Data-driven approach for identifying regime transitions

Cristina Masoller
Universitat Politecnica de Catalunya, Physics Department
Terrassa, Barcelona, Spain

Cristina.masoller@upc.edu

www.fisica.edu.uy/~cris

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

Ibersinc II Meeting
Madrid, April 2017

ICREA

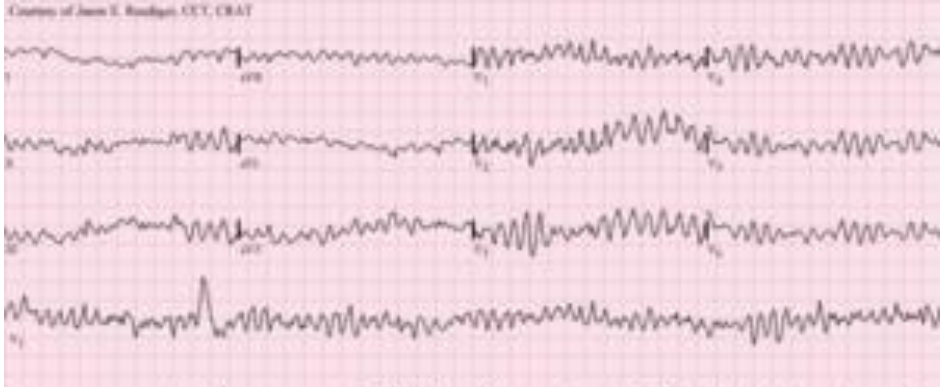
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

Dangerous regime transitions

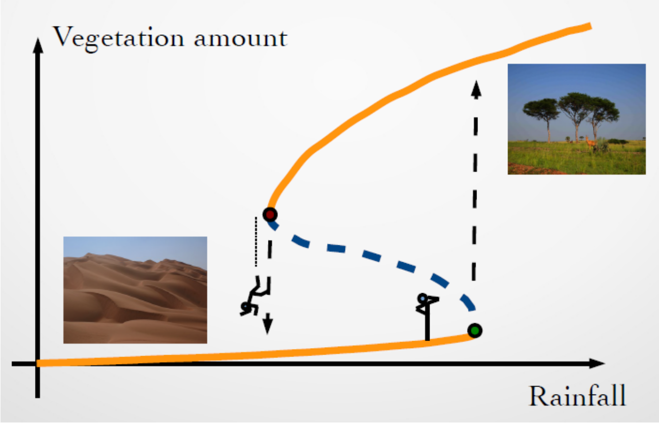
Electroencephalographs - EEGs

Source: www.epilepsysociety.org.uk

Cardiac arrhythmia

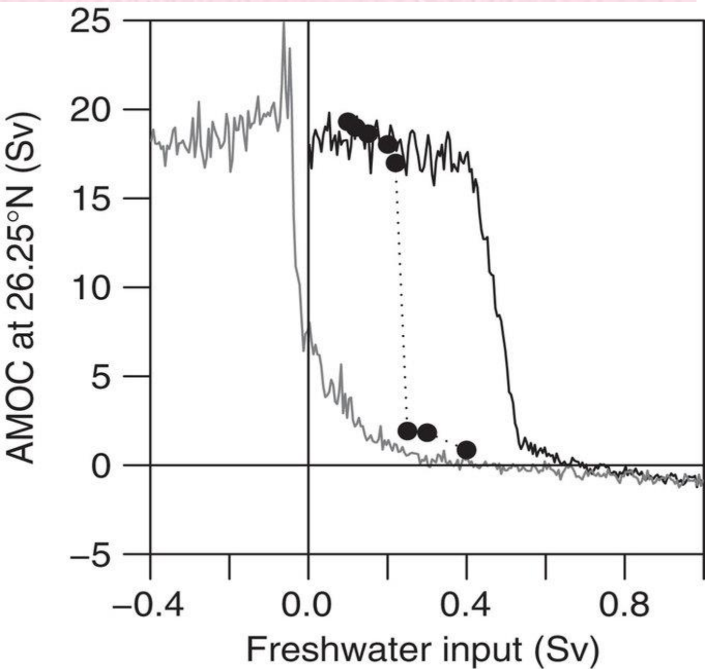


Source: Wikipedia



Collapse of Atlantic Meridional Overturning Circulation

Nat. Comm. 2014



Motivation and outline

To develop **diagnostic tools** to identify and characterize regime transitions.

- Lasers: optical chaos emerges from noise

  Empirical data, from a semiconductor ***laser*** *with feedback*

- Neuronal spikes: emergence of temporal correlations

  Synthetic data, generated with the ***neuron*** *FitzHugh-Nagumo model*

- Climate: how to quantify regional climate change

  Surface air ***temperature*** *(semi-empirical data)*

How optical chaos emerges from noise?

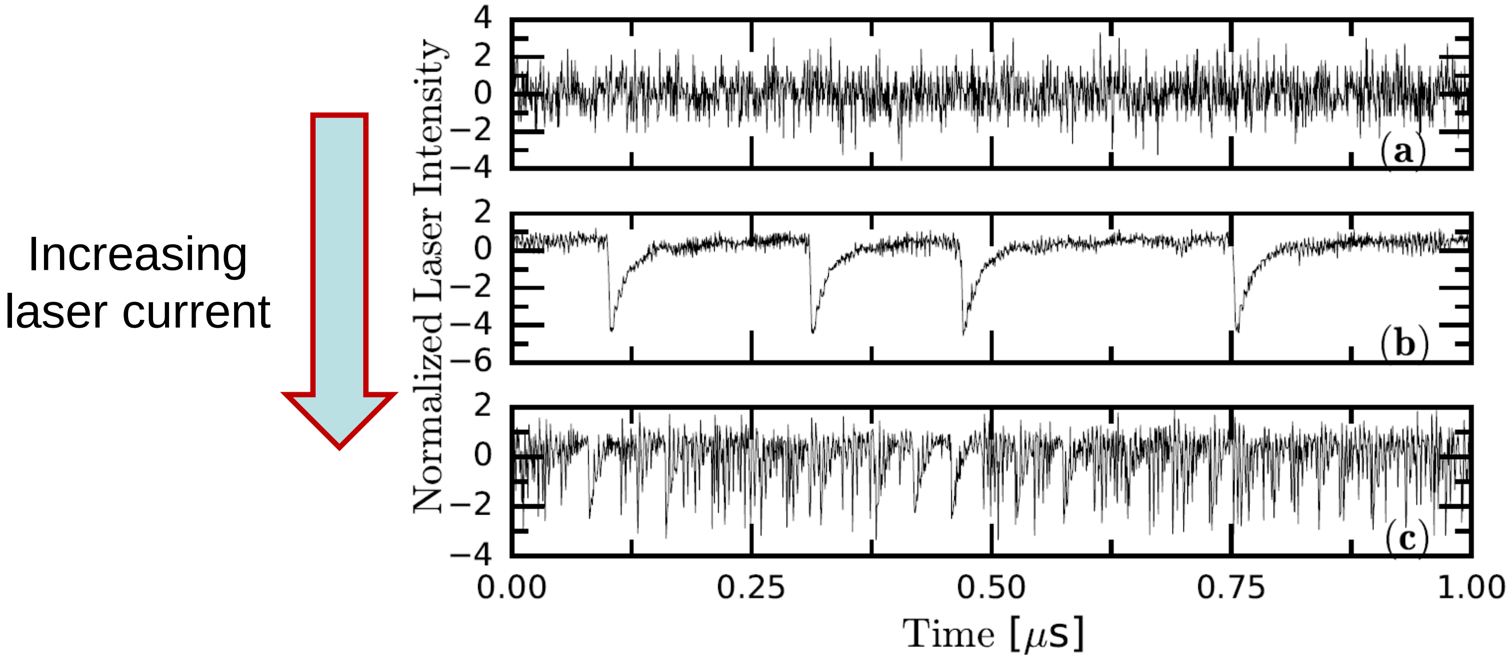
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

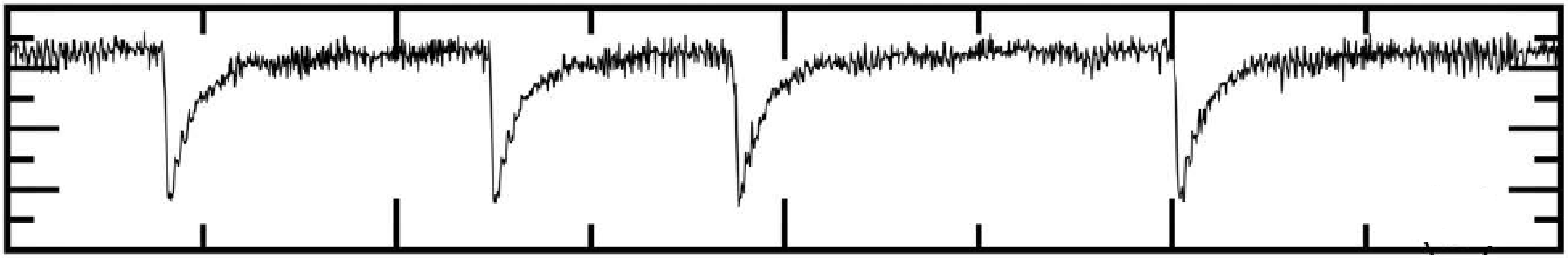
Campus d'Excel·lència Internacional

video

Low-frequency fluctuations (LFFs) and Coherence Collapse (CC)

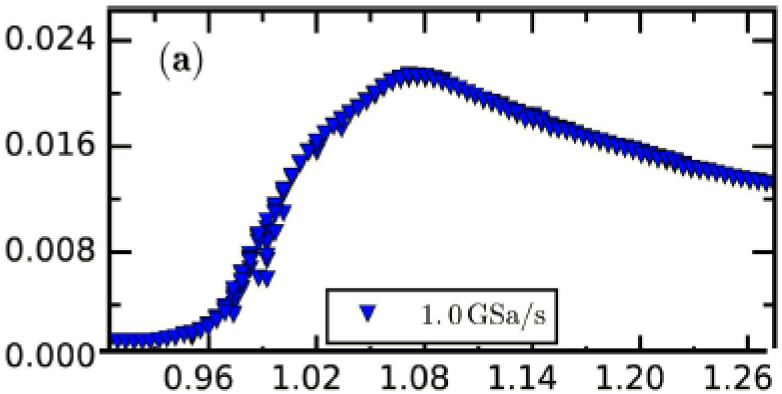
- LFF and CC regimes have been intensively studied.
- Two different regimes?
- If yes, can we identify the onset of each regime?



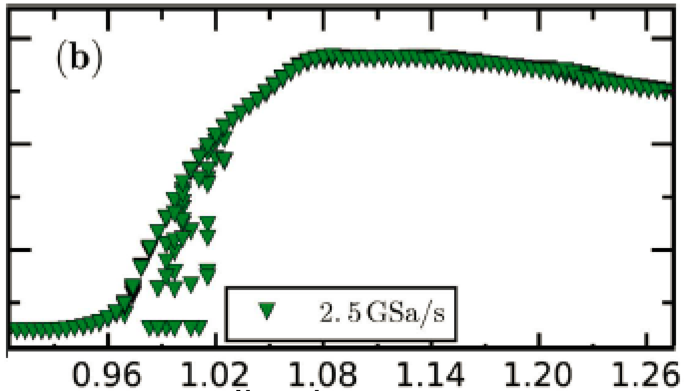


LFF slow-fast dynamics: intensity PDF depends on the oscilloscope sampling time

σ



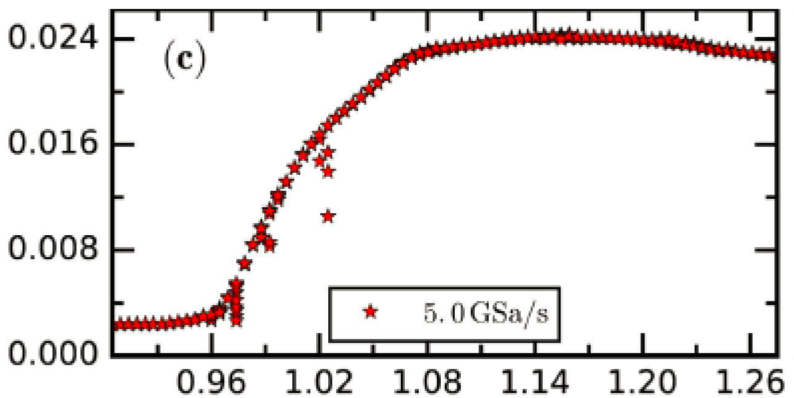
Normalized current



Normalized current

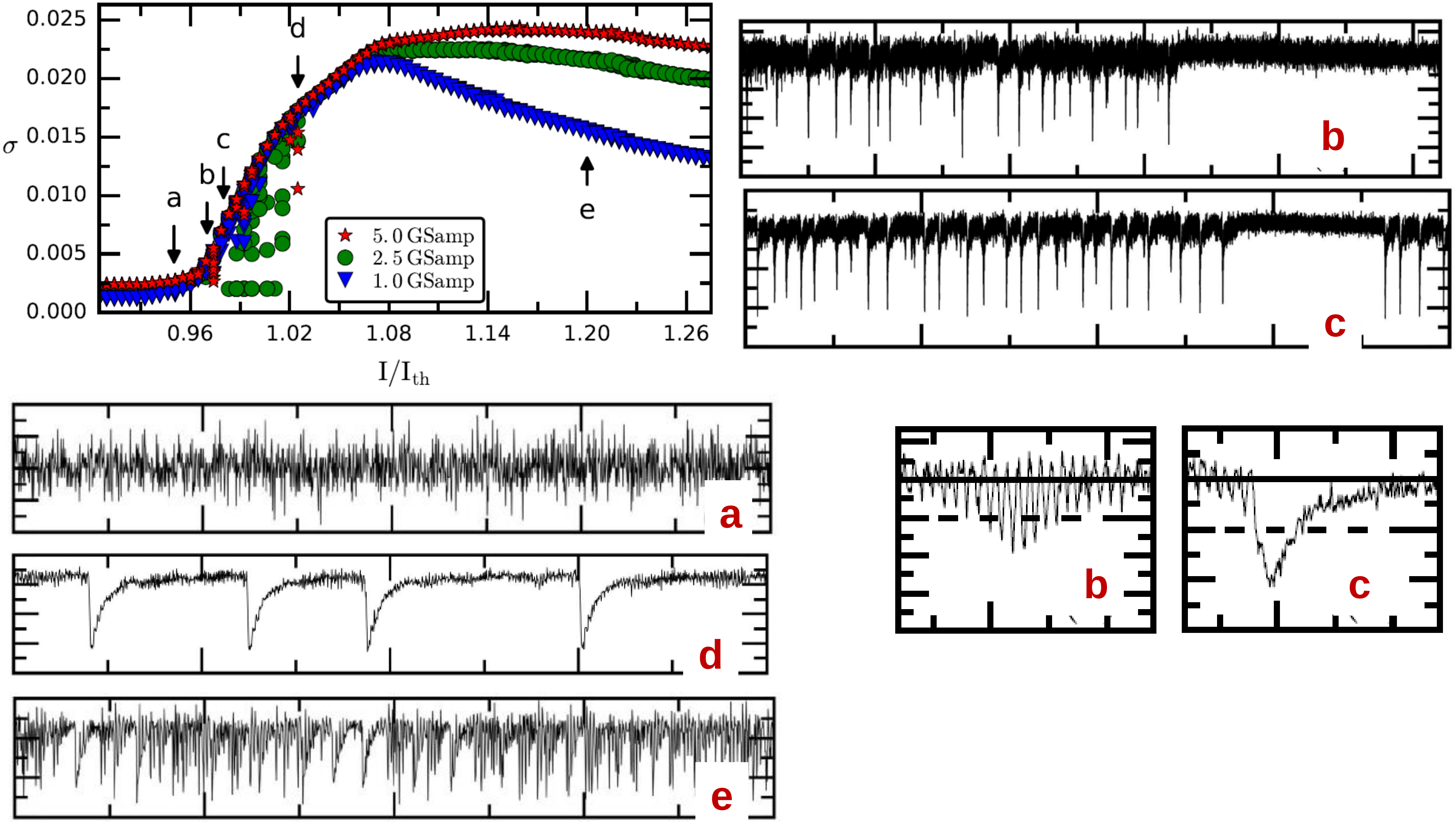
Threshold current=1

σ



Normalized current

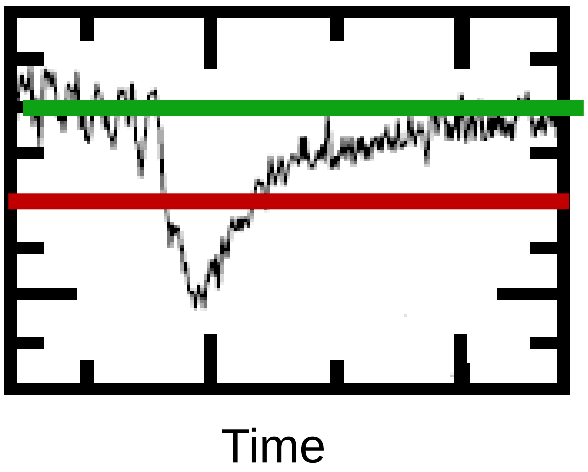
Identifying regime transition points

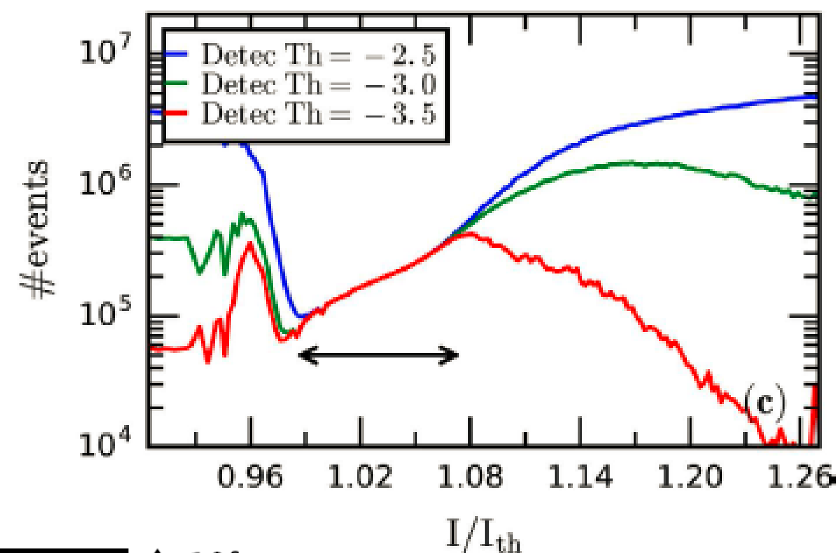


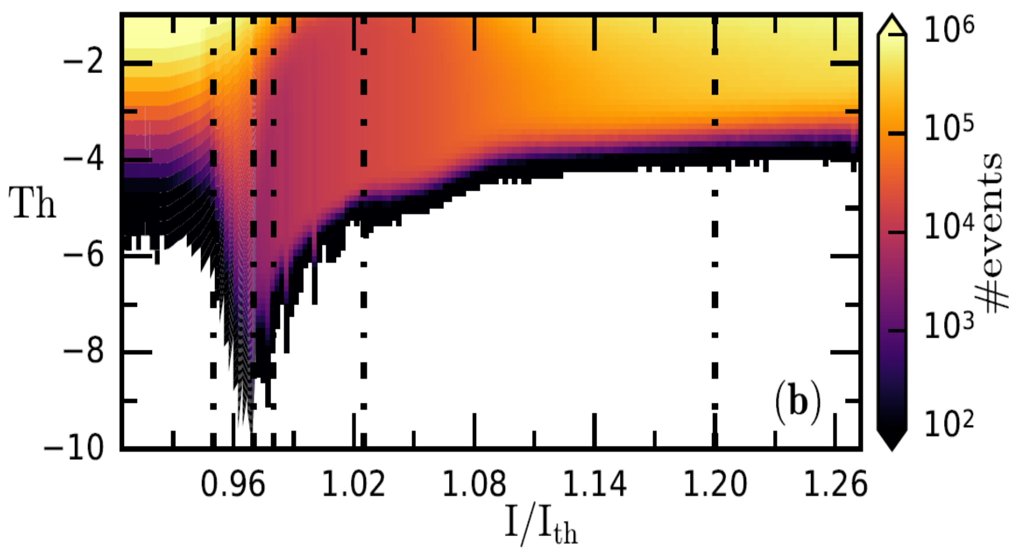
C. Quintero-Quiroz et al, Sci. Rep. 6 37510 (2016)

Number of threshold-crossing events

Each time series is first normalized to <x>=0 and σ=1







We can identify regions of:

- Regular LFFs
- "Extreme" LFFs

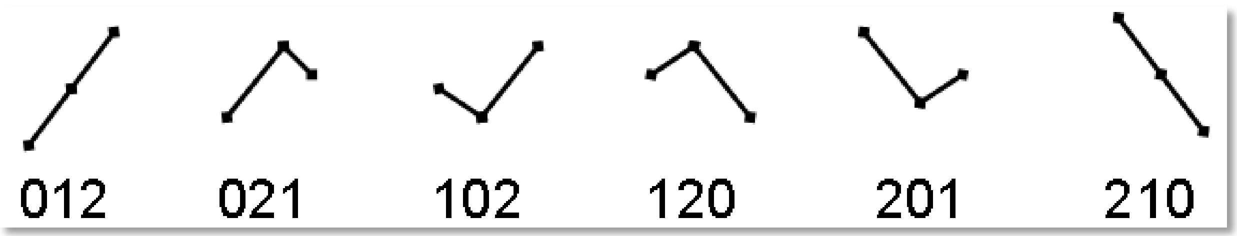
Temporal correlations: method of symbolic ordinal analysis

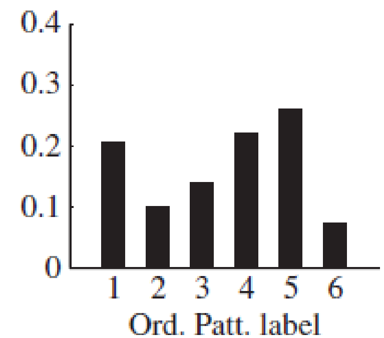
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

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

Brandt & Pompe, PRL 88, 174102 (2002)





The OP probabilities allow identifying <u>more expressed and/or infrequent patterns</u> in the <u>*order*</u> of the sequence of data values.

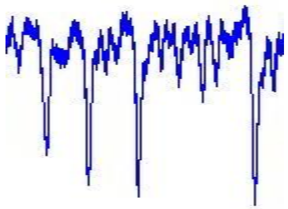
Random data?

- Advantage: the probabilities uncover temporal correlations.
- Drawback: we lose information about the actual values.

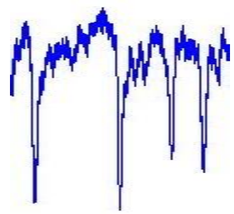
$\Rightarrow$ Ordinal analysis provides complementary information to that gained with other analysis tools.

Read more: M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, Entropy 14, 1553 (2012)

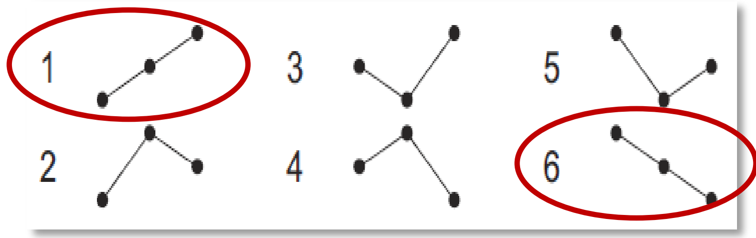
- **D=3**: correlations among 3 ***inter-spike-intervals*** (ISIs).



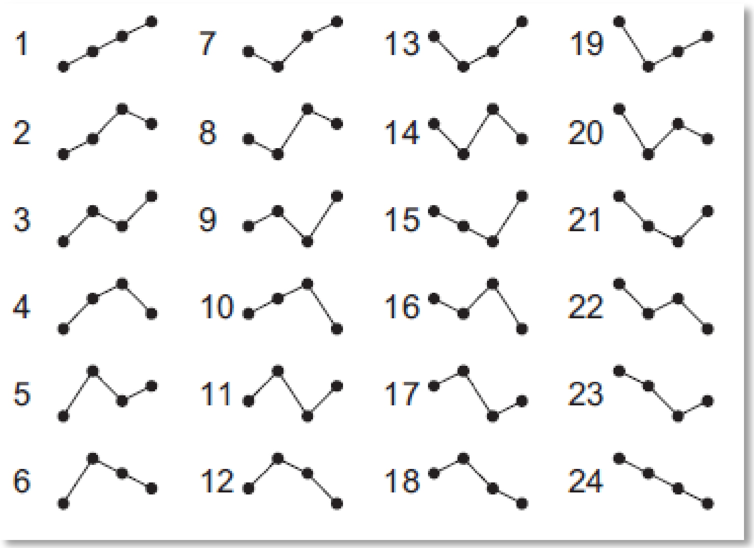
⇒**012**



⇒**210**



- The number of patterns grows as **D!**



- How to quantify the information?
  - Permutation entropy

$$s_p = -\sum p_i \log p_i$$

- How to select optimal D? depends on:
  - The length of the data.
  - The length of the correlations

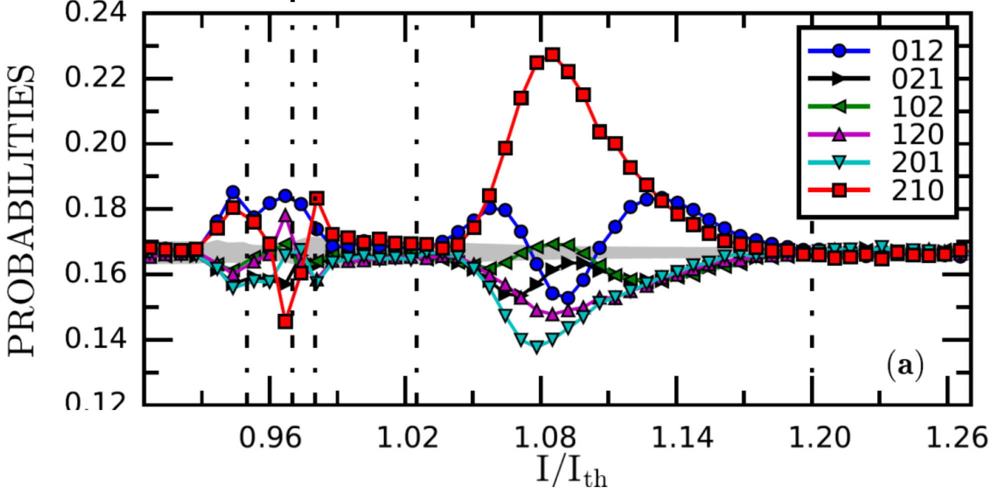
U. Parlitz et al. / Computers in Biology and Medicine 42 (2012) 319–327

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

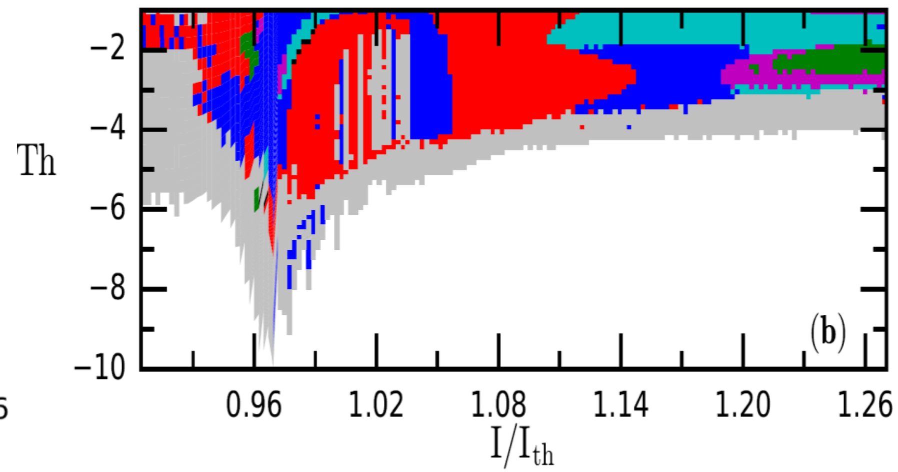
Campus d'Excel·lència Internacional

Ordinal analysis

Spike detection threshold = -3

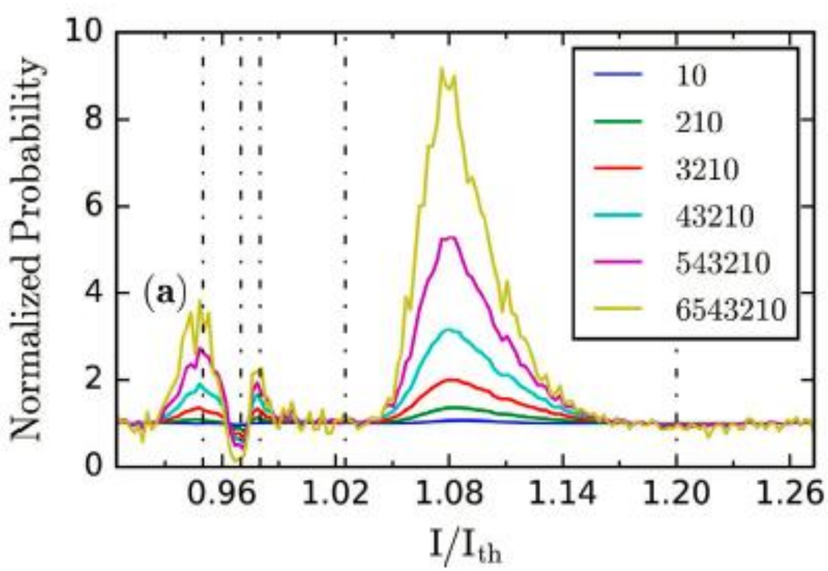


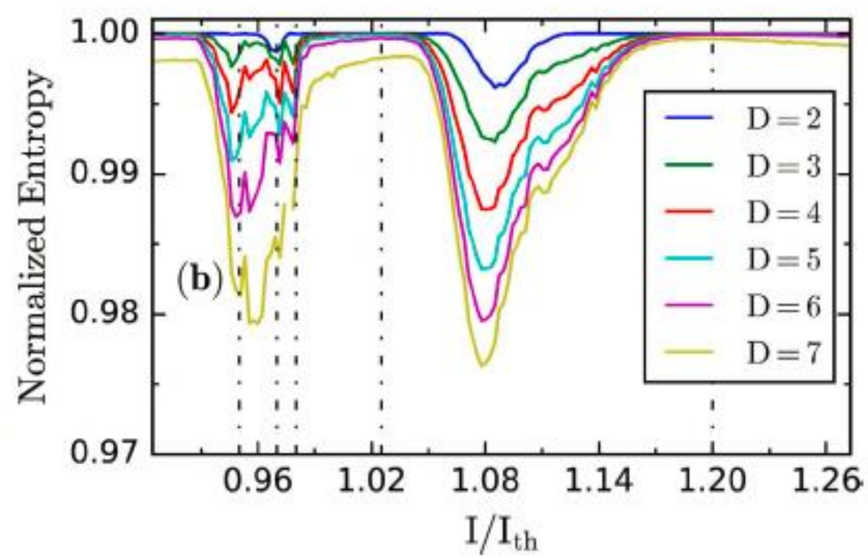
Most probable pattern



Gray region = equal probabilities

Gray region= not enough data





What did we learn?

- Transition to optical chaos:

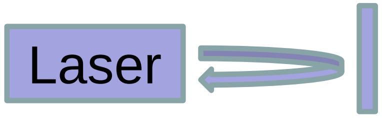
  - As the control parameter (pump current) increases, the low frequency fluctuations (LFFs) and the coherence collapse (CC) can be *quantitatively* distinguished.

  - By using three diagnostic tools:
    - σ of intensity pdf as a function of the sampling time;
    - number of threshold-crossings as function of the threshold;
    - ordinal probabilities as a function of the threshold.

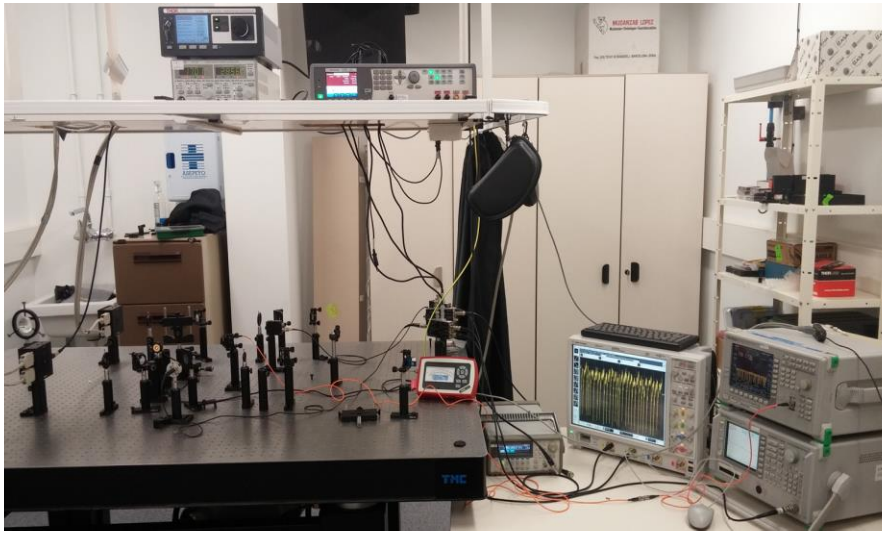
Emergence of temporal correlations in neuronal spikes

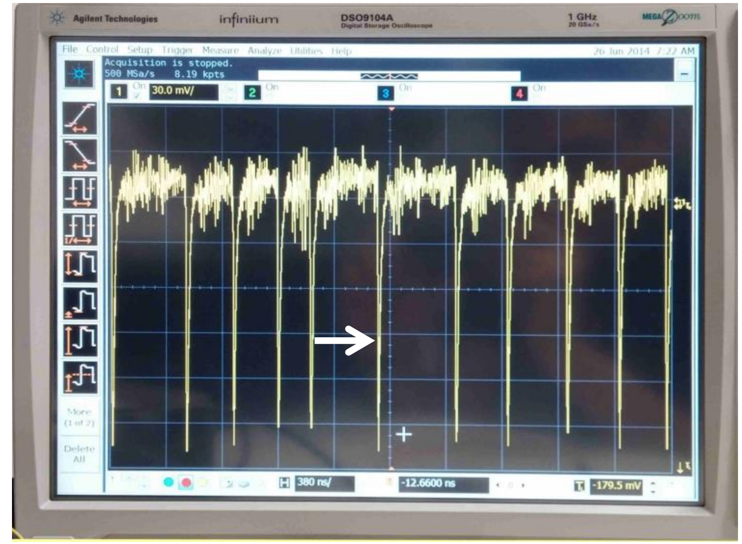
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

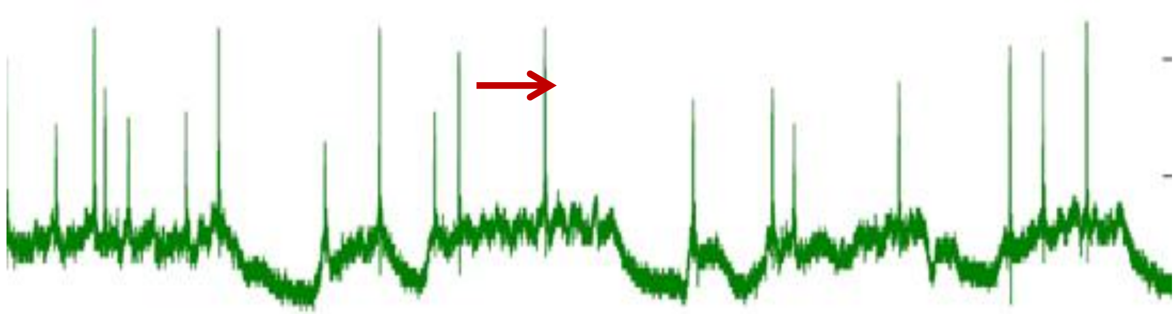
Campus d'Excel·lència Internacional



WHAT DO LASERS HAVE TO DO WITH NEURONS?







Similar statistics of inter-spike intervals?

MOTIVATION

Science 345, 668 (2014)

"a computer that is **inspired** by the brain."

Neuro-synaptic architecture allows to do things like image classification at a very low power consumption.

- Spiking lasers: photonic neurons?
- potential building blocks of brain-inspired computers.
- Ultra fast ! (micro-sec vs. mili-sec)

HOW SIMILAR NEURONAL AND OPTICAL SPIKES ARE?

Response to periodic stimulation

Neuron inter-spike interval (ISI) distribution

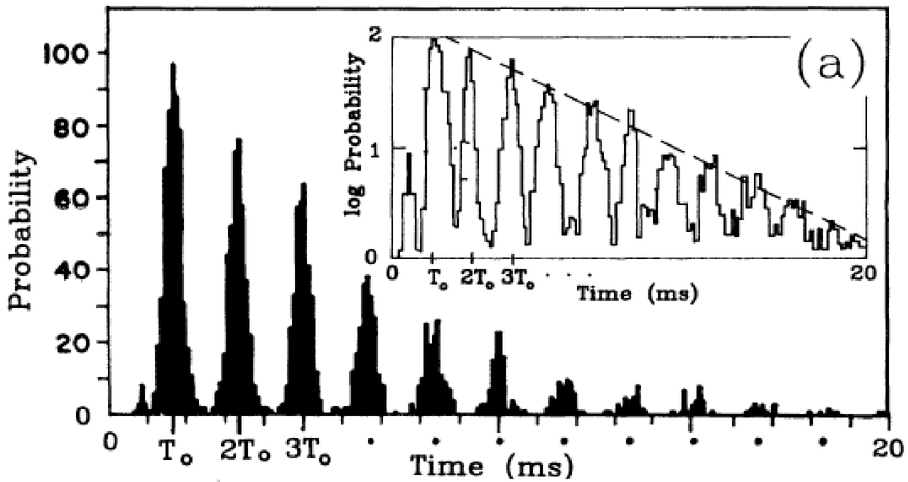
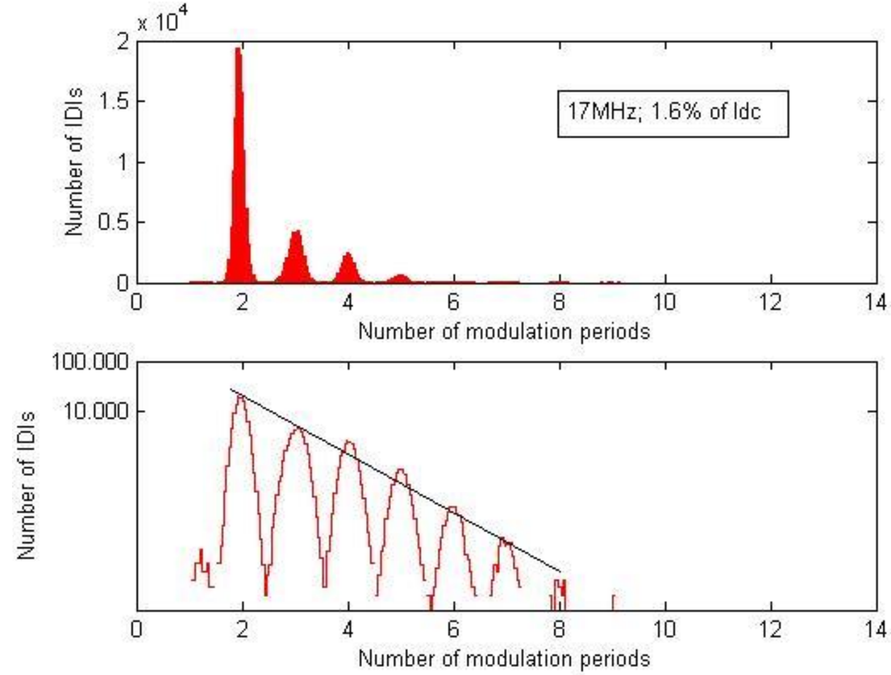


FIG. 1. (a) An experimental ISIH obtained from a single auditory nerve fiber of a squirrel monkey with a sinusoidal 80-dB sound-pressure-level stimulus of period $T_0 = 1.66$ ms applied at the ear. Note the modes at integer multiples of T_0. Inset:

A. Longtin et al, PRL 67 (1991) 656

Optical ISI distribution, data collected in our lab



when a sinusoidal signal is applied to the laser current

Neuronal ISIs

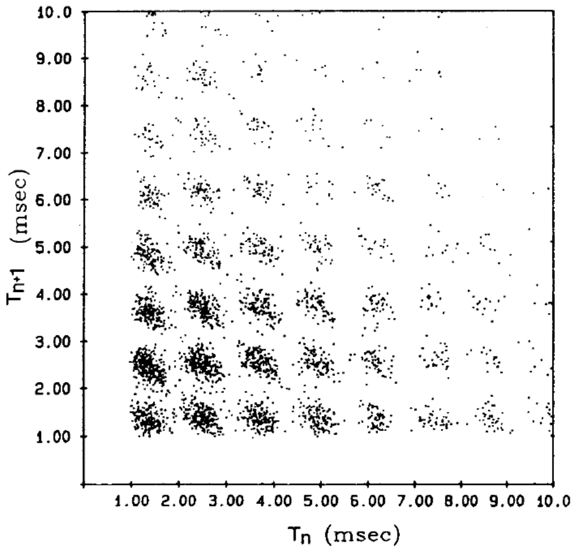
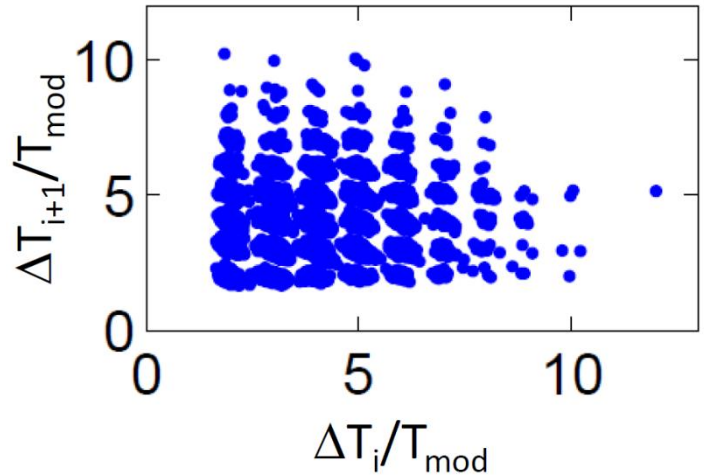
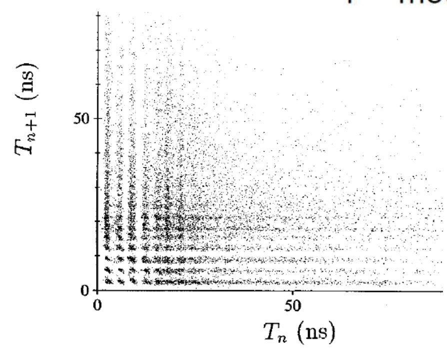


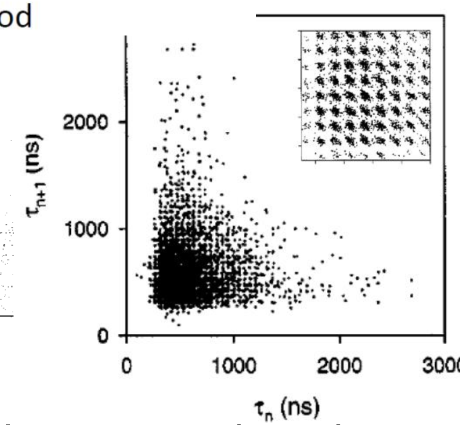
Fig. 4. Scatter plot of spike train data obtained from extracellular measurements of cat auditory fiber activity in response to an 800 Hz 60 dB sound pressure level pure tone presented to the outer ear. The stimulus is discontinuous (see

A. Longtin IJBC 3 (1993) 651

Optical ISIs

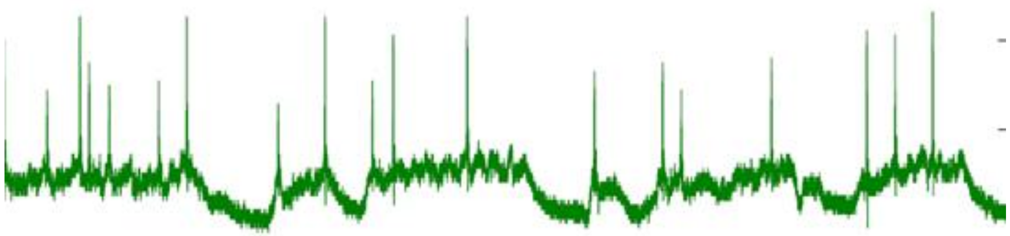






A. Aragoneses et al, Opt. Exp. (2014)
M. Giudici et al, PRE 55, 6414 (1997)
D. Sukow and D. Gautheir, JQE (2000)

How neurons encode information?



- In the spike rate?

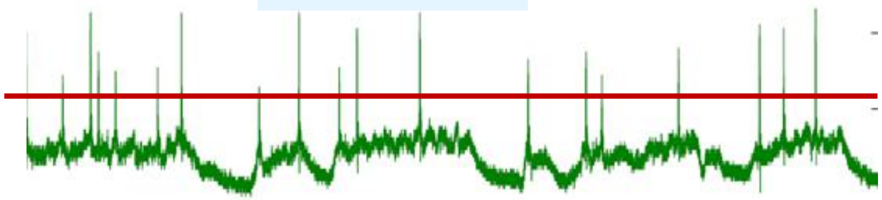
- Is the **timing** of the spikes relevant?
  - Rate-based information encoding is slow.
  - Temporal codes transmit more information.

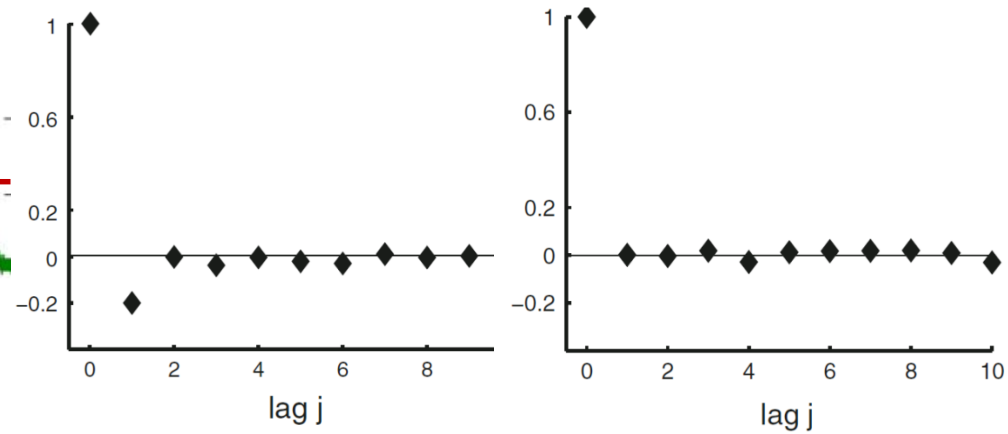
HOW TEMPORAL CORRELATIONS CAN BE IDENTIFIED AND QUANTIFIED?

Inter-spike-intervals serial correlation coefficients

$$\{\ldots I_{i-1}, I_i, I_{i+1} \ldots\} \qquad C_j = \frac{\langle (I_i - \langle I \rangle)(I_{i-j} - \langle I \rangle) \rangle}{\sigma^2}$$

$$I_i = t_{i+1} - t_i$$



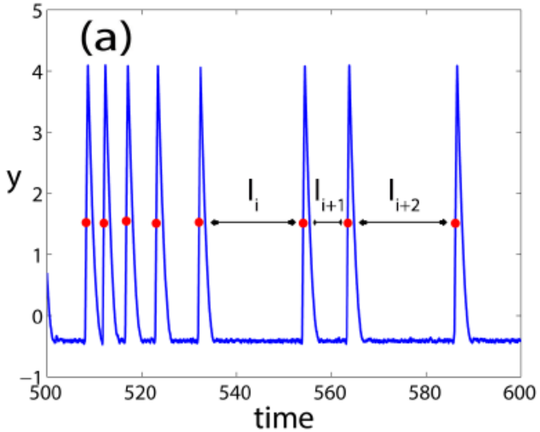


Exp Brain Res (2011) 210:353–371

HOW TO INDENTIFY TEMPORAL STRUCTURES?
RECURRENT / INFREQUENT PATTERNS?

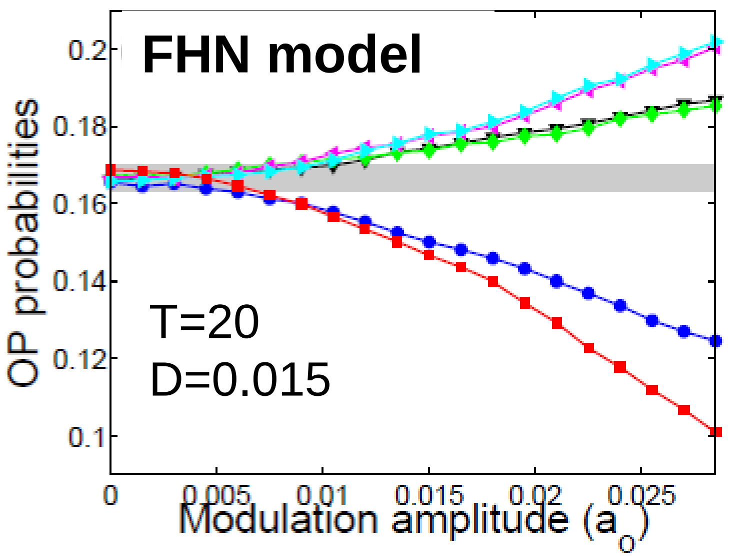
FHN model

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

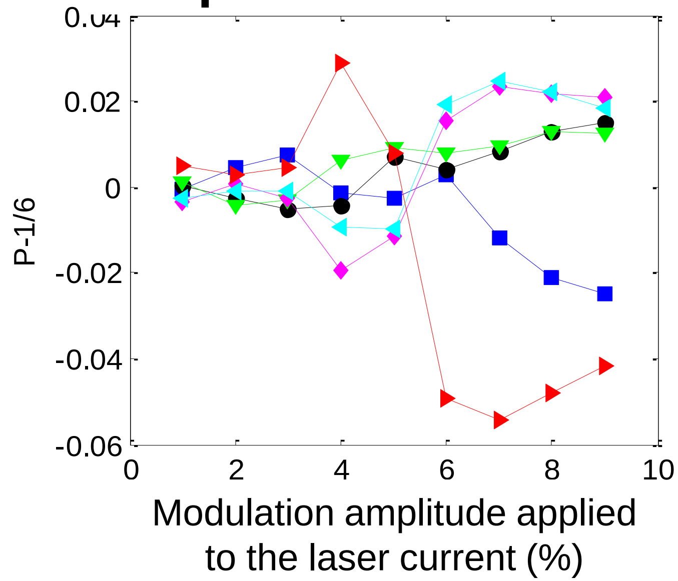


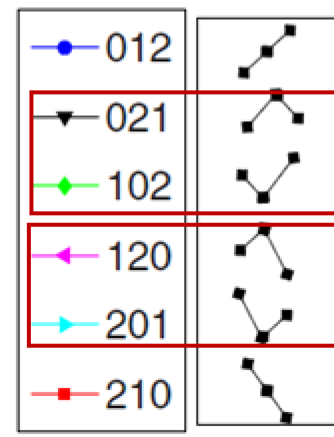
Gaussian white noise and *subthreshold* (weak) modulation: **a_0** and **T** such that *spikes are only noise-induced*.

Time series with 100,000 ISIs simulated.



Empirical laser data



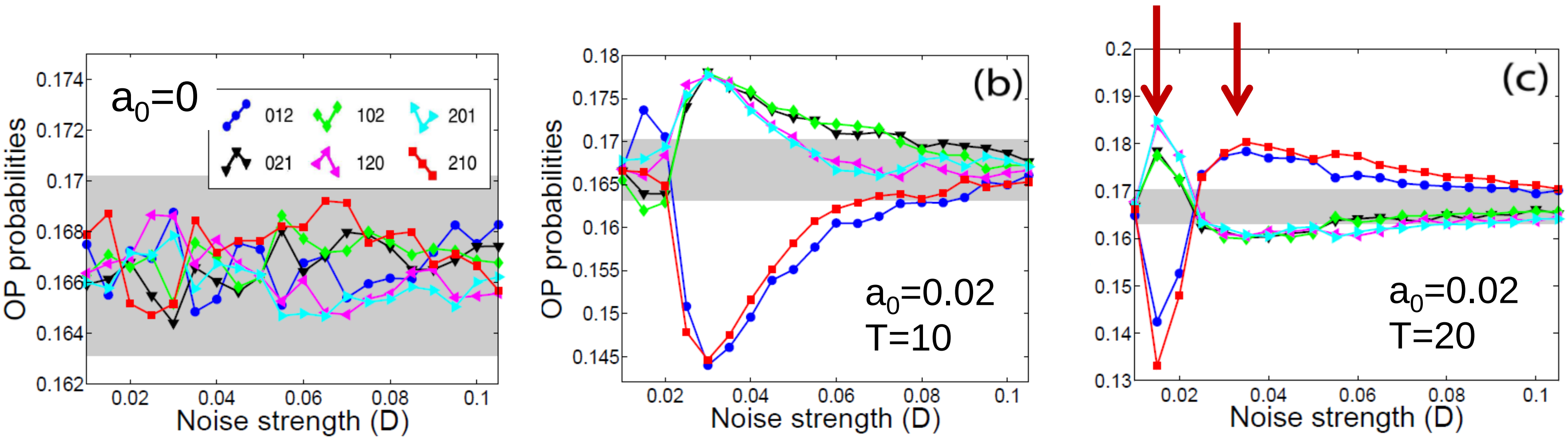


$\Rightarrow$ Good qualitative agreement

J. A. Reinoso et al PRE 94, 032218 (2016)

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

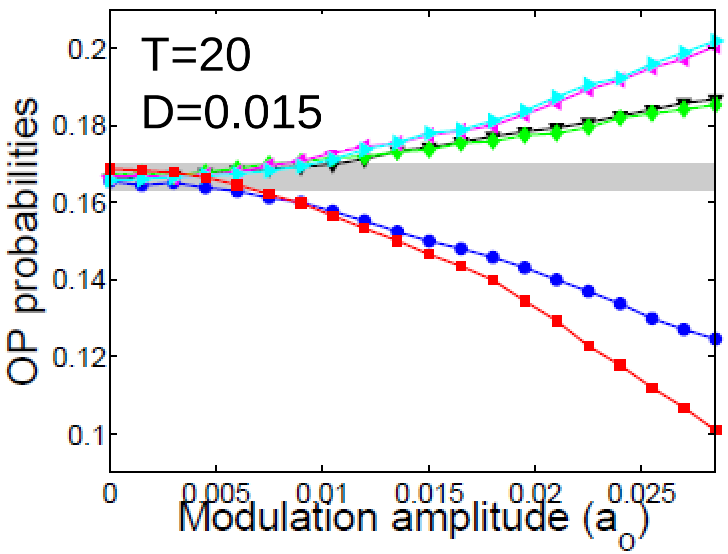
FHN model: role of the noise strength

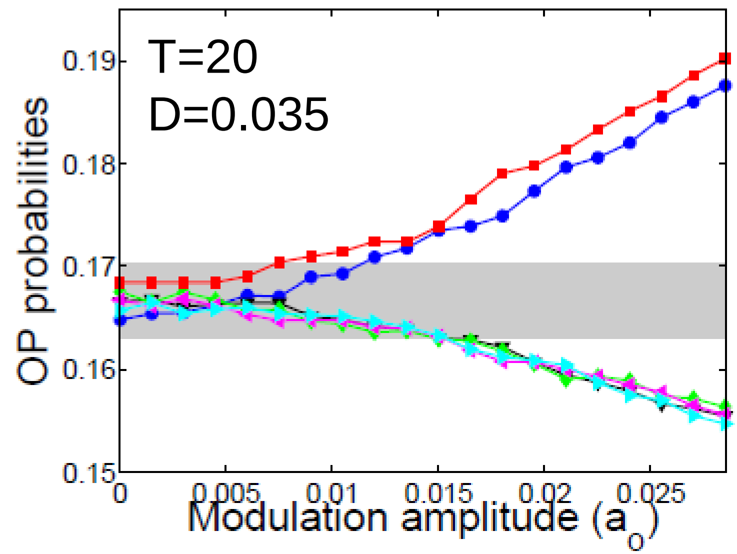


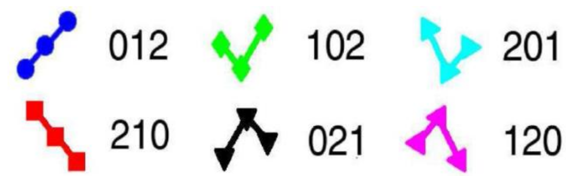
- Without external input: no temporal ordering.
- External periodic input induces temporal ordering.
- Preferred ordinal patterns depend on the noise strength.
- Resonant-like behavior.

J. A. Reinoso, M. C. Torrent and CM, PRE 94, 032218 (2016)

Role of the modulation amplitude







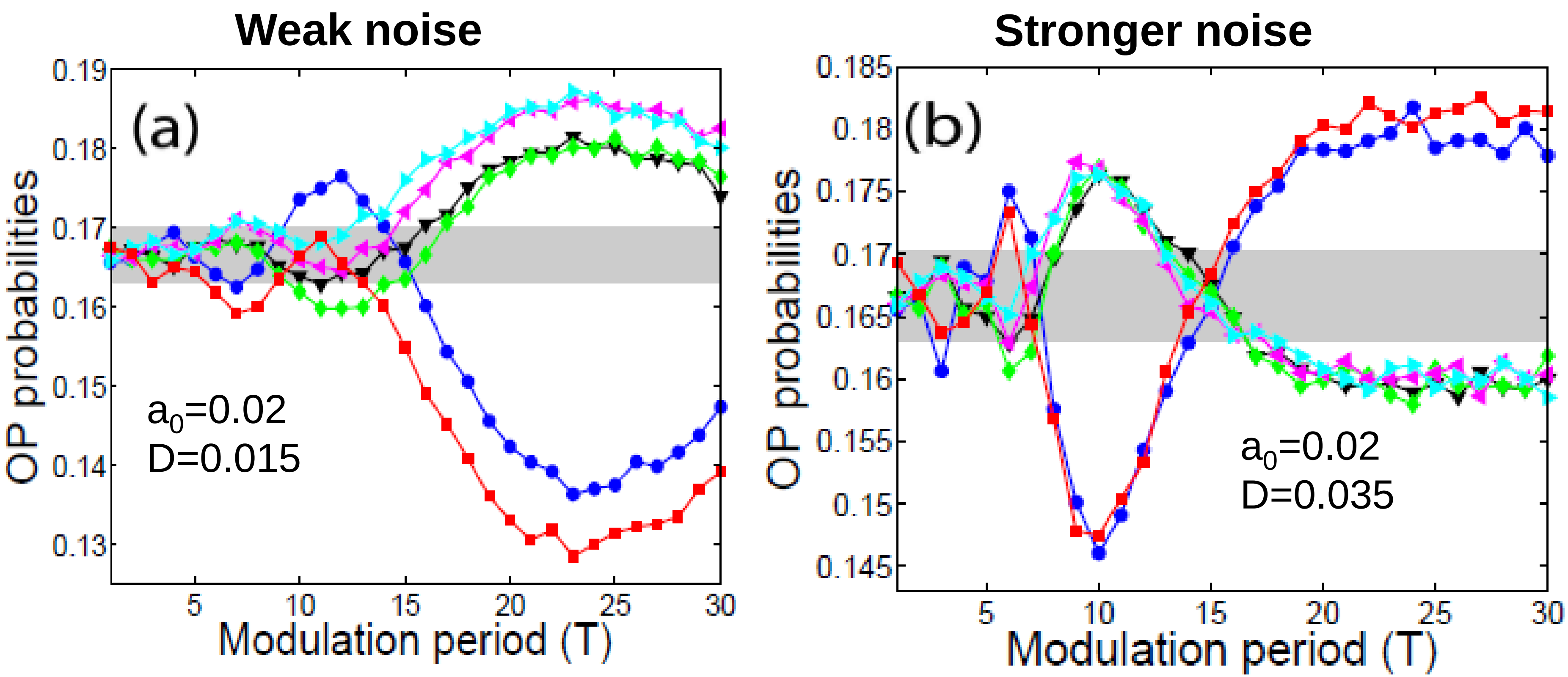
- The amplitude of the (weak) modulation does not modify the preferred and the infrequent patterns.

J. A. Reinoso, M. C. Torrent and CM, PRE 94, 032218 (2016)

Role of the modulation period

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

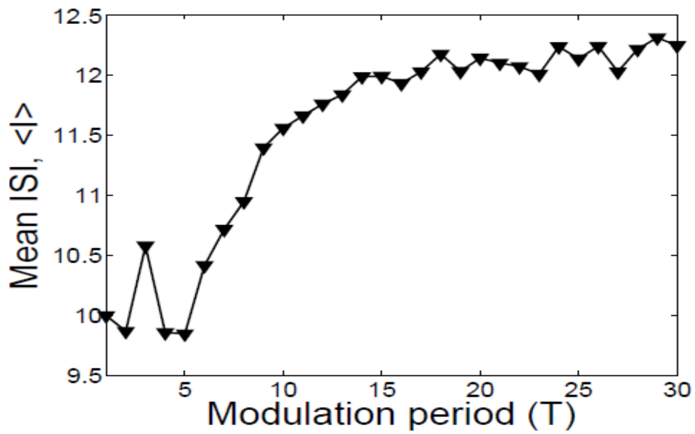


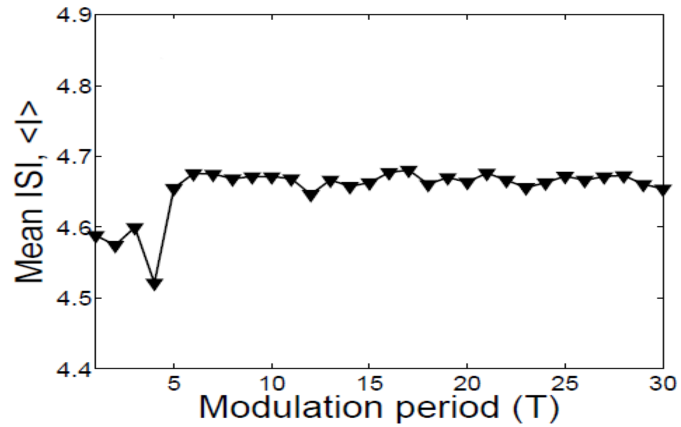
More probable patterns depend on the period of the external signal and on the noise strength.

J. A. Reinoso, M. C. Torrent and CM, PRE 94, 032218 (2016)

Underlying mechanism?

- change of the spike rate? (i.e., mean ISI)

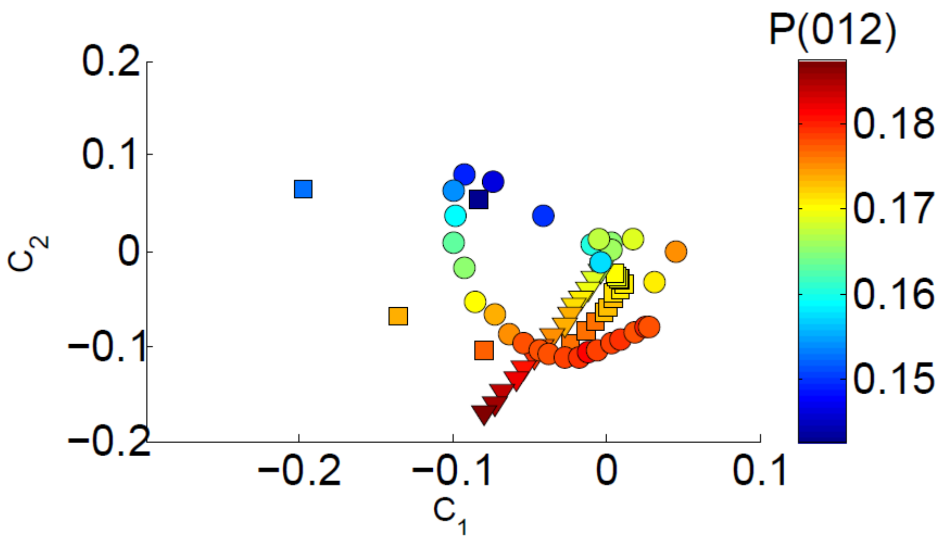




$\Rightarrow$ No direct relation.

- Serial correlation coefficients?

$$C_j = \frac{\langle (I_i - \langle I \rangle)(I_{i-j} - \langle I \rangle) \rangle}{\sigma^2}$$



$\Rightarrow$ there is a relation but it is not univocal.

Ongoing work

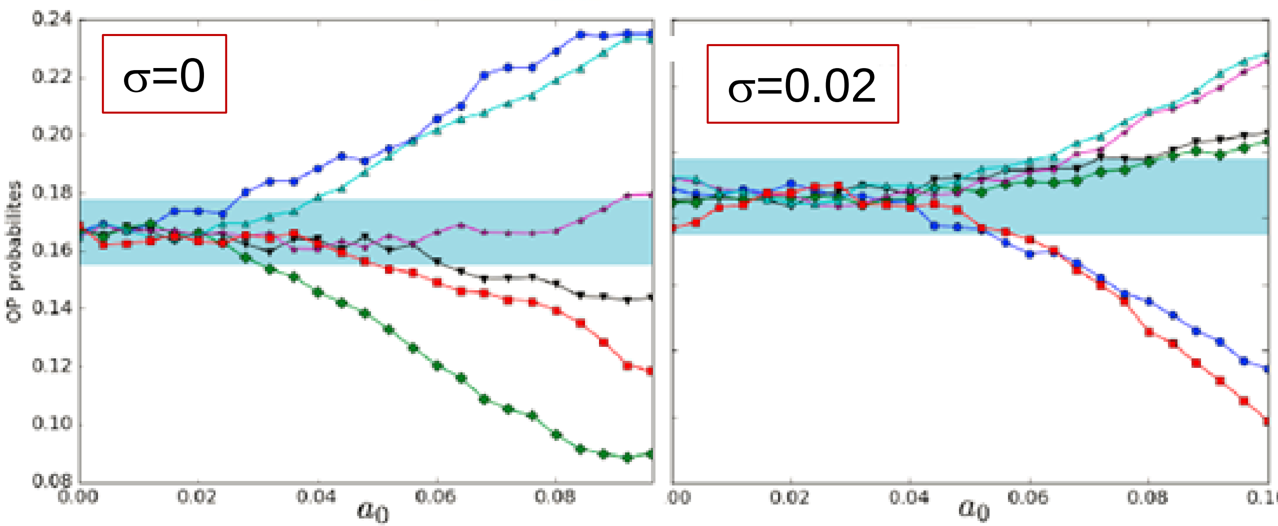
- Role of weak coupling? induce preferred/infrequent patterns?

$$\epsilon \dot{u}_1 = u_1 - \frac{u_1^3}{3} - v_1 + \overbrace{\boxed{a_0 \cos(2\pi t/T)}}^{\text{modulation}} + \overbrace{\boxed{\sigma u_2}}^{\text{Coupling}} + \sqrt{2D}\xi_1(t)$$

$$\dot{v}_1 = u_1 + a,$$

$$\epsilon \dot{u}_2 = u_2 - \frac{u_2^3}{3} - v_2 + \boxed{\sigma u_1} + \sqrt{2D}\xi_2(t)$$

$$\dot{v}_2 = u_2 + a$$



What did we learn?

- Temporal correlations induced by periodic modulation: good qualitative agreement in optical & neuronal spikes.
- FHN model:
  - Preferred ordinal patterns depend on the noise strength and on the period of the input signal, but not on (weak) amplitude of the signal.
  - Resonance-like behavior: certain periods and noise levels maximize the probabilities of the preferred patterns, enhancing temporal order.

- Open issues:
  - Hierarchical & clustered structure: universal feature?
  - Mathematical insight: can we calculate the probabilities?
  - Empirical data? (single-neuron ISI sequences)

How to quantify regional climate change?

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

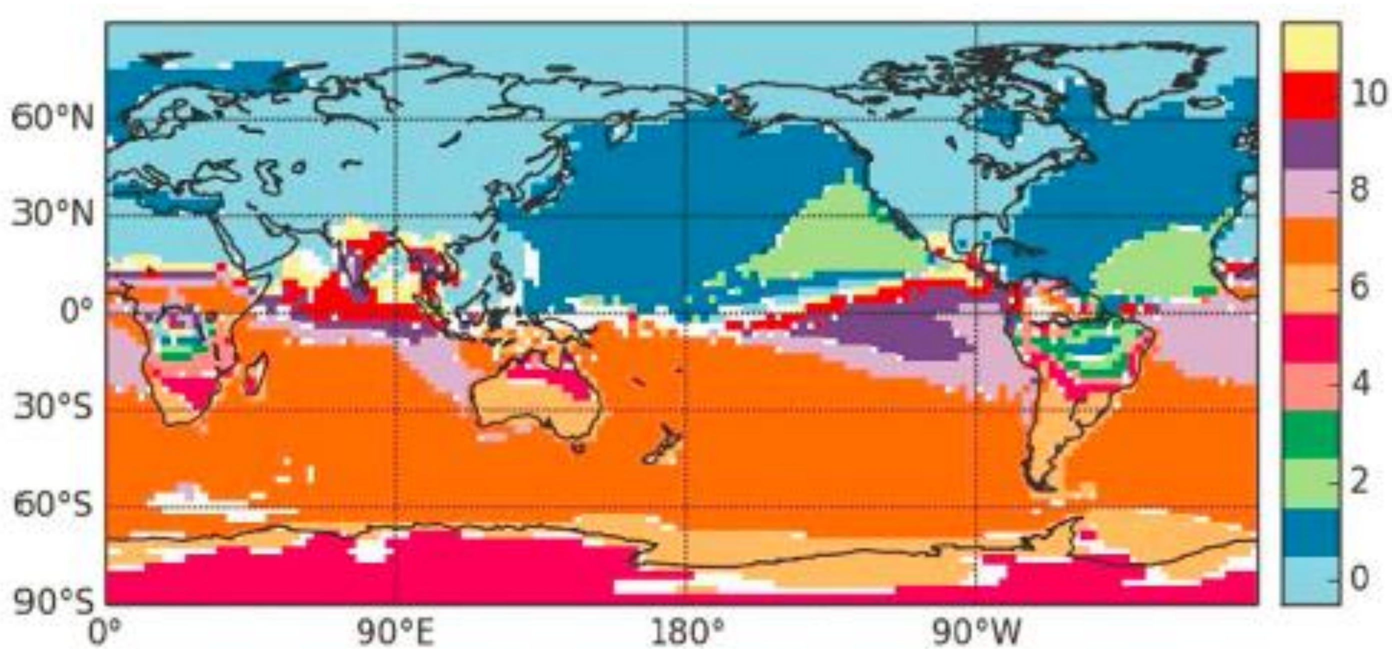
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

Initial motivation

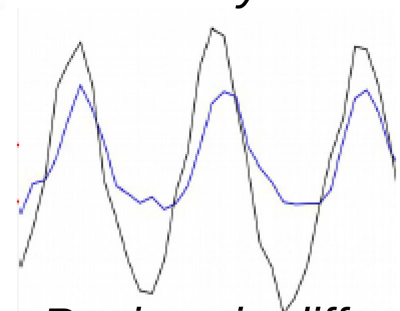
- To detect phase / frequency synchronization in climate data.

- Communities of regions with “in-phase” seasonal cycle.



G. Tirabassi and C. M., Sci. Rep. 2016

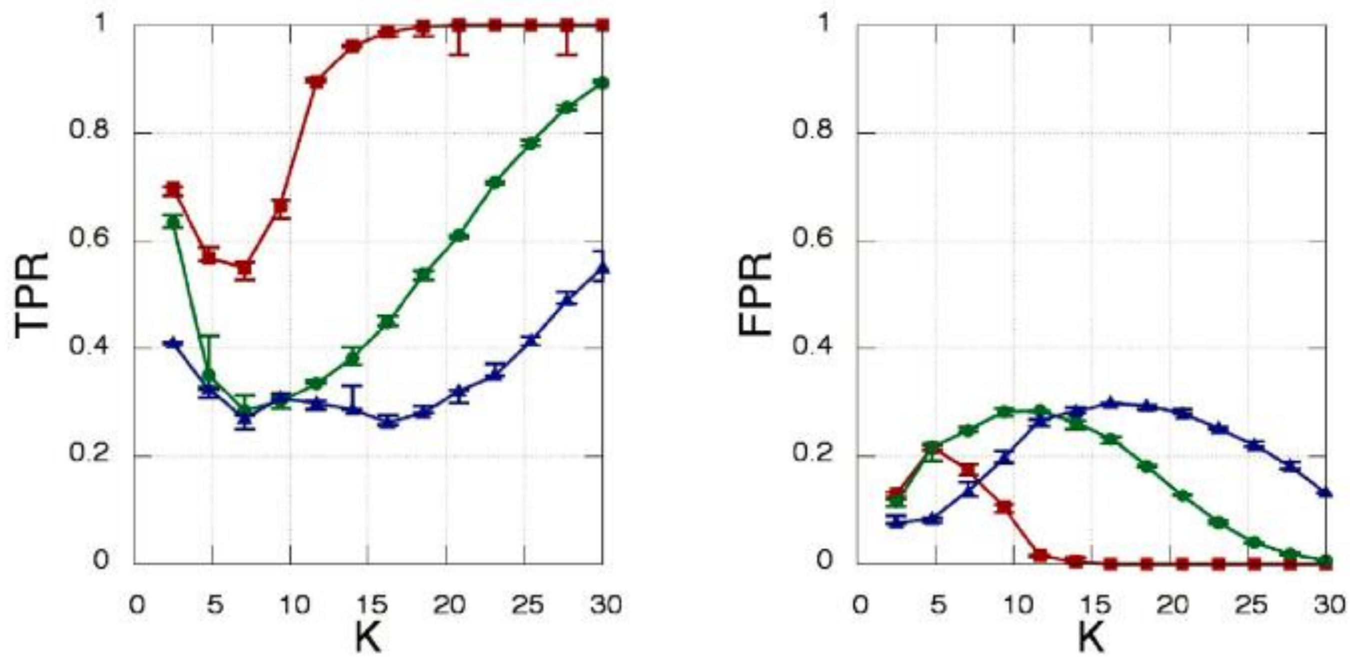
Regions in the same community



Regions in different communities

Also a motivation

- In coupled Kuramoto oscillators, the best way to infer the network from data is by similarity analysis of instantaneous frequencies.



CC
MI
MIOP

G. Tirabassi et al, Sci. Rep. 2015

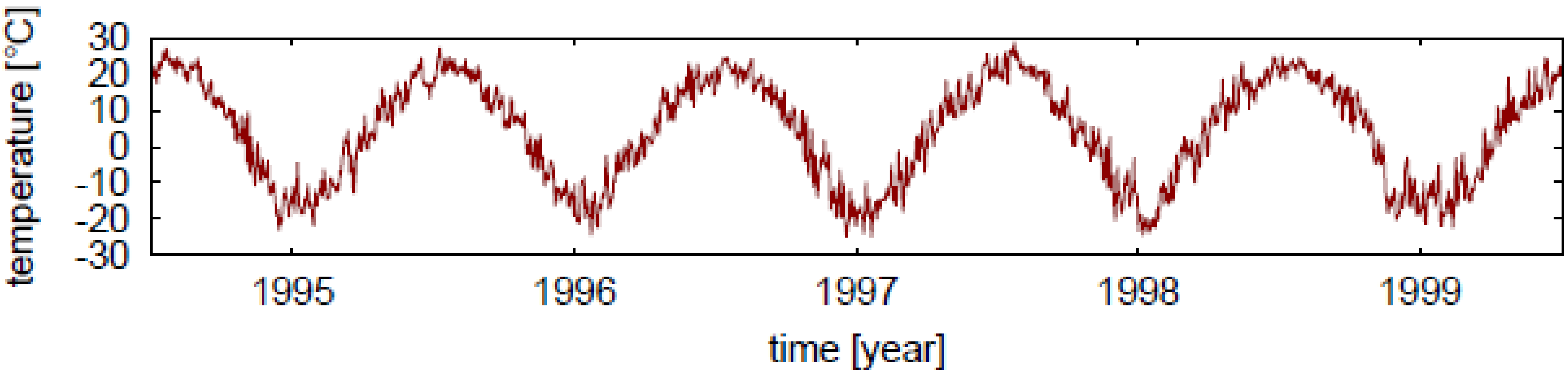
The data: surface air temperature

- Spatial resolution 2.5 x 2.5 $\Rightarrow$ 10512 time series
- Daily resolution January 1979 to June 2016 $\Rightarrow$ 12328 data points

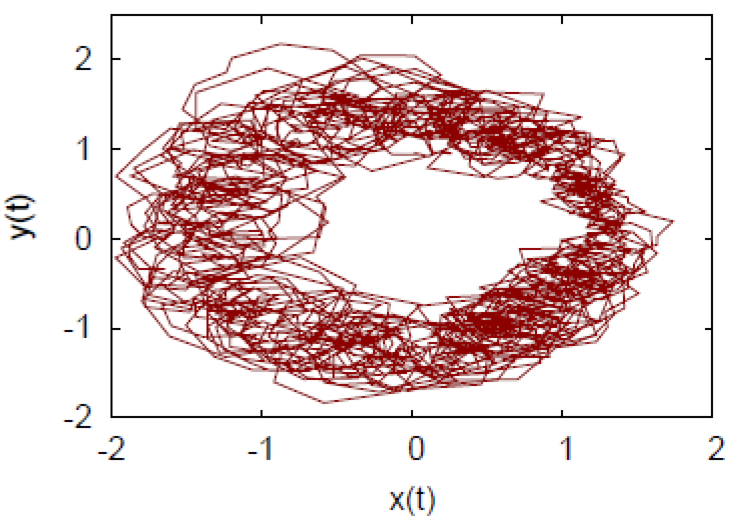
Where does the data come from?

- Freely available (ERA-Interim by European Centre for Medium-Range Weather Forecasts, ECMRF)
- Reanalysis = run a sophisticated model of general atmospheric circulation and feed the model (data assimilation) with empirical data, where and when available.
- This process restricts the solution of the model to one as close to reality as possible in regions/times where there are data available, and to a solution physically "plausible" in regions/times where no data is available.

Hilbert analysis



- De-trend and normalize each SAT time-series: $\langle x(t)\rangle=0$, $\sigma_x=1$
- Then apply Hilbert transform (***H*(sin ωt) = cos ωt**)
- **No filter is applied**.



$$a_j(t) = \sqrt{[x_j(t)]^2 + [y_j(t)]^2}$$

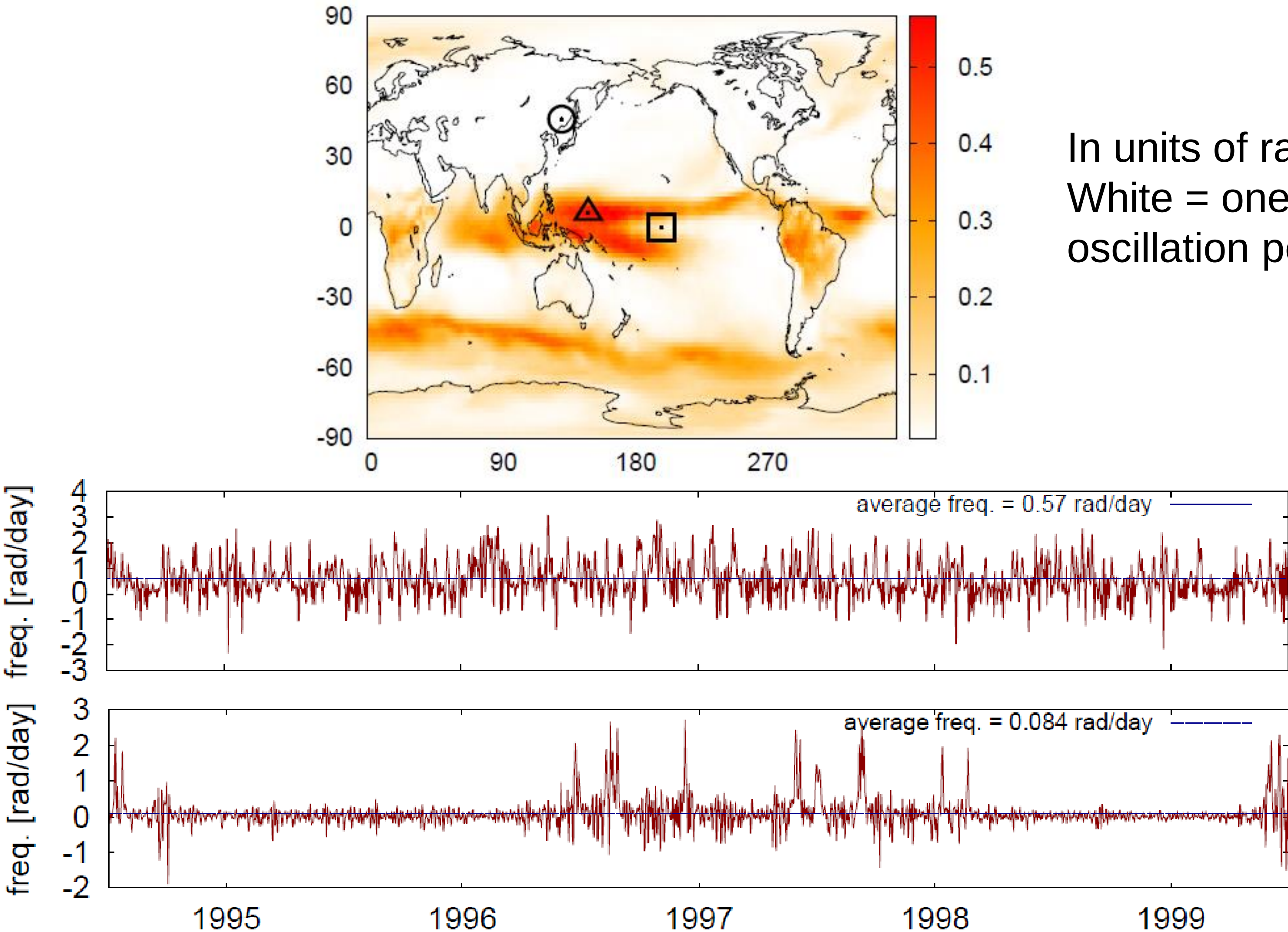
$$\varphi_j(t) = \arctan \frac{y_j(t)}{x_j(t)} \quad \Rightarrow \omega_j(t) = \dot{\varphi}_j(t)$$

$$\boxed{x_j(t) = a_j(t)\cos\big(\omega_j(t)t\big)}$$

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

Time-averaged instantaneous frequency



In units of rad/day
White = one oscillation per year

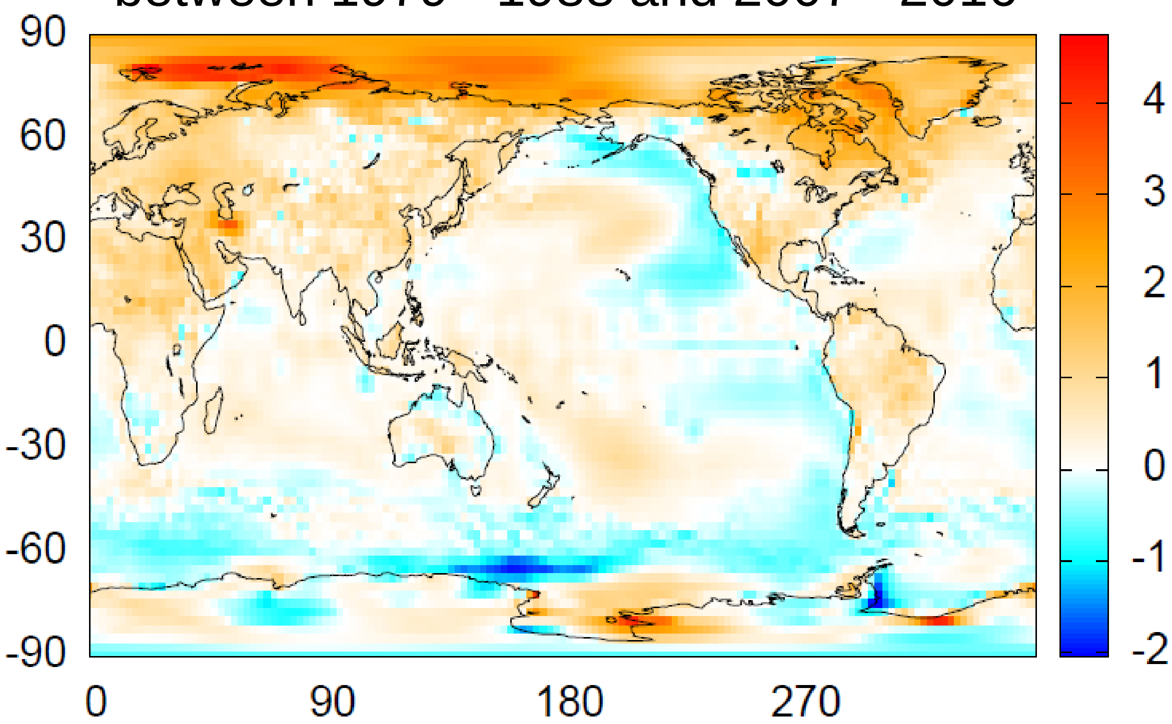
▲

■

D. A. Zappala, M. Barreiro, and CM, Entropy 2016

Global warming

Change in the average temperature (in C)
between 1979 - 1988 and 2007 - 2016



Monte Perdido (Pyrenees)

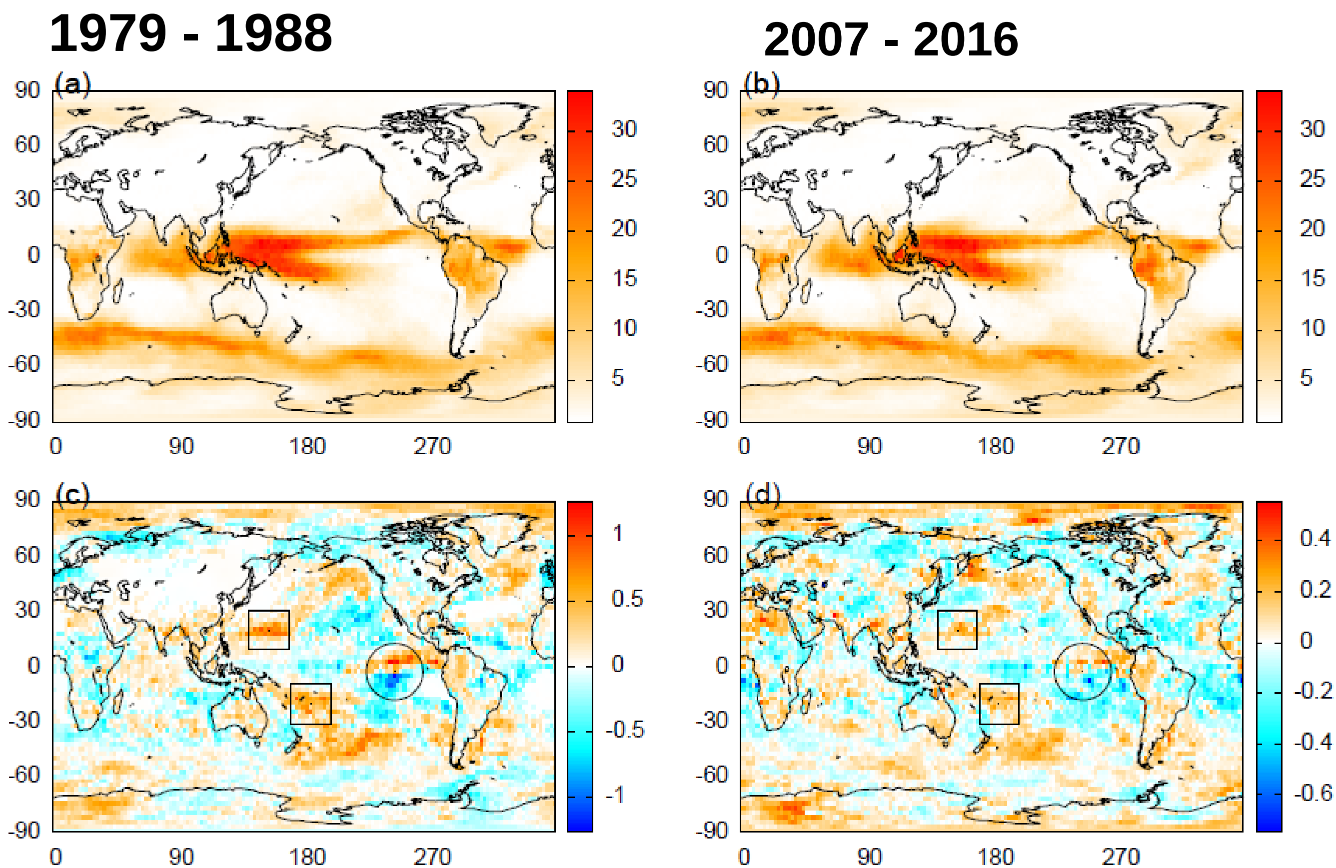
Change in averaged Hilbert frequency

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

White: one oscillation per year

$$\Delta\omega / \langle\omega\rangle$$



Zero-crossings

Frequency variations capture a northward shift of the inter-tropical convergence zone (ITCZ) and a widening of the rainfall band in the western Pacific Ocean

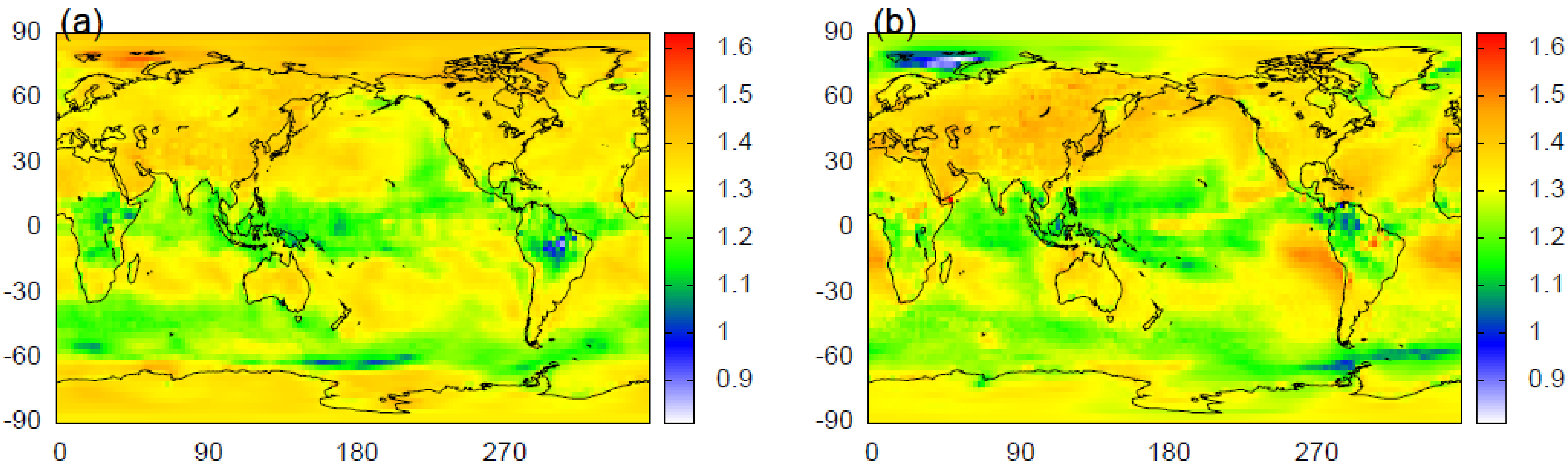
D, A. Zappala, M, Barreiro, and CM, submitted 2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

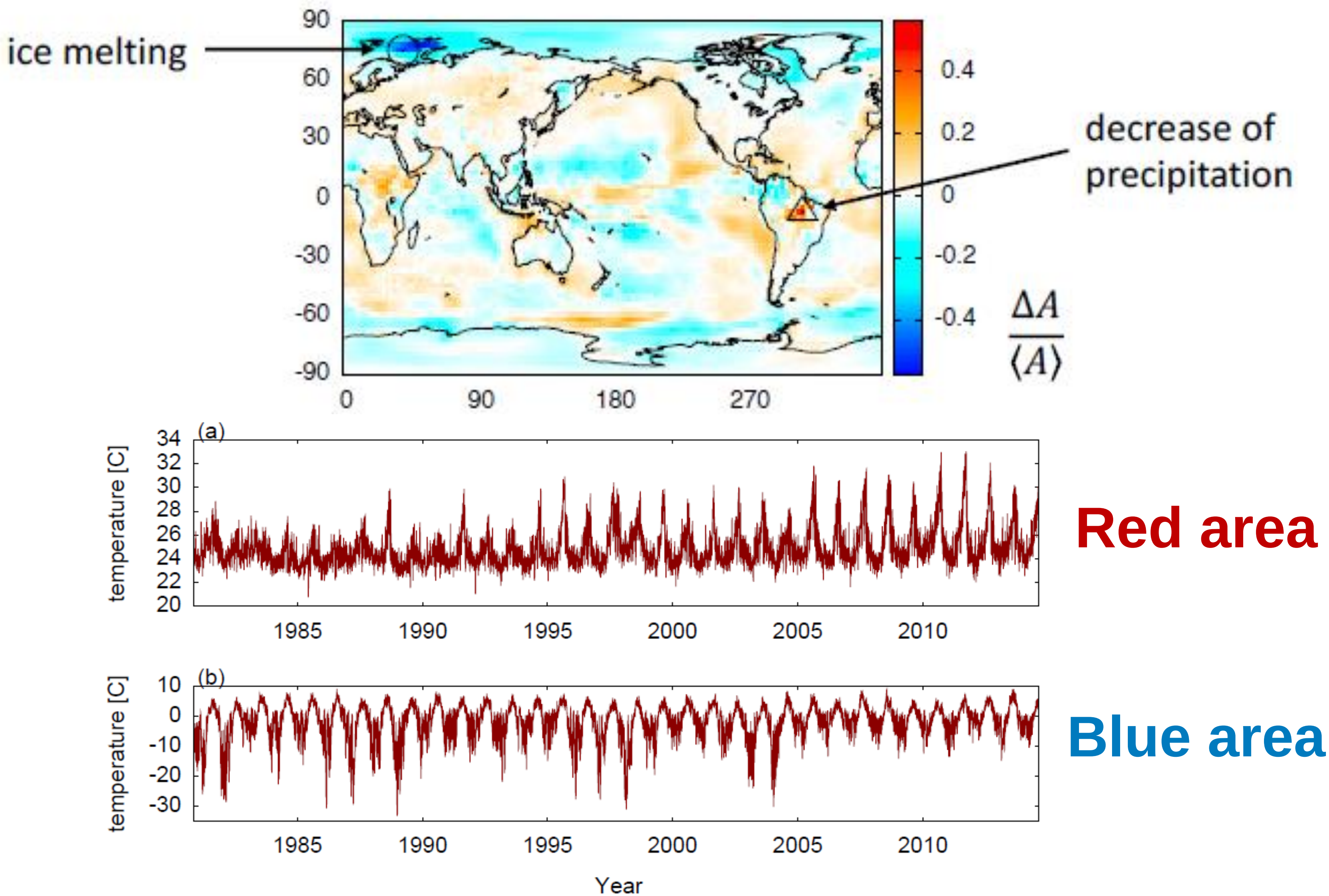
Campus d'Excel·lència Internacional

Time-averaged Hilbert amplitude

1979 - 1988 **2007 - 2016**

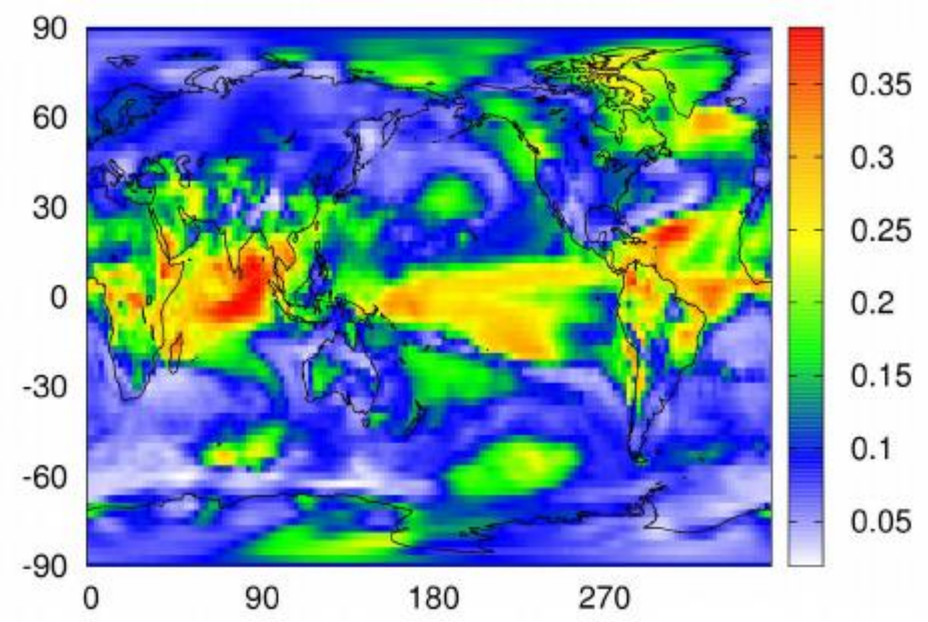


Variation and time series in "hotspots"

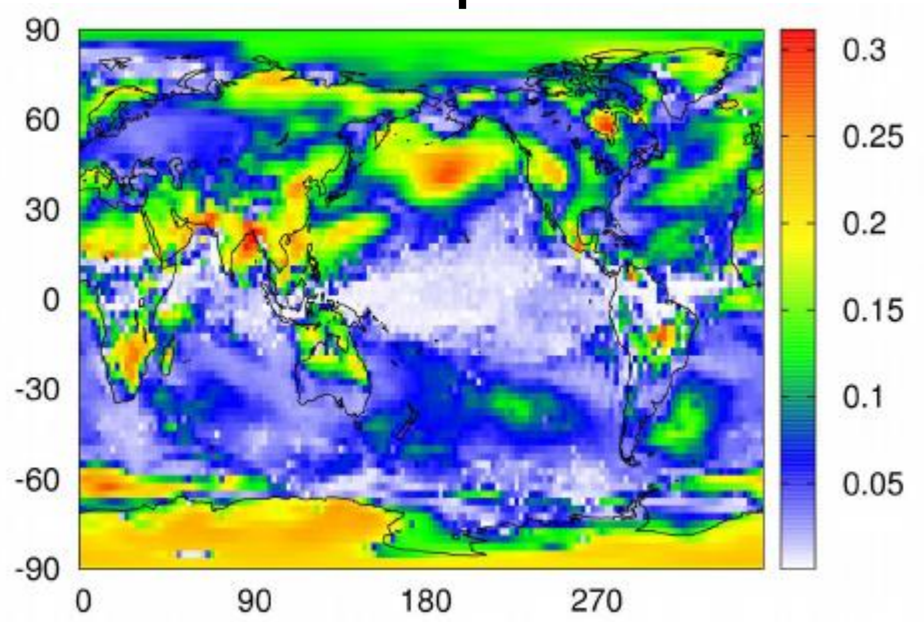


Ongoing work: frequency synchronization

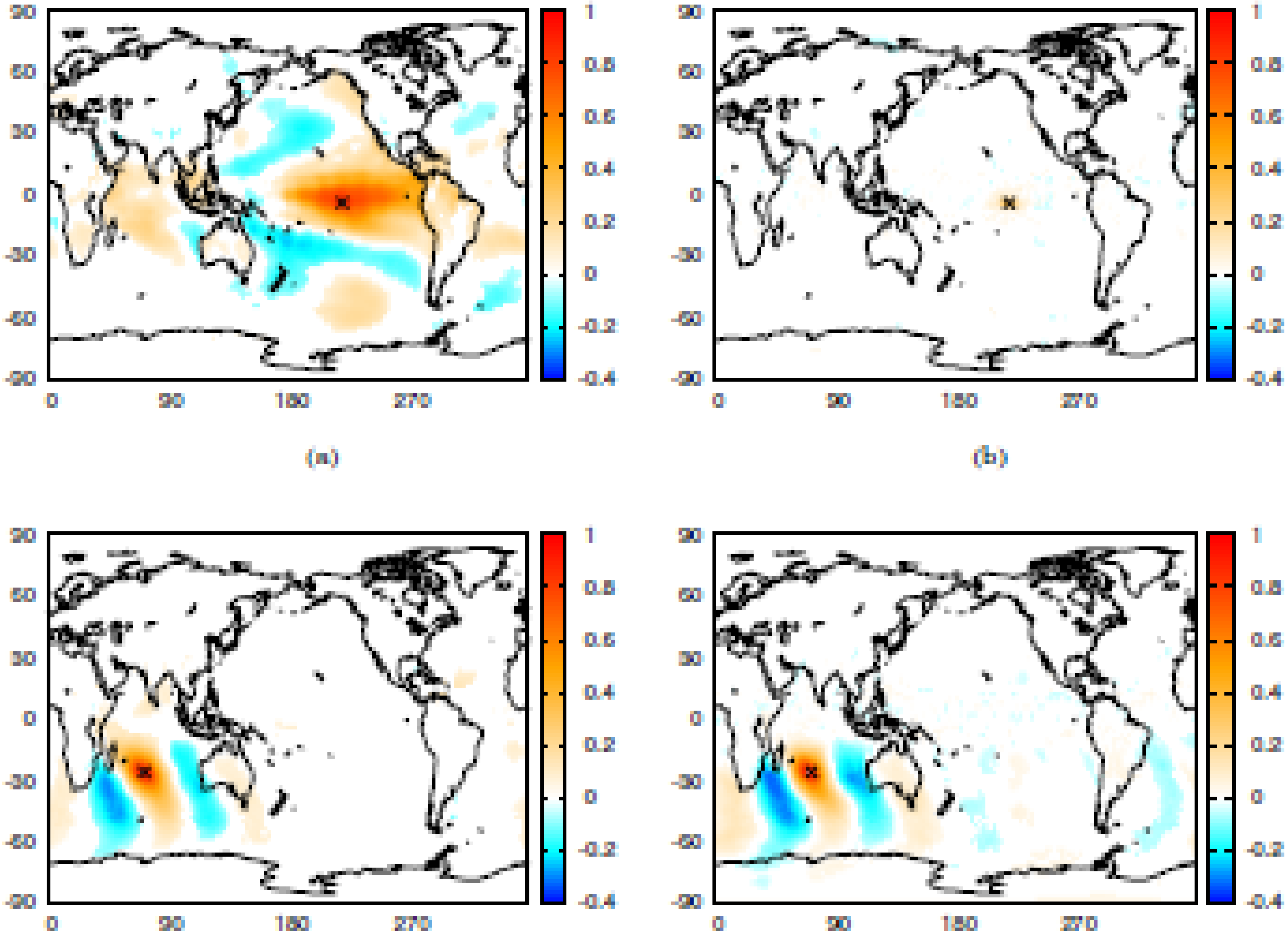
Network constructed from correlation analysis of SAT anomalies



Network constructed from correlation analysis of Hilbert frequencies



Why? Connectivity maps (cross-correlation)



(a)

(b)

What did we learn?

- Hilbert analysis uncovers inter-decadal changes in SAT daily time series.
- Large variations of Hilbert amplitude (more than 50%) in the Arctic and in Amazonia were interpreted respectively as due to ice melting and precipitation decrease.
- Hilbert frequency also uncovered areas of large changes (more than 100%).
- No filter used (Hilbert directly applied to data).
- Powerful tool to detect transitions, directly from the observed data.

Work done in collaboration with

At UPC

- Carlos Quinteros
- Dario Zappala
- Maria Masoliver
- J. M. Aparicio Reinoso
- Jordi Tiana
- M. C. Torrent

At Universidad de la Republica (Uruguay)

- Marcelo Barreiro

Predicting Transitions in Complex Systems (PRETRA 2018)

- **23-27 April 2018**
- To be held at the *Max Plank Institute for Complex Systems* (Dresden)
- Topics: data- and theory-driven approaches to detect, characterize and predict transitions (physical, biomedical, social and economic systems, geosciences, etc.)
- Organizers
  - Jaroslav Hlinka (Czech Academy of Sciences)
  - Klaus Lehnertz (University of Bonn, Germany)
  - CM (UPC)

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

THANK YOU FOR YOUR ATTENTION !

<cristina.masoller@upc.edu>

Papers at **http://www.fisica.edu.uy/~cris/**

- C. Quintero-Quiroz et al, "*Quantitative identification of dynamical transitions in a semiconductor laser with optical feedback*", Sci. Rep. 6, 37510 (2016).
- J. A. Reinoso et al, "*Emergence of spike correlations in periodically forced excitable systems*", Phys. Rev. E 94, 032218 (2016).
- D. A. Zappala et al, "*Hilbert analysis unveils interdecadal changes in large-scale patterns of surface air temperature variability*", submitted (2017).

Funding: ICREA

Agència de Gestió d'Ajuts Universitaris i de Recerca AGAUR

ITN LINC
ITN NETT

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN