# Nonlinear data analysis tools for complex systems research

Cristina Masoller

Departament de Física, Campus Terrassa, Barcelona

Universitat Politecnica de Catalunya

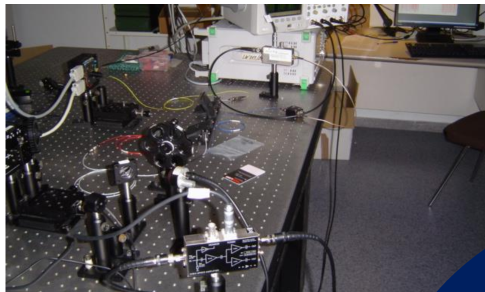
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

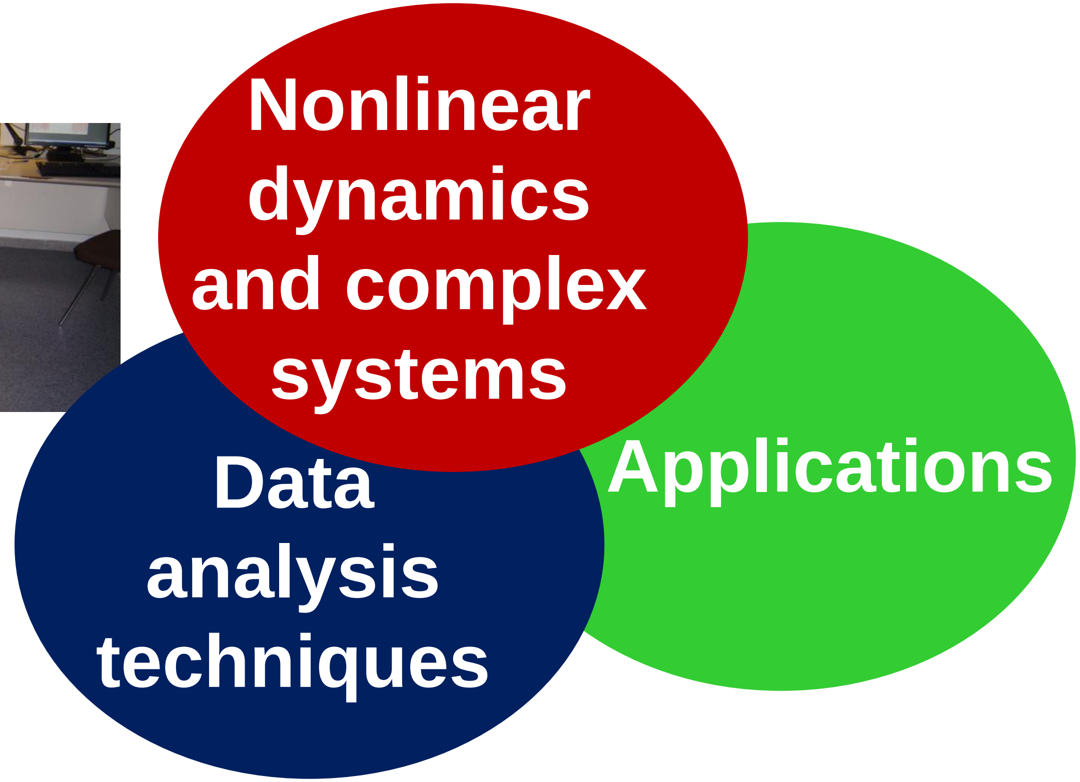
Campus d'Excel·lència Internacional

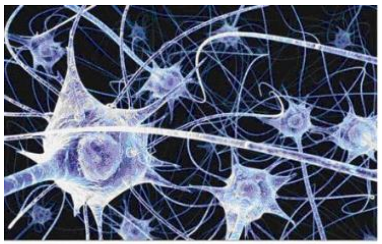
*XXV Congreso de Física Estadística (FisEs'25)*

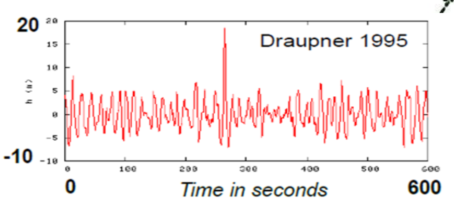
*Santiago de Compostela, June 19, 2025*

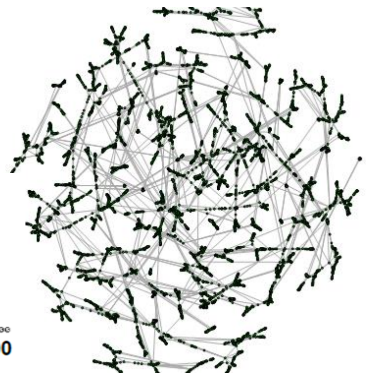
# Research lines

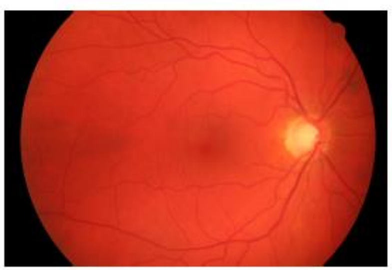


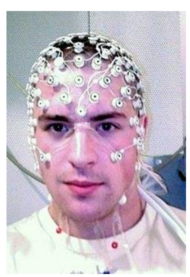


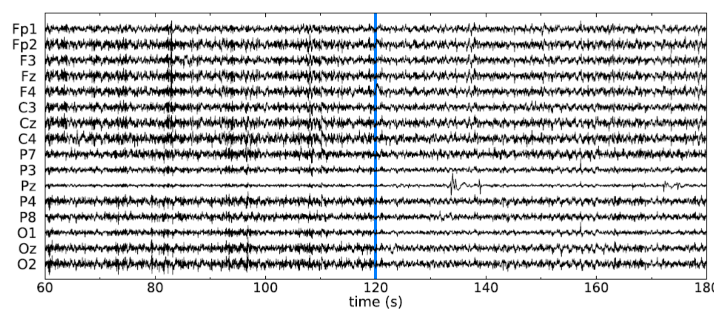



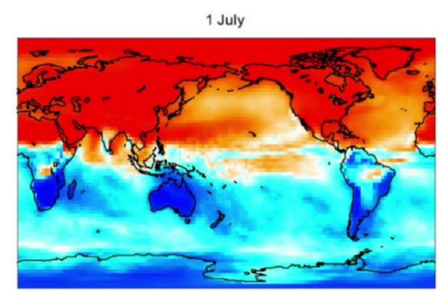






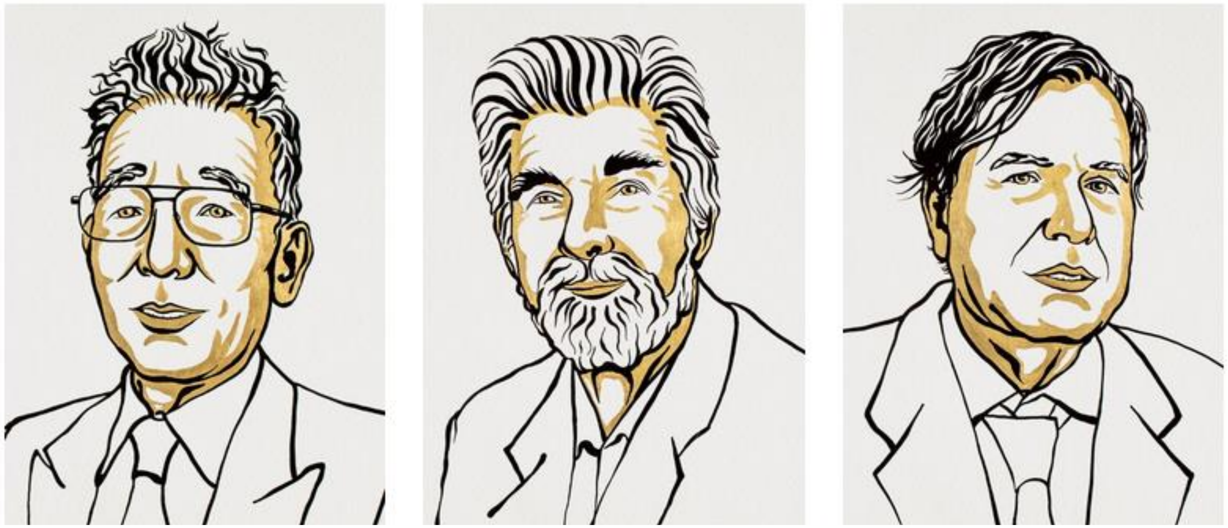





cristina.masoller@upc.edu @cristinamasoll1

The Nobel Prize in Physics 2021



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

½ Syukuro Manabe and Klaus Hasselmann ½ Giorgio Parisi

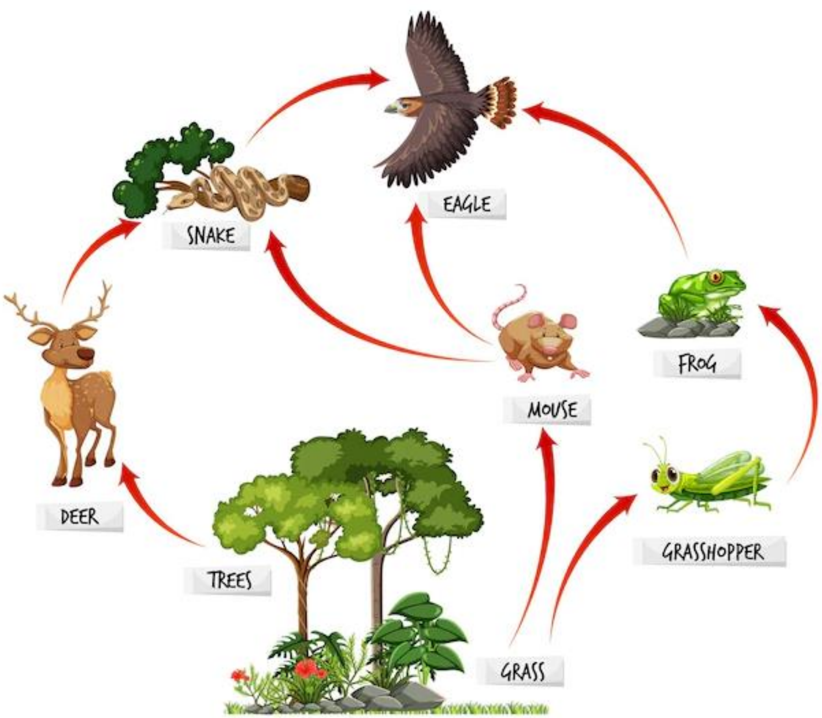
# What is a complex system?

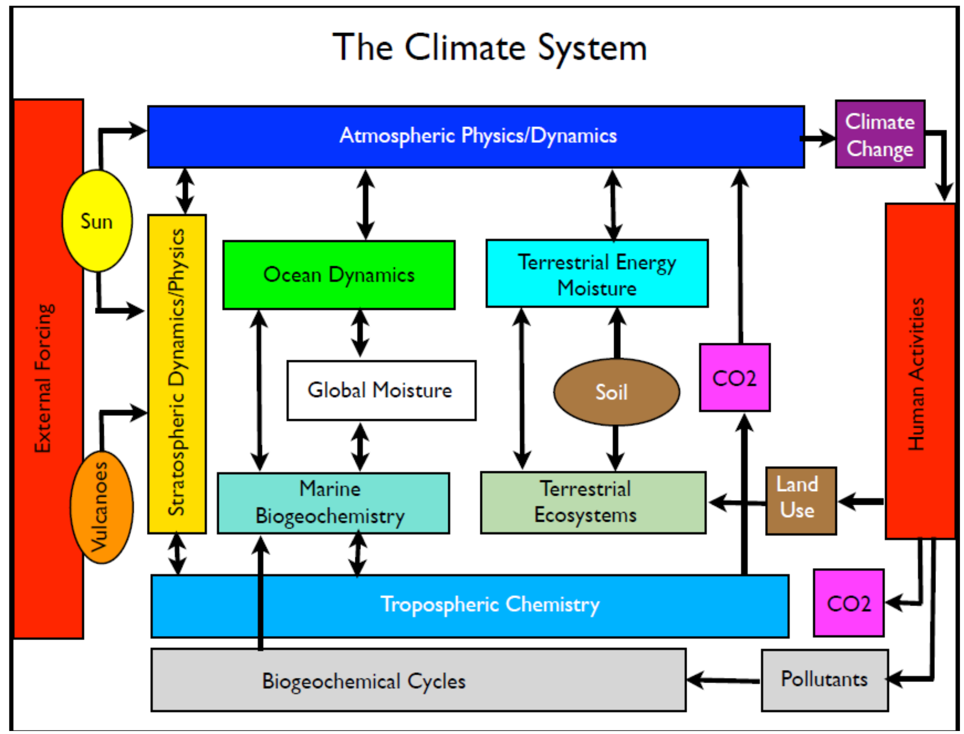


cristina.masoller@upc.edu @cristinamasoll1

# Complex systems are

- High-dimensional (huge number of variables or elements)
- Nonlinear (the elements and/or the interactions are nonlinear)
- Heterogeneous, multiscale, have memory and adapt.




*Courtesy of Henk Dijkstra (Universidad de Ultrech)*



cristina.masoller@upc.edu @cristinamasoll1

# Complex systems display:

Emergent phenomena -synchronization
Gradual or abrupt transitions
Extreme fluctuations

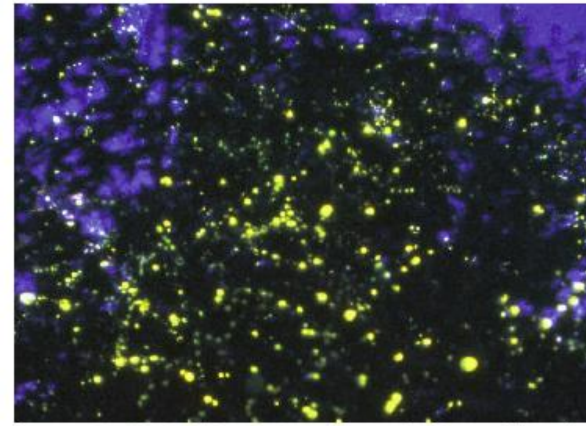
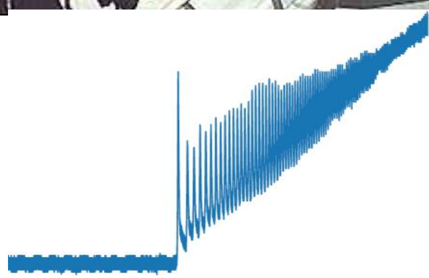
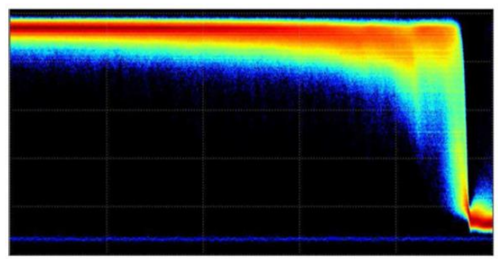


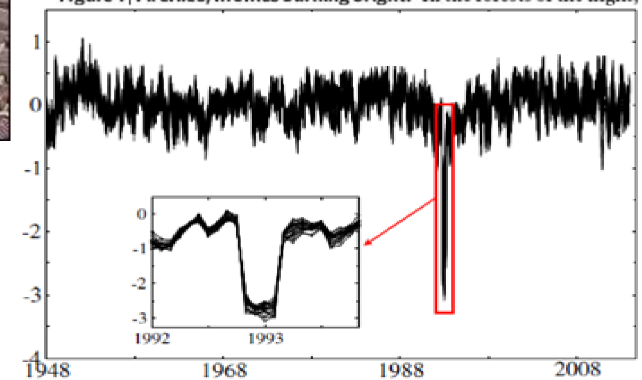
Figure 1| Fireflies, fireflies burning bright. In the forests of the night,

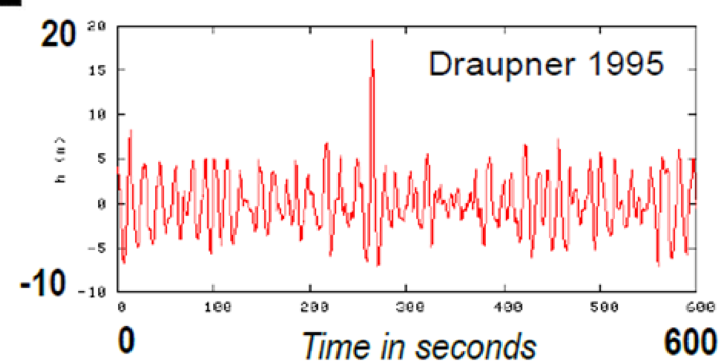


Time



Time




For the sake of clarity, what is NOT, in my opinion, a complex system:

Any linear system (no matter how big).
Low dimensional system.



cristina.masoller@upc.edu @cristinamasoll1

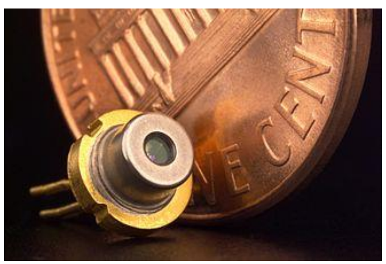
# Controlled experiments with complex systems are rare

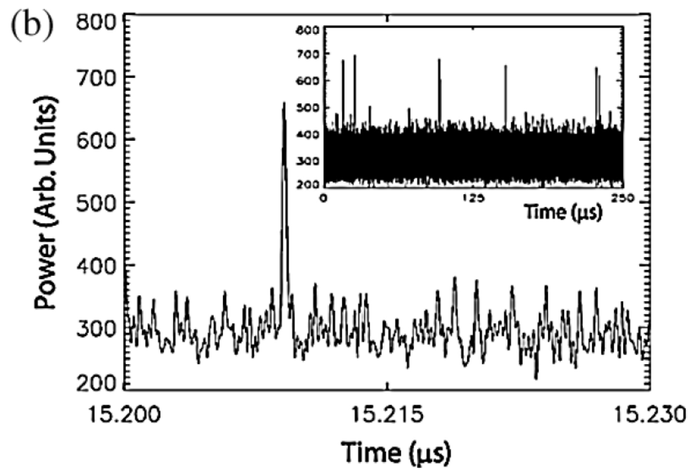
Time scales of photonic systems allow recording long data sets in short time.

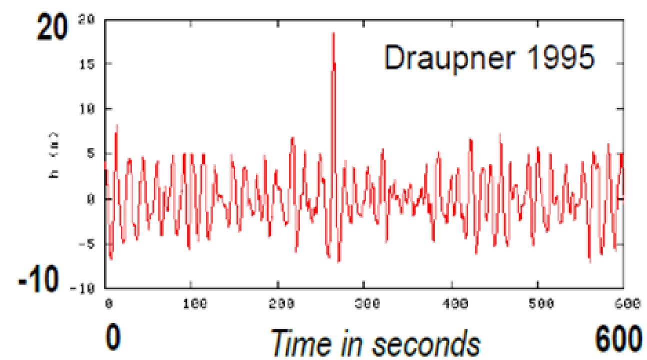
## The power of analogies

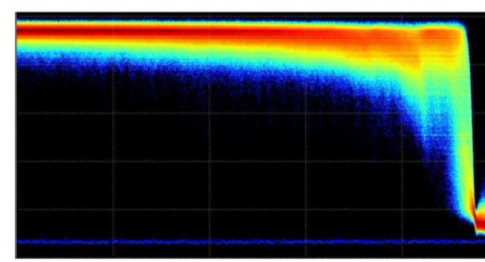
NATURE PHOTONICS | VOL 8 | JANUARY 2014 |

Black holes, gravitational lenses, turbulence, chaotic flow and rogue waves are just a few examples of complex physical phenomena that can be conveniently modelled using photonics.



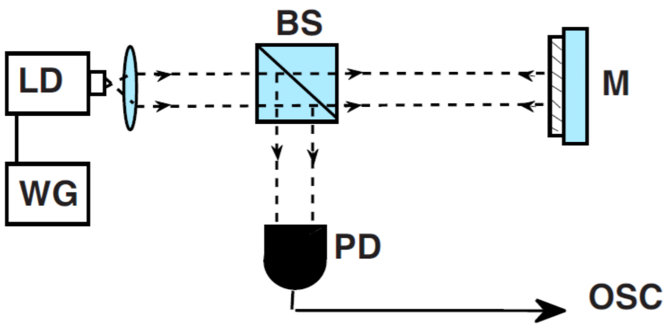
*Time (0.1 μs)*

*C. Bonatto et. al, Phys. Rev. Lett. 107, 053901 (2011)*

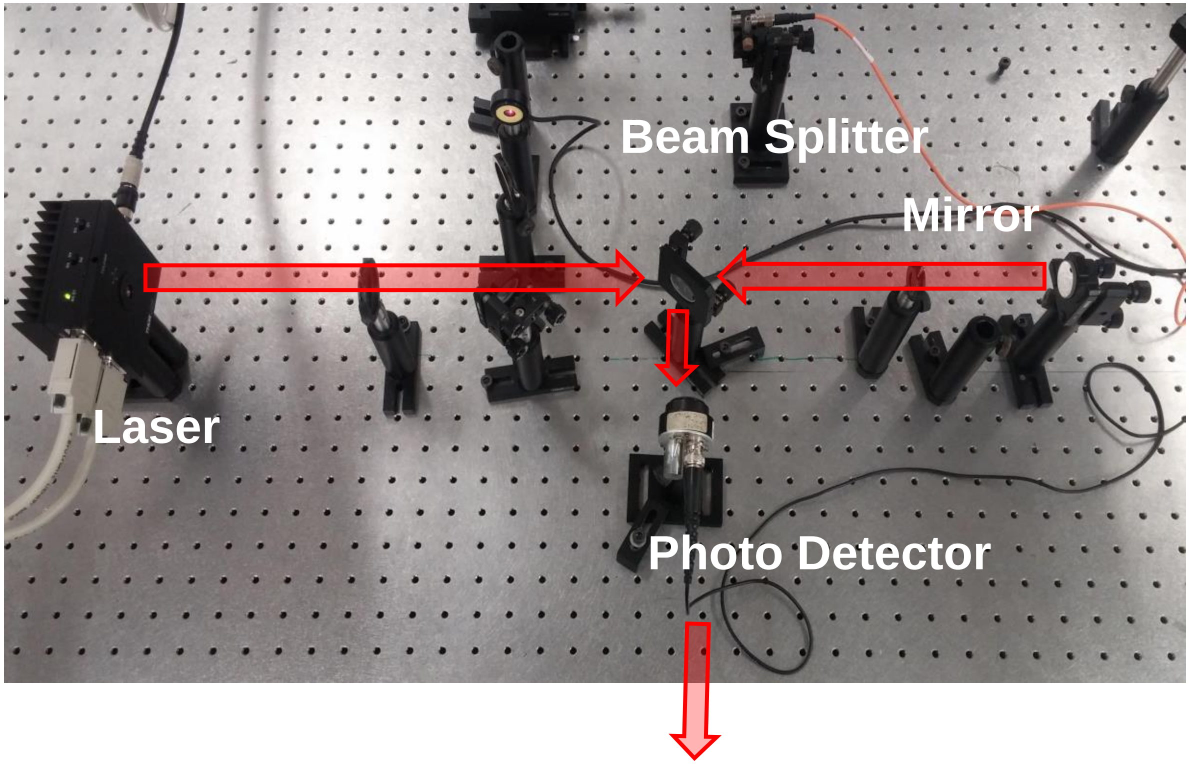


cristina.masoller@upc.edu @cristinamasoll1

# In our lab: experiments with diode lasers with optical feedback




WG: wave generator used to modulate the laser current.




Oscilloscope

**What do we see in the oscilloscope?**
**(with 1 GHz resolution)**

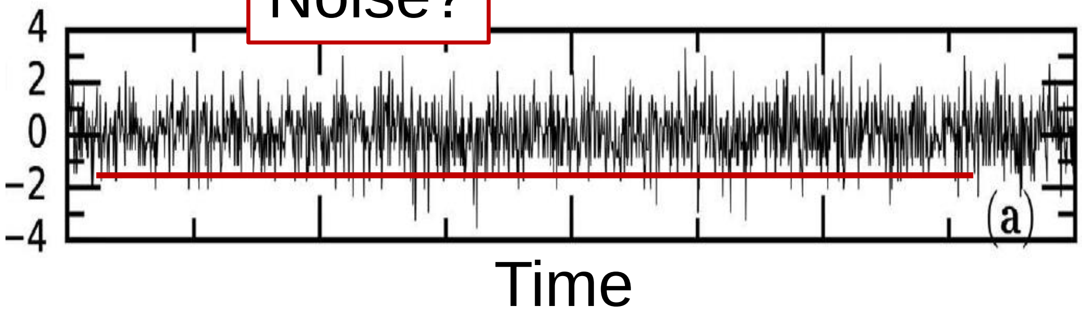


cristina.masoller@upc.edu @cristinamasoll1

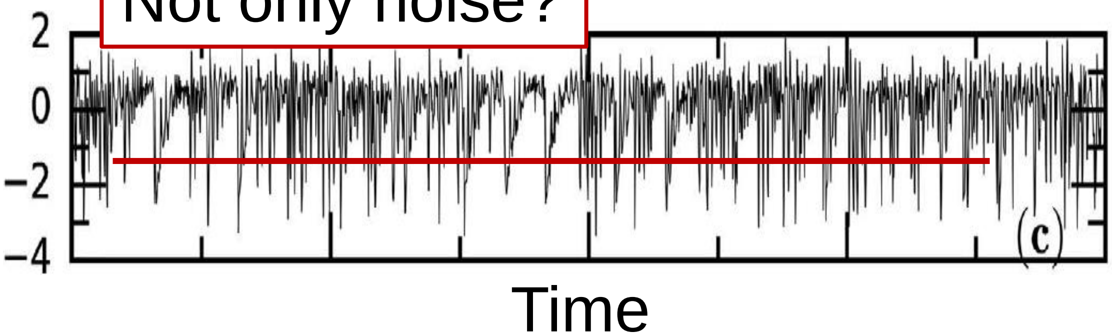
# By counting the number of "threshold crossings" we can distinguish dynamical regimes

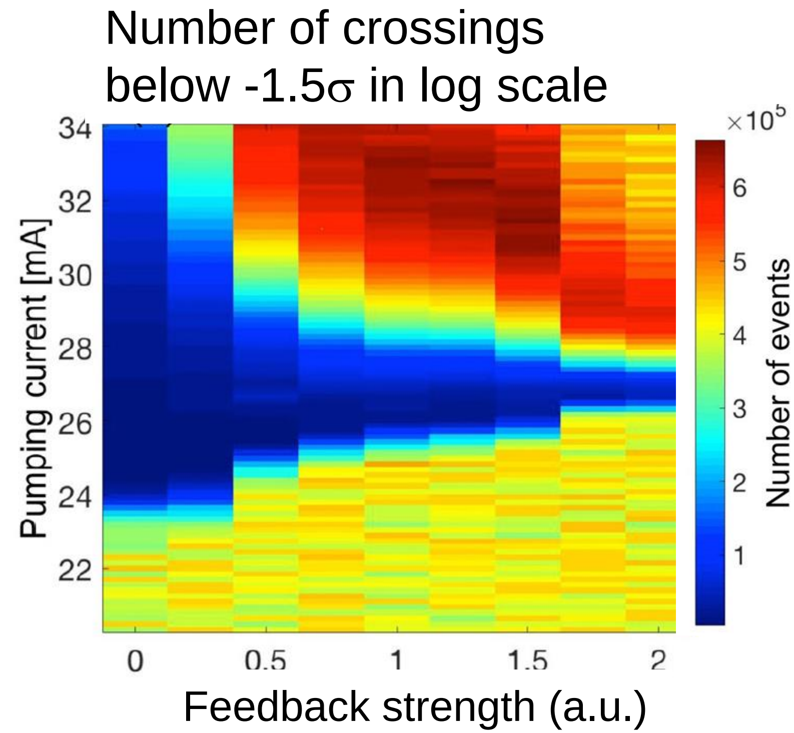
$$Z = \frac{X - \mu}{\sigma}$$

Noise?

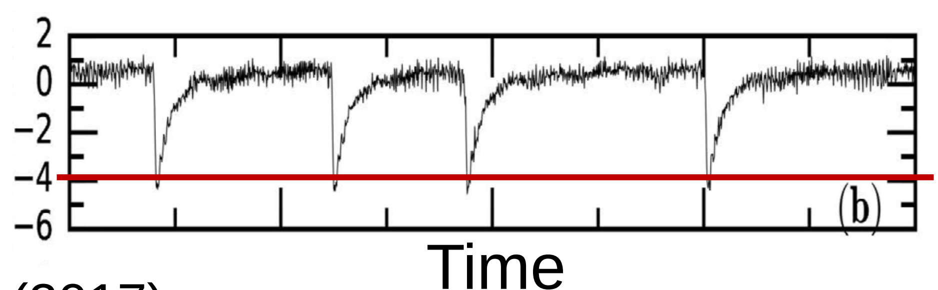



Not only noise?




Blue region? Optical Spikes

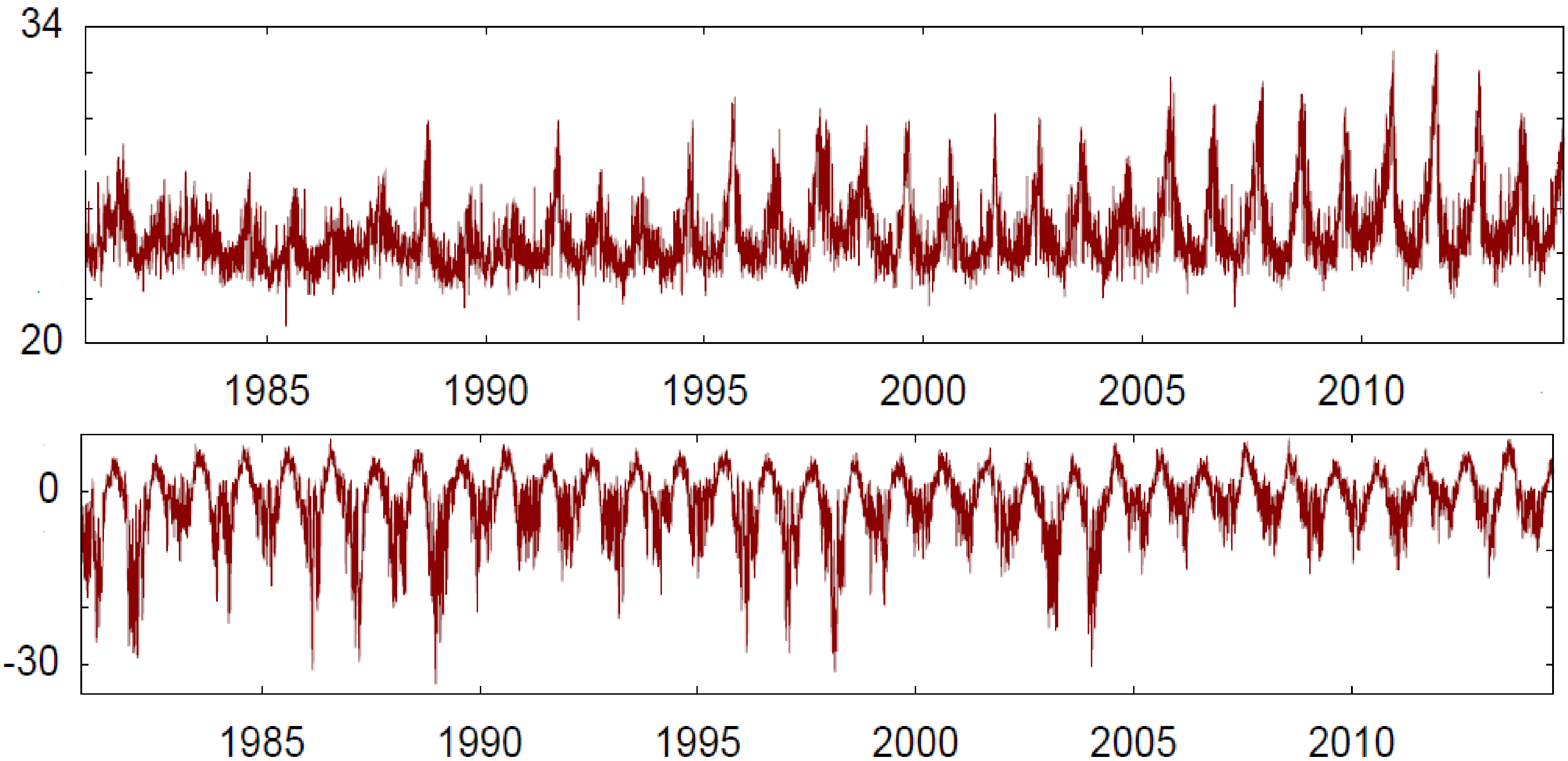



M. Panozzo et al, Chaos 27, 114315 (2017).



cristina.masoller@upc.edu @cristinamasoll1

# An example of a gradual change in behavior: Surface Air Temperature (SAT) in two geographical regions



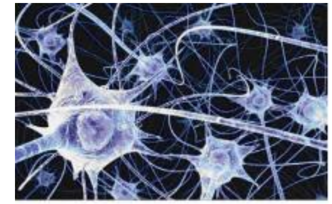

**How to quantify the differences?**

Dario A. Zappala, M. Barreiro and C. Masoller, Earth Syst. Dynam. 9, 383 (2018)

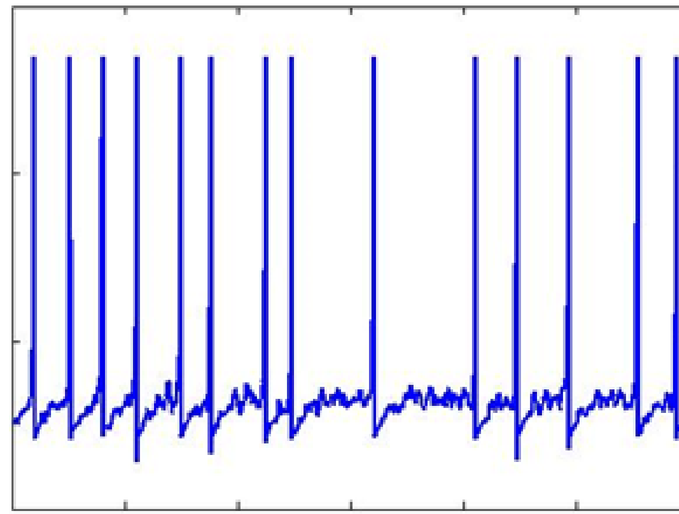


cristina.masoller@upc.edu @cristinamasoll1

# Data analysis methods can discover similarities in different systems




Time $10^{-9}$ s



Time $10^{-3}$ s

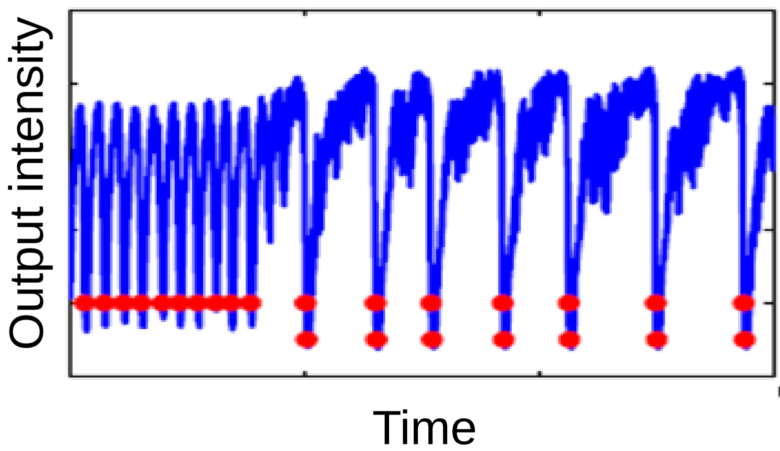
**Interdisciplinary research (at the border of different disciplines) creates new knowledge.**

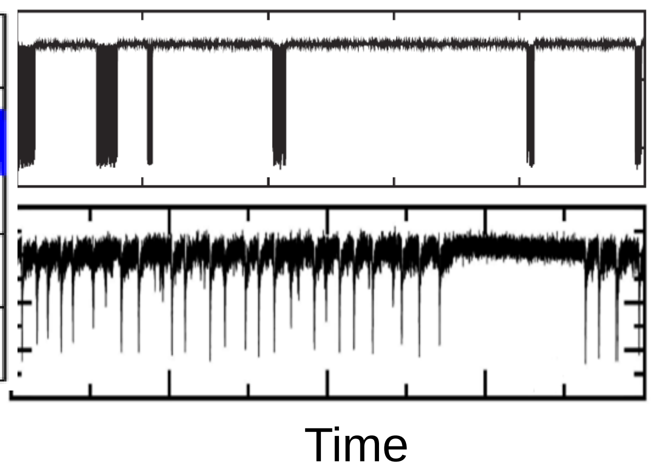


cristina.masoller@upc.edu @cristinamasoll1

## Uncovering similarities between neuron and lasers… Interesting but relevant?

- Data centers, AI and HPC systems consume a lot of power.
- Big concern in the context of climate change.
- The human brain works with only 20 Watts.
- **Laser-based neurons** should work as neurons, but
  - Much faster
  - With much less energy consumption.




**Statistically similar to neurons?**

A. Aragoneses et al., Sci. Rep. **4**, 4696 (2014).
C. Quintero-Quiroz et al., Sci. Rep. **6** 37510 (2016).

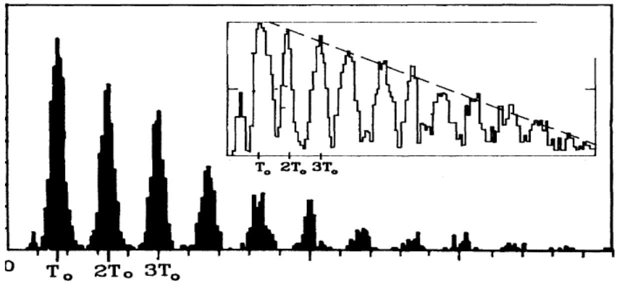


cristina.masoller@upc.edu @cristinamasoll1

# Real neuron ISI distribution

(spikes in the auditory nerve when a monkey hears a pure tone sound)

$P(\Delta T_i)$

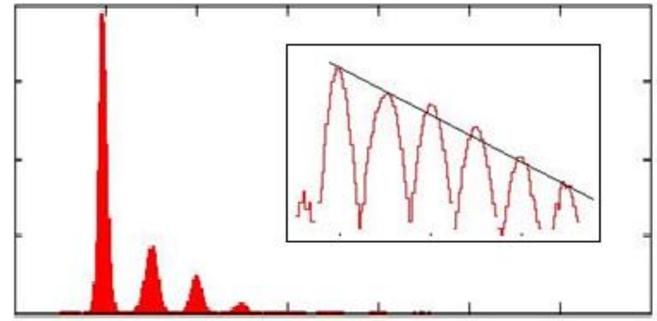



Inter-spike-interval, $\Delta T_i$

# Diode laser ISI distribution

(when the laser is sinusoidally modulated)

$P(\Delta T_i)$

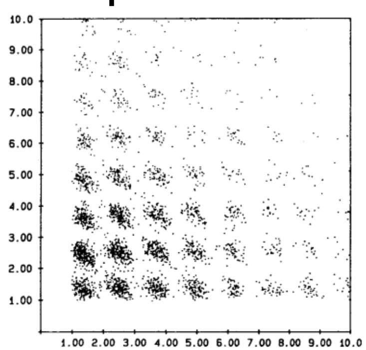


**$2T_0$ $4T_0$**

Inter-spike-interval, $\Delta T_i$

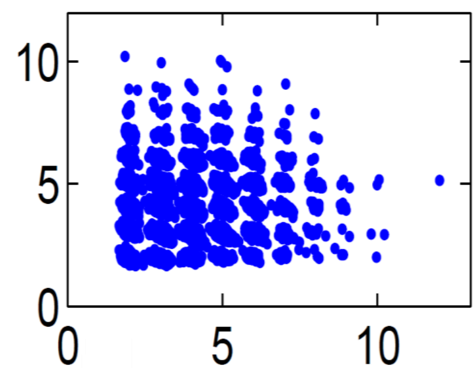
Spikes in the auditory nerve when a cat hears a pure tone sound:

$\Delta T_{i+1}$




$\Delta T_i$

$\Delta T_{i+1}$




$\Delta T_i$

**How to detect similar temporal order in the ISI sequences?**

Andre Longtin et al. PRL (1991), Int. J. Bif. Chaos (1993).

Andres Aragoneses et al. Optics Express (2014).



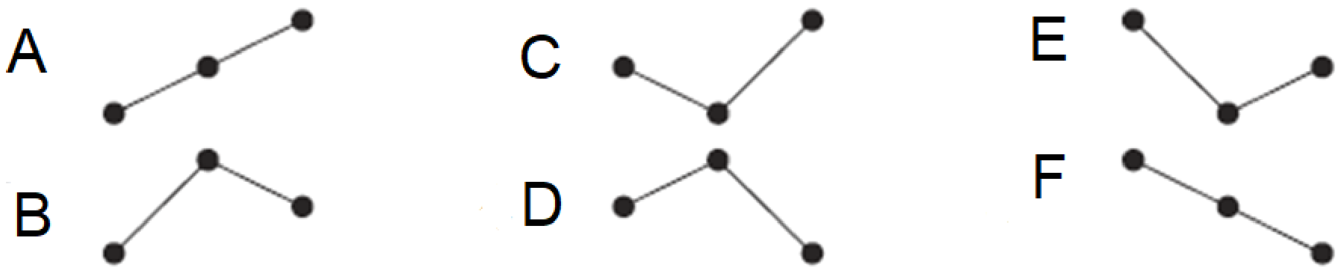
cristina.masoller@upc.edu @cristinamasoll1

# First analysis method: ordinal analysis

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

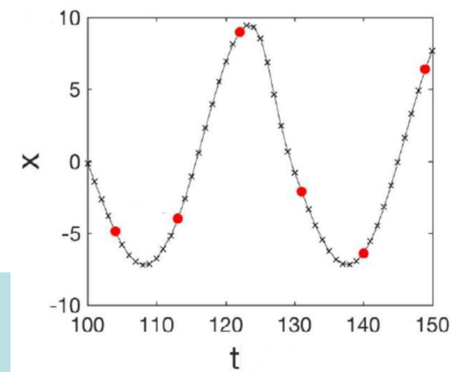
Possible order relations among three numbers (e.g., 2, 5, 7)

{…2, 5, 7…} {…5, 2, 7…} {…7, 2, 5 …}




{…2, 7, 5…} {…5, 7, 2…} {…7, 5, 2…}

**Which is the code?**




A B F C

Drawback! Information about the values is lost.

(5,7,2) and (5,70,2) are both represented by symbol “D”.

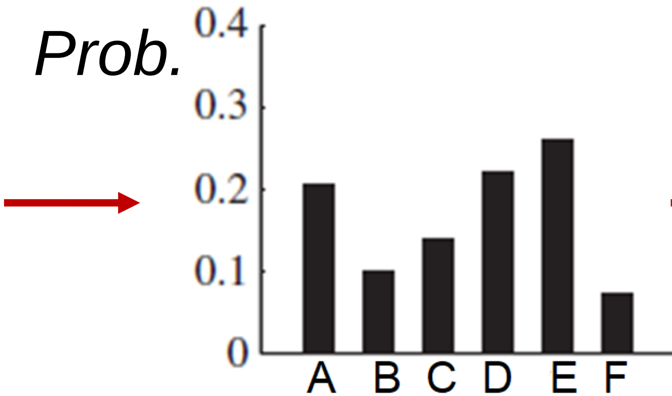
C. Bandt and B. Pompe, Phys. Rev. Lett. 88, 174102 (2002).



cristina.masoller@upc.edu @cristinamasoll1

# From a sequence of data points, by counting the different patterns we can estimate the "ordinal probabilities"

→




→ ?

**What to do next?**

1. Permutation Entropy:

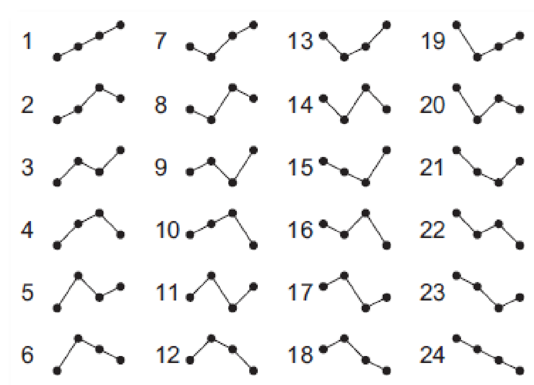
$p_i = p_j$ for all $i, j \Rightarrow H=1$
$p_i = 1,\ p_j = 0$ for all $j \neq i \Rightarrow H=0$

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

(Nonlinear dimensionality reduction)

2. Analyze all the probabilities

Another drawback: The number of possible patterns (N) increases with the length (D) of the pattern as D!



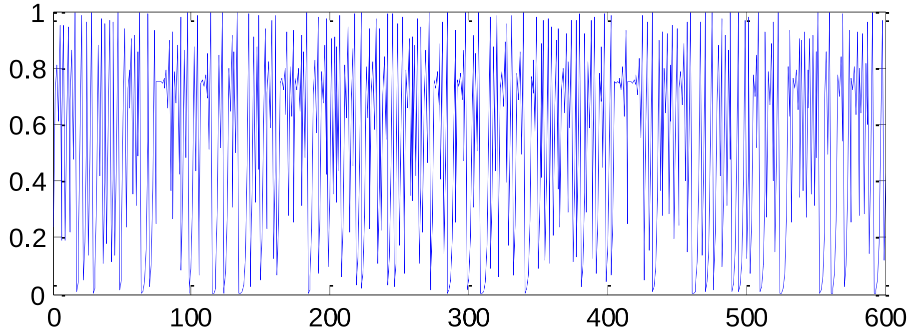



cristina.masoller@upc.edu @cristinamasoll1

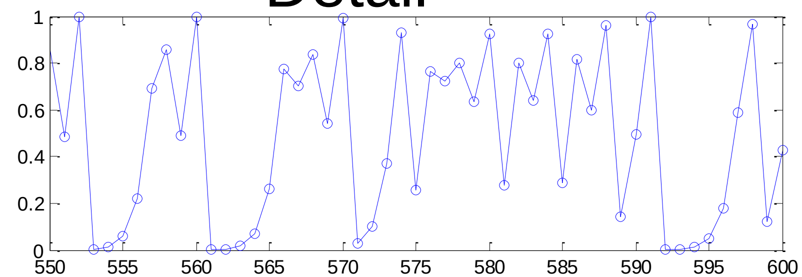
# Example: chaotic time series generated with the Logistic map

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

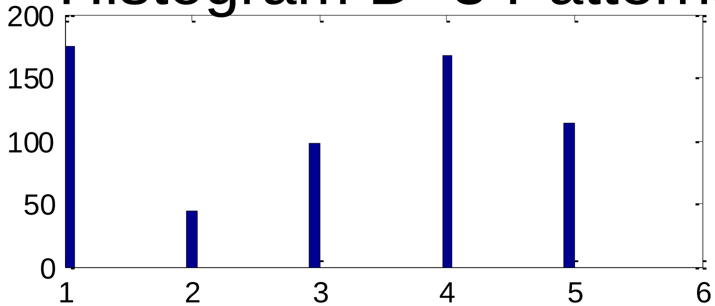
Time series



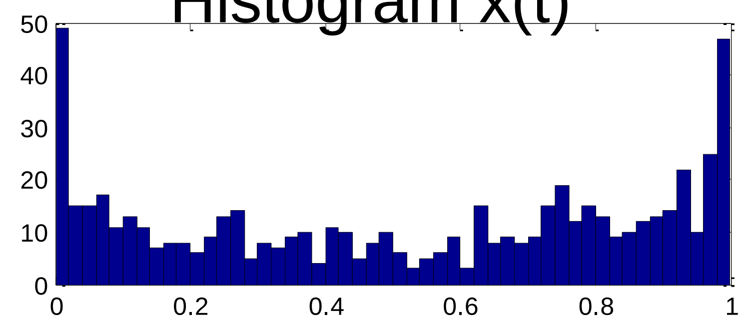

Detail

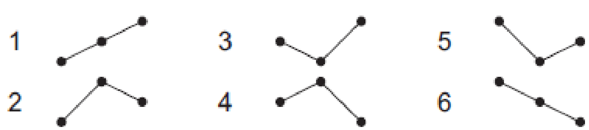



Histogram D=3 Patterns




Histogram x(t)



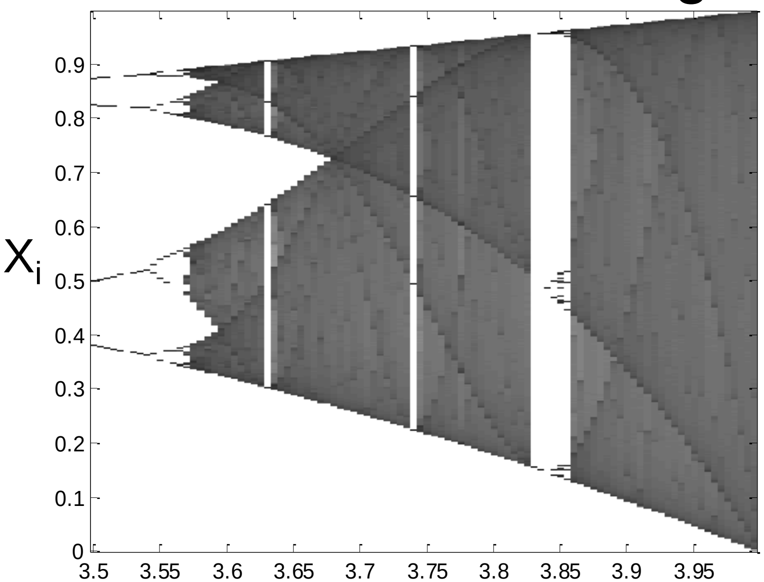

↑ **forbidden**



cristina.masoller@upc.edu @cristinamasoll1

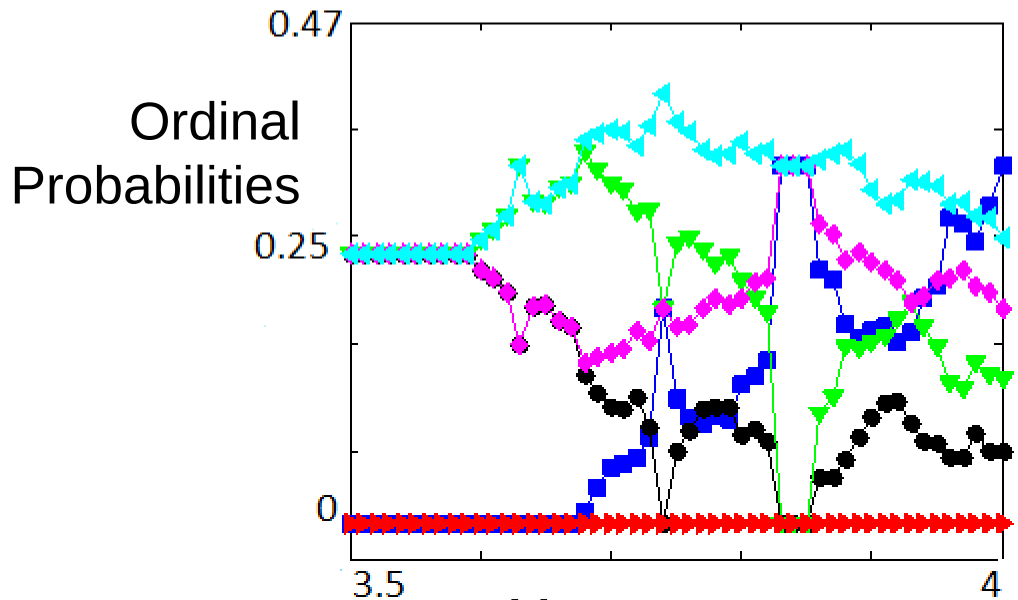
# “Normal” and “Ordinal” bifurcation diagrams of the Logistic map

Normal bifurcation diagram




Map parameter

Ordinal diagram with D=3




Map parameter, r

**012 021 102 120 201 210**

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



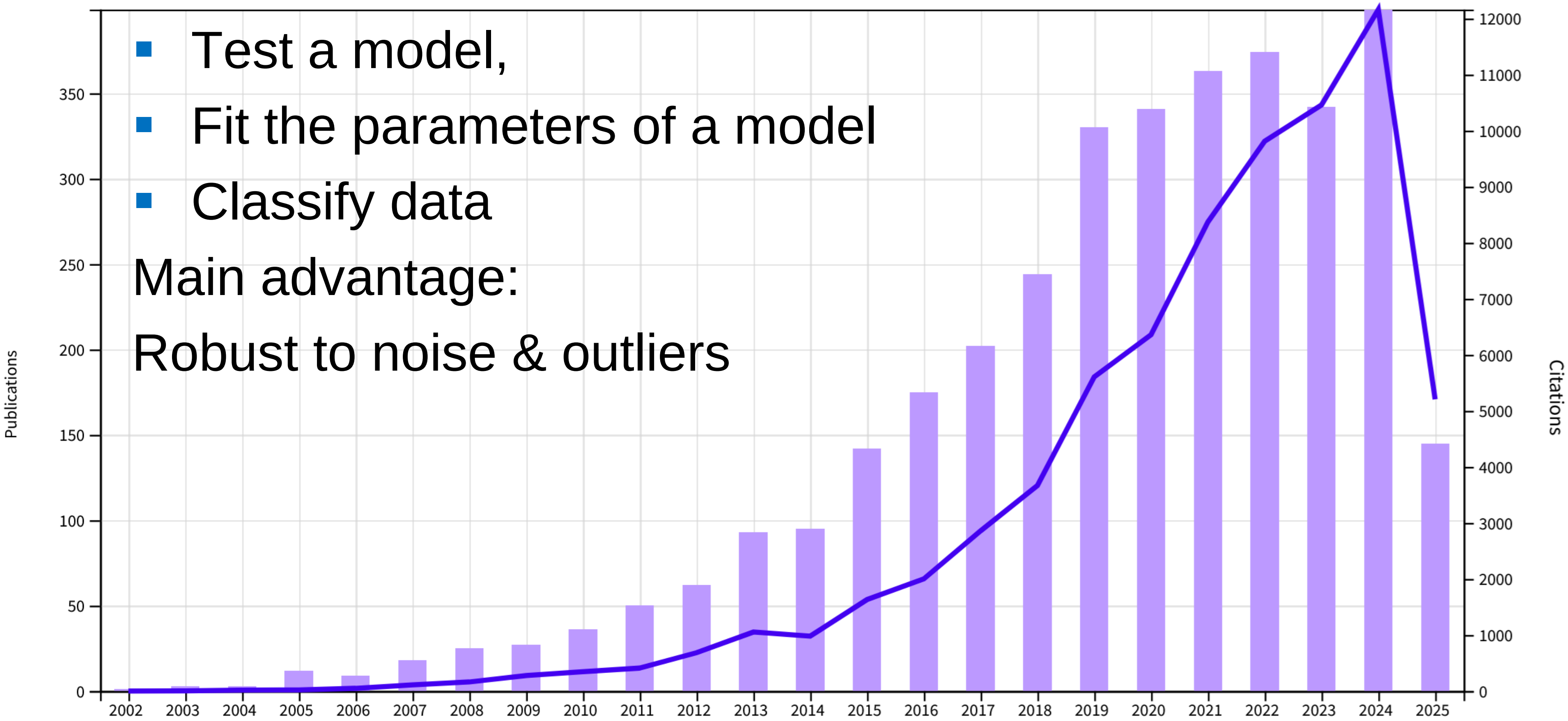
cristina.masoller@upc.edu @cristinamasoll1

# Ordinal analysis is a popular technique to:

- Test a model,
- Fit the parameters of a model
- Classify data

Main advantage:

Robust to noise & outliers

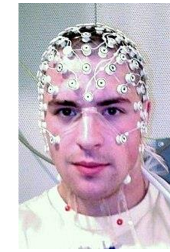



I. Leyva, J. M. Martinez, C. Masoller, O. A. Rosso, M. Zanin, *"20 Years of Ordinal Patterns: Perspectives and Challenges"*, EPL 138, 31001 (2022).

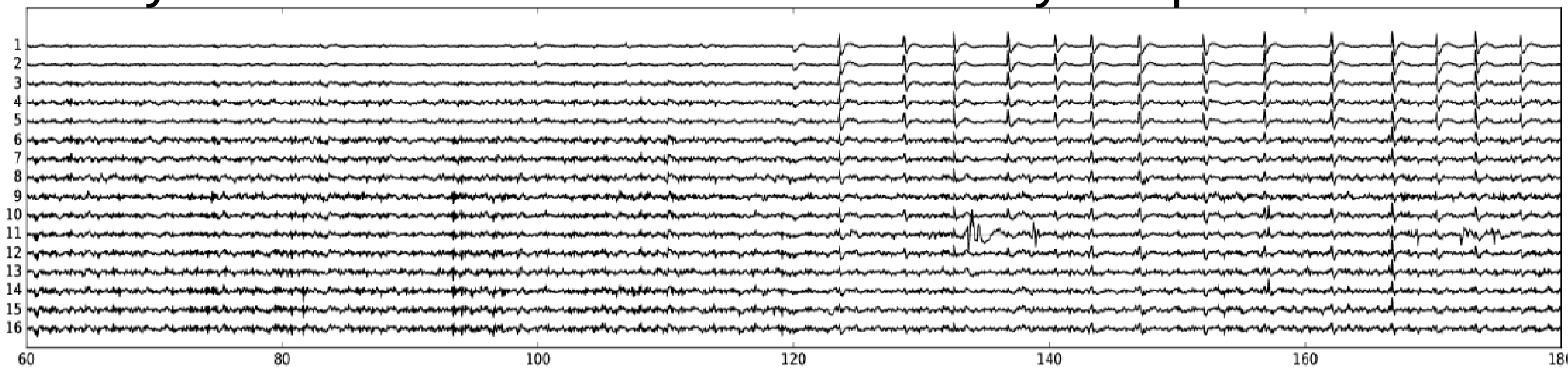


cristina.masoller@upc.edu @cristinamasoll1

# First example: entropy analysis of EEG recordings can distinguish eyes-closed and eyes-open states?



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 (healthy) | 71 | 109 |

DTS1: Bitbrain (Zaragoza)
DTS2: Physionet

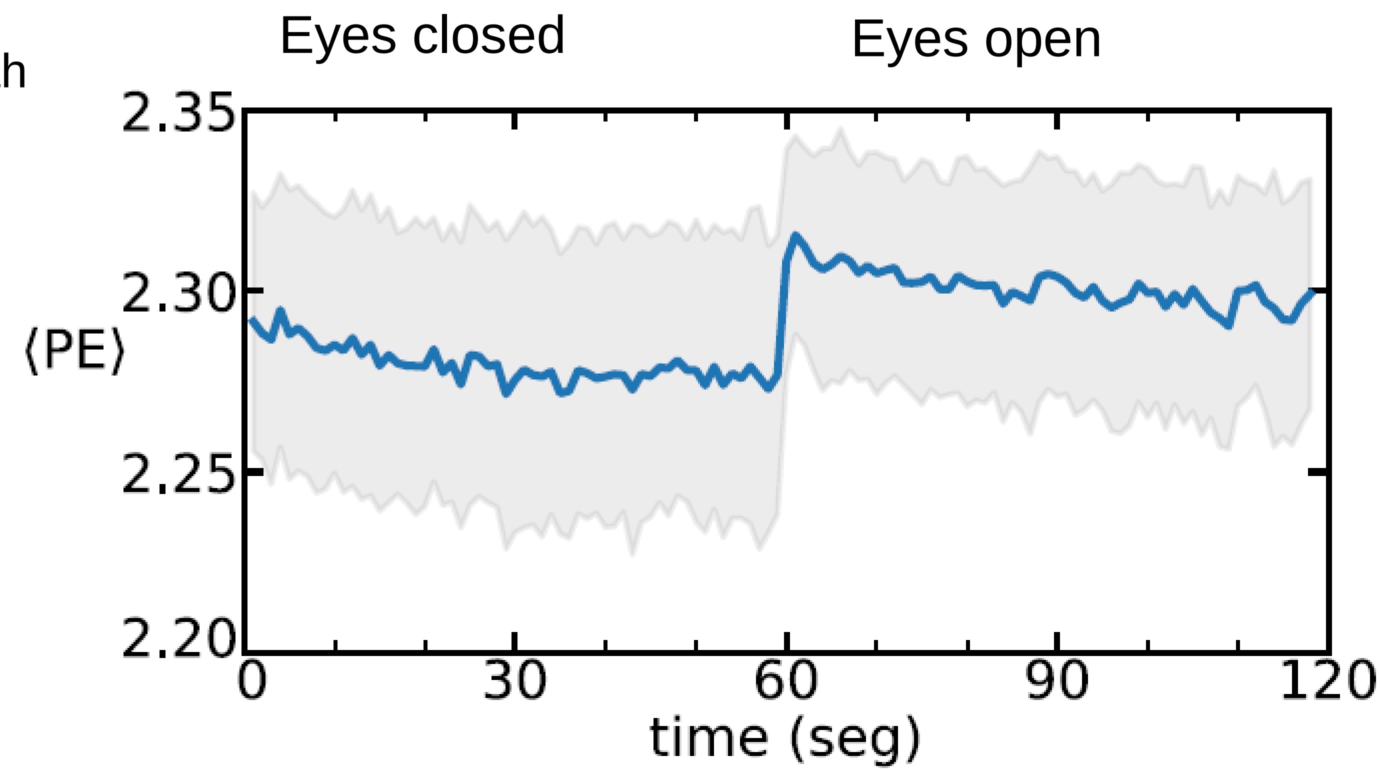


cristina.masoller@upc.edu @cristinamasoll1

# 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



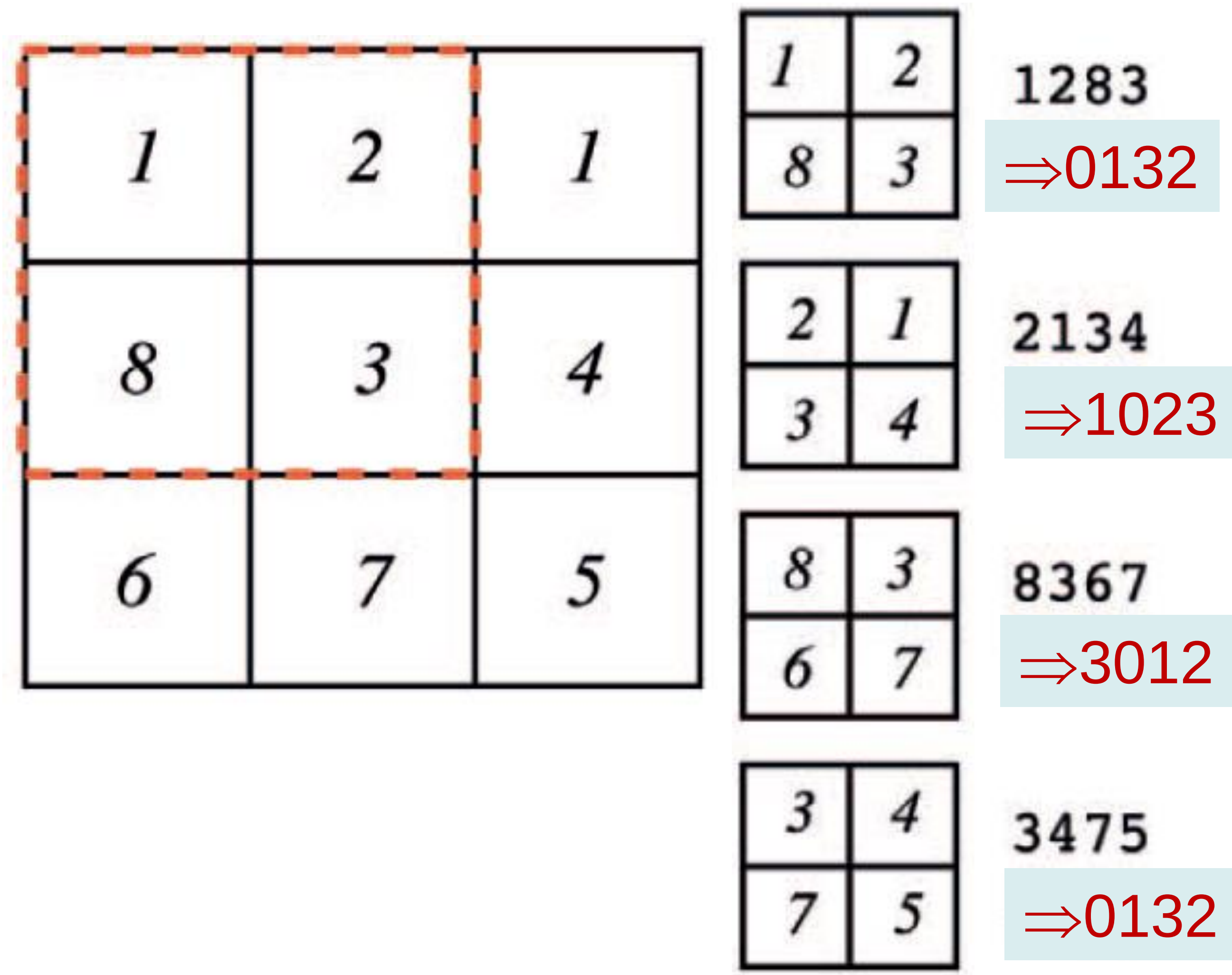

Gray region: Standard deviation of ⟨PE⟩ across subjects

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



cristina.masoller@upc.edu @cristinamasoll1

# Spatial data ⇒ Spatial Permutation Entropy (SPE)




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

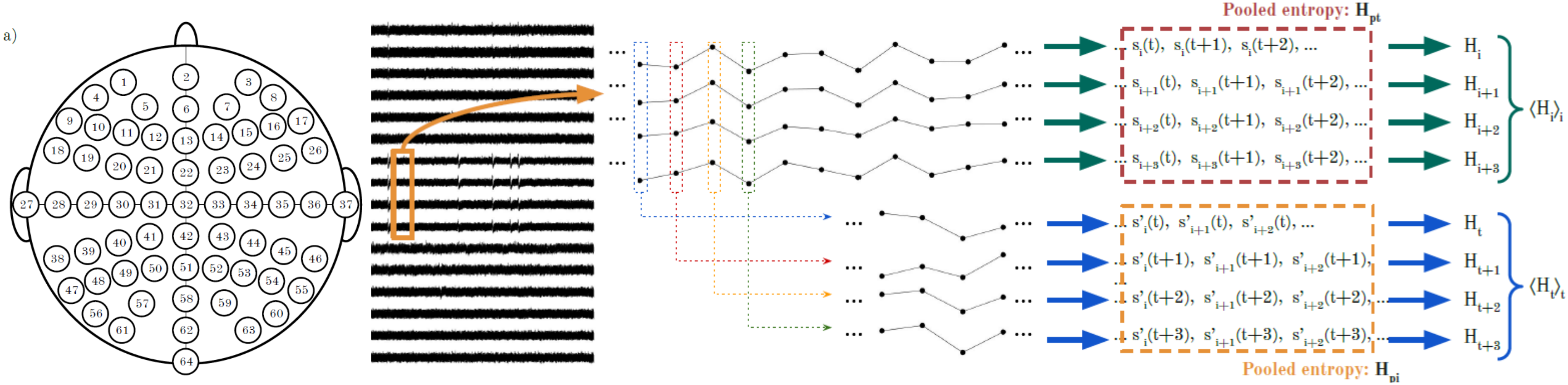
2x2 pixels:
24 possible
patterns

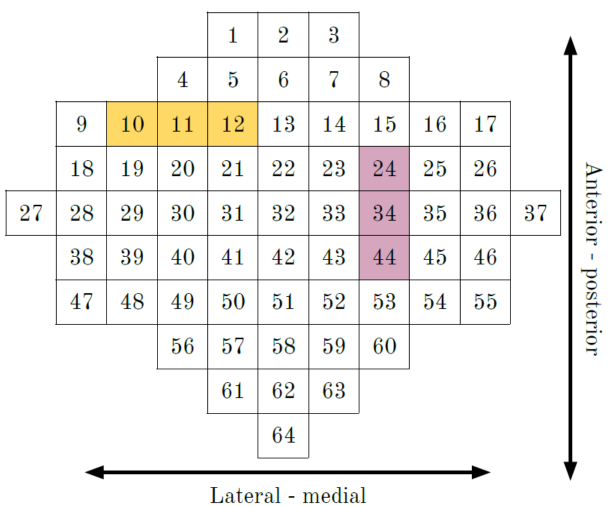
*Haroldo V. Ribeiro and coworkers, PLoS ONE 7, e40689 (2012)*

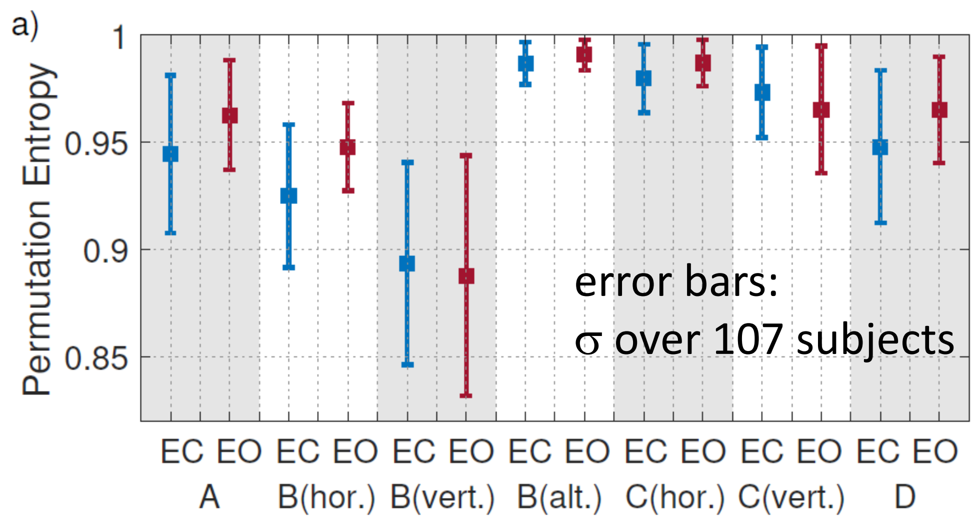


cristina.masoller@upc.edu @cristinamasoll1

# Different approaches to calculate permutation entropy



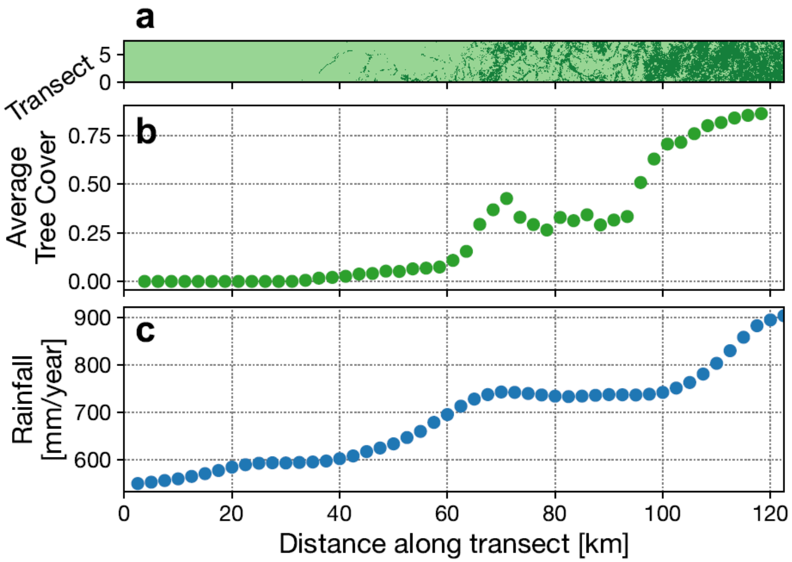

Juan Gancio, C. Masoller, G. Tirabassi, Chaos 34, 043130 (2024).

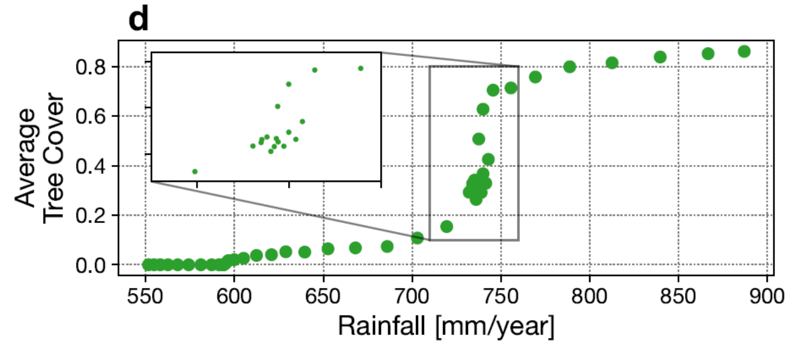
A: temporal coding
B: spatial coding (horizontal, vertical, alt.)
C: spatial pooling (horizontal and vertical symbols)
D: temporal pooling

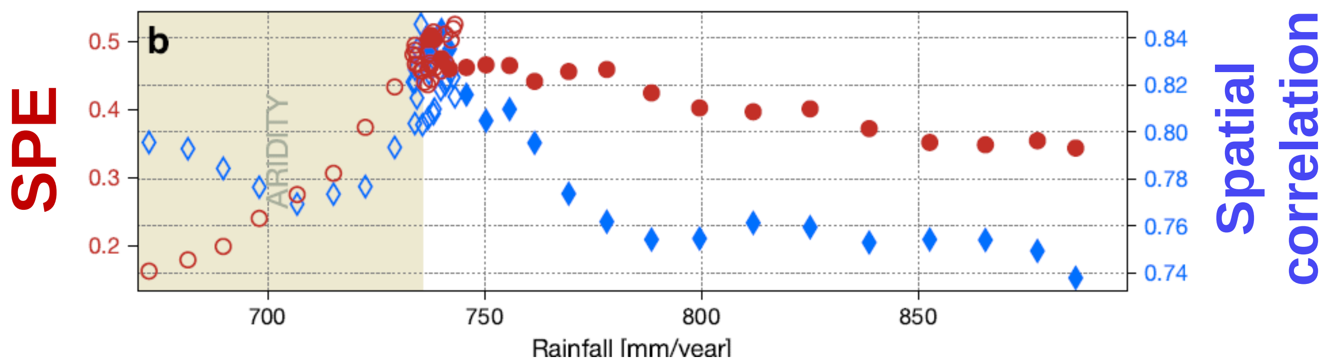


cristina.masoller@upc.edu @cristinamasoll1

# The spatial permutation entropy (SPE) can be an early indicator of a vegetation transition (tipping point)



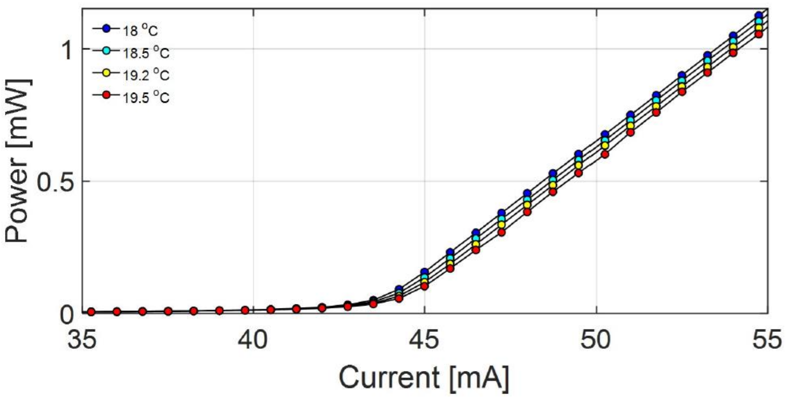

**Can we test this indicator in controlled experimental data?**

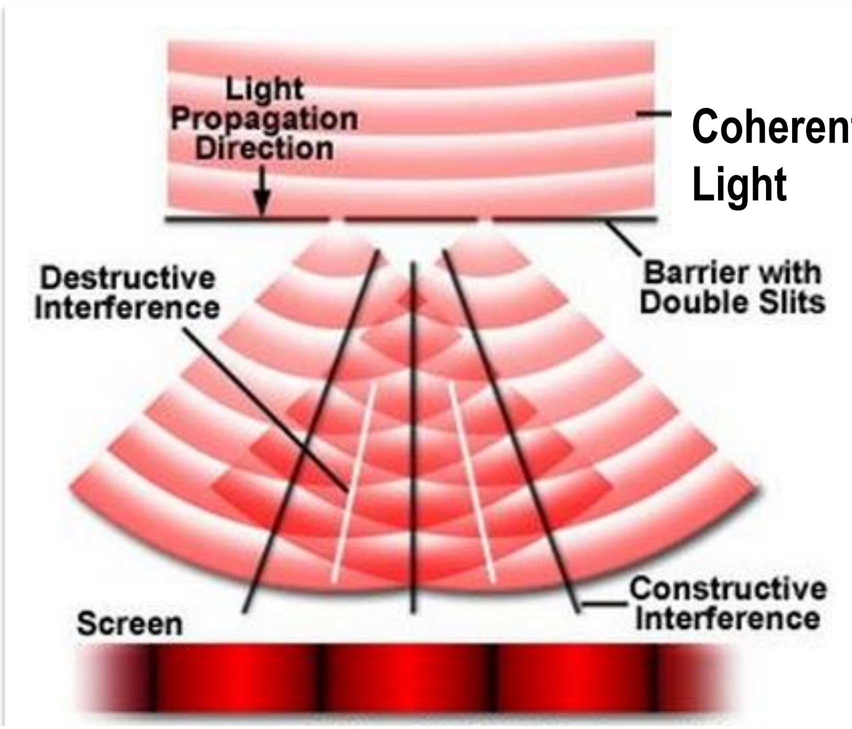
Giulio Tirabassi and C. Masoller, "*Entropy-based early detection of critical transitions in spatial vegetation fields*", PNAS 120, e2215667120 (2023).



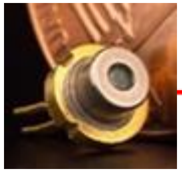
cristina.masoller@upc.edu @cristinamasoll1

# Quick reminder: laser threshold and interference of coherent waves

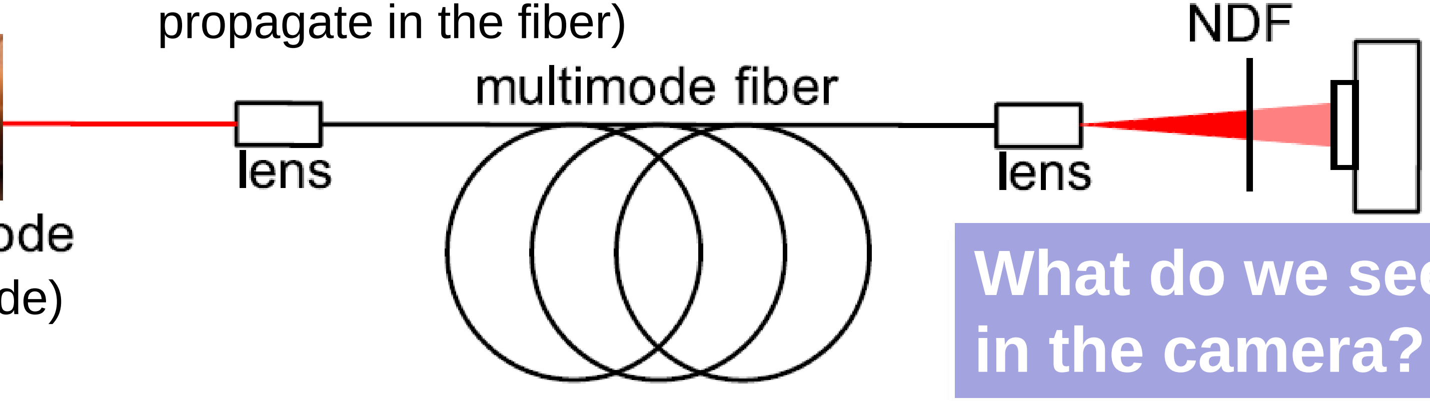



(many modes can propagate in the fiber)



laser diode
(multimode)




**What do we see in the camera?**

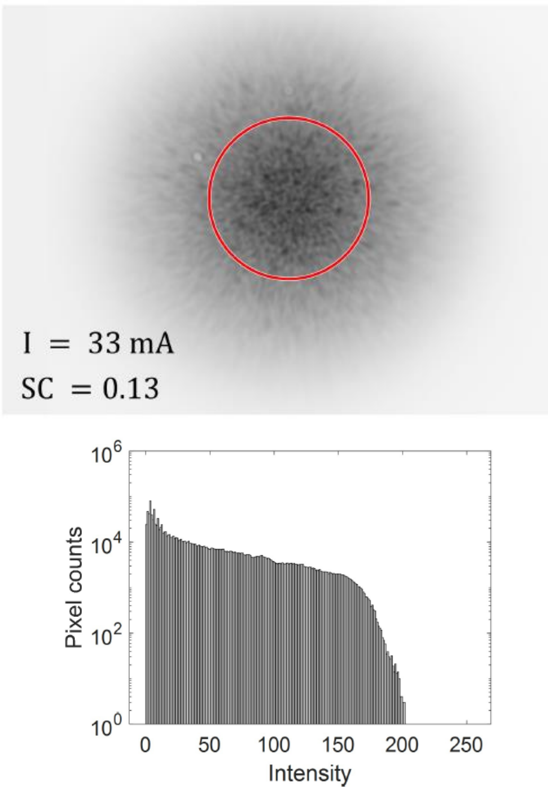


cristina.masoller@upc.edu @cristinamasoll1

# Speckle pattern

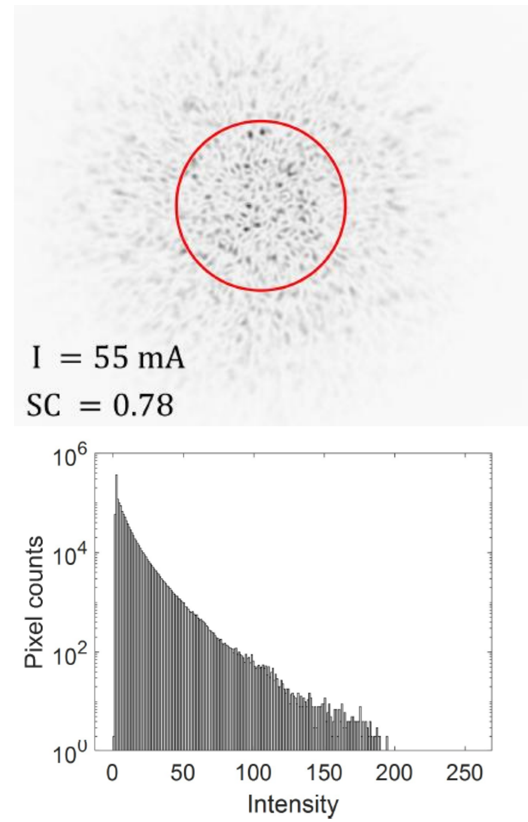
## Below lasing threshold

Low coherence → low speckle contrast




## Above lasing threshold

High coherence→ high speckle contrast




**How to quantify the degree of speckle?**

Speckle contrast (coefficient of variation, CV)

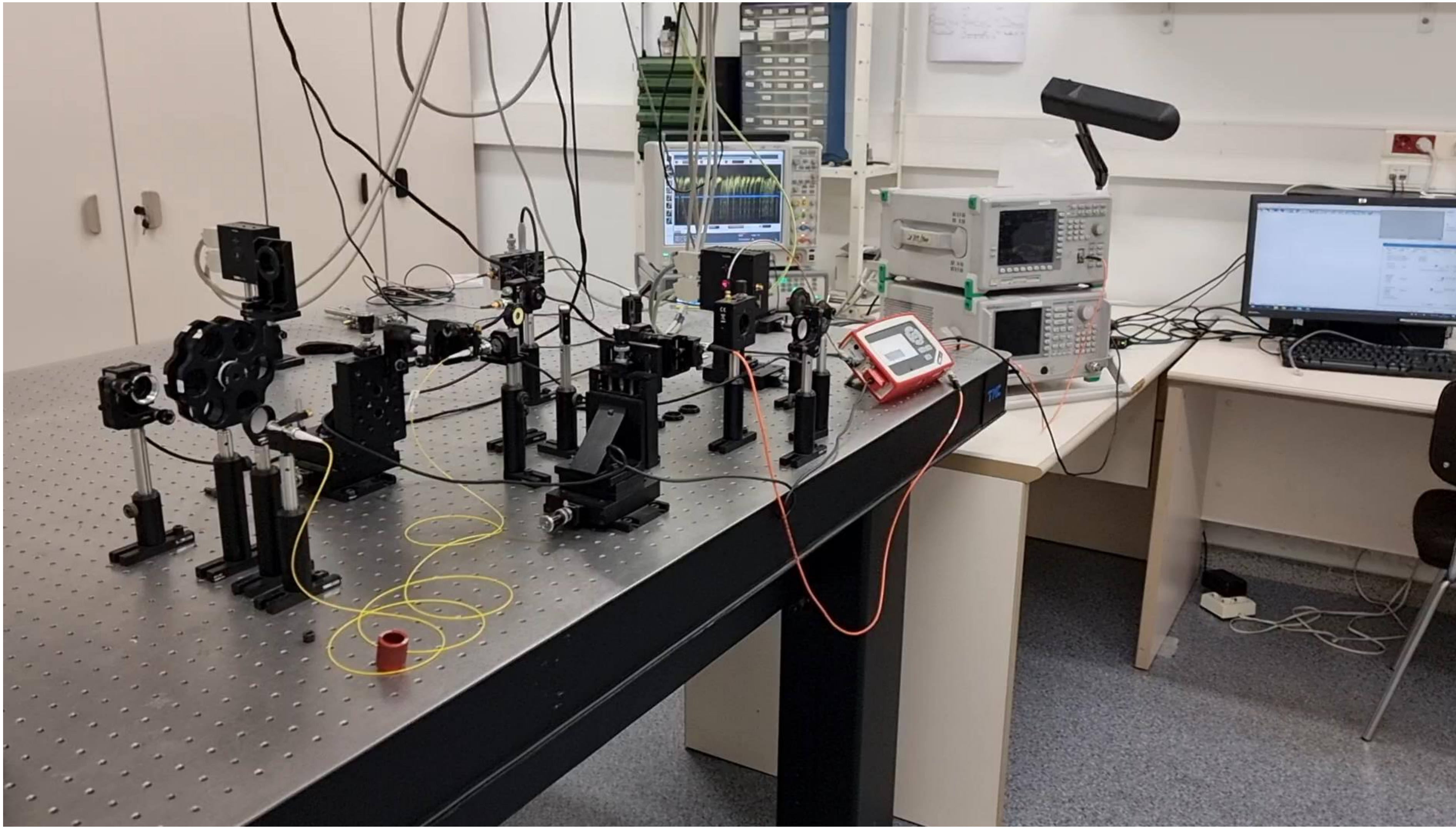
$$\boxed{SC = \sigma/\langle I \rangle}$$



cristina.masoller@upc.edu @cristinamasoll1

Three different diffusive media are used to generate speckle:

Multimode fiber --- Multimode fiber and diffuser --- Single mode fiber and diffuser





cristina.masoller@upc.edu @cristinamasoll1

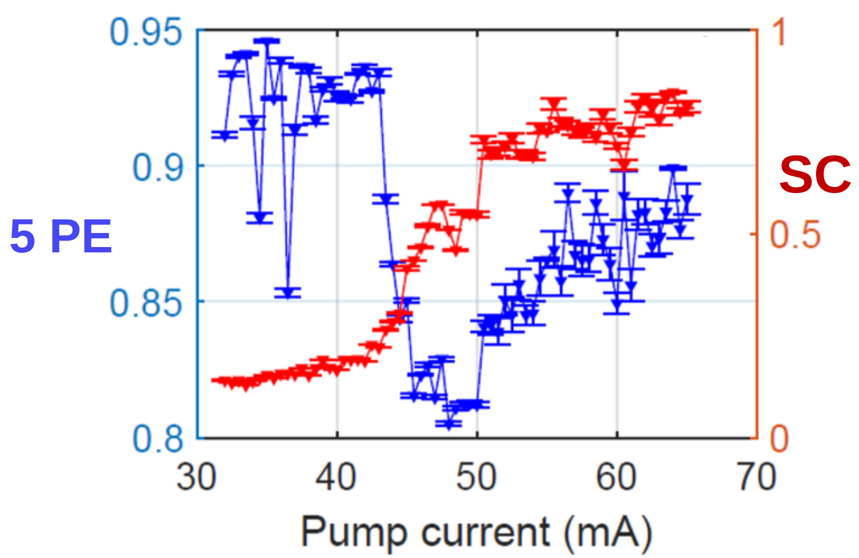
## Can we identify, from the entropy of the patterns, the laser threshold? and infer which medium generated the speckles?

$$\text{PE} = -\frac{1}{\ln N!}\sum_{i=1}^{N} p_i \ln p_i$$

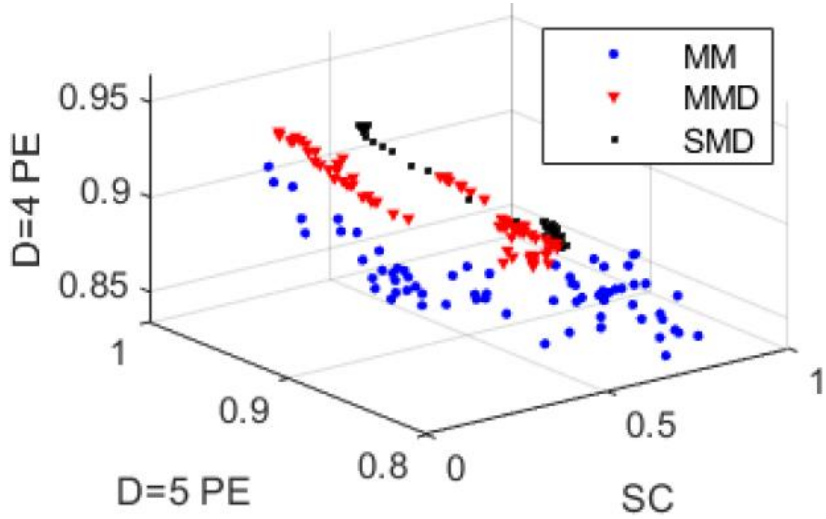
$$\text{SC} = \sigma/\langle I \rangle$$

4 PE : x x
x x

5 PE : x
x x x
x




Diffusive Medium:
Multimode fiber
Multimode fiber and diffuser
Single mode fiber and diffuser




Accuracy of the random forest classifier: 99.4 % ± 0.4 %

**Are the ordinal probabilities also informative?**

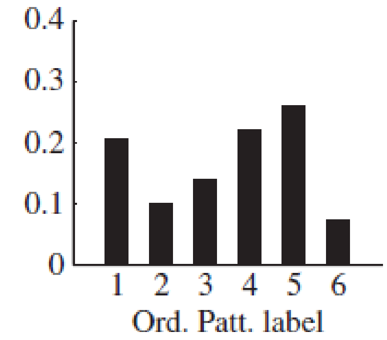
Giulio Tirabassi et al., APL Photonics 8, 126112 (2023).



cristina.masoller@upc.edu @cristinamasoll1

## Analysis of the ordinal probabilities

Are the D! ordinal patterns equally probable?



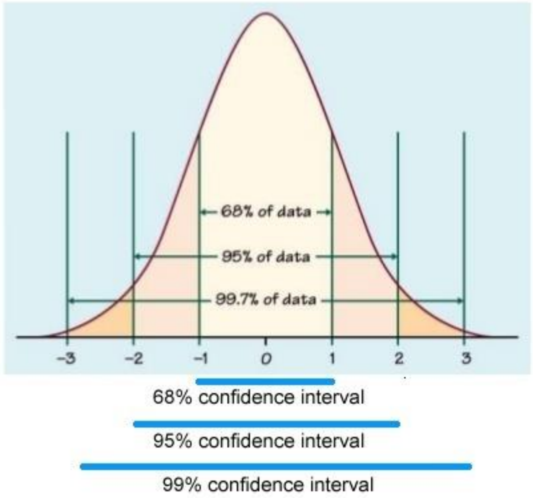

- **Null hypothesis**:

  $p_i = p = 1/D!$ for all $i = 1 \ldots D!$

- If at least one probability **is not** in the interval $p \pm 3\sigma$ with $\sigma = \sqrt{p(1-p)/N}$ and $N$ the number of ordinal patterns:

  We **reject** the NH with 99.74% confidence level.

- Else, we **fail to reject** the NH with 99.74% confidence level.




**Back to question: How to detect similar temporal order in ISI sequences of neurons and lasers?**

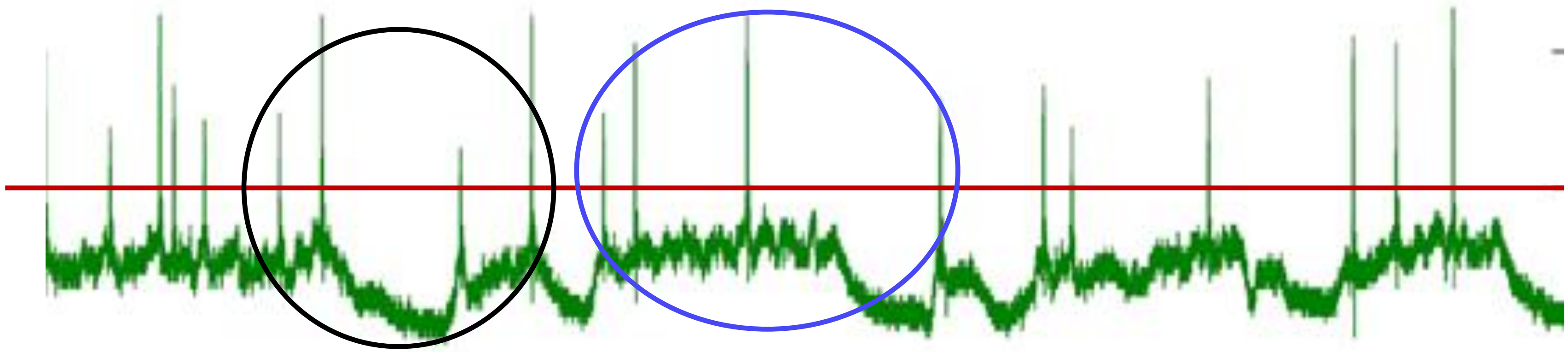


cristina.masoller@upc.edu @cristinamasoll1

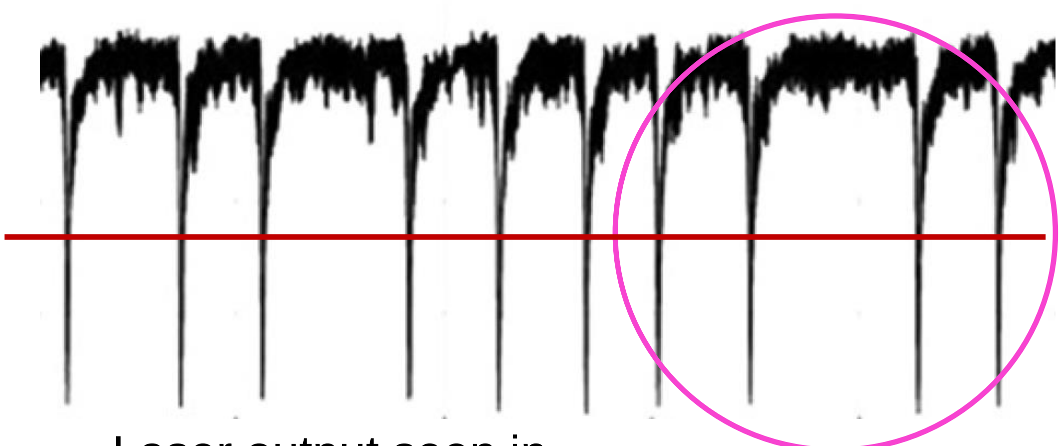
# Sequence of inter-spike-intervals (ISIs) ⇒ sequence of ordinal patterns

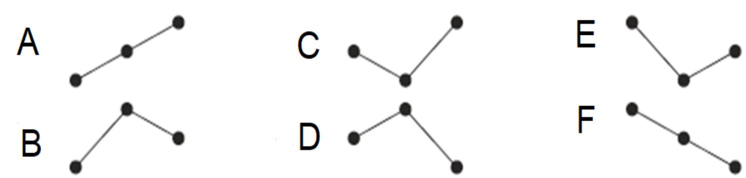
**D=3**

Intracellular recording



**021=B** **012=A**






Laser output seen in
1 GHz oscilloscope

**120=D**



cristina.masoller@upc.edu @cristinamasoll1

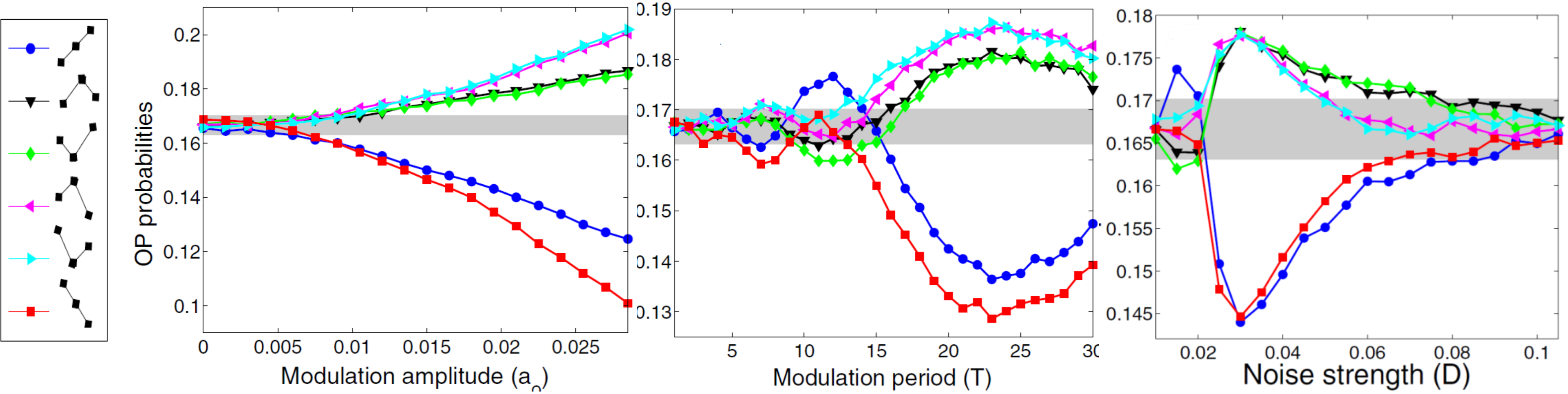
# Analysis of inter-spike-sequences (ISIs) simulated with a simple neuron model (FitzHugh-Nagumo)

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

$$\frac{dy}{dt} = x + a + \boxed{a_o \cos(2\pi t/T)} + D\xi(t),$$

Weak, **subthreshold** input

Gaussian white noise




$\Rightarrow$ More and less probable patterns depend on the period and amplitude of the input and the level of noise.

Gray region: NH with 99.74% confidence level

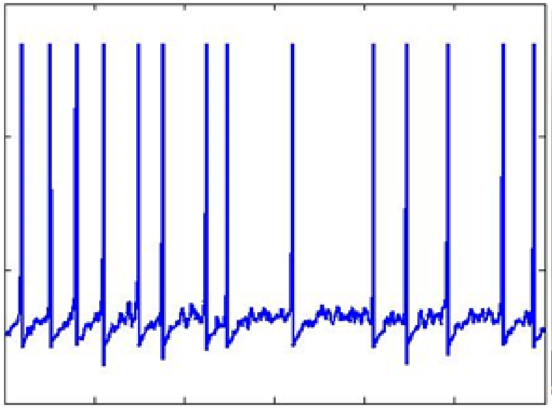
J. A. Reinoso, M. C. Torrent, and C. Masoller, Phys. Rev. E. 94, 032218 (2016).



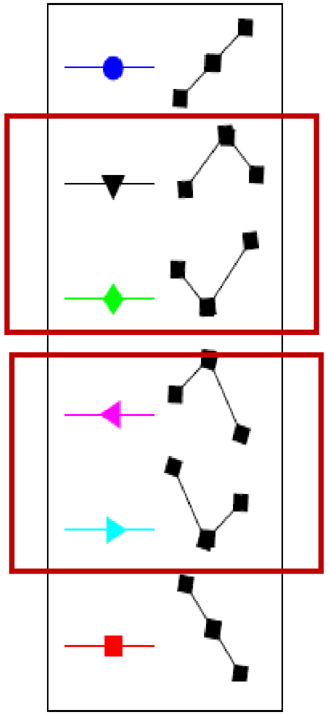
cristina.masoller@upc.edu @cristinamasoll1

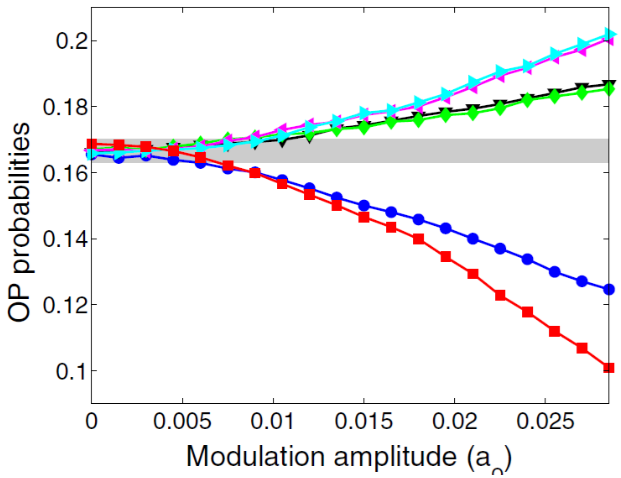
# Comparing the spike timing of a neuron and a laser when they are weakly forced with a sinusoidal input

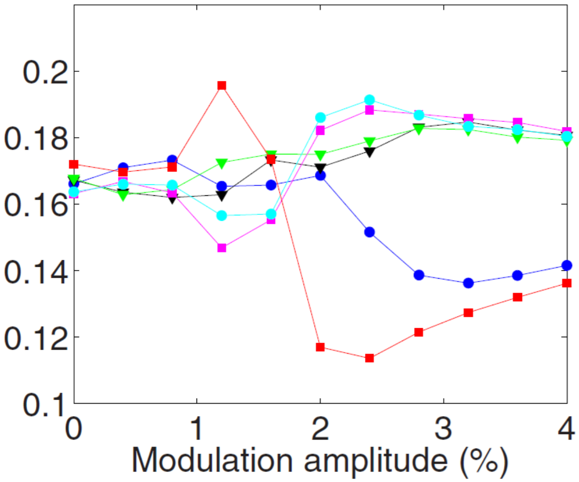
## Neuron model



## Diode laser with feedback






J. M. Aparicio-Reinoso et al PRE 94, 032218 (2016)

A. Aragoneses et al, Sci. Rep. **4**, 4696 (2014)



cristina.masoller@upc.edu @cristinamasoll1

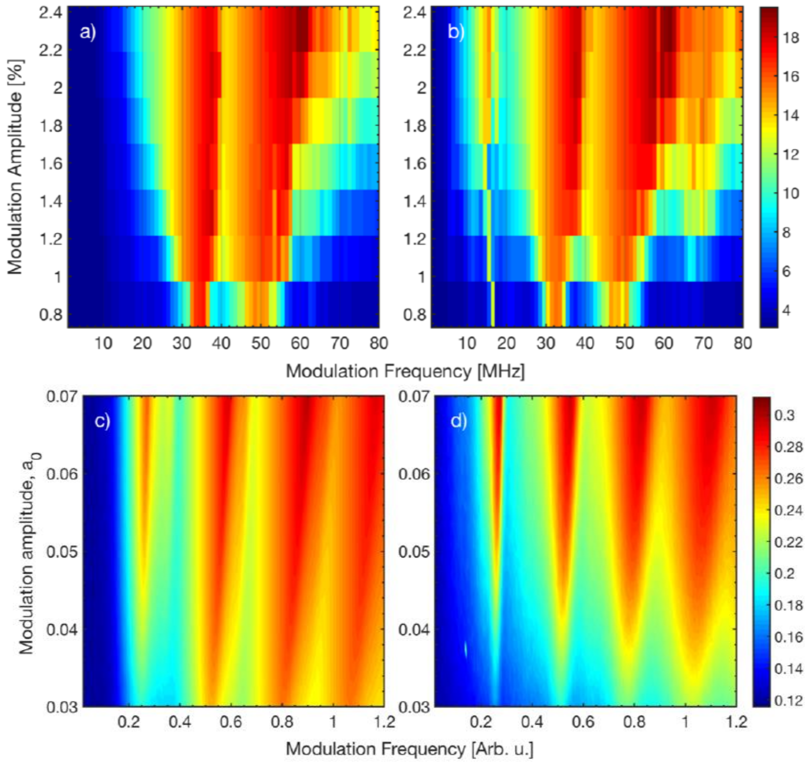
# Laser-neuron comparison: spike rate

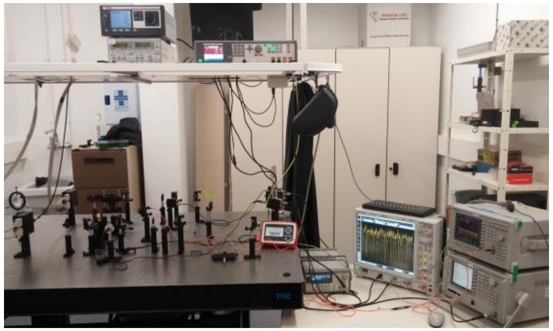
## Spike rate in color code

Sinusoidal    Pulsed signal

Experiments modulating the laser current

Neuron model with the same (re-scaled) input



**How about temporal order in the ISI sequences?**

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) (2019).

J. Tiana-Alsina and C. Masoller, Sci. Rep. 12, 4914 (2022).



cristina.masoller@upc.edu @cristinamasoll1

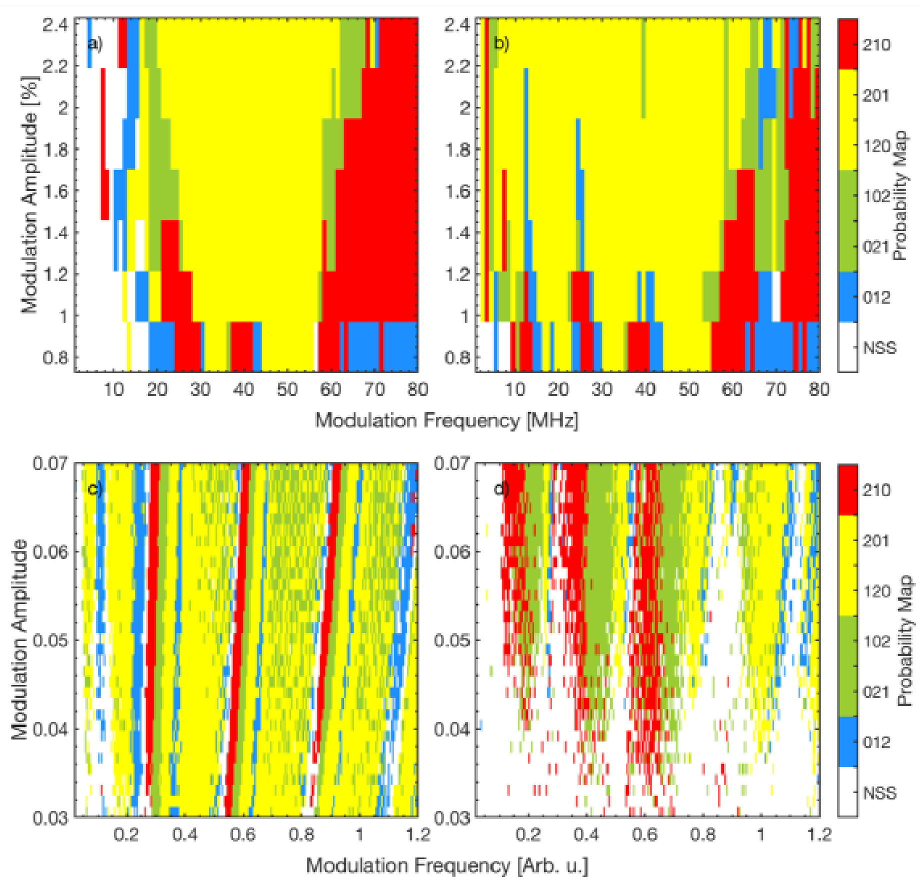
# Ordinal analysis uncovers differences in ordinal probabilities

## Most probable pattern in color code

Sinusoidal    Pulsed signal

Diode laser with optical feedback

FitzHugh-Nagumo model

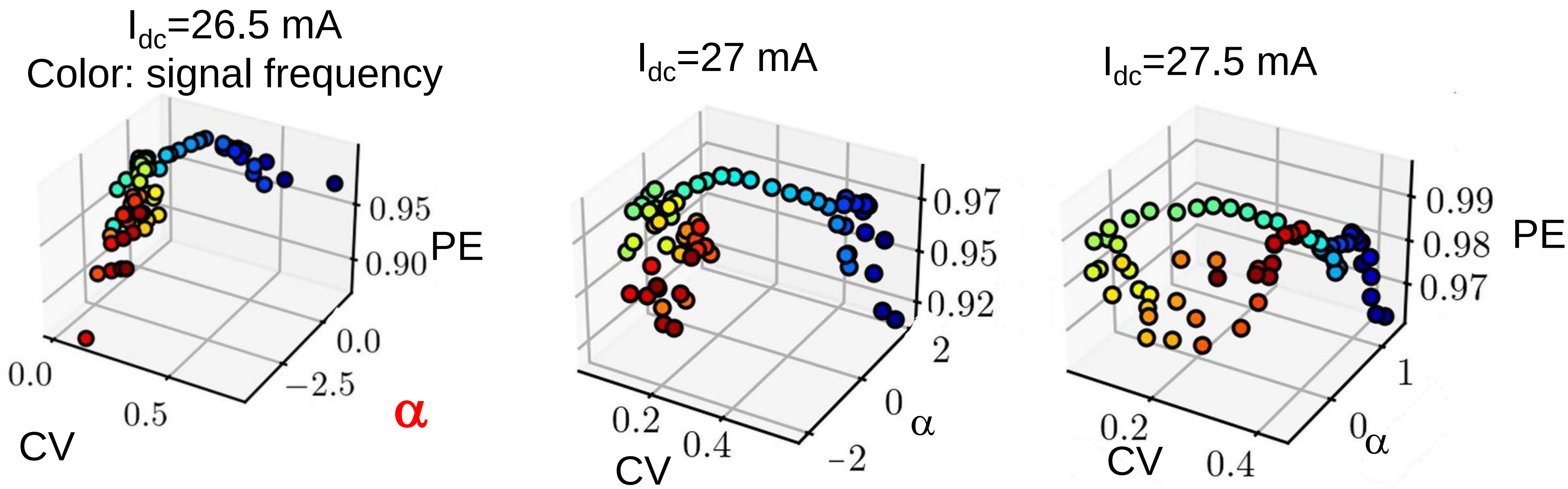



J. Tiana-Alsina, C. Quintero-Quiroz and C. Masoller, New J. of Phys. 21, 103039 (2019).



cristina.masoller@upc.edu @cristinamasoll1

# Can we "decode" information of the signal applied to the laser, from the analysis of the laser output?




CV=σ/⟨ISI⟩

Machine learning fit of the spectrum of the ISI time series to flicker noise, $P(f)=1/f^{\alpha}$

Bruno R. R. Boaretto, R. C. Budzinski, K. L. Rossi, T. L. Prado, S. R. Lopes, C. Masoller, "*Discriminating chaotic and stochastic time series using permutation entropy and artificial neural networks*", Sci. Rep. 11, 15789 (2021).
Bruno R. R. Boaretto, E. E. N. Macau, C. Masoller, "*Characterizing the spike timing of a chaotic laser by using ordinal analysis and machine learning*", Chaos 34, 043108 (2024).



cristina.masoller@upc.edu @cristinamasoll1

# Single-neuron vs ensemble encoding

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

subthreshold input

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

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

$k_i = \sum_j a_{ij}$

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

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

Random with prob. ***p***

Maria 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).

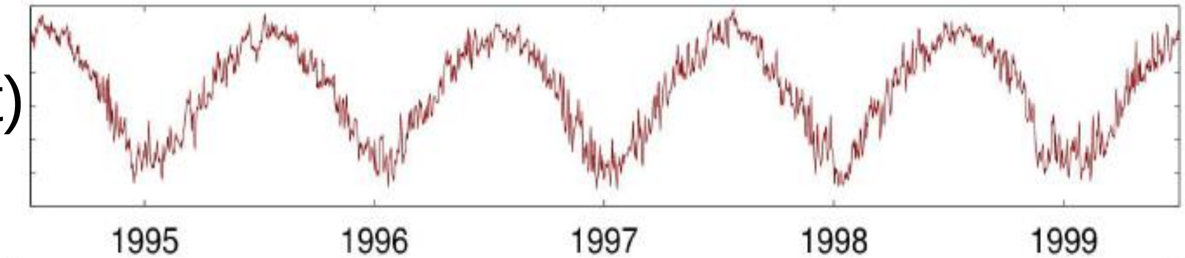


cristina.masoller@upc.edu @cristinamasoll1

# Second data analysis method: Hilbert analysis

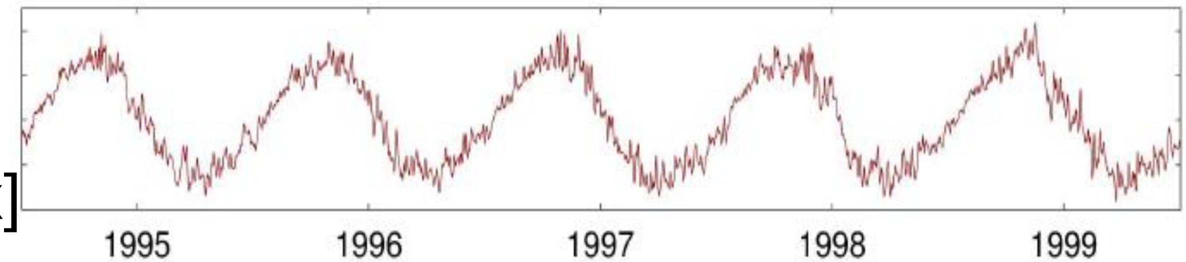
Surface Air Temperature (SAT) in a region

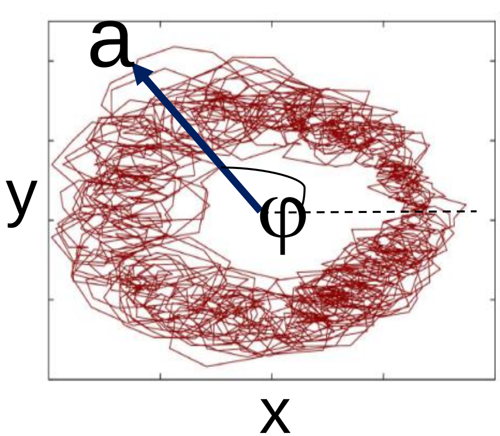
x(t)




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

y(t)= HT[x]




*Instantaneous* amplitude and phase

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

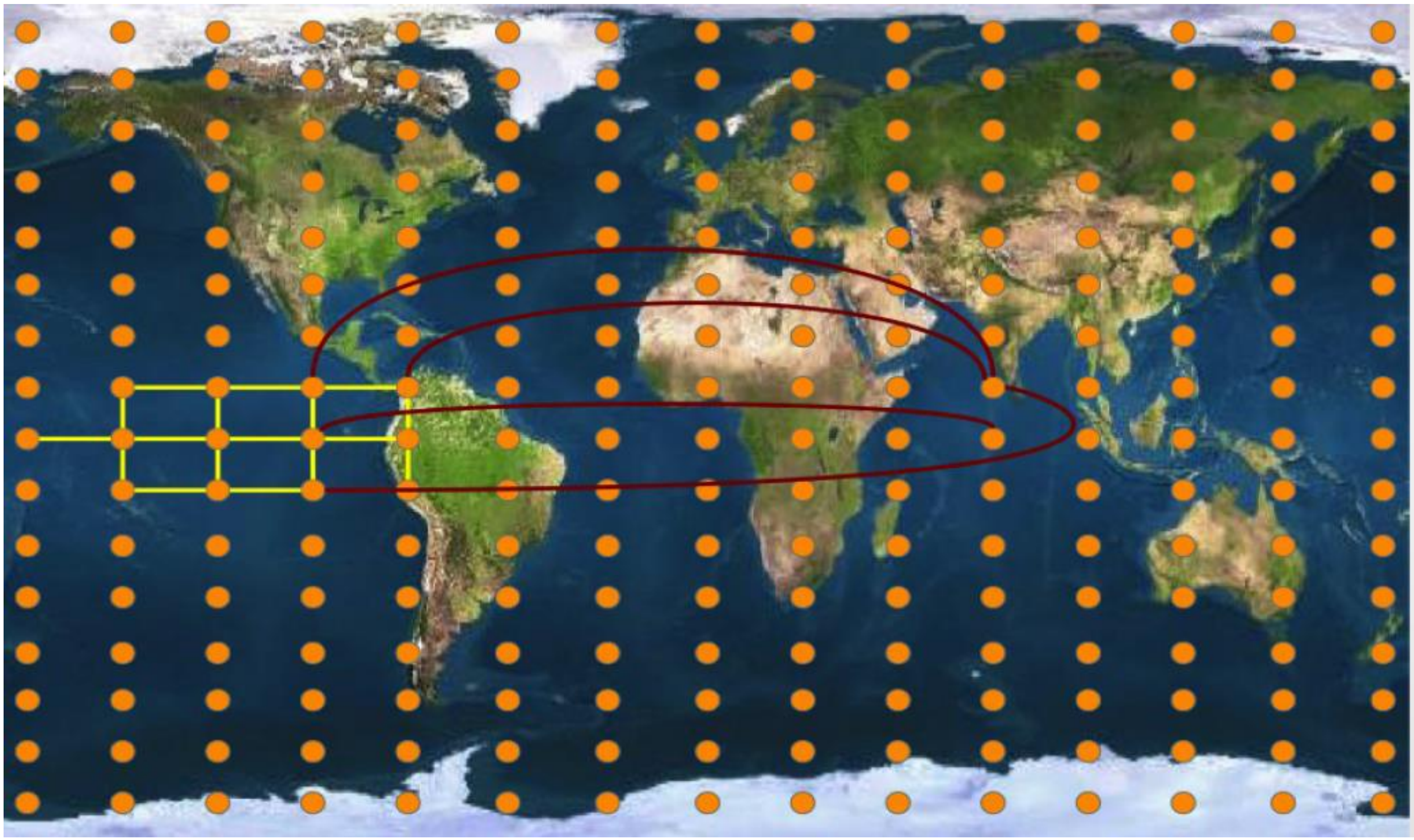
$$\varphi(t) = \arctan[y(t)/x(t)]$$

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



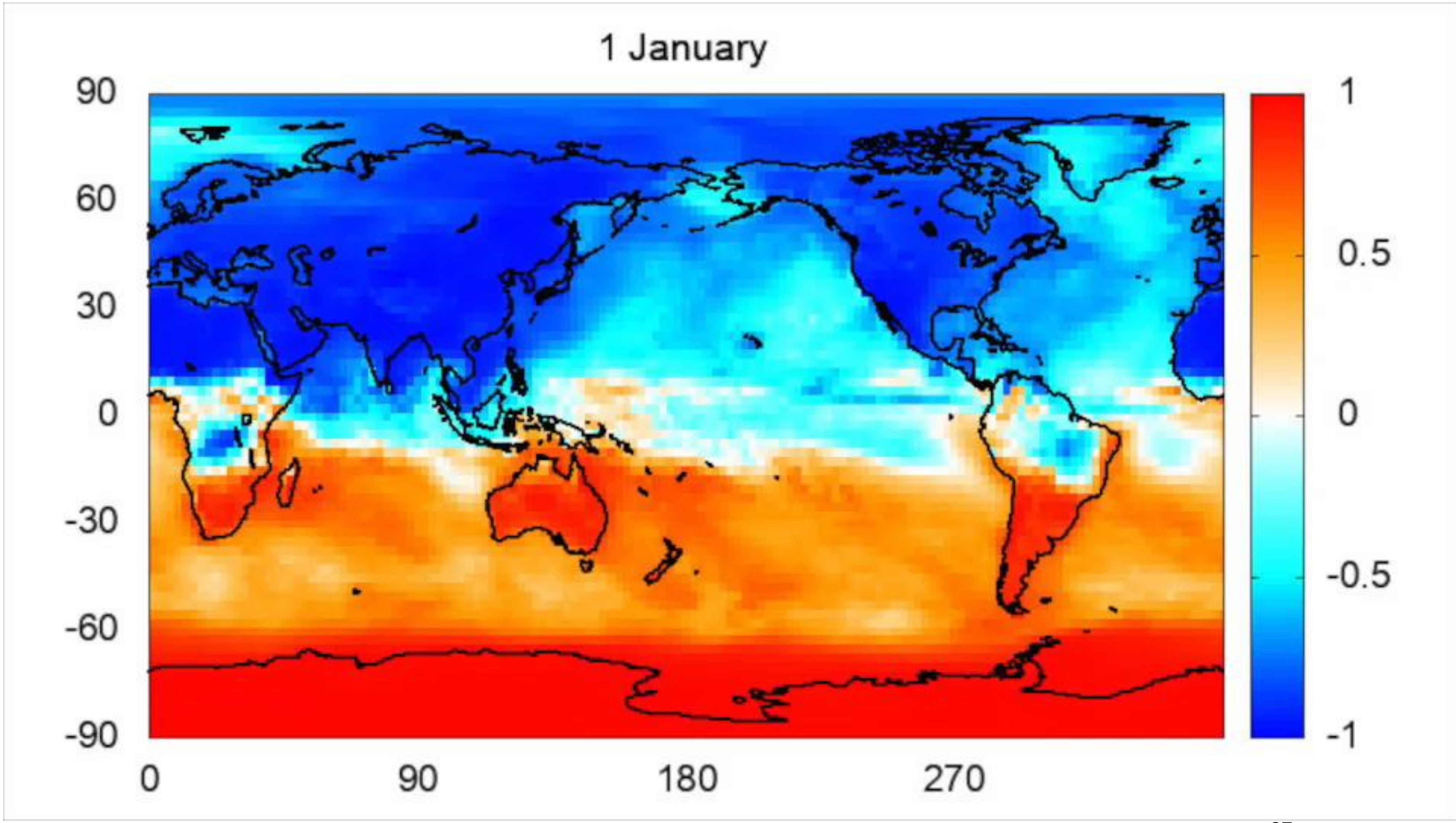
73 x 144 = 10 512 geographical sites, in each site the SAT time series has 13696 days



cristina.masoller@upc.edu @cristinamasoll1

# Average of the cosine of Hilbert phase of surface air temp.

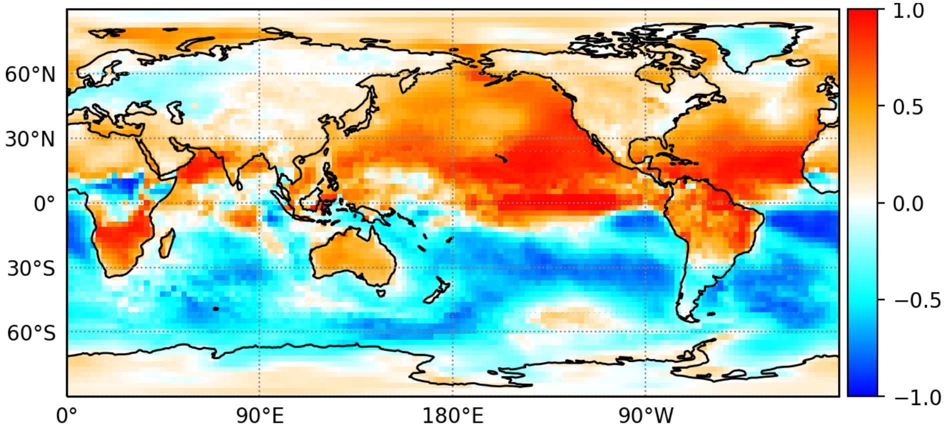
**Can we visualize the passing of the seasons?**



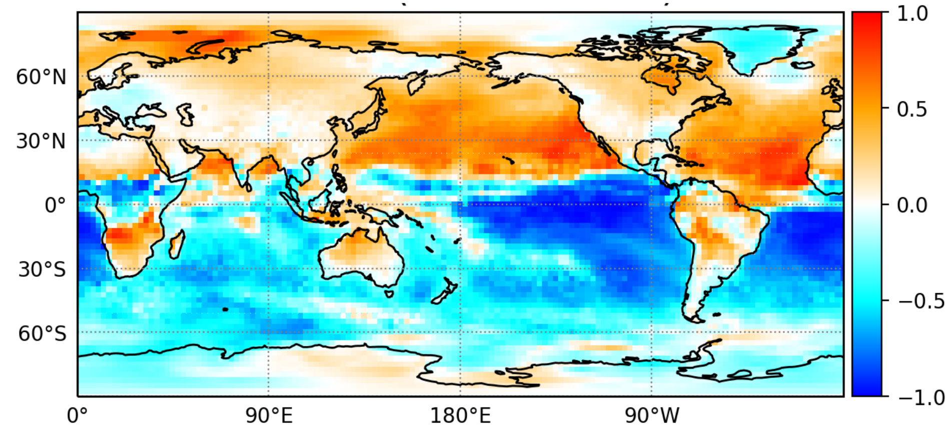



cristina.masoller@upc.edu @cristinamasoll1

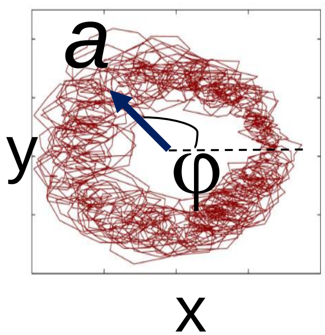
## Cosine of Hilbert phase during an *El Niño* year




## Cosine of Hilbert phase during a *La Niña* year




**Can we detect significant changes in the last decades in the amplitude of the oscillations of surface air temperature?**




In each grid point we calculate:

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

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

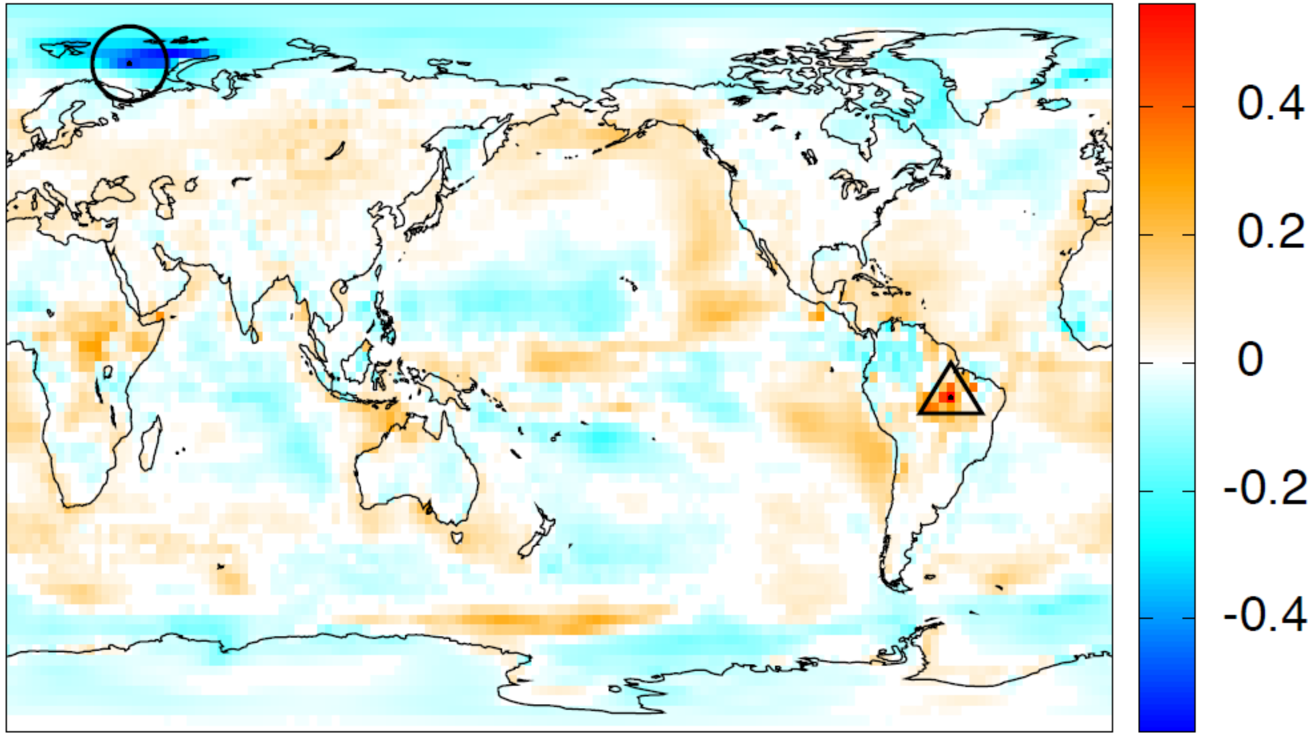
(100 surrogates)

cristina.masoller@upc.edu @cristinamasoll1

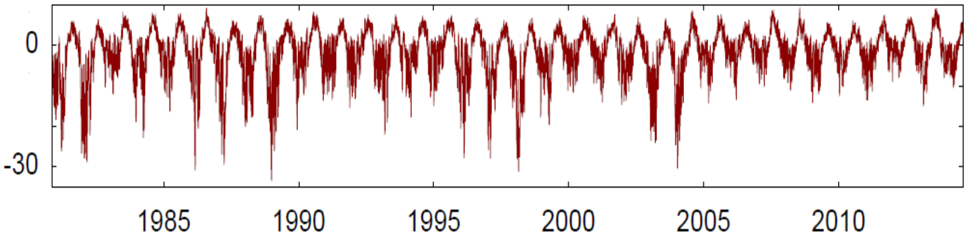
Color code:

$$\frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

White: not significant

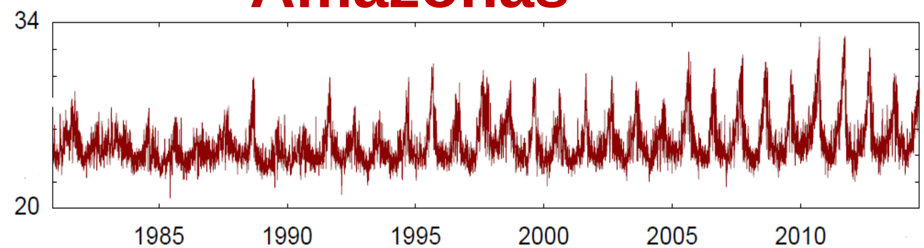



**Artic**




Sea ice melting

**Amazonas**




Deforestation

Dario A. Zappala, M. Barreiro and C. Masoller, Earth Syst. Dynam. 9, 383 (2018)



cristina.masoller@upc.edu @cristinamasoll1

# Take home messages

- Nonlinear data analysis techniques are useful to uncover patterns and relationships in data generated by complex systems.

- Different methods often provide *complementary* information.

- “Surrogate” tests are needed to analyze statistical significance.

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

- Prof. Holger Kantz: “*Every data set bears its own difficulties: data analysis is never routine*”.



cristina.masoller@upc.edu @cristinamasoll1

# Funding and references

ICREA

AGAUR

GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA E INNOVACIÓN

PGC2018-099443-B-I00
PID2021-123994NB-C21

- J. Tiana-Alsina et. al, "*Comparing the dynamics of periodically forced lasers and neurons*", New J. of Phys. 21, 103039 (2019).
- B. R. R. Boaretto, E. E. N. Macau, C. Masoller, "*Characterizing the spike timing of a chaotic laser by using ordinal analysis and machine learning*", Chaos 34, 043108 (2024).
- 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).
- G. Tirabassi and C. Masoller, "*Entropy-based early detection of critical transitions in spatial vegetation fields*", PNAS 120, e2215667120 (2023).
- D. A. Zappala et. al, "*Quantifying changes in spatial patterns of surface air temperature dynamics over several decades*", Earth Syst. Dynam. 9, 383 (2018).

## Muchas gracias por su atención!



cristina.masoller@upc.edu @cristinamasoll1