BE·LIGHT Improving biomedical diagnosis through light-based technologies and machine learning

New indicators for early detection of critical transitions

Cristina Masoller,
Departament de Física, UPC

1st Be-Light School and Mid-term meeting
October 2, 2024
Göttingen, Germany

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

ICREA

AGAUR

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

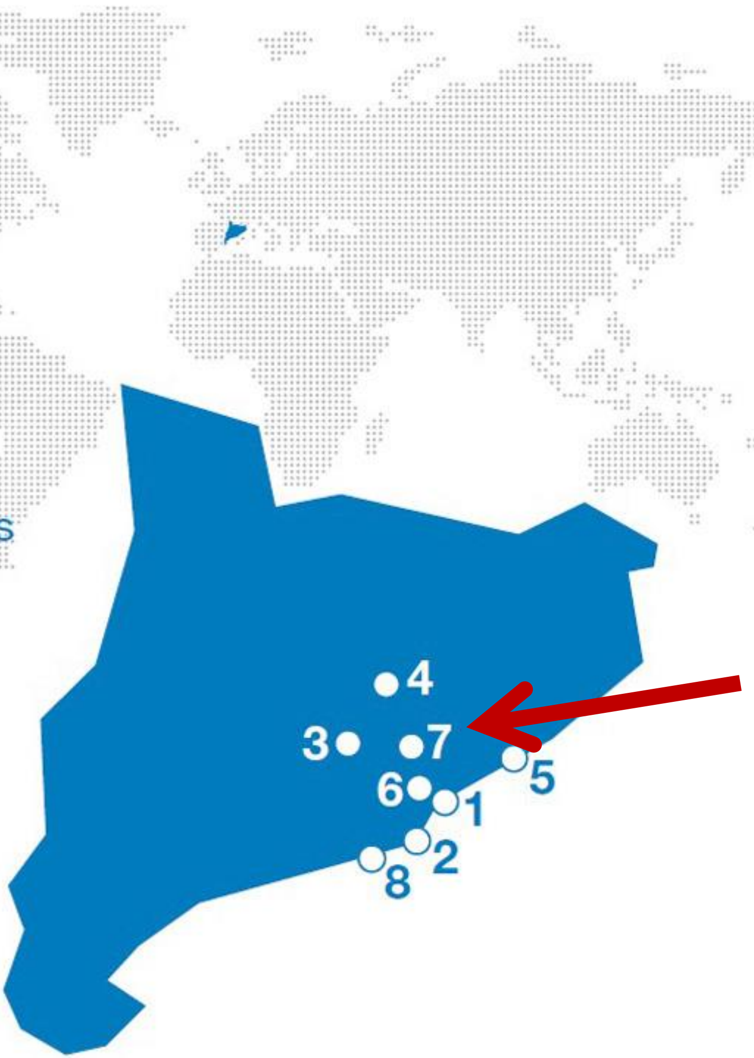
Presentation

- Bachelor & Master in Physics (1986-1991)
- PhD in Physics (1999, Bryn Mawr College, PA, USA)
- At Physics Department UPC since 2004 (Prof. Catedratica since 2018)

cristina.masoller@upc.edu @cristinamasoll1

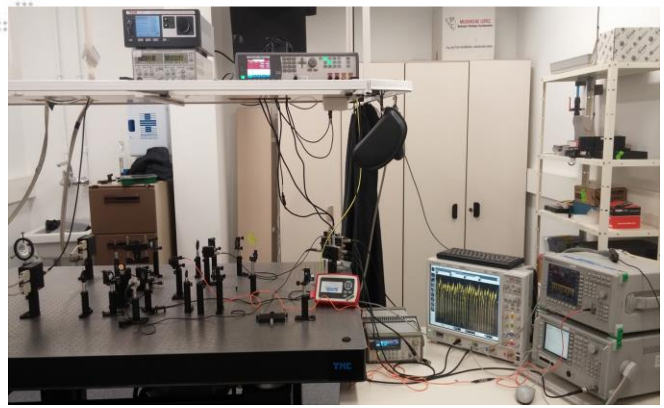
Where are we? UPC Campus Terrassa

1. Barcelona
2. Castelldefels
3. Igualada
4. Manresa
5. Mataró
6. Sant Cugat del Vallès
7. Terrassa
8. Vilanova i la Geltrú



Viernes, 25 de septiembre de 2009 Diari de Terrassa

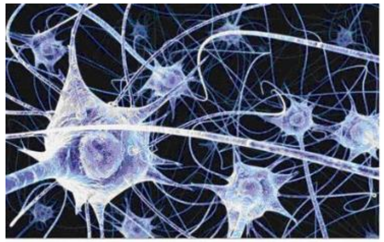
El edificio Gaia centraliza grupos científicos consolidados y emergentes.

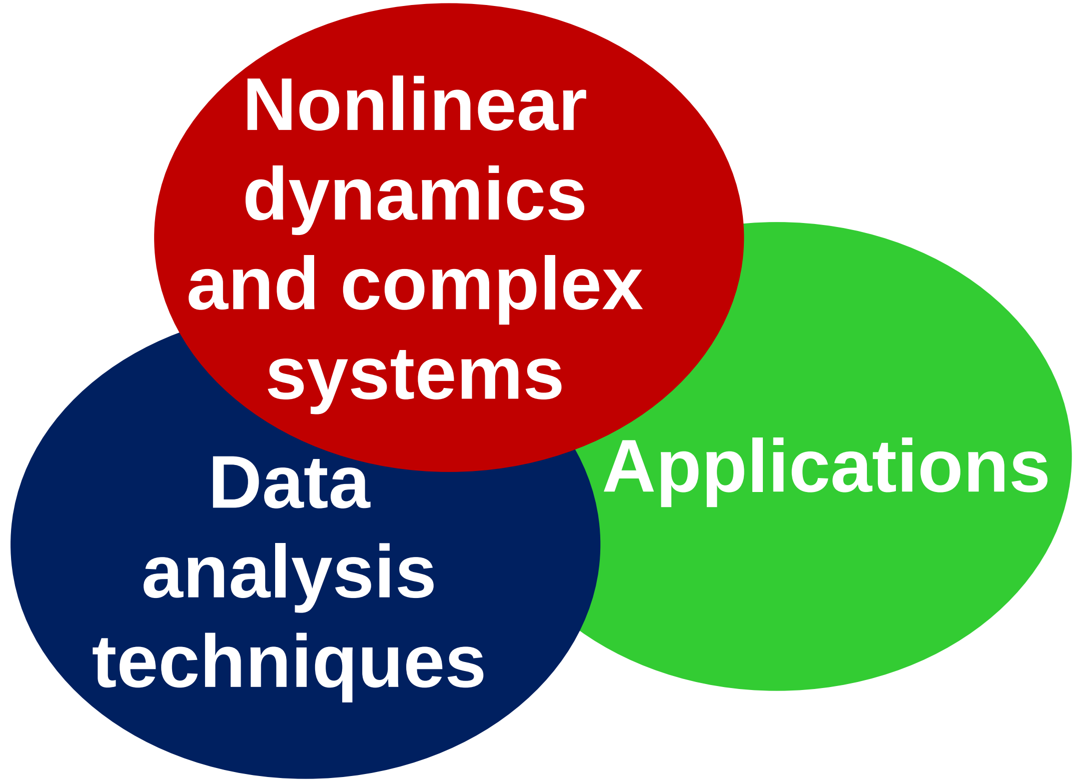


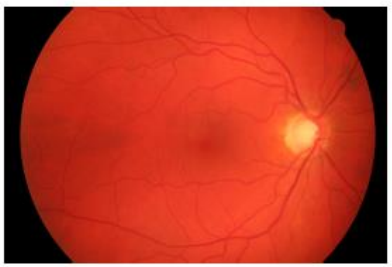
Laser lab in Gaia Building, UPC Campus Terrassa

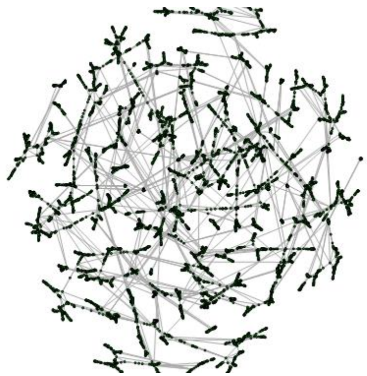
cristina.masoller@upc.edu @cristinamasoll1

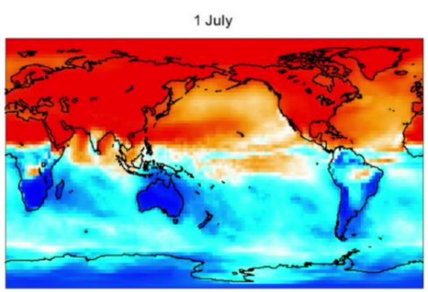
Research lines

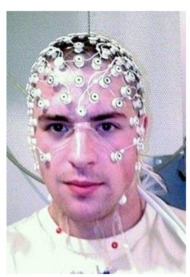


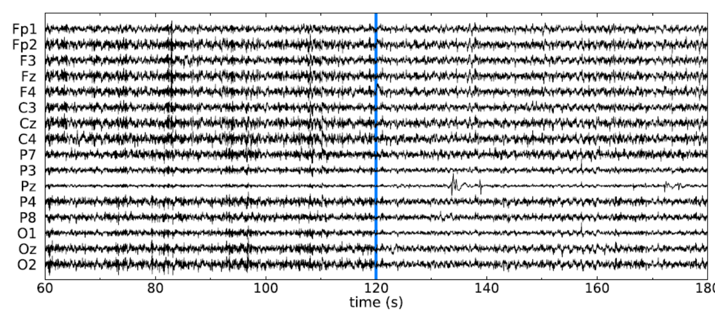


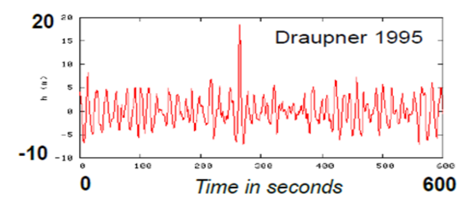












cristina.masoller@upc.edu @cristinamasoll1

Data analysis method: ordinal analysis

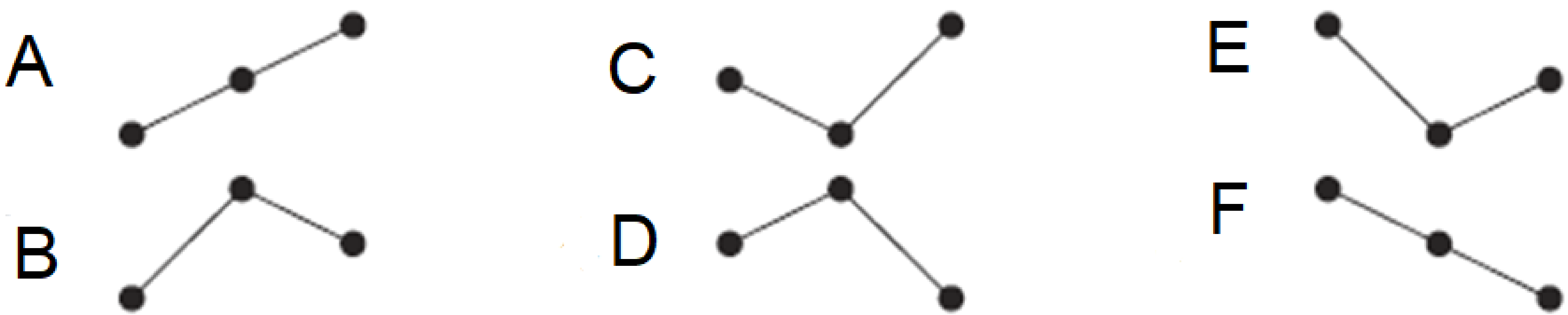
$$\{\ldots x_i,\ x_{i+1},\ x_{i+2},\ \ldots\}$$

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

{…2, 5, 7…}

{…5, 2, 7…}

{…7, 2, 5 …}



{…2, 7, 5…}

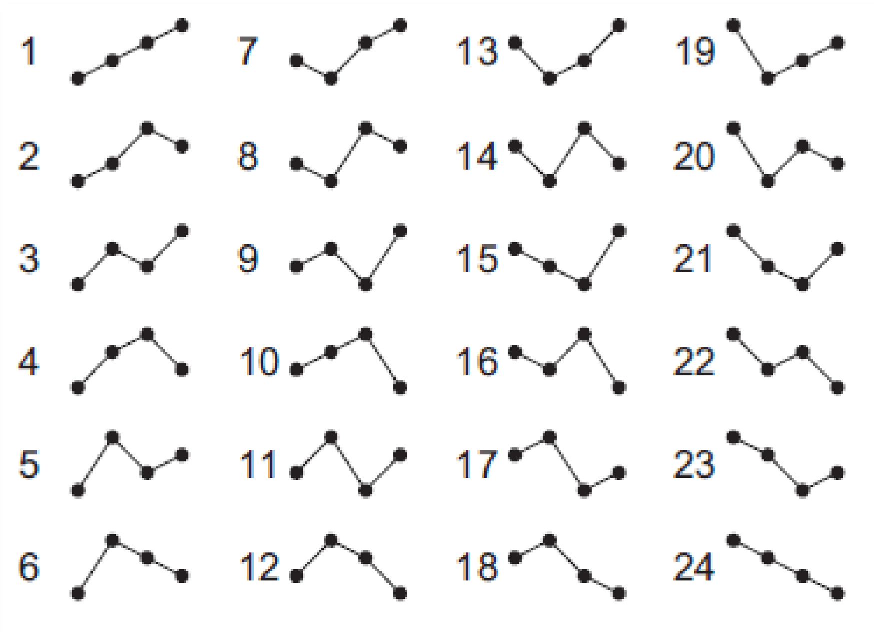
{…5, 7, 2…}

{…7, 5, 2…}

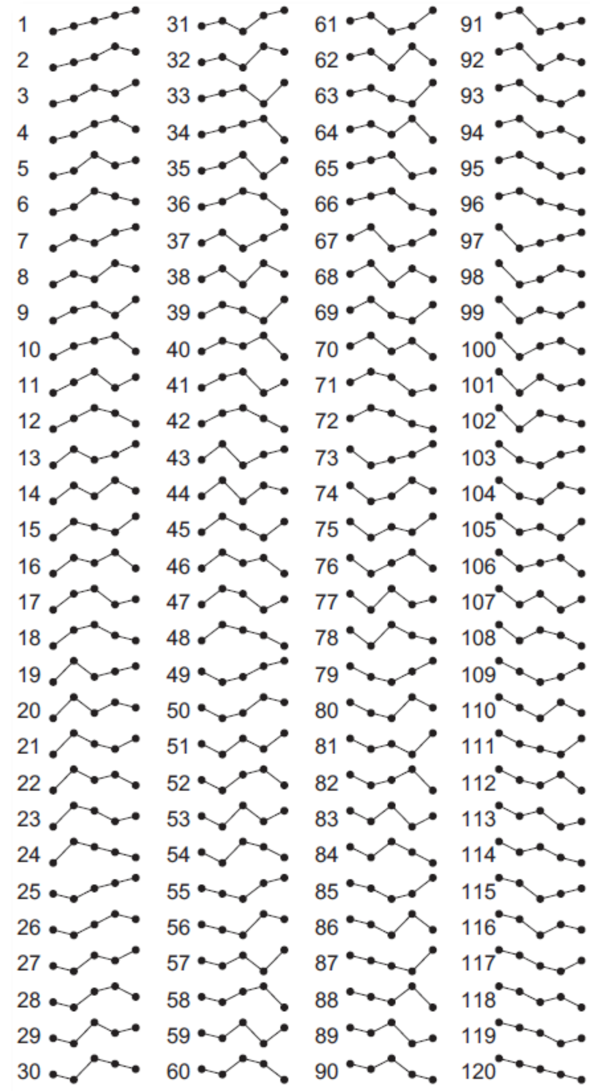
Bandt and Pompe: Phys. Rev. Lett. 2002

cristina.masoller@upc.edu @cristinamasoll1

The number of ordinal patterns increases as D!



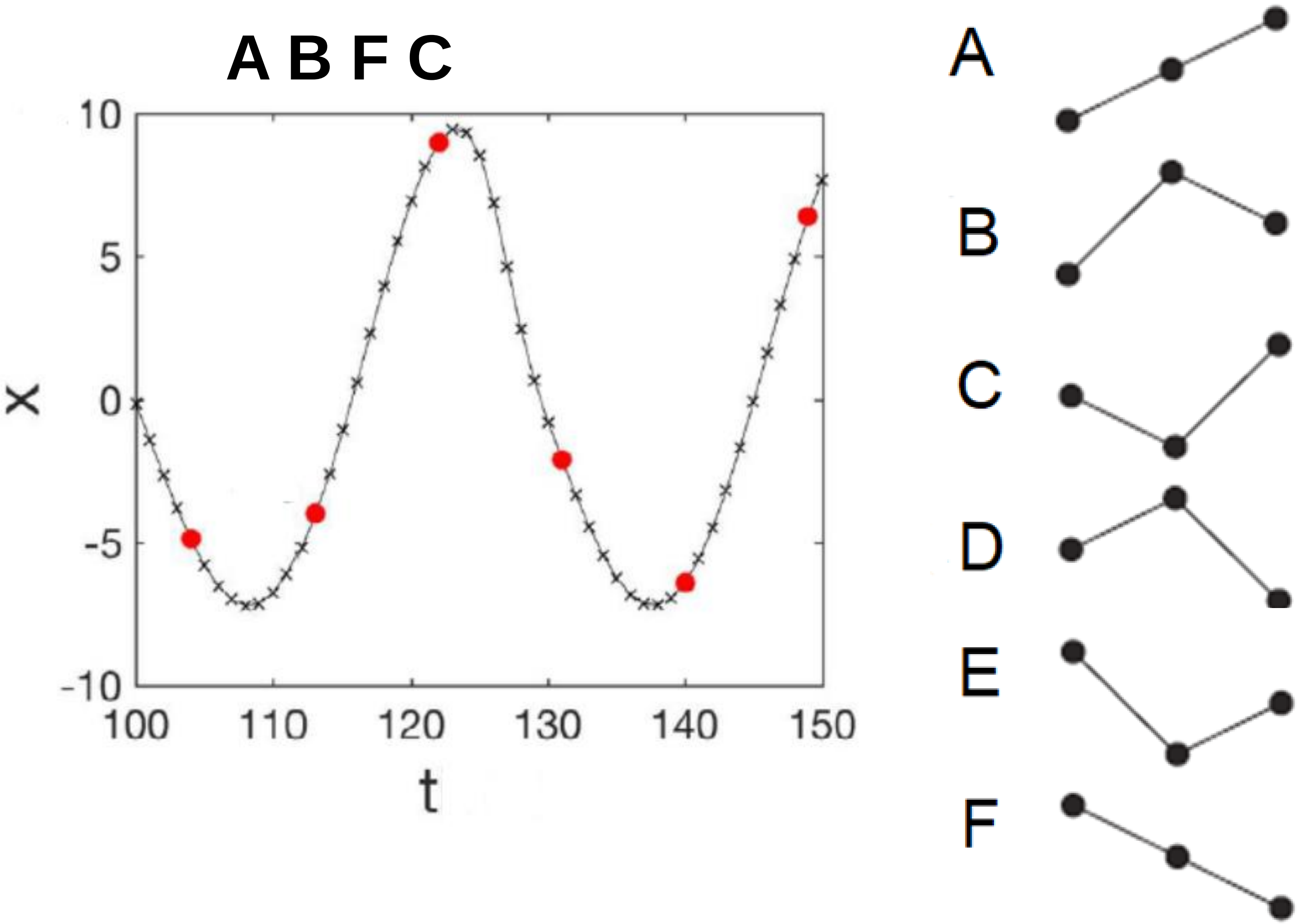
A problem for short datasets.



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

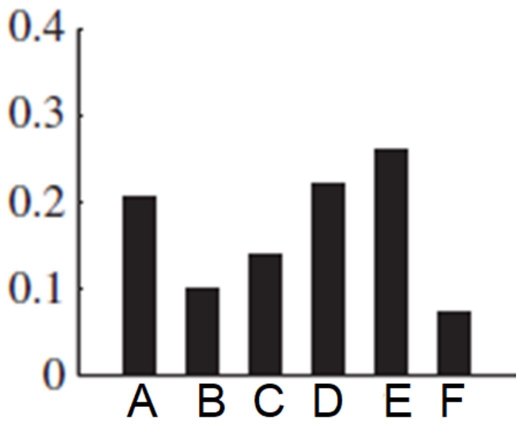
cristina.masoller@upc.edu @cristinamasoll1

Using the "ordinal code", which is the message?



cristina.masoller@upc.edu @cristinamasoll1

From the frequency of occurrence of the patterns, we calculate the “ordinal probabilities”



?

A. Analyze the probability values; use them as features for ML algorithms
B. Analyze “information theory measures” (e.g. entropy)— a form of nonlinear dimensionality reduction.

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

Robust to outliers.

Ordinal analysis has been extensively used:

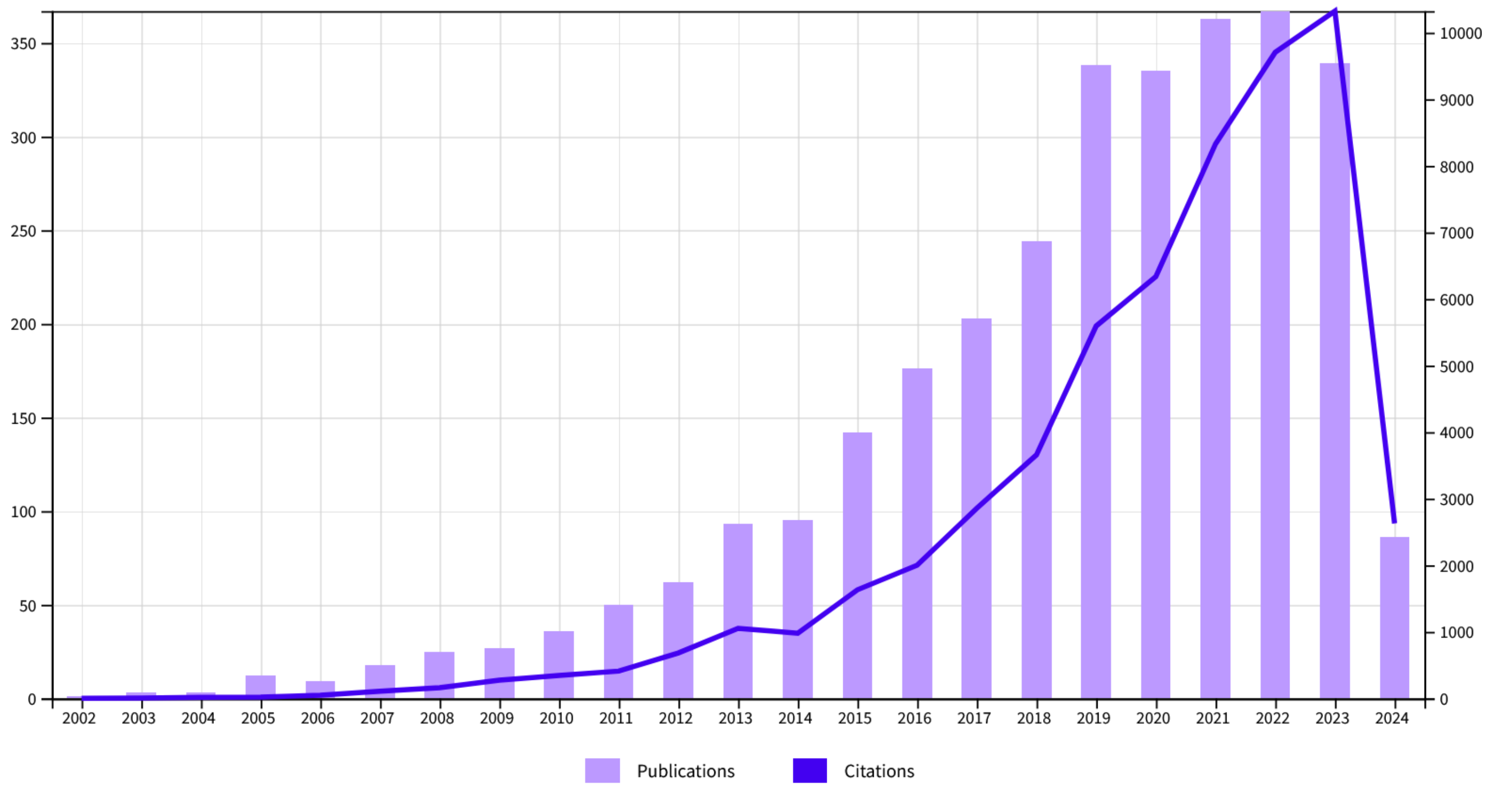
- to test if a model is good for the data,
- to fit the model’s parameters,
- to classify different types of data based on similarities of probabilities of ordinal patterns.

cristina.masoller@upc.edu @cristinamasoll1

VOLUME 88, NUMBER 17 PHYSICAL REVIEW LETTERS 29 APRIL 2002

Permutation Entropy: A Natural Complexity Measure for Time Series

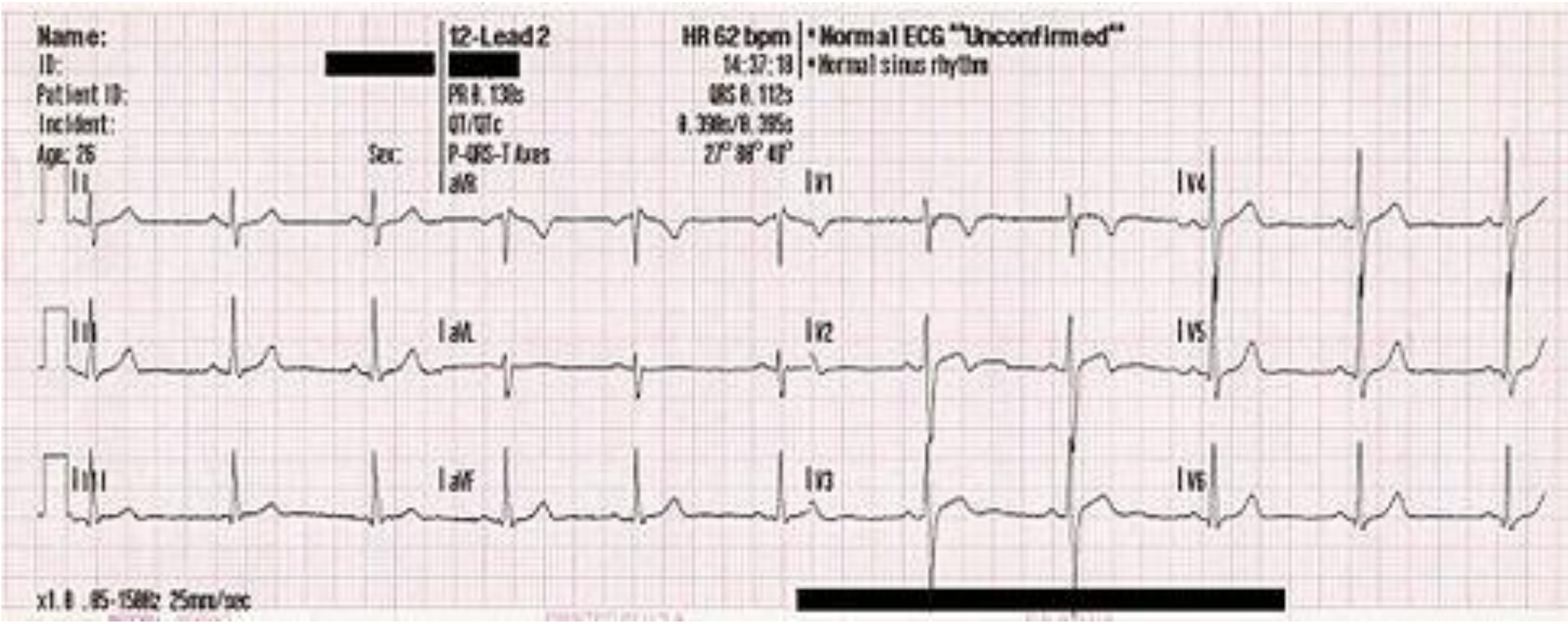
Christoph Bandt and Bernd Pompe
Institute of Mathematics and Institute of Physics, University of Greifswald, Greifswald, Germany
(Received 19 June 2001; revised manuscript received 20 December 2001; published 11 April 2002)



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

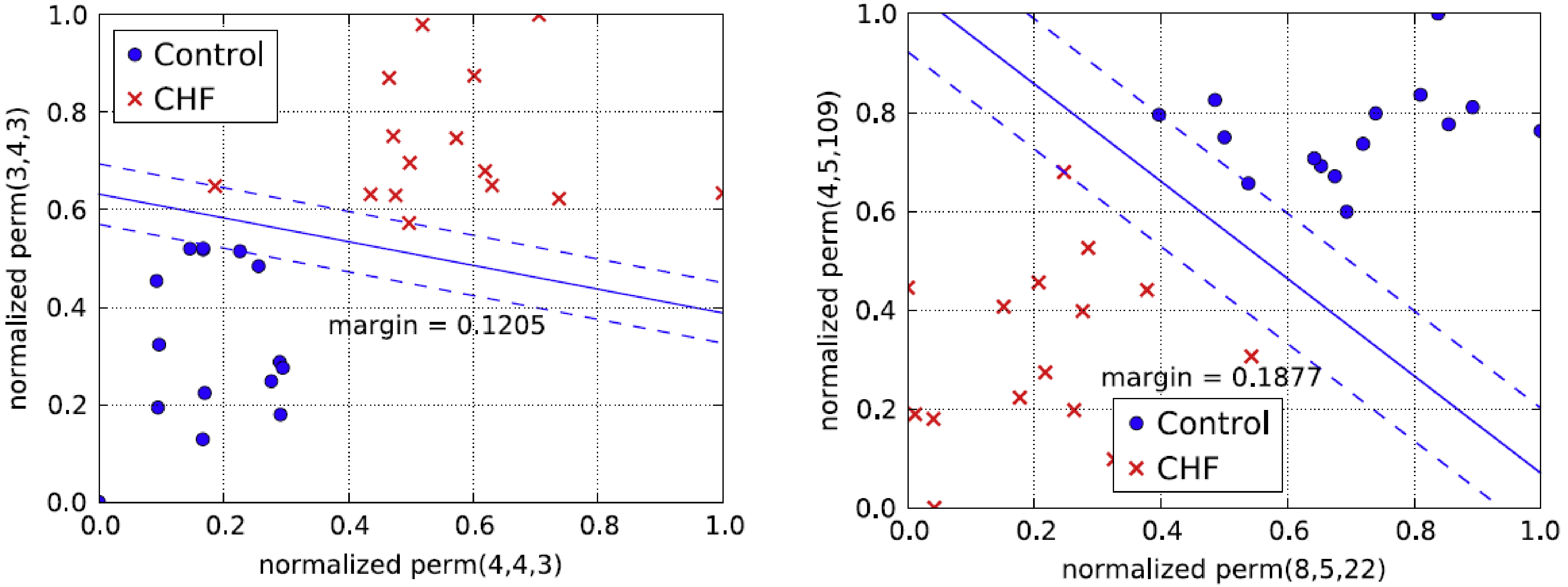
cristina.masoller@upc.edu @cristinamasoll1

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 anomalies - outliers)
- The probabilities are normalized with respect to the smallest and largest values occurring in the data set.

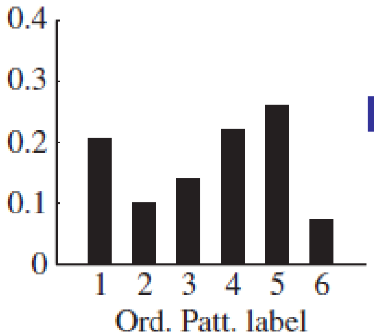
U. Parlitz et al. Computers in Biology and Medicine 42, 319 (2012)

cristina.masoller@upc.edu @cristinamasoll1

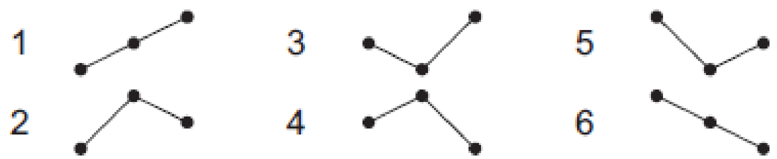
Permutation entropy: Shannon's entropy computed from ordinal probabilities

Time series

Ordinal probabilities



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



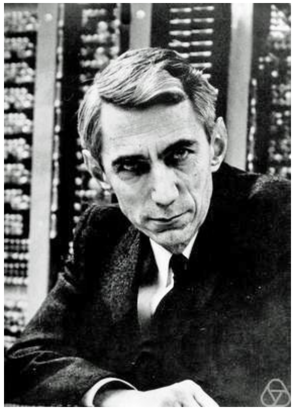
Permutation entropy

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

cristina.masoller@upc.edu @cristinamasoll1

Shannon entropy

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



Claude Shannon

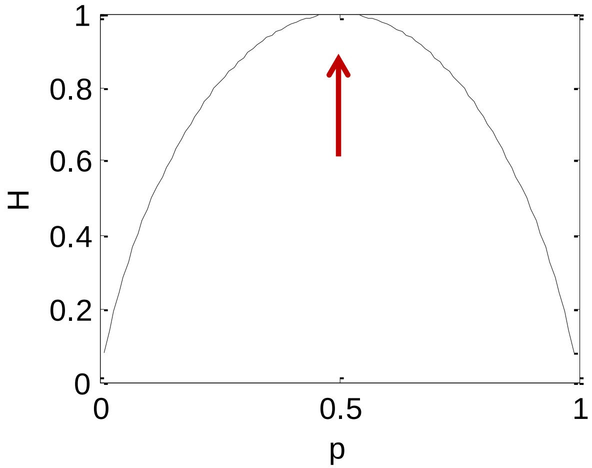
- Interpretation: "*quantity of* ***surprise*** *one should feel upon reading the result of a measurement* ".

- 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).

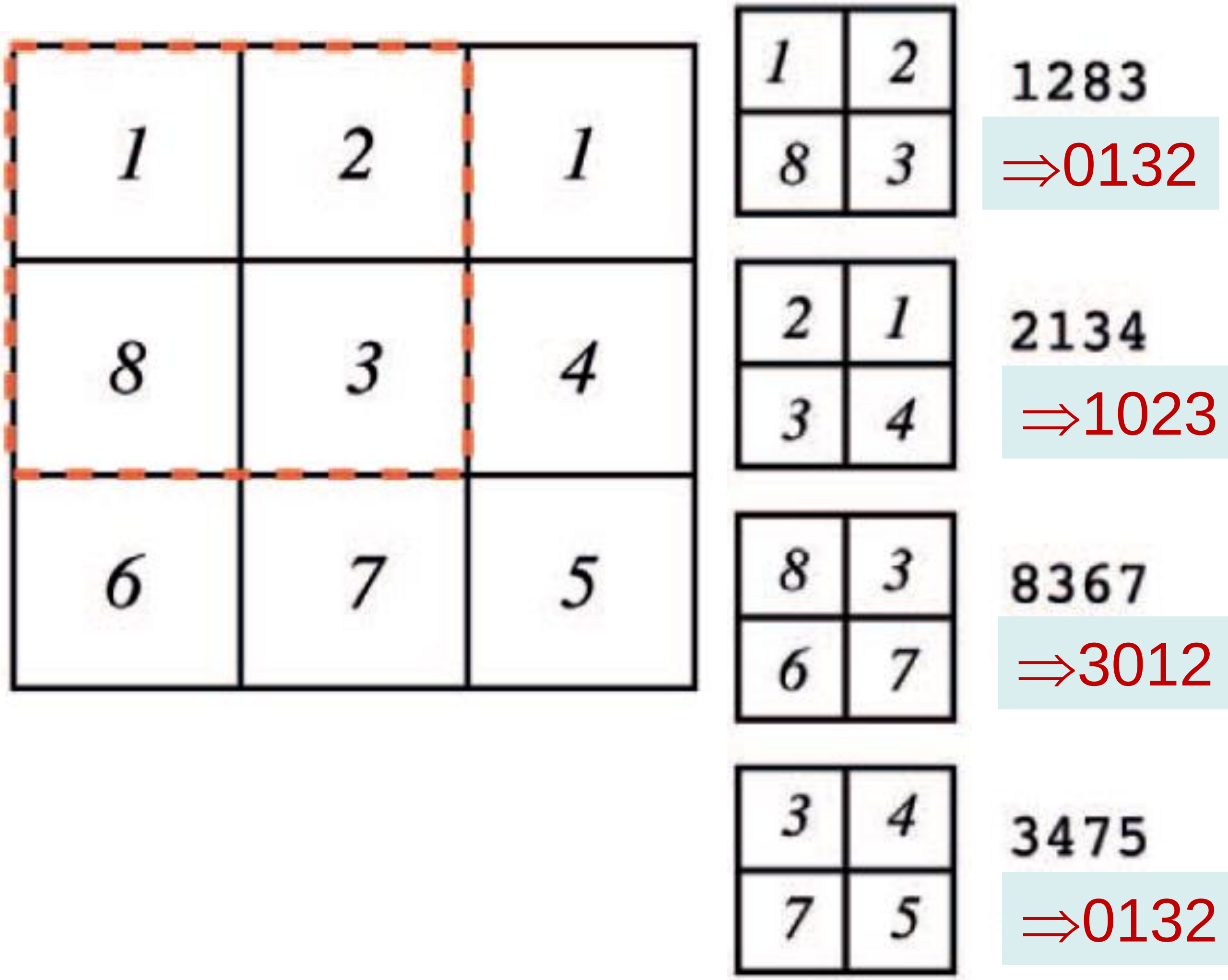
$\Rightarrow$ p=0.5: Maximum **unpredictability.**



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

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).

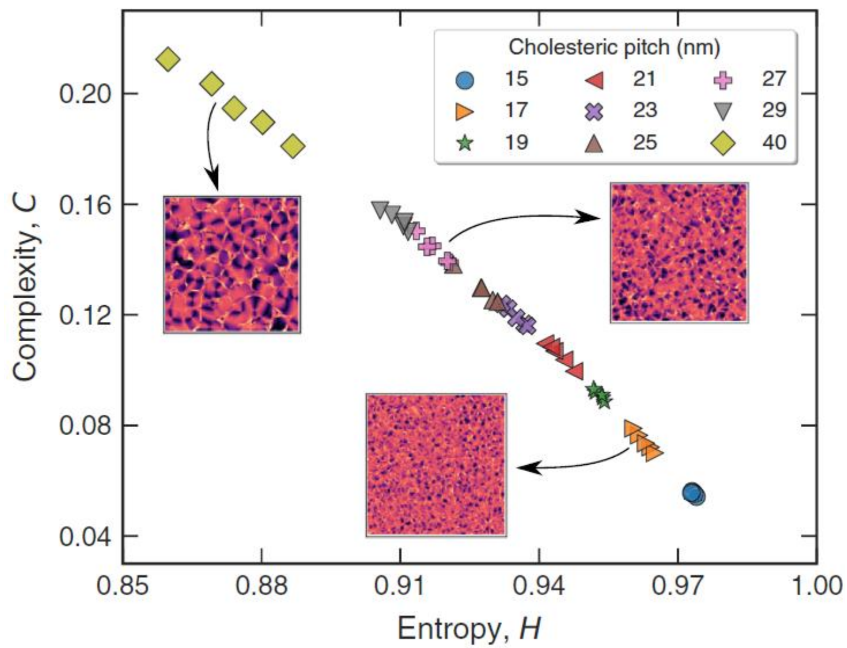
cristina.masoller@upc.edu @cristinamasoll1

The “spatial” permutation entropy has been used 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

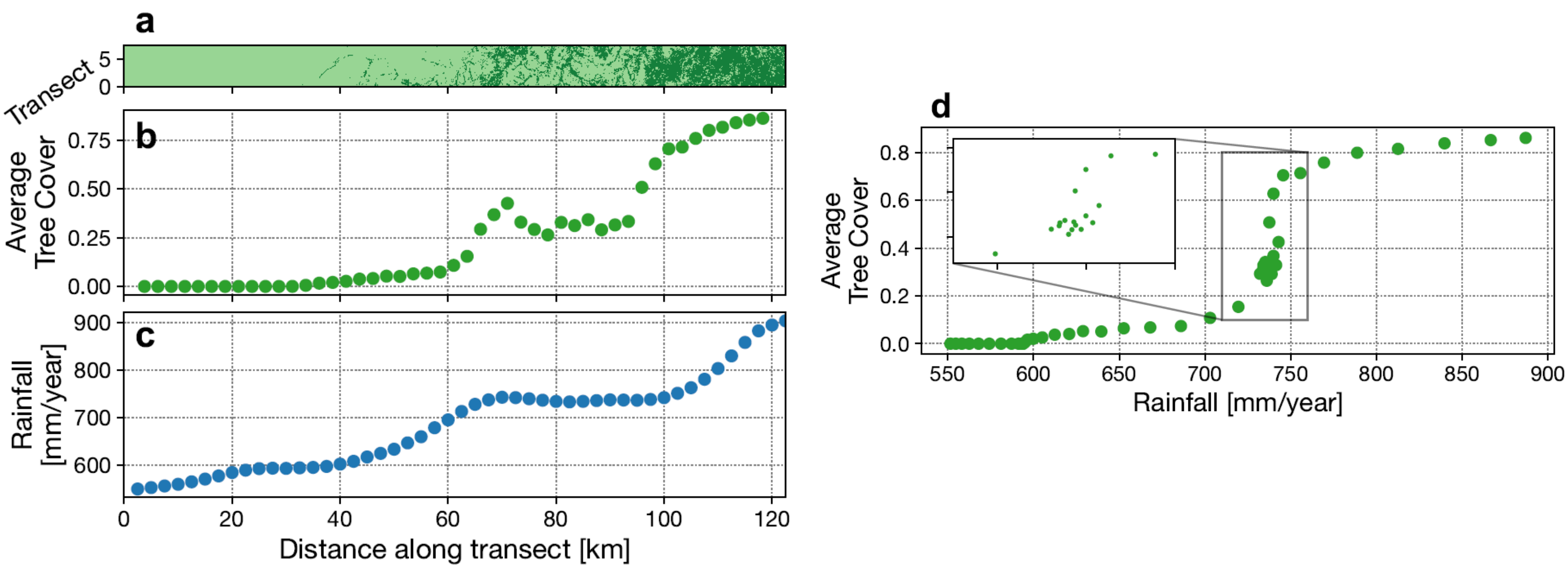
H. Y. D. Sigaki,[1] R. F. de Souza,[1] R. T. de Souza,[1,2] R. S. Zola,[1,2,*] and H. V. Ribeiro[1,†]
[1]*Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil*
[2]*Departamento de Física, Universidade Tecnológica Federal do Paraná, Apucarana, PR 86812-460, Brazil*



cristina.masoller@upc.edu @cristinamasoll1

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

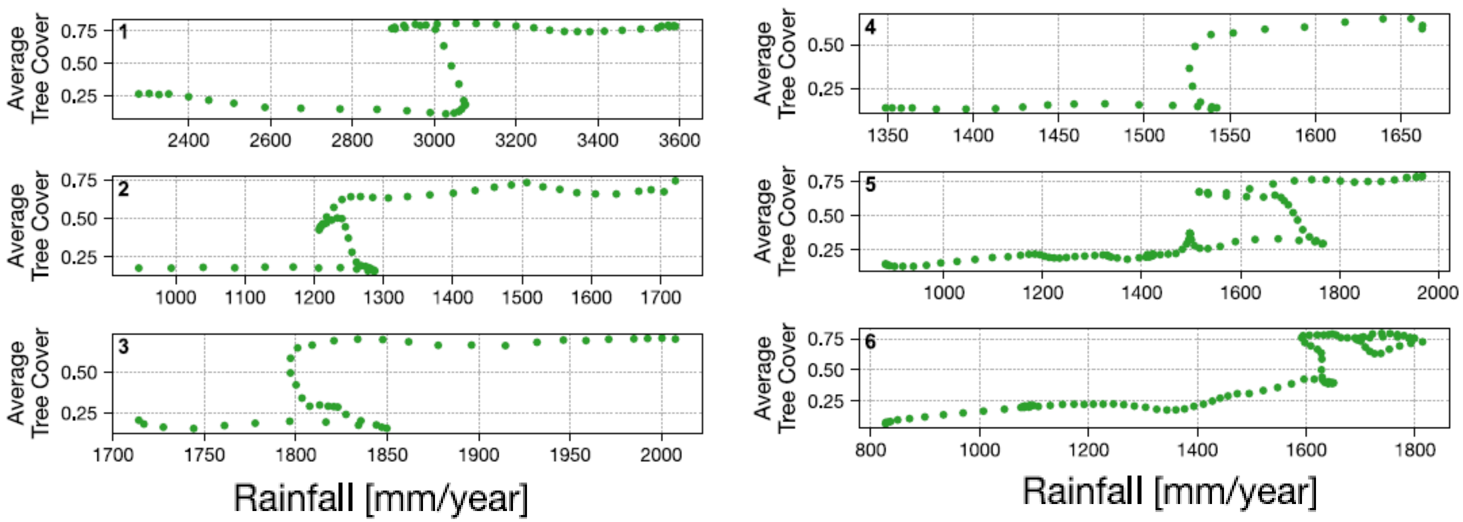
High-resolution vegetation data from the Serengeti–Mara ecosystem in northern Tanzania and southern Kenya.



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)

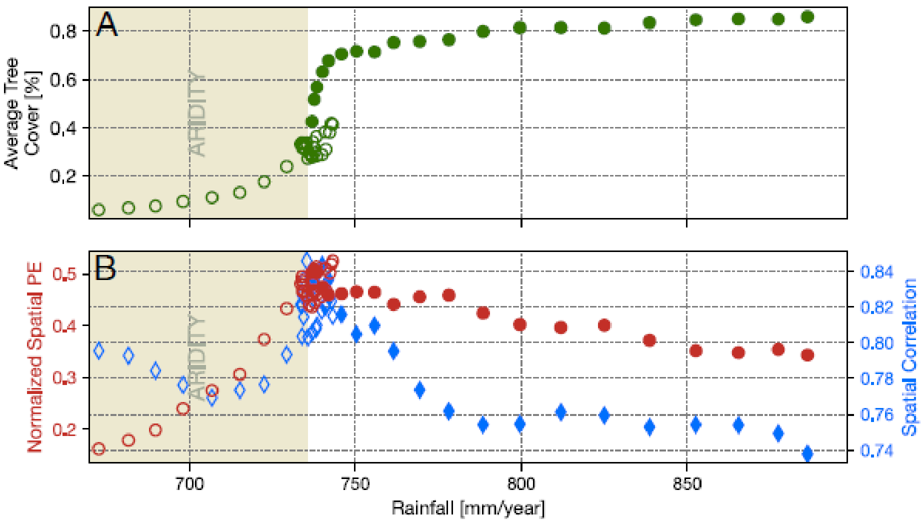
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

Spatial correlation

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

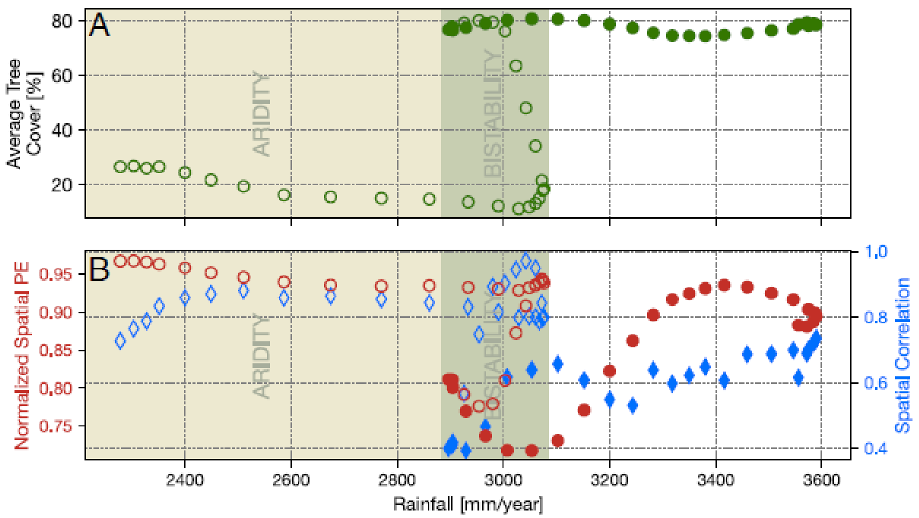
w_{ij}=1 if i, j first neighbors, else 0

High-resolution data



Low-resolution data

(transect 1; large variability across transects)

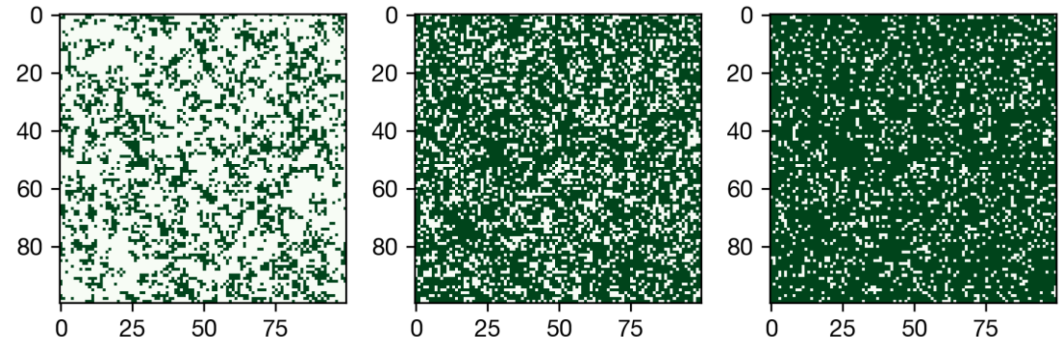


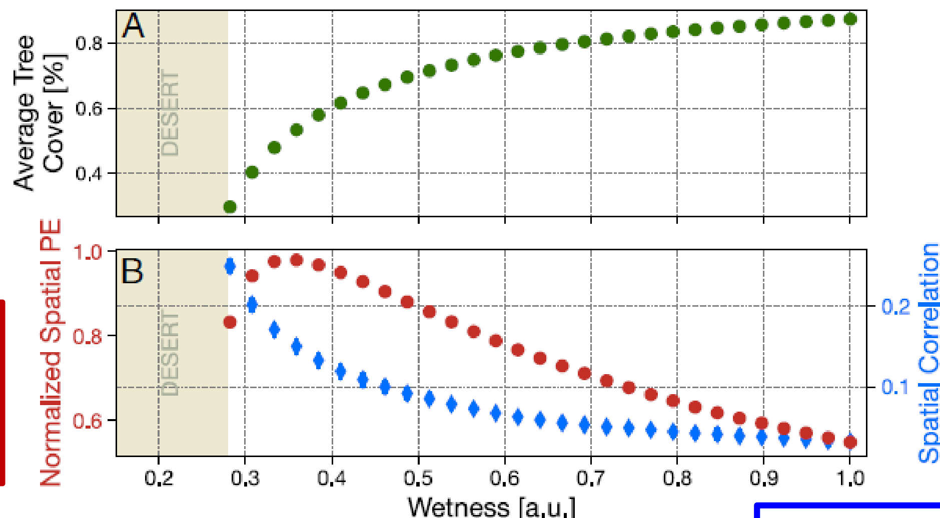
G. Tirabassi, C. Masoller, "Entropy-based early detection of critical transitions in spatial vegetation fields", PNAS 120, e2215667120 (2023).

cristina.masoller@upc.edu @cristinamasoll1

To gain insight: simulations of vegetation models

A) Cellular automata model





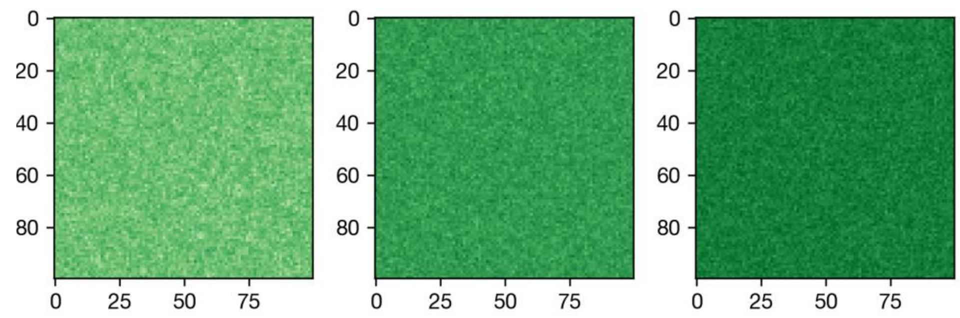
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

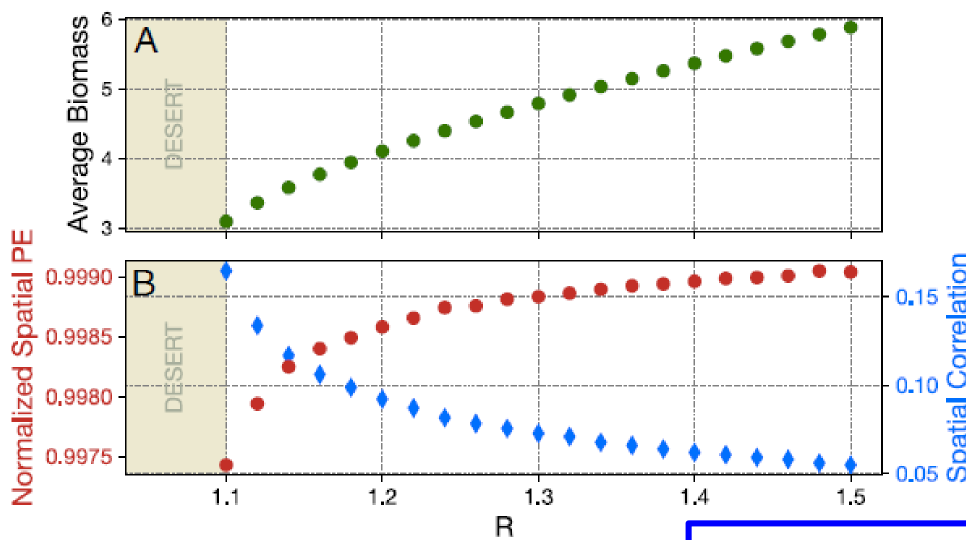
$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_i (u_i - \bar{u})^2}$$

cristina.masoller@upc.edu @cristinamasoll1

To gain insight: simulations of vegetation models

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





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

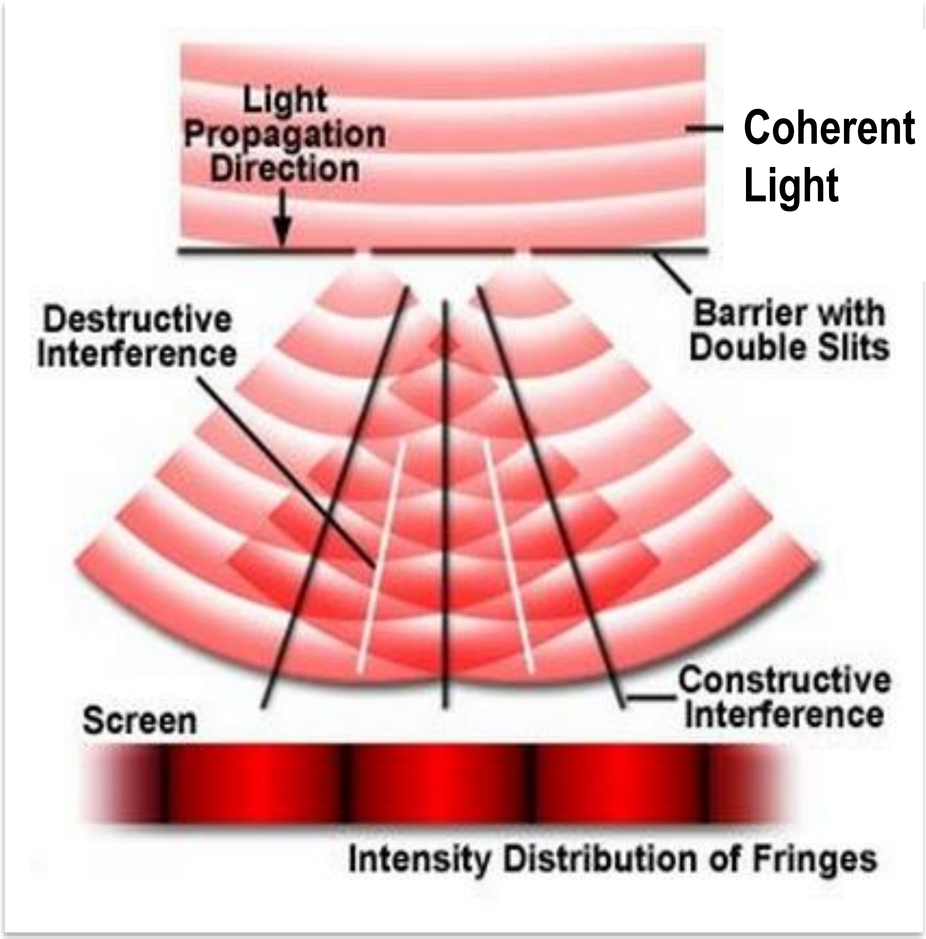
$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_i (u_i - \bar{u})^2}$$

cristina.masoller@upc.edu @cristinamasoll1

Diode laser experiments

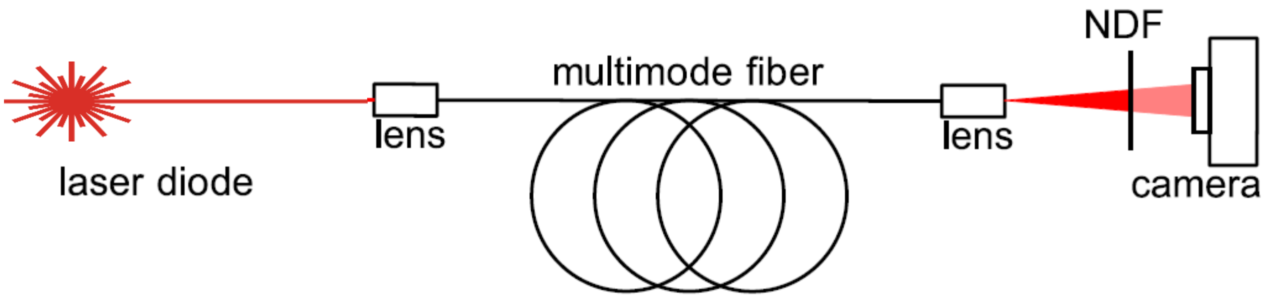
Transition from low-coherence emission (stochastic quantum spontaneous emission) to coherent emission (laser turn-on stimulated emission).

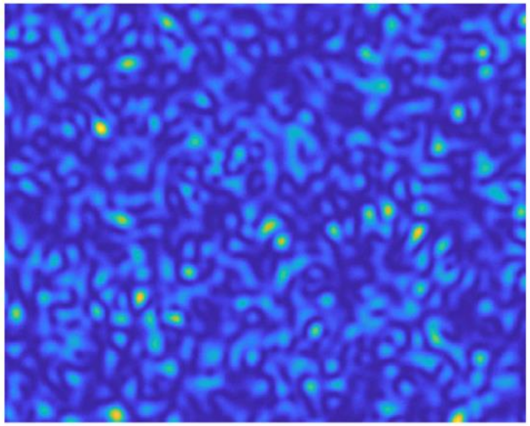
Quick review on the interference of coherent waves



cristina.masoller@upc.edu @cristinamasoll1

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

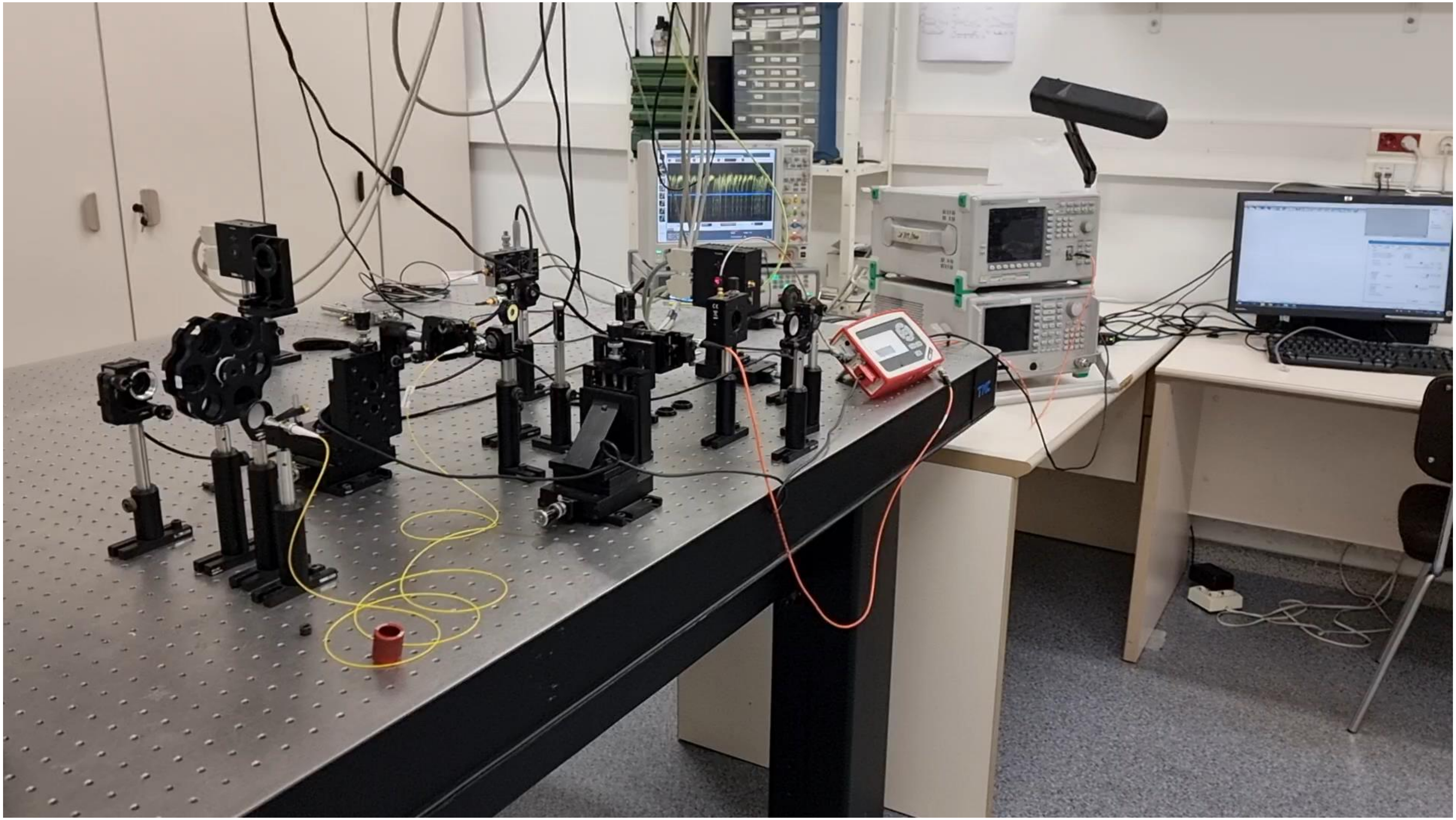




Many applications. Two main types
- Extract information of the light (wavemeters)
- Extract information of the medium that generates the speckle (speckle-based spectroscopy)

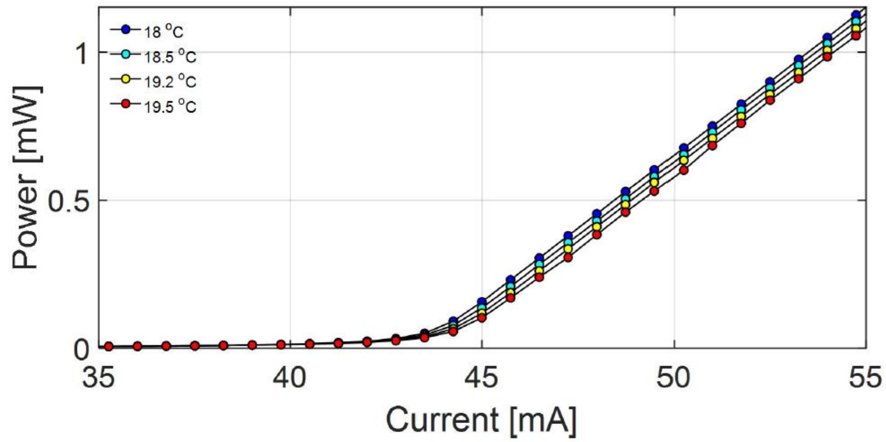
But
Speckle is a drawback in laser-based illumination and imaging application.

cristina.masoller@upc.edu @cristinamasoll1

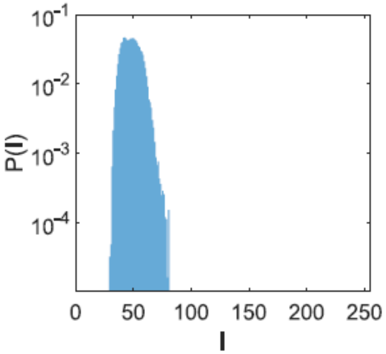


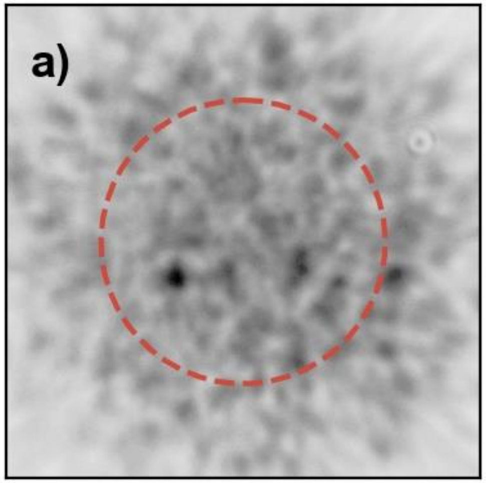
cristina.masoller@upc.edu @cristinamasoll1

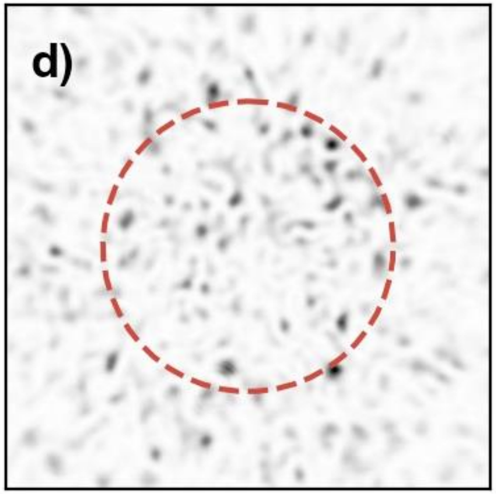
Analysis of Speckle Patterns using Permutation Entropy

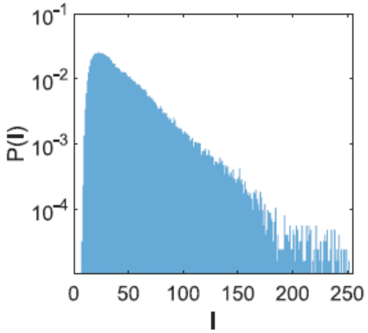


Below threshold **Above threshold**









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

cristina.masoller@upc.edu @cristinamasoll1

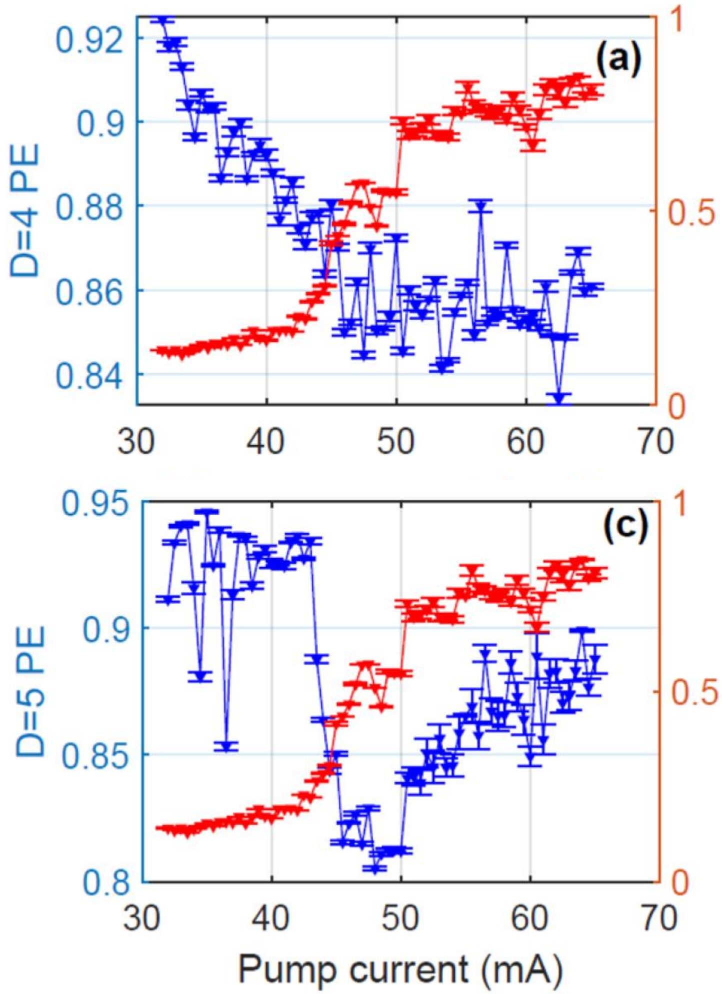
Results

Pattern:
x x
x x

SC = σ/⟨I⟩

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

Pattern: x
x x x
x



Speckle Contrast

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

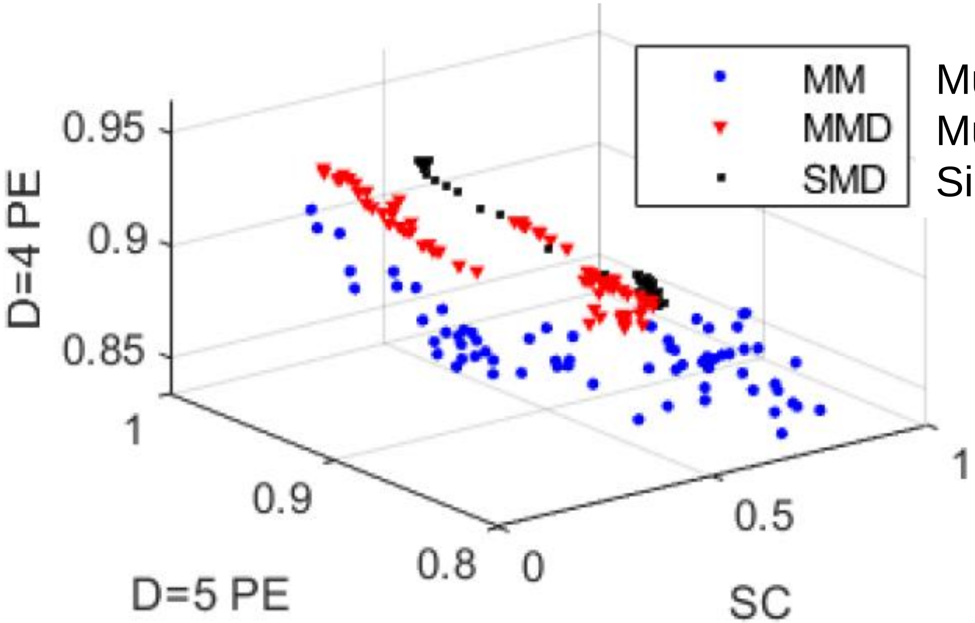
Three features allow to differentiate the speckle patterns according to the type of medium that generated the speckles

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

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

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

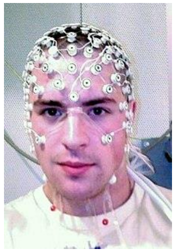
Solitary laser: 99.4 % ± 0.4 %

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).

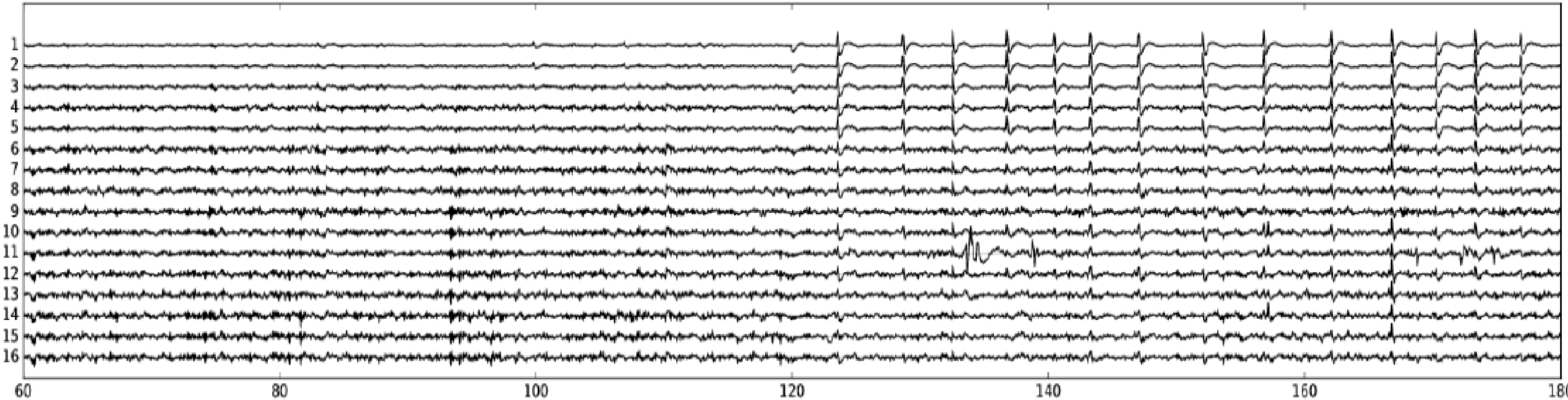
cristina.masoller@upc.edu @cristinamasoll1

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

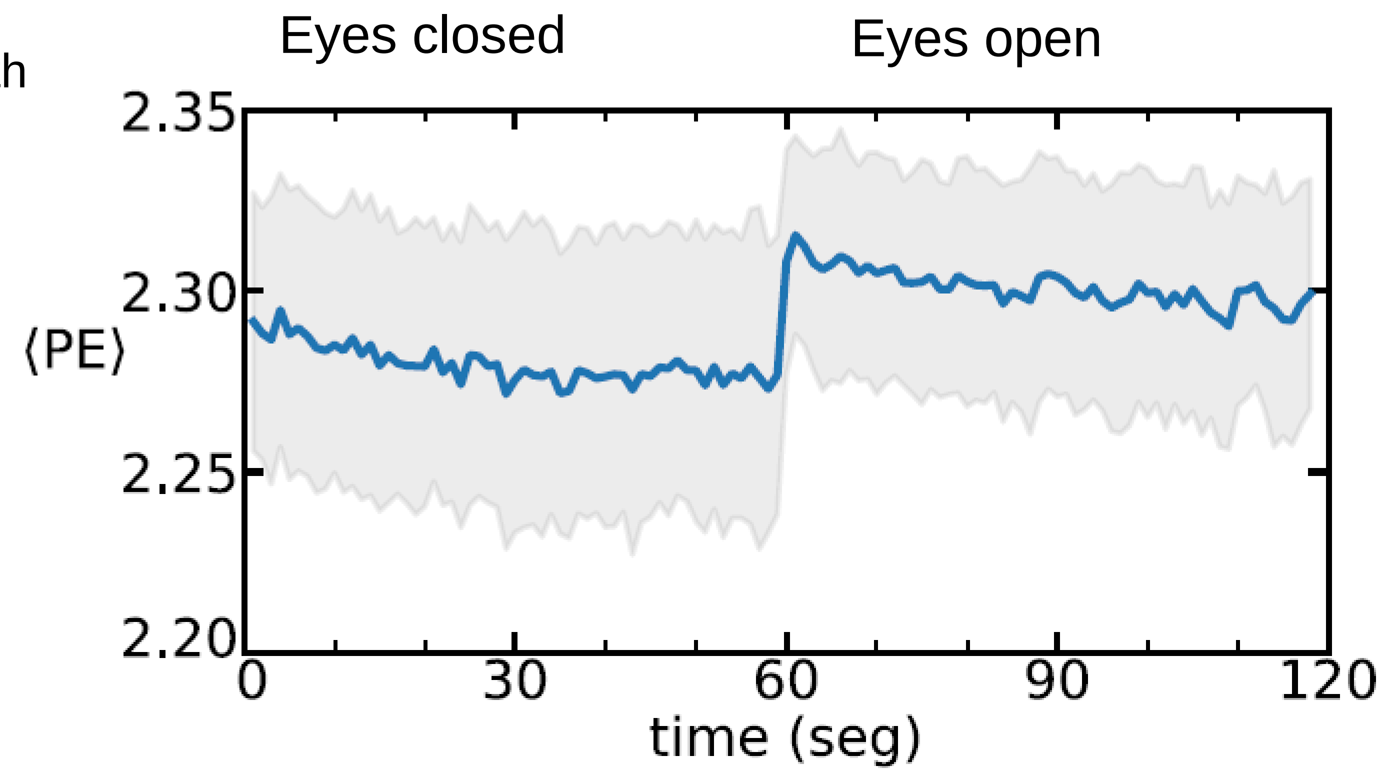
cristina.masoller@upc.edu @cristinamasoll1

The Permutation Entropy increases in the eyes open state

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

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

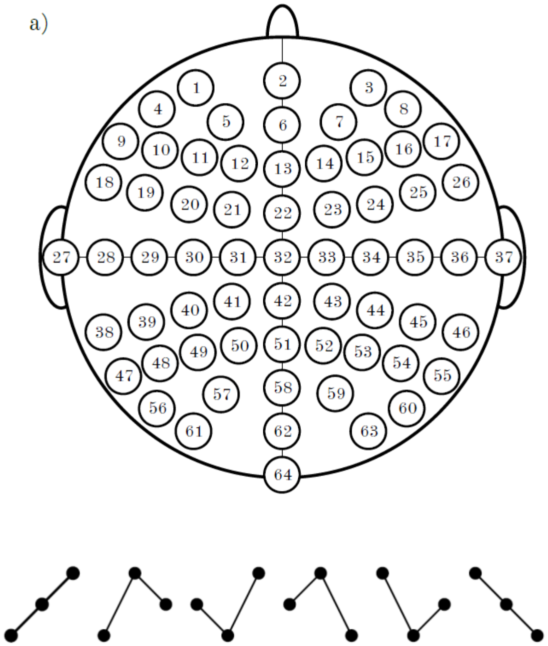
Gray region: σ of ⟨PE⟩ values across subjects

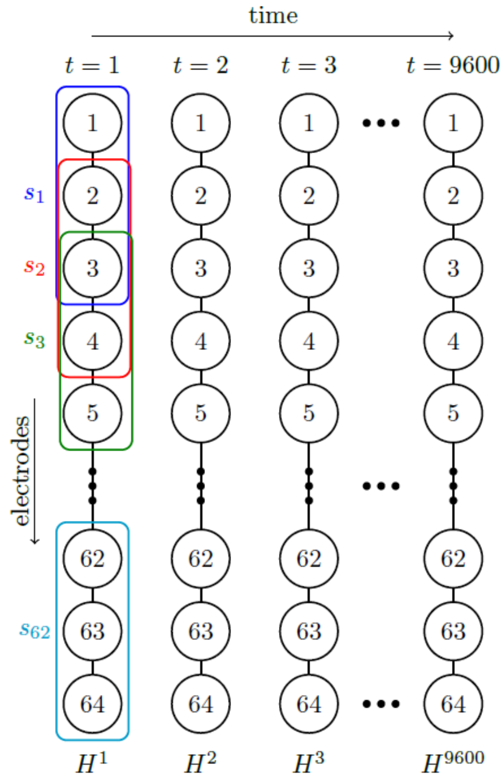


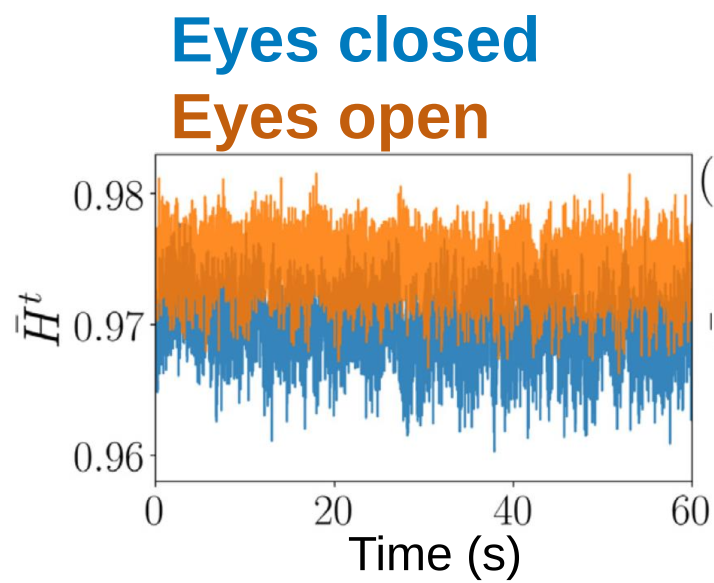
C. Quintero-Quiroz et al., "Differentiating resting brain states using ordinal symbolic analysis", Chaos 28, 106307 (2018).

cristina.masoller@upc.edu @cristinamasoll1

Spatial approach to compute the Permutation Entropy







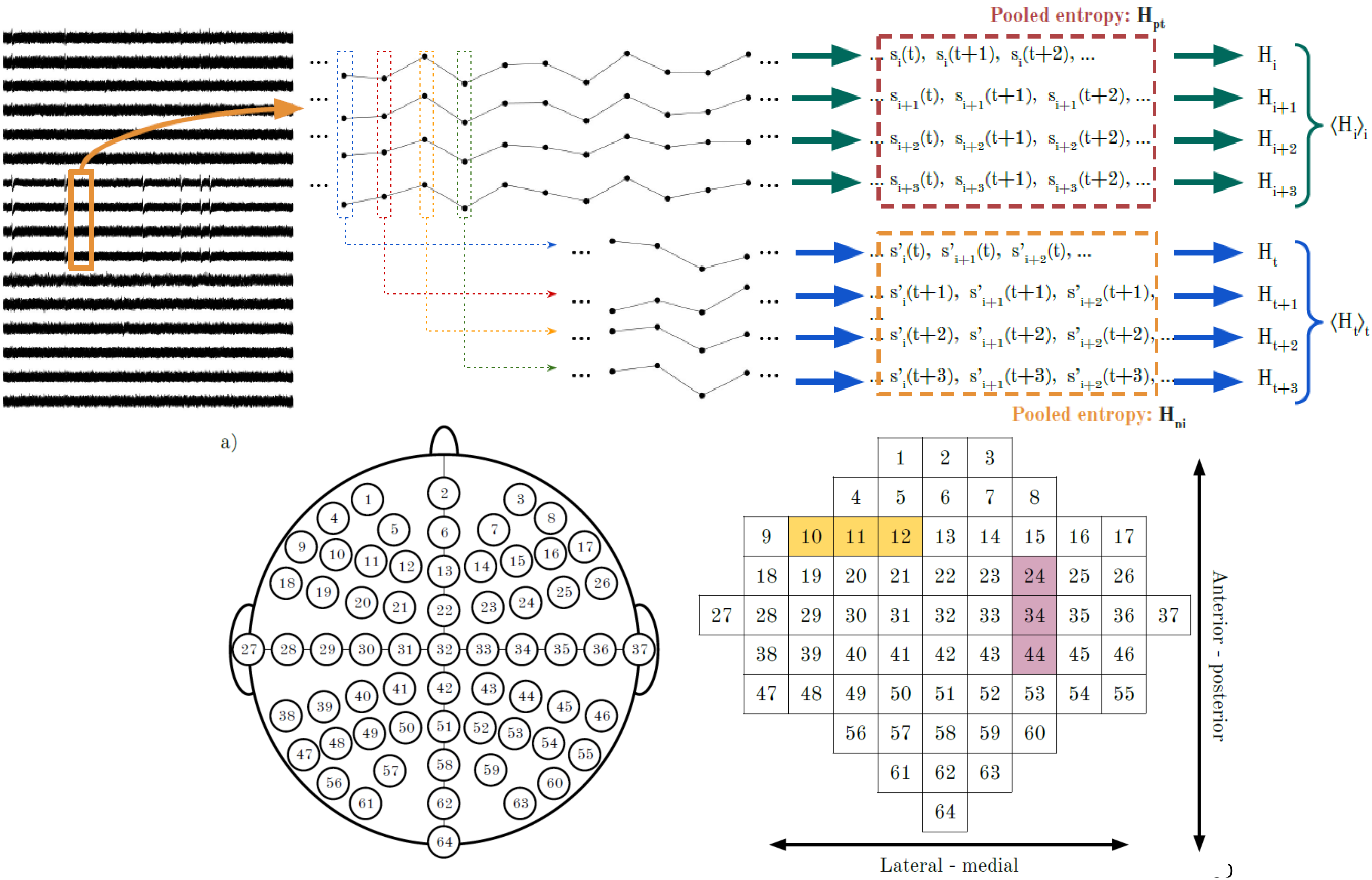
(H_t is averaged over subjects)

At each time: data values of 64 channels $\Rightarrow$ 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

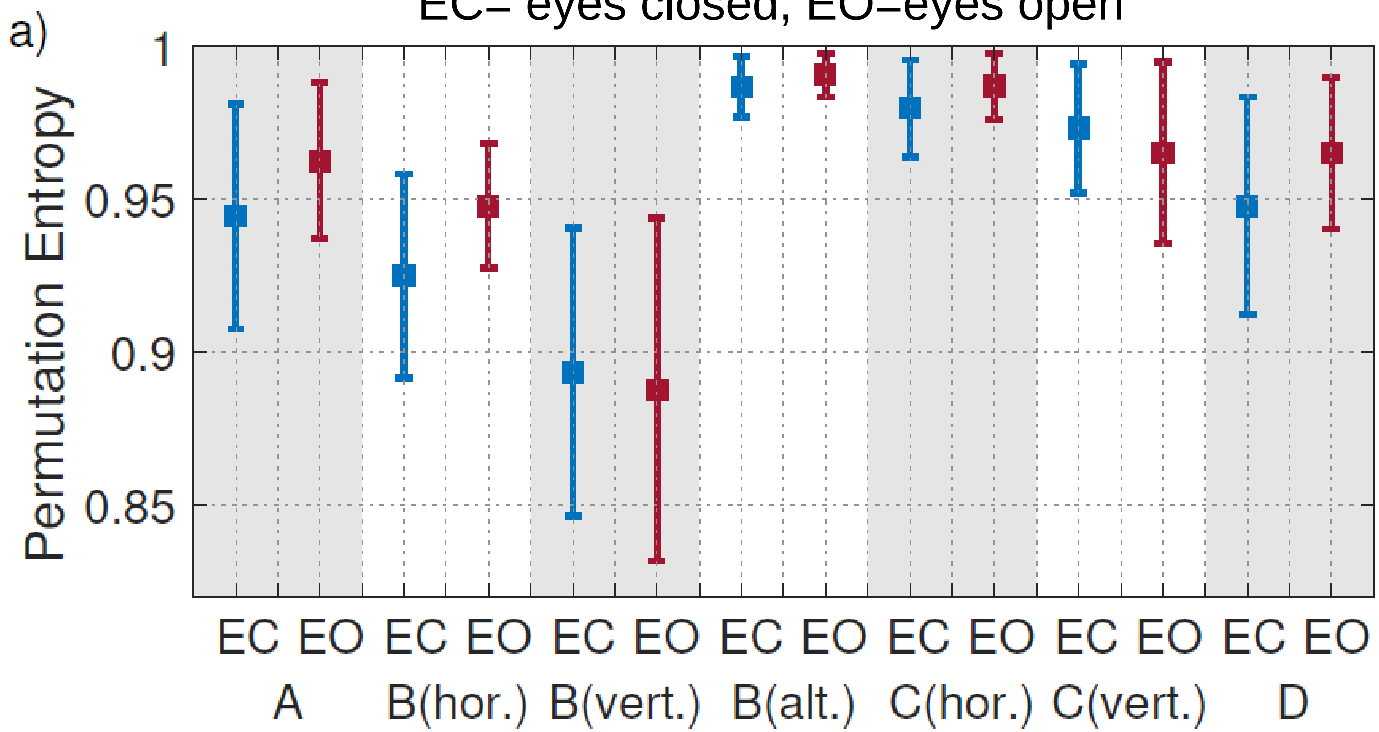
Four approaches to calculate the permutation entropy



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

Take home messages

- 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.

- "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

- G. Tirabassi and C. Masoller, "*Entropy-based early detection of critical transitions in spatial vegetation fields*", PNAS 120, e2215667120 (2022).
- 197. G. Tirabassi, M. Duque-Gijon, J. Tiana-Alsina, C. Masoller, *"Permutation entropy-based characterization of speckle patterns generated by semiconductor laser light*", APL Photonics 8, 126112 (2023).
- 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).

Thank you for your attention!

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