using D=3 ordinal patterns (6 possible patterns, 36 possible combinations for p_{ij}) we need at least 400 data points in each time series.

Using lagged points to define the patterns allows to select the time scale of the analysis, very useful for seasonal data

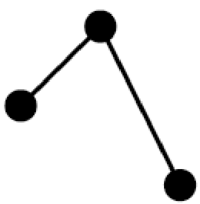
Intra-season

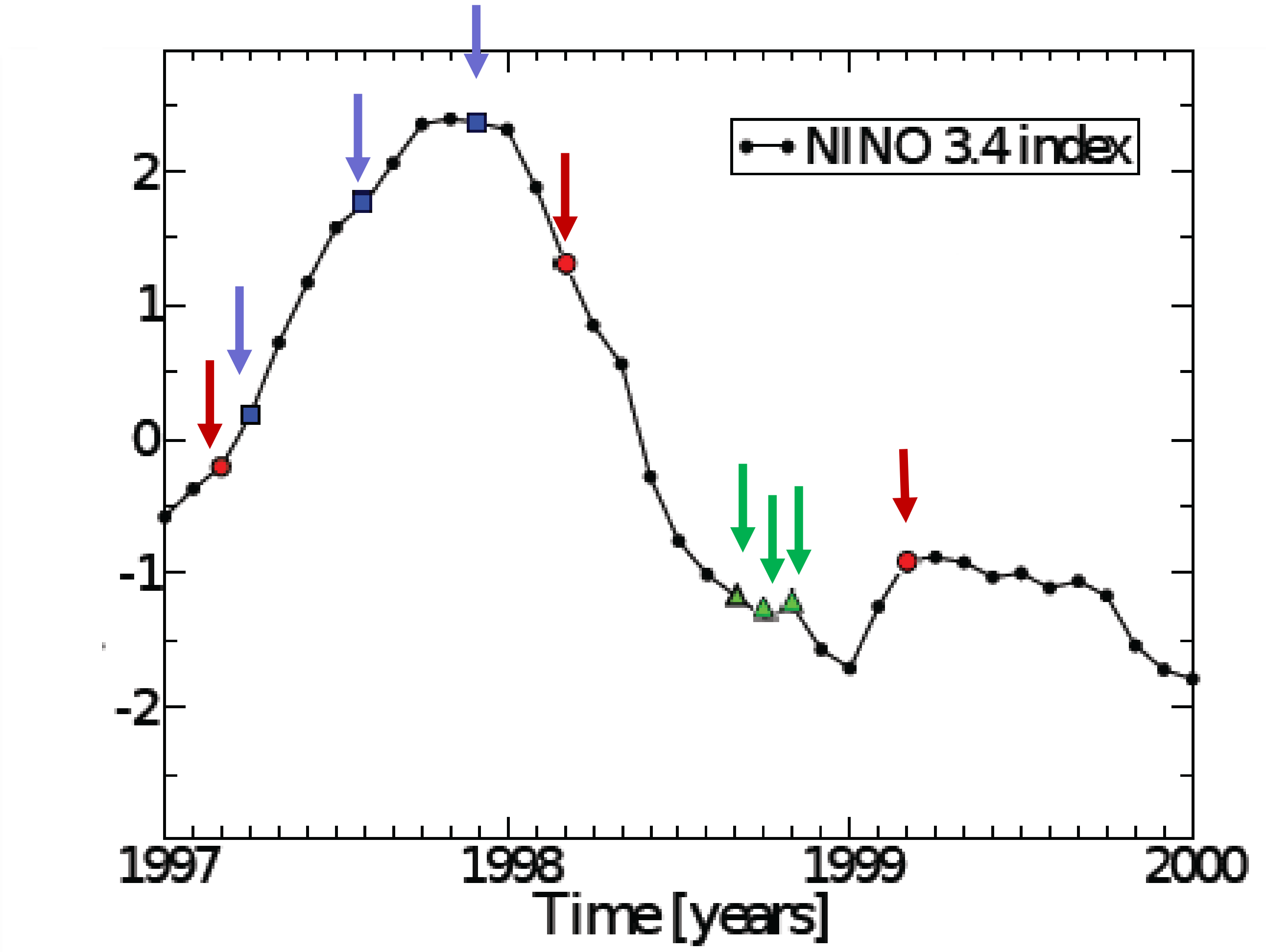


Intra-annual



Inter-annual



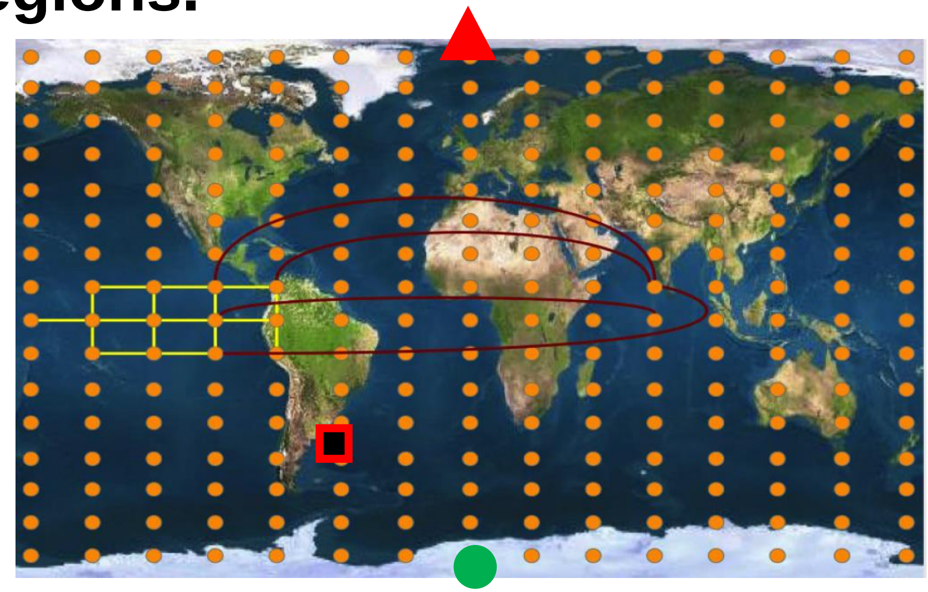


cristina.masoller@upc.edu @cristinamasoll1

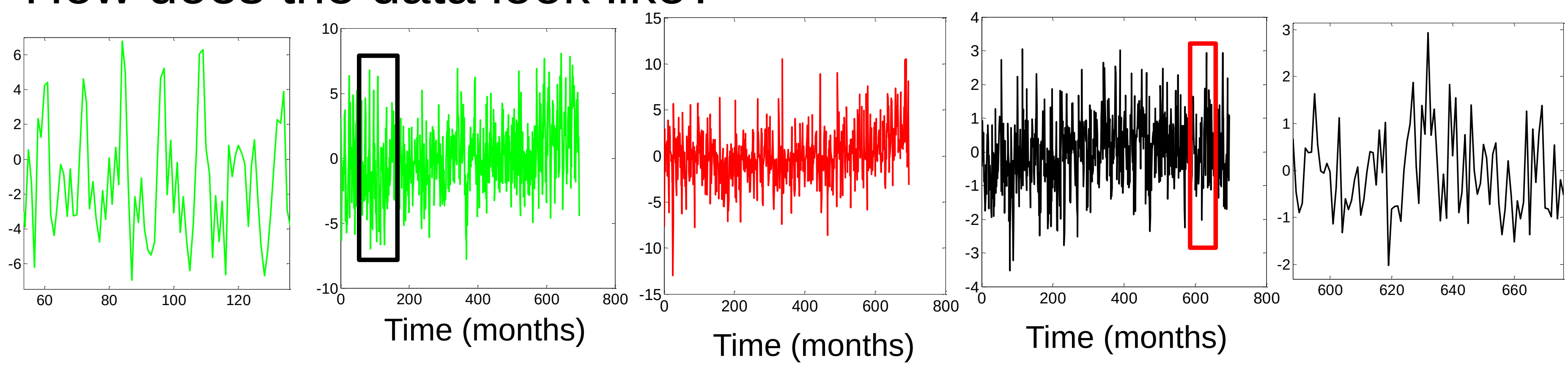
Application: analysis of surface air temperature (SAT) anomaly in two geographical regions.

Anomaly = annual solar cycle removed

Reanalysis (data assimilation)
2.5º x 2.5º =10226 grid points.
In each point 696 anomaly values
(1949-2006: 58 years x 12 months)



How does the data look like?

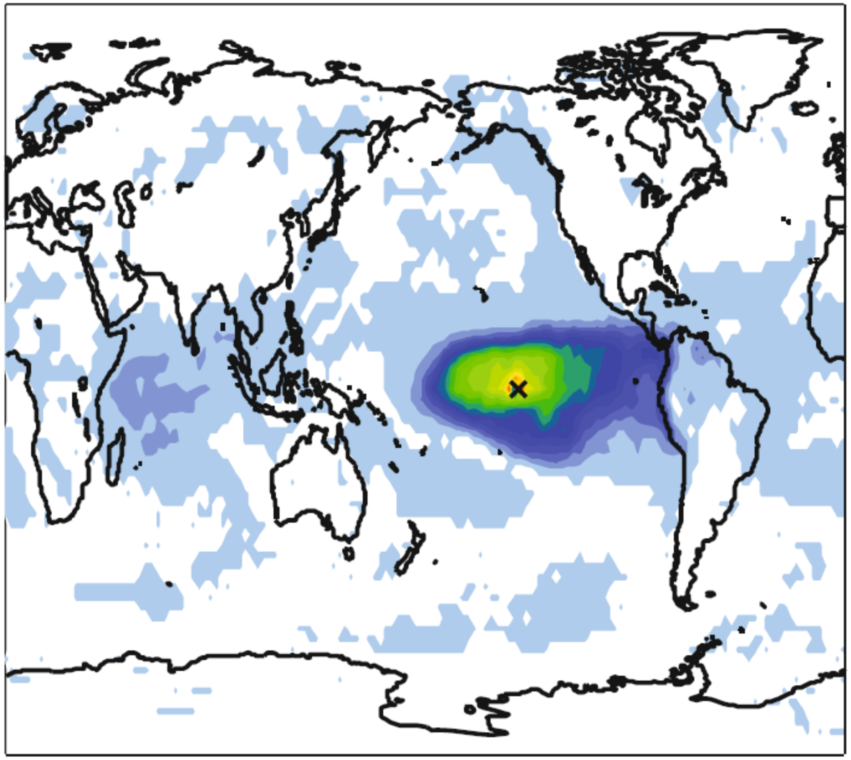


cristina.masoller@upc.edu @cristinamasoll1

Mutual Information (color code) of SAT anomaly in El Niño region and other regions (white: MI not significant)

$$M_{ij} = \sum_{m,n} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)p_j(n)}$$

MI from probabilities of SAT anomaly values



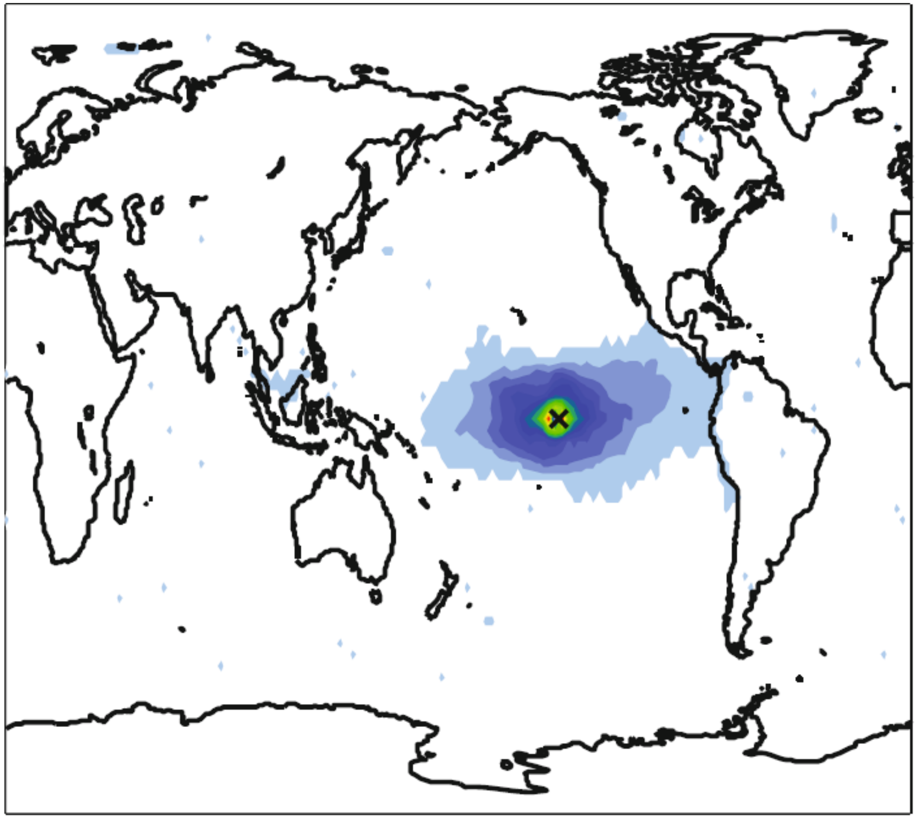
J. I. Deza, M. Barreiro, C. Masoller, "Inferring interdependencies in climate networks constructed at inter-annual, intra-season and longer time scales", Eur. Phys. J. ST 222, 511 (2013).

cristina.masoller@upc.edu @cristinamasoll1

Mutual Information (color code) of SAT anomaly in El Niño region and other regions (white: MI not significant)

$$M_{ij} = \sum_{m,n} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)p_j(n)}$$

MI from probabilities of ordinal patterns defined by values in 3 consecutive months.



J. I. Deza, M. Barreiro, C. Masoller, "Inferring interdependencies in climate networks constructed at inter-annual, intra-season and longer time scales", Eur. Phys. J. ST 222, 511 (2013).

cristina.masoller@upc.edu @cristinamasoll1

Comparison

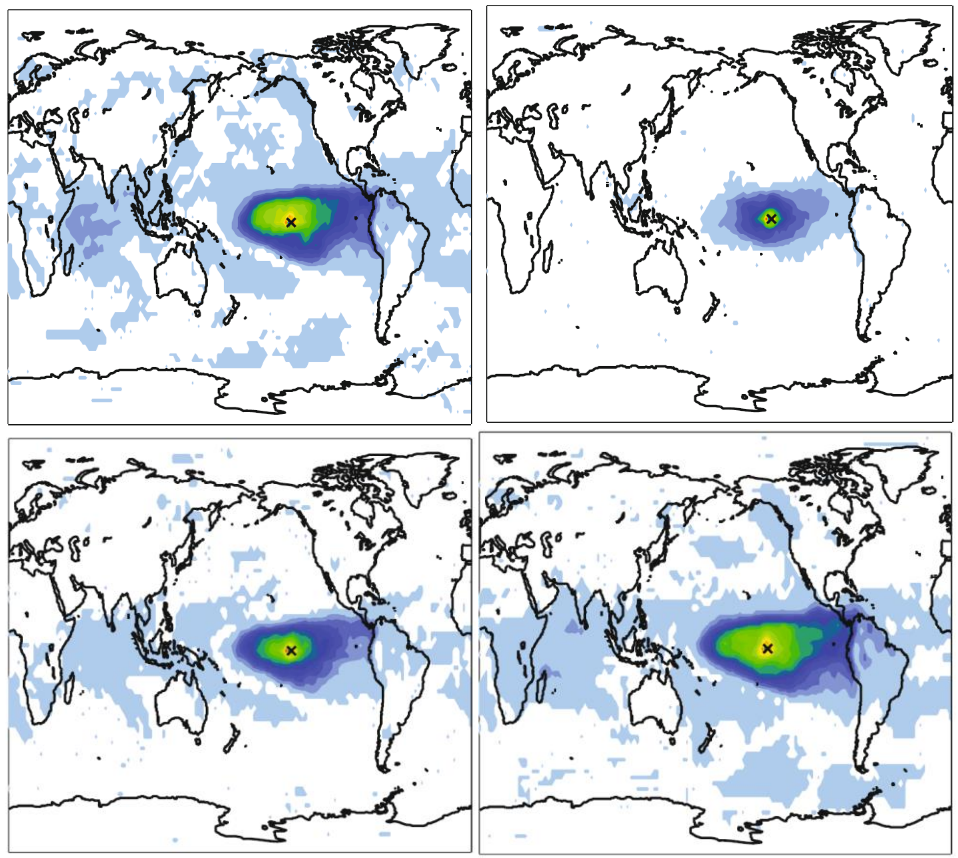
$$M_{ij} = \sum_{m,n} p_{ij}(m,n) \log \frac{p_{ij}(m,n)}{p_i(m)p_j(n)}$$

probabilities of SAT values

probabilities of ordinal patterns defined by values in 3 consecutive months.

probabilities of patterns defined by 3 values in a year.

probabilities of patterns defined by values in 3 consecutive years.



J. I. Deza, M. Barreiro, C. Masoller, "Inferring interdependencies in climate networks constructed at inter-annual, intra-season and longer time scales", Eur. Phys. J. ST 222, 511 (2013).

cristina.masoller@upc.edu @cristinamasoll1

Outline

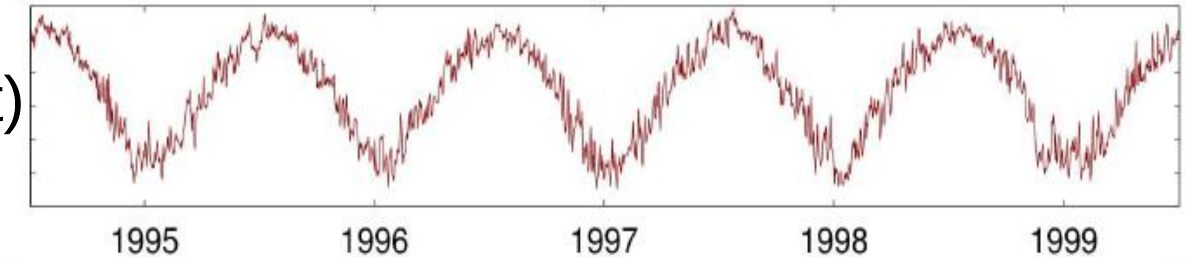
- Complex systems and time series analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: laser, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

cristina.masoller@upc.edu @cristinamasoll1

Hilbert Transform applied to Surface Air Temperature (SAT)

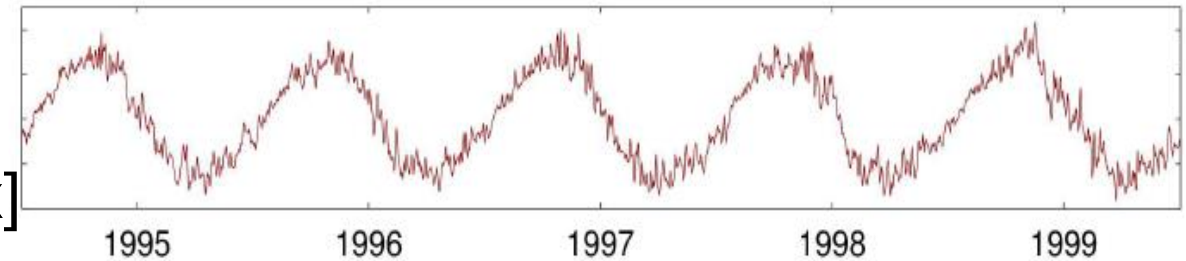
SAT in a geographical region

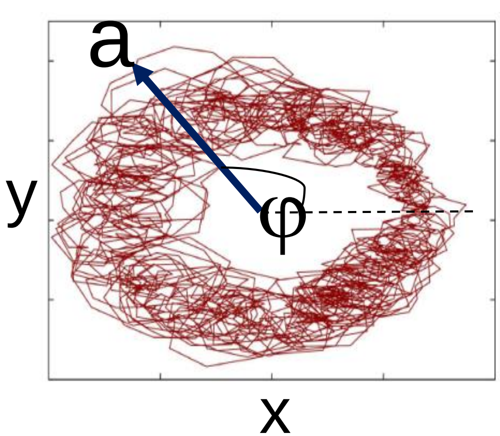
x(t)



HT[sin(x)]=cos(x)

y(t)= HT[x]





Instantaneous amplitude and phase

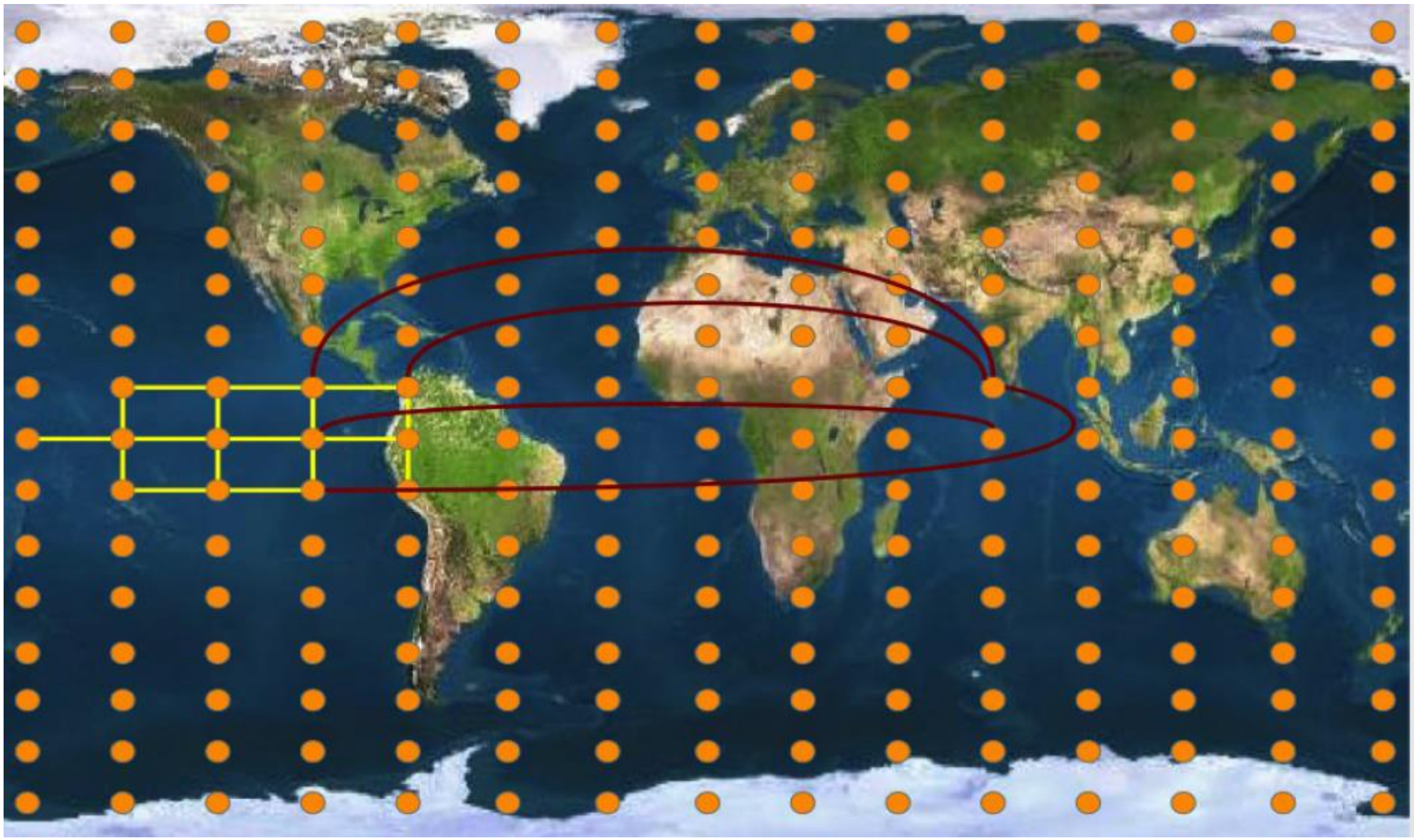
$$a(t) = \sqrt{[x(t)]^2 + [y(t)]^2}$$

$$\varphi(t) = \arctan[y(t)/x(t)]$$

Clear physical meaning only if *x(t)* is a narrow-band signal. Then, *a(t)* coincides with the envelope of *x(t)* and $\omega(t)=d\varphi/dt$, coincides with the main frequency in the spectrum.

cristina.masoller@upc.edu @cristinamasoll1

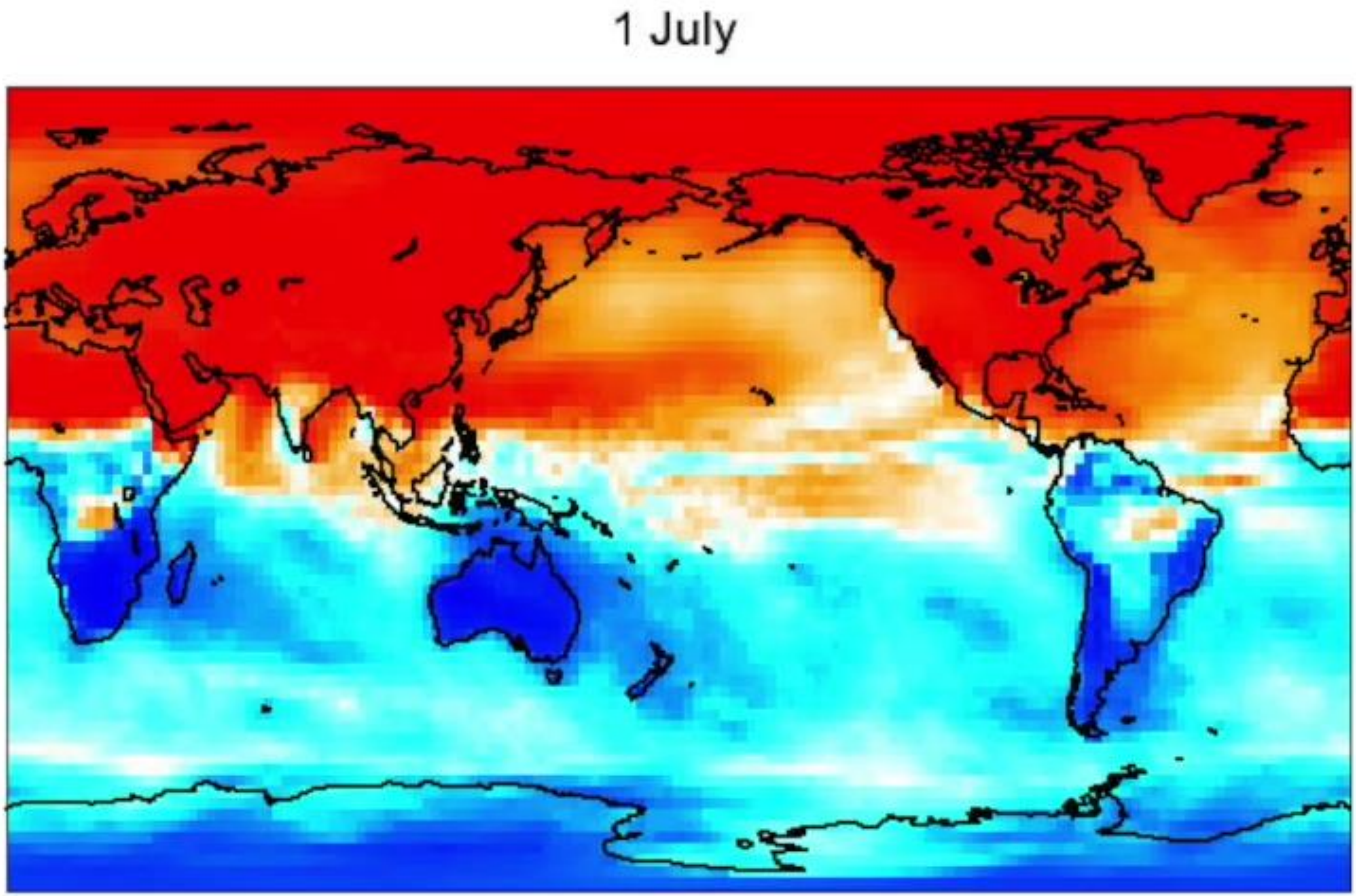
Using the HT we analyzed "re-analysis data" from the *European Centre for Medium-Range Weather Forecasts*, with high spatial and temporal resolution in the period 1979-2016



73 x 144 = 10 512 geographical sites, in each site the SAT time series has 13696 days

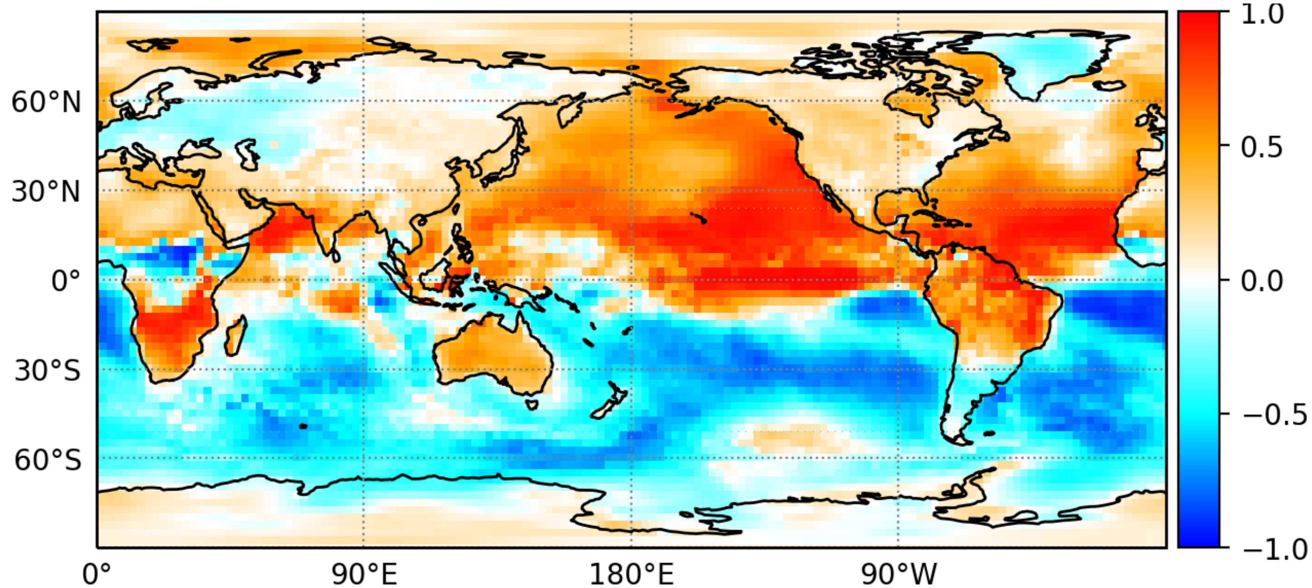
cristina.masoller@upc.edu @cristinamasoll1

Average of the cosine of the Hilbert phase

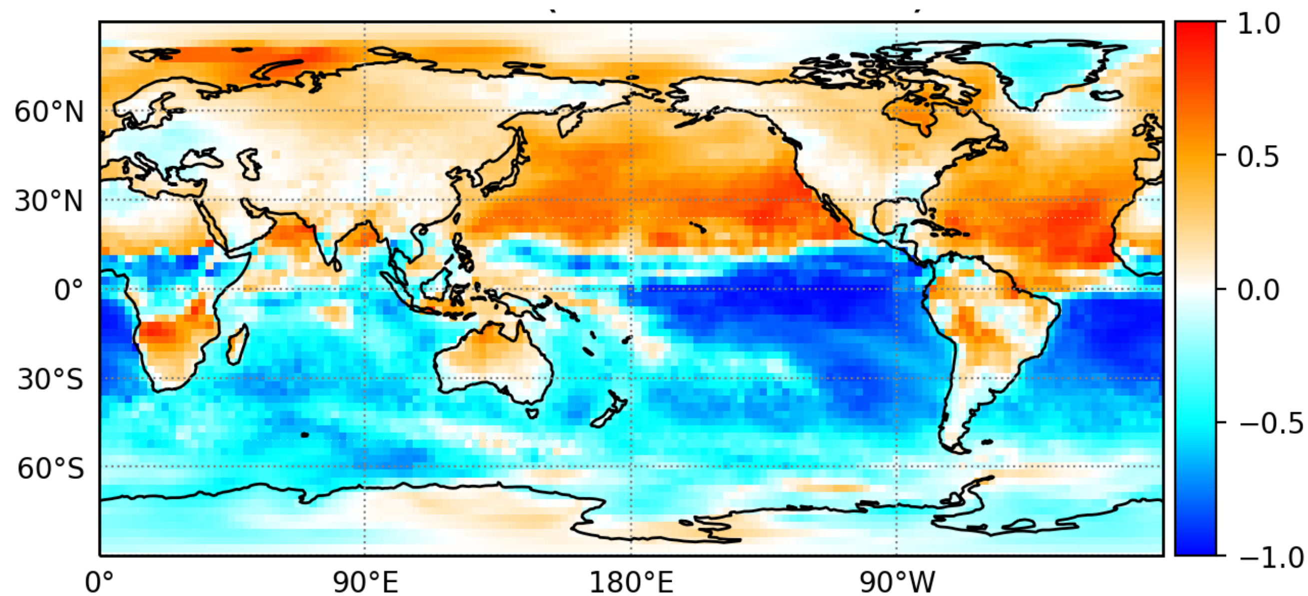


cristina.masoller@upc.edu @cristinamasoll1

Cosine of Hilbert phase during an *El Niño* period (October 2011)



Cosine of Hilbert phase during a *La Niña* period (October 2011)



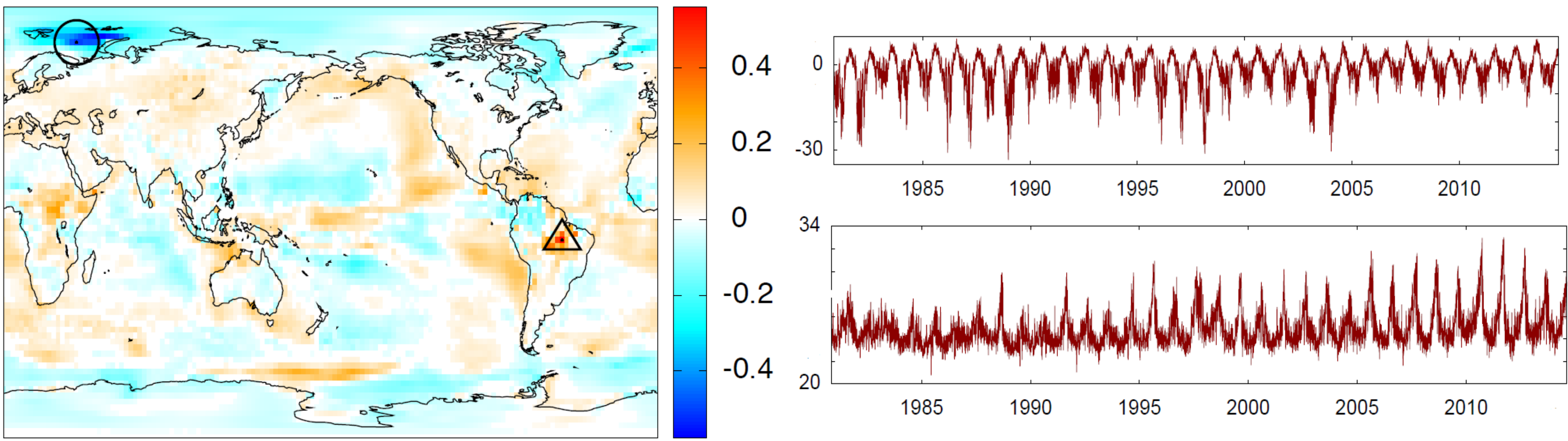
cristina.masoller@upc.edu @cristinamasoll1

How to detect significant changes in the last 30 years?

$$\Delta a = \langle a \rangle_{2016-2007} - \langle a \rangle_{1988-1979} \qquad \frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

Significant if: $\frac{\Delta a}{\langle a \rangle} \geq \langle . \rangle_s + 2\sigma_s$ or $\frac{\Delta a}{\langle a \rangle} \leq \langle . \rangle_s - 2\sigma_s$

with σ_s computed from 100 “surrogates”



D. A. Zappala, M. Barreiro, C. Masoller, “Quantifying changes in spatial patterns of surface air temperature dynamics over several decades”, Earth Syst. Dynam. 9, 383–391 (2018).

cristina.masoller@upc.edu @cristinamasoll1

Outline

- Complex systems and time series analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: laser, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

cristina.masoller@upc.edu @cristinamasoll1

Granger Causality

Hypothesis: X_1 and X_2 can be described by stationary autoregressive linear models.

Sir. Clive Granger

$$X_1(t) = \underbrace{\sum_{j=1}^{p} A_{11,j}\, X_1(t-j)}_{\text{past of } X_1} + \underbrace{E_1(t)}_{\text{Residual error}}$$

$$X_1(t) = \underbrace{\sum_{j=1}^{p} A_{11,j}\, X_1(t-j)}_{\text{past of } X_1} + \underbrace{\sum_{j=1}^{p} A_{12,j} X_2(t-j)}_{\text{past of } X_2} + \underbrace{E'_1(t)}_{\text{Residual error}}$$

If $\langle E'_1(t)\rangle < \langle E_1(t)\rangle$ ⟹ $X_2 \rightarrow X_1$

C. W. J. Granger *Investigating causal relations by econometric models and cross-spectral methods*. *Econometrica* 37, 424–438 (1969).

cristina.masoller@upc.edu @cristinamasoll1

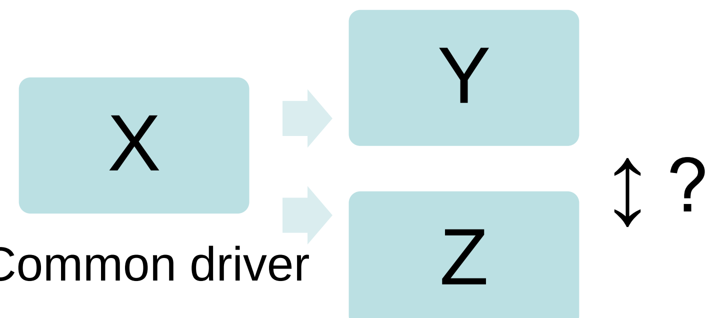
Transfer Entropy (TE)

- Measures the amount of transfer of information between two random processes.
- TE: *Conditional* Mutual Information, given the past of one of the variables.

$$TE\ (x,y) = MI\ (x,\ y|x_\tau) \qquad TE\ (y,x) = MI\ (y,\ x|y_\tau)$$

- TE and GC are equivalent for Gaussian processes.
- Problems of TE and GC:

Direct or indirect link X→Z?



Common driver

Thomas Schreiber, Measuring information transfer, Phys. Rev. Lett. 85, 461 (2000).

cristina.masoller@upc.edu @cristinamasoll1

A “simple” solution

Use an analytical expression of the Transfer Entropy that is valid for Gaussian processes.

Does this work?

Sometimes.

cristina.masoller@upc.edu @cristinamasoll1

Data generating processes and significance analysis

DGPs: We know whether X and Y are independent or not.

	Model
Y X	M0, M1, M2
Y → X	M3, M4, M5, M6, M7, M8, M9, M10, M11, M12
Y ⇆ X	M13, M14

$$x_t = (0.01 + 0.5\, x_{t-1}^2)^{0.5} + E_{1t} \qquad y_t = 0.5\, y_{t-1} + E_{2t}$$

$$x_t = 0.6\, x_{t-1} + 0.5\, y_{t-1} + E_{1t} \qquad y_t = 0.5\, y_{t-1} + E_{2t}$$

M11: two Rössler systems coupled by the first variable

$$x_t = 0.15\, x_{t-1} + 0.7\, y_{t-1} + E_{1t}$$

$$y_t = 0.1\, y_{t-1} + 0.8\, x_{t-1} + E_{2t}$$

Significance analysis: time-shifted surrogates (cheap for causality testing)

Quiroga et al., Phys. Rev. E 65, 041903 (2002).

cristina.masoller@upc.edu @cristinamasoll1

Results

Power: there is causality and we find causality (True Positives)

Size: there is no causality but we find causality (False Positives)

	Model	pTE $Y \to X$	pTE $X \to Y$	
Y X	M0	3.8	3.9	✓
	M1	2.3	2.6	
	M2	4.2	4.7	
$Y \to X$	M3	100	4.5	✓
	M4	80.7	3.8	
	M5	100	2.2	
	M6	100	1.8	
	M7	100	2.8	
	M8	100	4.5	
	M9	100	0.1	
	M10	62.6	3.1	✗
	M11	46.1	43.1	
	M12	99.9	1.0	
$Y \leftrightarrows X$	M13	100	100	✓
	M14	100	100	

cristina.masoller@upc.edu @cristinamasoll1

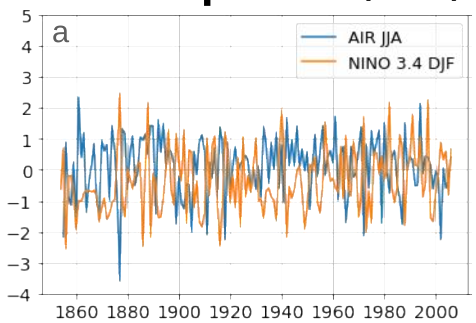
Comparison with Granger Causality and Transfer Entropy

	Model	pTE		GC		TE	
		Y → X	X → Y	Y → X	X → Y	Y → X	X → Y
Y X	M0	3.8	3.9	5.1	5.0	4.4	4.4
	M1	2.3	2.6	3.3	3.1	100	100
	M2	4.2	4.7	5.5	5.9	4.7	4.9
Y → X	M3	100	4.5	100	4.8	70.2	5.6
	M4	80.7	3.8	84.2	4.9	96.0	4.7
	M5	100	2.2	100	3.1	100	3.8
	M6	100	1.8	100	2.8	100	4.3
	M7	100	2.8	100	3.4	100	4.0
	M8	100	4.5	100	5.6	100	100
	M9	100	0.1	100	0.1	100	100
	M10	62.6	3.1	67.3	4.3	12.2	4.5
	M11	46.1	43.1	53.1	49.8	37.8	45.0
	M12	99.9	1.0	100	0.9	100	0
Y ⇆ X	M13	100	100	100	100	100	100
	M14	100	100	100	100	100	100

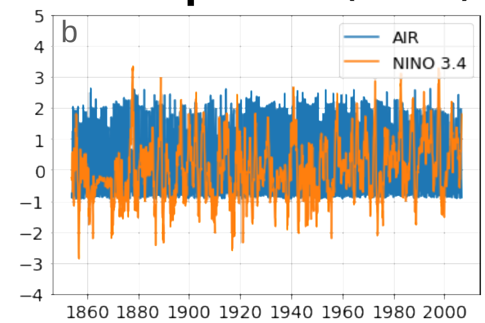
cristina.masoller@upc.edu @cristinamasoll1

Application to climate data NINO3.4 ← → All India Rainfall

Yearly sampled (152)



Monthly sampled (1836)



pTE

NINO3.4 ⟶ AIR — 0.04 s

NINO3.4 ⟵ AIR — 0.5 s

GC

NINO3.4 ⟶ AIR — 0.4 s

NINO3.4 ⟵ AIR — 0.9 s

TE

NINO3.4 ⇆ AIR — 1 s

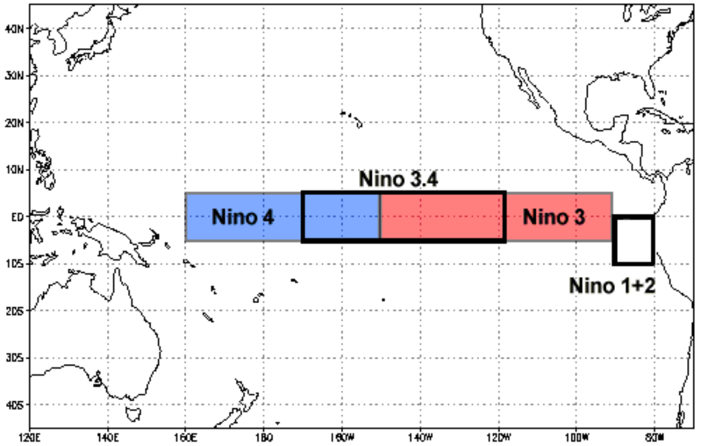
NINO3.4 ⇆ AIR — 68 s

cristina.masoller@upc.edu @cristinamasoll1

How much time we save by using "pseudo Transfer Entropy"?

For two time-series of 500 data points (1 data point per month, 40 years):

TE:**112 ms** but pTE: **4 ms**



8000 grid points (high resolution)
$\Rightarrow$ 64 x 10^6 pairs

$\Rightarrow$ **829 days** (TE) vs. **29 days** (pTE).

(without "surrogate" analysis)

But, there is a price to pay, no "free lunch".

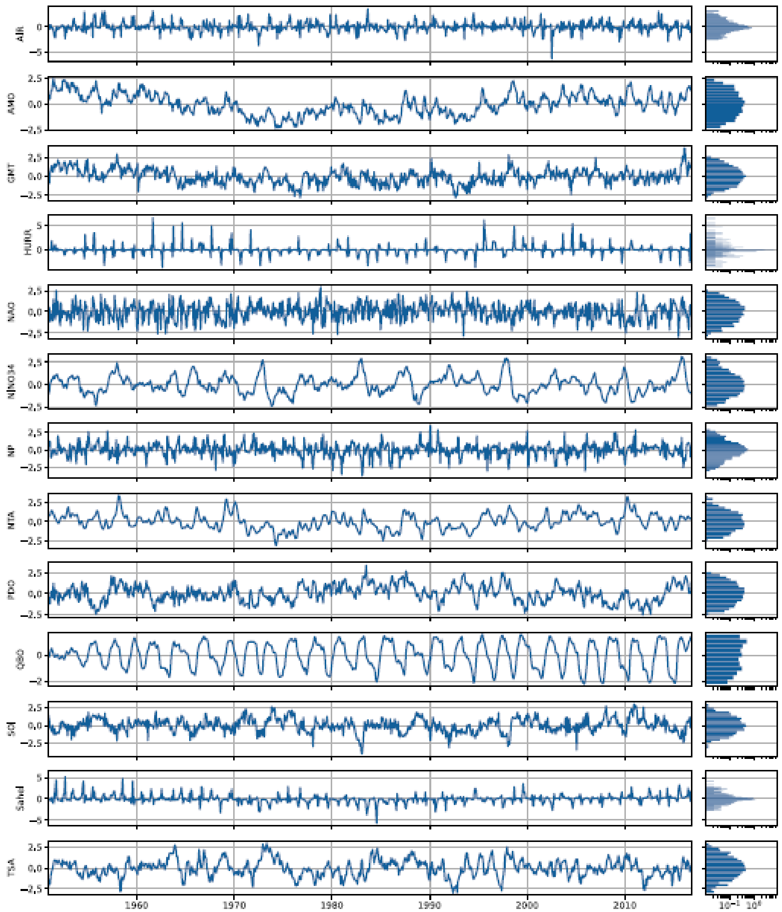
https://github.com/riccardosilini/pTE

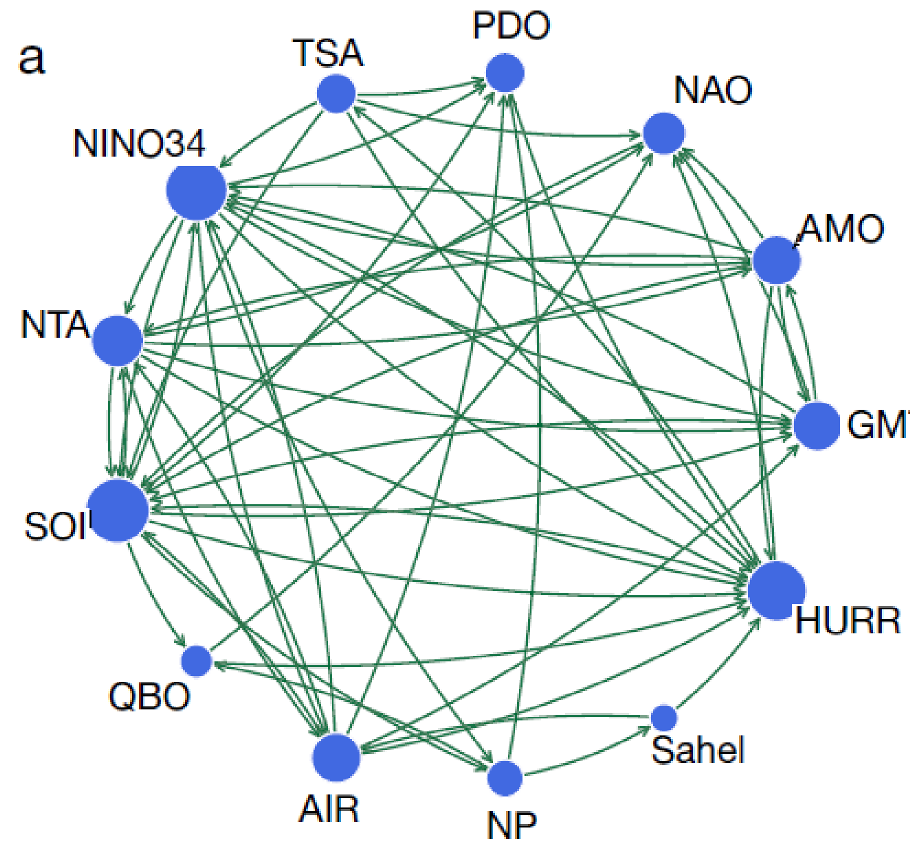
R. Silini and C. Masoller "*Fast and effective pseudo transfer entropy for bivariate data-driven causal inference*", Sci. Rep. 11, 8423 (2021).

cristina.masoller@upc.edu @cristinamasoll1

Directed network of climatic indices

Constructed using pTE with different lags





R. Silini, G. Tirabassi, M Barreiro, L. Ferranti, C. Masoller, "*Assessing causal dependencies in climatic indices*", Climate Dynamics 10.1007/s00382-022-06562-0 (2022).

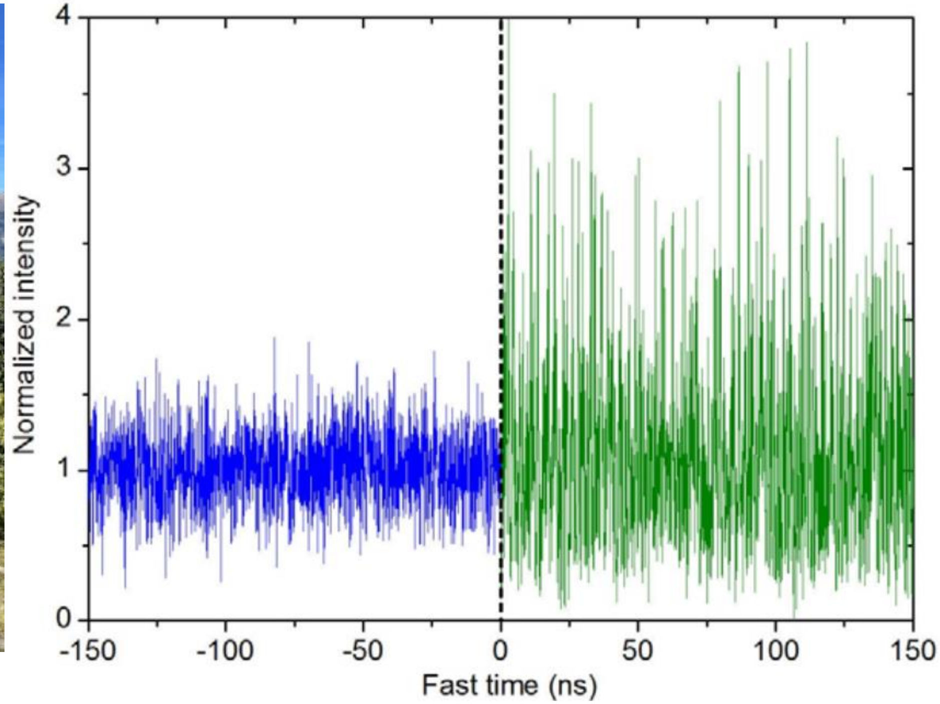
cristina.masoller@upc.edu @cristinamasoll1

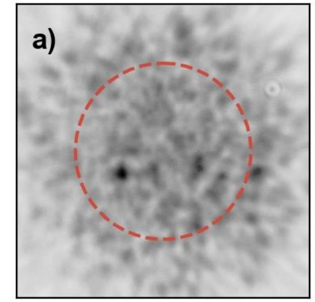
Outline

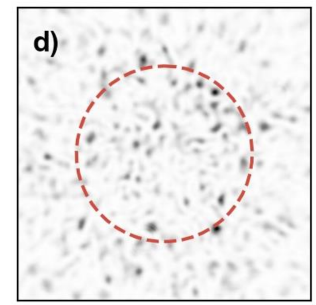
- Complex systems and time series analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: Lasers, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

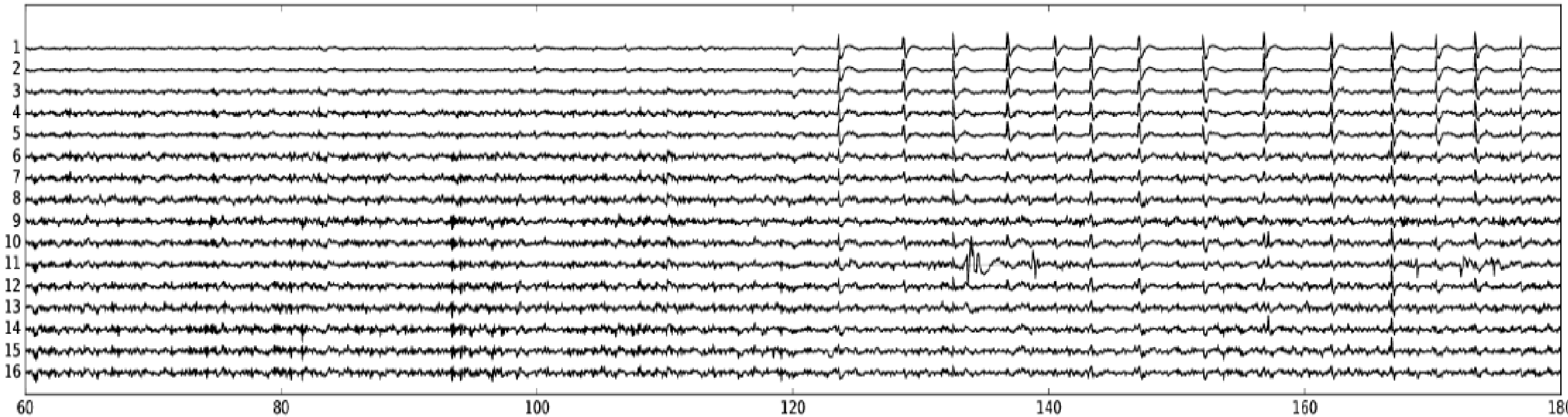
cristina.masoller@upc.edu @cristinamasoll1

Regime transitions in complex systems









cristina.masoller@upc.edu @cristinamasoll1

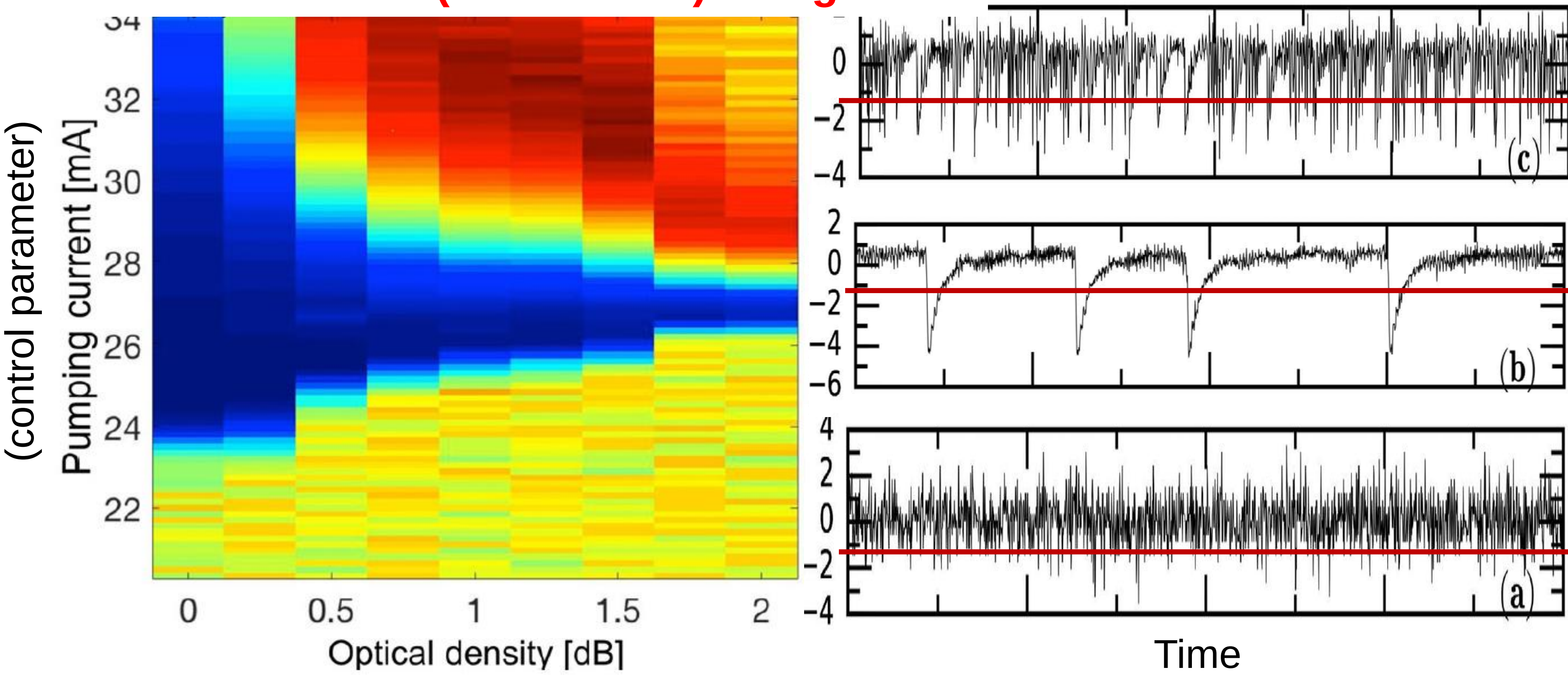
How to identify, characterize and predict regime transitions?

cristina.masoller@upc.edu @cristinamasoll1

Counting the number of extreme values allows to distinguish different dynamical regimes

$$Z = \frac{X - \mu}{\sigma}$$

Number of events (below -1.5σ) in log scale

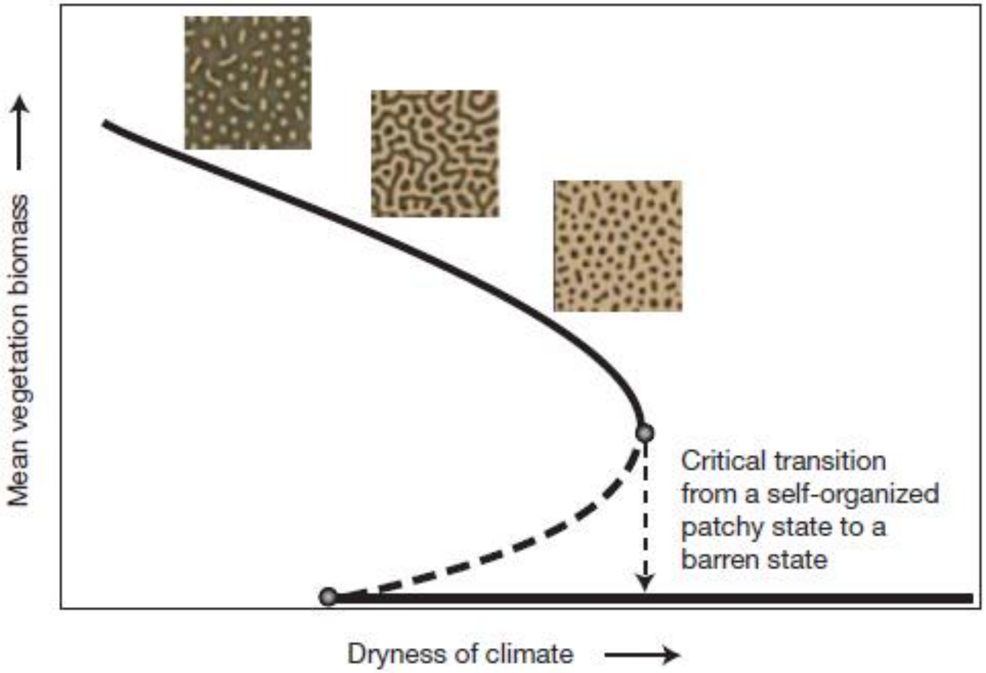


(control parameter)

Panozzo et al, Chaos 27, 114315 (2017)

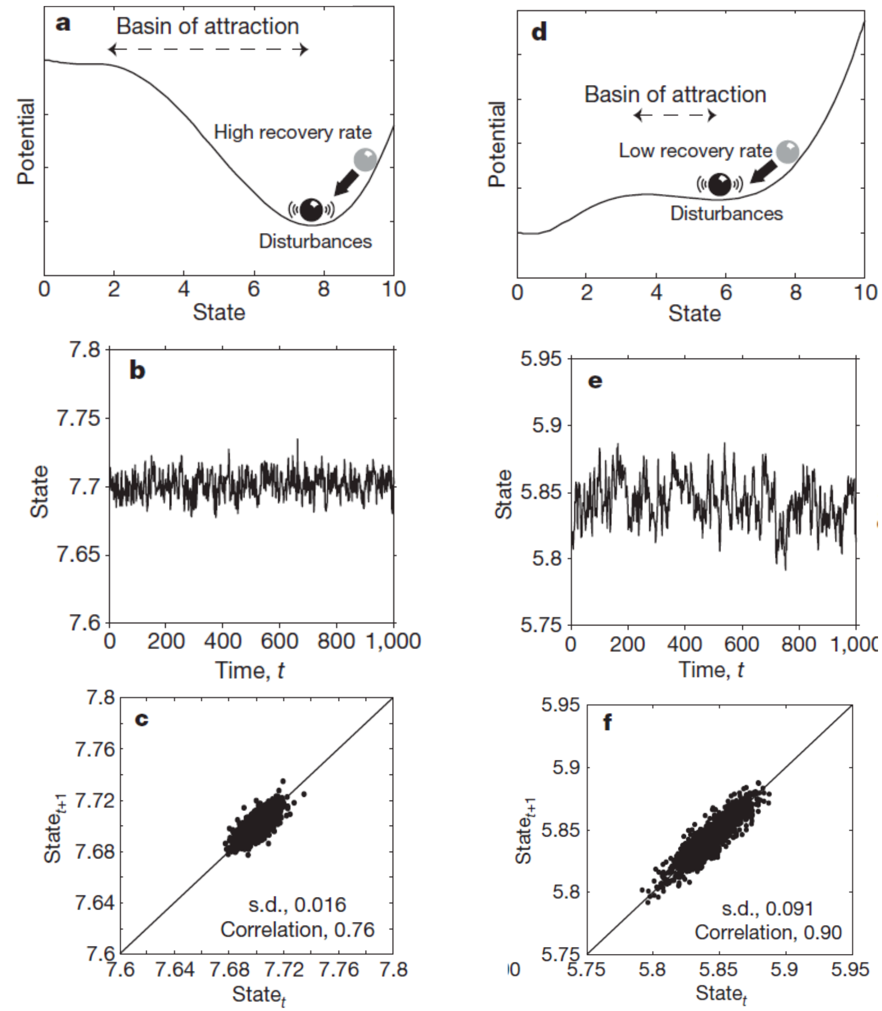
cristina.masoller@upc.edu @cristinamasoll1

Classical indicators of approaching critical transitions



M. Scheffer et al., Nature 461, 53 (2009)

Far from bif. | Close to bif.



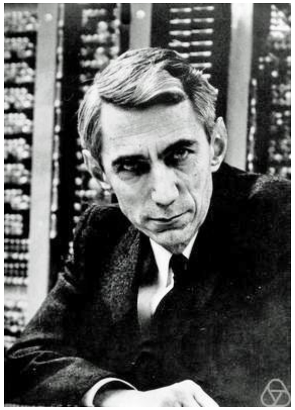
⇒ increase of variance and autocorrelation (*critical slowing down*)

cristina.masoller@upc.edu @cristinamasoll1

Shannon entropy

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

$$\boxed{H = -\sum_{i=1}^{N} p_i \ln p_i}$$



Claude Shannon

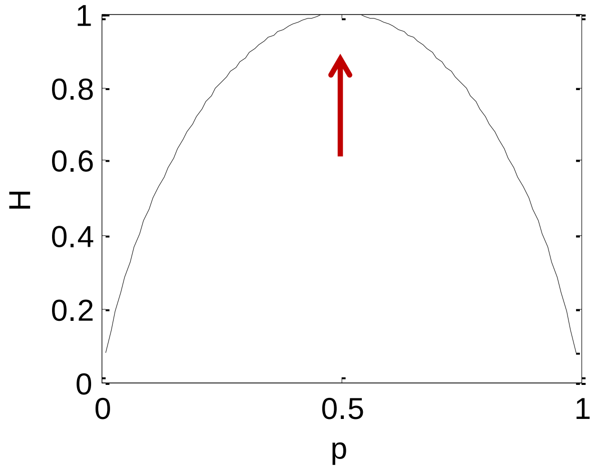
- Interpretation: "*quantity of* ***surprise*** *one should feel upon reading the result of a measurement* ".

- Example: a random variable takes values 0 or 1 with probabilities:

  p(0) = p, p(1) = 1 − p.

  H = −p ln(p) − (1 − p) ln(1 − p).

  ⇒ p=0.5: Maximum **unpredictability.**



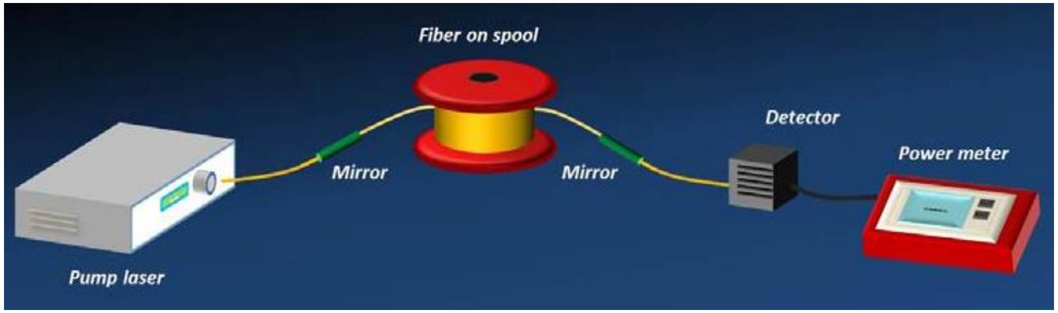
C. Shannon, "A Mathematical Theory of Communication", Bell System Technical Journal. 27 (3): 379–423 (1948). Bell System Technical Journal. 27 (4): 623–656 (1948).

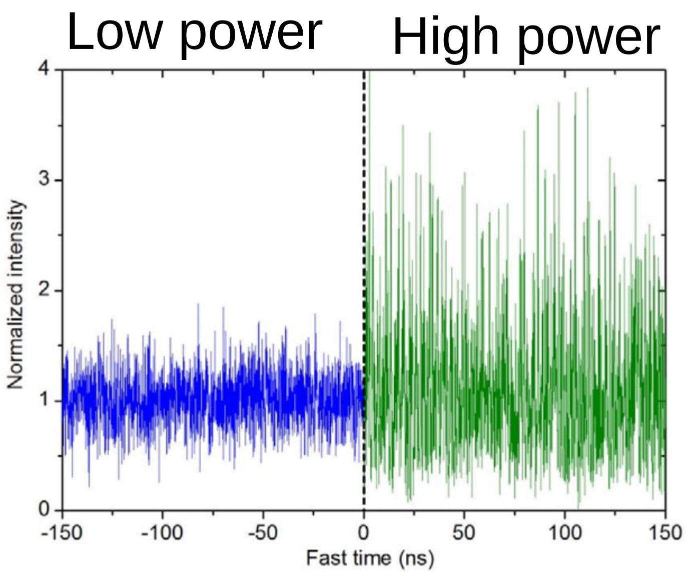
cristina.masoller@upc.edu @cristinamasoll1

Transition laminar → optical turbulence in a fiber laser

(governing equations similar to hydrodynamics)

Control parameter: power of pump laser





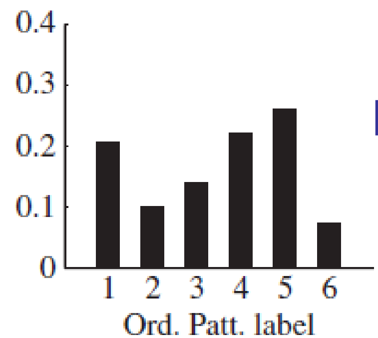
E. G. Turitsyna et. al, Nat. Photonics 7, 783 (2013).

cristina.masoller@upc.edu @cristinamasoll1

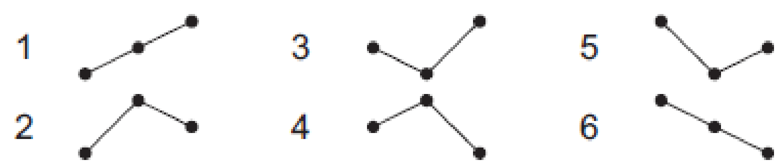
Permutation entropy: Shannon's entropy computed from ordinal probabilities

Time series

Ordinal probabilities



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



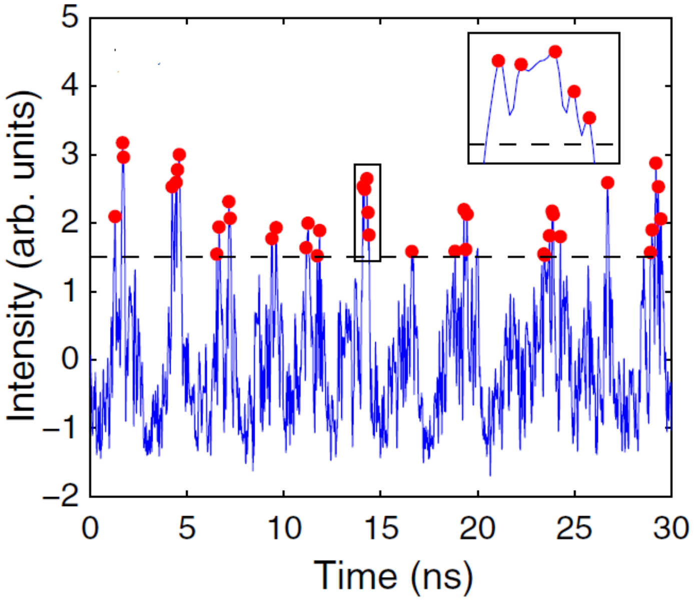
Permutation entropy

$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

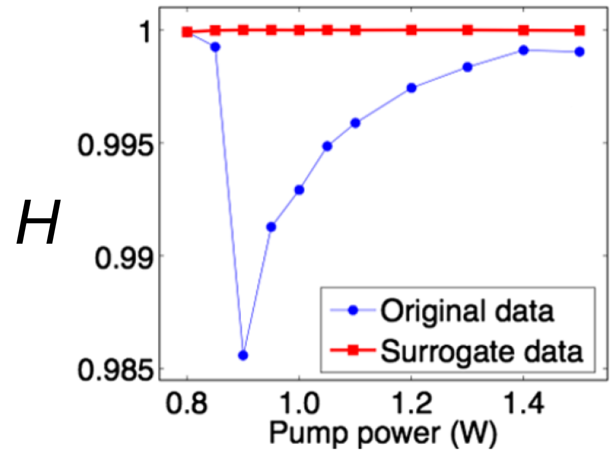
cristina.masoller@upc.edu @cristinamasoll1

Entropy characterization of the transition

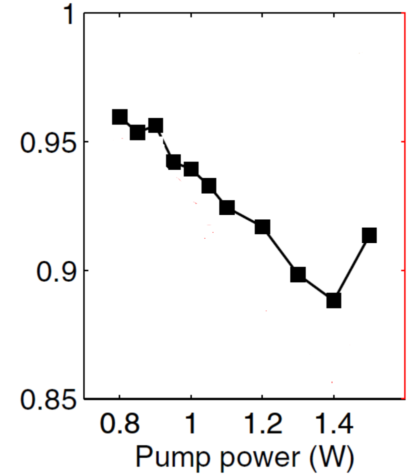
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$



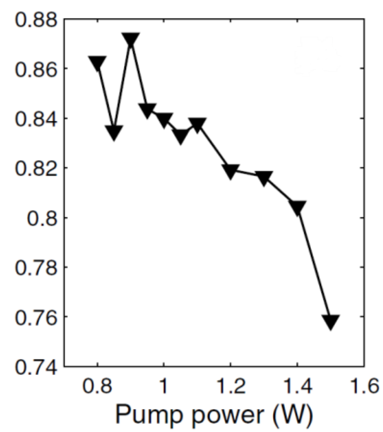
Thresholded data



All data



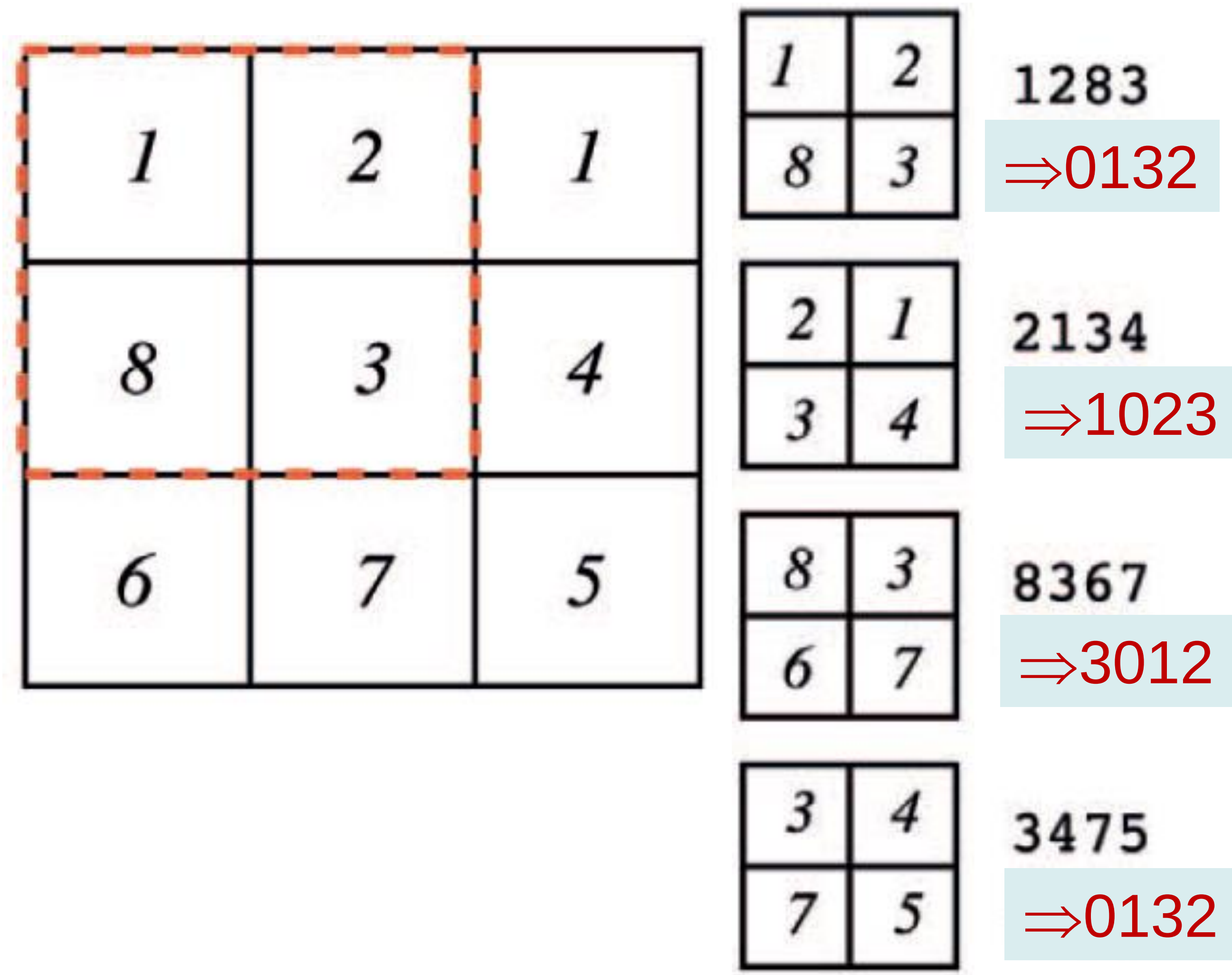
Entropy of the distribution of values. The distribution develops a tail (extreme events) and H decreases.



A. Aragoneses et al., "Unveiling temporal correlations characteristic of a phase transition in the output intensity of a fiber laser", PRL 116, 033902 (2016).

cristina.masoller@upc.edu @cristinamasoll1

Ordinal analysis of two-dimensional patterns



Spatial permutation entropy

$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

2x2 pixels: 24 possible patterns

H. V. Ribeiro et. al, PLoS ONE 7, e40689 (2012).

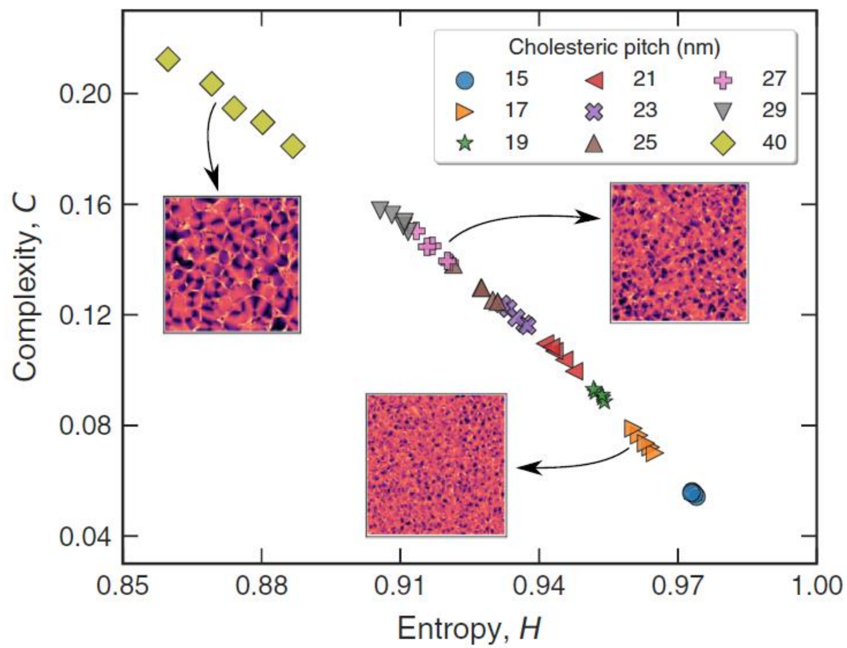
cristina.masoller@upc.edu @cristinamasoll1

The “spatial” permutation entropy was proposed to characterize 2D patterns, “textures” and images.

PHYSICAL REVIEW E **99**, 013311 (2019)

Estimating physical properties from liquid crystal textures via machine learning and complexity-entropy methods

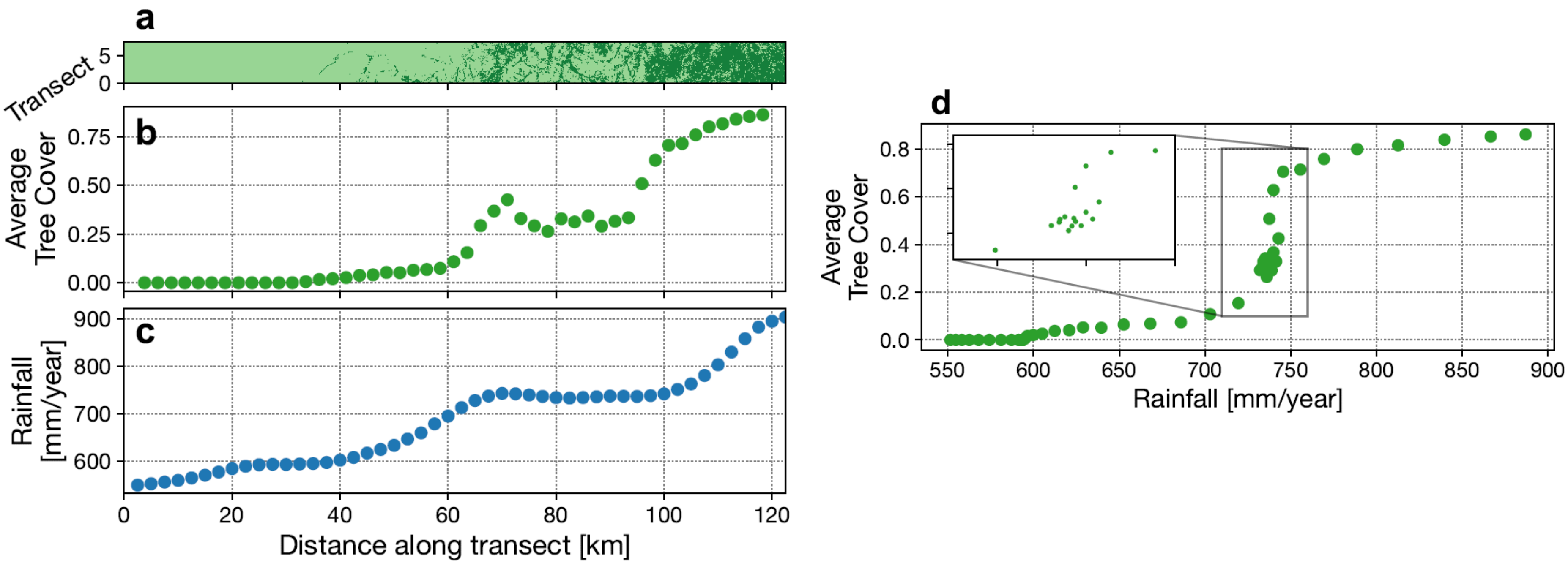
H. Y. D. Sigaki,[1] R. F. de Souza,[1] R. T. de Souza,[1,2] R. S. Zola,[1,2,*] and H. V. Ribeiro[1,†]
[1]*Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil*
[2]*Departamento de Física, Universidade Tecnológica Federal do Paraná, Apucarana, PR 86812-460, Brazil*



cristina.masoller@upc.edu @cristinamasoll1

The variation of the spatial permutation entropy can give an early indicator of a vegetation transition.

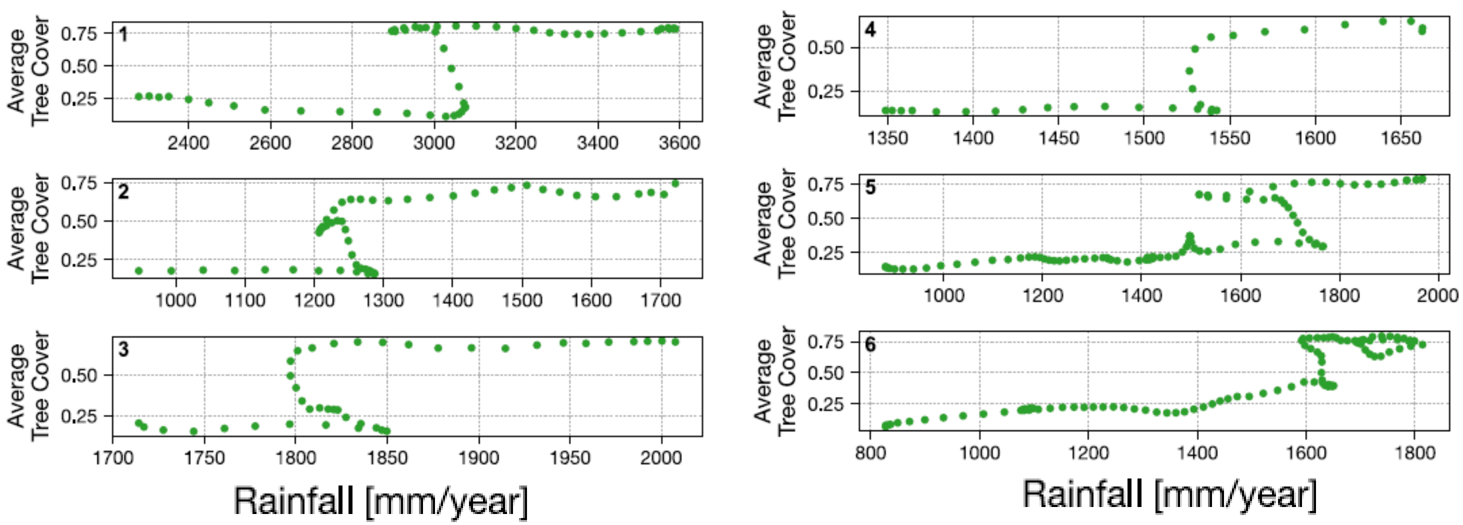
High-resolution vegetation data from the Serengeti–Mara ecosystem in northern Tanzania and southern Kenya.



G. Tirabassi, C. Masoller, "Entropy-based early detection of critical transitions in spatial vegetation fields", PNAS 120, e2215667120 (2023).

cristina.masoller@upc.edu @cristinamasoll1

We also analyzed **low-resolution** satellite (MODIS) vegetation data, combined with data from the Tropical Rainfall Measuring Mission (TRMM)



G. Tirabassi, C. Masoller, "Entropy-based early detection of critical transitions in spatial vegetation fields", PNAS 120, e2215667120 (2023).

cristina.masoller@upc.edu @cristinamasoll1

Results

Permutation entropy
(ordinal patterns defined by the values of 2x2 pixels)

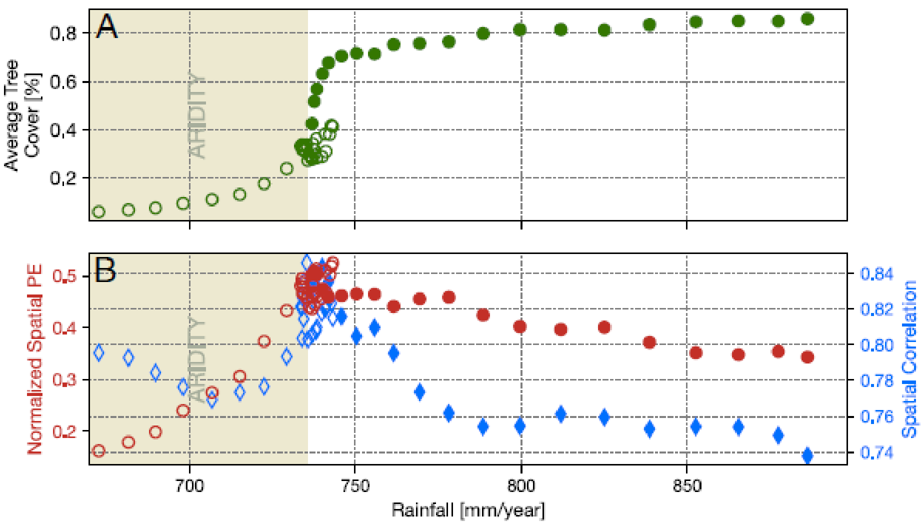
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

Spatial correlation

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_i (u_i - \bar{u})^2}$$

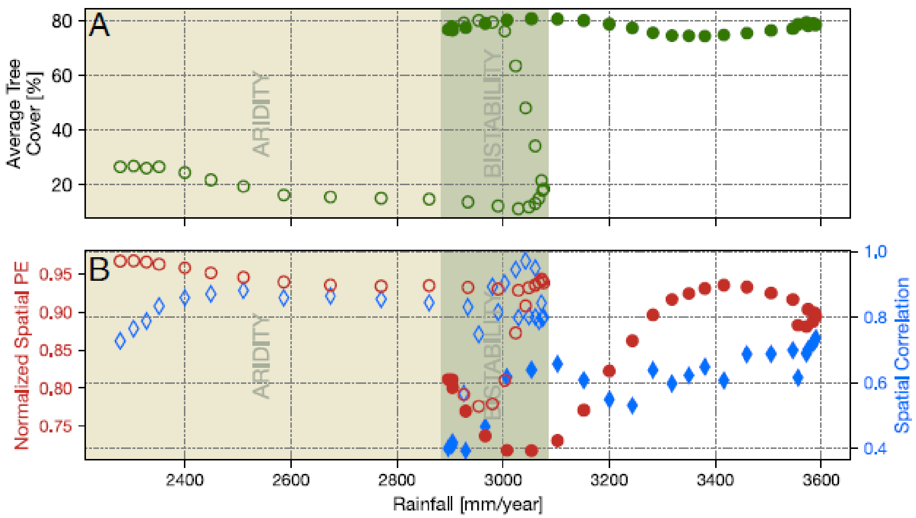
w_{ij}=1 if i, j first neighbors, else 0

High-resolution data



Low-resolution data

(transect 1; large variability across transects)

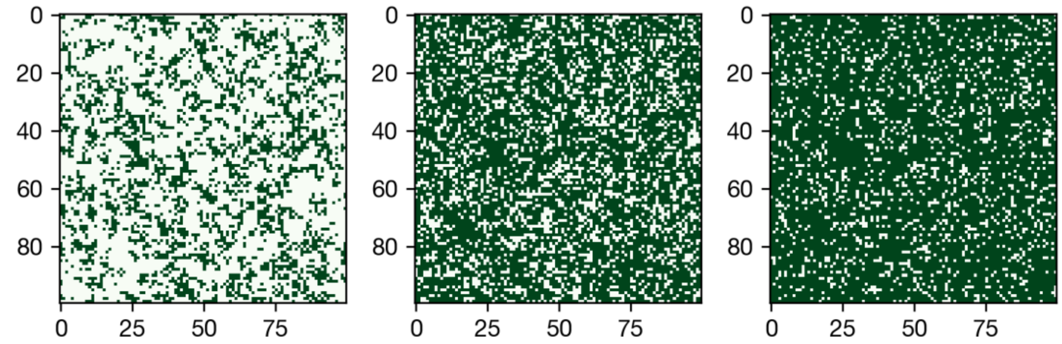


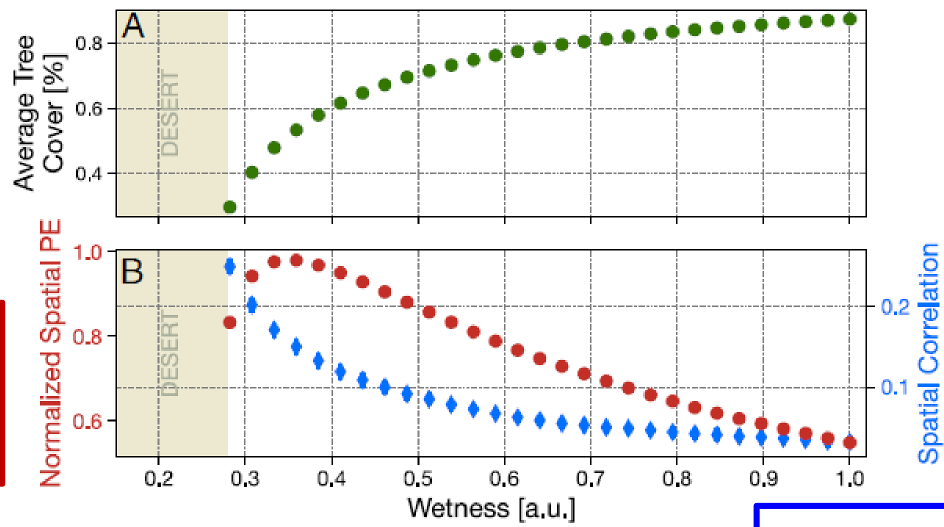
G. Tirabassi, C. Masoller, "Entropy-based early detection of critical transitions in spatial vegetation fields", PNAS 120, e2215667120 (2023).

cristina.masoller@upc.edu @cristinamasoll1

To gain insight: simulations of vegetation models

A) Cellular automata model



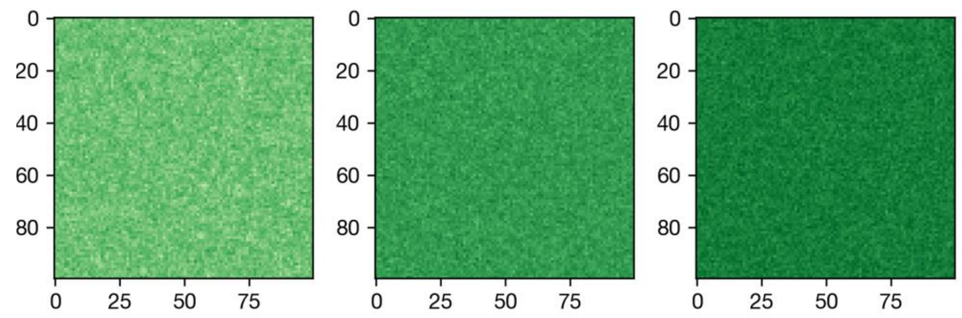


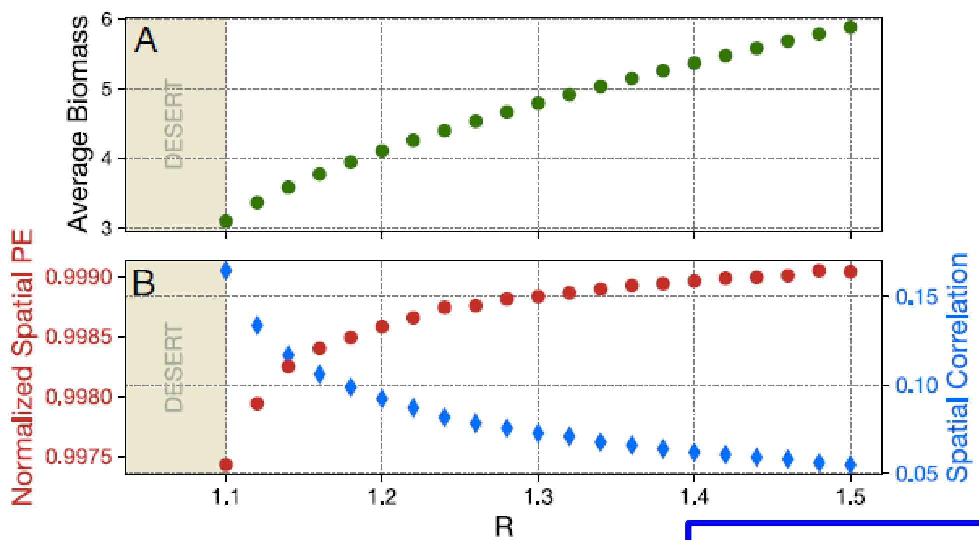
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_i (u_i - \bar{u})^2}$$

To gain insight: simulations of vegetation models

B) Local Positive Feedback model
(two partial differential equations)





$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

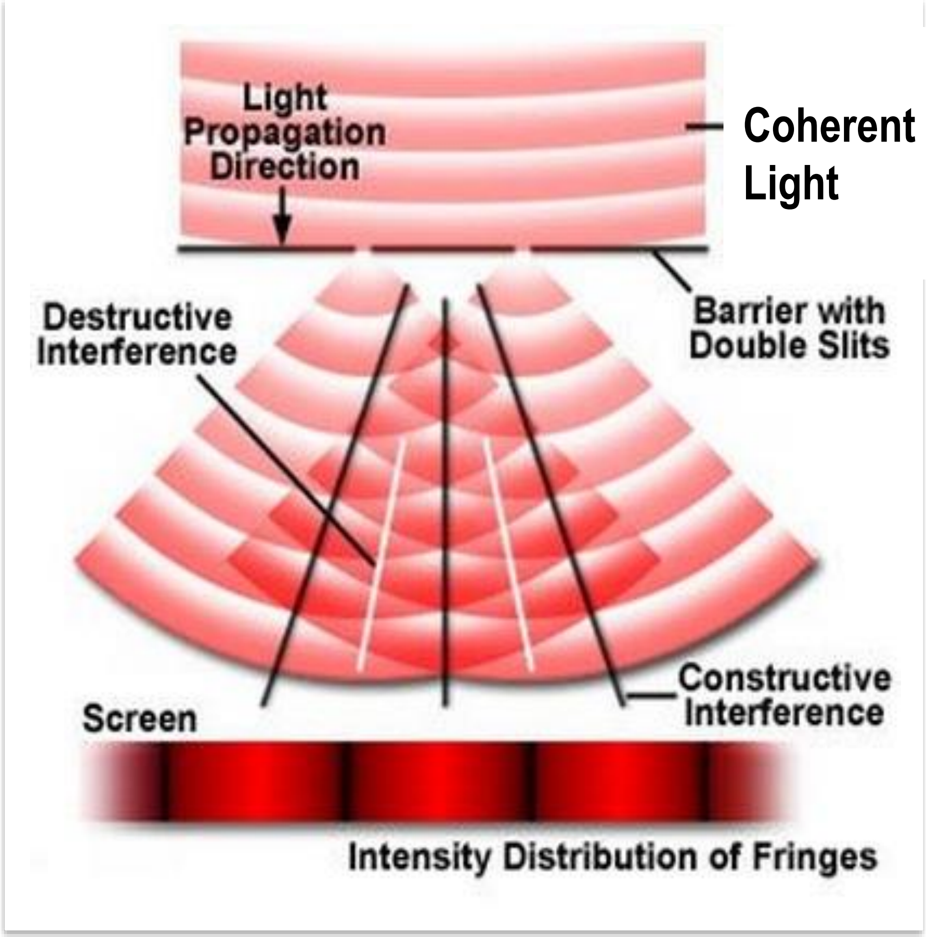
$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_i (u_i - \bar{u})^2}$$

cristina.masoller@upc.edu @cristinamasoll1

Diode laser experiments

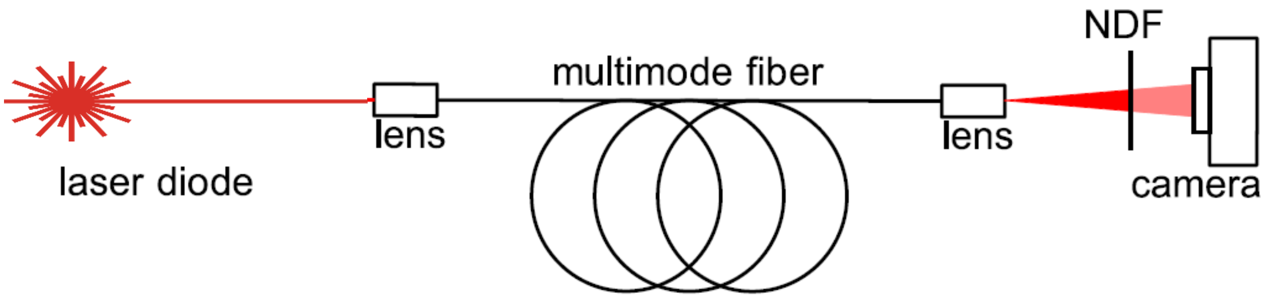
Transition from low-coherence emission (stochastic quantum spontaneous emission) to coherent emission (laser turn-on stimulated emission).

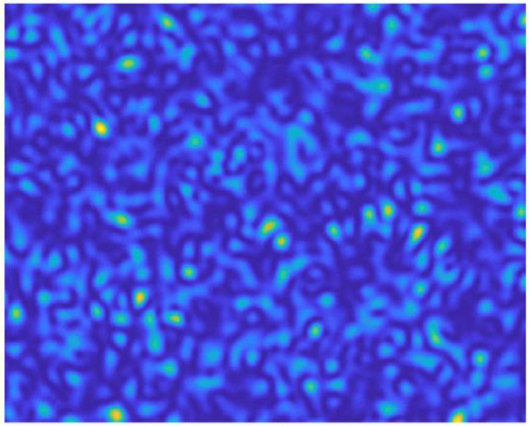
Quick review on the interference of coherent waves



cristina.masoller@upc.edu @cristinamasoll1

Speckle pattern: generated by random interference / scattering of coherent waves





Many applications. Two main types

- Extract information of the light (wavemeters)
- Extract information of the medium that generates the speckle (speckle-based spectroscopy)

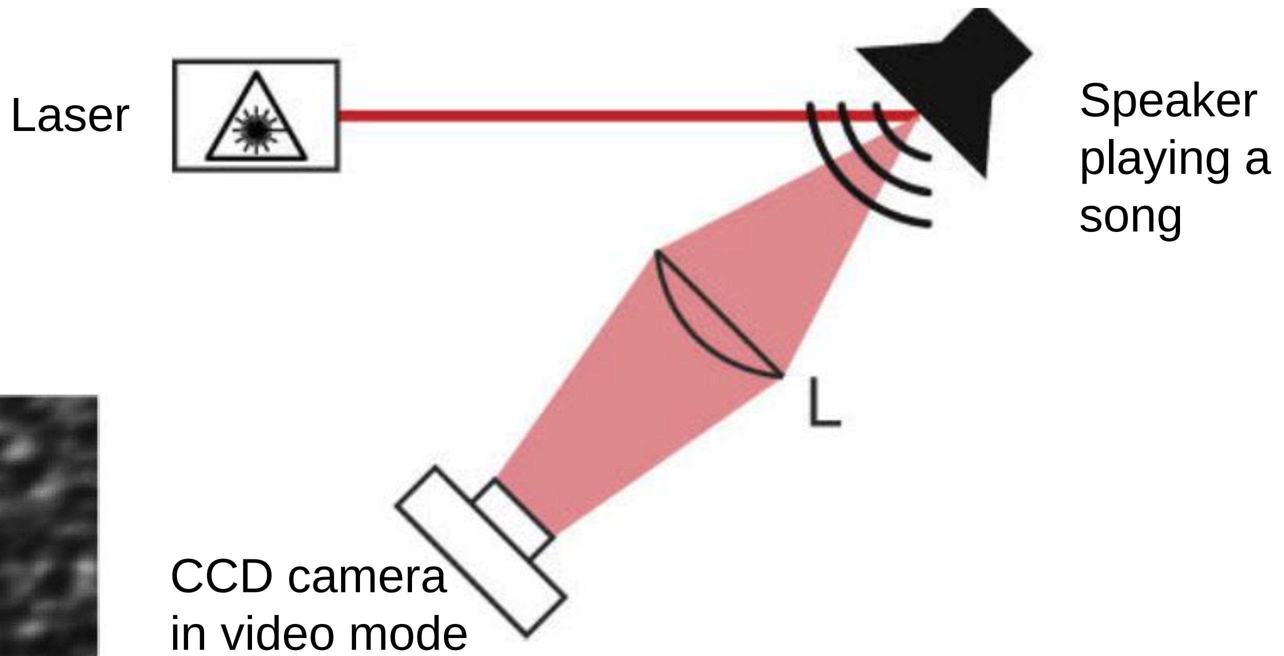
But

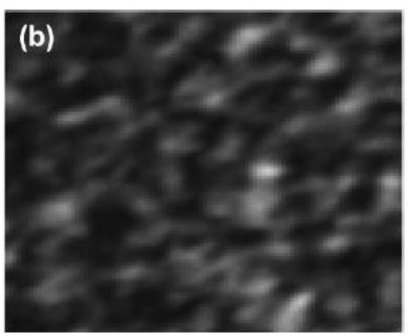
Speckle is a drawback in laser-based illumination and imaging application.

cristina.masoller@upc.edu @cristinamasoll1

Example of application of speckle analysis in our lab

Recovery of audio signals from silent videos of speckle patterns



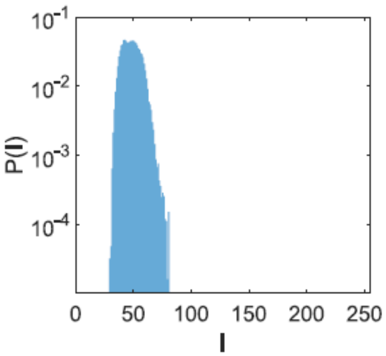


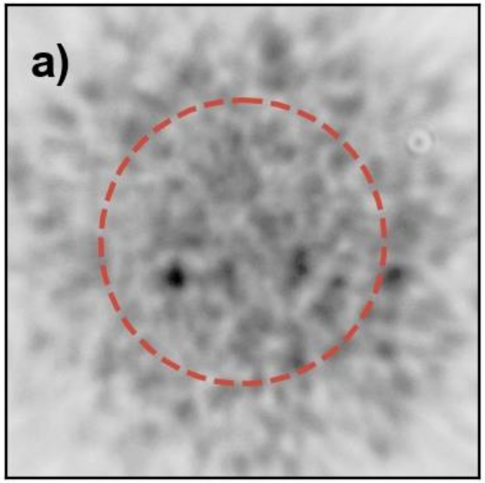
C. Barcellona, D. Halpaap, P. Amil, A. Buscarino, L. Fortuna, J. Tiana, C. Masoller, "*Remote recovery of audio signals from videos of optical speckle patterns: a comparative study of signal recovery algorithms*", Opt. Exp. 28, 8716 (2020).

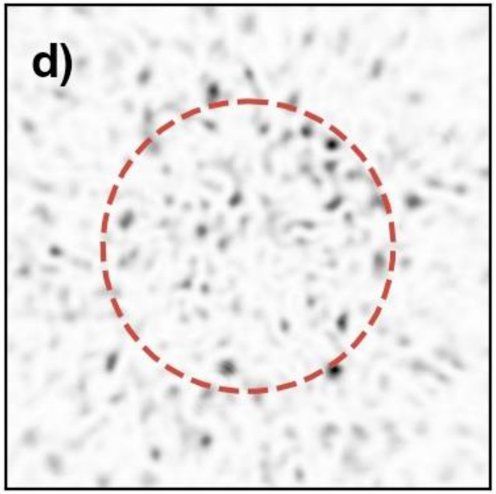
cristina.masoller@upc.edu @cristinamasoll1

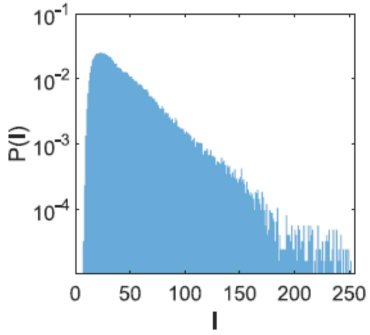
Analysis of Speckle Patterns using Permutation Entropy

Below threshold **Above threshold**









Quantification of speckle contrast: $SC = \sigma / \langle I \rangle$

cristina.masoller@upc.edu @cristinamasoll1

Results

Pattern:
x x
x x

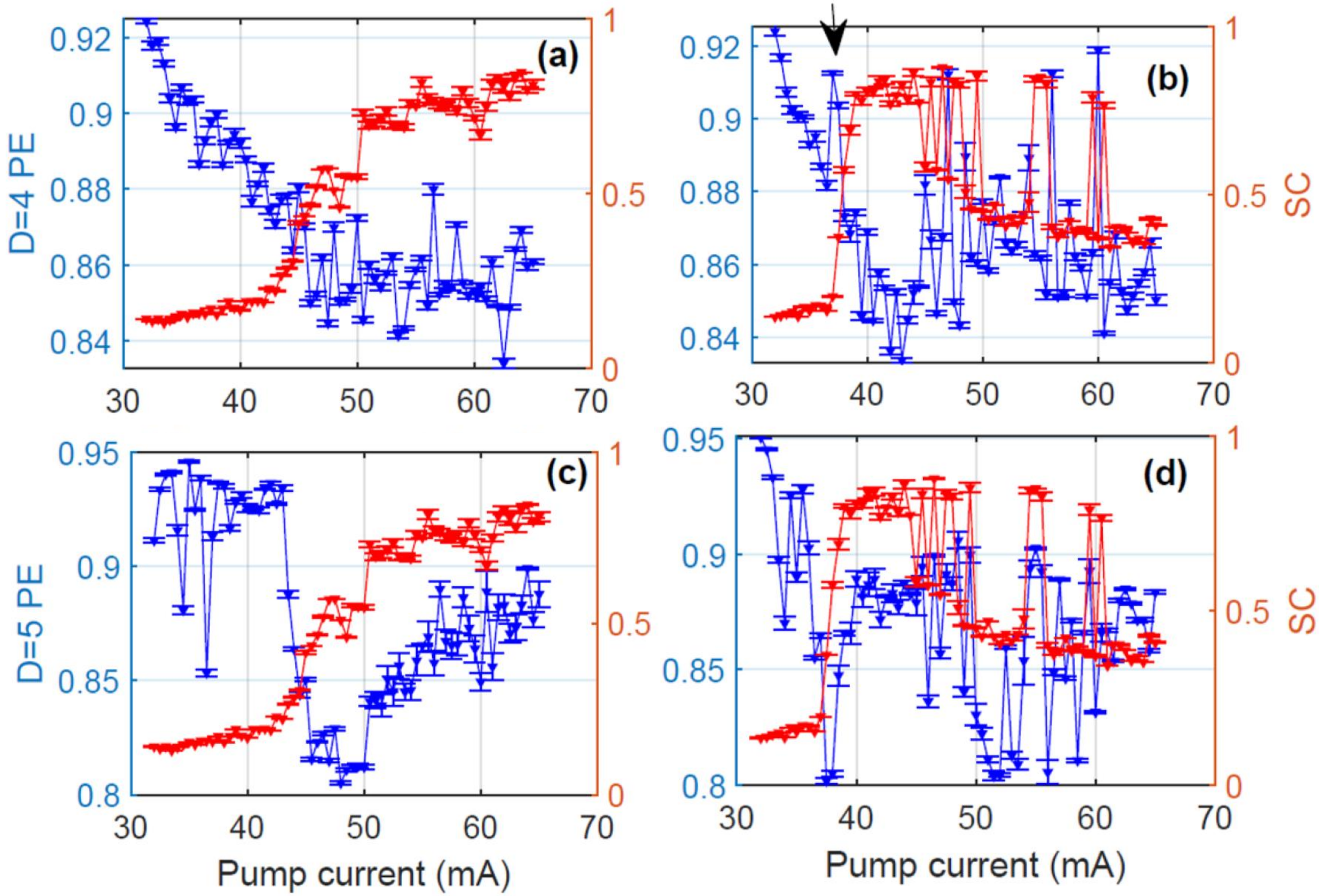
SC = σ/⟨I⟩

$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

Pattern: x
x x x
x

Laser without feedback

Laser with optical feedback



G. Tirabassi et al., "Permutation entropy-based characterization of speckle patterns generated by semiconductor laser light", APL Photonics 8, 126112 (2023).

cristina.masoller@upc.edu @cristinamasoll1

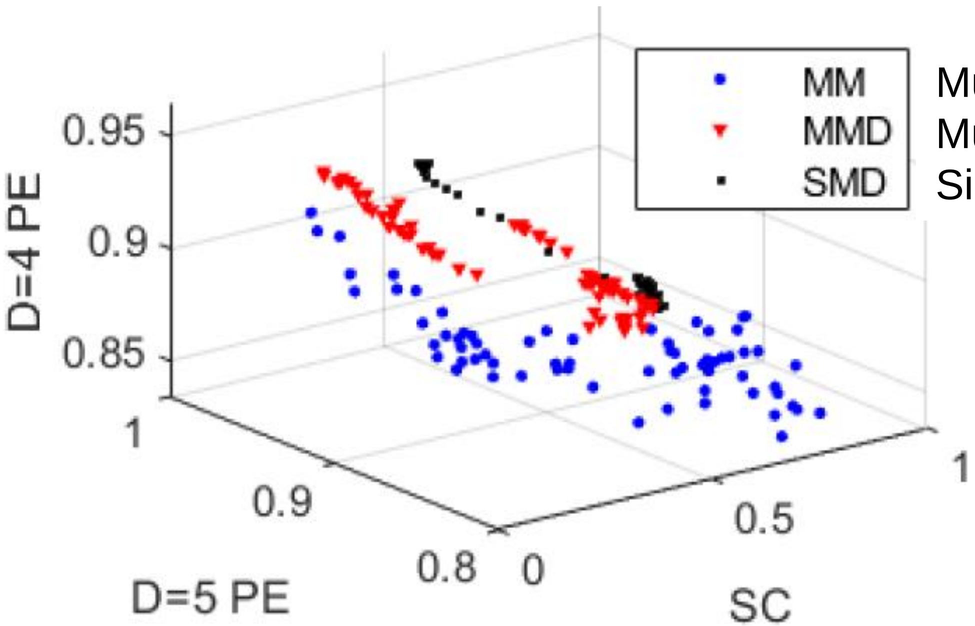
Three features allow to differentiate the speckle patterns according to the type of medium that generated the speckles

$SC = \sigma/\langle I \rangle$

$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

Pattern: x x
x x

Pattern: x
x x x
x



Multimode fiber
Multimode fiber and diffuser
Single mode fiber and diffuser

Accuracy of a random forest classifier

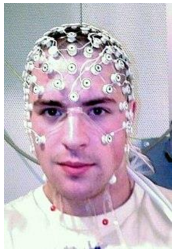
Solitary laser: 99.4 % ± 0.4 %

Laser with optical feedback: 97.1 % ± 1.3 %

G. Tirabassi et al., "Permutation entropy-based characterization of speckle patterns generated by semiconductor laser light", APL Photonics 8, 126112 (2023).

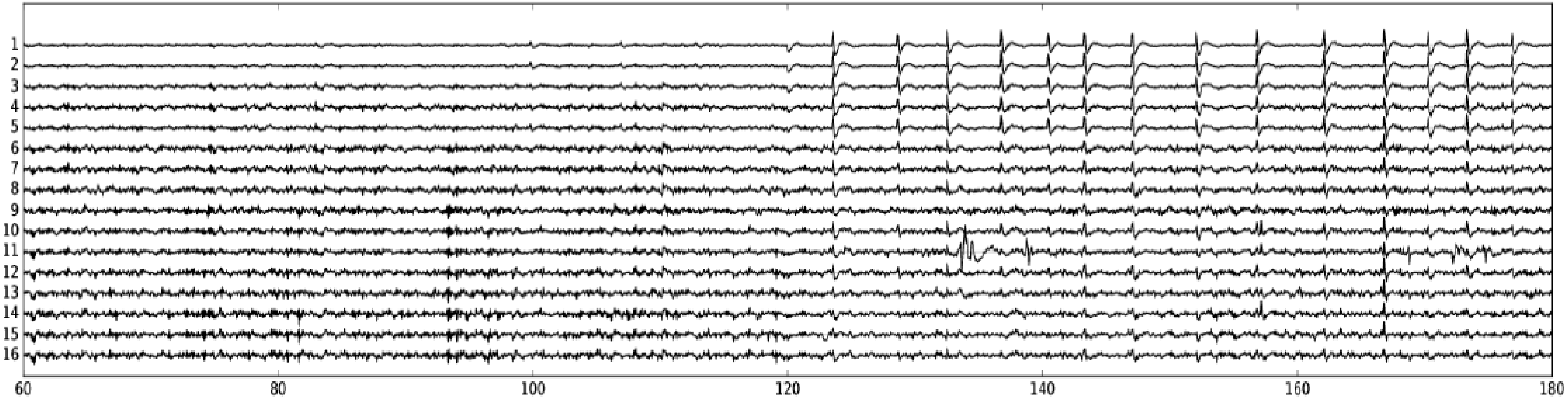
cristina.masoller@upc.edu @cristinamasoll1

Permutation Entropy analysis of EEG signals recorded from healthy subjects.



Eyes closed

Eyes open



Time (seconds)

TABLE I. Description of the datasets used.

	DTS1	DTS2
Sampling rate (Hz)	256	160
Time task (seg)	120	60
Total points	30 720	9600
Number of electrodes	16	64
Number of subjects	71	109

DTS1: Britbrain (Zaragoza)
DTS2: Physionet

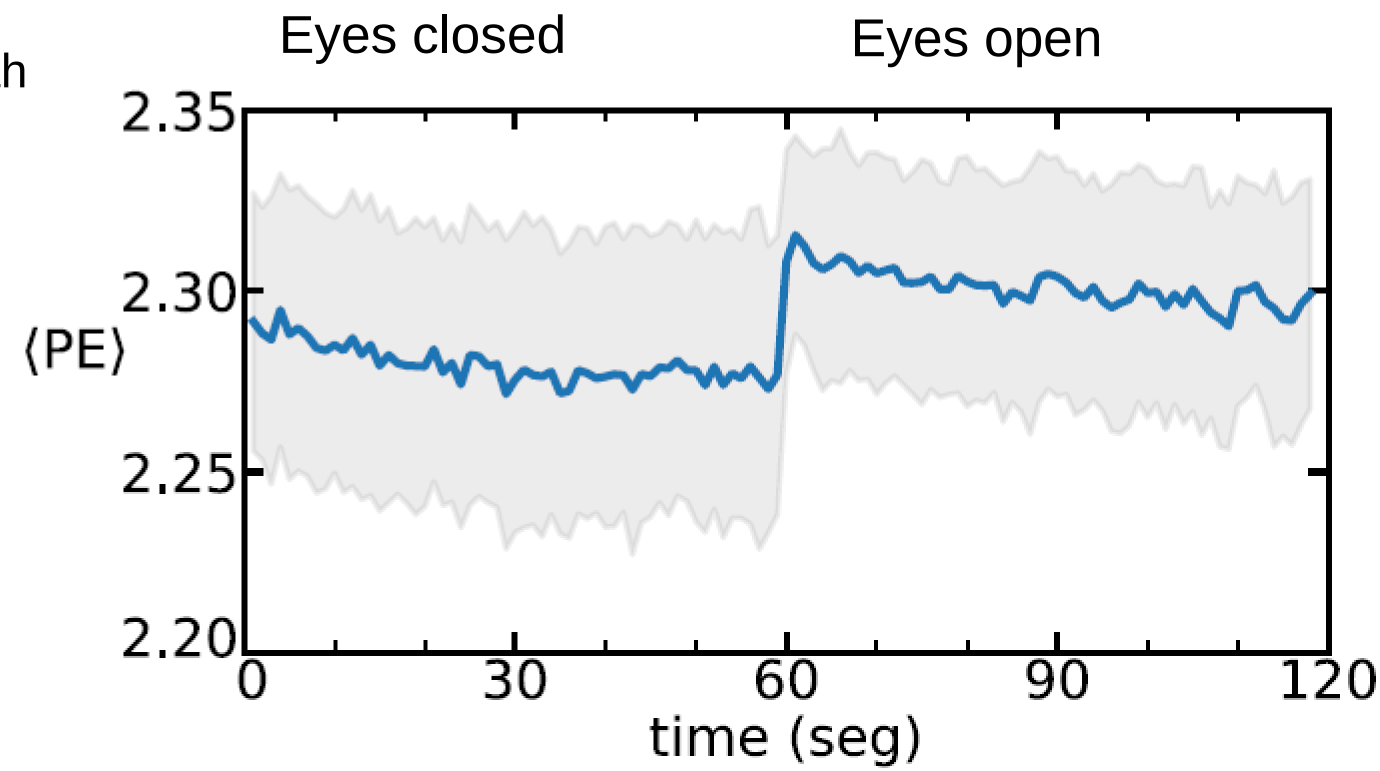
cristina.masoller@upc.edu @cristinamasoll1

The Permutation Entropy increases in the eyes open state

$$\langle \mathrm{PE} \rangle = \frac{1}{N[\text{electrodes}]} \sum_i \mathrm{PE}^i$$

PE was calculated with patterns of length 4 (# of possible patterns 24) in time windows containing >4000 patterns

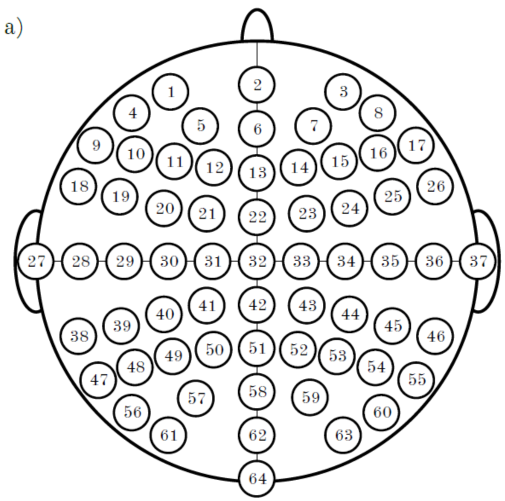
Gray region: σ of ⟨PE⟩ values across subjects

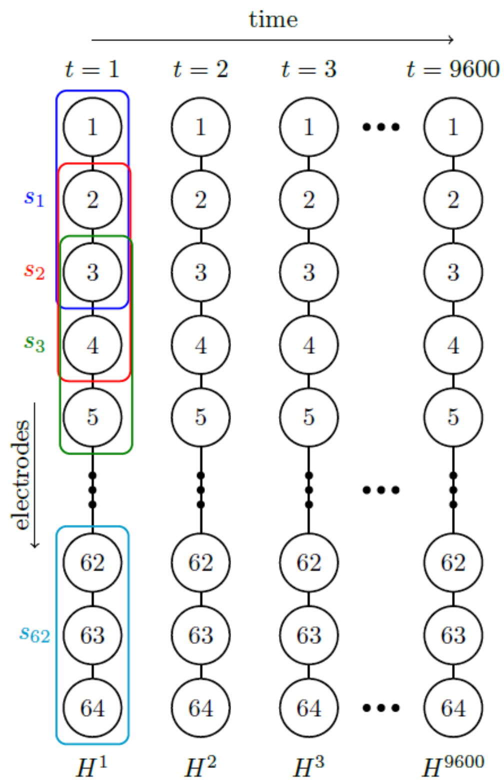


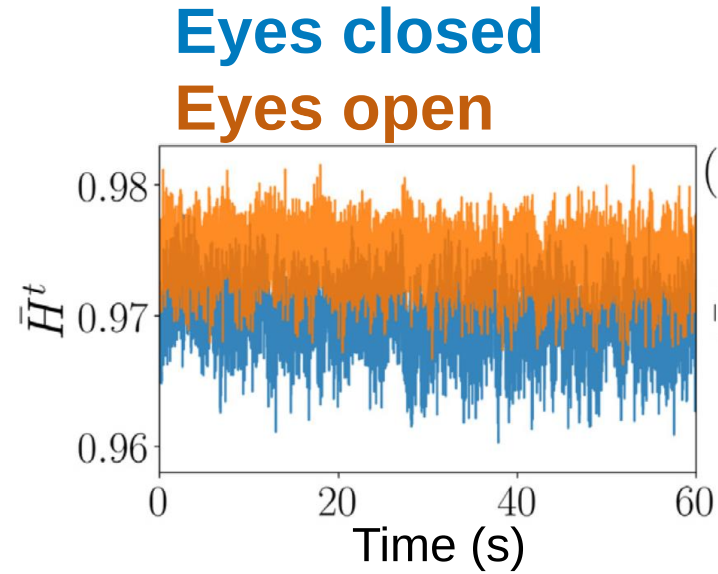
C. Quintero-Quiroz et al., "Differentiating resting brain states using ordinal symbolic analysis", Chaos 28, 106307 (2018).

cristina.masoller@upc.edu @cristinamasoll1

Spatial approach to compute the Permutation Entropy







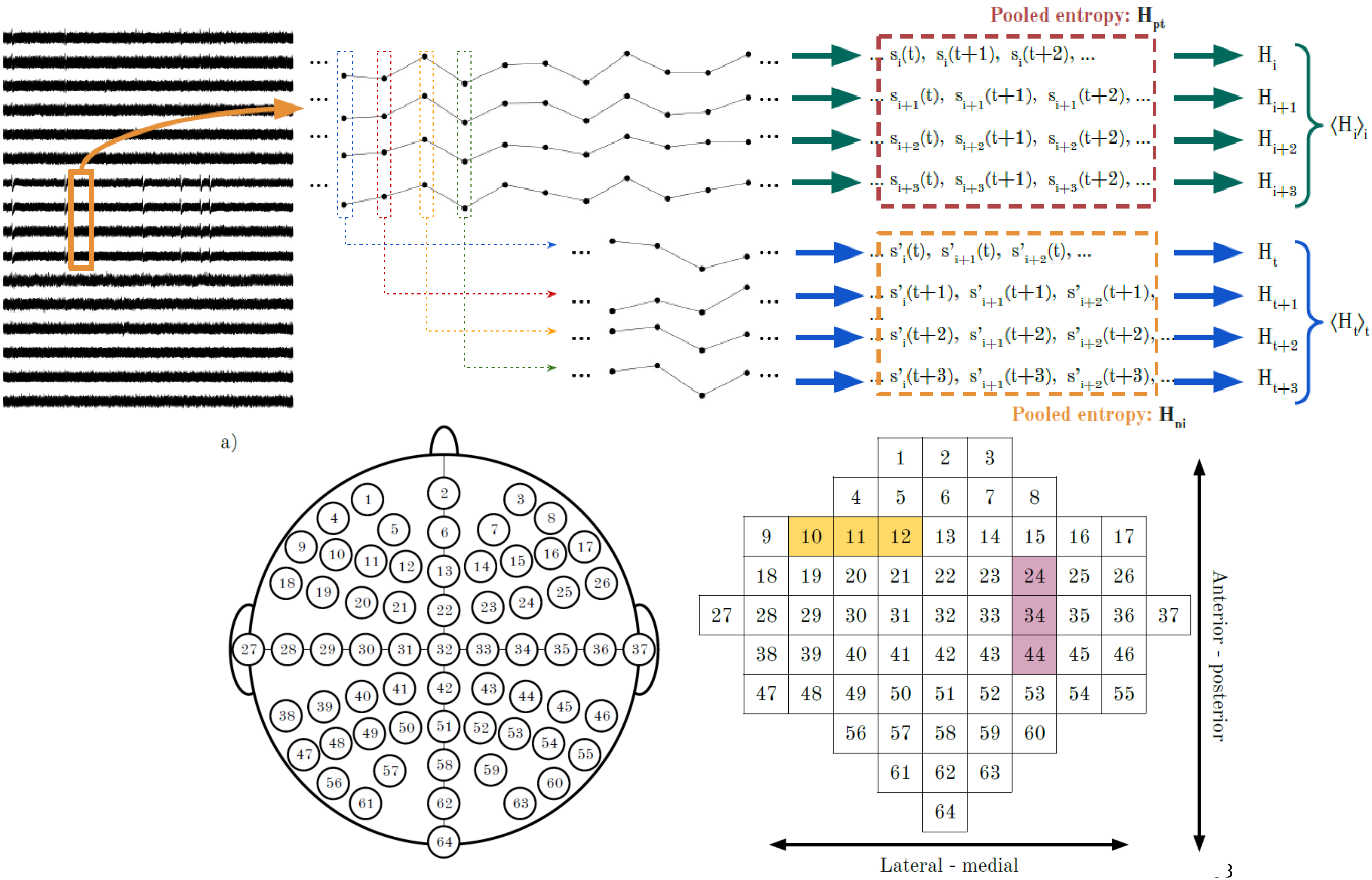
(H_t is averaged over subjects)

At each time: data values of 64 channels ⇒ 62 ordinal patterns to calculate 6 probabilities.

B. R. R. Boaretto et al, "Spatial permutation entropy distinguishes resting brain states", Chaos, Solitons & Fractals 171, 113453 (2023).

cristina.masoller@upc.edu @cristinamasoll1

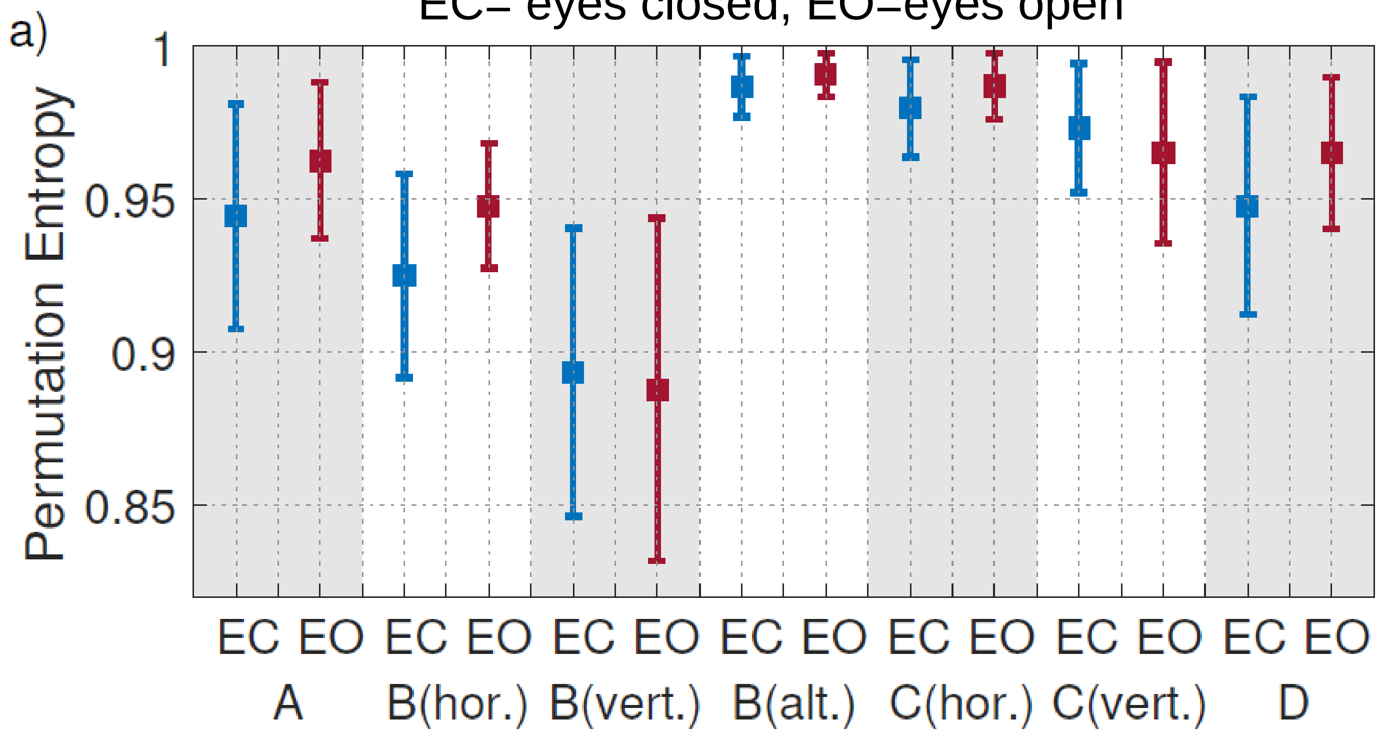
Four approaches to calculate the permutation entropy



cristina.masoller@upc.edu @cristinamasoll1

Results

EC= eyes closed, EO=eyes open



error bars:
σ over 107
subjects

A: temporal coding
B: spatial coding (horizontal, vertical, and alt. symbols)
C: spatial pooling (horizontal and vertical symbols)
D: temporal pooling

J. Gancio, C. Masoller, G. Tirabassi, "Permutation entropy analysis of EEG signals for distinguishing eyes-open and eyes-closed brain states: Comparison of different approaches", Chaos 34, 043130 (2024).

cristina.masoller@upc.edu @cristinamasoll1

Random forest classification of eyes open-eyes closed states

		Accuracy	F1 score	Precision	Recall	Specificity
Horizontal	$\langle H_t^s \rangle_t$	61 ± 7	59 ± 10	63 ± 10	57 ± 16	65 ± 16
	$\sigma\left(H_t^s\right)$	66 ± 7	65 ± 9	66 ± 9	$\mathbf{67 \pm 15}$	65 ± 15
	H_{pi}^s	58 ± 8	54 ± 12	61 ± 11	50 ± 16	66 ± 15
Vertical	$\langle H_t^s \rangle_t$	54 ± 9	55 ± 12	54 ± 10	59 ± 17	50 ± 15
	$\sigma\left(H_t^s\right)$	56 ± 9	59 ± 10	56 ± 9	64 ± 15	48 ± 16
	H_{pi}^s	55 ± 9	56 ± 11	55 ± 10	59 ± 17	51 ± 16
Temporal	$\langle H_i^s \rangle_i$	63 ± 8	56 ± 13	70 ± 15	49 ± 16	77 ± 15
	$\sigma\left(H_i^s\right)$	$\mathbf{69 \pm 8}$	$\mathbf{66 \pm 10}$	$\mathbf{73 \pm 12}$	62 ± 14	76 ± 13
	H_{pt}^s	64 ± 8	58 ± 13	72 ± 14	51 ± 16	$\mathbf{78 \pm 14}$

Using filtered data tends to improve the performance.

Performance is as good as that of other statistical measures.

J. Gancio, C. Masoller, G. Tirabassi, "Permutation entropy analysis of EEG signals for distinguishing eyes-open and eyes-closed brain states: Comparison of different approaches", Chaos 34, 043130 (2024).

cristina.masoller@upc.edu @cristinamasoll1

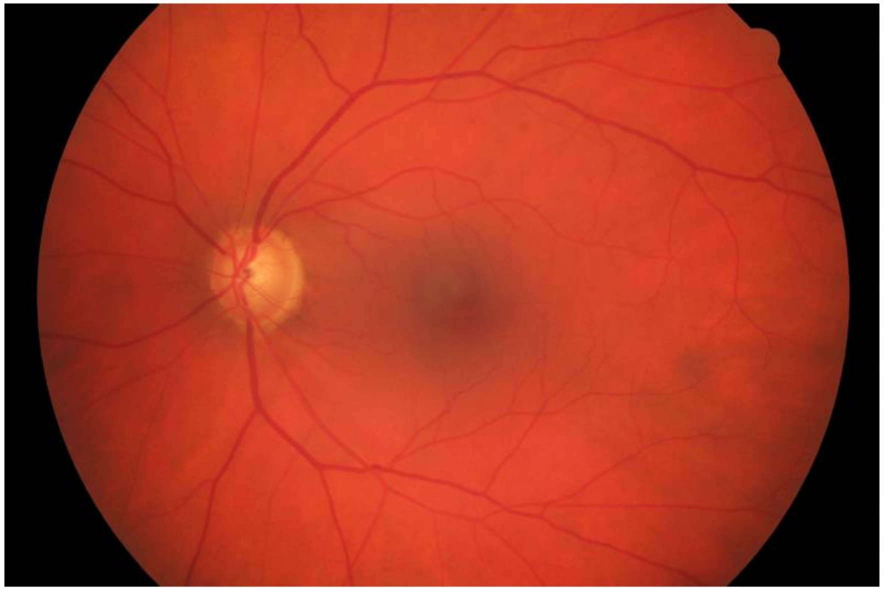
Outline

- Complex systems and time series analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: Lasers, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

cristina.masoller@upc.edu @cristinamasoll1

Analysis of retina fundus images

- For the diagnosis of eye diseases & follow up of treatments.
- Biometric identity identification.
- Opportunity to detect other diseases (alterations in retina network may reflect alterations in other arterial systems).



BE-OPTICAL

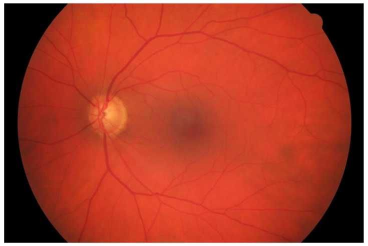
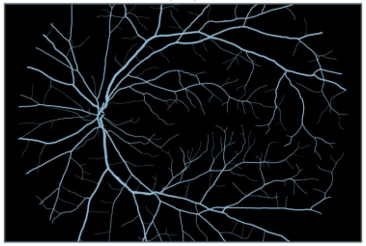
Advanced Biomedical Optical Imaging and Data Analysis

H2020-675512

cristina.masoller@upc.edu @cristinamasoll1

Data and image analysis steps

- 45 high resolution images (3504 × 2336 pixels)
  15 healthy subjects
  15 glaucoma
  15 diabetic retinopathy
- For every subject we had:
  - fundus photography
  - manual segmentation done by an expert ophthalmologist.

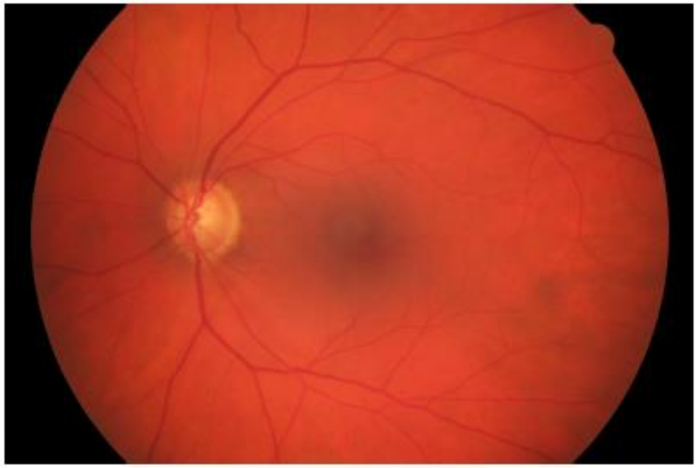

Steps:

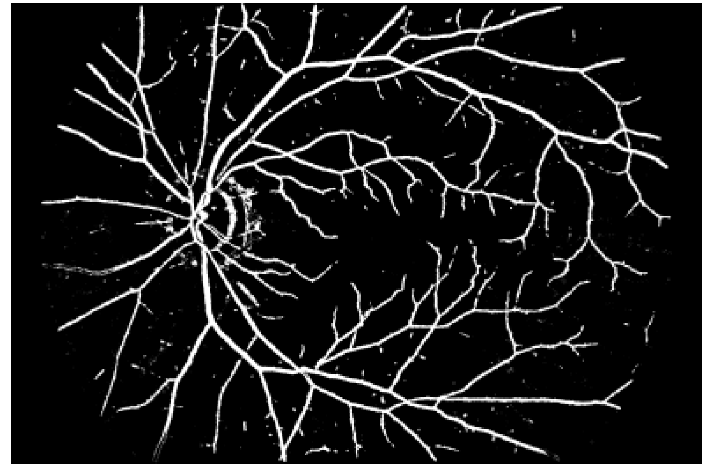
1. Pre-process and un-supervisely, segment the images.
2. Extract network.
3. Compare networks obtained from different images.
4. Classify the images.

https://www5.cs.fau.de/research/data/fundus-images/

cristina.masoller@upc.edu @cristinamasoll1

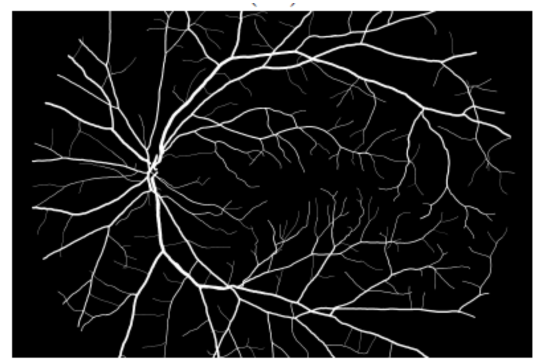
Step 1: Pre-process and segmentation





We adapted an *unsupervised* algorithm, originally developed for segmenting images of cultured neuronal networks.

Manual segmentation

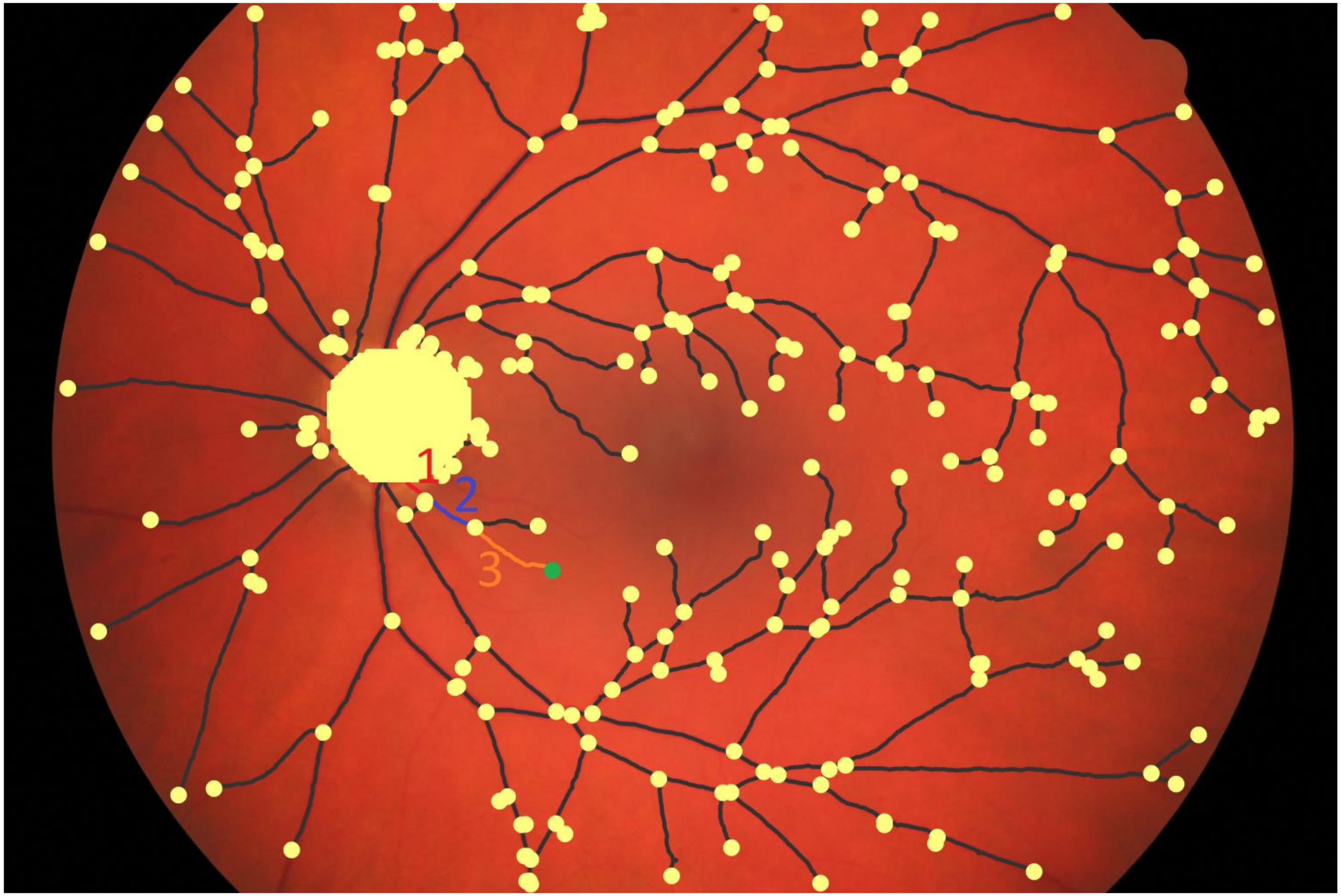


D. Santos-Sierra, I. Sendiña-Nadal, I. Leyva et al. Cytometry Part A. 87, 513 (2015).

P. Amil, F. Reyes-Manzano, L. Guzmán-Vargas, I. Sendiña-Nadal, C. Masoller, "*Network-based features for retinal fundus vessel structure analysis*", PLoS ONE 14, e0220132 (2019).

cristina.masoller@upc.edu **@cristinamasoll1**

Step 2: extract the network (identification of the optical nerve, nodes and links and assign weights to the links).

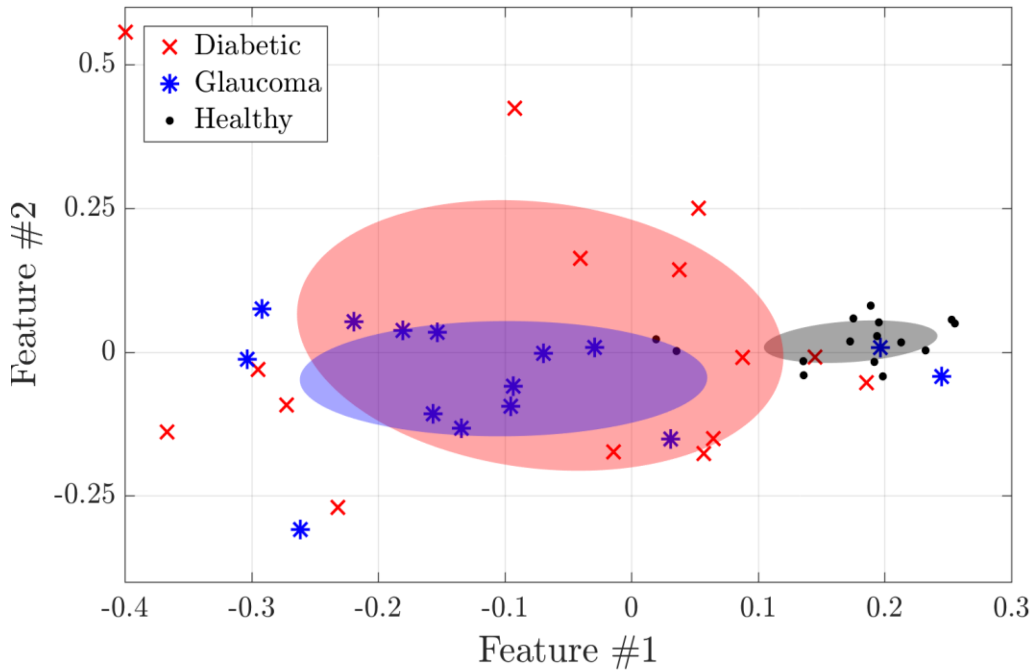


cristina.masoller@upc.edu @cristinamasoll1

Steps 3 and 4: Compare the networks extracted from different images and classify the images.

- $\{p_{i,j}\}$: distances between probability distributions that characterize the networks obtained from images i and j.
- We used nonlinear dimensionality reduction (*Isomap*) to reduce the set of 45x45 $\{p_{i,j}\}$ values to only two features.

Distance distribution to the central node in the *manual* segmentation

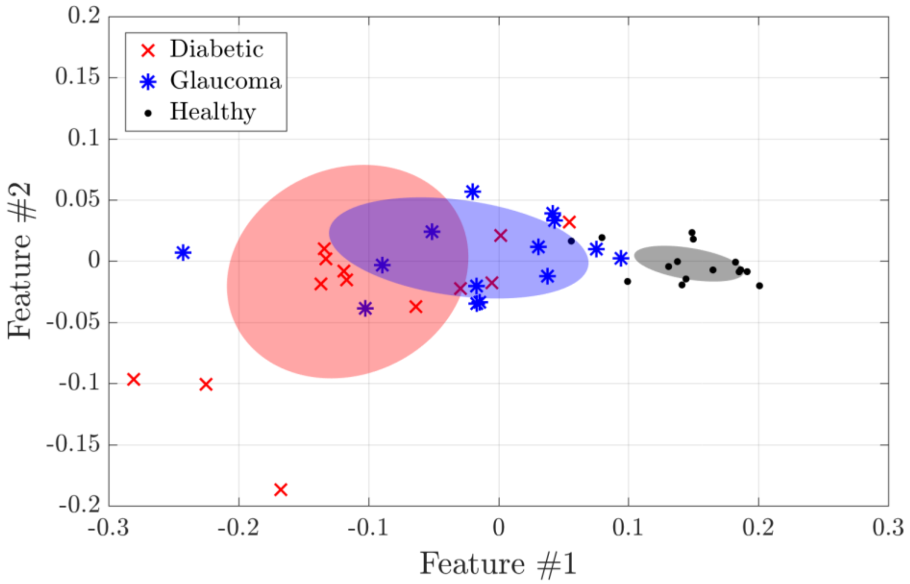


P. Amil et al, Network-based features for retinal fundus vessel structure analysis, PLoS ONE 14 e0220132 (2019).

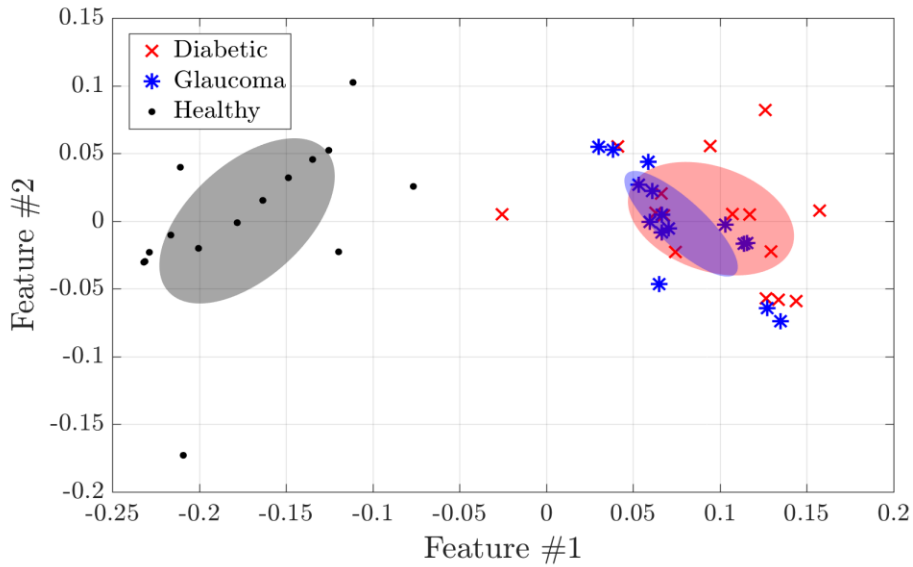
cristina.masoller@upc.edu @cristinamasoll1

Performance of network features in the *manual* segmentation

Distribution of weights along the shortest path to central node



Distribution of weighted degrees

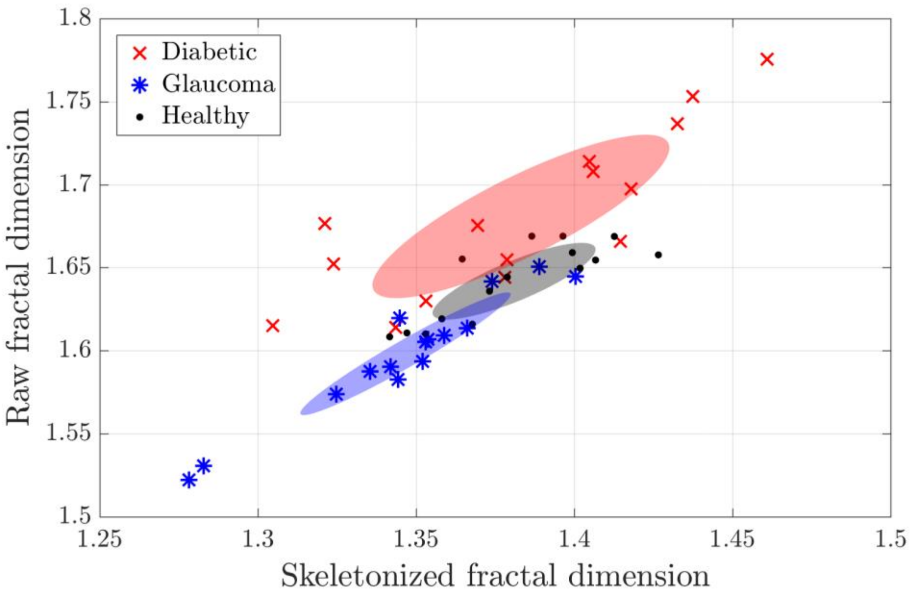


P. Amil et al, Network-based features for retinal fundus vessel structure analysis, PLoS ONE 14 e0220132 (2019).

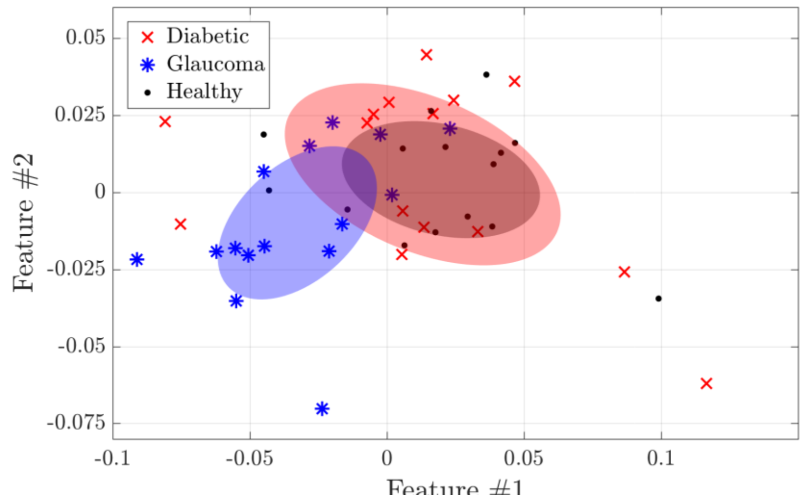
cristina.masoller@upc.edu @cristinamasoll1

In the automated segmentation

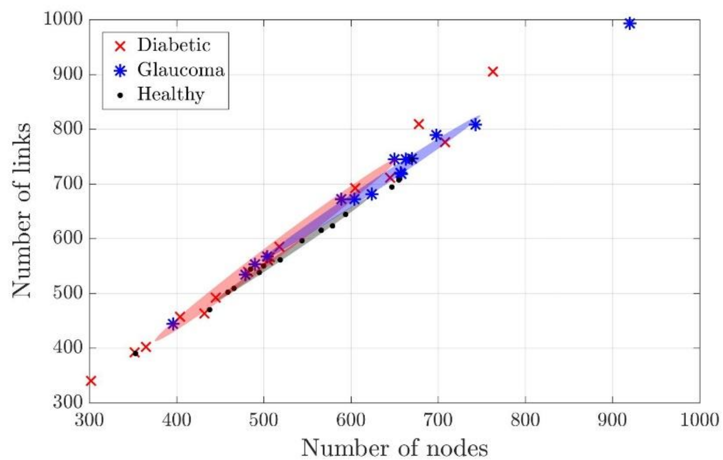
Fractal dimension $D = \lim_{\varepsilon \to 0} \frac{\log(N(\varepsilon))}{\log(1/\varepsilon)}$



Mean weight distribution along the shortest path to central node



Simple network features do not differentiate



P. Amil et al, Network-based features for retinal fundus vessel structure analysis, PLoS ONE 14 e0220132 (2019).

cristina.masoller@upc.edu @cristinamasoll1

Outline

- Complex systems and time series analysis
- Ordinal analysis: Lasers and neurons and climate data
- Hilbert analysis: Climate data
- Causal inference: Synthetic and climate data
- Regime transitions: Lasers, EEG and vegetation data
- Network analysis: Retina fundus images
- Take home messages

cristina.masoller@upc.edu @cristinamasoll1

- Data analysis methods allow us to uncover patterns and relationships in data, which characterize (and sometimes predict) the behavior of complex systems.

- Different methods provide *complementary* information.

- Even when the data does not meet the mathematical or algorithmic requirements, the results can give useful info.

- "Surrogate" tests are needed to determine if the numerical values are statistically significant.

- Data analysis is an interdisciplinary field -many applications.

Holger Kantz: "*Every data set bears its own difficulties: data analysis is never routine*"

cristina.masoller@upc.edu @cristinamasoll1

ICREA AGAUR GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

- J. Tiana-Alsina et. al, "*Comparing the dynamics of periodically forced lasers and neurons*", New J. of Phys. 21, 103039 (2019).
- M. Masoliver and C. Masoller, "*Neuronal coupling benefits the encoding of weak periodic signals in symbolic spike patterns*", Commun. Nonlinear Sci. Numer. Simulat. 88, 105023 (2020).
- D. A. Zappala et. al, "*Quantifying changes in spatial patterns of surface air temperature dynamics over several decades*", Earth Syst. Dynam. 9, 383–391 (2018).
- G. Tirabassi and C. Masoller, "*Entropy-based early detection of critical transitions in spatial vegetation fields*", PNAS 120, e2215667120 (2022).
- J. Gancio et. al, "Permutation entropy analysis of EEG signals for distinguishing eyes-open and eyes-closed brain states: Comparison of different approaches", Chaos 34, 043130 (2024).
- P. Amil et. al, "*Network-based features for retinal fundus vessel structure analysis*", PLoS ONE 14, e0220132 (2019).

Thank you for your attention!

cristina.masoller@upc.edu @cristinamasoll1