

# Nonlinear time series analysis:

## Univariate analysis

Cristina Masoller

Universitat Politècnica de Catalunya, Terrassa, Barcelona, Spain

Cristina.masoller@upc.edu

[www.fisica.edu.uy/~cris](http://www.fisica.edu.uy/~cris)

[http://www.fisica.edu.uy/~cris/teaching/curitiba\\_masoller\\_uni.pdf](http://www.fisica.edu.uy/~cris/teaching/curitiba_masoller_uni.pdf)



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH

*Campus d'Excel·lència Internacional*

## ■ Introduction

- Historical developments: from dynamical systems to complex systems

## ■ Univariate analysis

- Methods to extract information from a time series.
- Applications.

## ■ Bivariate analysis

- Correlation, directionality and causality.
- Applications.

## ■ Multivariate analysis

- Many time series: complex networks.
- Network characterization and analysis.
- Applications.

# Methods of time-series analysis

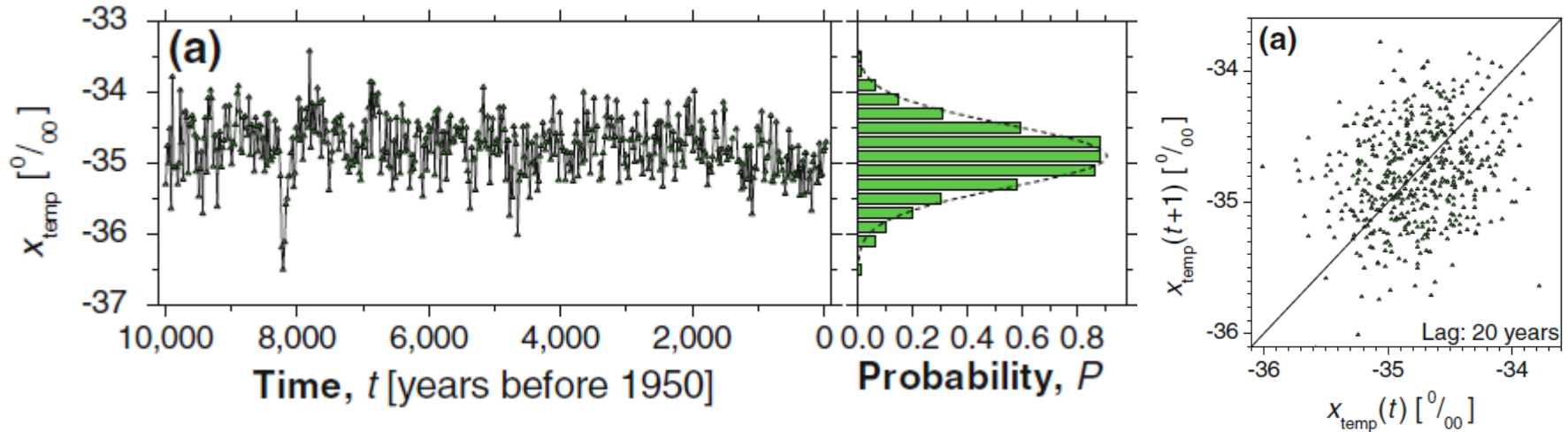
- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
- Spatio-temporal representation of a time-series
- Instantaneous phase and amplitude

# To begin with the analysis of a time-series

$$X = \{x_1, x_2, \dots, x_N\}$$

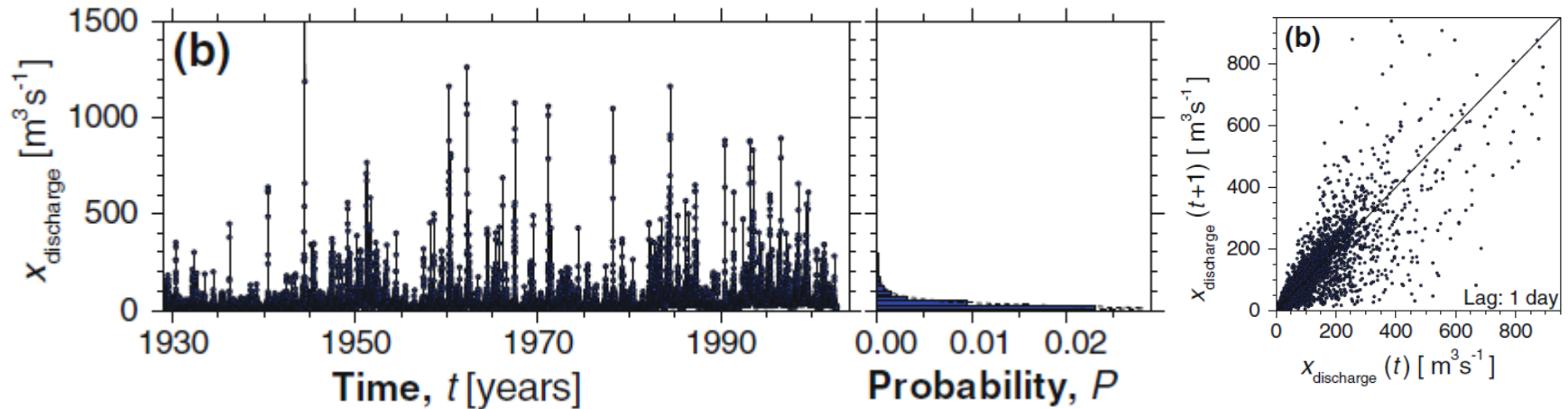
- First step: Look at the data.
- Examine simple properties:
  - Return map: plot of  $x_i$  vs.  $x_{i+\tau}$
  - Distribution of data values
  - Auto-correlation
  - Fourier spectrum

# First example of a geophysical time series



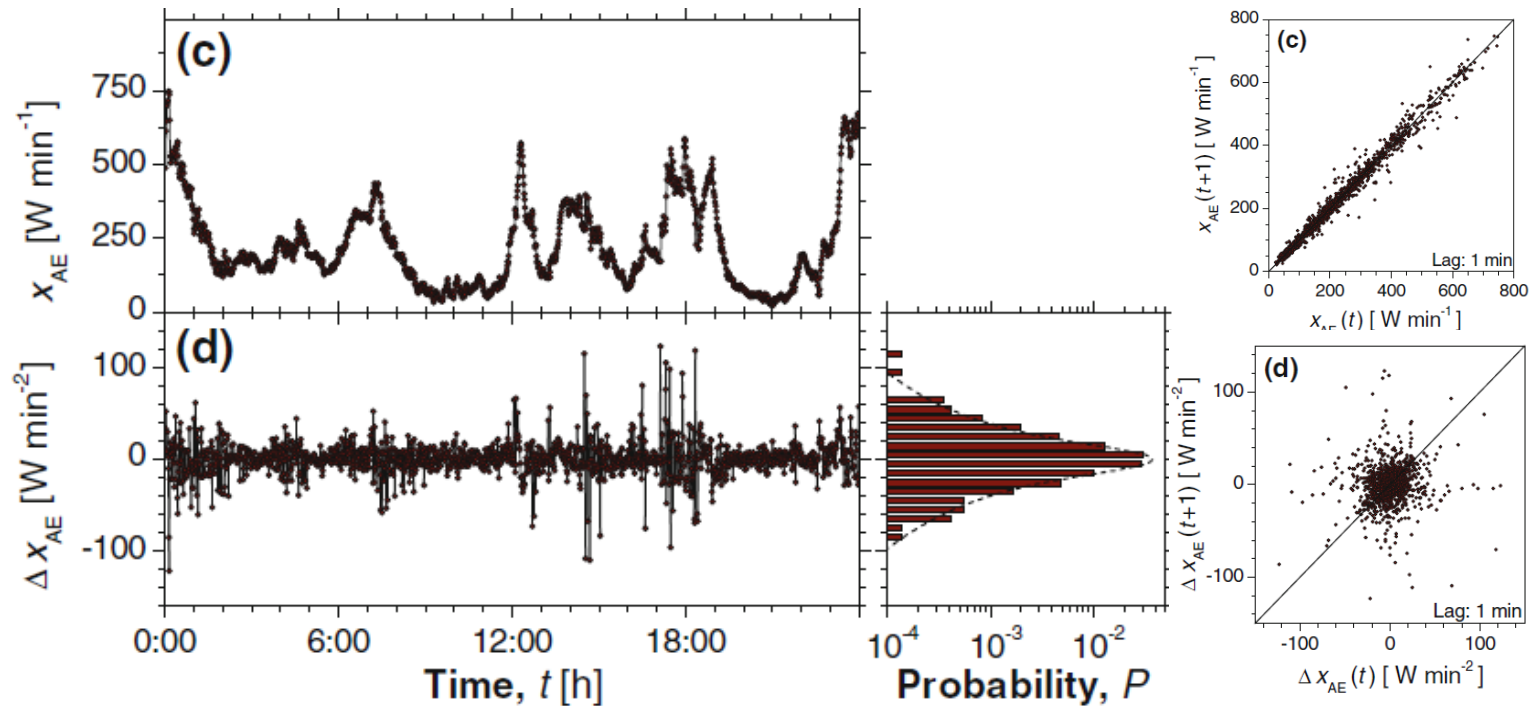
Bi-decadal oxygen isotope data set d18O (proxy for palaeotemperature) from Greenland Ice Sheet Project Two (GISP2) for the last 10,000 years with 500 values given at 20 year intervals.

## Second example



Discharge of the Elkhorn river (at Waterloo, Nebraska, USA) sampled daily for the period from 01 January 1929 to 30 December 2001.

# Third example



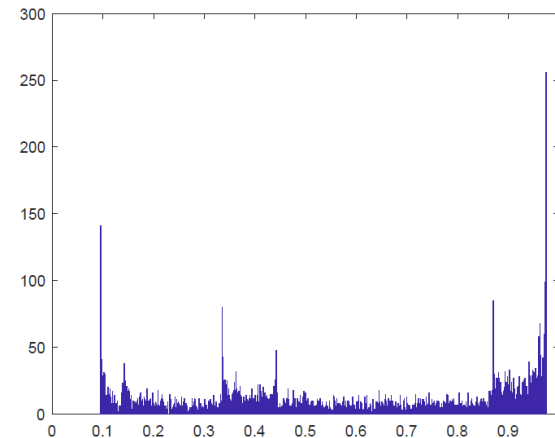
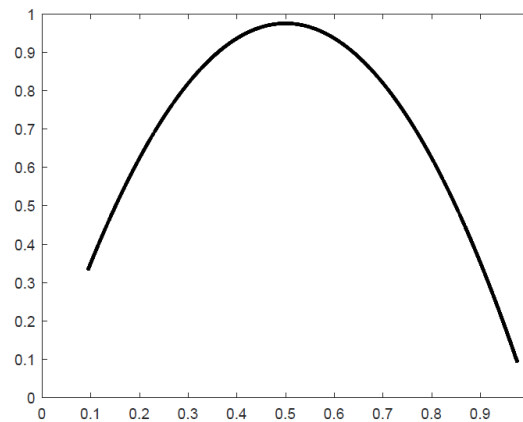
The geomagnetic auroral electrojet (AE) index sampled per minute for the 24 h period of 01 February 1978 and the differenced index:  $\Delta x_{AE}(t) = x_{AE}(t) - x_{AE}(t - 1)$

# Exercise 1: The Logistic map

Calculate the return maps and the distribution of data values for different values of the control parameter,  $r \in [3, 4]$ , starting from a random initial condition,  $x(1) \in (0, 1)$ .

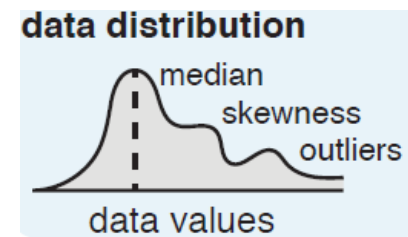
$$x(i + 1) = r x(i)[1 - x(i)]$$

$r=3.9$





# How to characterize a distribution of data values?



- **Mean** (expected value of  $X$ ):  $\mu = E[X] = \sum_{i=1}^k x_i p_i = x_1 p_1 + x_2 p_2 + \dots + x_k p_k$   
 $\mu = \langle x(t) \rangle$
- **Variance**:  $\sigma^2 = \text{Var}(X) = E[(X - \mu)^2]$

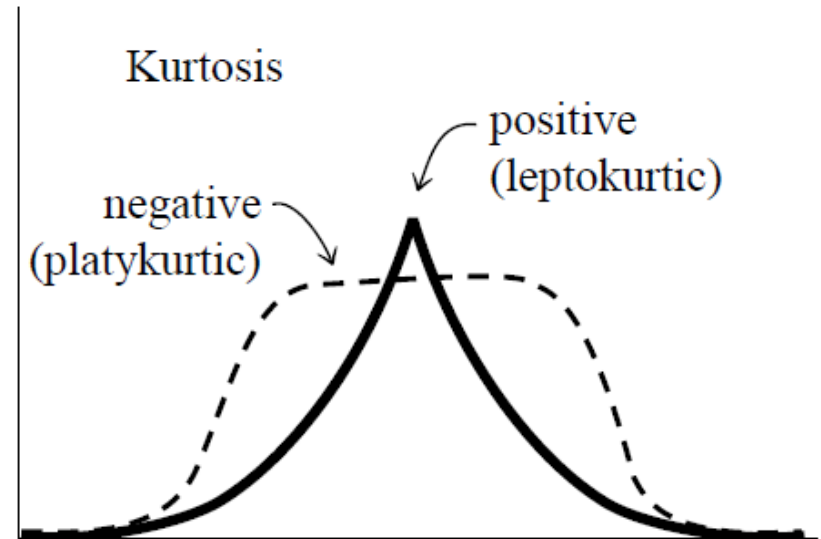
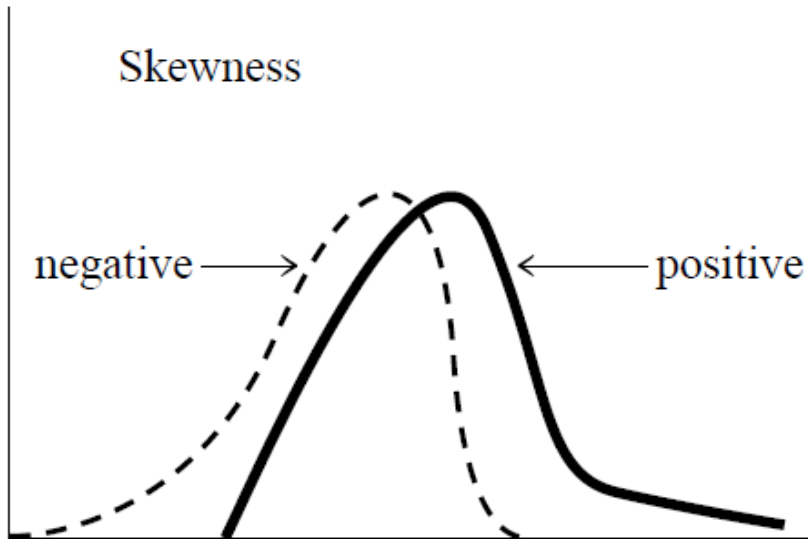
- **Skewness**: “measures” the asymmetry of the distribution

$$Z = \frac{X - \mu}{\sigma} \quad S = E[Z^3]$$

- **Kurtosis**: measures the “tailedness” of the distribution. For a normal distribution  $K=3$ .

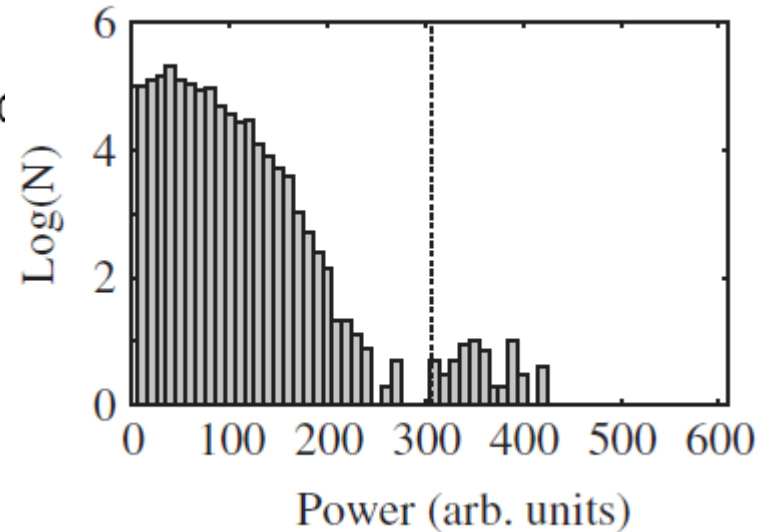
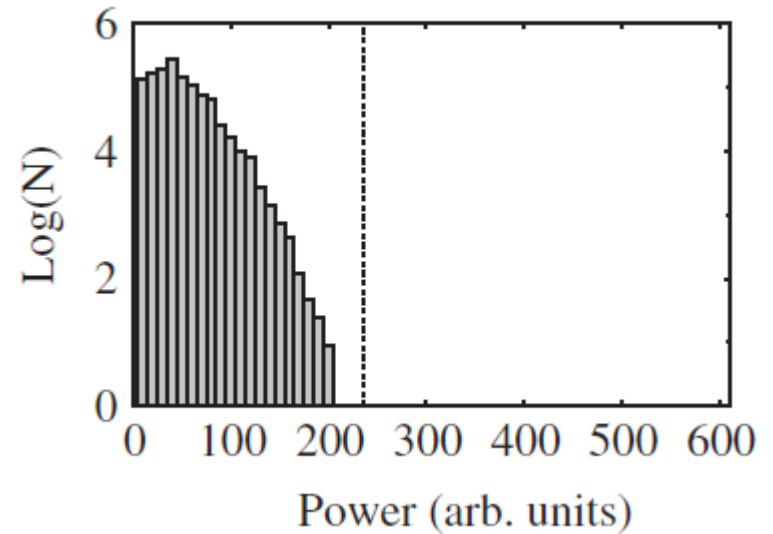
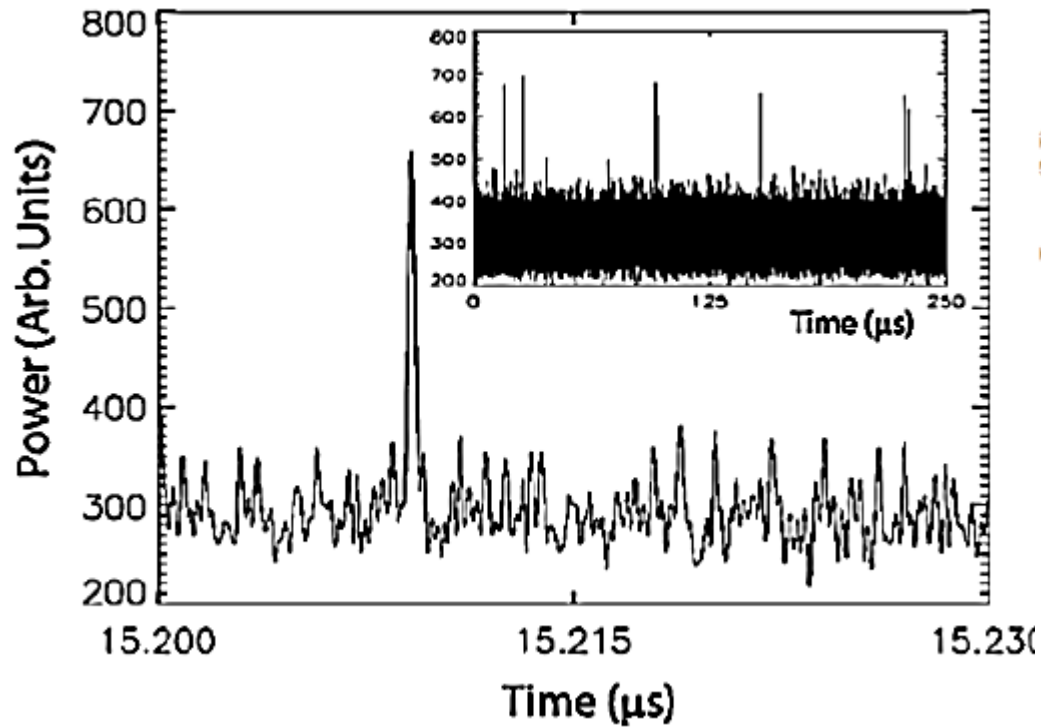
$$K = E[Z^4]$$

- **Coefficient of variation**: normalized measure of the width of the distribution.  
 $C_v = \sigma / |\mu|$



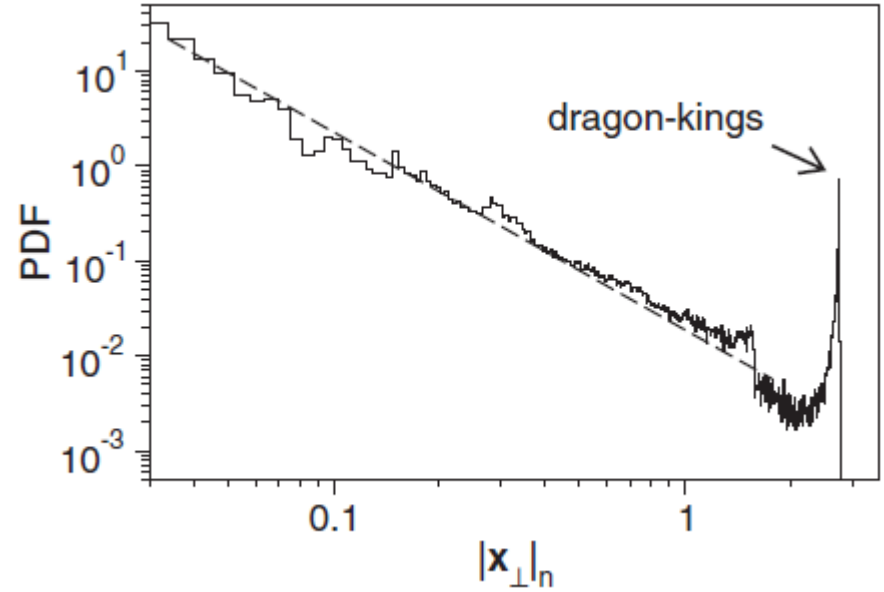
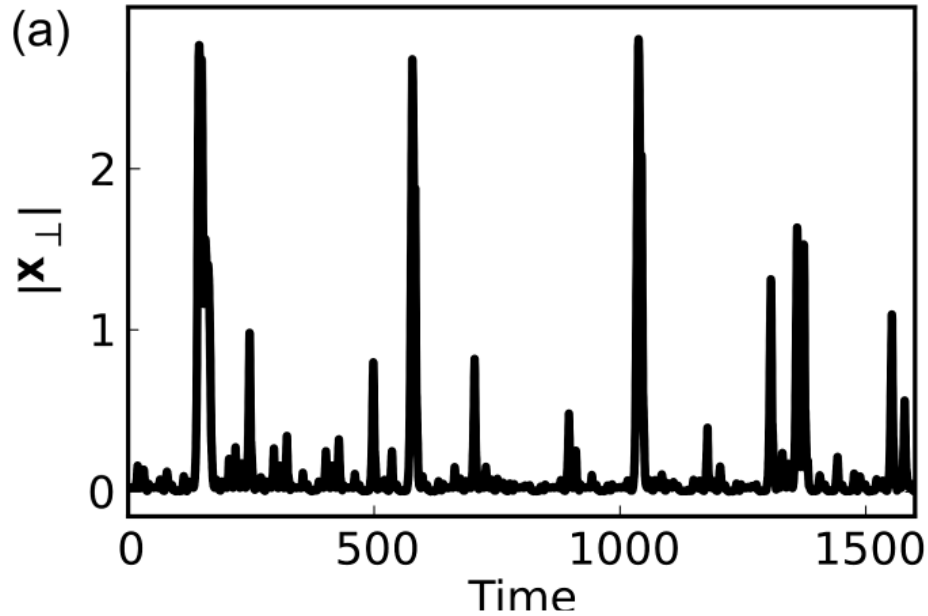
Source: Press WH et al. *Numerical recipes: the art of scientific computing* (Cambridge University Press)

# Long tailed distribution? Outliers?



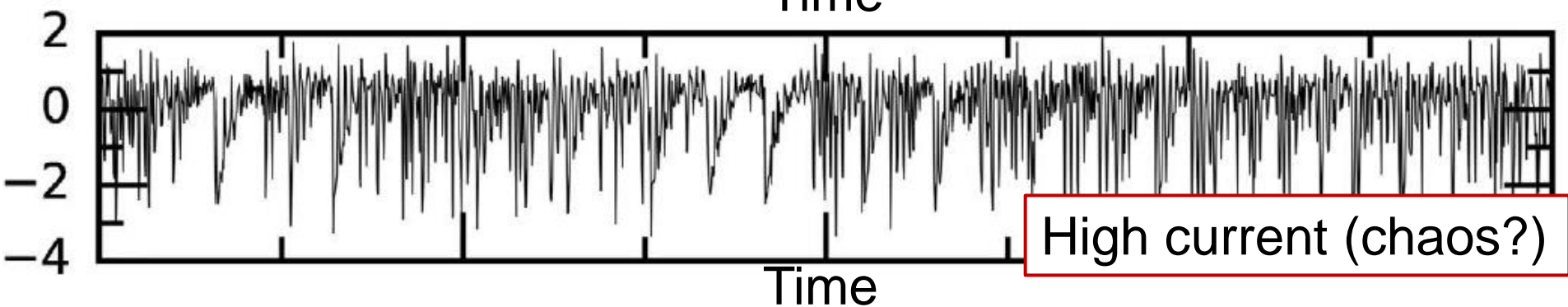
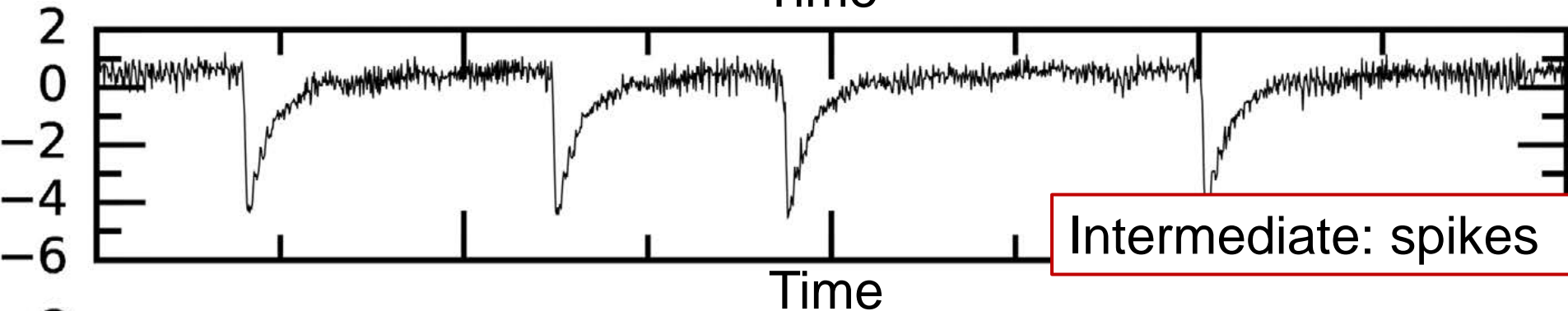
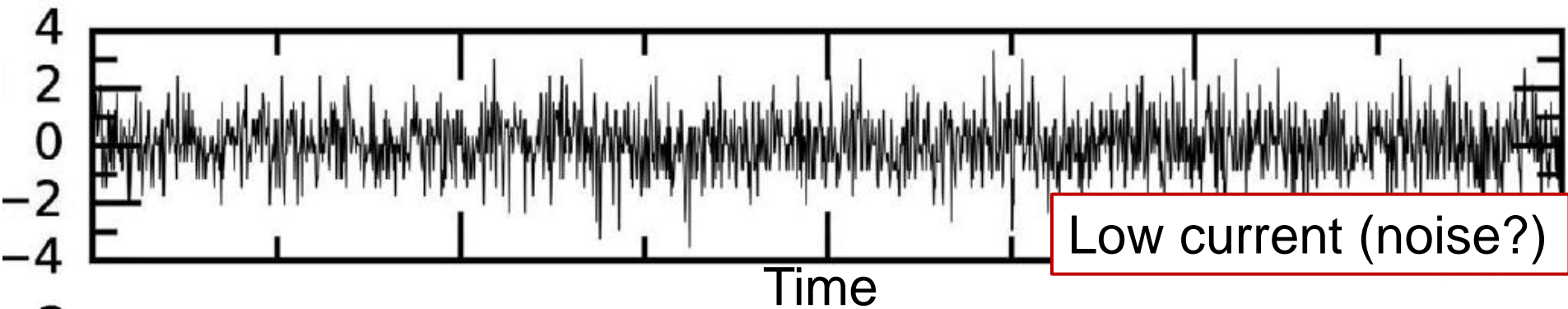
Bonatto et al. PRL 107, 053901 (2011)

# “Dragon kings”: extreme outliers

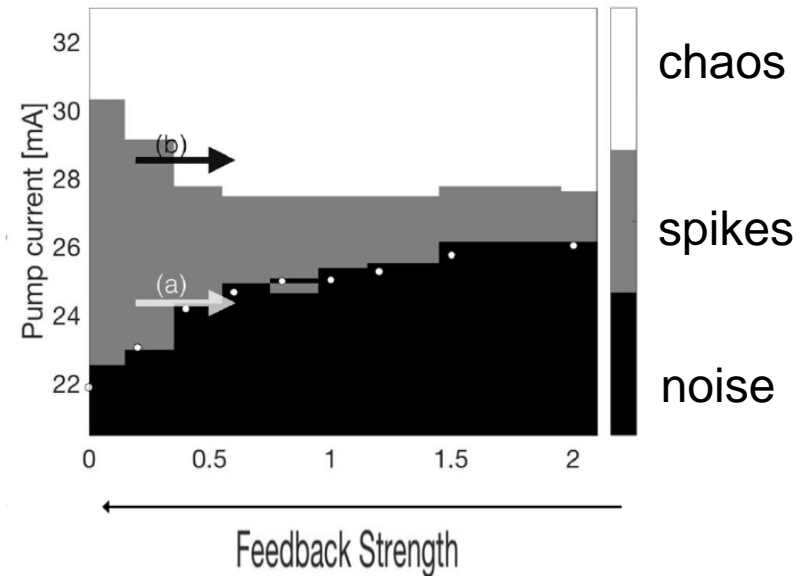
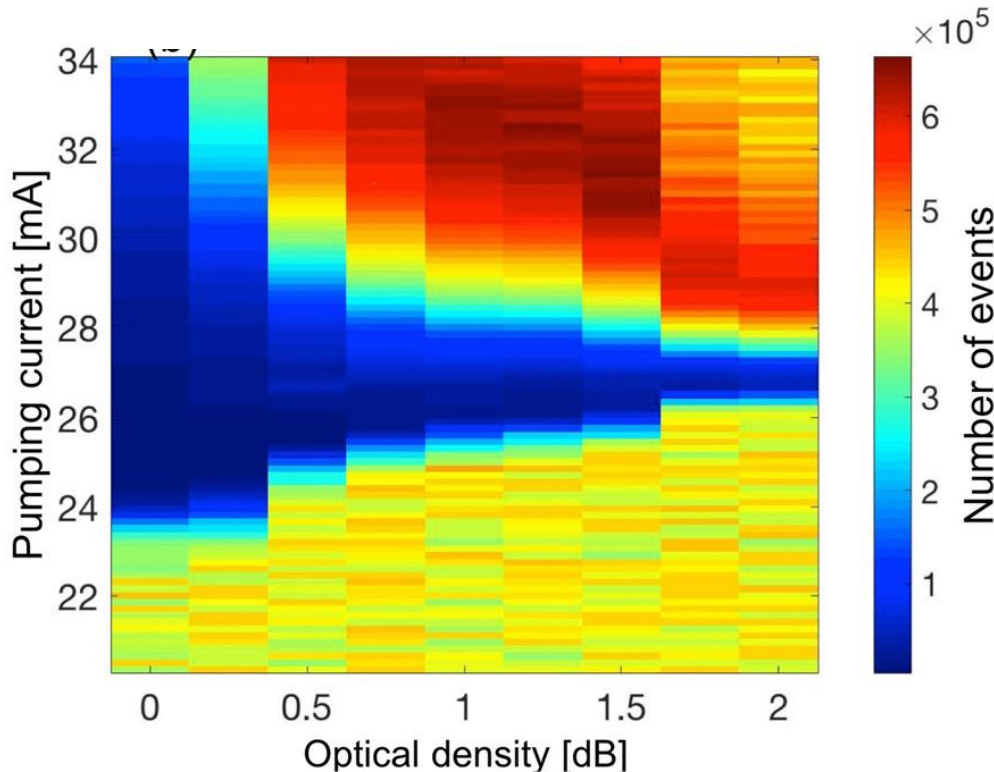
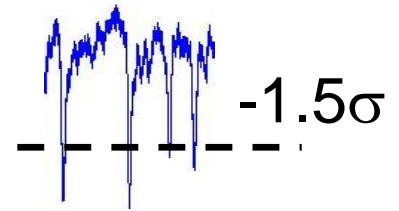


Hugo L. D. de S. Cavalcante et al PRL 111, 198701 (2013)

# Example: intensity emitted by a diode laser with feedback, as the pump current increases



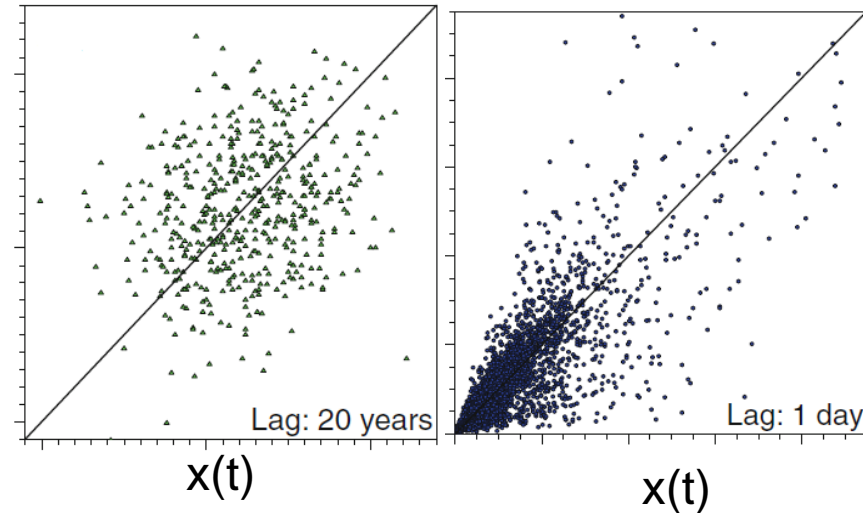
By counting the number of times the intensity falls below a threshold we can distinguish the three regimes



# Autocorrelation function (ACF)

- The return map uncovers correlations.
- How to quantify?

$x(t+\tau)$



- The autocorrelation function

$$C(\tau) = \frac{\langle [x(t) - \mu][x(t - \tau) - \mu] \rangle}{\sigma^2}$$

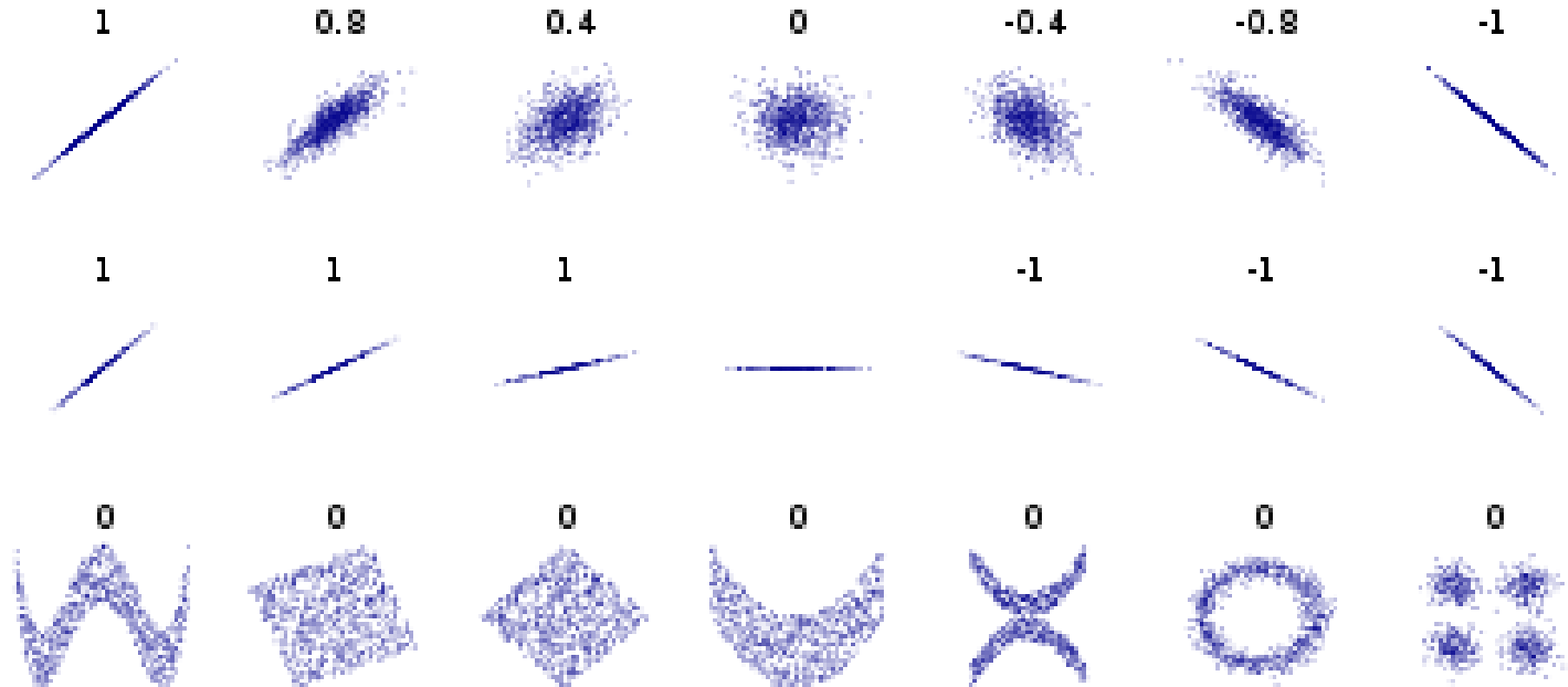
$$C(0) = 1$$

- **Normalization:**  $\mu=0, \sigma=1$        $C(\tau) = \langle x(t)x(t - \tau) \rangle$

- For a stationary process:  $C(-\tau) = C(\tau)$

( $\mu$  and  $\sigma$  are constant in time)

# Auto-correlation detects linear relationships only

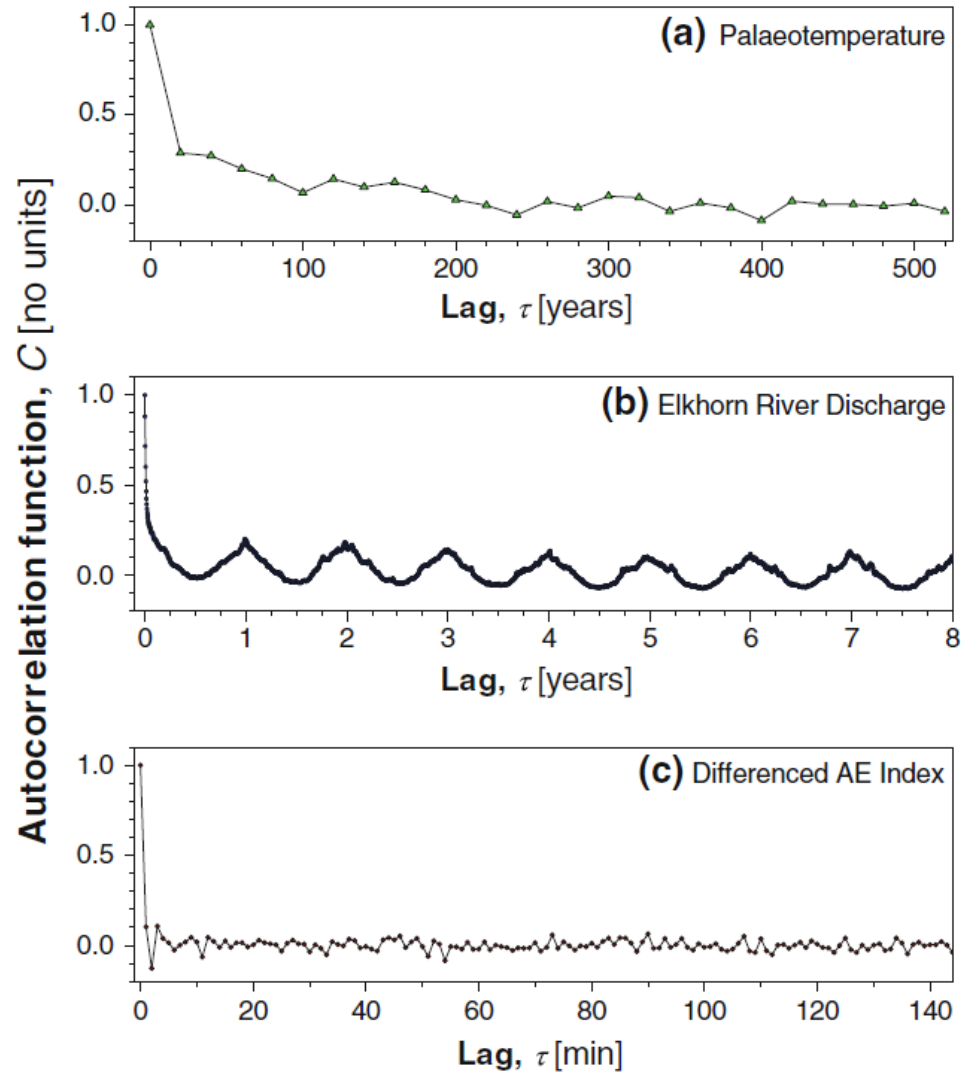




# Correlation properties

- Persistence: large values tend to follow large ones, and small values tend to follow small ones.
- Anti-persistence: large values tend to follow small ones and small values large ones.
- Short-range correlations: values are correlated with one another at short lags in time  $|C(\tau)| \leq \kappa_0 \exp(-\kappa\tau)$
- Long-range correlations: values are correlated with one another at very long lags in time (all or almost all values are correlated with one another)

# Back to the three examples of geophysical time series

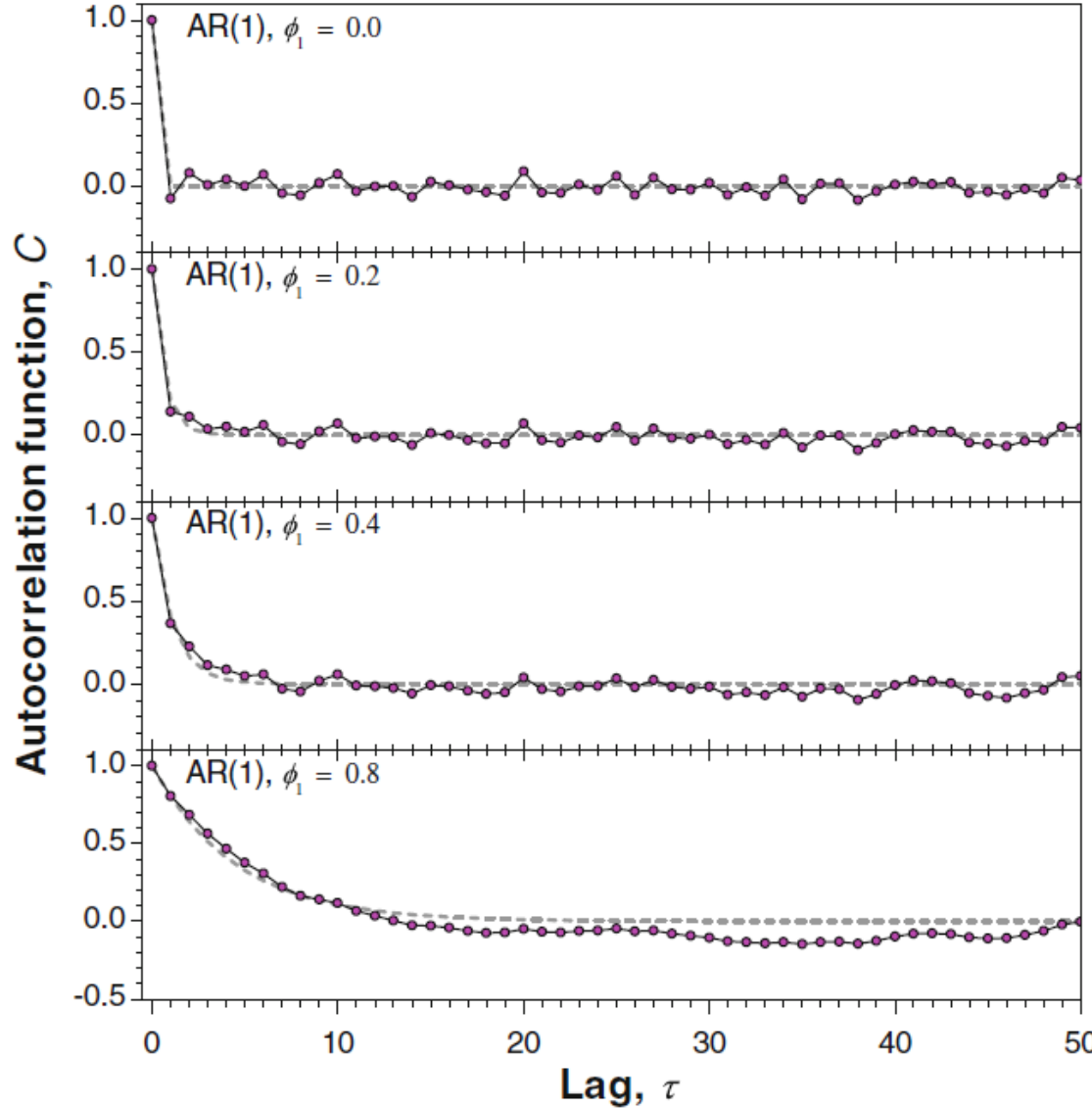


# Synthetic model: first-order autoregressive (AR(1)) process

$$x_t = \phi_1 x_{t-1} + \varepsilon_t$$

$-1 < \phi_1 < 1$

White noise  
(uncorrelated  
random  
numbers,  
 $\mu=0, \sigma=1$ )



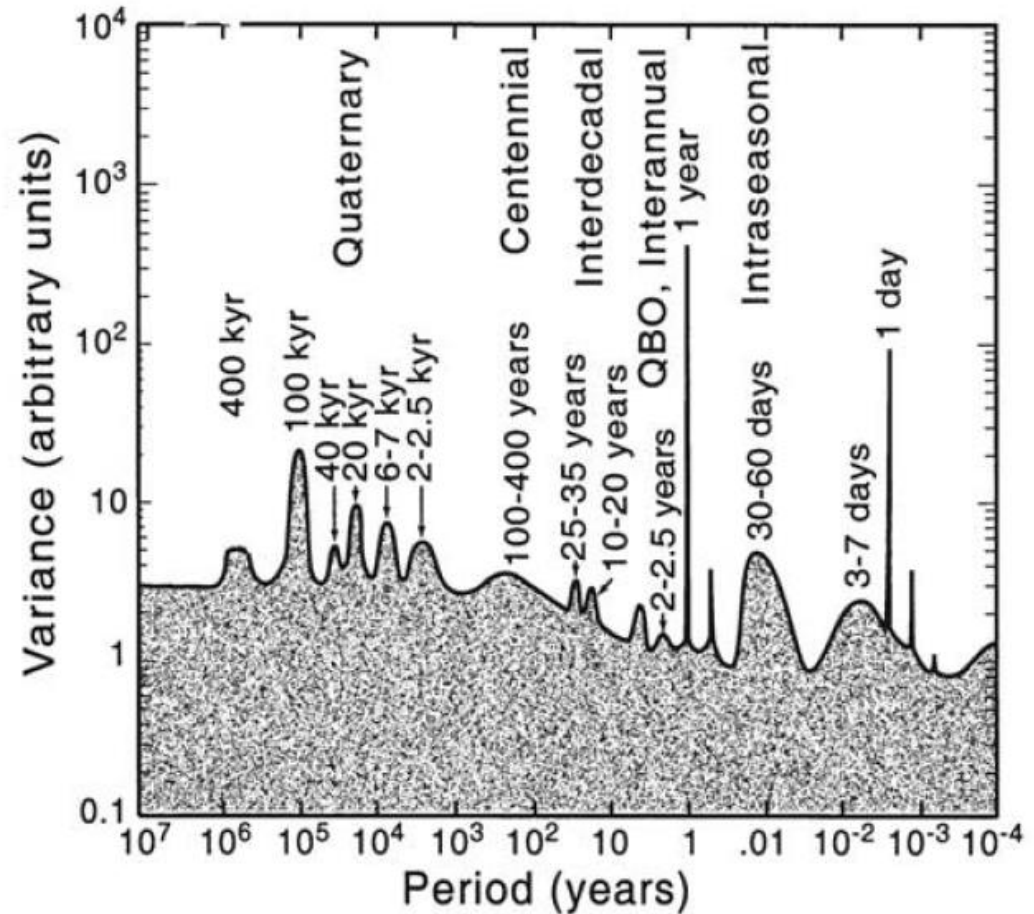
A. Witt and B. D. Malamud,  
Surv Geophys (2013) 34:541–651

In matlab `r = normrnd(mu,sigma)` generates a random number from the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .

# Fourier analysis

- Describes a signal in terms of its oscillatory components.
- Example: climatic time series involve oscillations in a wide range of scales
  - hours to days,
  - months to seasons,
  - decades to centuries,
  - and even longer...

An “artist’s representation” of the power spectrum of climate variability (Ghil 2002).



# The Power Spectrum (or power spectral density, PSD)

- Discrete Fourier Transform

$$\Delta = \text{sampling interval} \quad X_k = \Delta \sum_{t=1}^N x_t e^{2\pi i t k / N}, \quad k = 1, 2, \dots, N$$

N = number of data points

$X_k$  = complex Fourier coeff. of freq.  $f_k = k/(N\Delta)$

- The PSD gives the intensity of each frequency component:

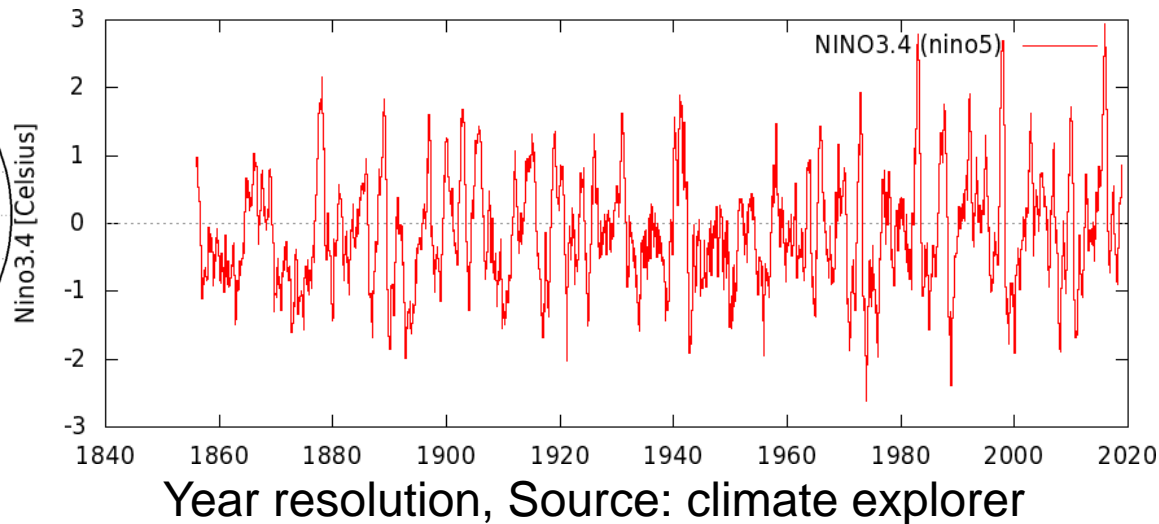
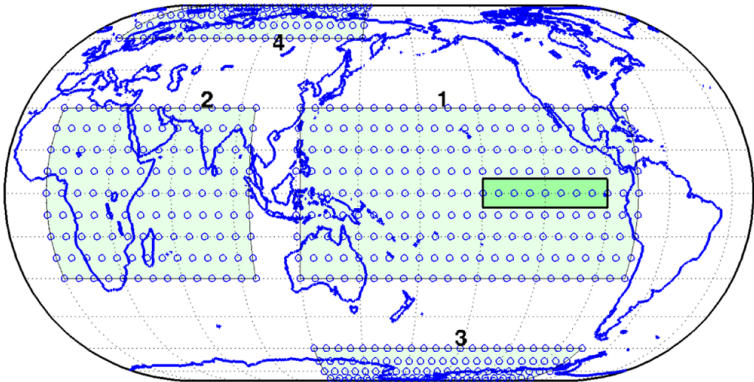
$$S_k = \frac{2|X_k|^2}{N\Delta}, \quad k = 1, 2, \dots, \frac{N}{2}$$

- For stationary signals the inverse FT of the PSD is  $|C(\tau)|$  (Wiener–Khinchin theorem).
- How to calculate? Fast Fourier Transform (FFT).  $N=2m$ .
- With Matlab:  $z=\text{fft}(x)$ ;  $q=z.\text{conj}(z)$ ;  $v=\text{ifft}(q)$ .

Further reading: Press WH et al. *Numerical recipes in C/Fortran: the art of scientific computing* (Cambridge University Press)

# Example: El Niño 3.4 index

## Sea surface temperature (SST) anomaly in the Eastern tropical Pacific Ocean

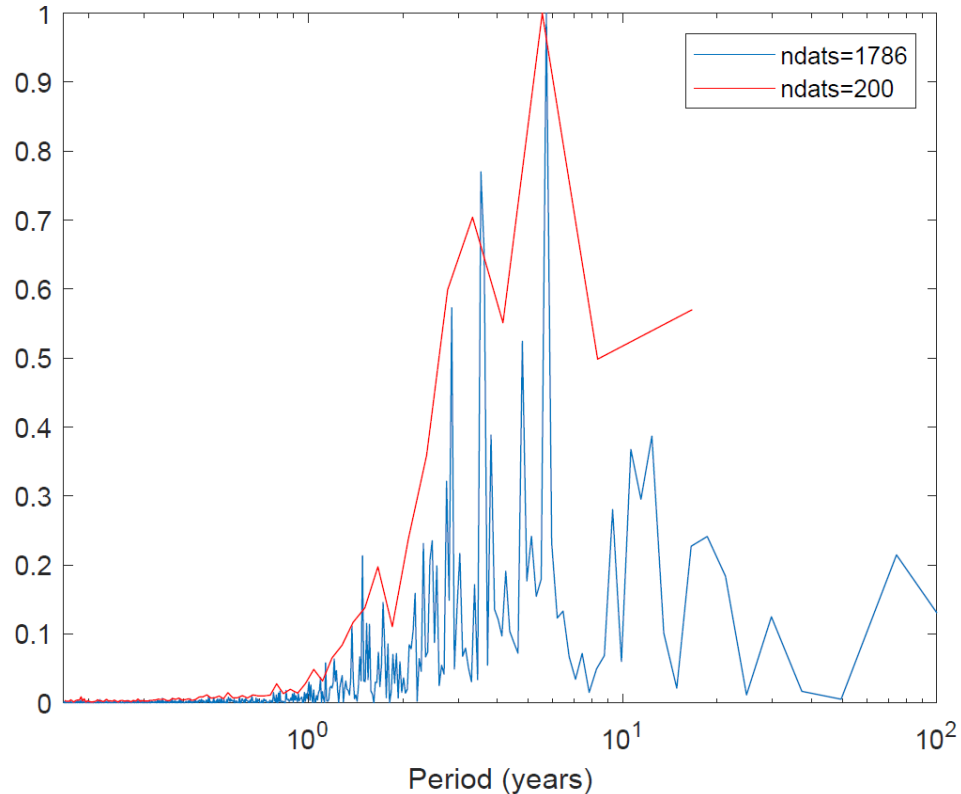


- From 5°N to 5°S and from 170°W to 120°W
- An El Niño (La Niña) event is identified if the 5-month running-average of the NINO3.4 index exceeds +0.4°C (-0.4°C) for at least 6 consecutive months.

# “Periodogram” of NINO3.4 index

dataset:

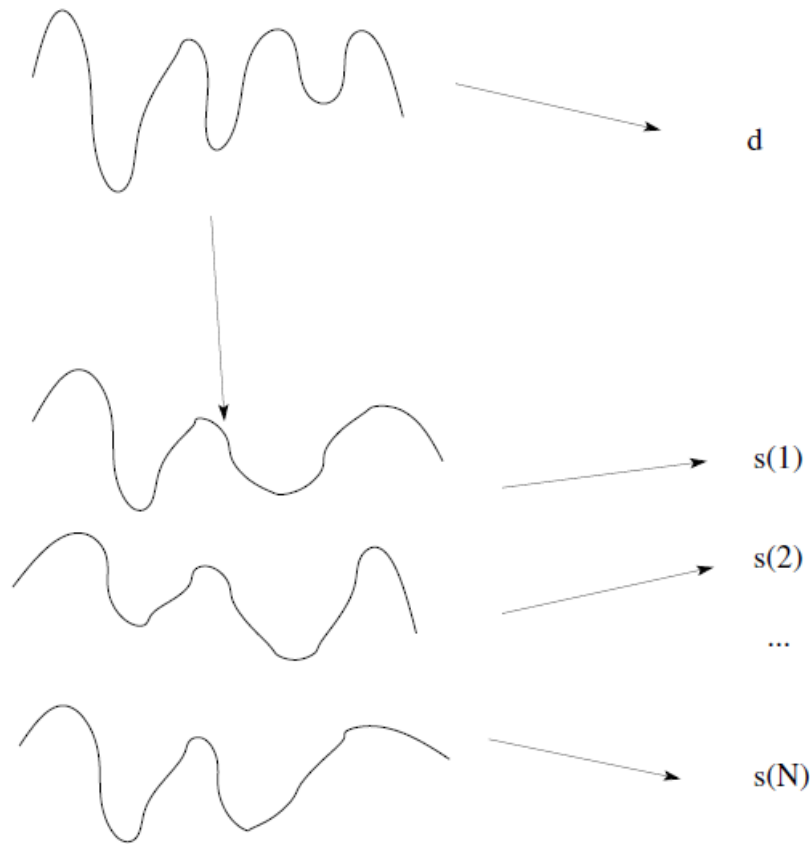
[https://www.esrl.noaa.gov/psd/gcos\\_wgsp/Timeseries/Data/nino34.long.anom.data](https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino34.long.anom.data)



Peaks at 5.7 and 3.5 years

**Are these peaks  
“significant”?**

# The method of surrogate data

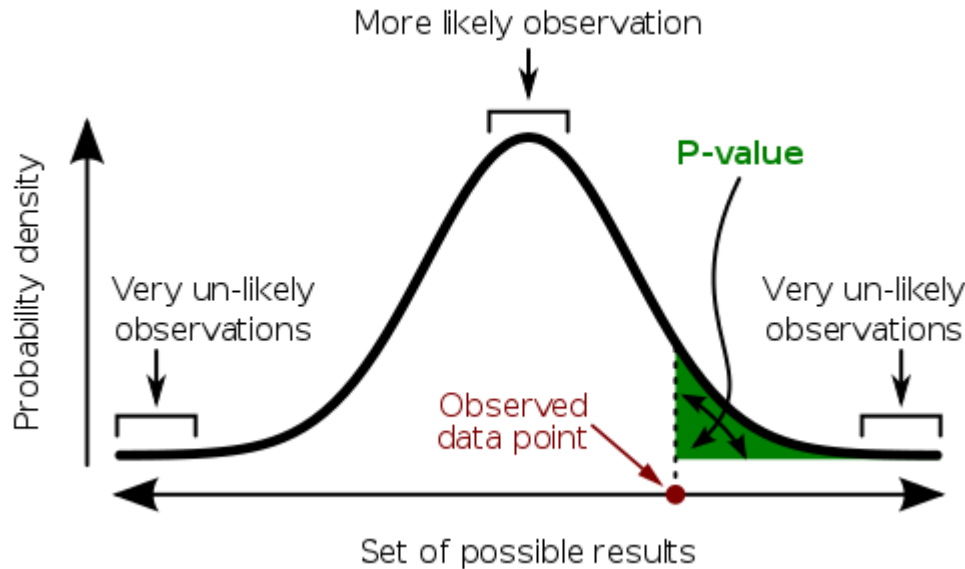


Taken from  
M. Small, *Applied Nonlinear Time  
Series Analysis* (World Scientific,  
2005)

- Real observed time series.
- Generate an ensemble of “surrogate” time series that are both “similar” to the original and also consistent with the specific **null hypothesis** (NH) that we want to test.
- Measure an statistical property: “ $d$ ” in the original series and “ $s(i)$ ” in the ensemble time series.
- Is “ $d$ ” consistent with the distribution of “ $s(i)$ ” values?
  - No! we **reject** the NH.
  - Yes! we **“fail to reject”** the NH.



# p value

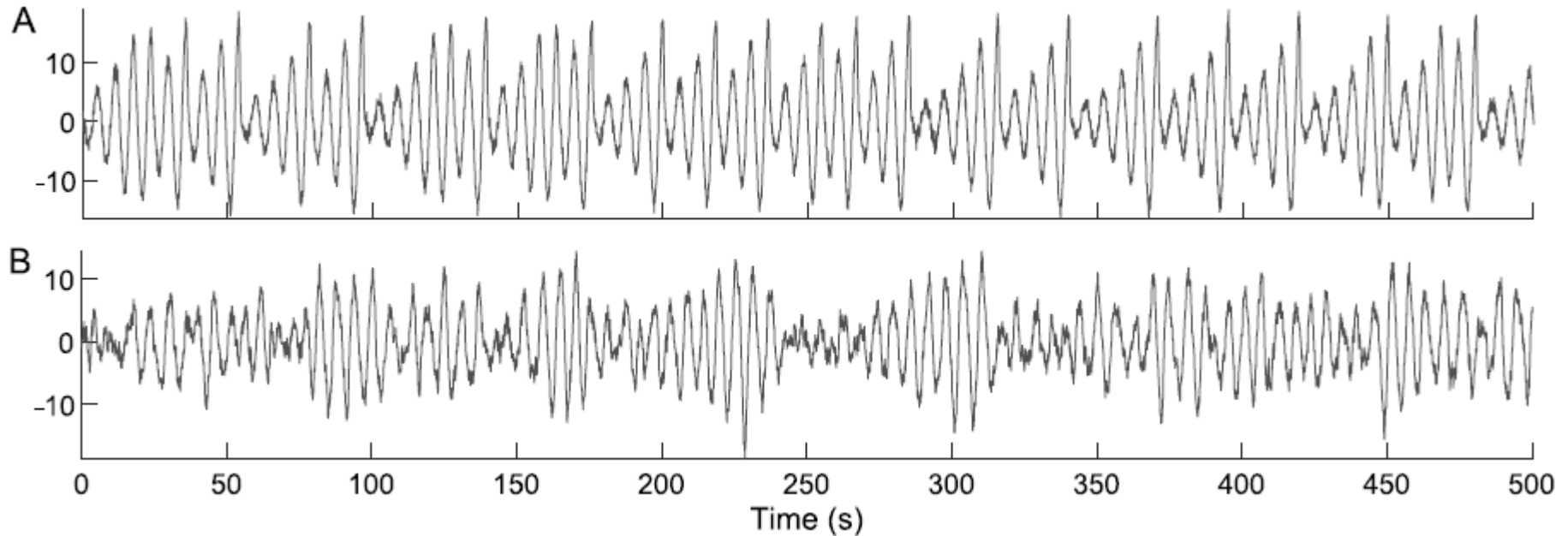


Source: wikipedia

The p-value only measures the compatibility of an observation with a hypothesis, not the truth of the hypothesis.

Altman, N. and Krzywinski, M. *Interpreting P values*. Nature Methods 14, 213 (2017).

# Surrogate test for nonlinearity



A: Rossler with  
 $a = 0.165$ ,  $b = 0.2$   
and  $c = 10$

$$\dot{x} = -y - z$$

$$\dot{y} = x + ay$$

$$\dot{z} = b + z(x - c),$$

B: High-order (linear) autoregressive process

$$s_n = 1.625s_{n-1} + 0.284s_{n-2} + 0.355s_{n-3} + \eta_n - 0.960\eta_{n-1}$$

Proper surrogates detect nonlinearity in  
A (reject NH) but not in B (fail to reject NH)

Further reading: G. Lancaster et al, “*Surrogate data for hypothesis testing of physical systems*”, Physics Reports 748 (2018) 1–60.

## **Examples of applications:**

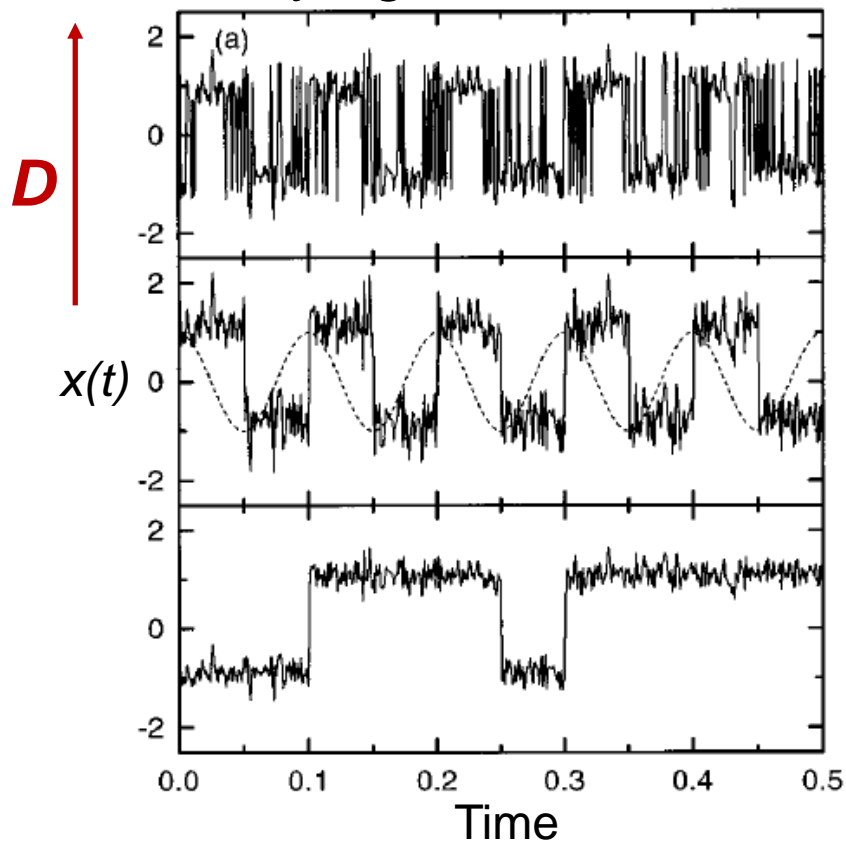
- Stochastic resonance in bistable systems**
- Coherence resonance in excitable systems**

# Bistable system with sinusoidal forcing and noise

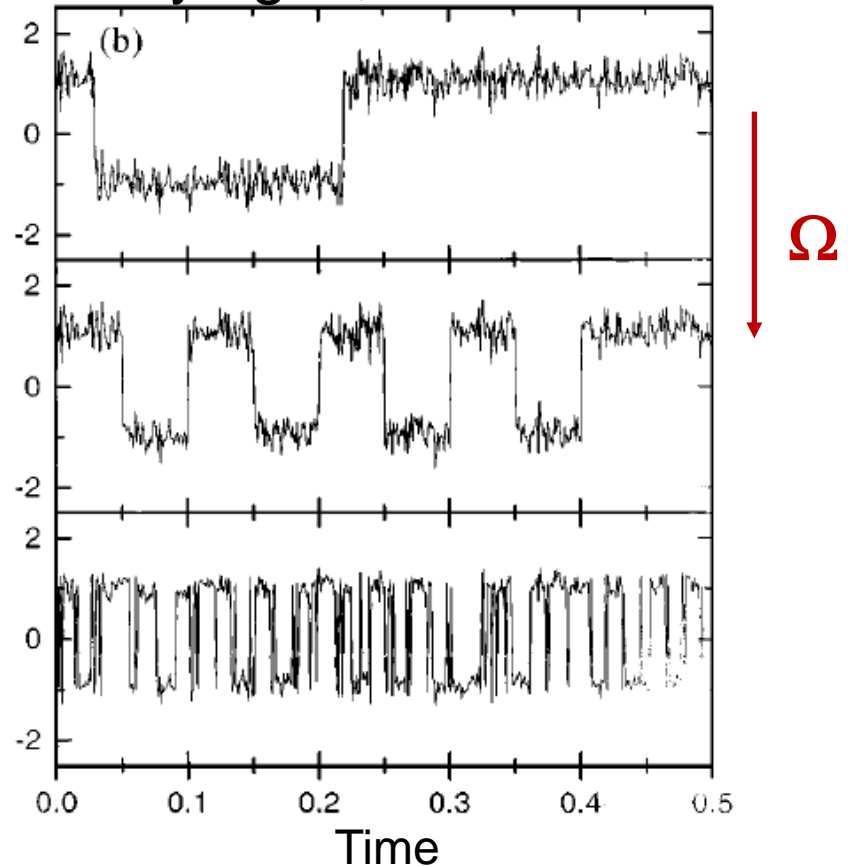
$$\dot{x}(t) = -V'(x) + A_0 \cos(\Omega t + \varphi) + \xi(t)$$

$$V(x) = -\frac{a}{2} x^2 + \frac{b}{4} x^4 \quad \langle \xi(t) \xi(0) \rangle = 2D \delta(t)$$

Varying  $D$ ;  $\Omega$  constant

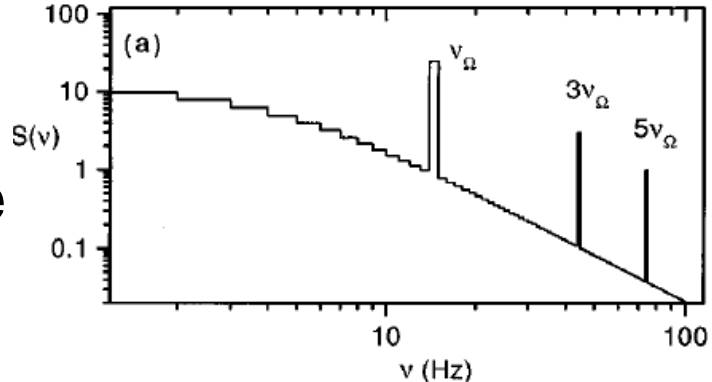


Varying  $\Omega$ ;  $D$  constant



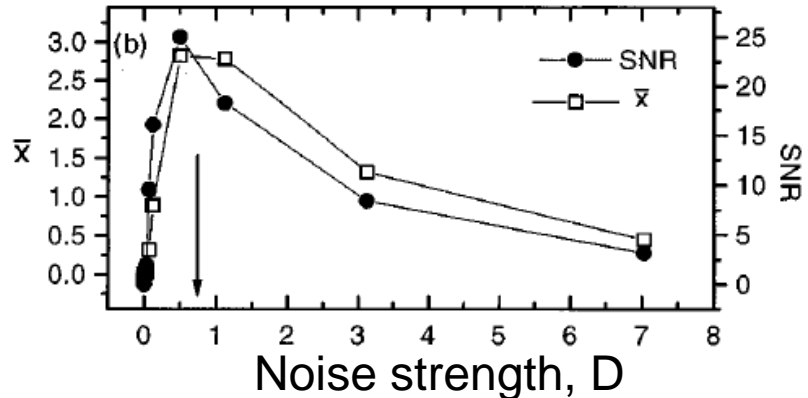
# Quantification of stochastic resonance with spectral analysis

- Phase-averaged power spectral density: average over many realizations of the noise and average over the input initial phase  $\varphi$ .

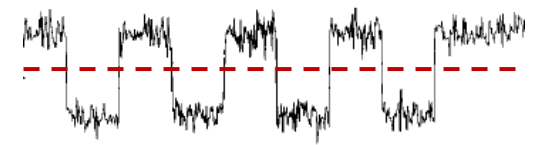


- Signal to noise ratio at  $\Omega$

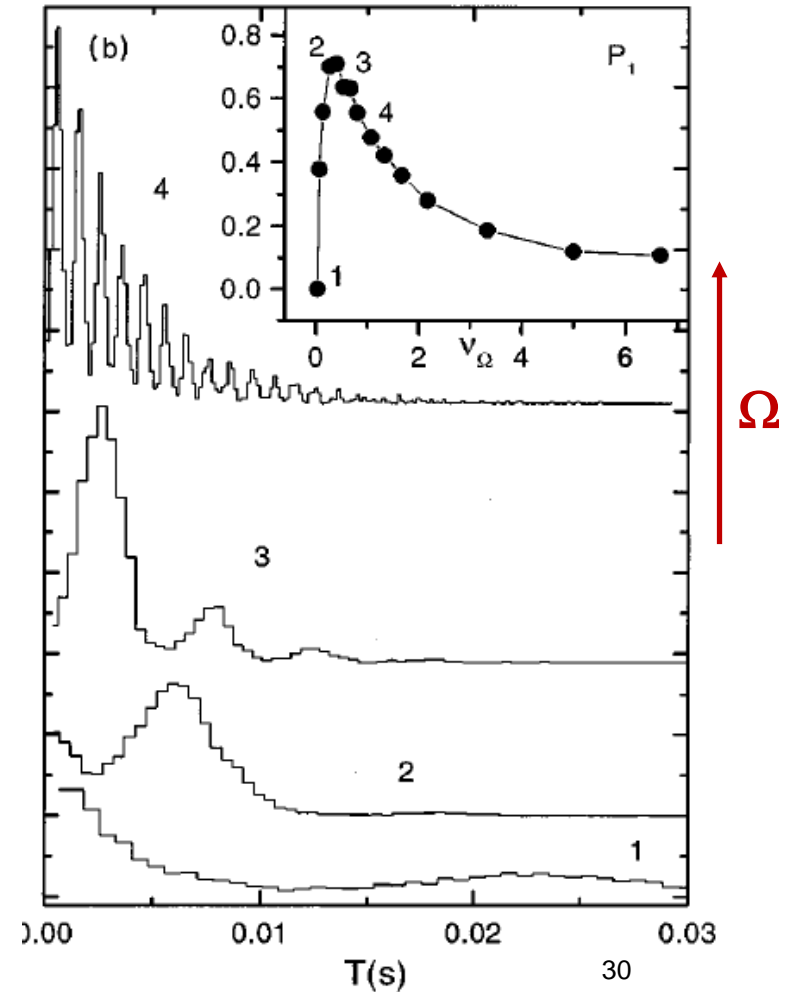
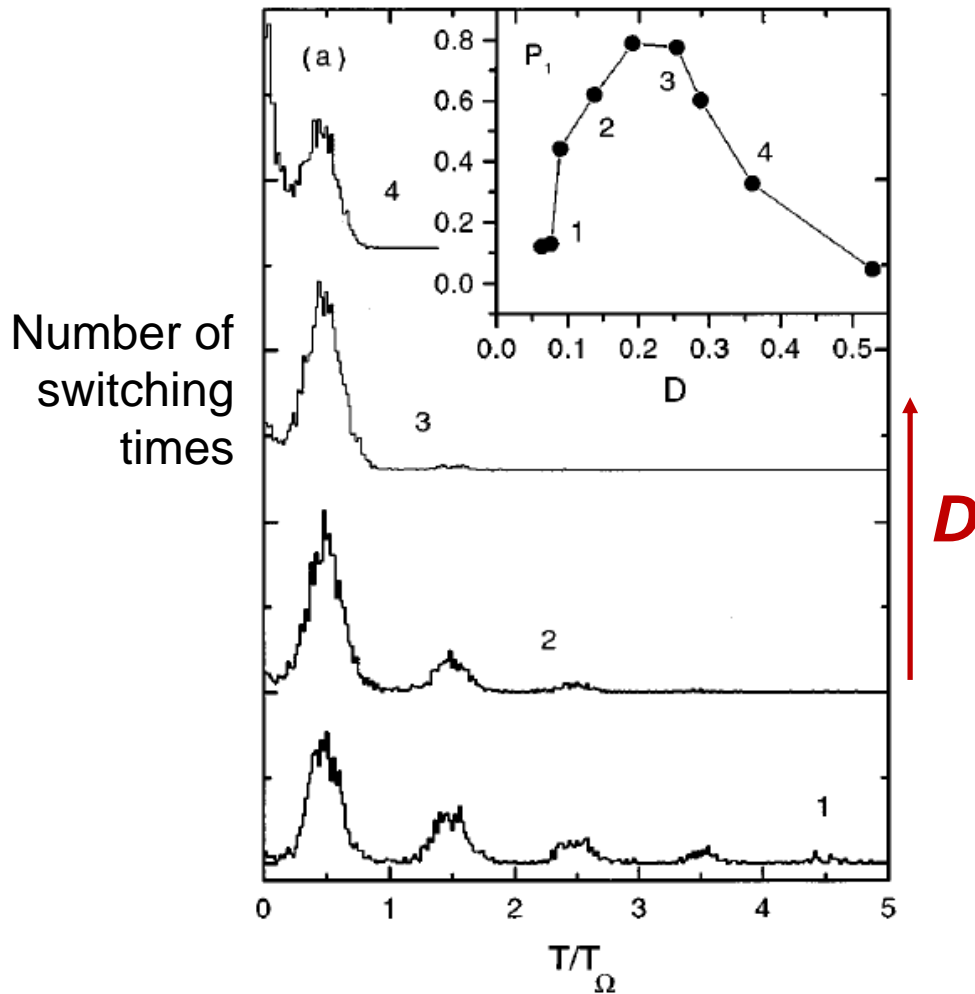
$$SNR(\Omega) = 2 \left[ \lim_{\Delta\omega \rightarrow 0} \int_{\Omega - \Delta\omega}^{\Omega + \Delta\omega} S(\omega) d\omega \right] / S_N(\Omega)$$



# Quantification of stochastic resonance with the *distribution of residence times*



Strength of the  $n$ th peak: area under the peak  $P_n = \int_{T_n - \alpha T_\Omega}^{T_n + \alpha T_\Omega} N(T) dT \quad 0 < \alpha \leq \frac{1}{4}$



# Excitable systems

- Nonlinear response to a perturbation:
  - Above a threshold, rapid and large variation.
  - Below the threshold, small change
- Refractory period: cannot support another event until a certain amount of time has passed
- Examples: neuronal spikes, cardiac beats, harmful algal blooms

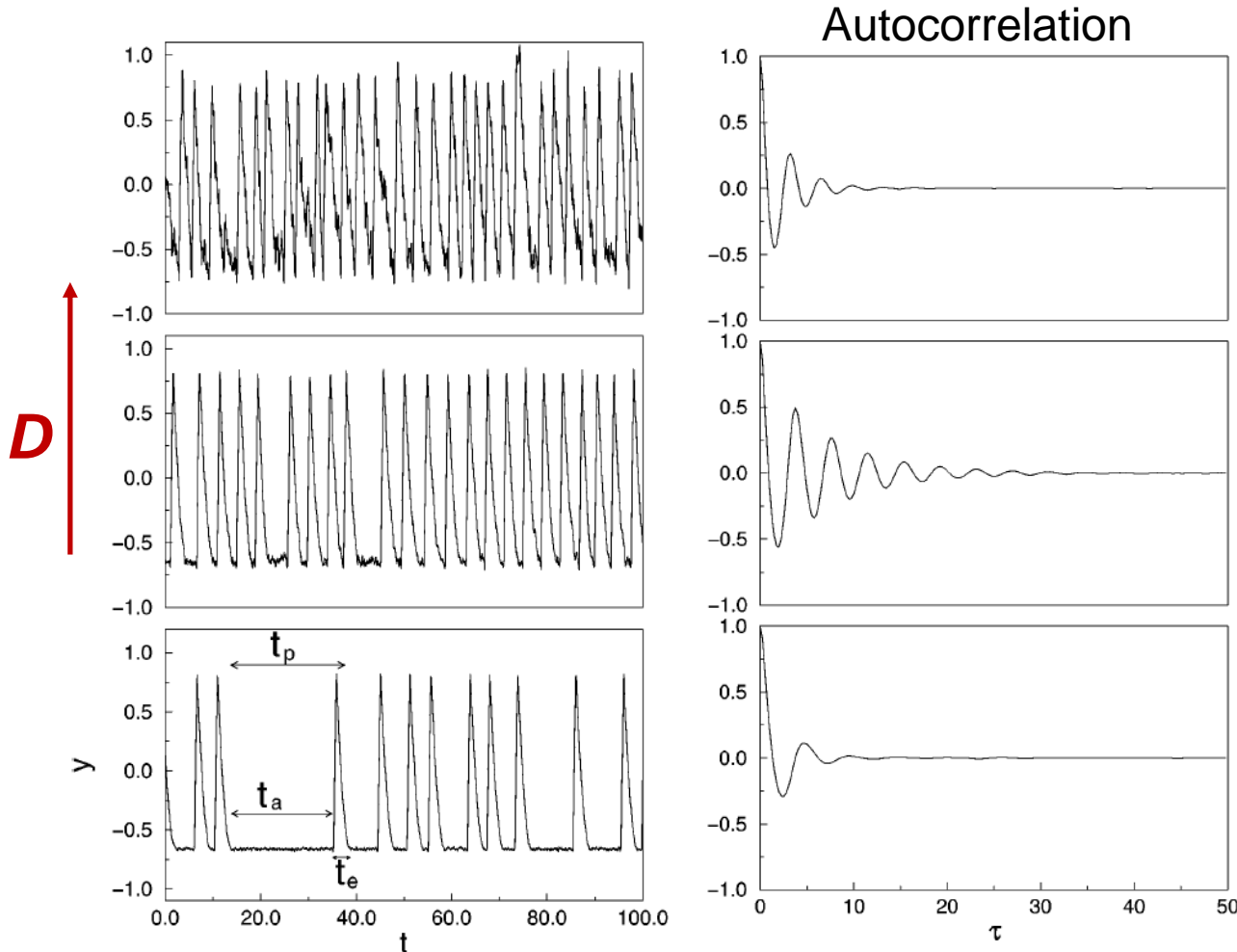


# Simple model of an excitable system

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

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

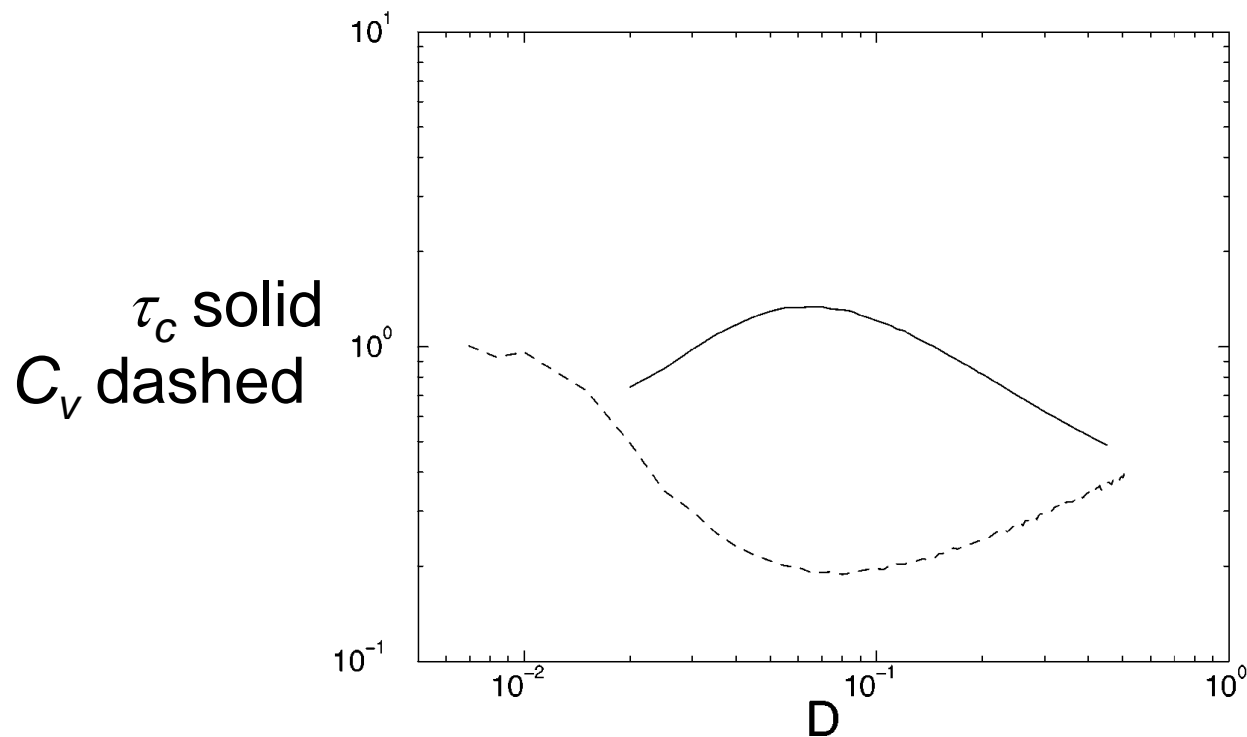
- Fitz Hugh–Nagumo model ( $\varepsilon=0.01$ ,  $a=1.05$ )





# How to quantify coherence resonance?

- Characteristic correlation time  $\tau_c = \int_0^\infty C^2(t) dt$
- Coefficient of variation of inter-spike-interval distribution  $C_V = \sigma/\mu$



# Methods of time-series analysis

- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
- Spatio-temporal representation of a time-series
- Instantaneous phase and amplitude

# What is an “attractor”?

## Lorentz model

$$\dot{x} = \delta(y - x)$$

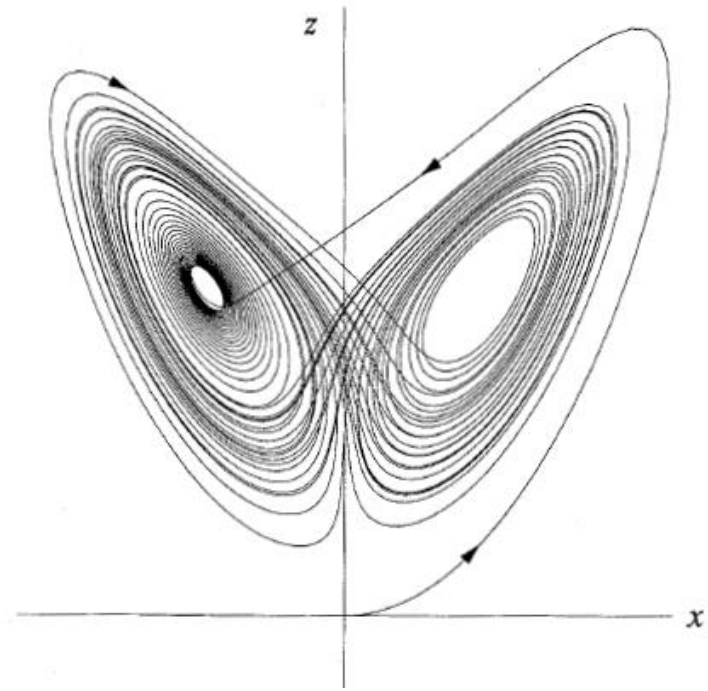
$$\dot{y} = x(r - z) - y$$

$$\dot{z} = xy - bz$$

$$\delta=10$$

$$b=8/3$$

$$r=28$$



## Why “attractor reconstruction” from a time series?

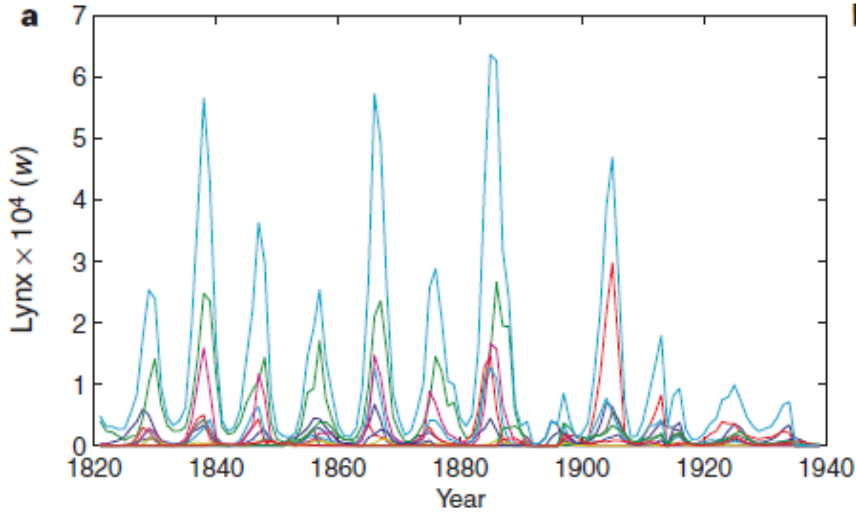
- In real systems “attractors” live in high-dimensional spaces.
- Models are too complex or unknown.

# Complex dynamics and phase synchronization in spatially extended ecological systems

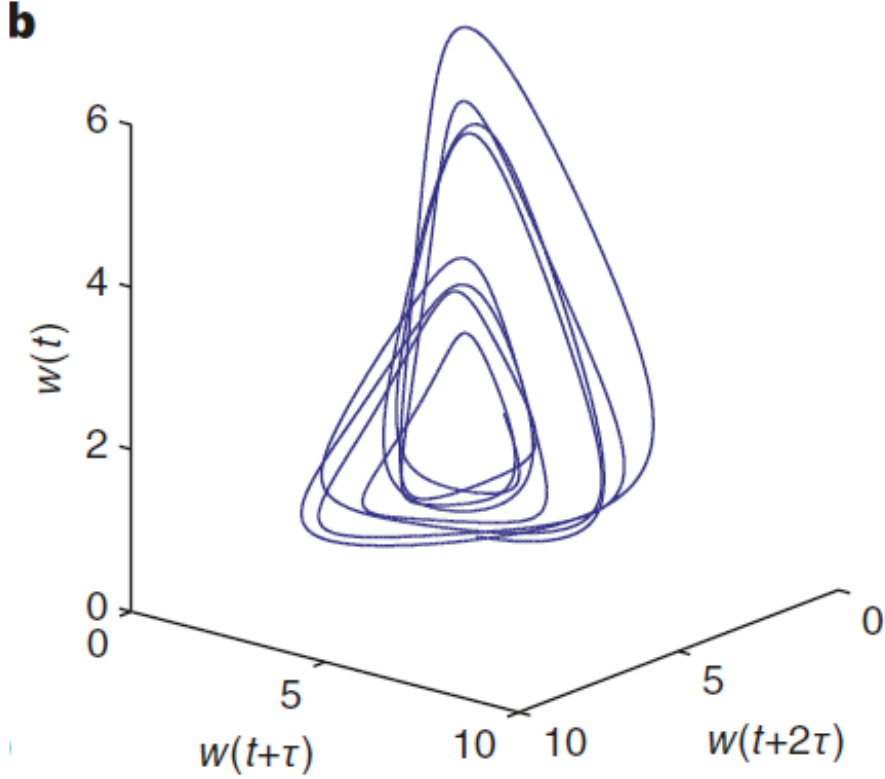
Bernd Blasius, Amit Huppert & Lewi Stone

The Porter Super-Center for Ecological and Environmental Studies & Department of Zoology, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel

NATURE | VOL 399 | 27 MAY 1999 |



lynx abundances from six regions in Canada



Reconstructed attractor obtained by spatially averaging all regional lynx data and embedding the resulting time-series  $w(t)$  using lagged coordinates,  $w(t)$  versus  $w(t+\tau)$  versus  $w(t+6\tau)$  after filtering and interpolating.

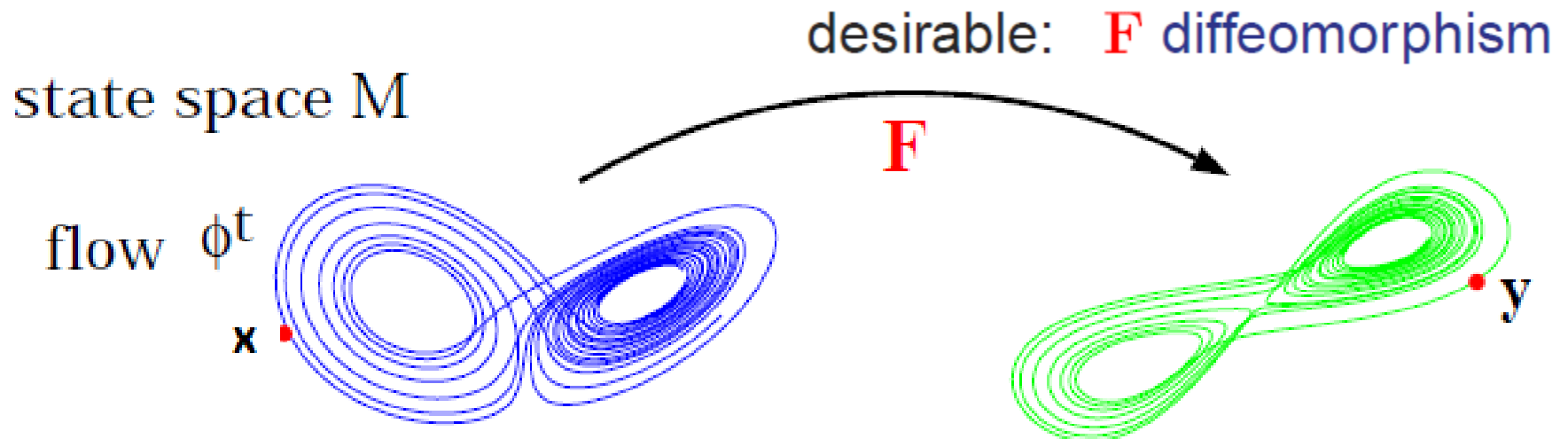


# How to identify (and quantify) chaos in observed data?

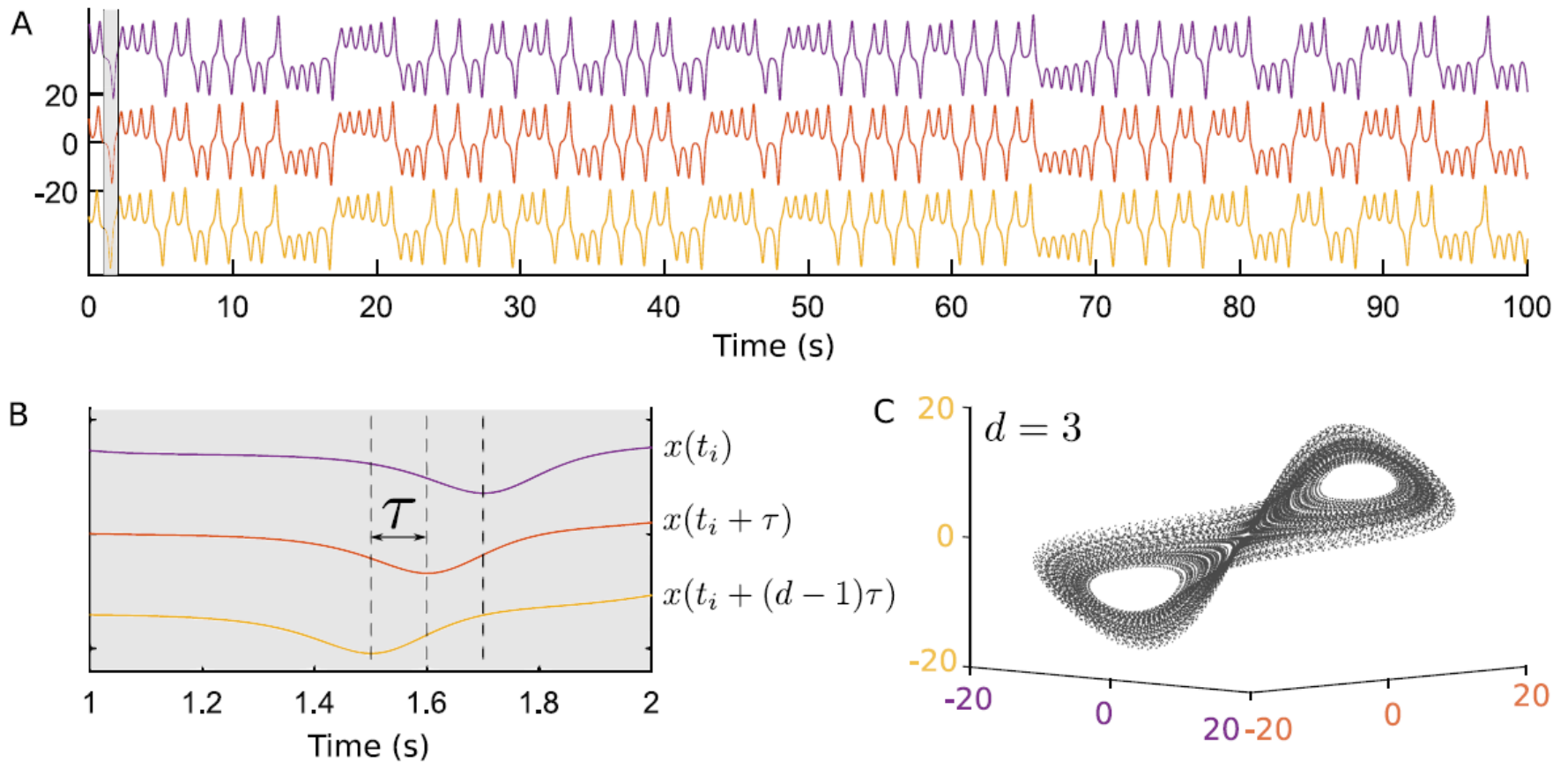
Observed time series  $S = \{s(1), s(2), \dots, s(t) \dots\}$

Attractor reconstruction: “embed” the time series in a phase-space of dimension  $d$  using delay  $\tau$  coordinates

$$y(t) = (s(t), s(t+\tau), \dots, s(t+(d-1)\tau))$$

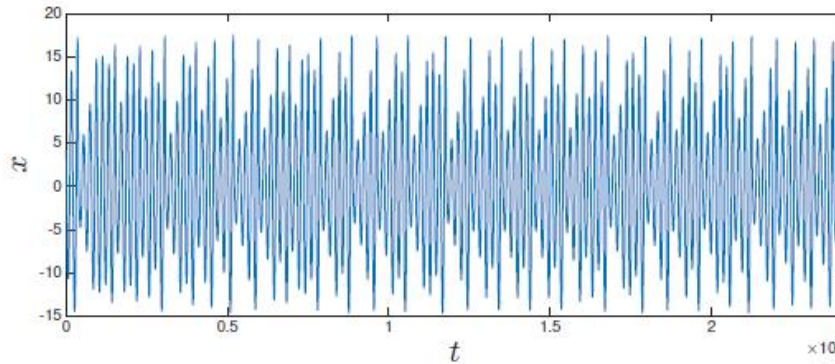


Adapted from U. Parlitz (Gottingen)

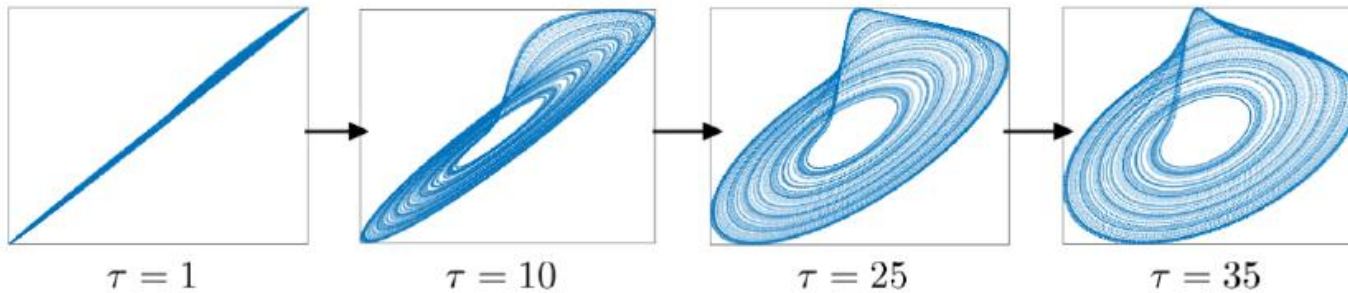


**Fig. 36.** Example of time delay embedding of the  $x$  component of the Lorenz system. (A) Following embedding, a series of delayed time series are obtained. Here the embedding dimension  $d$  is 3, and thus three time series are produced. Time series amplitude is offset for clarity. (B) Zoomed in view of the grey area shown in (A), to demonstrate the time delay  $\tau = 0.1$  s. (C) Reconstruction of the attractor in phase space is obtained by plotting the embedded signals against each other in phase space.

# Reconstruction using delay coordinates



$$\mathbf{X}(t_i) = \begin{bmatrix} x(t_i) \\ x(t_i + \tau) \\ \vdots \\ x(t_i + (d-1)\tau) \end{bmatrix}$$



A problem: how to choose the embedding parameters  
(lag  $\tau$ , dimension  $d$ )

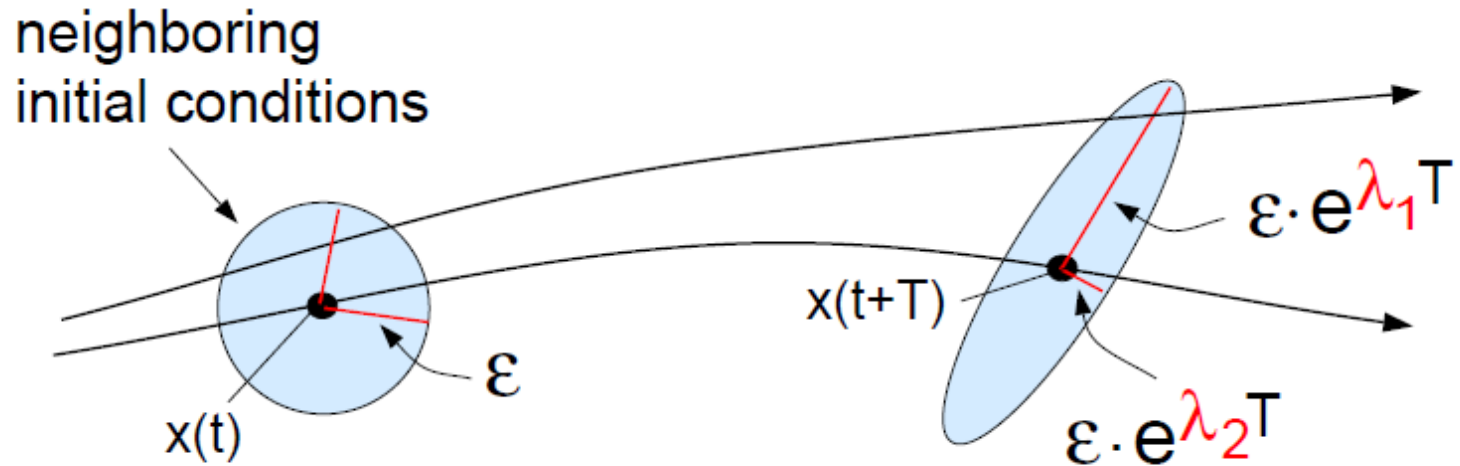
## How to choose the lag $\tau$ and the dimension $d$

- $\tau$  is chosen to maximize the spread of the data in phase space: the first zero of the autocorrelation function (or where  $|C(\tau)|$  is minimum)
- $d$  is often estimated with the *false nearest neighbors* technique that examines how close points in phase space remain close as the dimension is increased.
- Points that do not remain close are ‘false’ neighbors.
- The number of false neighbors decreases as the embedding dimension is increased.
- The first dimension for which the number of false neighbors decreases below a threshold provides the estimated  $d$ .

After reconstructing the attractor, we can characterize it by the **fractal dimension** and the **Lyapunov exponent**.



# Lyapunov exponents: measure how fast neighboring trajectories diverge



- A stable fixed point has negative  $\lambda$ s (since perturbations in any direction die out)
- An attracting limit cycle has one zero  $\lambda$  and negative  $\lambda$ s
- A chaotic attractor has at least one positive  $\lambda$ .

## Steps to compute the maximum LE

- Initial distance  $\delta_I = |s_i - s_j|$
- Final distance  $\delta_F = |s_{i+T} - s_{j+T}|$
- Local *exponential* grow  $\lambda_{\text{local}}^* = \frac{1}{T} \log(\delta_F / \delta_I)$
- The rate of grow is averaged over the attractor, which gives  $\lambda_{\text{max}}$

**A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, Physica D 16 (1985)**

A very popular method for detecting chaos in experimental time series.

## A word of warning!

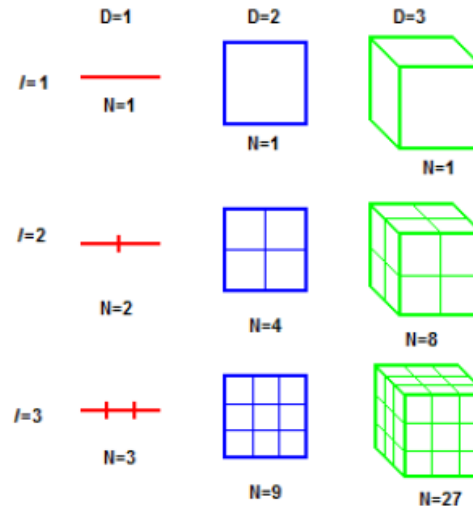
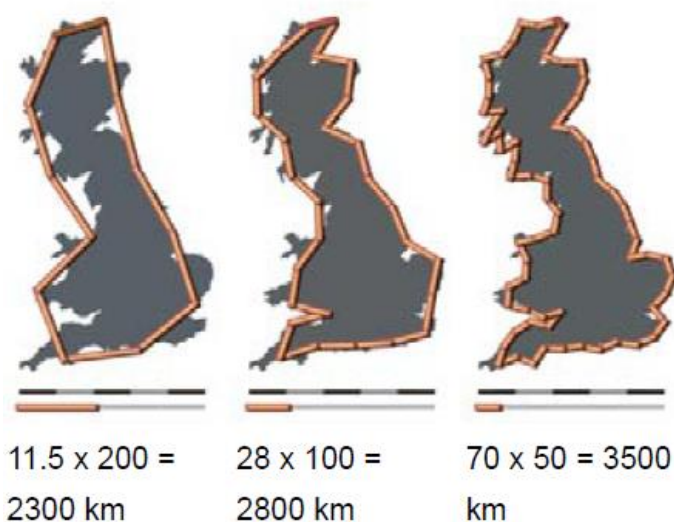
- The rate of grow depends on the direction in the phase space.
- The algorithm returns  $\lambda$  in the fastest expansion direction.
- Therefore, the algorithm always returns a positive number!
- This is a problem when computing the LE of noisy data.

### Further reading:

- F. Mitschke and M. Damming, *Chaos vs. noise in experimental data*, Int. J. Bif. Chaos 3, 693 (1993)
- A. Pikovsky and A. Politi, *Lyapunov Exponents* (Cambridge University Press, 2016)

# Fractal dimension

- Example: the fractal dimension of a coastline quantifies how the number of scaled measuring sticks required to measure the coastline changes with the scale applied to the stick.



Source: wikipedia

- Fractal dimension:  $N \propto \epsilon^{-D} \rightarrow D_0 = \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}}$

# Grassberger-Procaccia correlation dimension algorithm

- Another very popular method for detecting chaos in real-world data.
- Fractal dimension (box counting dimension):  $D_0 = \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}}$
- Problem: for time-series analysis,  $D_0$  does not distinguish between frequent and unfrequently visited boxes.

- An alternative: the *correlation dimension*.  
 $g$  is the number of pairs of points with distance between them  $< \epsilon$ .

For any set of  $N$  points in an  $m$ -dimensional space

$$\vec{x}(i) = [x_1(i), x_2(i), \dots, x_m(i)], \quad i = 1, 2, \dots, N$$

then the correlation integral  $C(\epsilon)$  is calculated by:

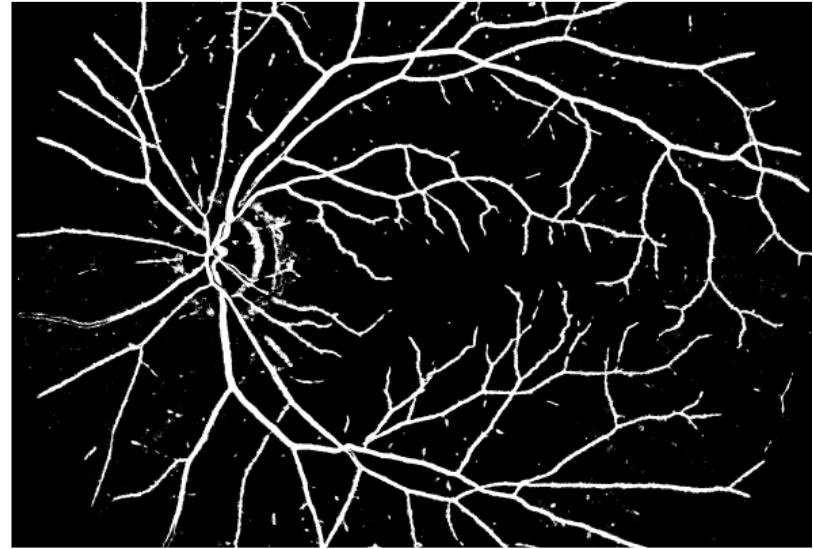
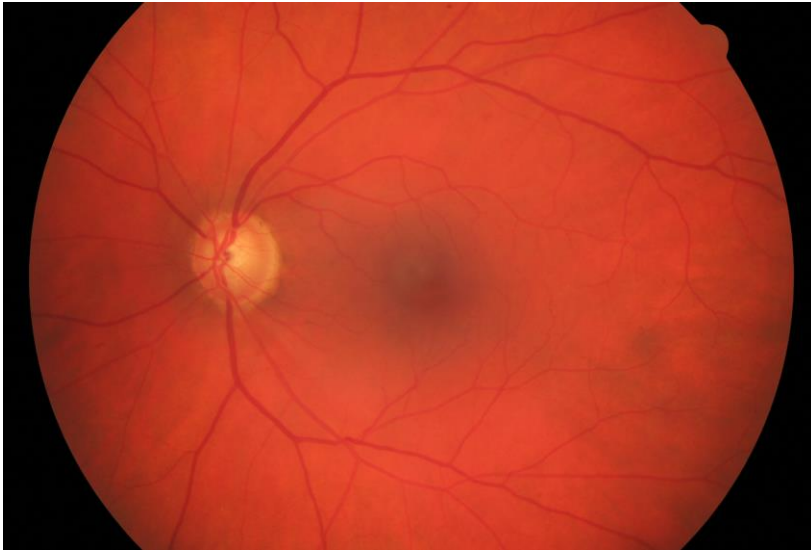
$$C(\epsilon) = \lim_{N \rightarrow \infty} \frac{g}{N^2} \quad \boxed{C(\epsilon) \sim \epsilon^\nu}$$

Further reading:

P. Grassberger and I. Procaccia, "Measuring the Strangeness of Strange Attractors".  
Physica D vol. 9, pp.189, 1983.

L. S. Liebovitch and T. Toth, "A fast algorithm to determine fractal dimensions by box counting," Physics Letters A, vol. 141, pp. 386, 1989.

# Example of application of fractal analysis: distinguishes between diabetic retinopathy and normal patients



Source: Pablo Amil, UPC

## The fractal dimension of the blood vessels

- in the normal human retina is  $\approx 1.7$
- tends to increase with the level of diabetic retinopathy
- varies considerably depending on the image quality and the technique used for measuring the fractal dimension

# **Symbolic methods to identify patterns and structure in time series**

# Can lasers mimic real neurons?

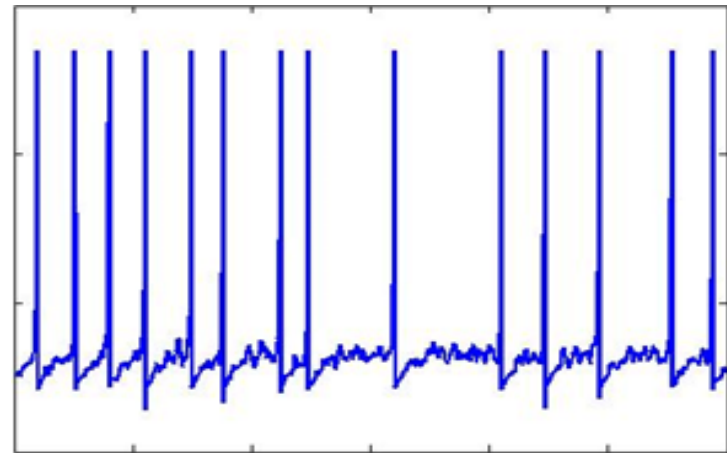
## Laser spikes



Time ( $\mu\text{s}$ )



## Neuronal spikes



Time (ms)





# Are there statistical similarities between neuronal spikes and optical spikes?

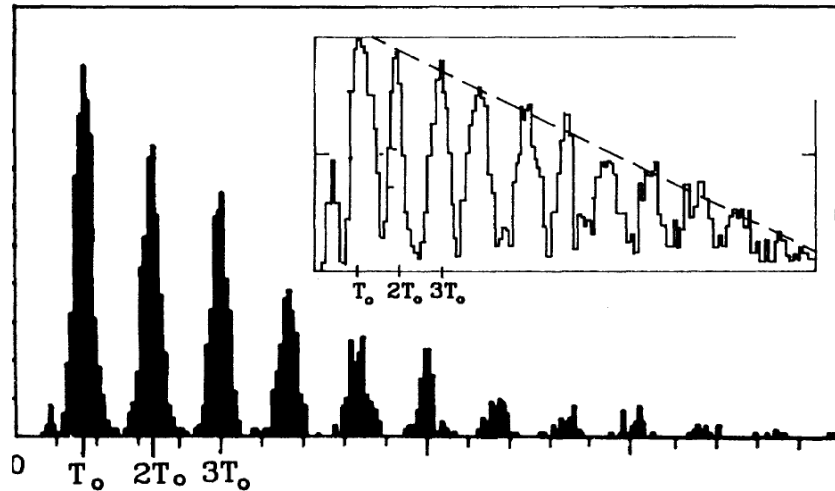
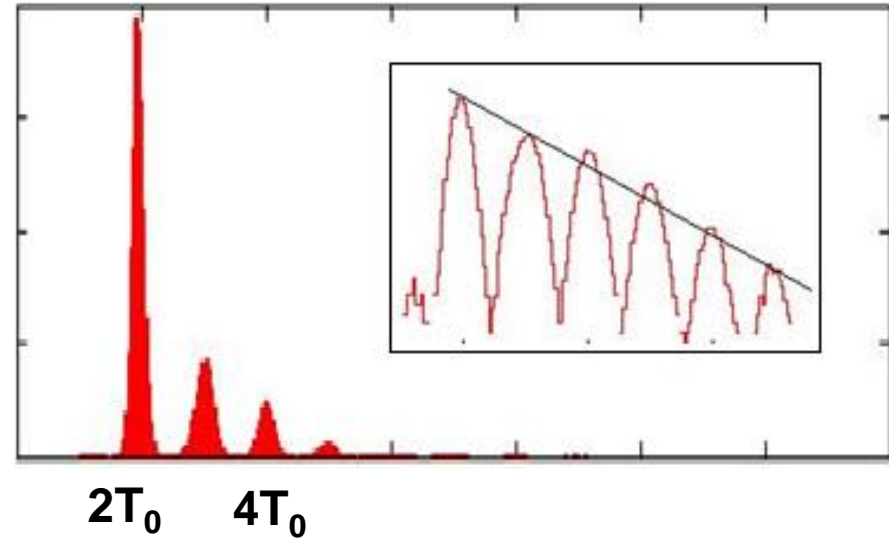


FIG. 1. (a) An experimental ISIH obtained from a single auditory nerve fiber of a squirrel monkey with a sinusoidal 80-dB sound-pressure-level stimulus of period  $T_0 = 1.66$  ms applied at the ear. Note the modes at integer multiples of  $T_0$ . Inset:

A. Longtin et al PRL (1991)

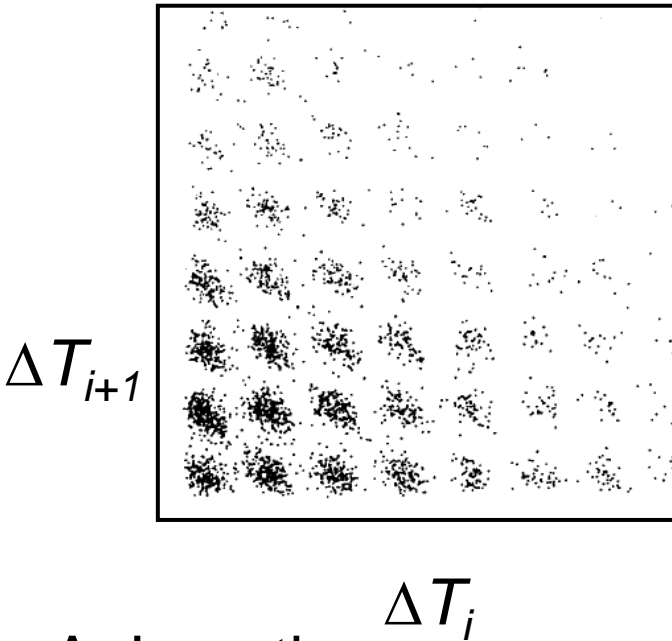


Experimental data when the laser current is modulated with a sinusoidal signal of period  $T_0$ .

[A. Aragonese et al](#)  
[Optics Express \(2014\)<sup>49</sup>](#)

# Return maps of inter-spike-intervals

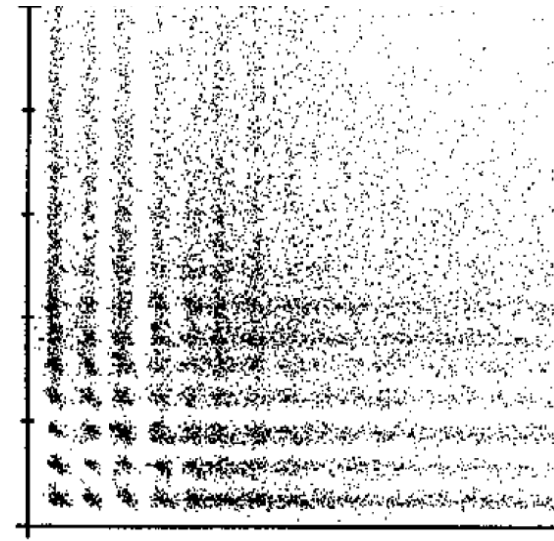
Neuronal ISIs



A. Longtin

Int. J. Bif. Chaos (1993)

Laser ISIs



M. Giudici et al PRE (1997)

[A. Aragonese et al  
Optics Express \(2014\)](#)

HOW TO IDENTIFY TEMPORAL ORDER?  
MORE/LESS EXPRESSED PATTERNS?

# Symbolic analysis

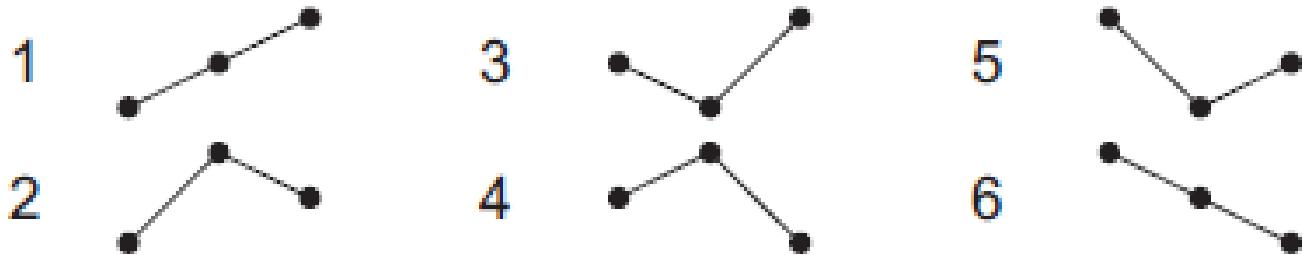
- The time series  $\{x_1, x_2, x_3, \dots\}$  is transformed (using an appropriated **rule**) into a sequence of symbols  $\{s_1, s_2, \dots\}$
- taken from an “**alphabet**” of possible symbols.
- Then consider “blocks” of  $D$  symbols (“**patterns**” or “**words**”).
- All the possible words form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

# Threshold transformation: “partition” of the phase space

- if  $x_i > x_{th} \Rightarrow s_i = 0$ ; else  $s_i = 1$   
transforms a time series into a sequence of 0s and 1s, e.g.,  
{011100001011111...}
- Considering “blocks” of  $D$  letters gives the sequence of words. Example, with  $D=3$ :  
{011 100 001 011 111 ...}
- The number of words (patterns) grows as  $2^D$
- More thresholds allow for more letters in the “alphabet” (and more words in the dictionary). Example:  
if  $x_i > x_{th1} \Rightarrow s_i = 0$ ;  
else if  $x_i < x_{th2} \Rightarrow s_i = 2$ ;  
else ( $x_{th2} < x_i < x_{th1}$ )  $\Rightarrow s_i = 1$ .

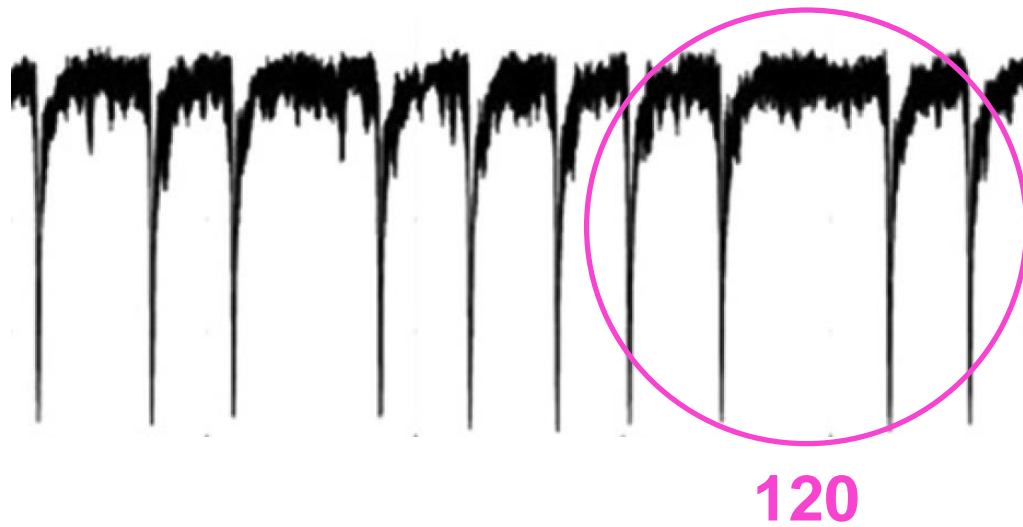
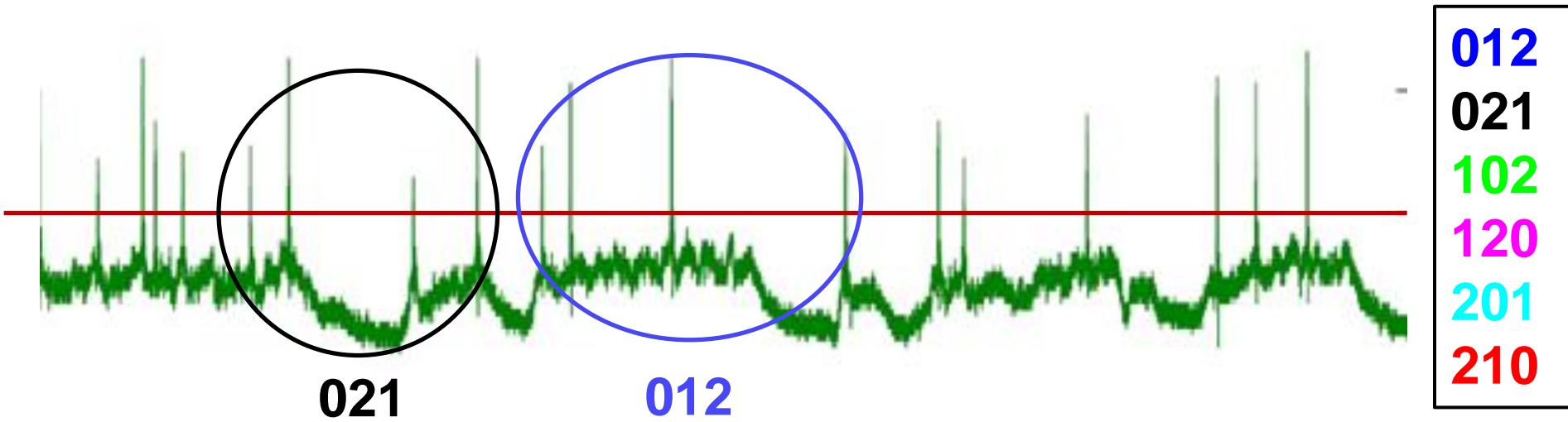
# Ordinal analysis: a method to find patterns in data

- Consider a time series  $X(t) = \{\dots, X_i, X_{i+1}, X_{i+2}, \dots\}$
- Which are the possible order relations among three data points?

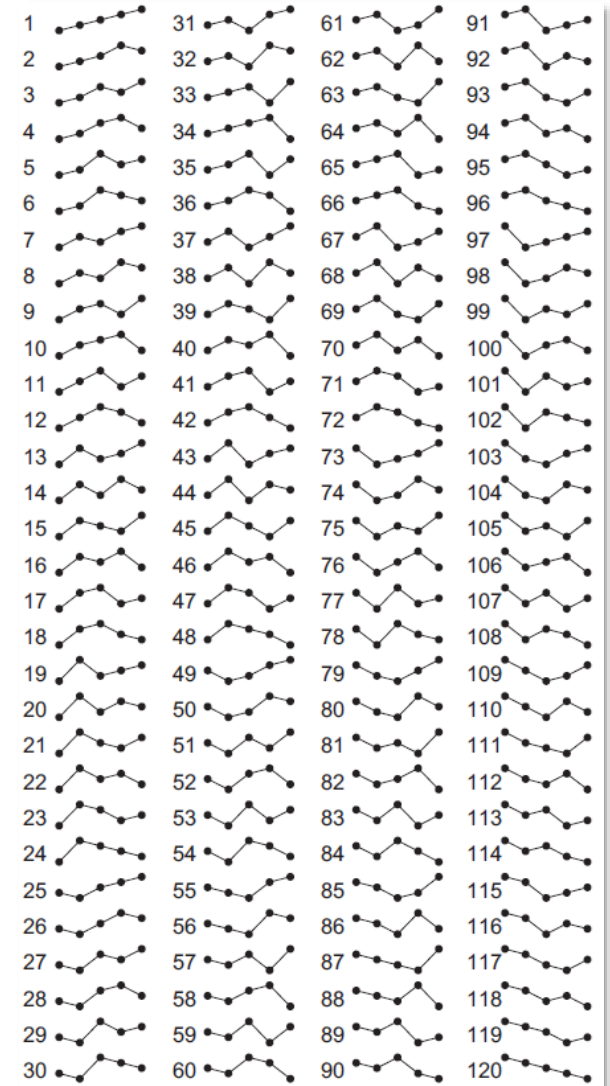
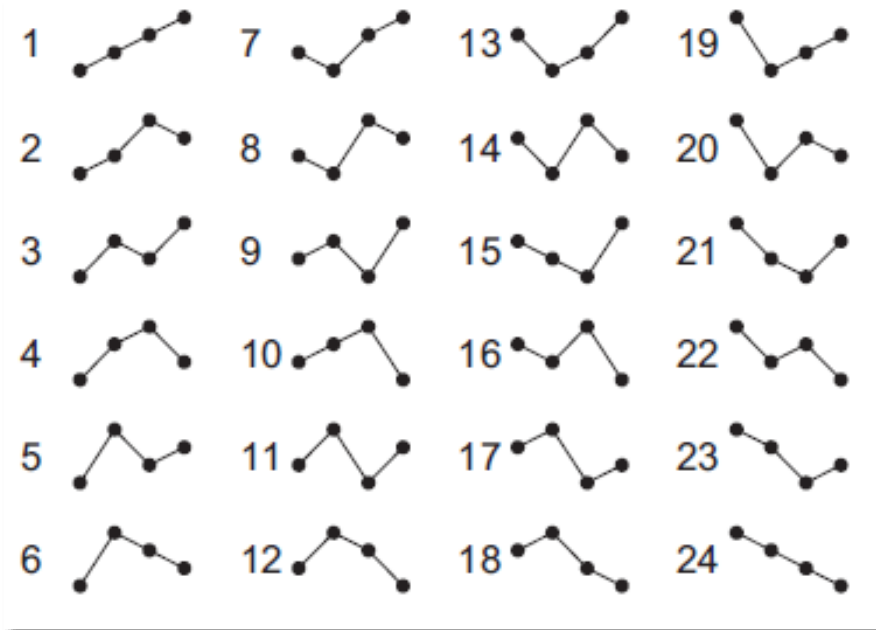


- Count how many times each “ordinal pattern” appears.
- Advantages: allows to identify temporal structures & is robust to noise.
- Drawback: information about actual data values is lost.

# Analysis of D=3 patterns in spike sequences



# The number of ordinal patterns increases as D!



- A problem for short datasets
- How to select optimal D?  
it depends on:
  - The length of the data
  - The length of the correlations

# Comparison

## Threshold transformation:

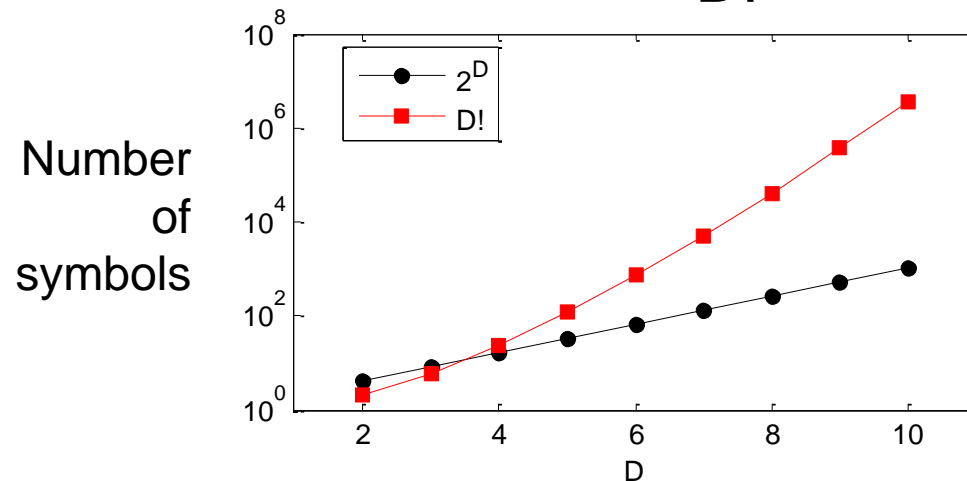
if  $x_i > x_{th} \Rightarrow s_i = 0$ ; else  $s_i = 1$

- Advantage: keeps information about the magnitude of the values.
- Drawback: how to select an adequate threshold (“partition” of the phase space).
- $2^D$

## Ordinal transformation:

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

- Advantage: no need of threshold; keeps information about the temporal order in the sequence of values
- Drawback: no information about the actual data values
- D!





# Are the $D!$ ordinal patterns equally probable?

- **Null hypothesis:**

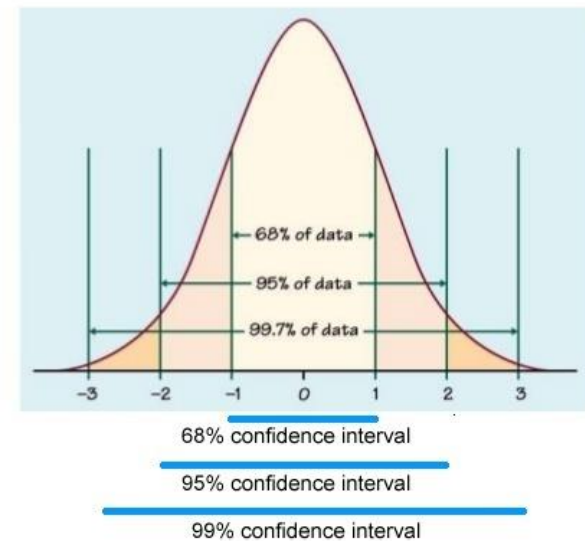
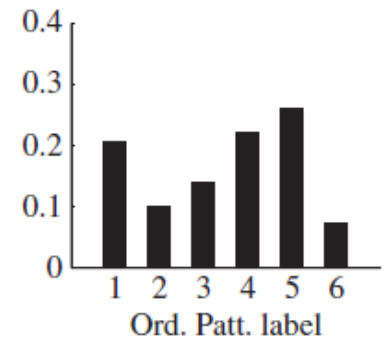
$$p_i = p = 1/D! \quad \text{for all } i = 1 \dots D!$$

- 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.

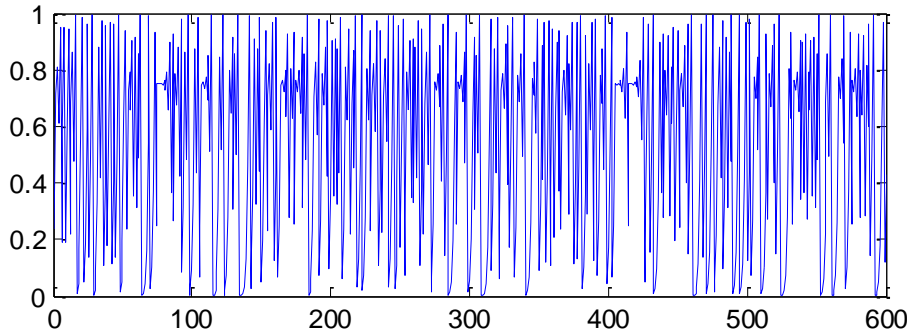
- Else

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

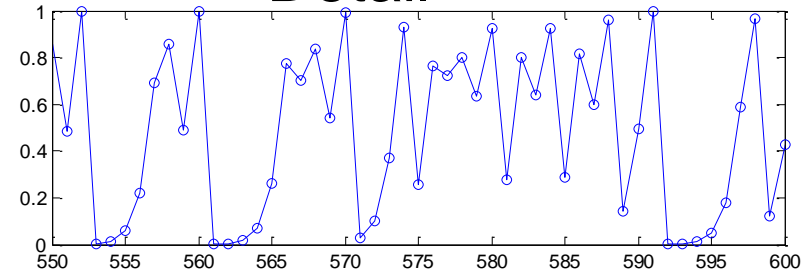


# Logistic map $x(i+1) = r x(i)[1 - x(i)]$

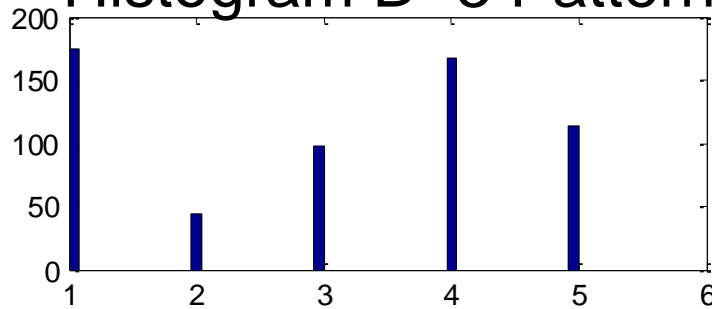
## Time series



## Detail

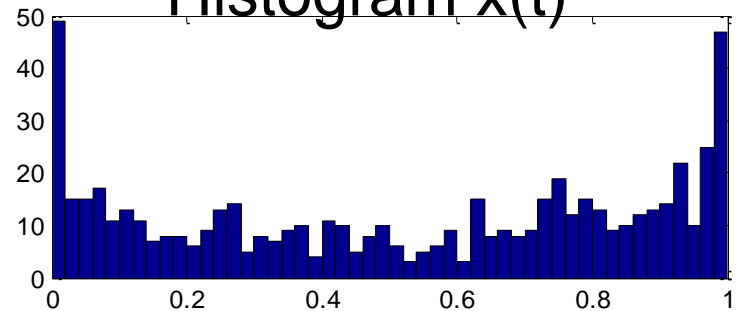


## Histogram D=3 Patterns



↑  
**forbidden**

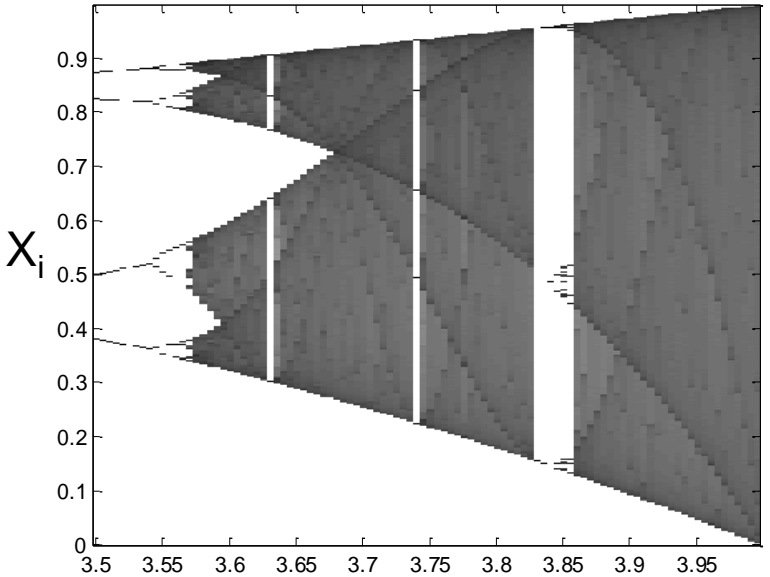
## Histogram $x(t)$



Ordinal analysis yields information about more and less expressed patterns in the data

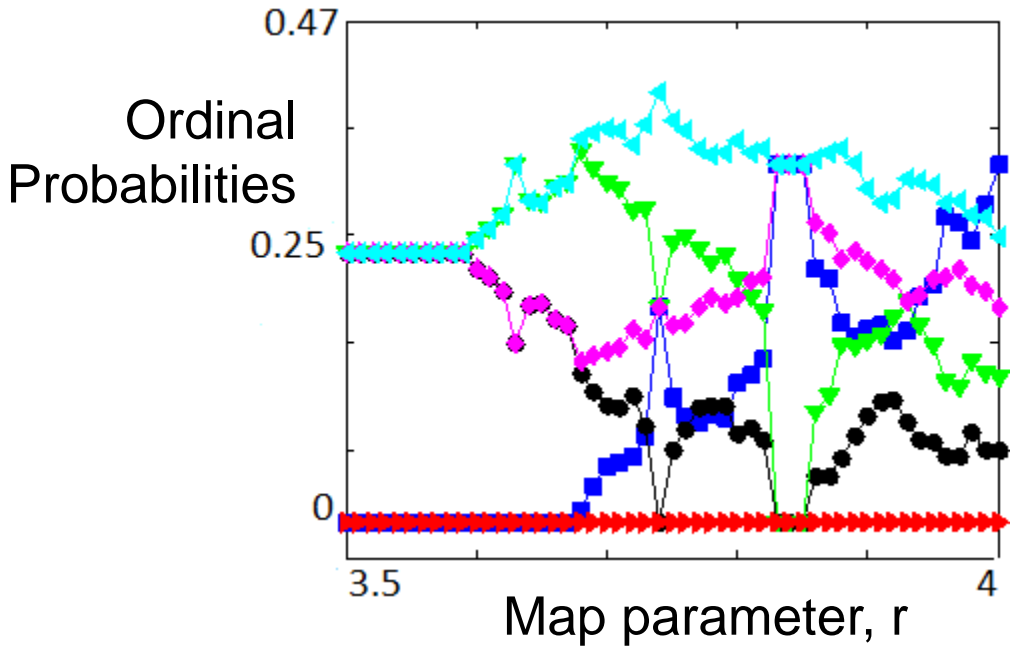
# “ordinal” bifurcation diagram of the Logistic map with D=3

Normal bifurcation diagram



Map parameter

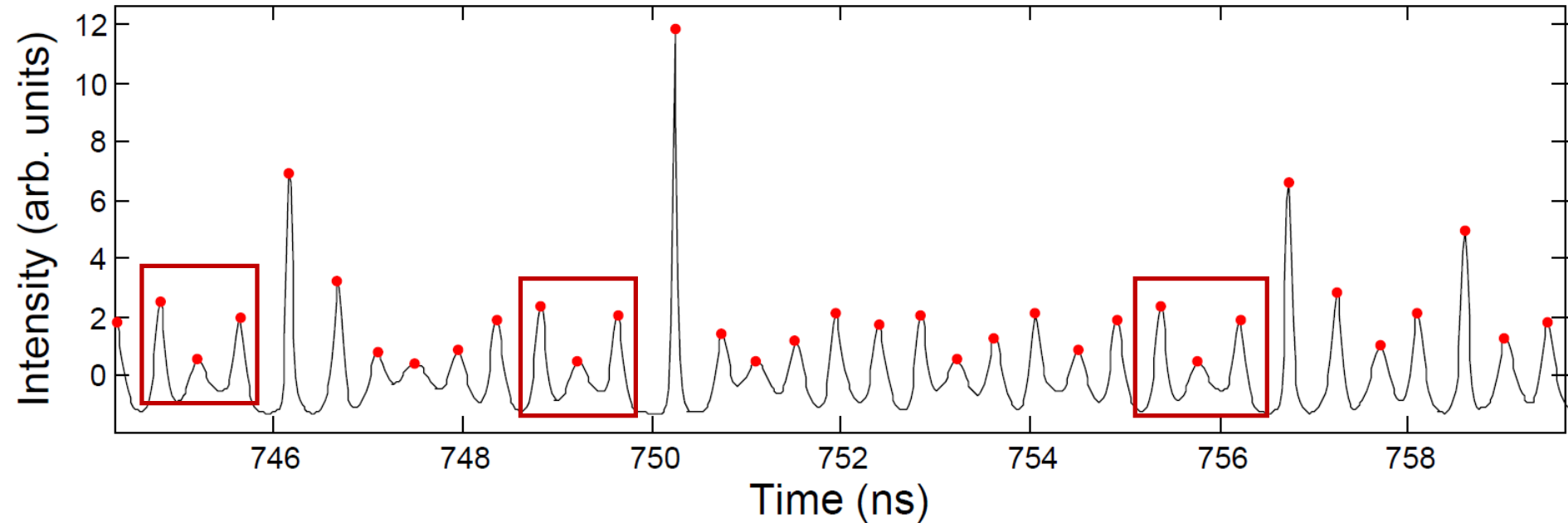
Ordinal bifurcation diagram



**012 021 102 120 201 210**

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

# Example: intensity pulses emitted by a chaotic laser



N. Martinez Alvarez, S. Borkar and C. Masoller, "Predictability of extreme intensity pulses in optically injected semiconductor lasers"  
Eur. Phys. J. Spec. Top. 226, 1971 (2017).

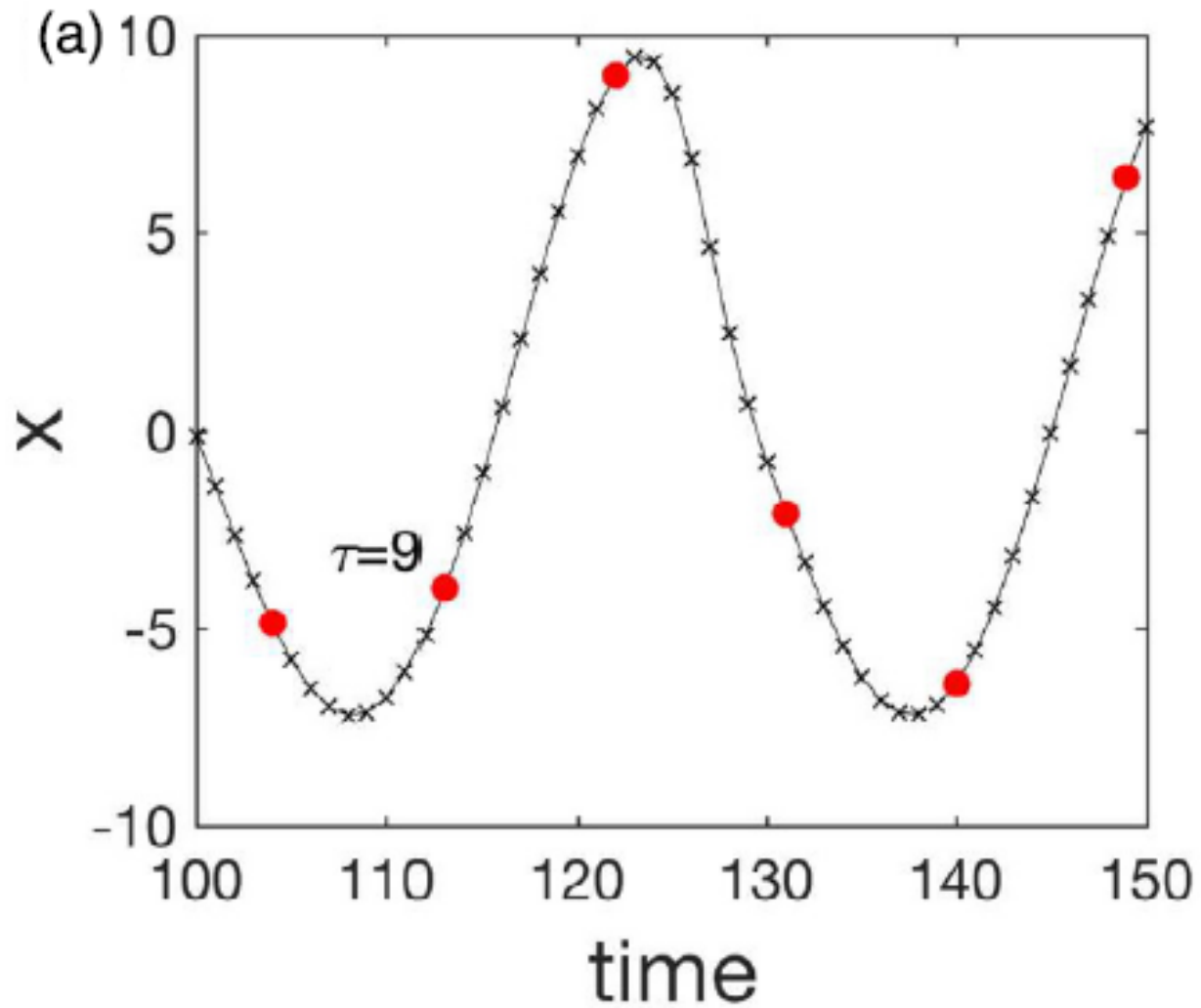
# How to detect longer temporal correlations?

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

- Problem: number of patterns increases as  $D!$ .
- Solution: a **lag  $\tau$**  allows considering long time-scales without having to use words of many letters

[...  $x(t)$ ,  $x(t + 2)$ ,  $x(t + 4)$ ,...]

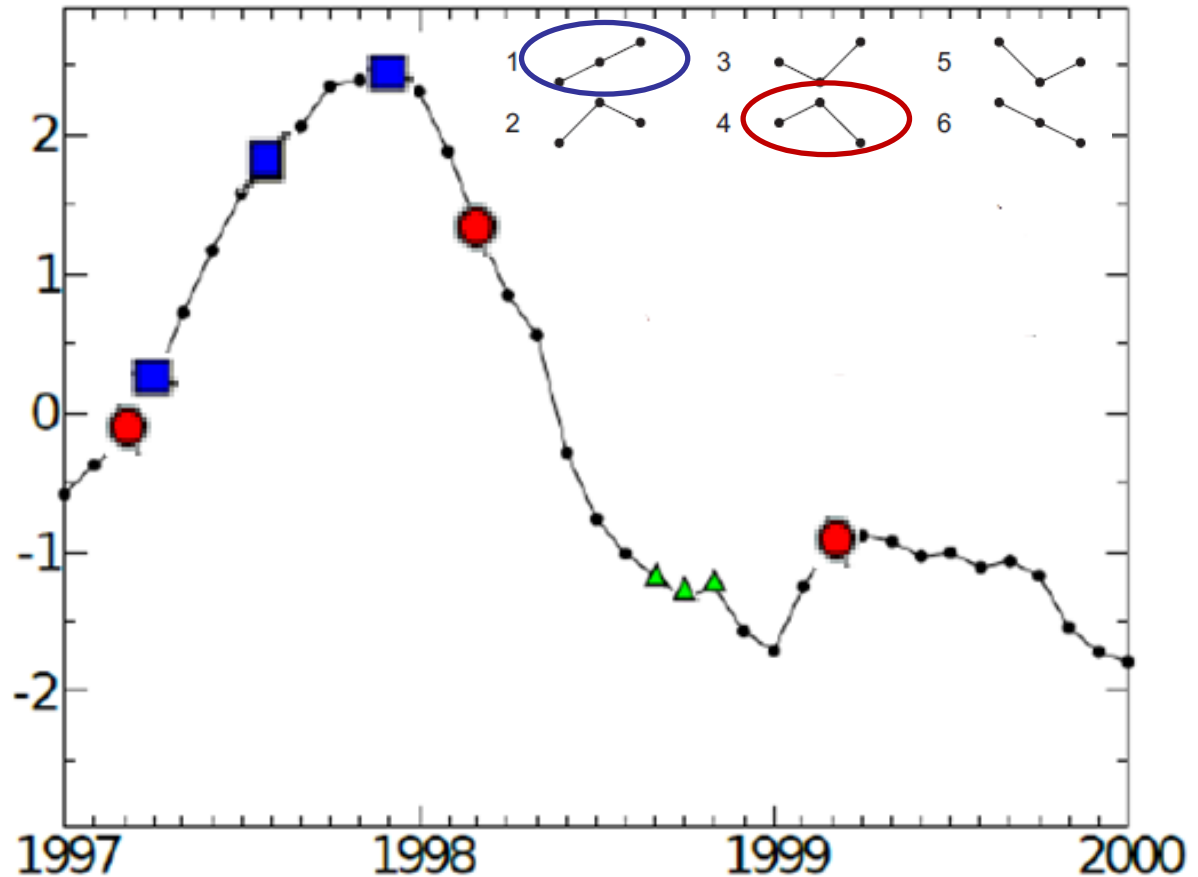
- Example: climatological data (monthly sampled)
  - Consecutive months: [... $x_i(t)$ ,  $x_i(t + 1)$ ,  $x_i(t + 2)$ ...]
  - Consecutive years: [... $x_i(t)$ ,...  $x_i(t + 12)$ ,...  $x_i(t + 24)$ ...]
- **Varying  $\tau$  = varying temporal resolution (sampling time)**



# Very useful for “seasonal” data: allows to select the time scale of the analysis

- **Green** triangles: intra-seasonal pattern,
- **blue** squares: intra-annual pattern
- **red** circles: inter-annual pattern

Example: el Niño index, monthly sampled

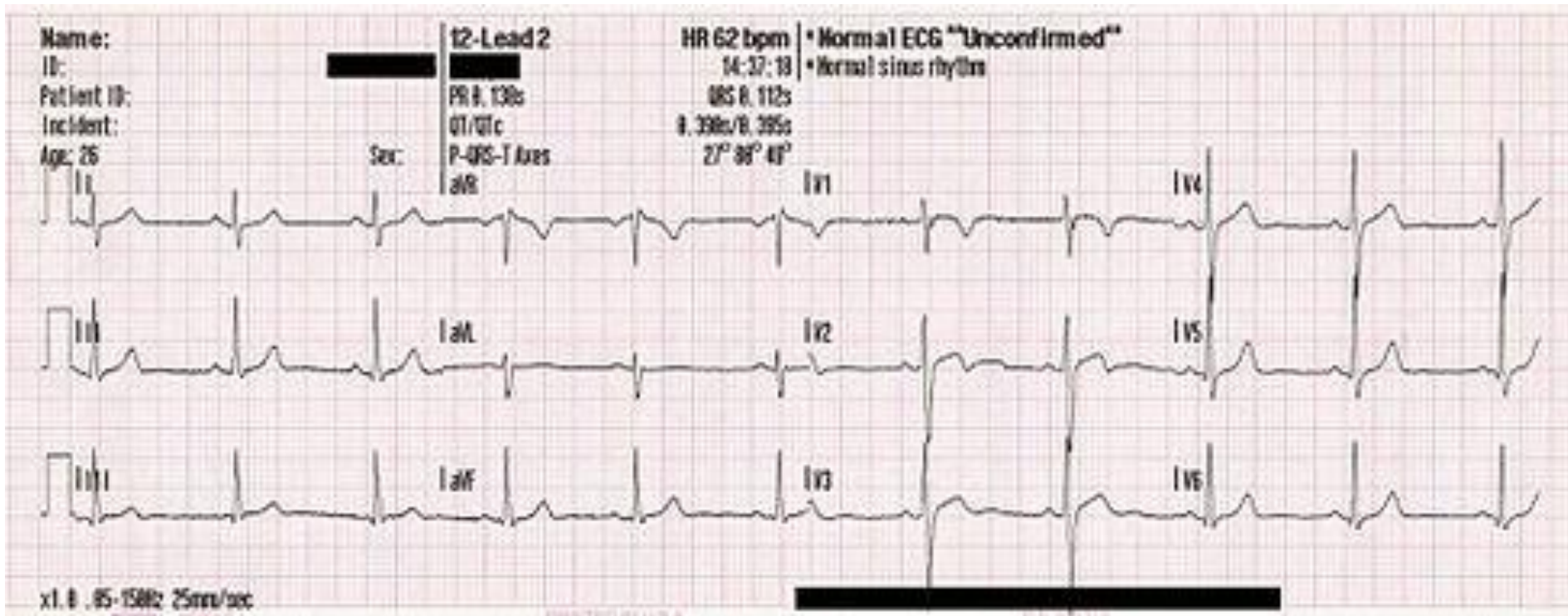


# **Examples of application**

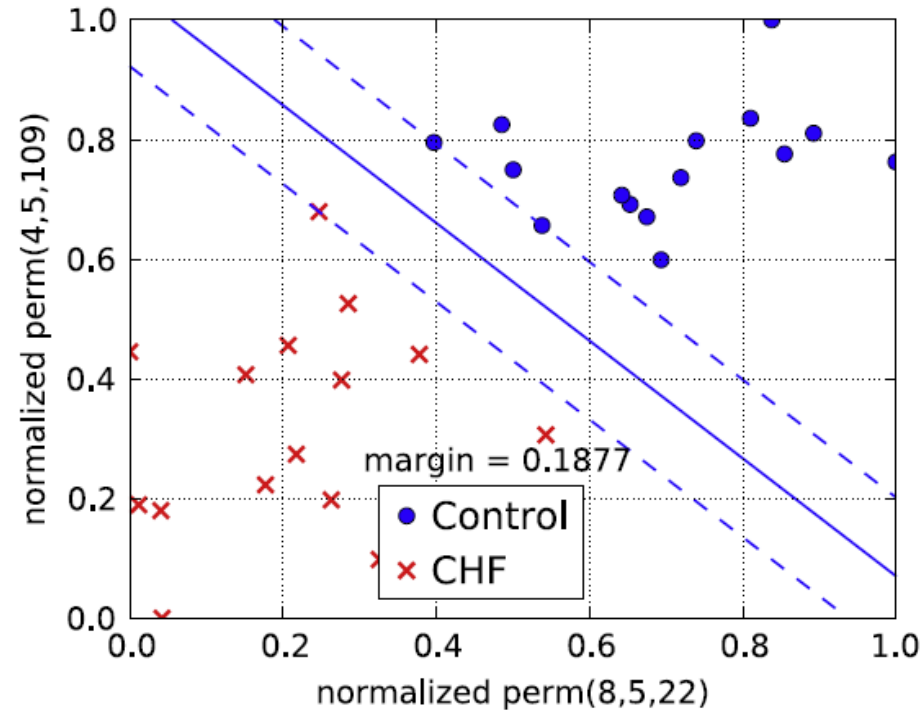
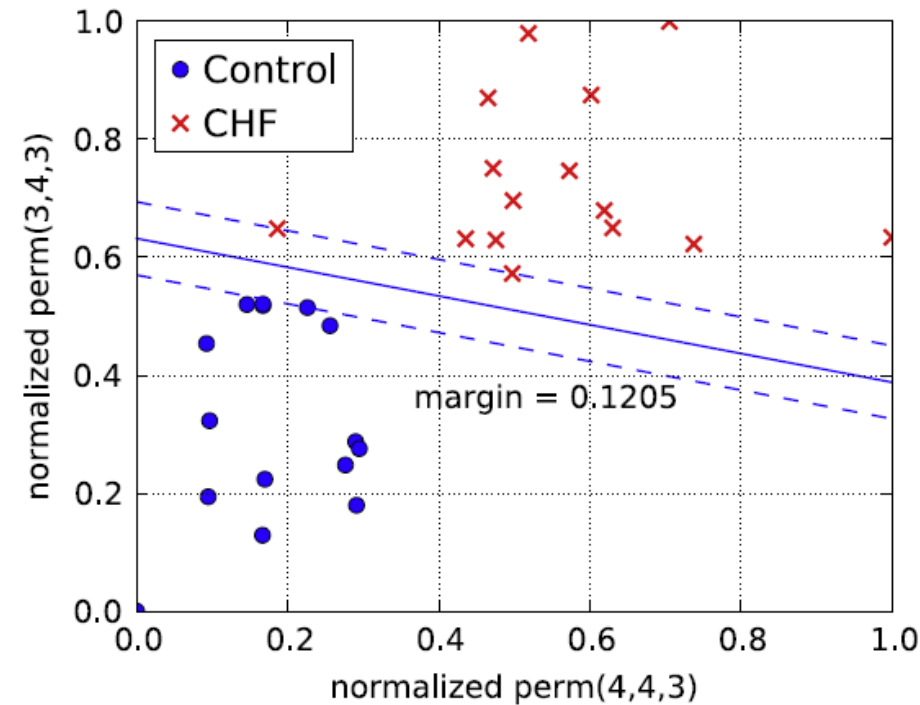
- cardiac signals**
- lasers & neurons**



# ECG signals: analysis of time series of **inter-beat intervals**



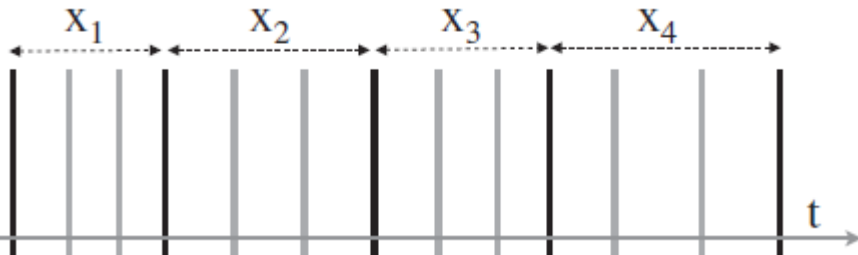
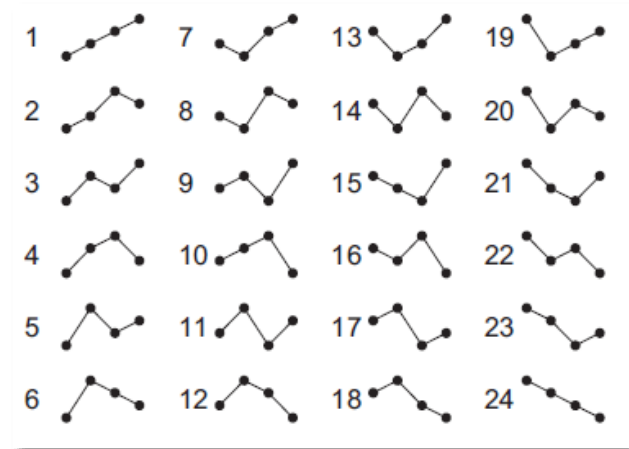
# Classifying ECG signals according to ordinal probabilities



- 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.

# 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 \(2012\)](#)

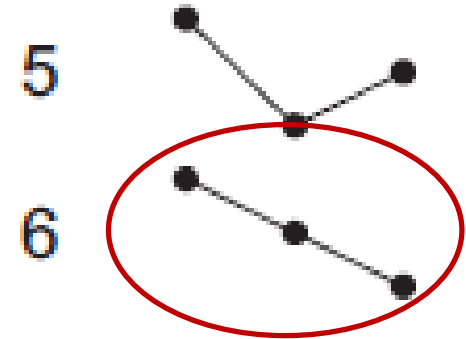
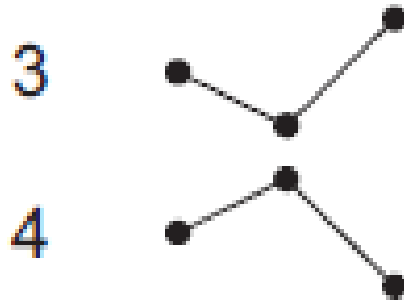
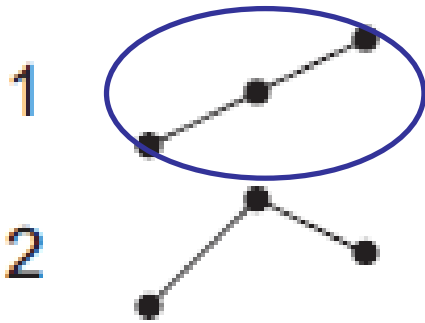
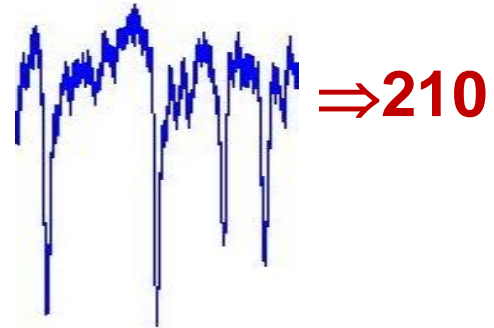
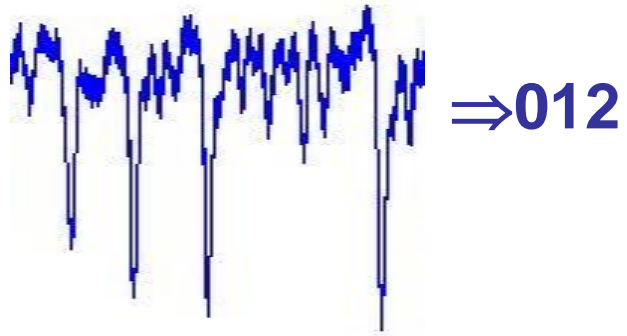


World length (wl): 4  
Lag = 3 (skip 2 points)  
Result:

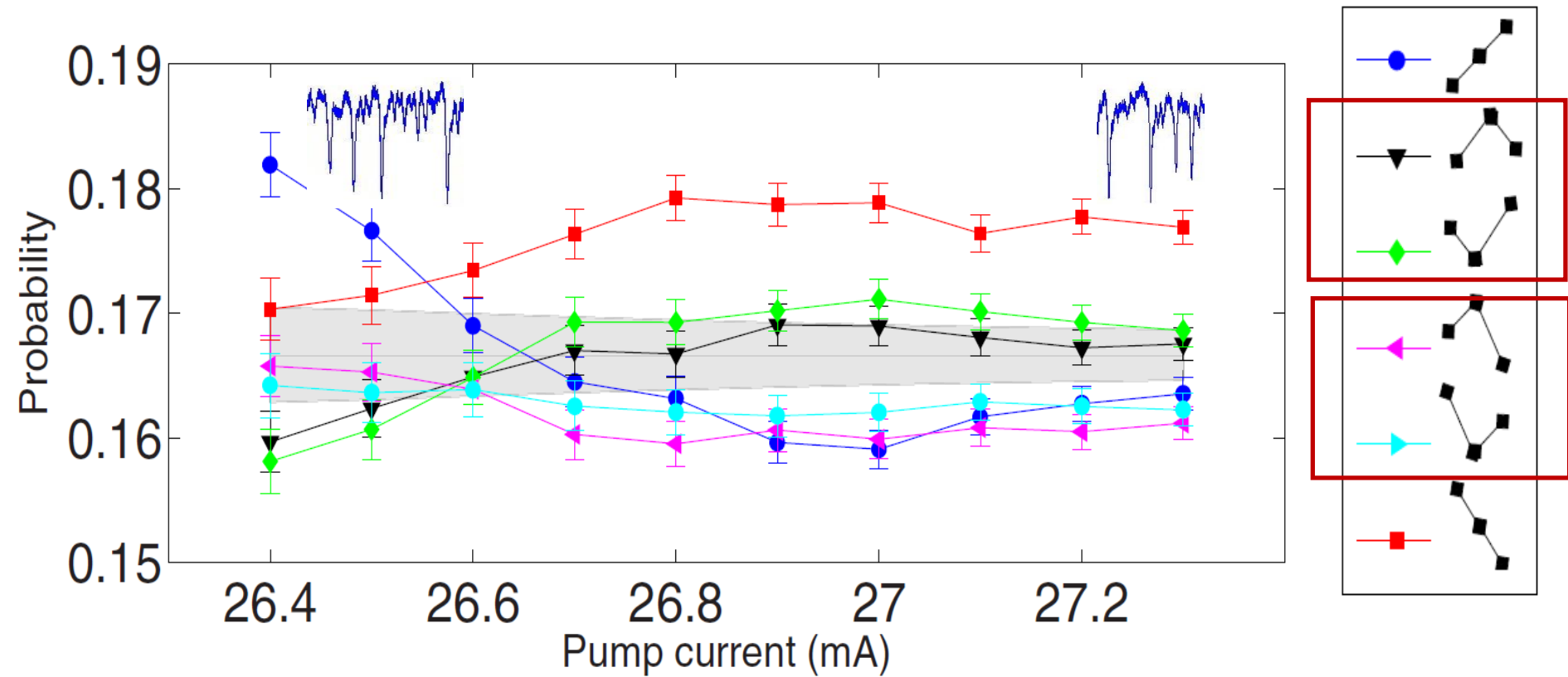
indcs = 3

```
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 ;
```

# Laser spikes: analysis of inter-spike intervals



# More probable pattern varies with the laser pump current

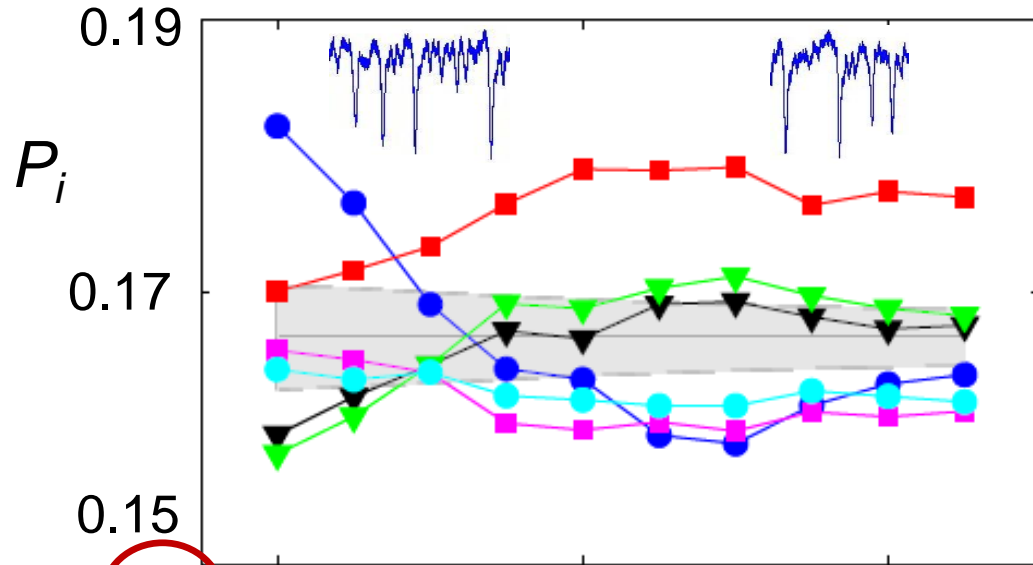


75,000 – 880,000 spikes

Gray region: probabilities consistent with 1/6

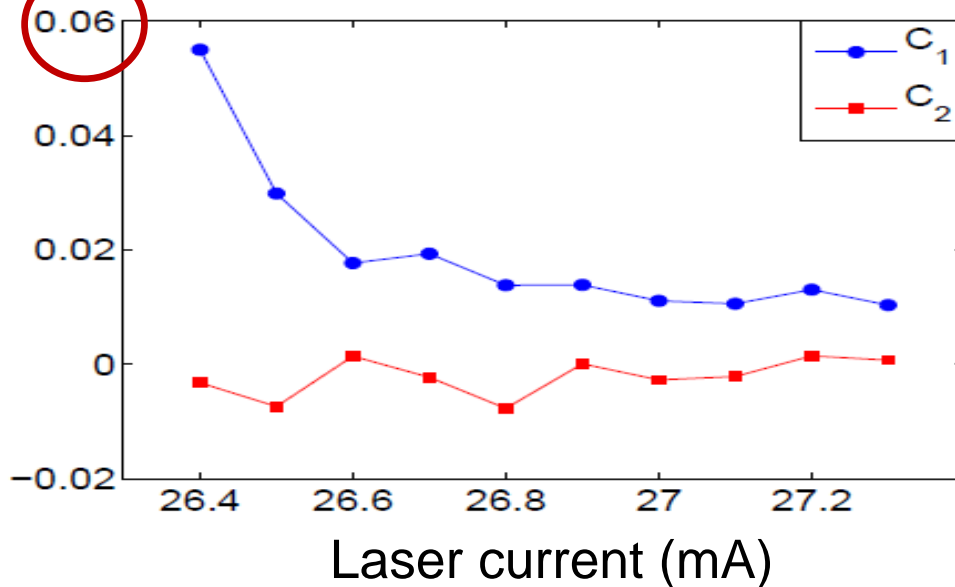
*A. Aragoneses et al, Sci. Rep. 4, 4696 (2014)*

# The variation is not captured by linear correlation analysis



$$C(\tau) = \frac{\langle [x(t) - \mu][x(t - \tau) - \mu] \rangle}{\sigma^2}$$

$$C_\tau = \frac{\langle (\Delta T_i - \langle \Delta T \rangle) (\Delta T_{i-\tau} - \langle \Delta T \rangle) \rangle}{\sigma^2}$$



# Minimal model? A modified circle map

$$\varphi_{i+1} = \varphi_i + \rho + \frac{K}{2\pi} \left[ \sin(2\pi\varphi_i) + \alpha_c \sin(4\pi\varphi_i) \right] + D\zeta$$

$$X_i = \varphi_{i+1} - \varphi_i$$

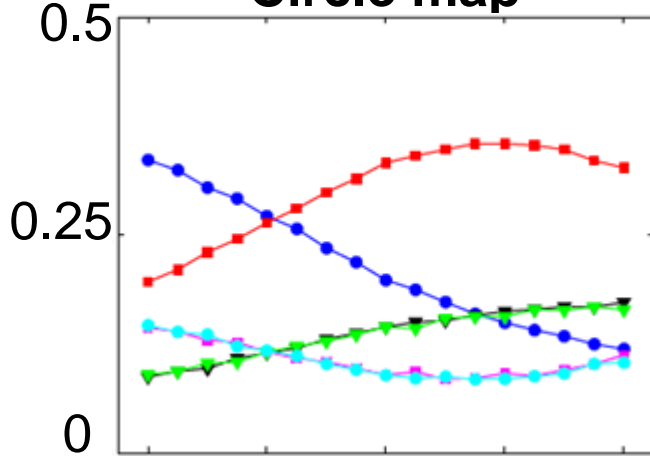
$\rho$  =  $\frac{\text{natural frequency}}{\text{forcing frequency}}$

K = forcing amplitude

D = noise strength

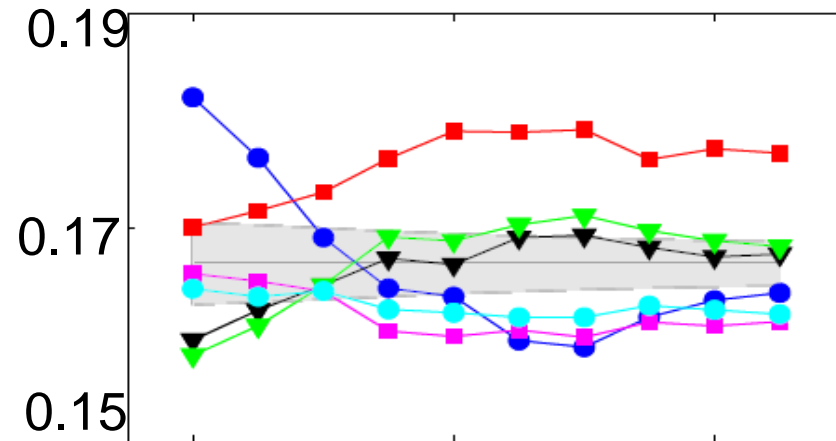
Prob

Circle map



Map parameter  $\alpha_c$

Laser spikes

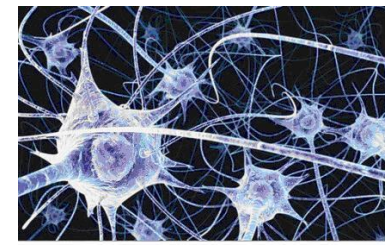


Laser current

Same “clusters” & same hierarchical structure

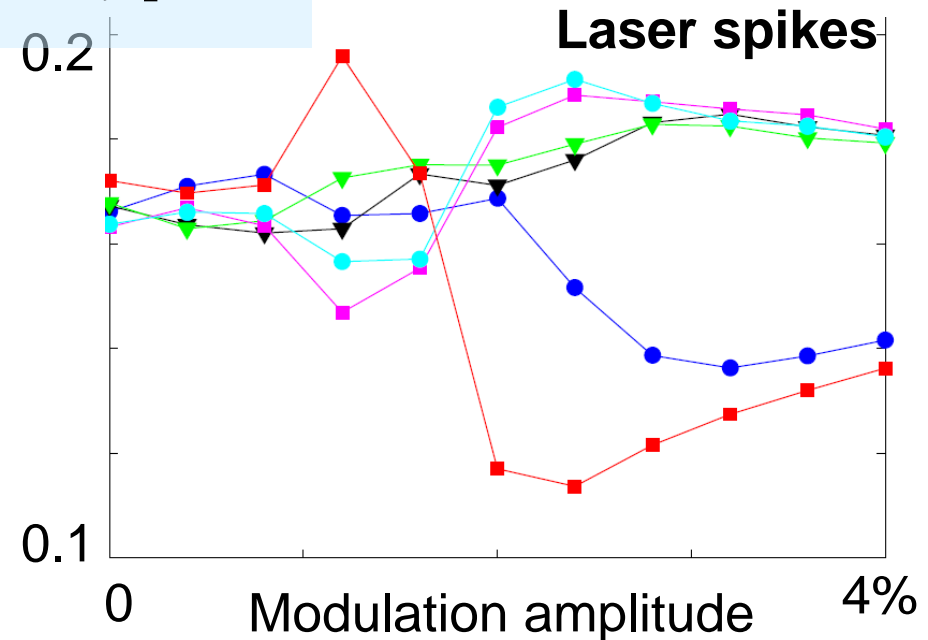
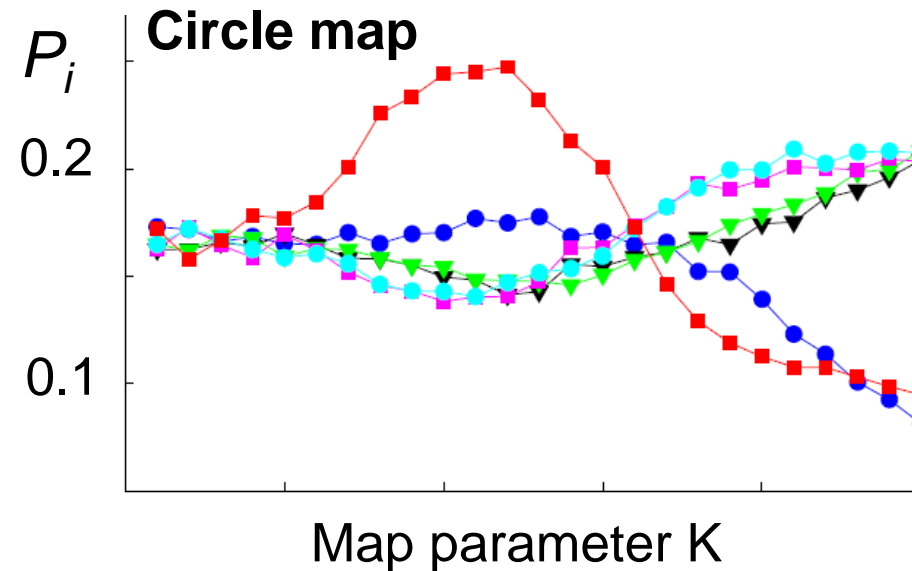
*A. Aragonese et al, Sci. Rep. 4, 4696 (2014)*

# Connection with neurons: the circle map describes many excitable systems



- Spike correlations in sensory neurons (Neiman and Russell, PRE 2005)
- Can we test its validity as a minimal model for the laser spikes?

$$\varphi_{i+1} = \varphi_i + \rho + \frac{K}{2\pi} [\sin(2\pi\varphi_i) + \alpha_c \sin(4\pi\varphi_i)] + D\zeta$$

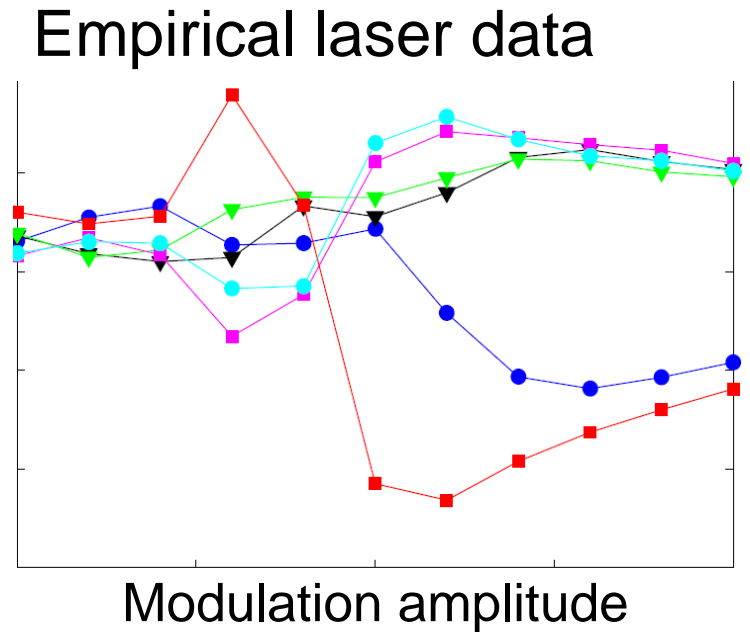
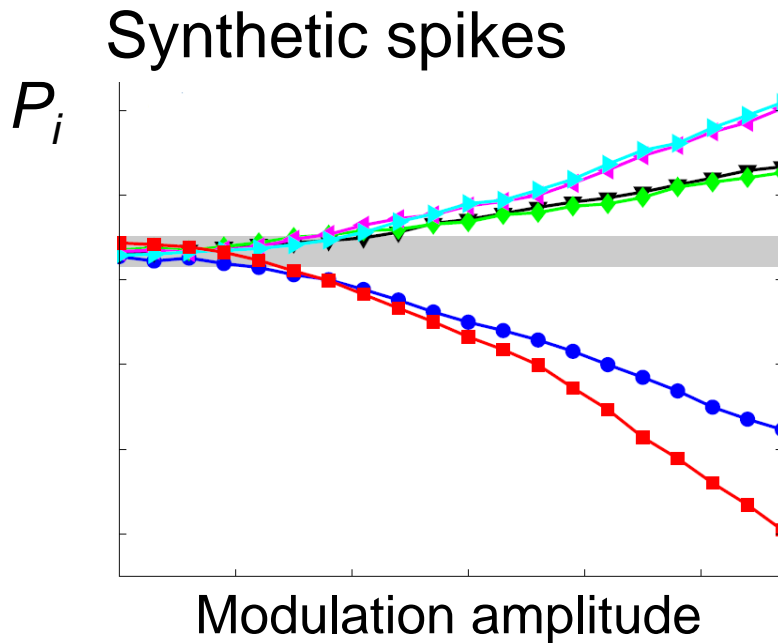




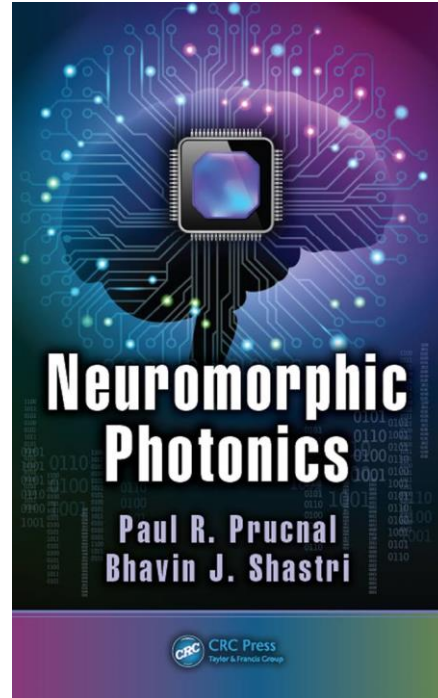
# Comparing with synthetic neuronal spikes: good agreement

Neuron model with Gaussian white noise and **weak** sinusoidal input: spikes are noise-induced

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



# Uncovering similarities can be interesting. But useful? Maybe...



- Excitable lasers: building blocks of **photonic neurons**.
- Diode lasers: very very low cost, highly energy efficient.
- Very very fast.
- Main challenge: understand how lasers and neurons encode a **weak periodic signal** in the presence of **noise**.

**How to quantify unpredictability  
and complexity?**

# Methods of time-series analysis

- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
- Spatio-temporal representation of a time-series
- Instantaneous phase and amplitude

# Information measure: Shannon entropy

- The time-series is described by a set of probabilities  $\sum_{i=1}^N p_i = 1$

- **Shannon** entropy:  $H = -\sum_i p_i \log_2 p_i$

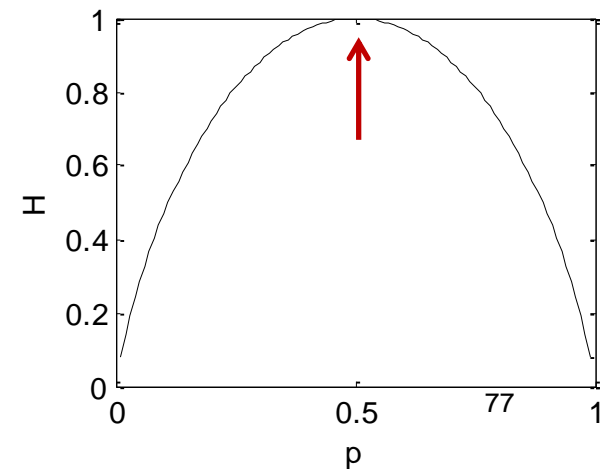
- Interpretation: “quantity of **surprise** one should feel upon reading the result of a measurement”

K. Hlavackova-Schindler et al, Physics Reports 441 (2007)

- Simple example: a random variable takes values 0 or 1 with probabilities:  $p(0) = p$ ,  $p(1) = 1 - p$ .

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

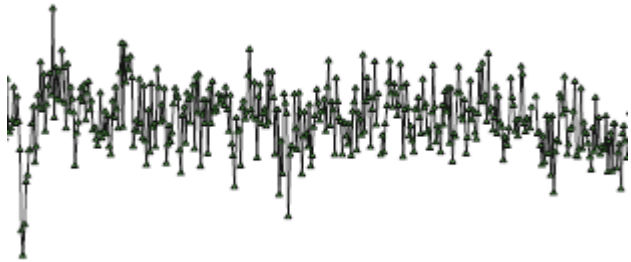
⇒  $p=0.5$ : Maximum **unpredictability**.



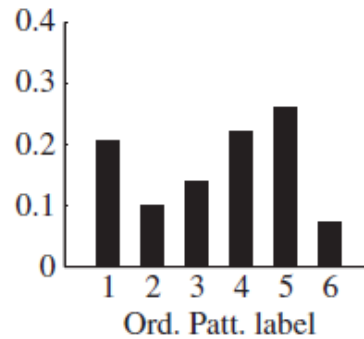
# Permutation entropy

- Entropy computed from ordinal probabilities.
- Number of bins = # of ordinal patterns (D!)

Time series



Ordinal probabilities



Permutation entropy

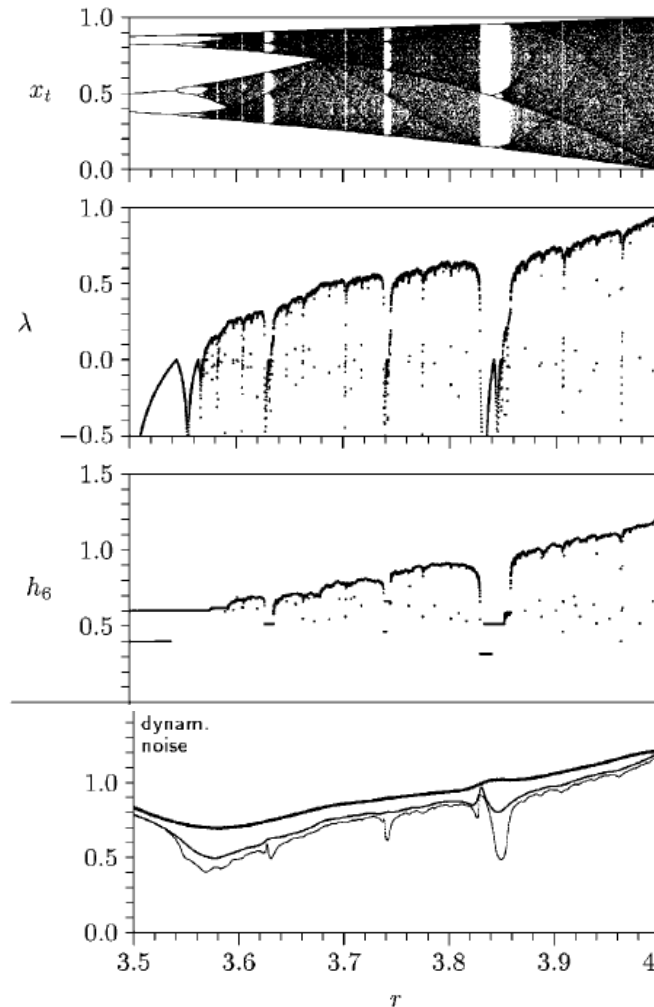
$$H = -\sum_i p_i \log_2 p_i$$

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



# Permutation entropy (PE) of the Logistic map

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



$$|\delta_n| \approx |\delta_0| e^{n\lambda}$$

$$\lambda = \lim_{n \rightarrow \infty} \left\{ \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)| \right\}$$

Entropy per symbol:

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

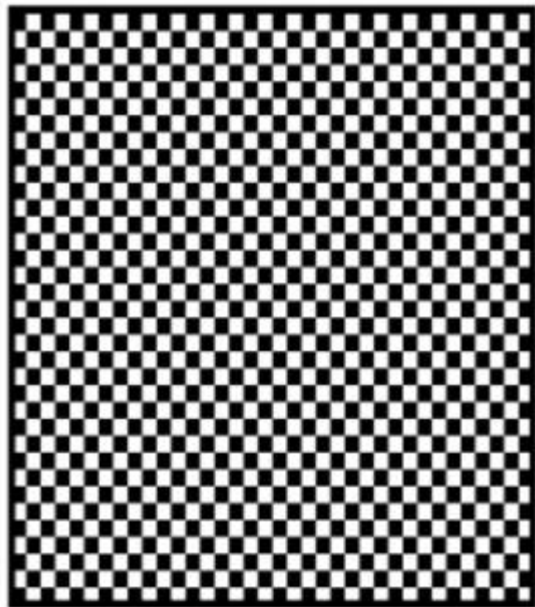
Robust to noise

*Bandt and Pompe  
Phys. Rev. Lett. 2002*

**Entropy: measures unpredictability or disorder.  
How to quantify Complexity?**

We would like to find a quantity “C” that measures **complexity**, as the entropy, “H”, measures **unpredictability**, and, for low-dimensional systems, the Lyapunov exponent measures **chaos**.

**Order**



$$H = 0$$
$$C = 0$$

**Chaos**



$$H \neq 0$$
$$C \neq 0$$

**Disorder**



$$H = 1$$
$$C = 0$$



*“A useful complexity measure needs to do more than satisfy the boundary conditions of vanishing in the high- and low-entropy limits.”*

*“Maximum complexity occurs in the region between the system’s perfectly ordered state and the perfectly disordered one.”*

## Information measures

- Assuming that we know the set of probabilities  $P=[p_i, i=1, Nbin]$  that characterizes a time series, several information measures have been proposed, a few popular ones are:

Shannon entropy  $I[P] = S_S[P] = -\sum_i p_i \ln p_i$

Tsallis entropy  $I[P] = S_T^q[P] = \frac{1}{q-1} \left[ 1 - \sum_i p_i^q \right]$

Renyi entropy  $I[P] = S_R^q[P] = \frac{1}{1-q} \ln \left[ \sum_i p_i^q \right]$

Fischer information  $I[P] = F[P] = F_0 \ln \left[ \sum_{i=1}^{Nbin-1} \left( p_{i+1}^{1/2} - p_i^{1/2} \right)^2 \right]$

## Normalization

$$H[P] = \frac{I[P]}{I_{\max}}$$

$$0 \leq H[P] \leq 1$$

where  $I_{\max} = I[P_e]$  and  $P_e$  is the equilibrium probability distribution (that maximizes the information measure).

Example: if  $I[P] = \text{Shannon entropy}$

then  $P_e = [p_i = 1/Nbin \text{ for } i=1, Nbin]$

and  $I_{\max} = \ln(Nbin)$

## Disequilibrium Q

Measures the “distance“ from  $P$  to the equilibrium distribution,  $P_e$

$$Q[P] = Q_0 D[P, P_e]$$

where  $Q_0$  is a normalization constant such that

$$0 \leq Q[P] \leq 1$$

# Distance between two distributions $P$ and $P_e$

Euclidean  $D_E[P, P_e] = \|P - P_e\|_E = \sum_i (p_i - p_{i,e})^2$

Kullback–Leibler divergence  
(relative entropy)

$$D_{KL}[P, P_e] = \sum_i p_i \log \frac{p_i}{p_{i,e}}$$

Jensen divergence  $D_J[P, P_e] = \frac{K[P|P_e] + K[P_e|P]}{2}$

**Read more:** [S-H Cha: Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions, Int. J of. Math. Models and Meth. 1, 300 \(2007\)](#)

# Statistical complexity measure C

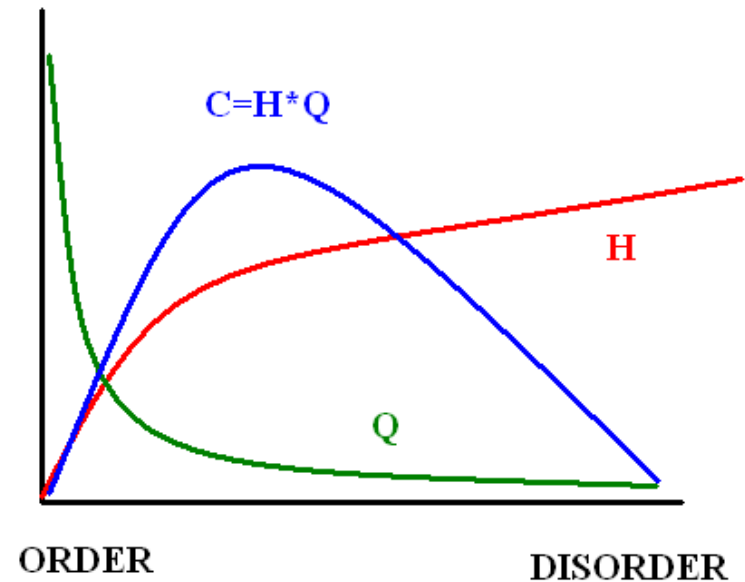
A family of complexity measures can be defined as:

$$C[P] = H_A[P] \cdot Q_B[P]$$

where

A = S, T, R (Shannon, Tsallis, Renyi)

B = E, K, J (Euclidean, Kullback, Jensen)



$$C_{LMC}[P] = H_S[P] \cdot Q_E[P]$$

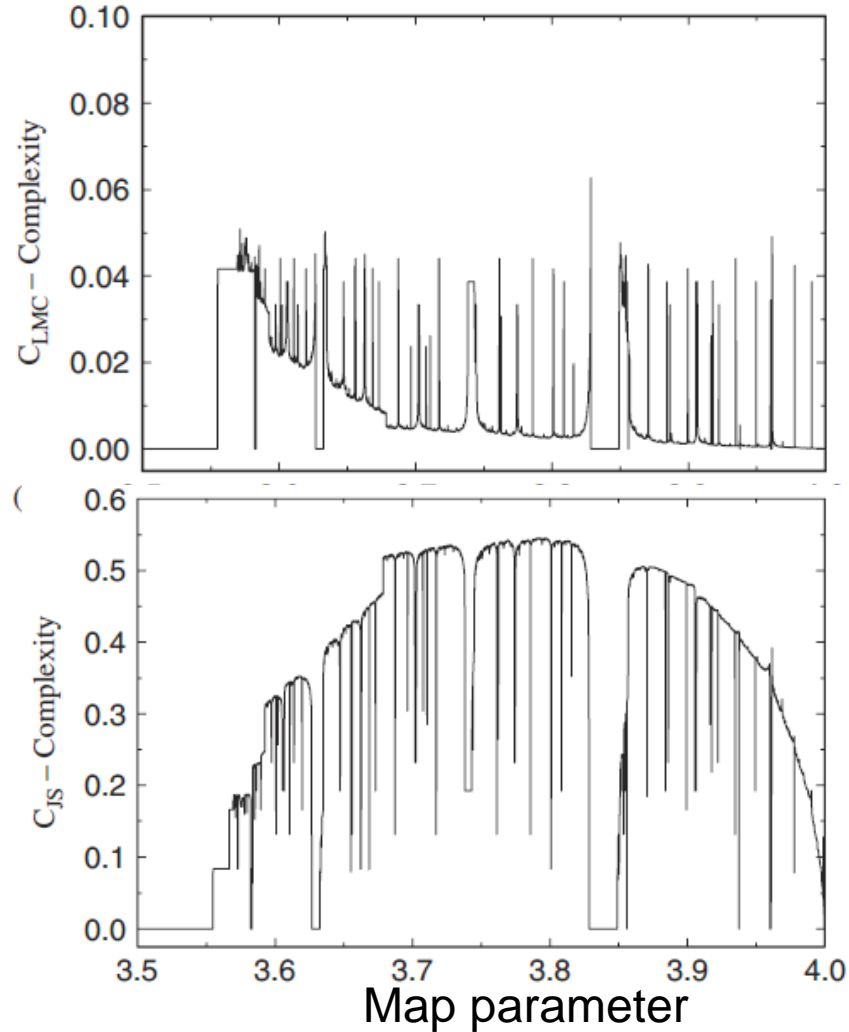
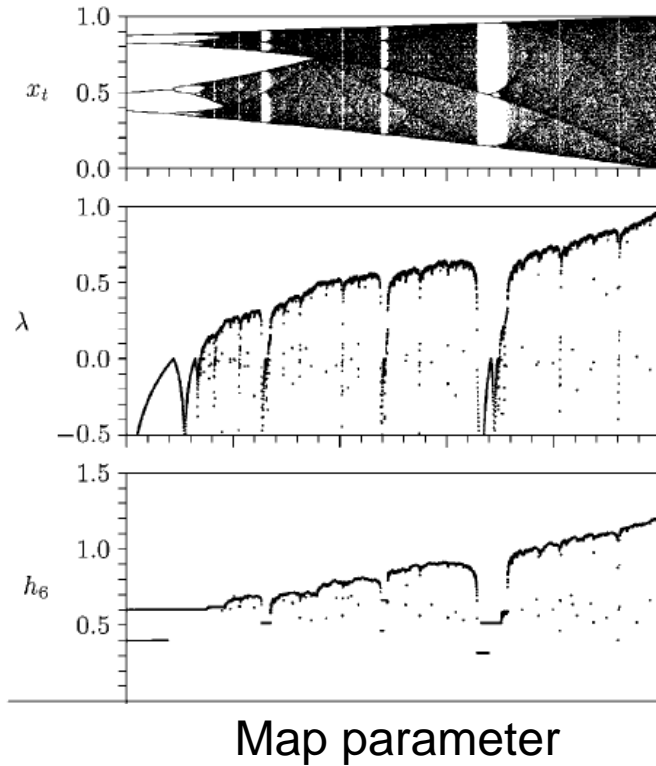
Lopez-Ruiz, Mancini & Calbet, Phys. Lett. A (1995).  
Anteneodo & Plastino, Phys. Lett. A (1996).

$$C_{MPR}[P] = H_S[P] \cdot Q_J[P]$$

Martín, Plastino & Rosso, Phys. Lett. A (2003).

# The complexity of the Logistic Map

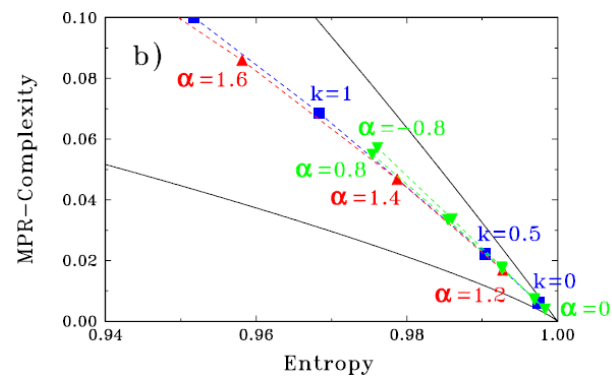
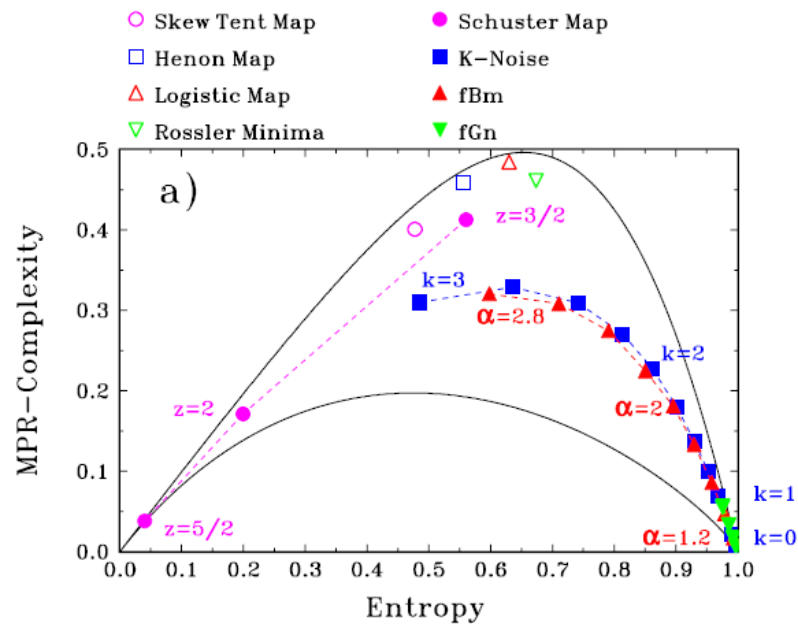
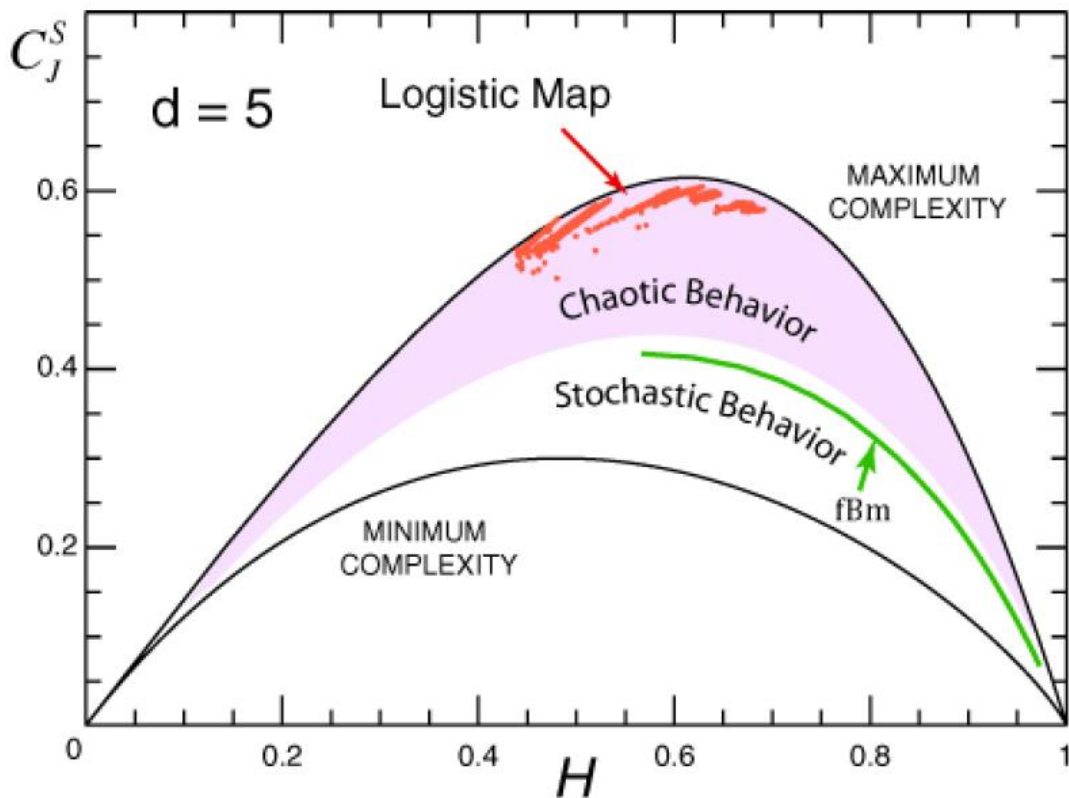
$$x(i+1) = r x(i)[1-x(i)]$$



Euclidian  
distance

Jensen  
distance

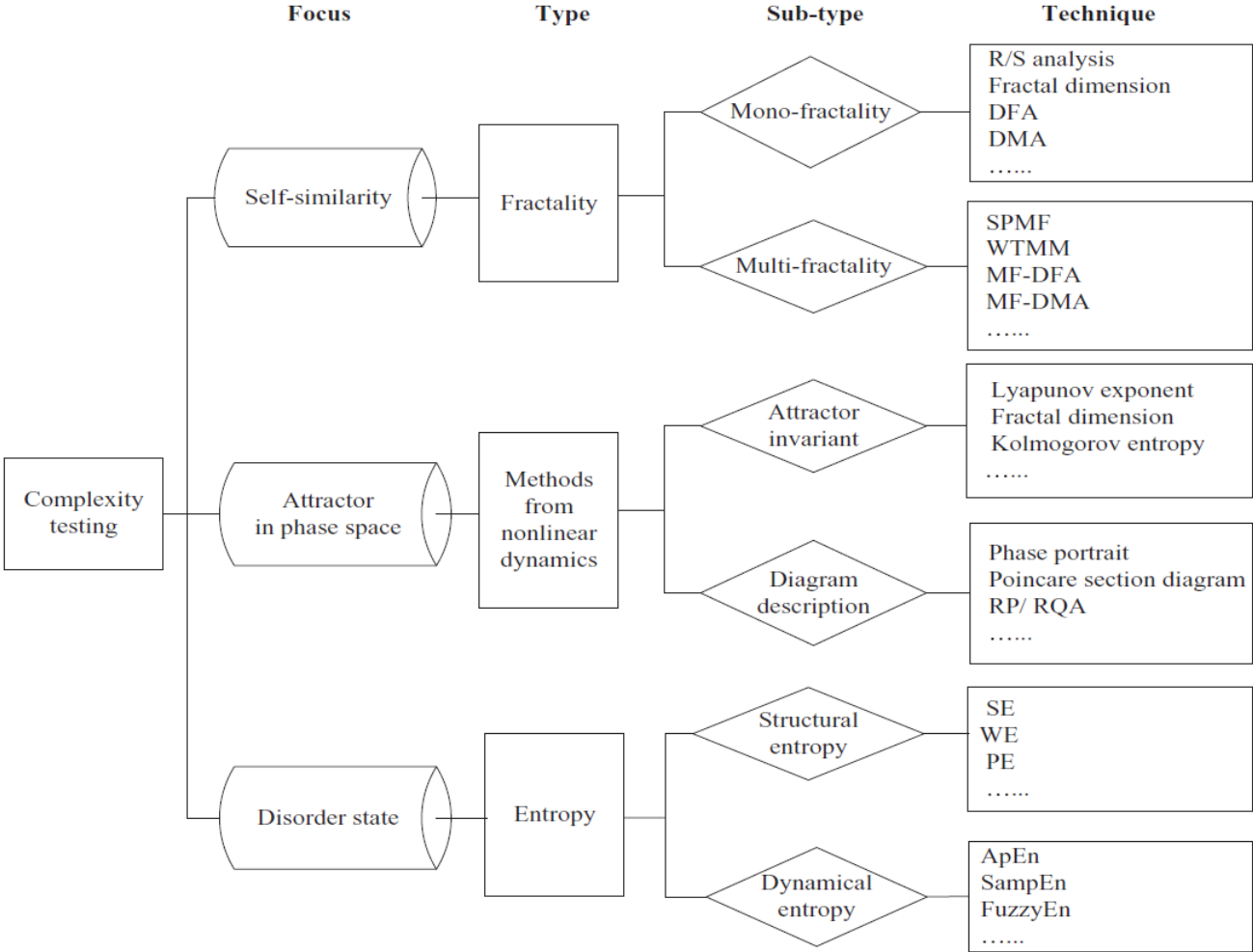
# The (entropy, complexity) plane: a useful tool to distinguish noise from chaos



O. A. Rosso et al, Phys. Rev. Lett. 99, 154102 (2007)



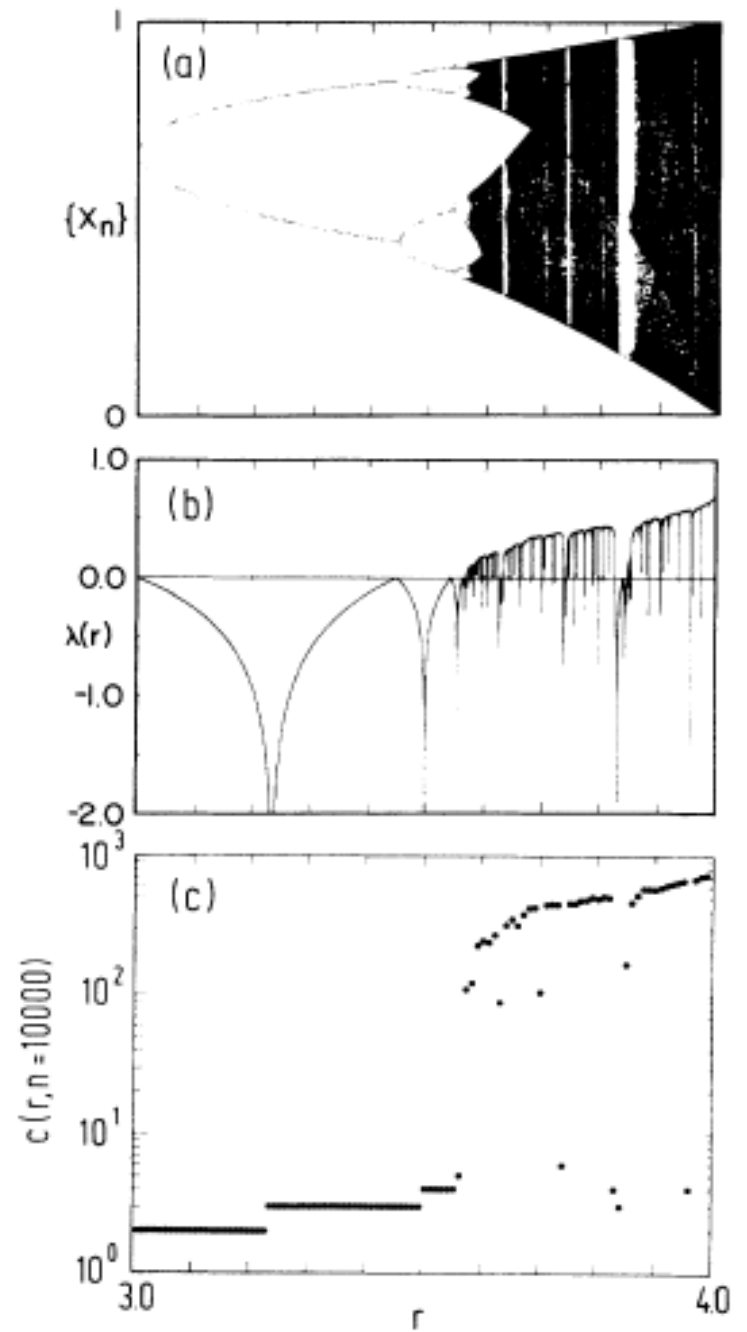
# Many complexity measures have been proposed



Further reading: L. Tang et al, “Complexity testing techniques for time series data: A comprehensive literature review”, Chaos, Solitons and Fractals 81 (2015) 117–135



# Lempel & Zip complexity of the Logistic Map



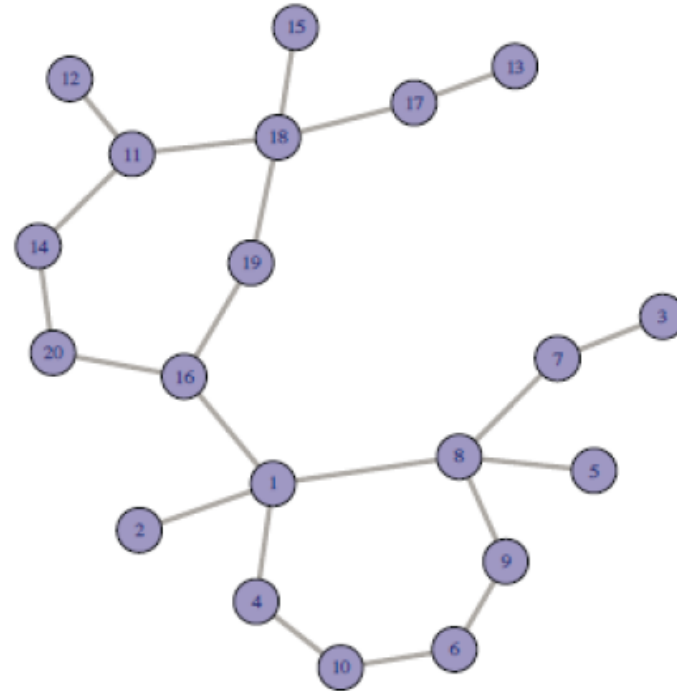
*Kaspar and Schuster, Phys Rev. A 1987*

# Methods of time-series analysis

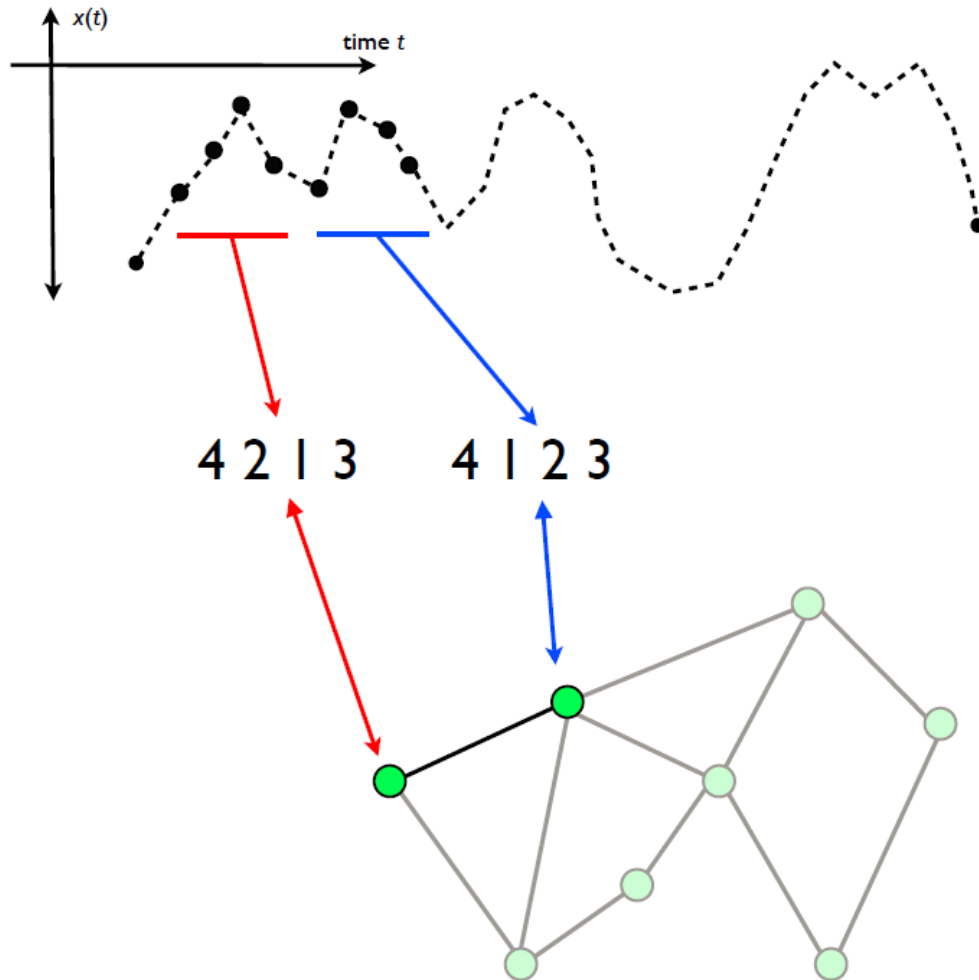
- Return maps
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Distribution of data values
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
- Spatio-temporal representation of a time-series
- Instantaneous phase and amplitude

# What is a network?

- A graph: a set of “nodes” connected by a set of “links”
- Nodes and links can be weighted or unweighted
- Links can be directed or undirected
- More in part 3 (multivariate time series analysis)



We use symbolic patterns as the *nodes* of the network.  
And the *links*? Defined as the transition probability  $\alpha \rightarrow \beta$



- In each node  $i$ :

$$\sum_j w_{ij}=1$$

- Weigh of node  $i$ : the probability of pattern  $i$

$$(\sum_i p_i=1)$$

⇒ **Weighted and directed network**

# Network-based diagnostic tools

- Entropy computed from node weights (**permutation entropy**)

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

- Average node entropy (entropy of the link weights)

$$s_n = \frac{1}{M} \sum_{i=1}^M s_i \quad s_i = -\sum w_{ij} \log w_{ij}$$

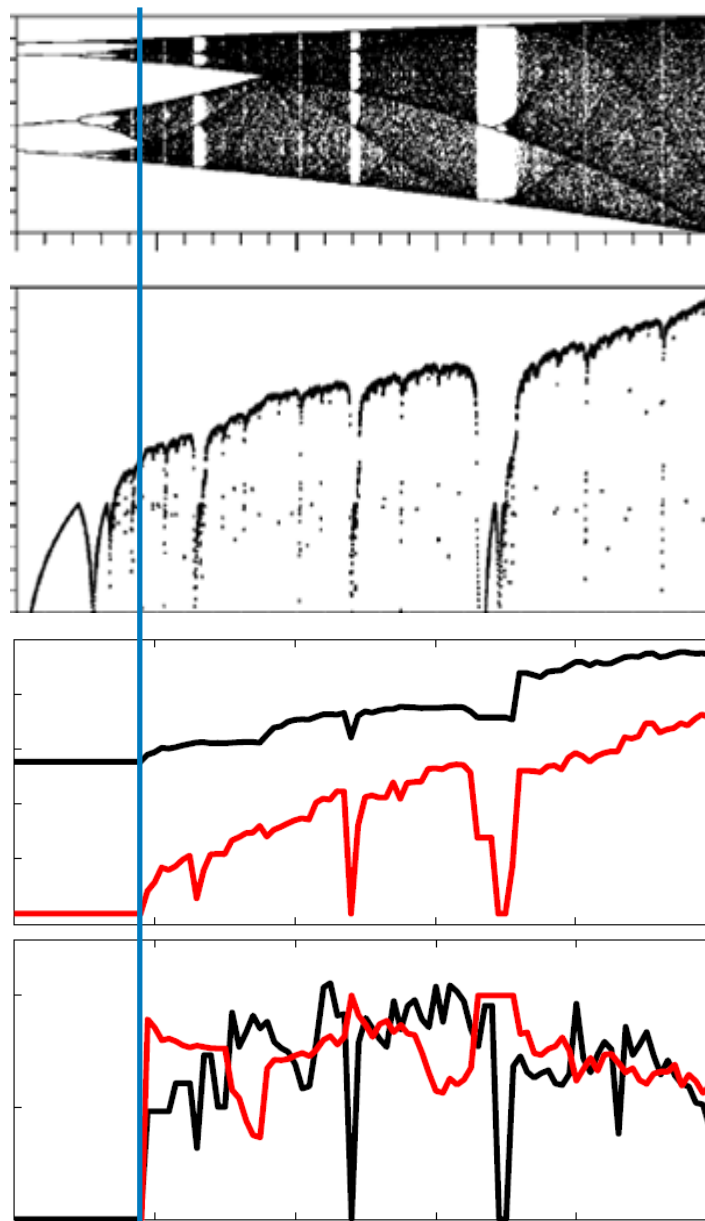
- Asymmetry coefficient: normalized difference of transition probabilities,  $P('01' \rightarrow '10') - P('10' \rightarrow '01')$ , etc.

$$a_c = \frac{\sum_i \sum_{j \neq i} |w_{ij} - w_{ji}|}{\sum_i \sum_{j \neq i} (w_{ij} + w_{ji})} \quad \begin{array}{l} (0 \text{ in a fully symmetric network;} \\ 1 \text{ in a fully directed network)} \end{array}$$

# A first test with the Logistic map

$D=4$

⇒ Detects the merging of four branches, not detected by the Lyapunov exponent.



Lyapunov exponent

$S_p = PE$   
 $S_n = S(TPs)$

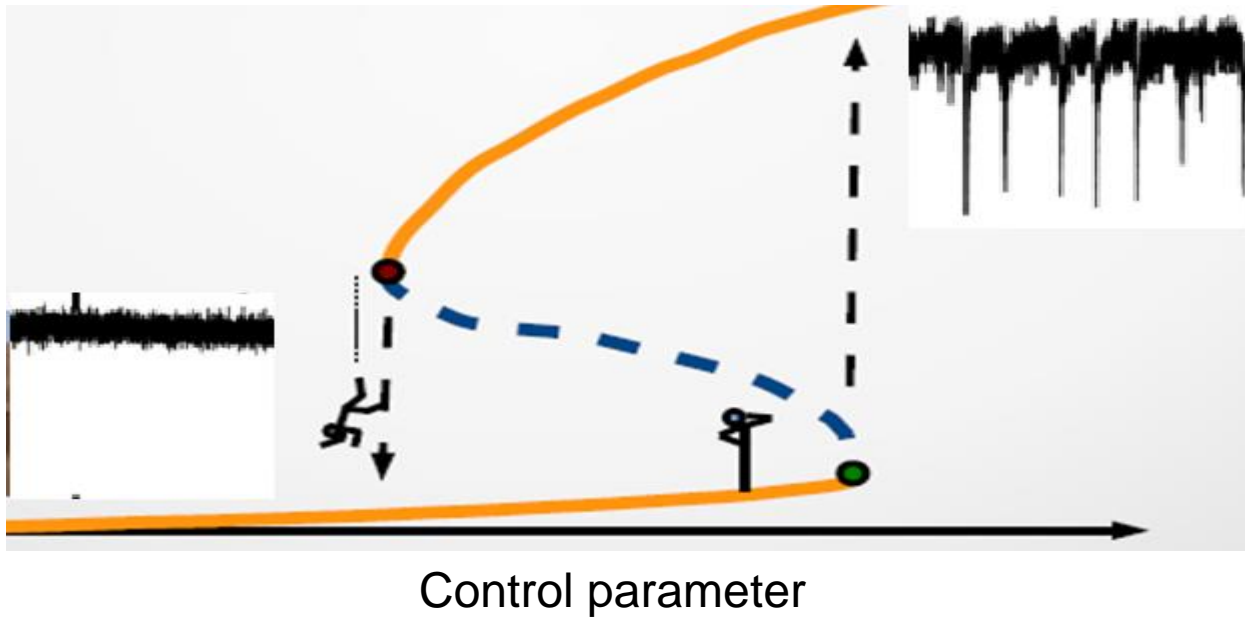
$S_{links}$   
 $a_c$

Map parameter

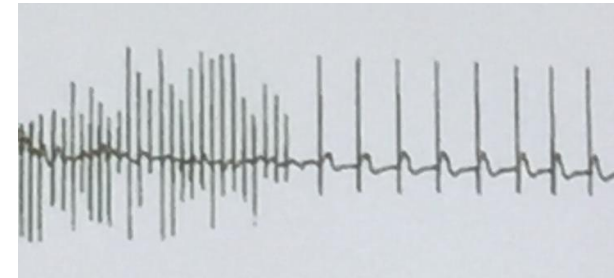
[C. Masoller et al, NJP \(2015\)](#)



# Approaching a “tipping point”



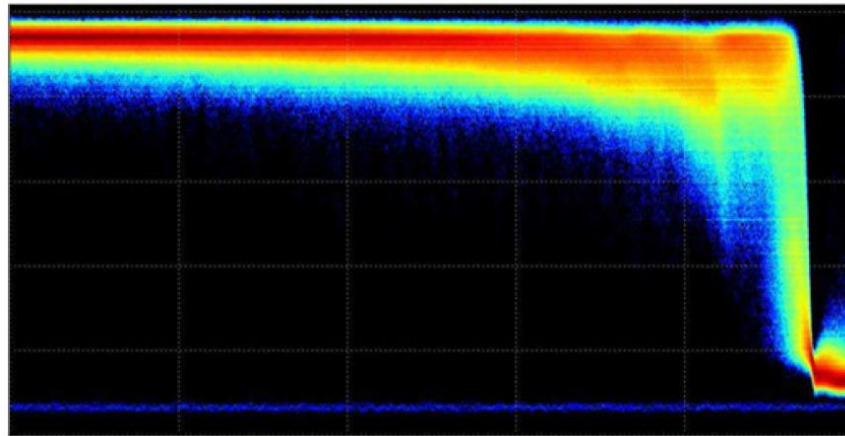
Can we use the ordinal network method to detect an *early warning signal* of a transition to a different dynamical regime?



Yes! Two examples: optical signals and brain signals

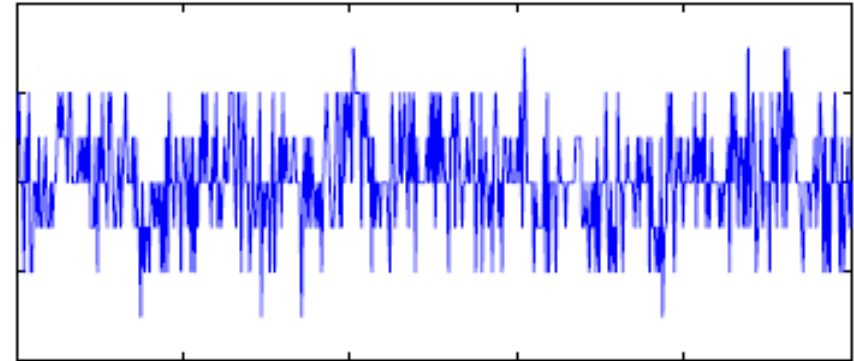
# Apply the ordinal network method to laser data

As the laser current increases



Time

Intensity @ constant current



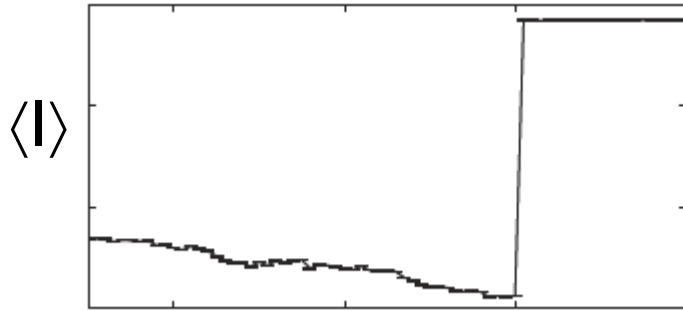
Time

- Two sets of experiments: intensity time series were recorded
  - keeping constant the laser current.
  - while increasing the laser current.
- We analyzed the polarization that turns on / turns off.

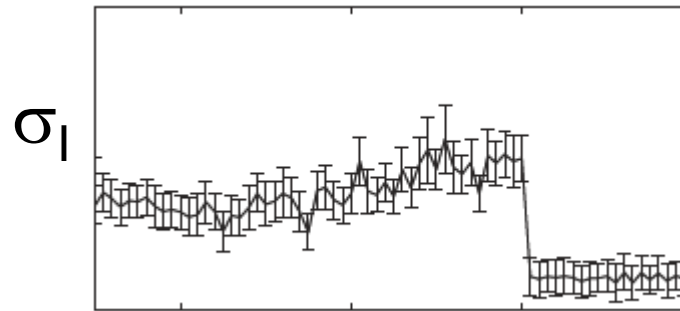
Is it possible to anticipate the switching?

No if the switching is fully stochastic.

**First set of experiments (the current is kept constant):  
despite of the stochasticity of the time-series, the node  
entropy “anticipates” the switching**



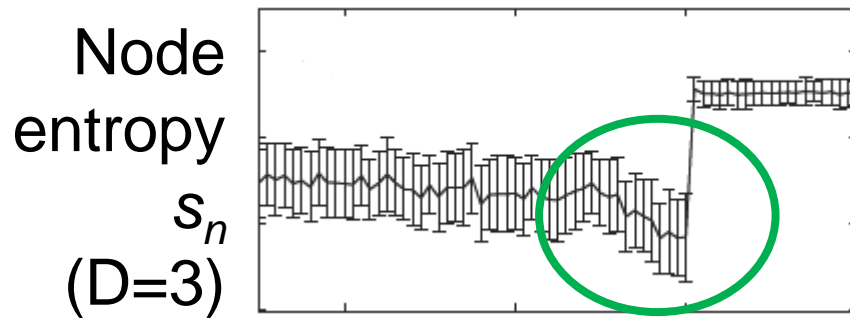
Laser current



Laser current

$L=1000$   
100 windows

$\Rightarrow$  **No  
warning**

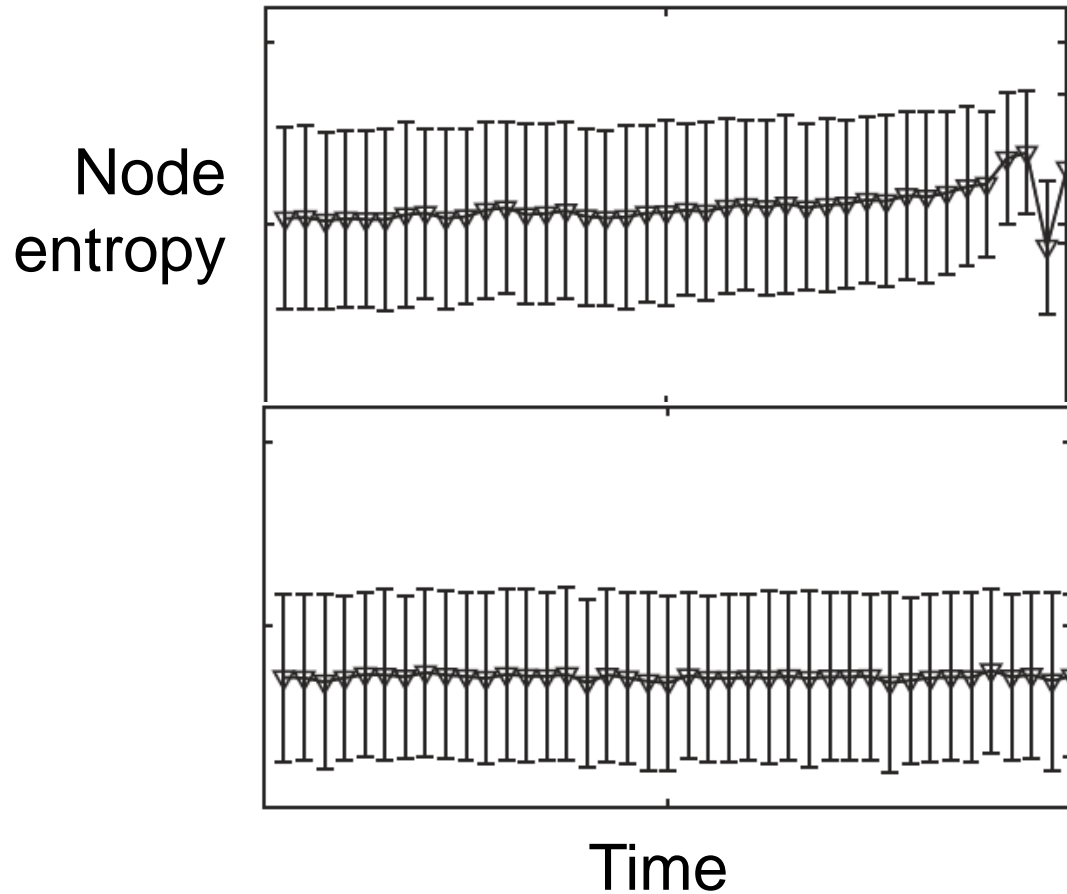


Laser current

**Early warning**

$\Rightarrow$  Deterministic mechanisms  
must be involved.

**In the second set of experiments (current increases linearly in time): an early warning is also detected**



$L=500, D=3$   
1000 time series

With slightly different experimental conditions: no switching.

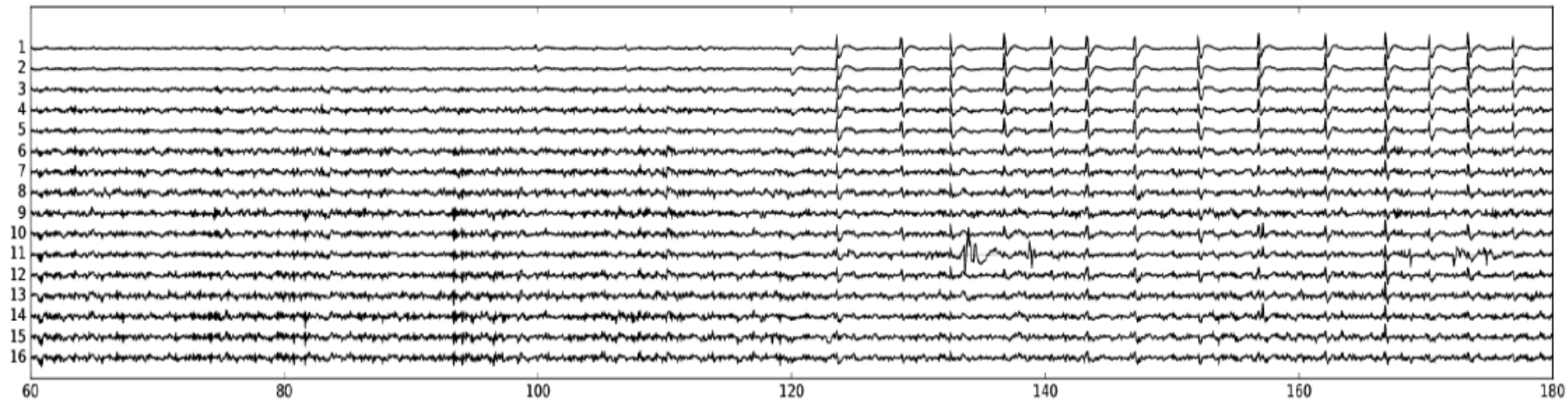
## Second application of the ordinal network method: distinguishing *eyes closed* and *eyes open* brain states

Analysis of two EEG datasets

	BitBrain	PhysioNet
	DTS1	DTS2
Sampling rate(Hz)	256	160
Time task(seg)	120	60
Total points	30720	9600
Number of electrodes	16	64
Number of subjects	70	109

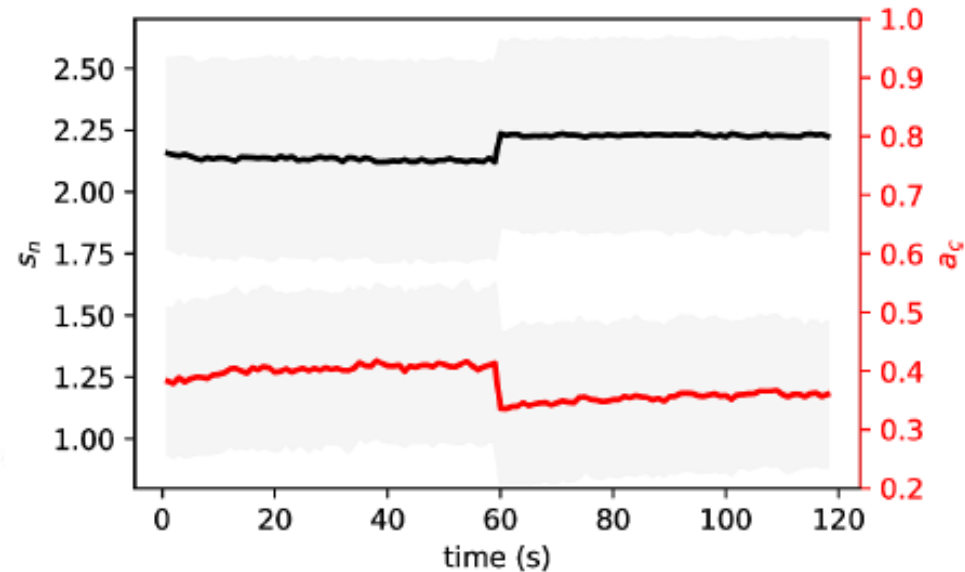
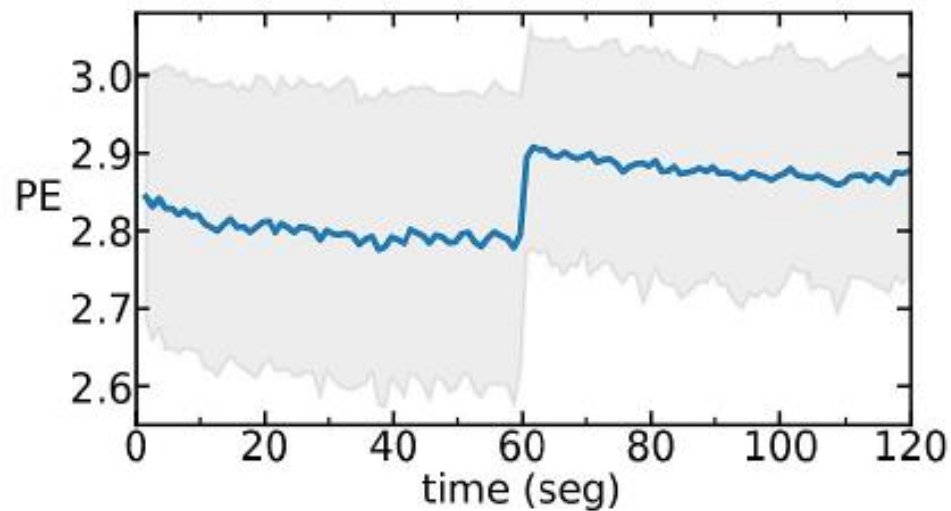
Eye closed

Eye open



- Symbolic analysis is applied to the **raw** data; similar results were found with **filtered** data using independent component analysis.

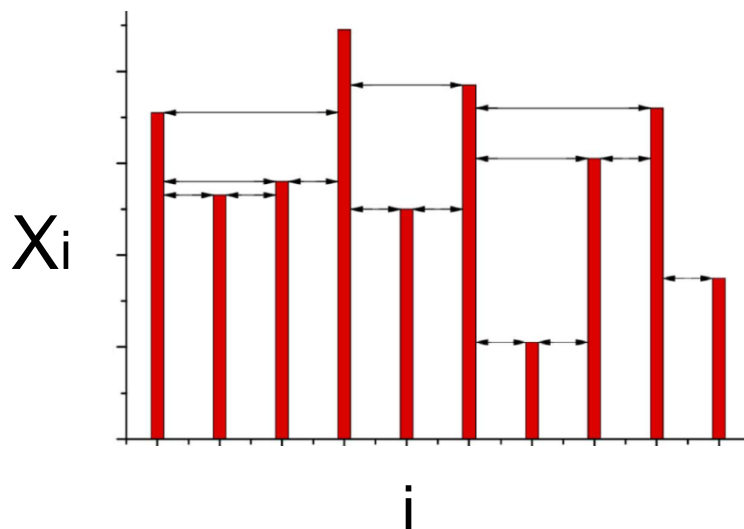
# “Randomization”: the entropies increase and the asymmetry coefficient decreases



Time window = 1 s  
(160 data points)

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

## Another way to represent a time series as a network: the horizontal visibility graph (HVG)



Rule: data points  $i$  and  $j$  are connected if there is “visibility” between them

⇒ **Unweighted and undirected graph**

Parameter free!

Luque et al PRE (2009); Gomez Ravetti et al, PLoS ONE (2014)



## Exercise

Consider the following time series:

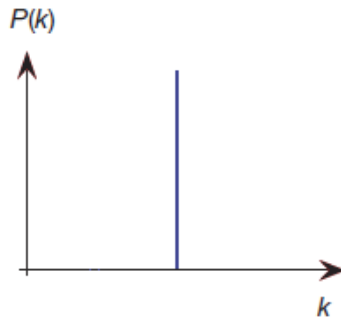
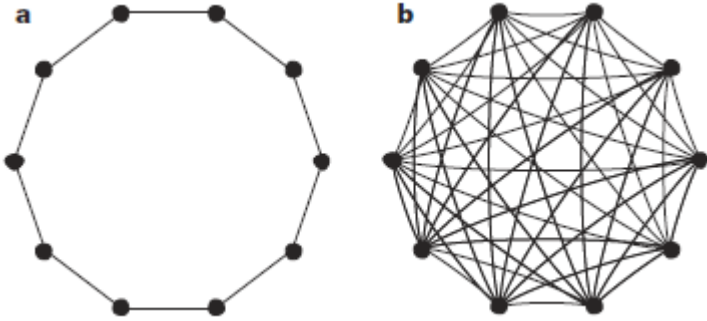
$$x = \{0.71, 0.53, 0.56, 0.89, 0.50, \\ 0.77, 0.21, 0.6, 0.72, 0.35\}.$$

How many links (“degree”) does each data point have?

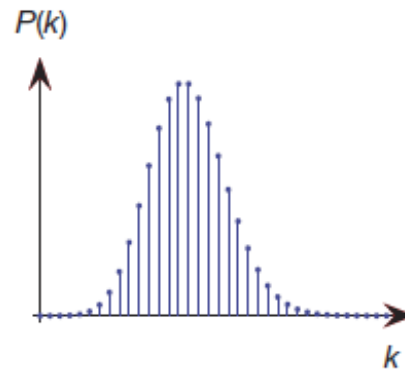
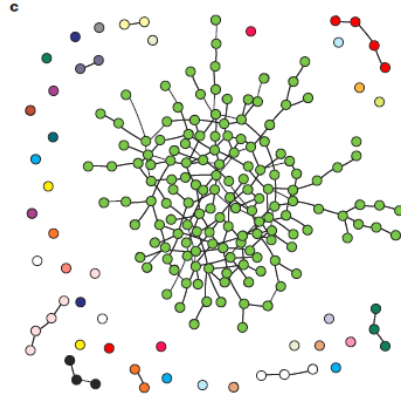
$$k = \{3, 2, 3, 4, 2, 5, 2, 3, 3, 1\}$$

# How to characterize the HV graph? The degree distribution

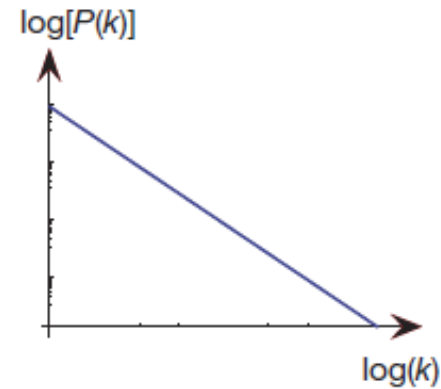
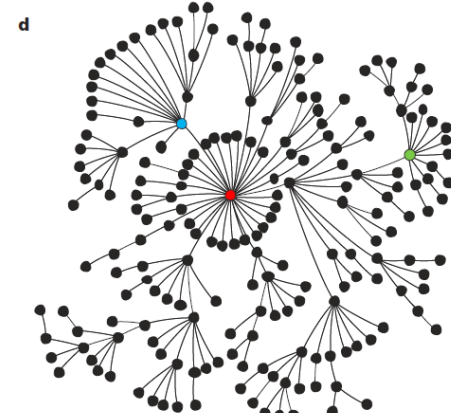
Regular



Random

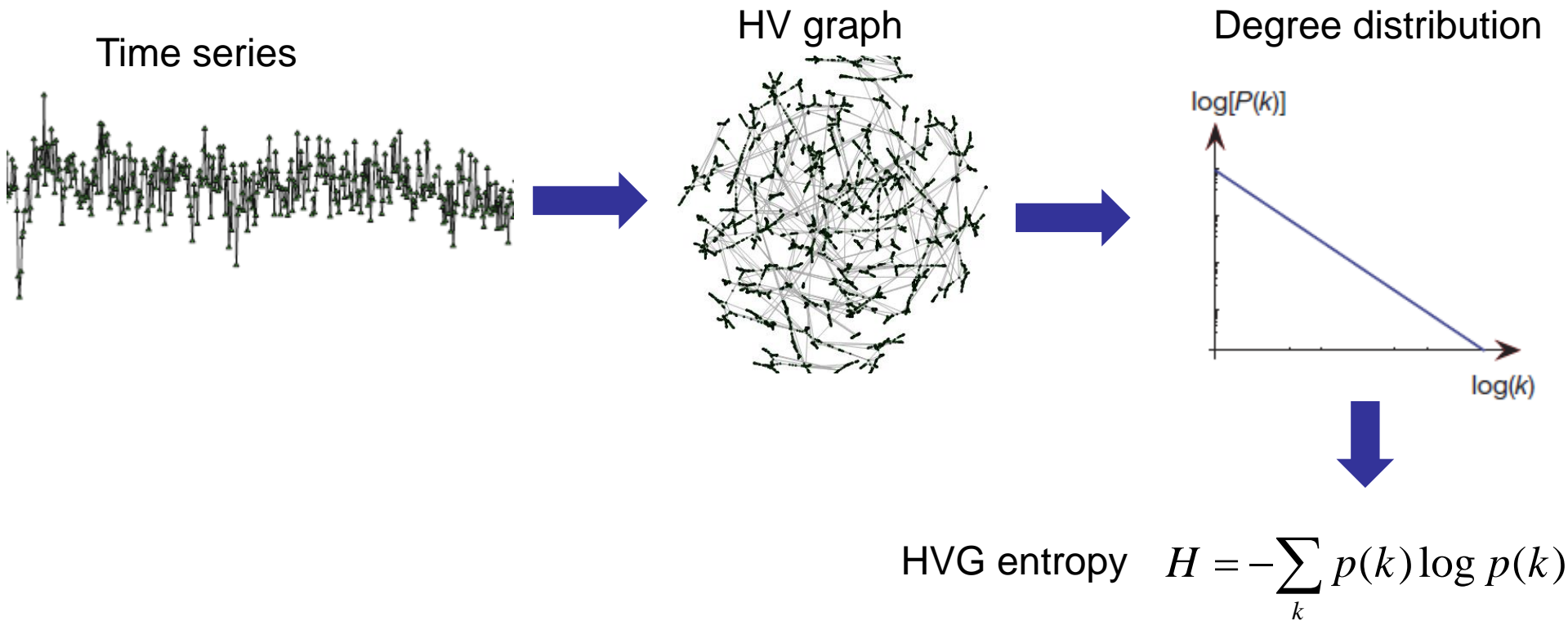


Scale-free

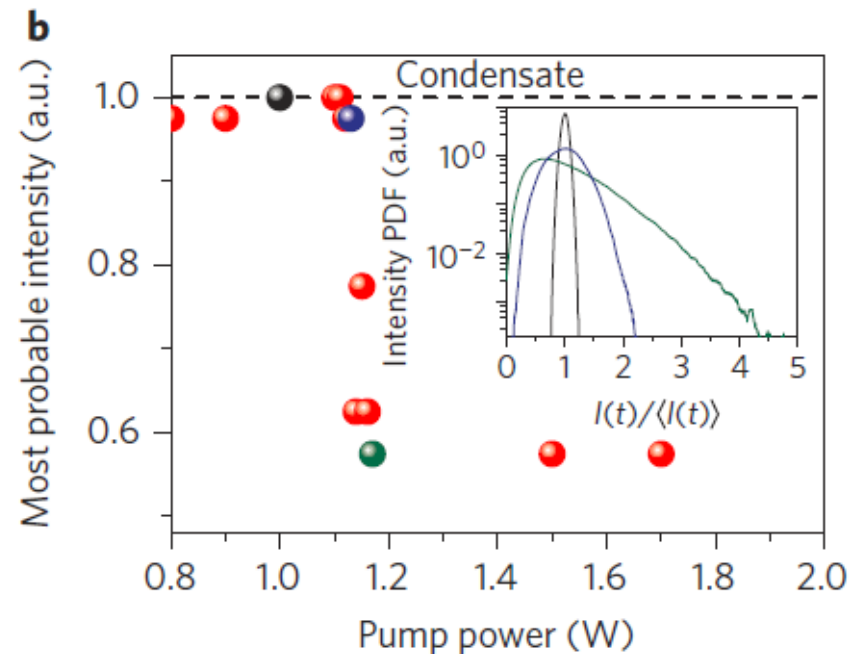
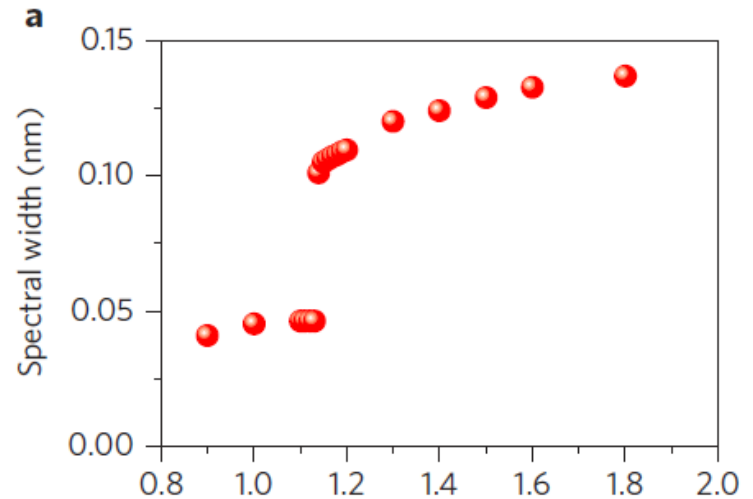
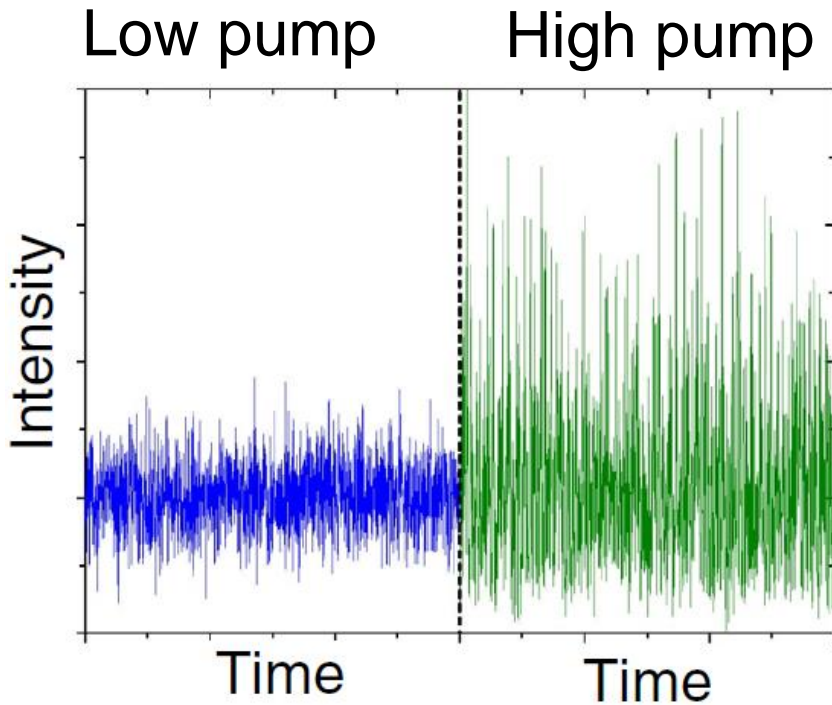


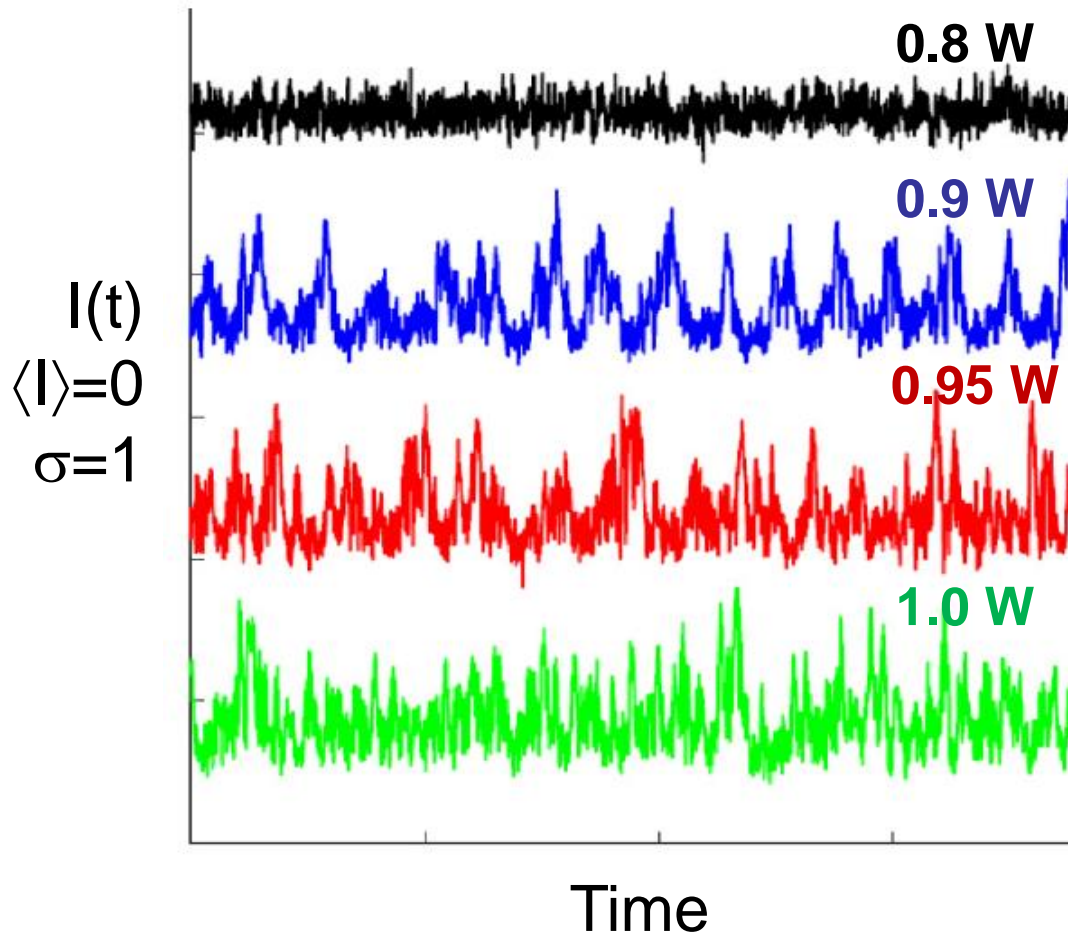
From the degree distribution of the horizontal visibility graph we calculate the entropy: **HVG entropy**

# HVG entropy: computed from the degree distribution



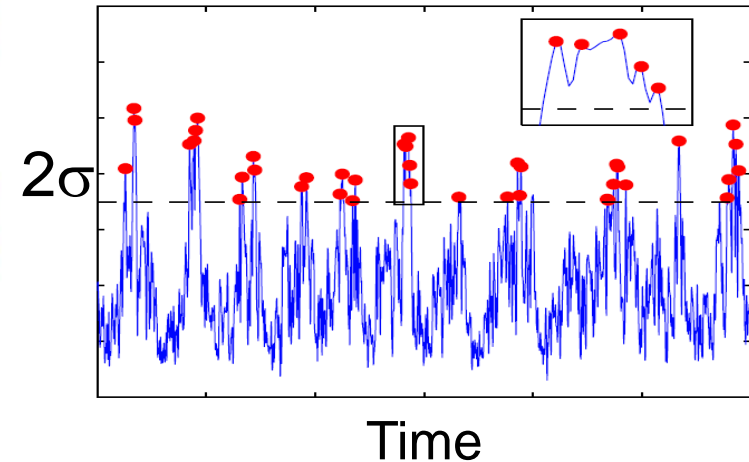
# Example of application: Laminar $\rightarrow$ Turbulence transition in a fiber laser as the pump (control parameter) increases





Nonlinear  
temporal  
correlations?

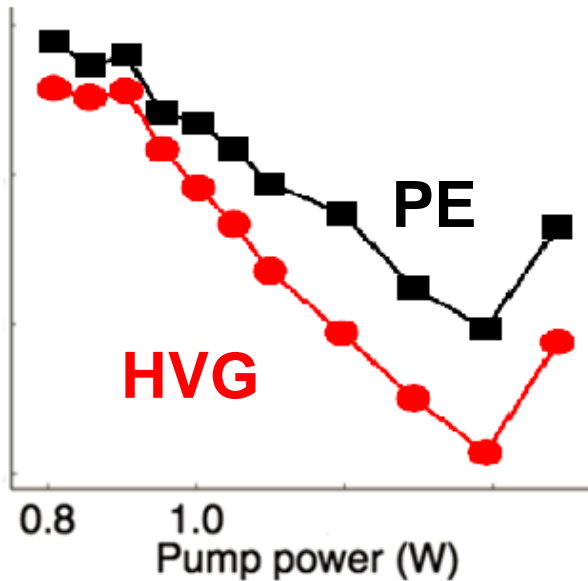
Raw and **thresholded** data



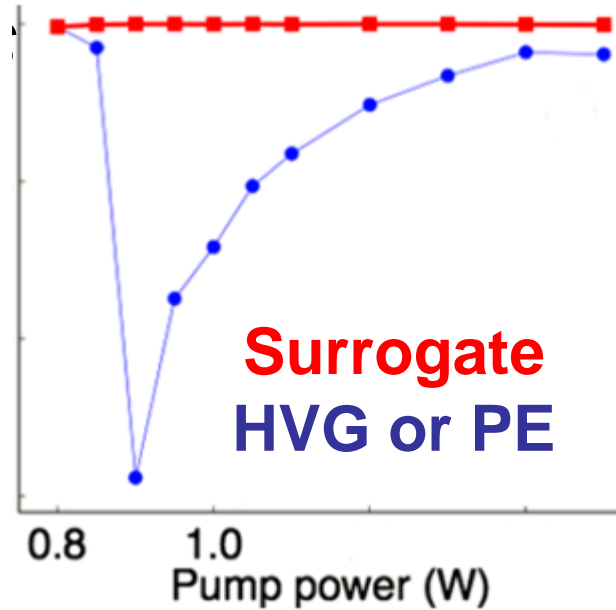
L. Carpi and C. Masoller, “*Persistence and stochastic periodicity in the intensity dynamics of a fiber laser during the transition to optical turbulence*”, Phys. Rev. A **97**, 023842 (2018).

# Four different ways to compute the entropy

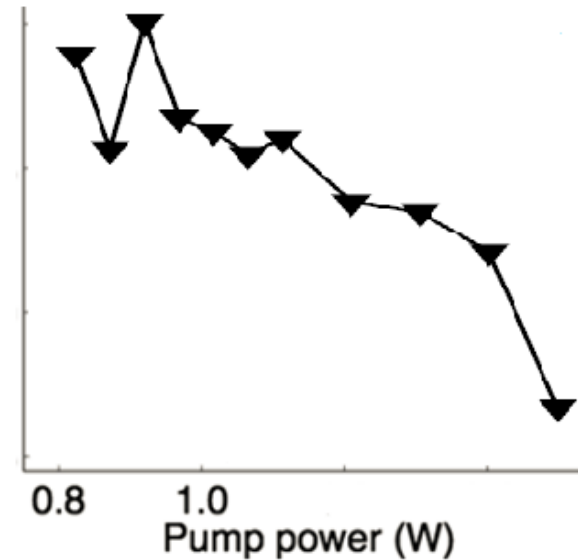
PE/HVG from  
“raw” data



“thresholded” data



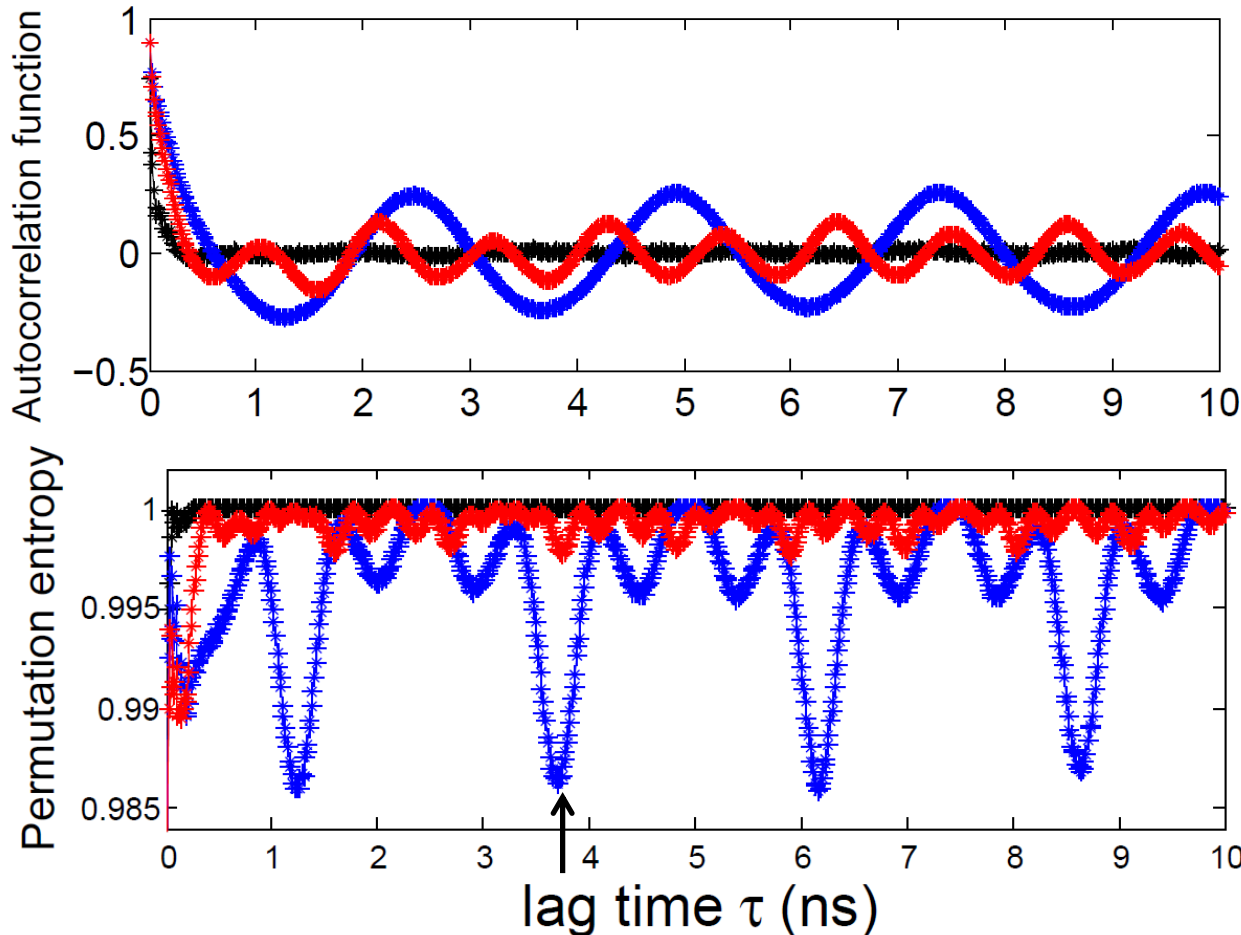
Histogram of  
“raw” values



(the abrupt transition is  
robust with respect to the  
selection of the threshold)

# Changing the sampling time $\tau$ identifies “hidden” time scales in the dynamics, undetected by correlation analysis

$$\{I_i, I_{i+\tau}, I_{i+2\tau}, \dots\};$$

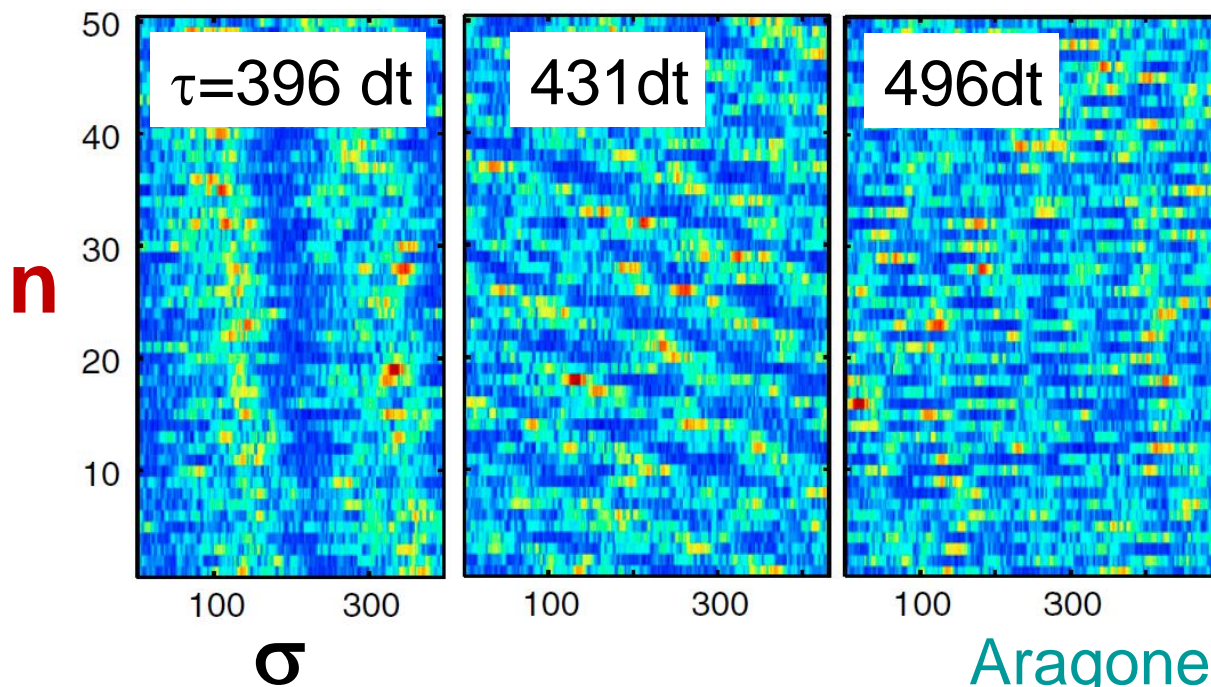
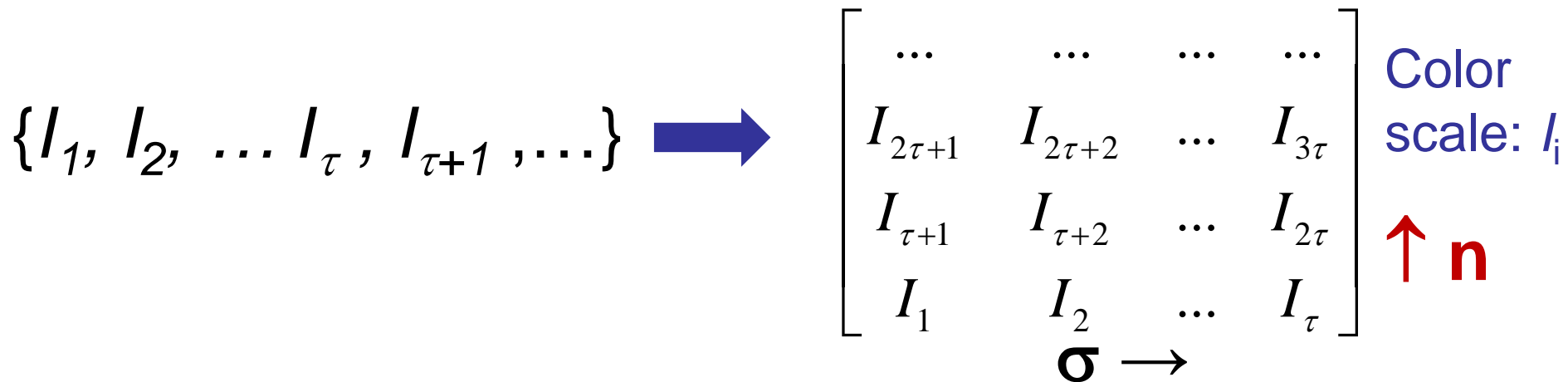


Below transition  
at transition  
above transition

# **Space-time representation of a time series**

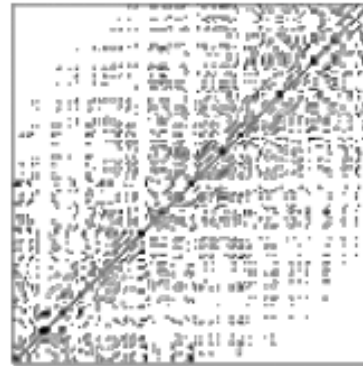
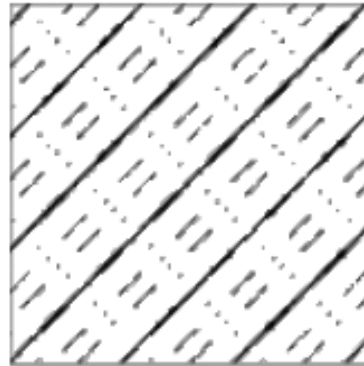
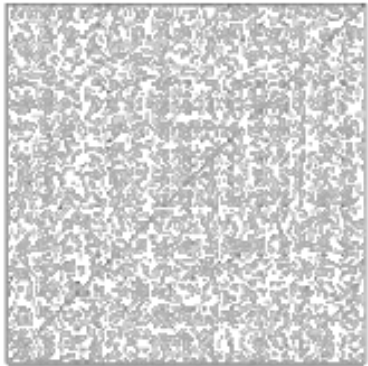
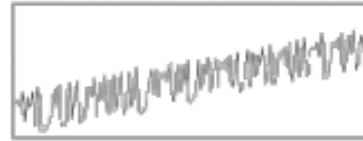
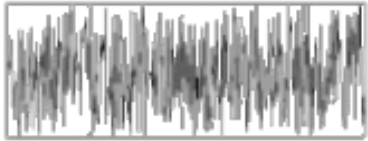


# The space-time representation of the intensity time series: a convenient way to visualize the dynamics



# Recurrence plots: another way to “visualize” a time series

- Embed the time series (find the embedding dimension and lag)
- Construct a binary matrix:  $A_{ij}=1$  if  $x_i$  and  $x_j$  are “close”, else  $A_{ij}=0$
- Plot  $A_{ij}$

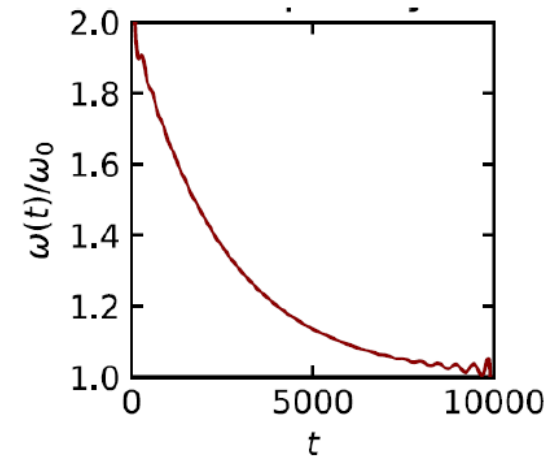
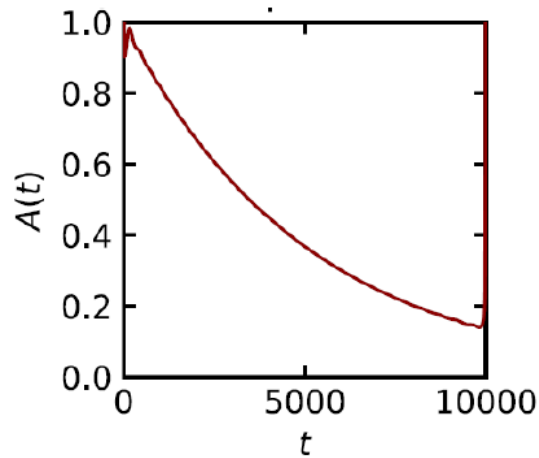
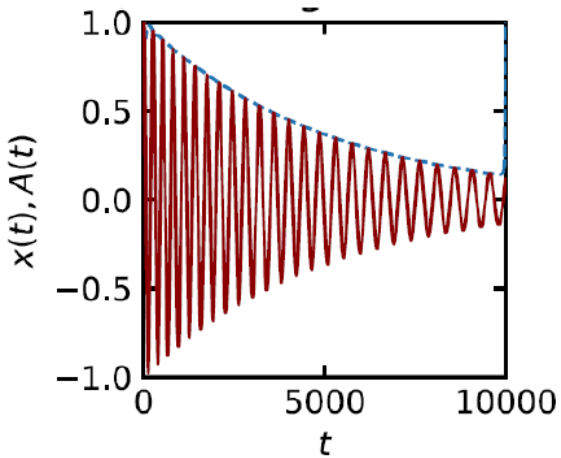


# Methods of time-series analysis

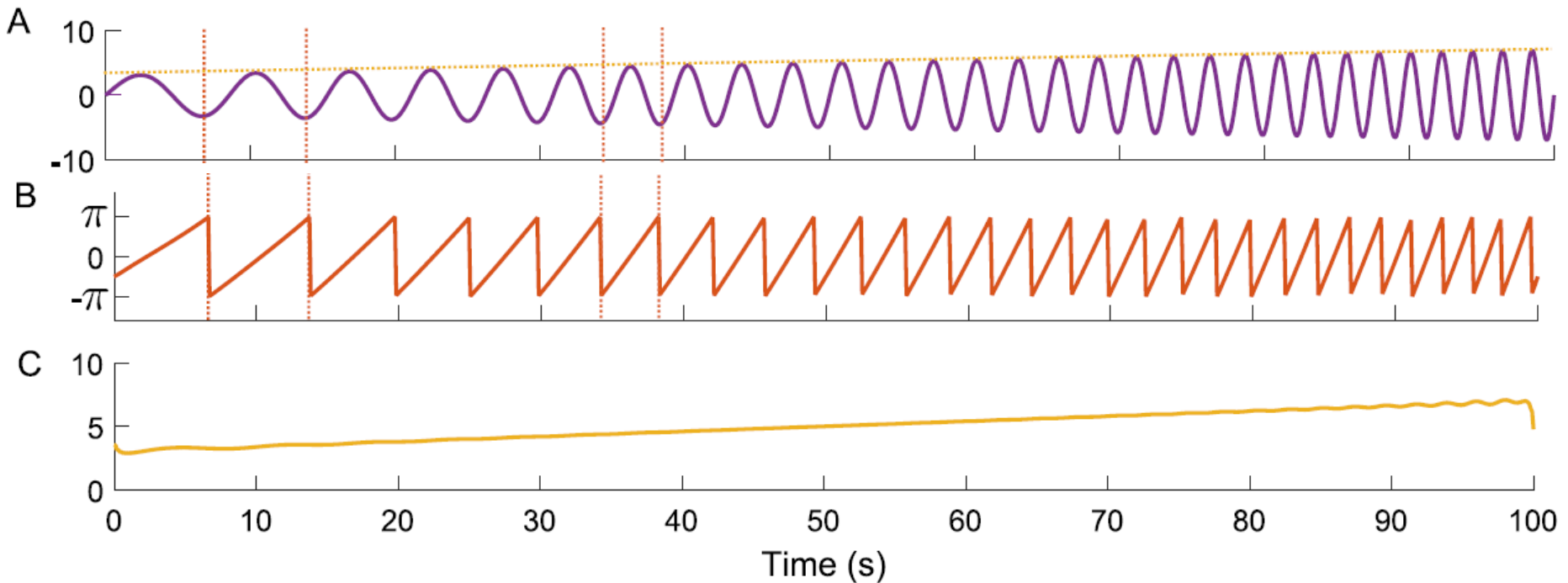
- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
- Spatio-temporal representation of a time-series
- **Instantaneous phase and amplitude**

# How to obtain instantaneous amplitude and frequency information from a time series?

$$x(t) = e^{-\alpha t} \cos[\omega_0(1 + e^{-2\alpha t})]$$

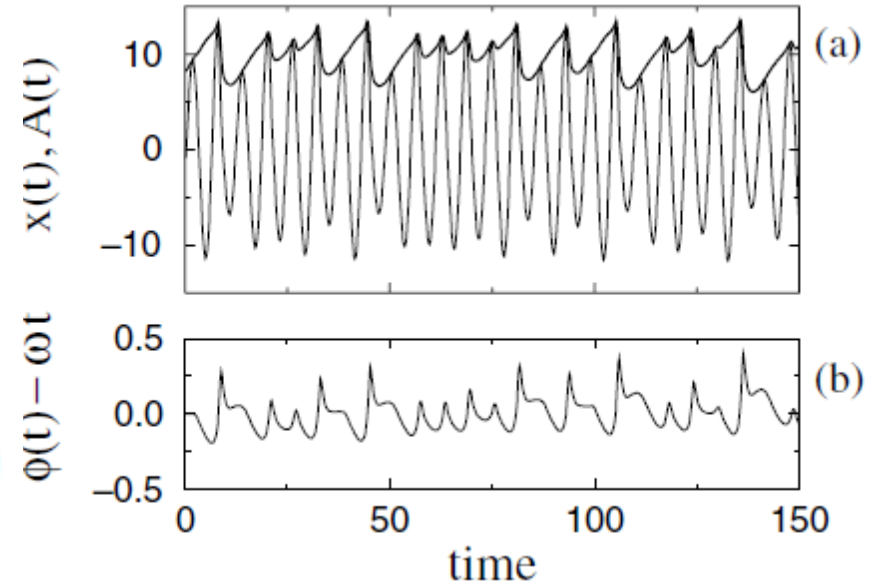
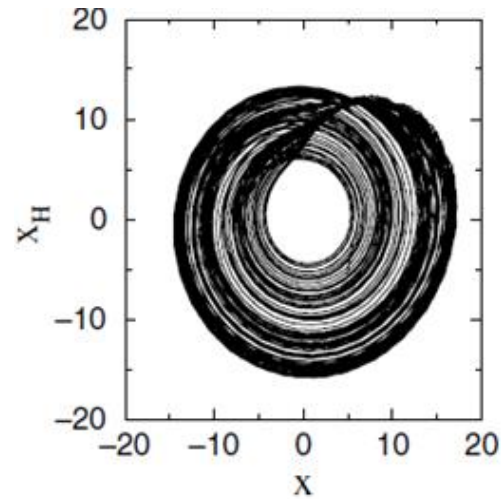
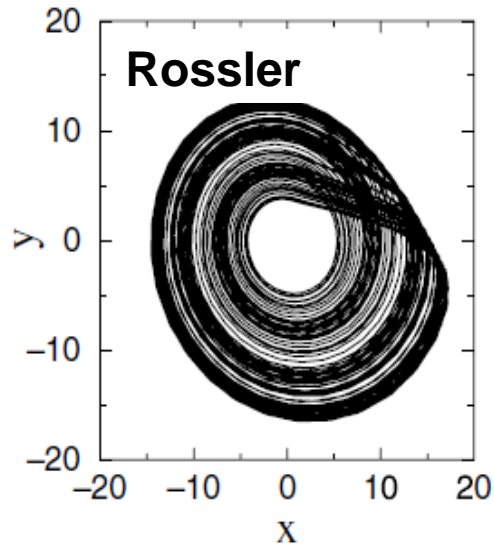


# Example: sine wave with increasing amplitude and frequency



- (A) The original signal. (B) The instantaneous phase extracted using the Hilbert transform. (C) The instantaneous amplitude.
- $A = C \cos(B)$ .

## Second example



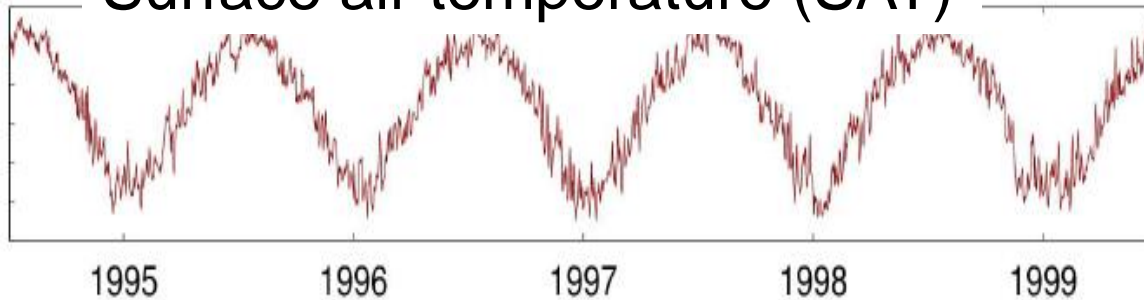
$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -y - z \\ x + ay \\ b + z(x - c) \end{pmatrix}$$

# Third example

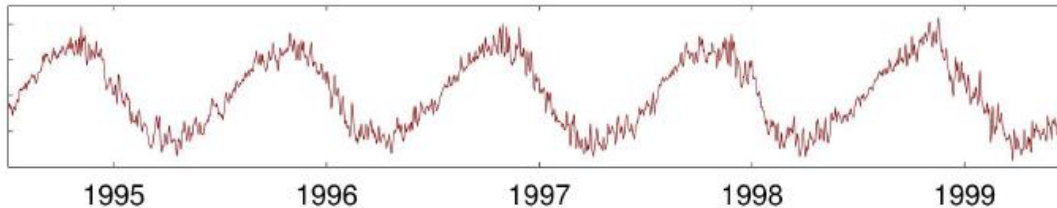
Surface air temperature (SAT)

**Normalization:**  $\mu=0, \sigma=1$

x

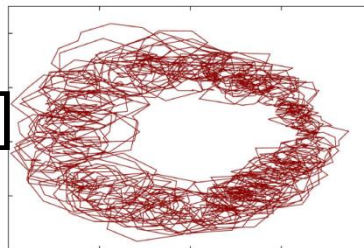


HT[x]



■  $HT[\sin(\omega t)] = \cos(\omega t)$

$y = HT[x]$



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



# Hilbert transform

- For a real time series  $x(t)$  defines an *analytic signal*

$$\zeta(t) = x(t) + iy(t) = a(t)e^{i\varphi(t)}$$
$$y(t) = H[x(t)] = \pi^{-1} \text{P.V.} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau$$

A word of warning:

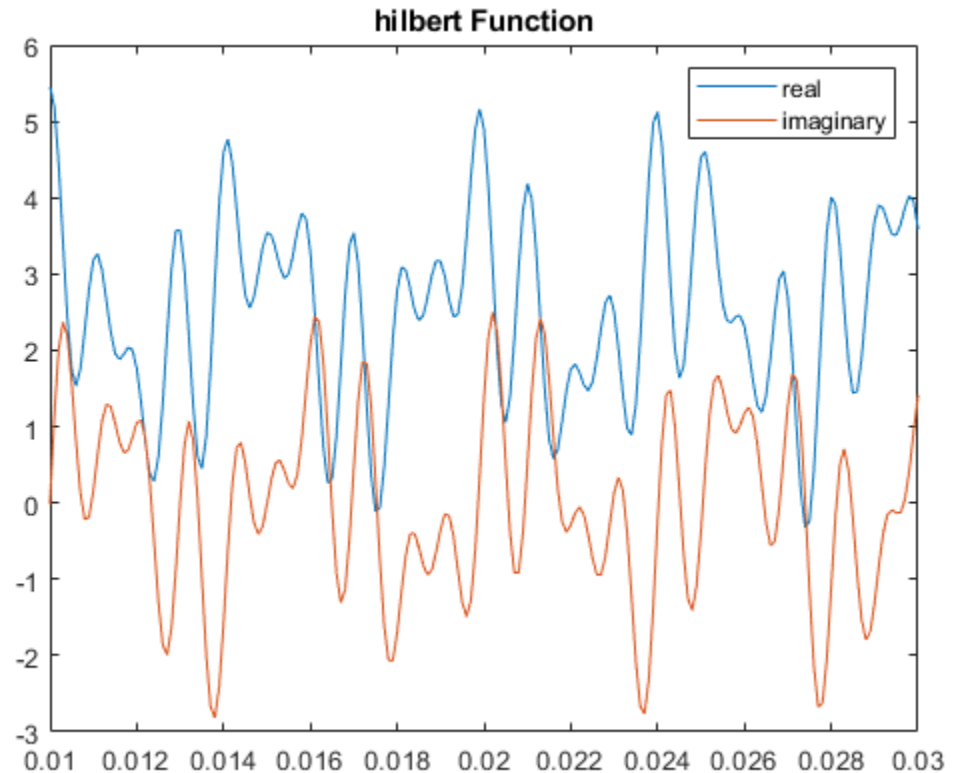
Although formally  $a(t)$  and  $\varphi(t)$  can be defined for any  $x(t)$ , **they have a clear physical meaning only if  $x(t)$  is a narrow-band oscillatory signal**: in that case, the  $a(t)$  coincides with the envelope of  $x(t)$  and the **instantaneous frequency,  $\omega(t)=d\varphi/dt$** , coincides with the dominant frequency in the power spectrum.



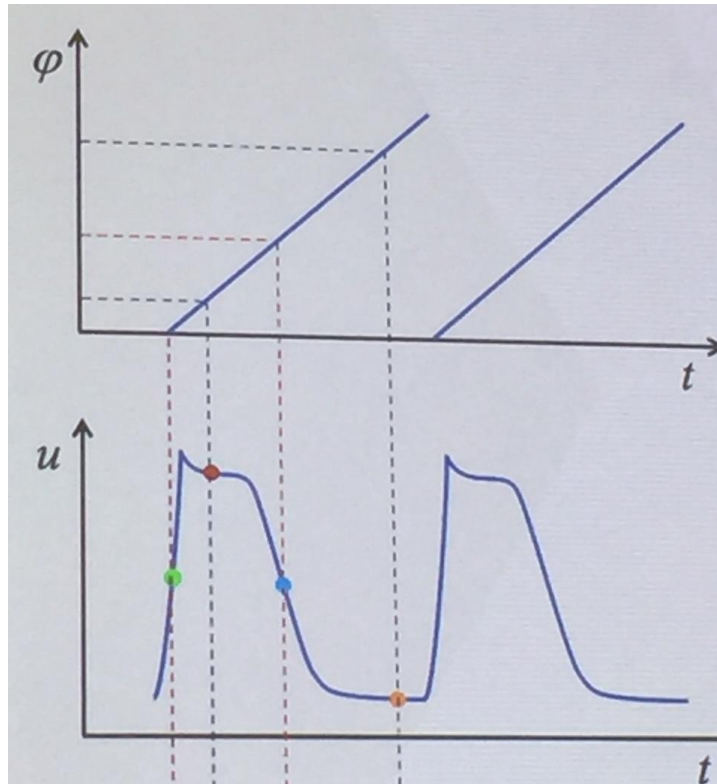
# Hilbert with matlab

```
x = 2.5 + cos(2*pi*203*t) + sin(2*pi*721*t) + cos(2*pi*1001*t);  
y = hilbert(x);  
plot(t,real(y),t,imag(y))  
xlim([0.01 0.03])  
legend('real','imaginary')  
title('hilbert Function')
```

The sampling rate must be chosen in order to have at least 20 points per characteristic period of oscillation.



# Another way to define a “phase”



Read more:

B. Kralemann, L. Cimponeanu, M. Rosenblum, A. Pikovsky, R. Mrowka, *Phase dynamics of coupled oscillators reconstructed from data*, Phys. Rev. E 77, 066205 (2008).

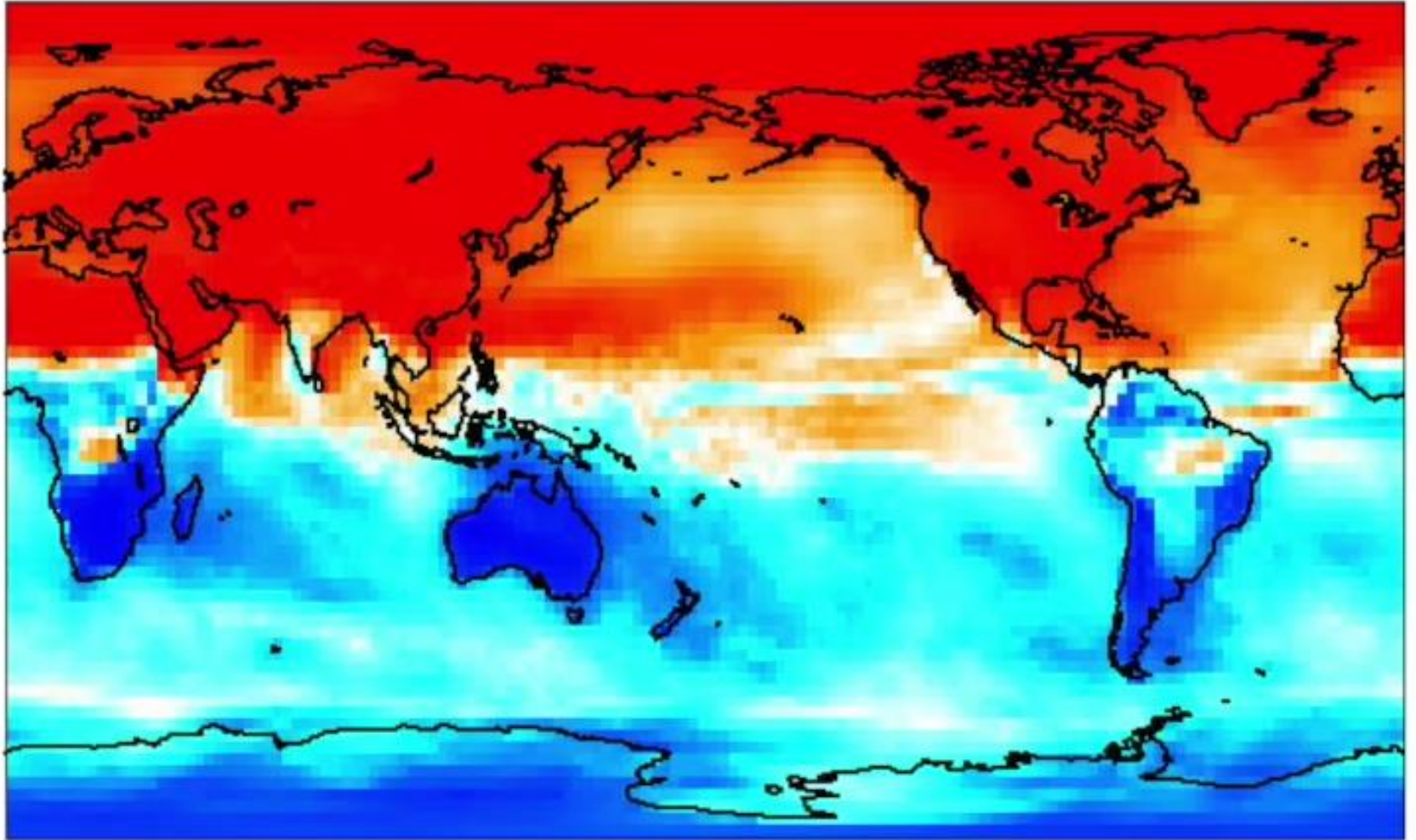
M. Chavez, M. Besserve, C. Adam, et al., *Towards a proper estimation of phase synchronization from time series*, J. Neurosci. Methods 154, 149-160 (2006).

# Application to climate data

- Can we use the Hilbert amplitude, phase, frequency, to :
  - Identify and quantify regional climate change?
  - Investigate synchronization in climate data?
- Problem: climate time series are not narrow-band.
- Usual solution (e.g. brain signals): isolate a narrow frequency band.
- However, the Hilbert transform applied to Surface Air Temperature time series yields meaningful insights.

# Cosine of Hilbert phase

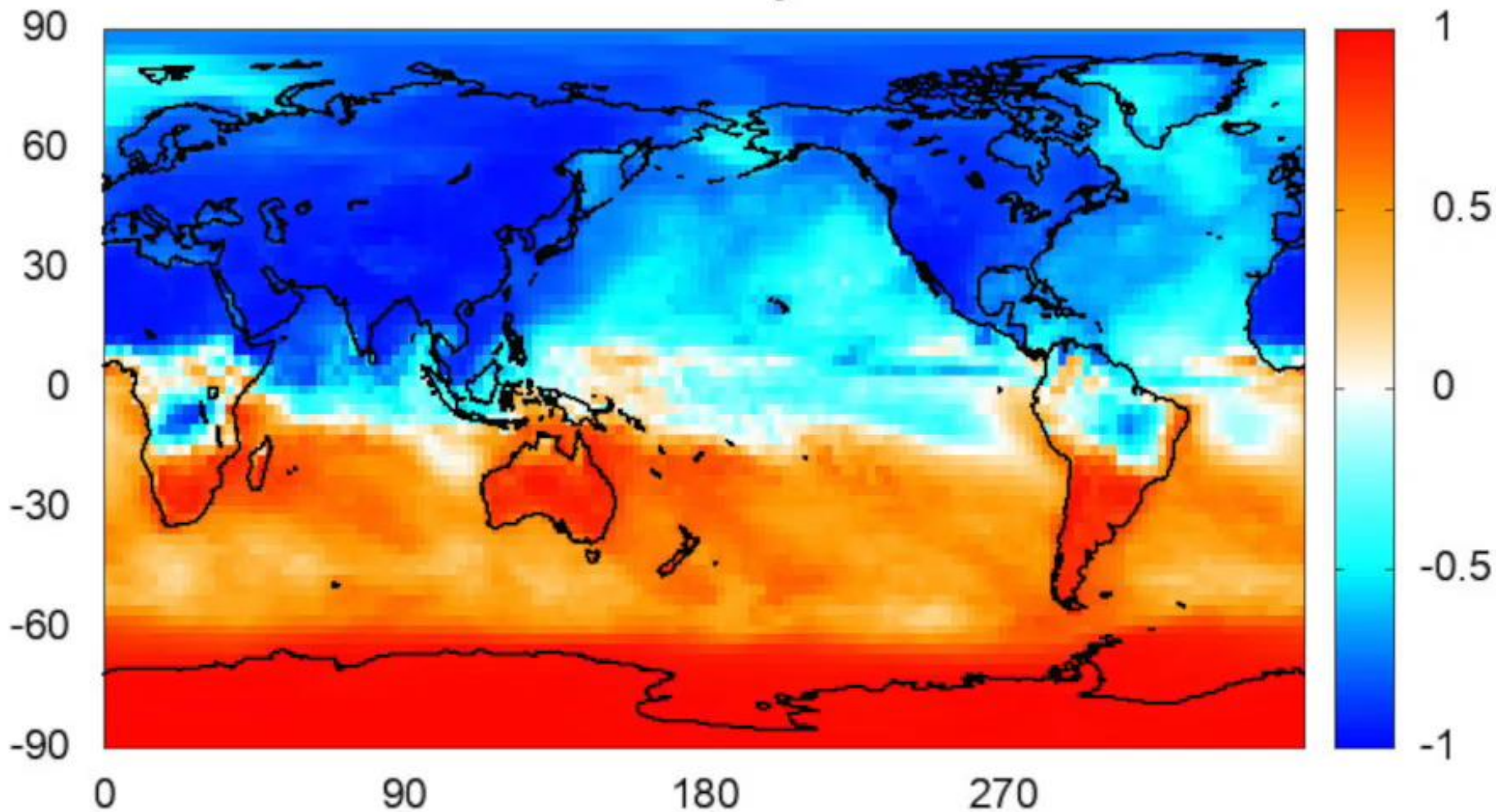
1 July



# How the seasons evolve?

## Temporal evolution of the cosine of the Hilbert phase

1 January





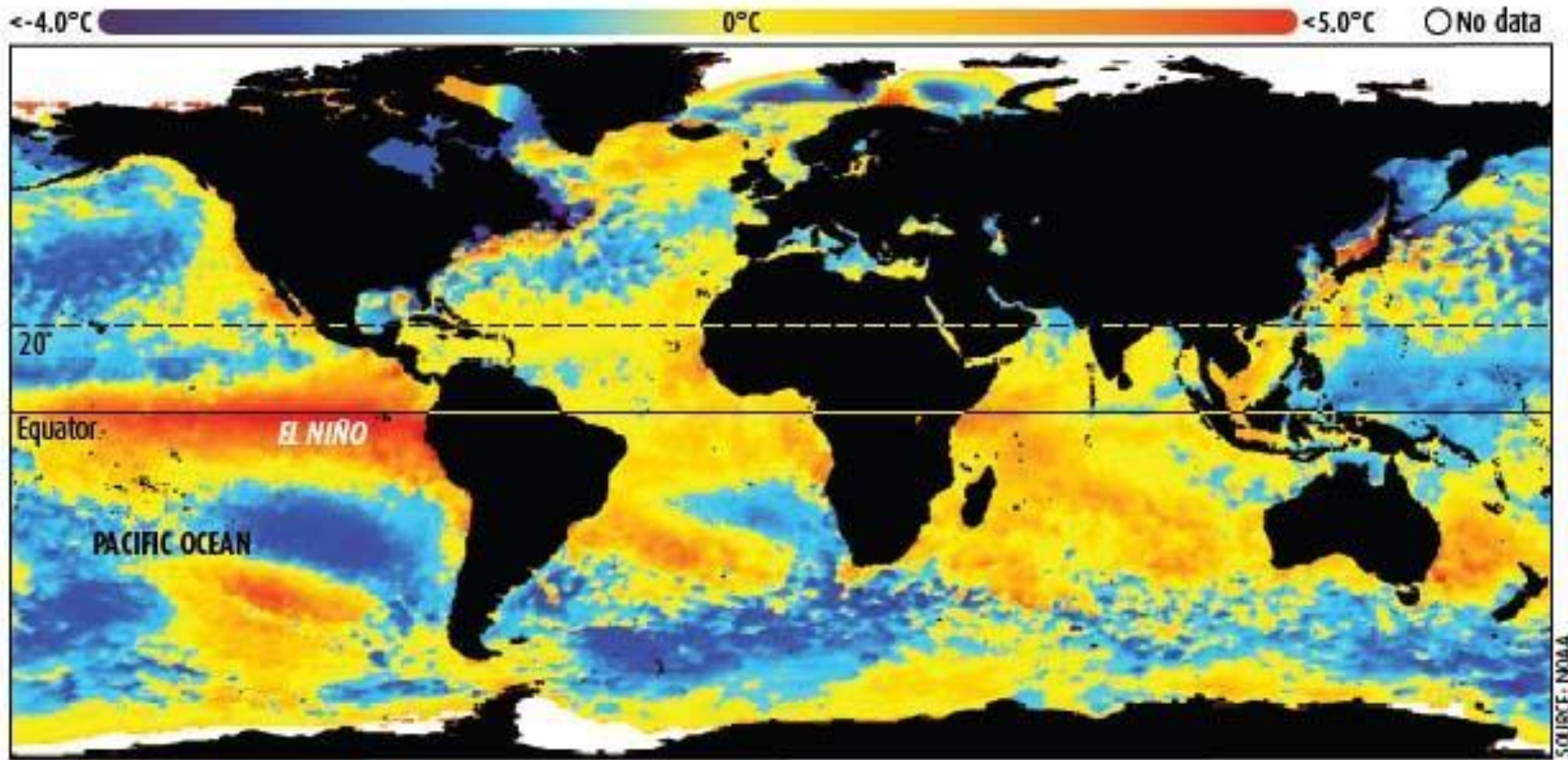
# El Niño/La Niña-Southern Oscillation (ENSO)

Is the most important climate phenomena on the planet

- Occurs across the tropical Pacific Ocean with  $\approx$  3-6 years periodicity.
- Variations in the surface **temperature** of the tropical eastern Pacific Ocean (warming: El Niño, cooling: La Niña)
- Variations in the air surface **pressure** in the tropical western Pacific (the Southern Oscillation).
- These two variations are coupled:
  - El Niño (ocean warming) -- high air surface pressure,
  - La Niña (ocean cooling) -- low air surface pressure.

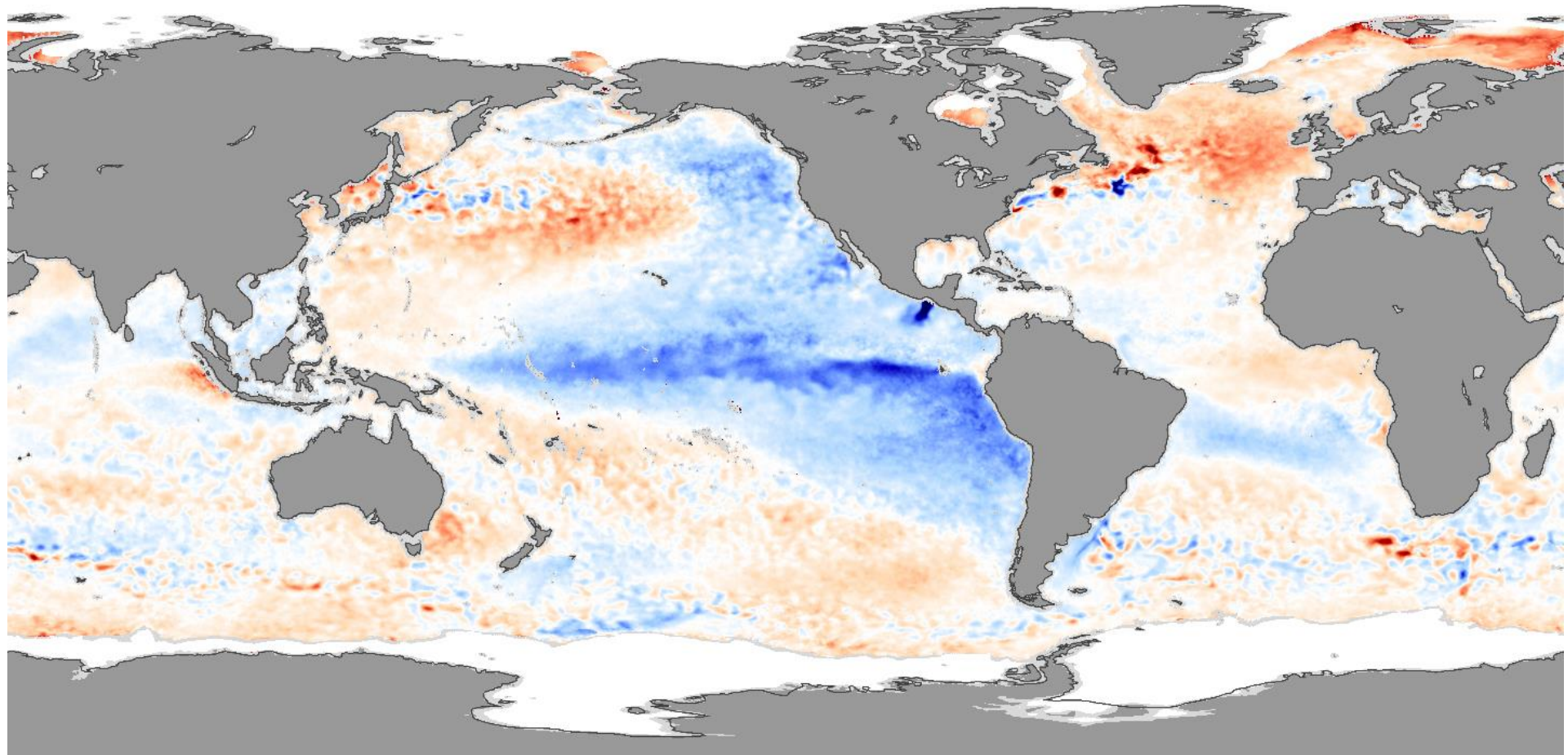
# January 1998: El Niño

## How ocean surface temperature differed from average



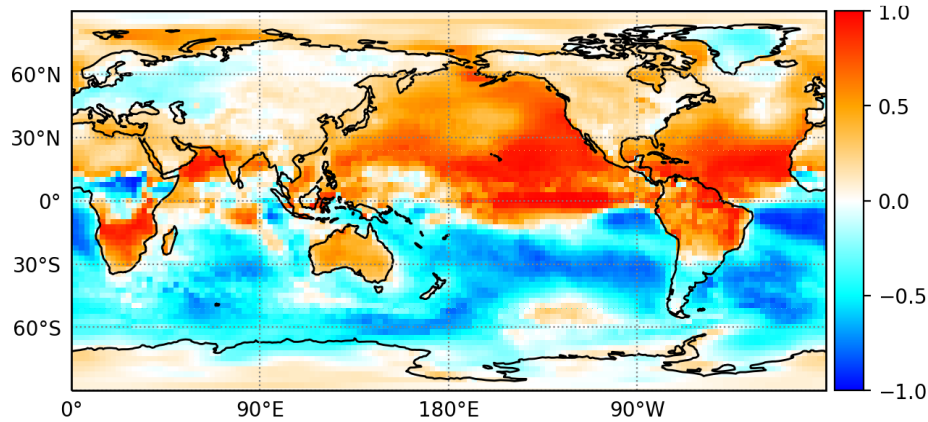
Satellite imagery of sea surface temperatures shows the strong El Niño that helped make 1998 one of the hottest years on record (Image: National Oceanic and Atmospheric Administration, US).

# Temperature anomalies during La Niña (November 2007)

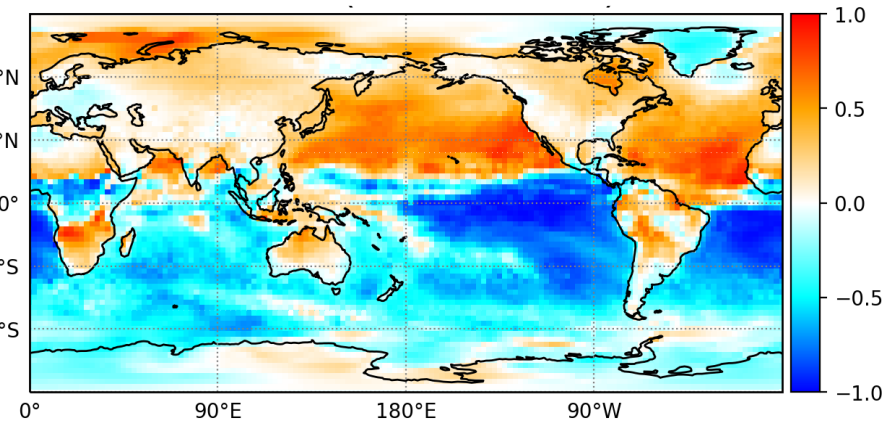




# Cosine of Hilbert phase during a El Niño period (October 2015)



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



# Changes in Hilbert amplitude and frequency detect inter-decadal variations in surface air temperature (SAT)

## The data:

- Spatial resolution  $2.5^{\circ} \times 2.5^{\circ} \Rightarrow 10226$  time series
- Daily resolution 1979 – 2016  $\Rightarrow 13700$  data points

## Where does the data come from?

- European Centre for Medium-Range Weather Forecasts (ECMWF, ERA-Interim).
- Freely available.

## “Features” extracted from each SAT time series

- Time averaged amplitude,  $\langle a \rangle$
- Time averaged frequency,  $\langle \omega \rangle$
- Standard deviations,  $\sigma_a$ ,  $\sigma_{\omega}$

## Relative decadal variations

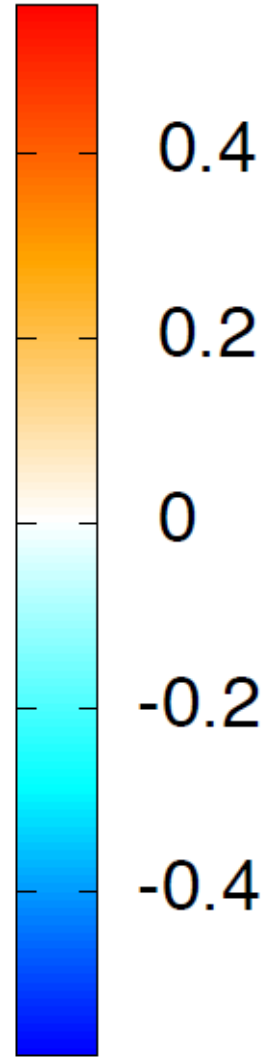
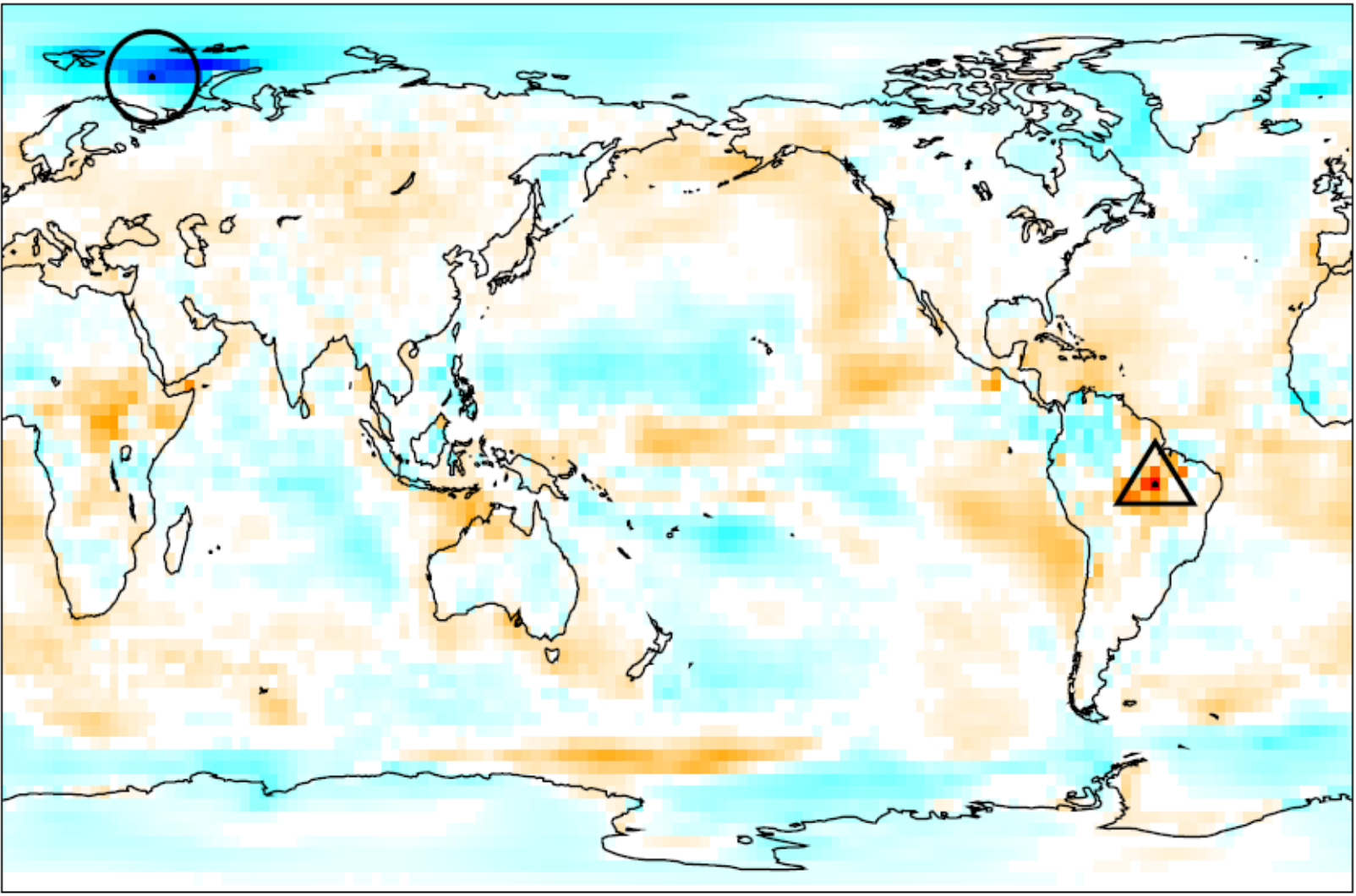
$$\Delta a = \langle a \rangle_{2016-2007} - \langle a \rangle_{1988-1979}$$
$$\frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

Relative variation is considered significant if:

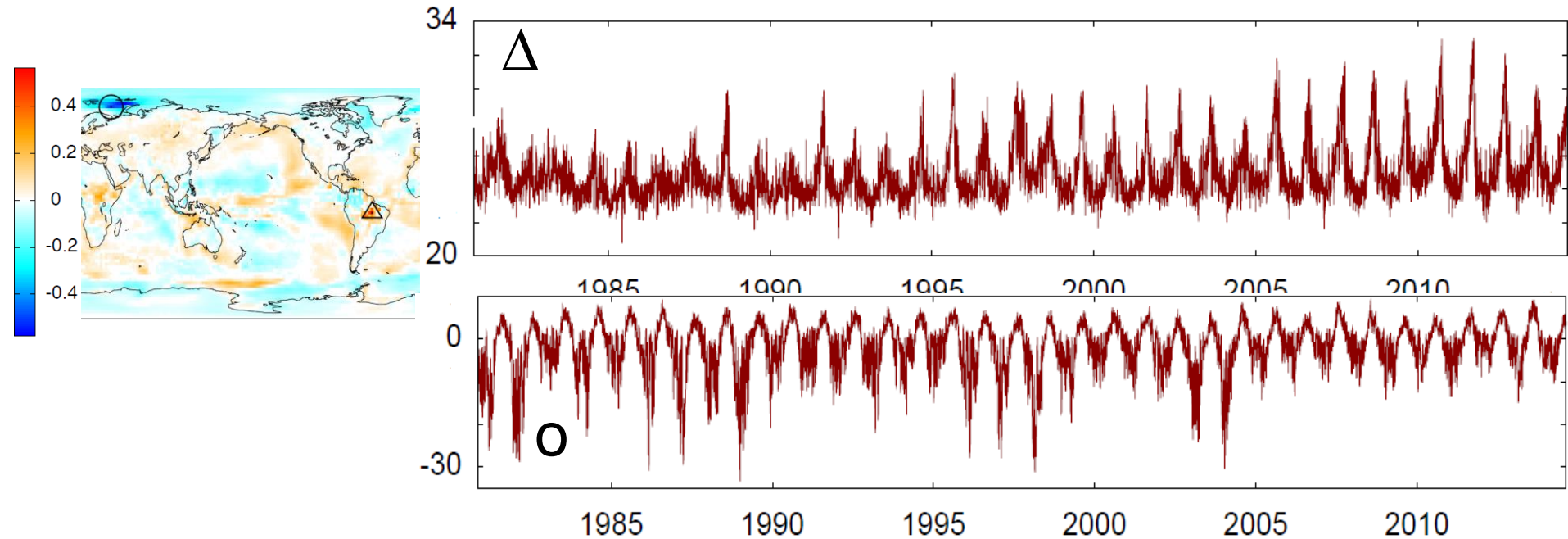
$$\frac{\Delta a}{\langle a \rangle} \geq \langle \cdot \rangle_s + 2\sigma_s \quad \text{or} \quad \frac{\Delta a}{\langle a \rangle} \leq \langle \cdot \rangle_s - 2\sigma_s$$

100 “block” surrogates

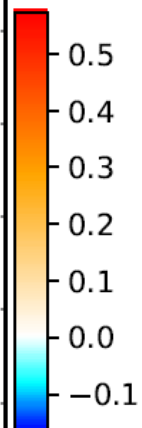
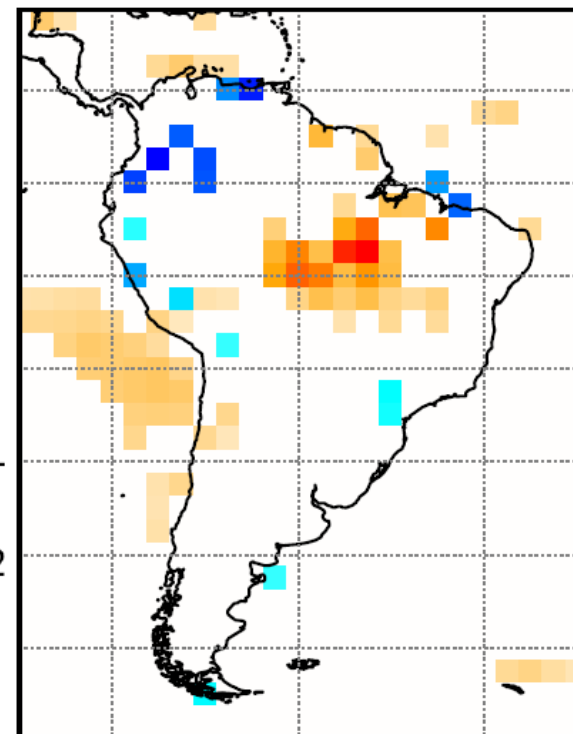
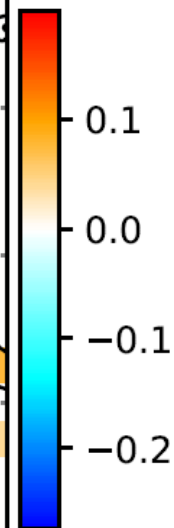
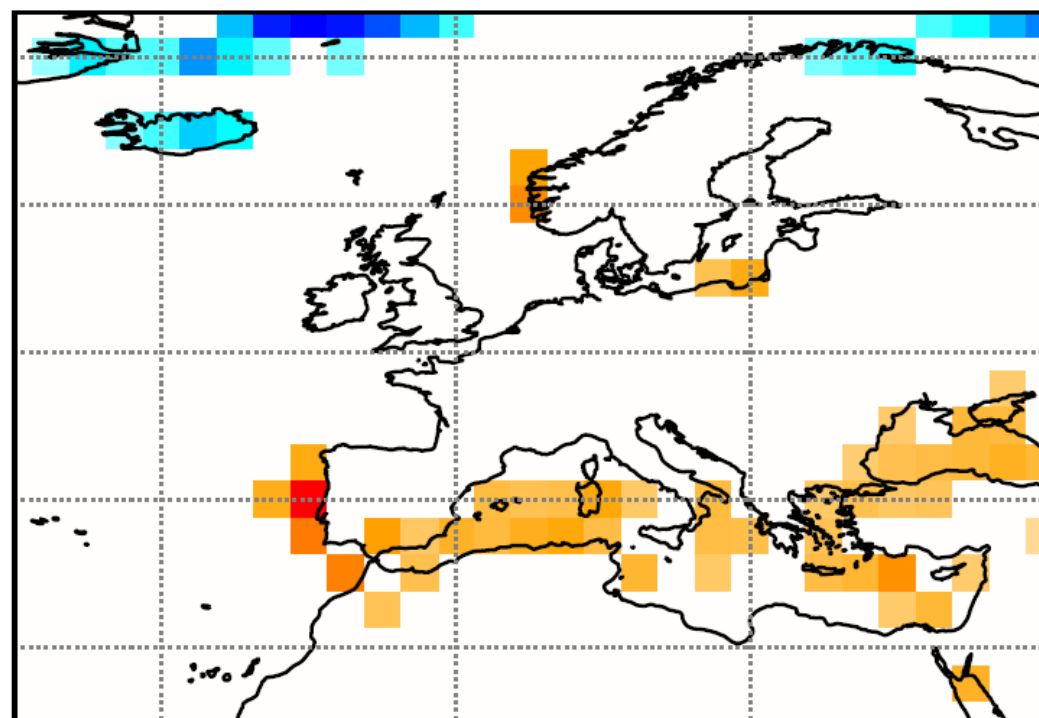
[D. A. Zappala, M. Barreiro and C. Masoller, “Quantifying changes in spatial patterns of surface air temperature dynamics over several decades”, Earth Syst. Dynam. \*\*9\*\*, 383 \(2018\)](#)



# Relative variation of average Hilbert amplitude uncovers regions where the amplitude of the seasonal cycle increased or decreased

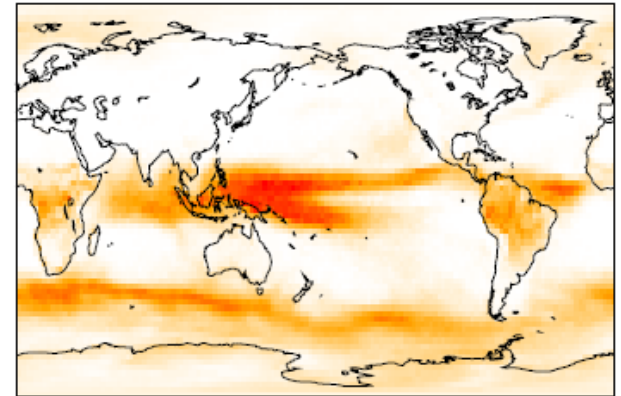


- **Decrease of precipitation:** the solar radiation that is not used for evaporation is used to heat the ground.
- **Melting of sea ice:** during winter the air temperature is mitigated by the sea and tends to be more moderated.

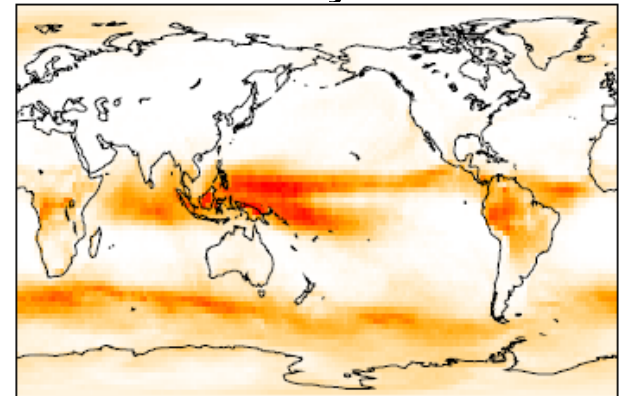


# Relative change of time-averaged Hilbert frequency consistent with a **north shift and enlargement of the intertropical convergence zone (ITCZ)**

First ten years



Last ten years



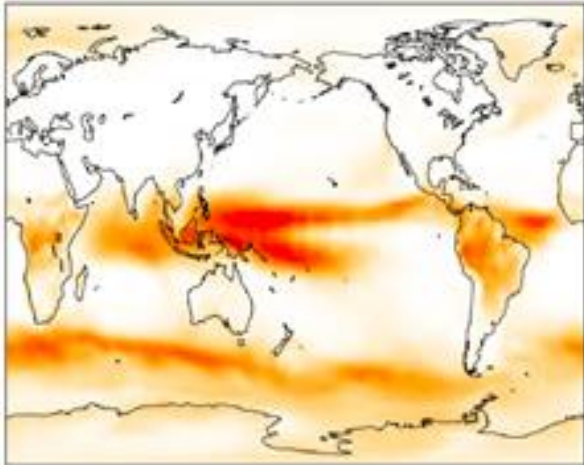


# Influence of pre-processing SAT time series by temporal averaging in a moving window: fast variability removed

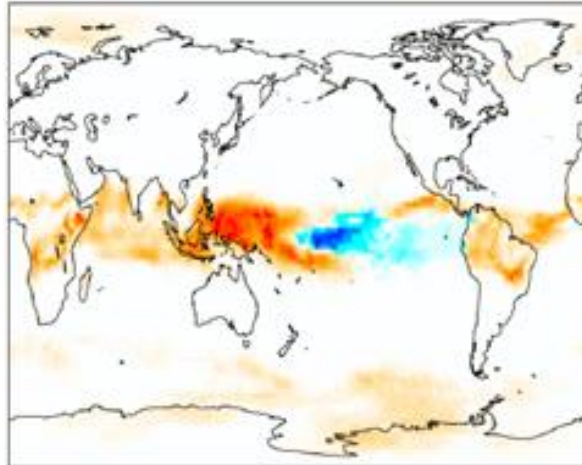
SAT  $\rightarrow$  average in a time window  $\rightarrow$  Hilbert

The colorcode shows the mean frequency (red fast, blue slow)

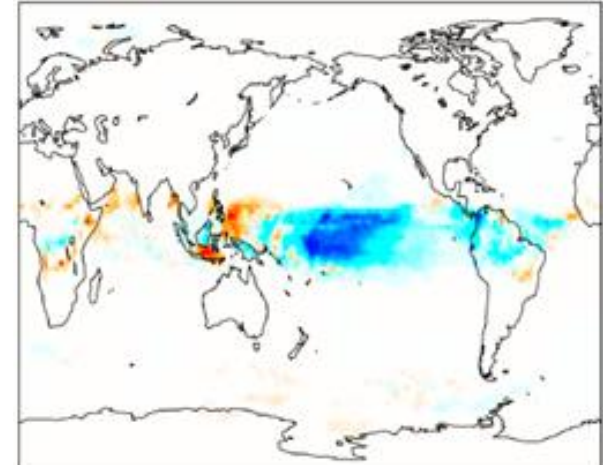
No filter



1 month



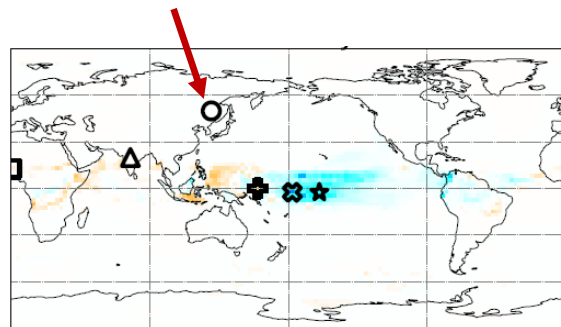
3 months



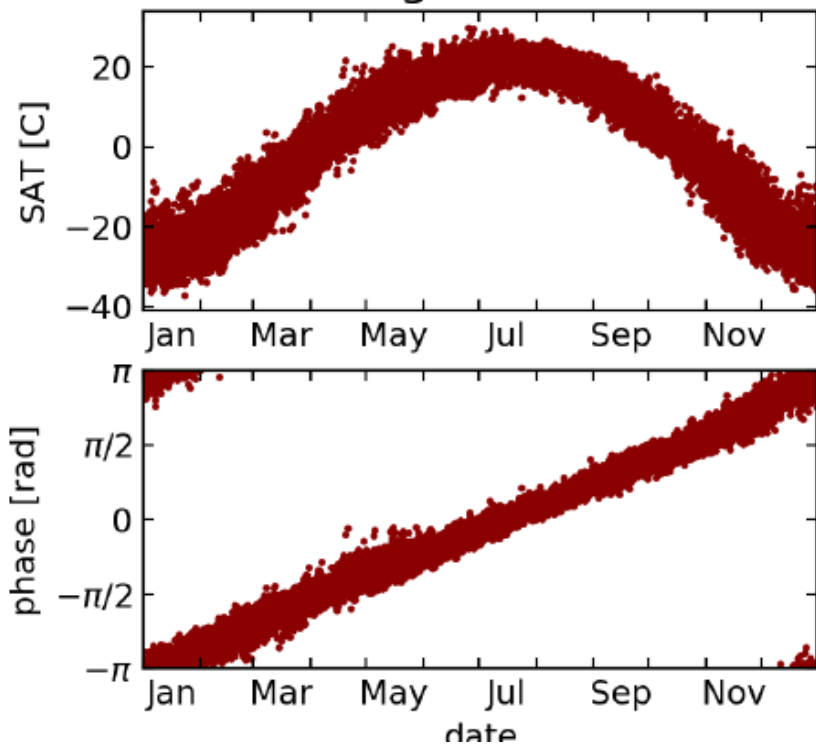
How does the phase dynamics depend on the length of the averaging window?



# Phase-date relation in a regular site

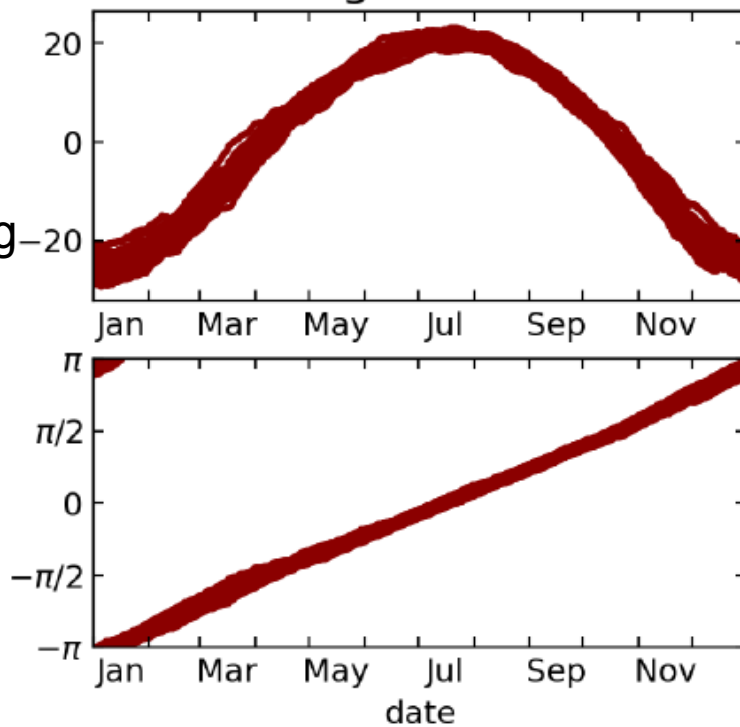


Regular site

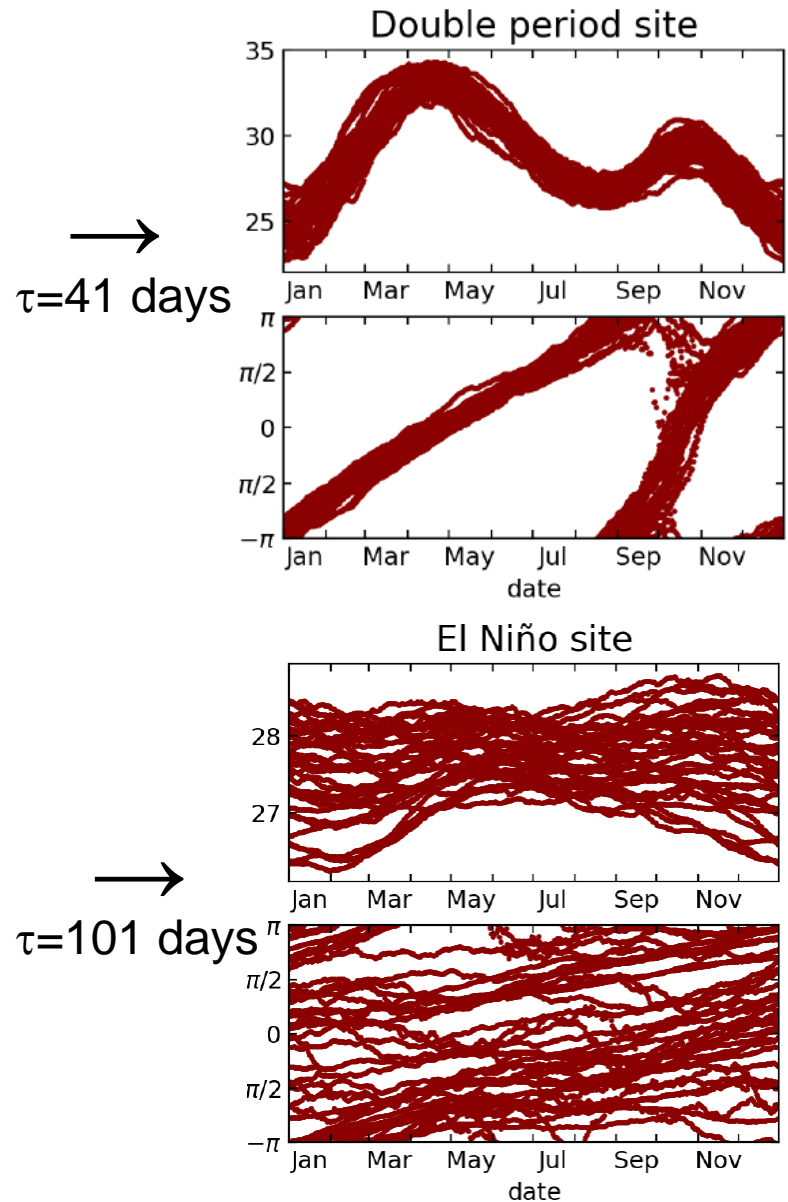
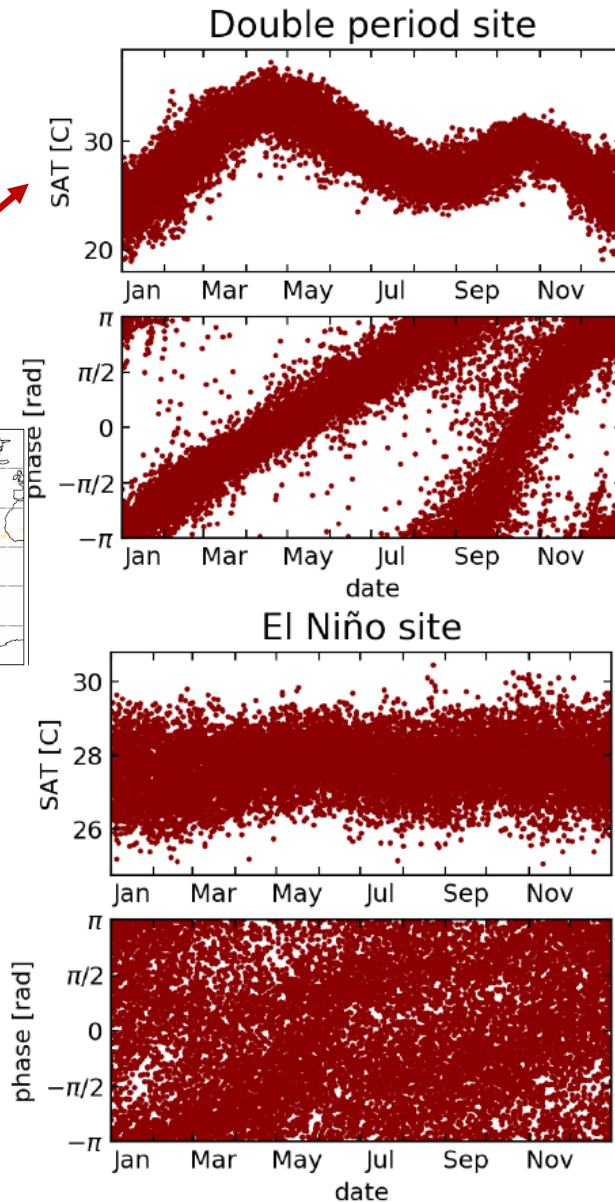
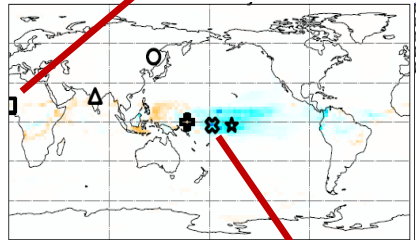


→  
Pre-filtering  
SAT time series in a  
moving window of  
 $\tau=41$  days

Regular site

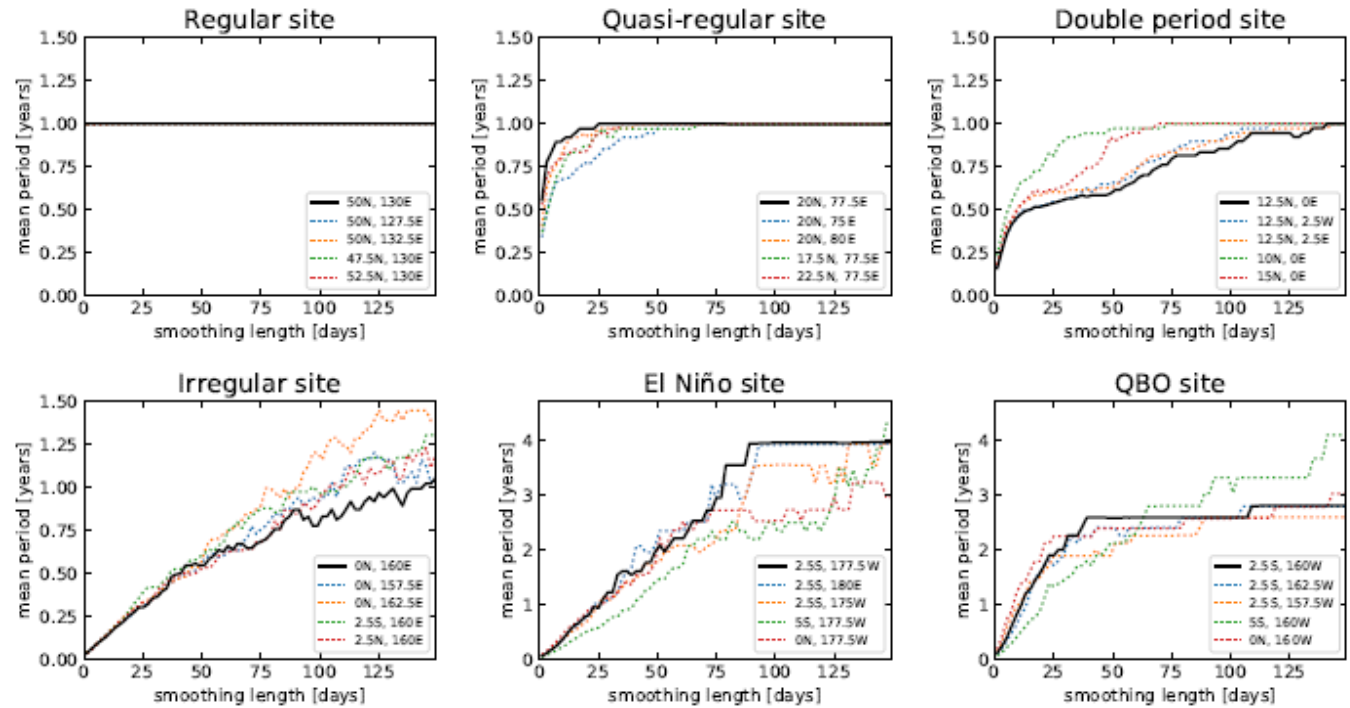


# In other sites: with appropriate pre-processing we uncover hidden regularity



# Variation of mean rotation period with the smoothing length

$$\bar{T}_j = 2\pi \frac{\Delta t}{\Delta \varphi_j}$$



- Regular, quasi-regular and double period sites: a plateau at 1 year is found when increasing the averaging window (sooner or latter)
- Irregular site: no plateau.
- In El Niño and QBO sites: plateau at 4 and  $\approx 2.5$  years respectively.

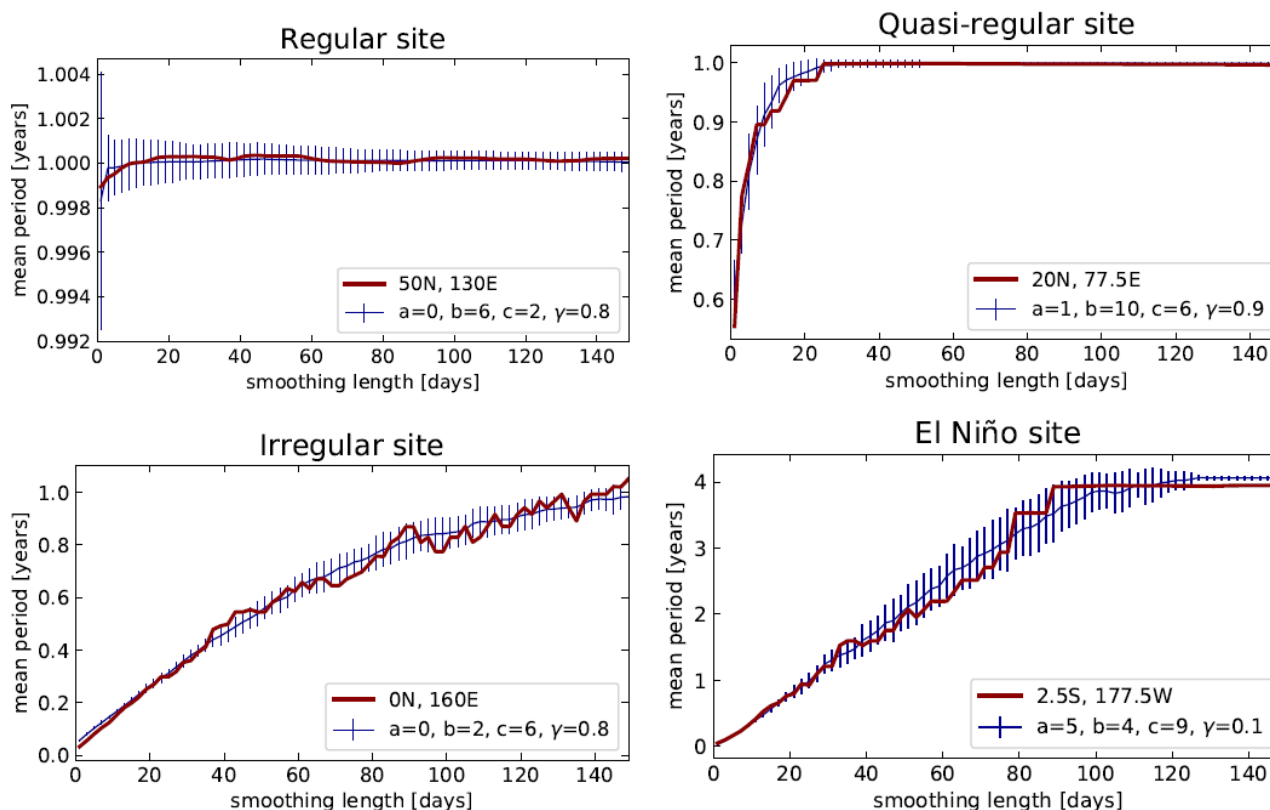
D. A. Zappala, M. Barreiro, and C. Masoller, “*Uncovering underlying regularities in climatological data through Hilbert phase analysis*”, *Chaos* 29, 051101 (2019)

**Can we understand the variation of  $T$  with  $\tau$  ?**

# Simple model: two sinusoidal oscillations (the annual cycle and a slower oscillation) and autoregressive noise

$$S(t) = \frac{a\sqrt{2} \cos(\omega_{sl}t) + b\sqrt{2} \cos(\omega_y t) + c\varepsilon(t)}{\sqrt{a^2 + b^2 + c^2}}$$

We generate synthetic data. We know the underlying equations and parameters and so can check the validity of our method.



# And many more TS analysis methods

- Wavelets
- Detrended fluctuation analysis
- Sample entropy, approximate entropy
- Multifractality
- Topological data analysis
- Etc. etc.

**There are MANY methods that return “features” that characterize time series.**

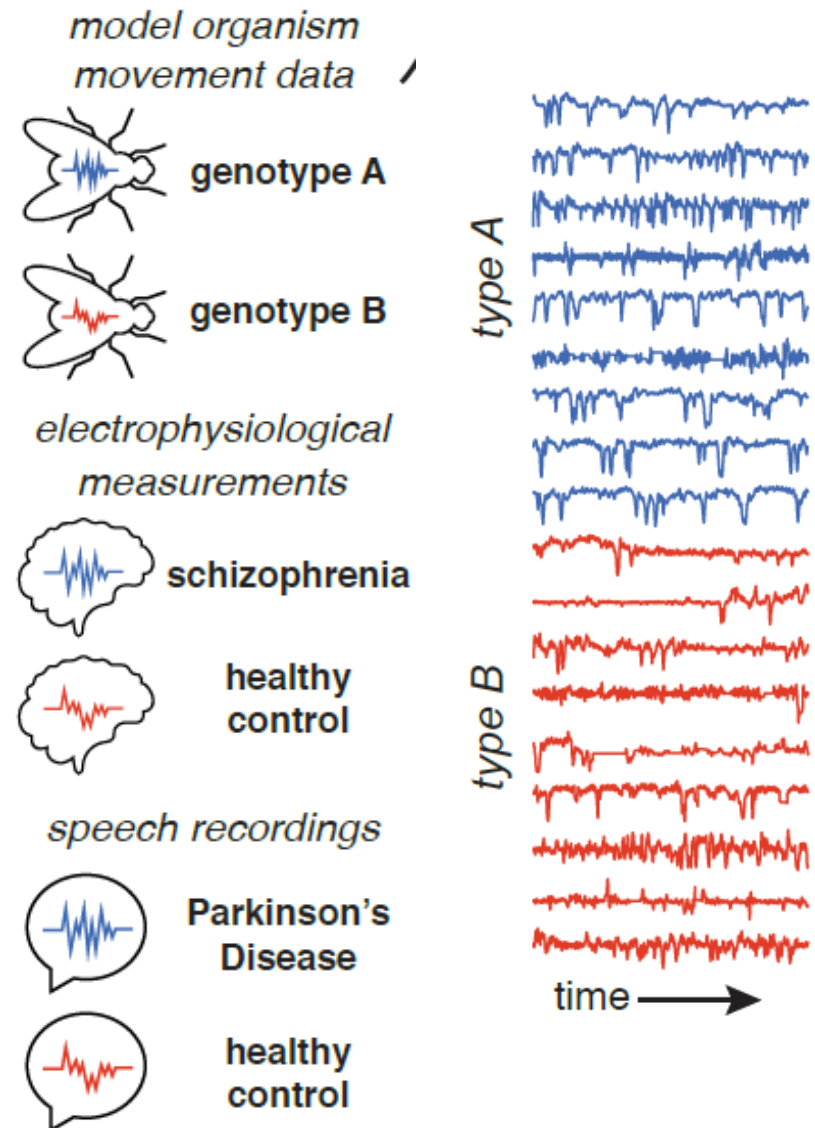
- How to identify similar time-series?
- How to compare different methods?
- How to identify similar methods?

[B. D. Fulcher, N. S. Jones: \*Automatic time-series phenotyping using massive feature extraction\*. Cell Systems 5, 527 \(2017\).](#)

# Systems producing time-series with different features (A/B)

Examples presented in this course:

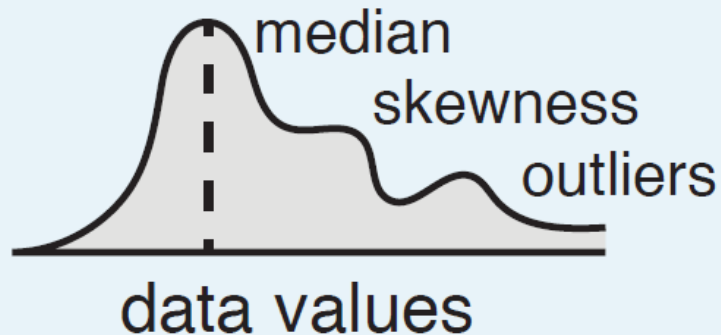
- Laser intensity (noise or chaos)
- EEG signals (eyes open or closed)
- ECG signals (congestive heart failure or healthy control)
- Etc.



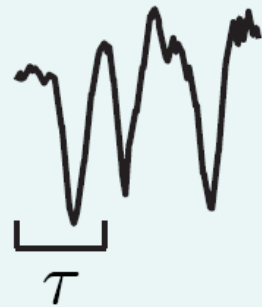


# Features extracted from each time series

## data distribution



## correlation properties



autocorrelation  
automutual information  
power spectral properties  
time-series entropy  
fluctuation analysis

## model fitting



linear autoregressive &  
nonlinear models  
model parameters  
goodness of fit

## others

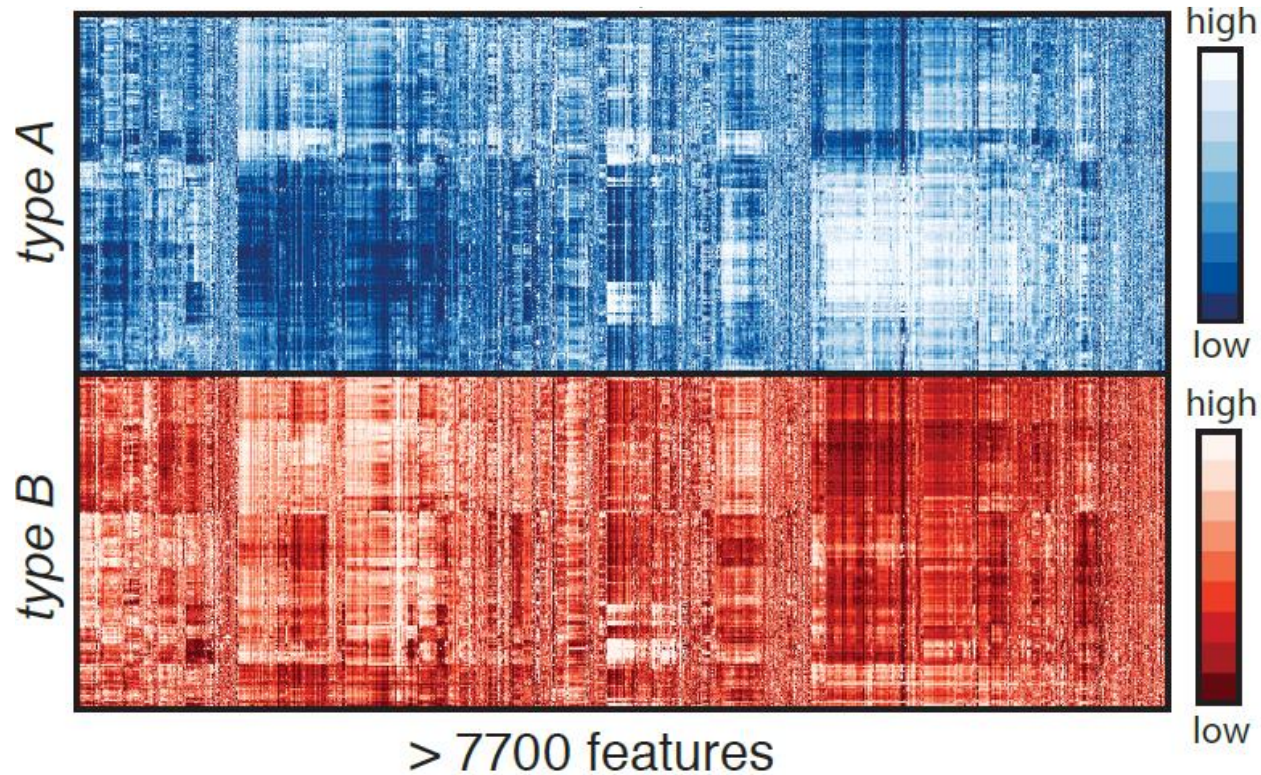
stationarity  
embedding dimension  
network properties

HCTSA (highly comparative time-series analysis) extracts, from each time series, over **7700 features**

Matlab code: [www.github.com/benfulcher/hctsa](http://www.github.com/benfulcher/hctsa)

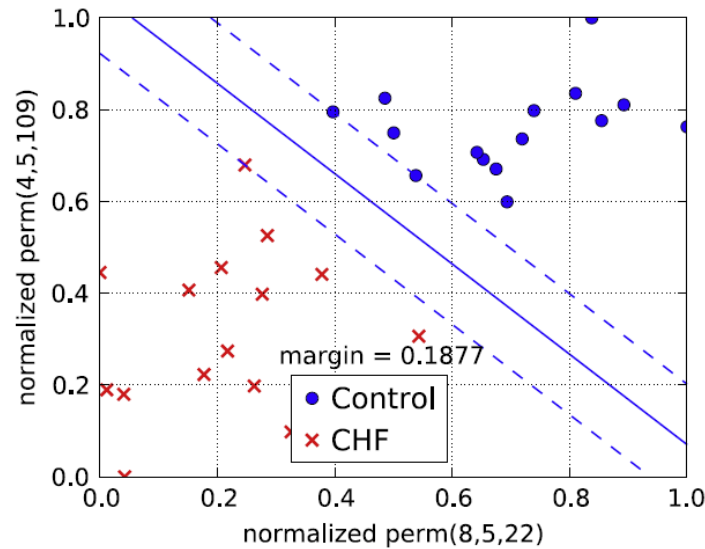
**What to do with more than  
7700 features?**

# Feature matrix



Each row represents a time series and each column represents a feature that encapsulates some property of that time series. Features normalized such that max. value =1, min. value 0.

# Example of “feature normalization”



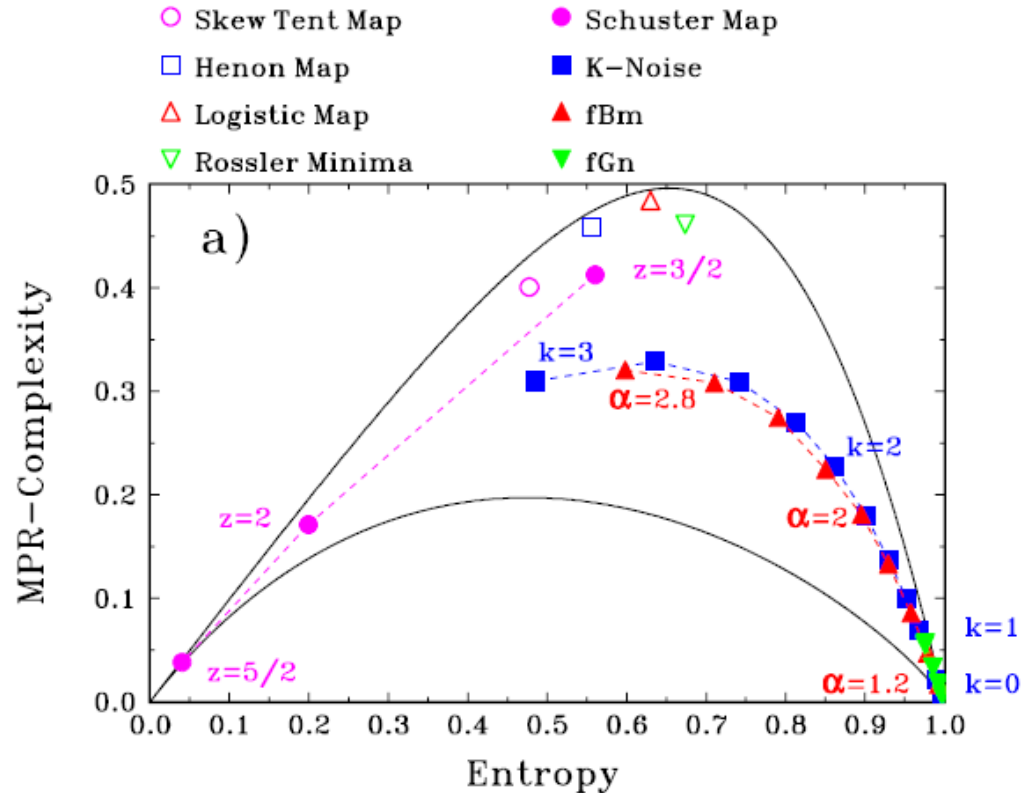
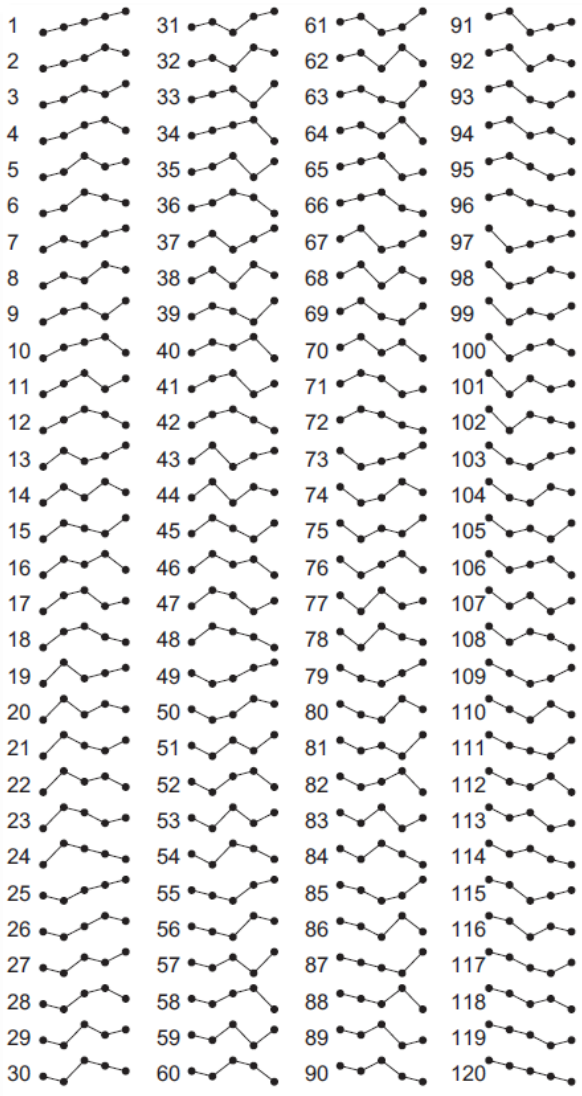
- Distribution of CHF cases (red crosses) and healthy subjects (blue filled circles) in two-dimensional feature space. The separating (solid) lines are computed using a linear support vector machine (SVM).
- The probabilities are normalized with respect to the smallest and the largest value occurring in the data set.

How to extract information from the feature matrix?

Let's take a step back.

Assuming we have computed 120 ordinal probabilities.

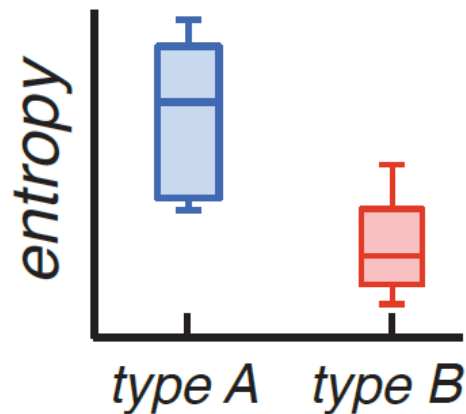
What we do next?



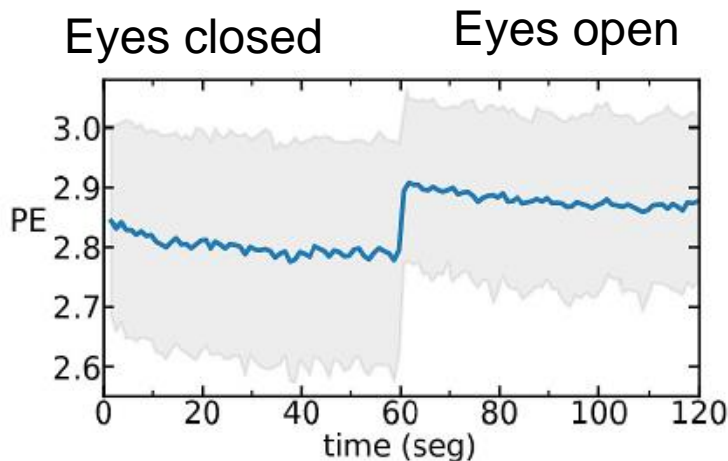
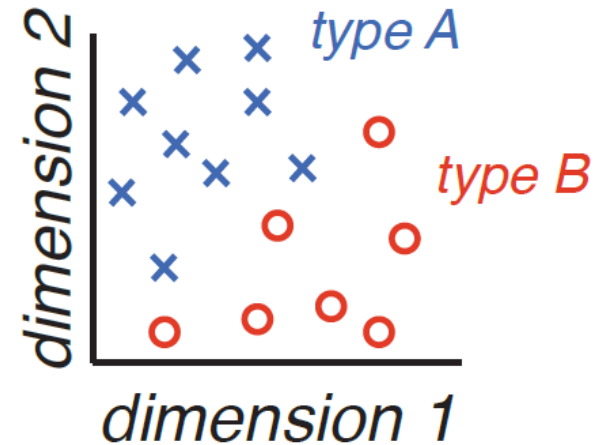
O. A. Rosso et al, Phys. Rev. Lett. 99, 154102 (2007)

# Extracting *meaningful insights* from feature matrix for classification (diagnose disease, gene function, etc.)

## Discriminative features



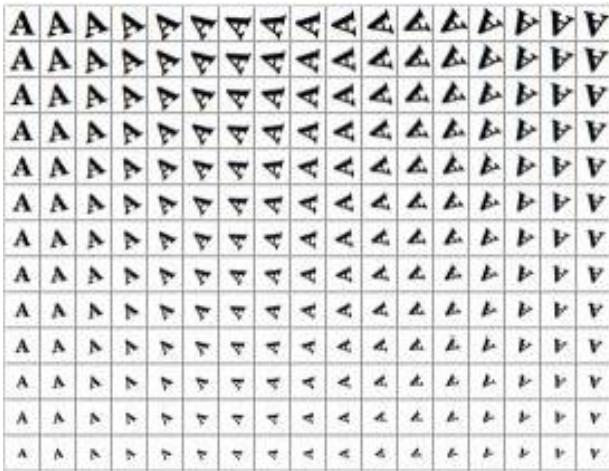
## Dimensionality reduction



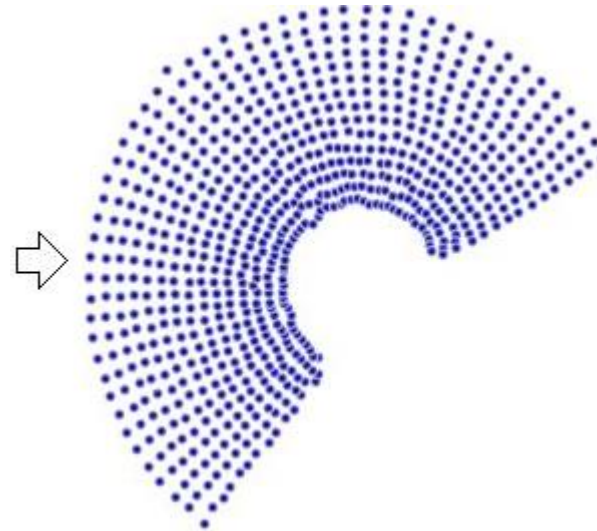


# Nonlinear dimensionality reduction (NLDR)

- Many methods for reducing high-dimensional data to a small number of dimensions.
- Example:



Each image of the letter 'A' has 32x32 pixels = vector of 1024 pixel values

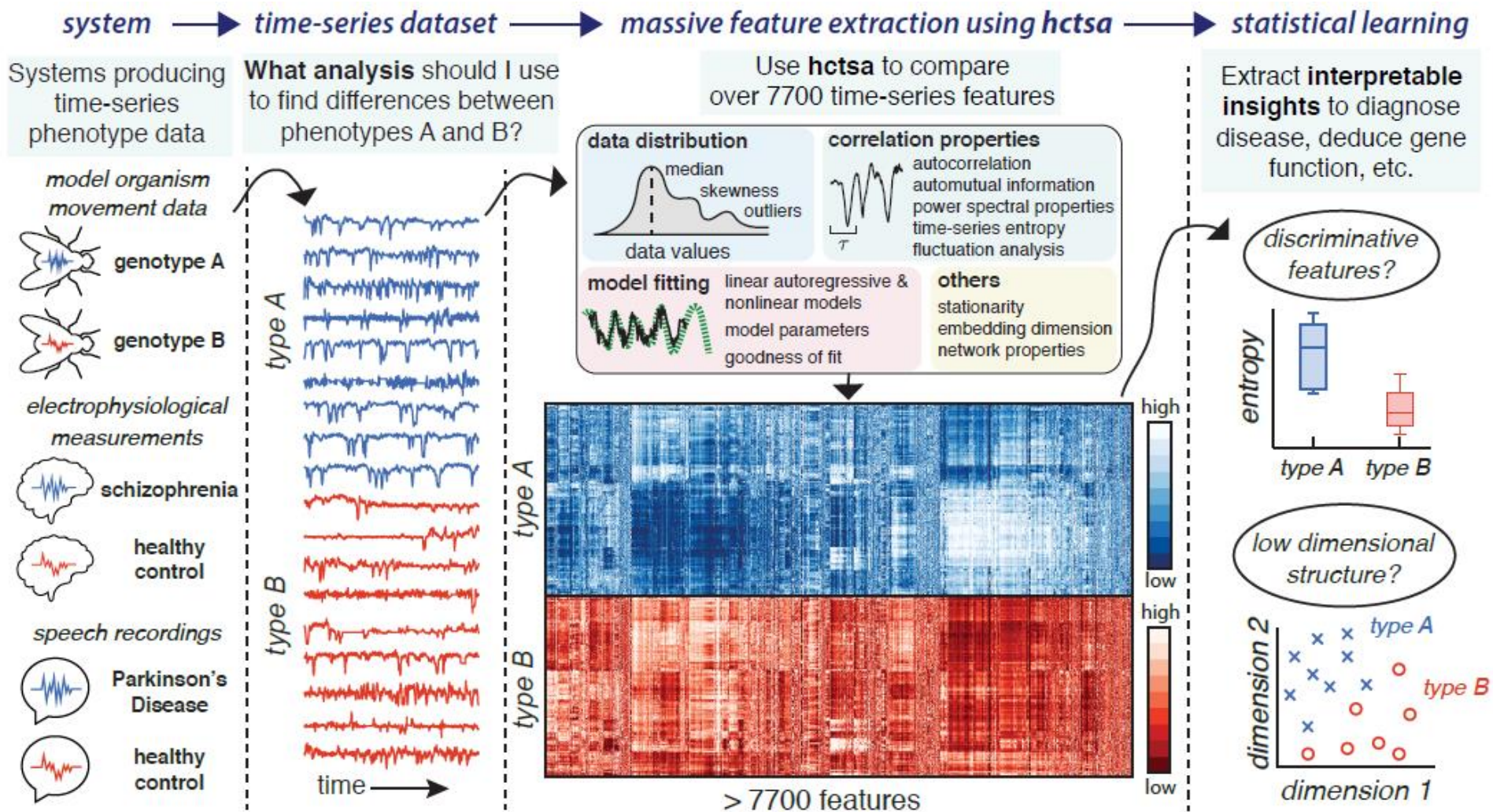


NLDR reduces the data into just two dimensions (rotation and scale)

Source:  
Wikipedia

- A popular method: *ISOMAP*
- A linear method: Principal Component Analysis (PCA)

# Summarizing: how to find time-series with statistically similar features



This can also be used to find **similar methods**



# Take home messages

- Symbolic analysis, network representation, spatiotemporal representation, etc., are useful tools for investigating complex signals.
- Different techniques provide *complementary* information.

“...nonlinear time-series analysis has been used to great advantage on thousands of real and synthetic data sets from a wide variety of systems ranging from roulette wheels to lasers to the human heart. Even in cases where the data do not meet the mathematical or algorithmic requirements, the results of nonlinear time-series analysis can be helpful in understanding, characterizing, and predicting dynamical systems...”

# References

- R. Hegger, H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (book and software [TISEAN](#)).
- [Bandt and Pompe, PRL 88, 174102 \(2002\)](#)
- [U. Parlitz et al., Computers in Biology and Medicine 42, 319 \(2012\)](#)
- [B. D. Fulcher and N. S. Jones, Cell Systems 5, 527 \(2017\).](#)
- [C. Masoller et al, NJP 17, 023068 \(2015\)](#)
- [A. Aragoneses et al, PRL 116, 033902 \(2016\)](#)
- [J. M. Aparicio-Reinoso et al PRE 94, 032218 \(2016\)](#)
- [Zappala, Barreiro and Masoller, Entropy 18, 408 \(2016\)](#)
- [Zappala, Barreiro and Masoller, Earth Syst. Dynam. 9, 383 \(2018\)](#)
- [D. A. Zappala, M. Barreiro, and C. Masoller, Chaos 29, 051101 \(2019\)](#)

<crisolina.masoller@upc.edu>

[http://www.fisica.edu.uy/~cris/teaching/curitiba\\_  
masoller\\_uni.pdf](http://www.fisica.edu.uy/~cris/teaching/curitiba_masoller_uni.pdf)