Nonlinear time series analysis
Bachelor's degree in Mathematics
Bachelor's degree in Statistics
Course 2020-2021

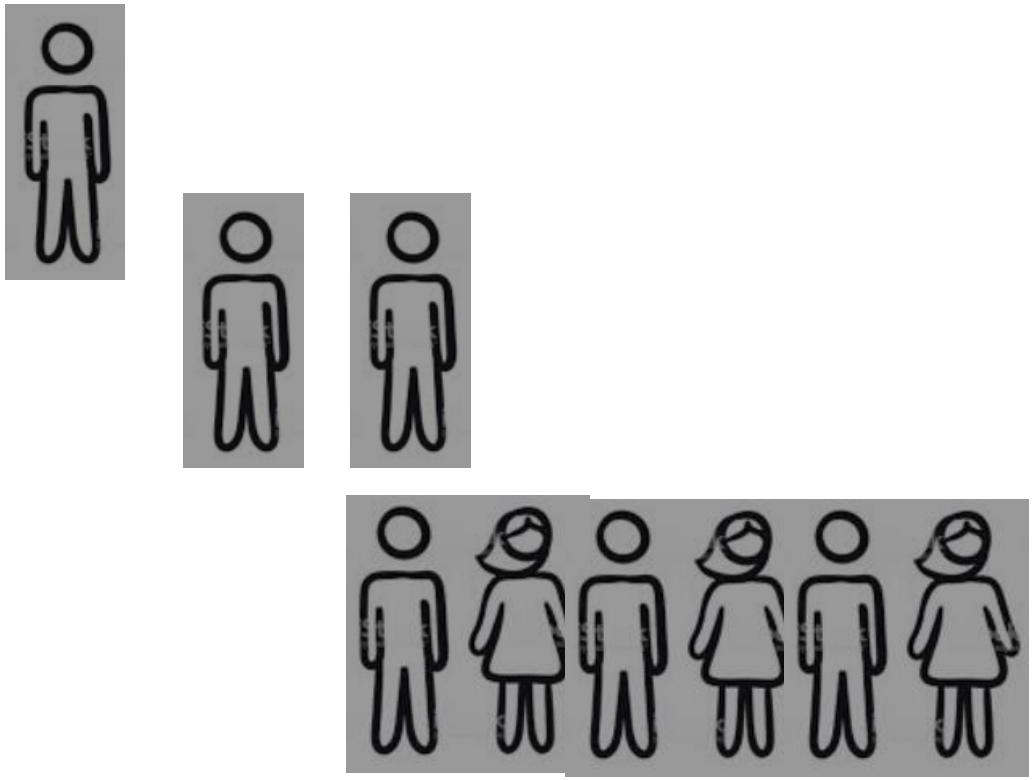
Univariate time series analysis

Prof. Cristina Masoller
Universitat Politècnica de Catalunya
cristina.masoller@upc.edu
www.fisica.edu.uy/~cris

50 ANYS
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
UPC
sense limits

Outline of the first block

- Introduction
- Univariate analysis
- Bivariate analysis
- Multivariate analysis



Methods of time series analysis

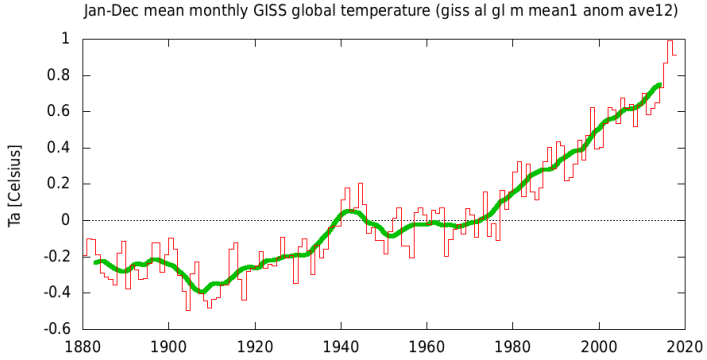
- Return maps
- Distribution of data values
- Autocorrelation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction, Lyapunov exponents, and fractal dimension
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series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

To begin with the analysis of a time series

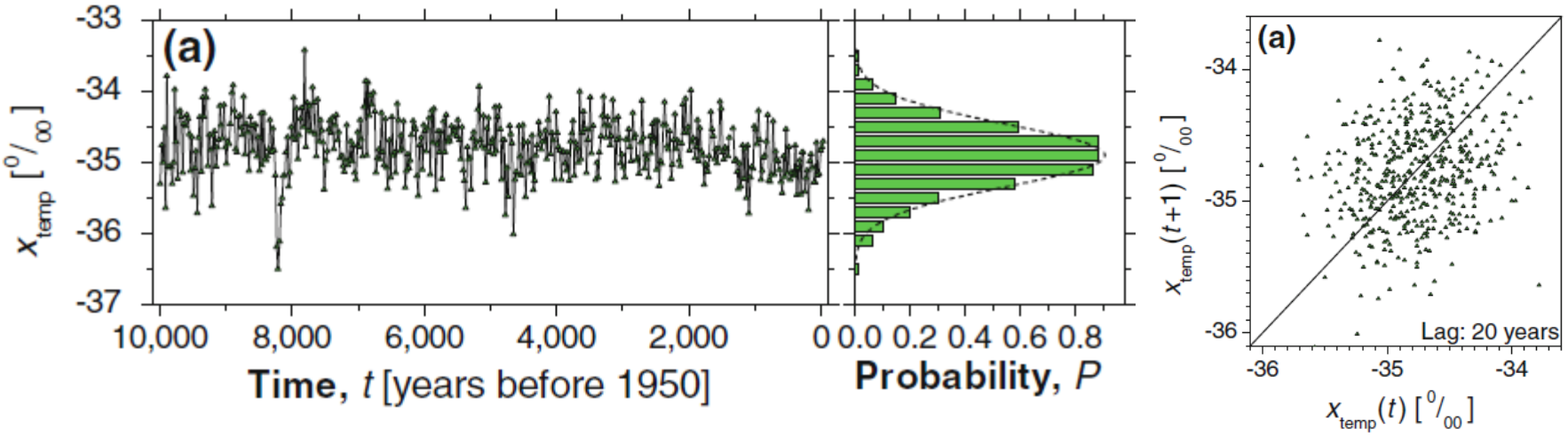
$X = \{x_1, x_2, \ldots x_N\}$

- First step: **Look at the data**.

- Examine simple properties:
  - Return map: plot of x_i vs. $x_{i+\tau}$
  - Distribution of data values (at least 10 data points per bin)
  - Auto-correlation
  - Fourier spectrum



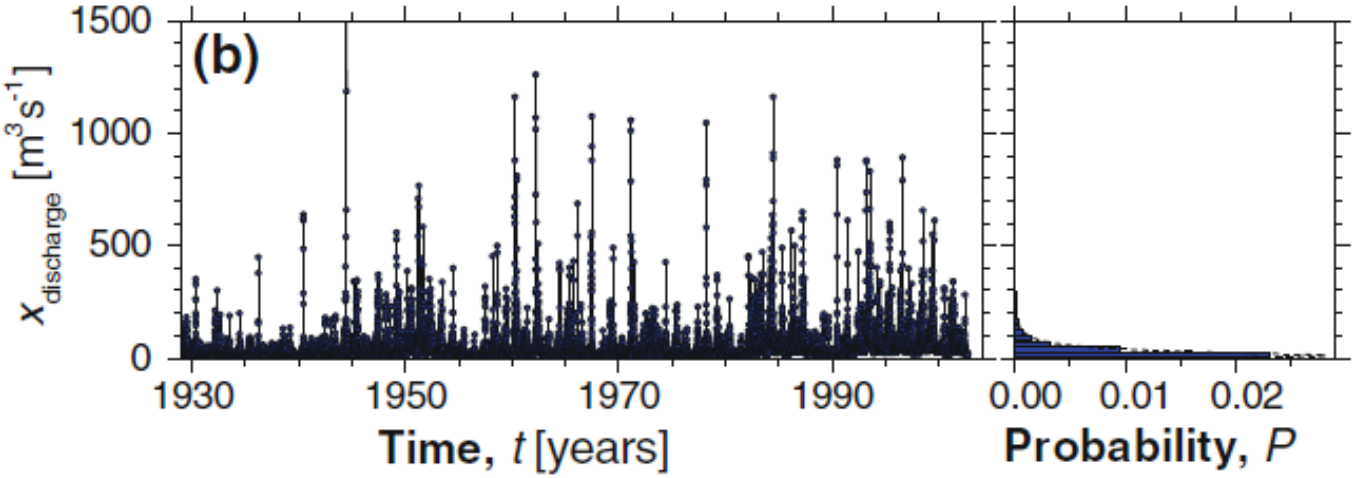
First example of a geophysical time series

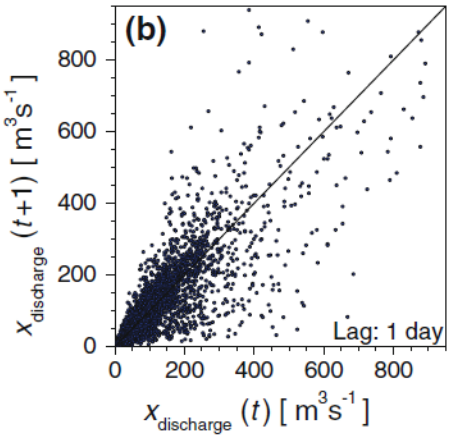


Bi-decadal oxygen isotope data set d18O (proxy for palaeotemperature) from Greenland Ice Sheet Project Two (GISP2) for the last 10,000 years with 500 values given at 20 year intervals.

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

Second example

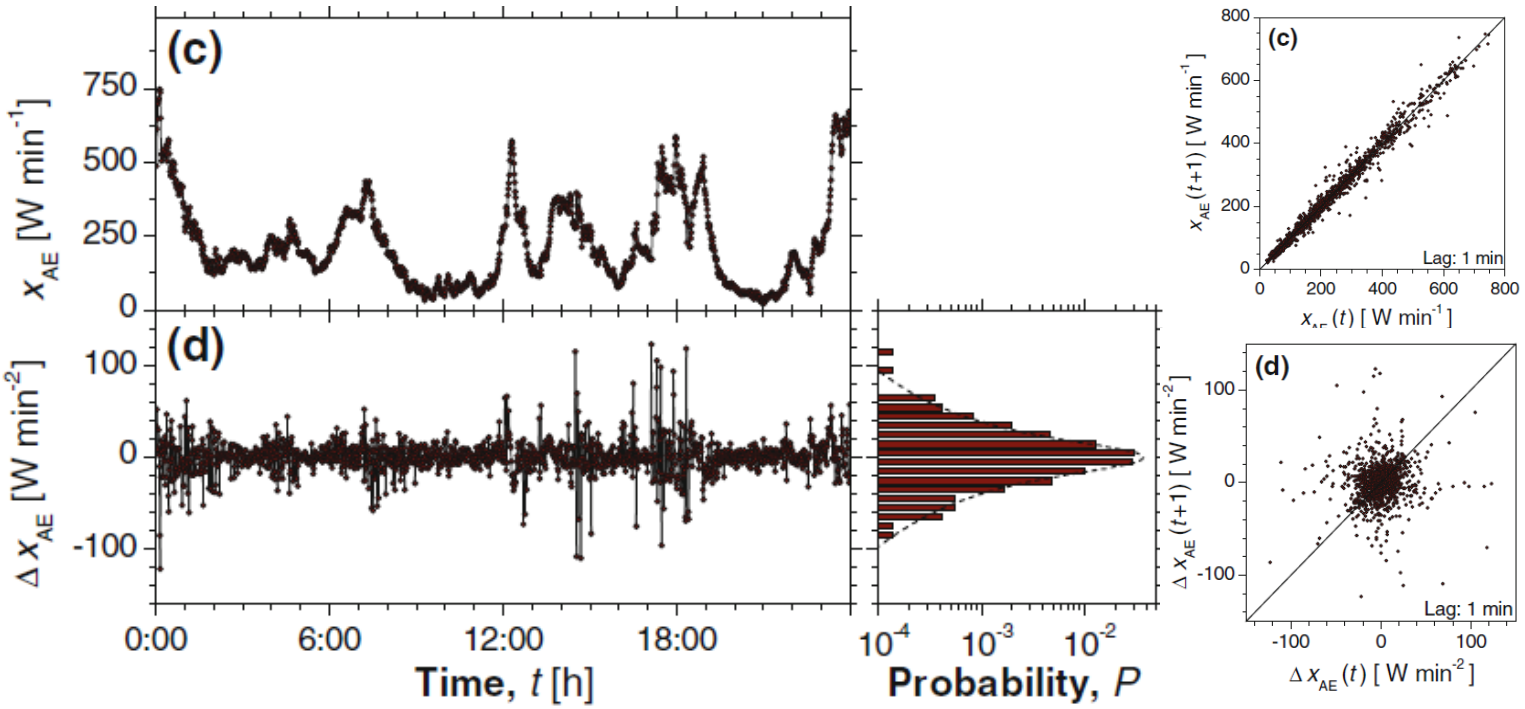




Discharge of the Elkhorn river (at Waterloo, Nebraska, USA) sampled daily for the period from 01 January 1929 to 30 December 2001.

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

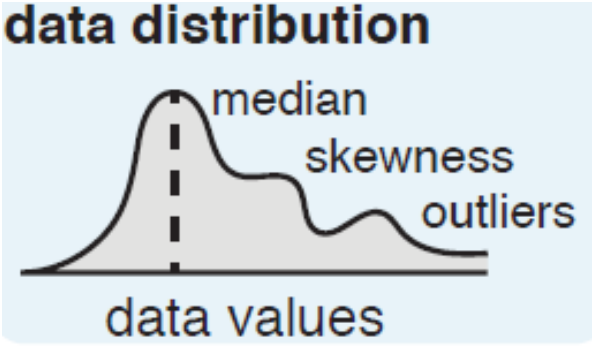
Third example



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

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

How to characterize the distribution of data values?



- **Mean** (expected value of *X*): $\mu = \langle x(t) \rangle$

$$\mathrm{E}[X] = \sum_{i=1}^{k} x_i \, p_i = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k$$

- **Variance**: $\sigma^2 = Var\,(X) = \mathrm{E}[(X-\mu)^2]$

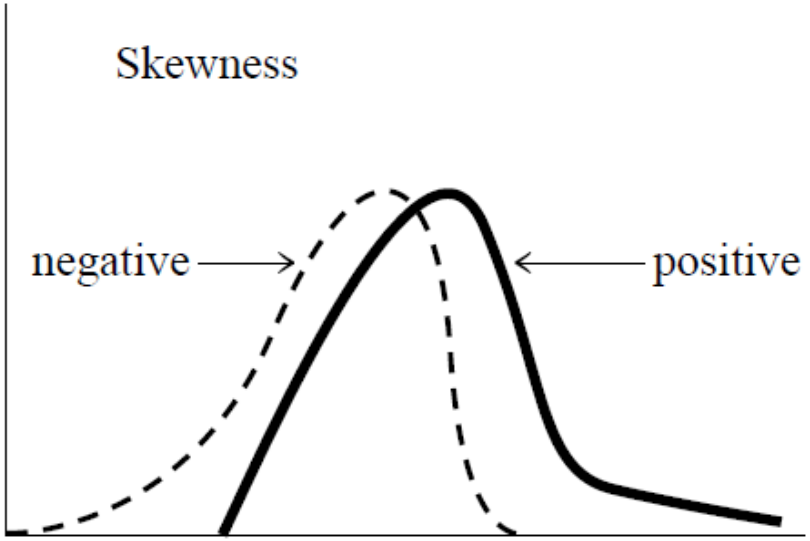
- **Skewness**: "measures" the asymmetry of the distribution

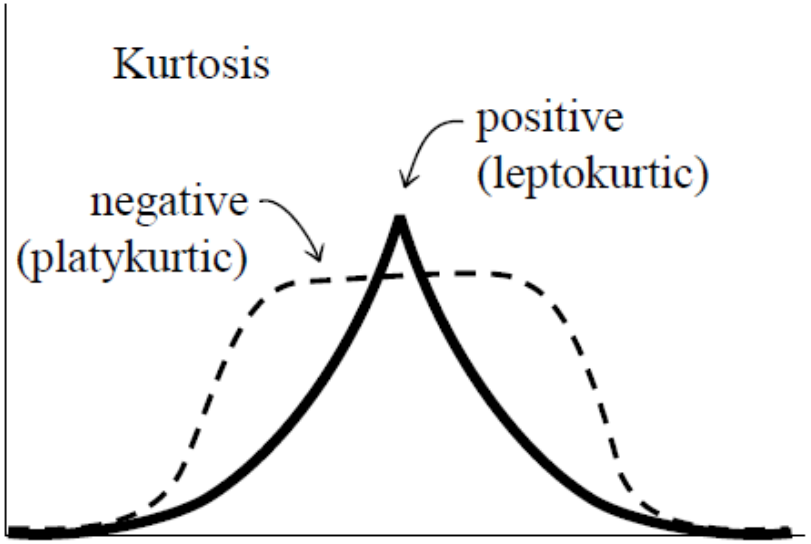
$$Z = \frac{X - \mu}{\sigma} \qquad S = \mathrm{E}[Z^3]$$

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

$$K = \mathrm{E}[Z^4]$$

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



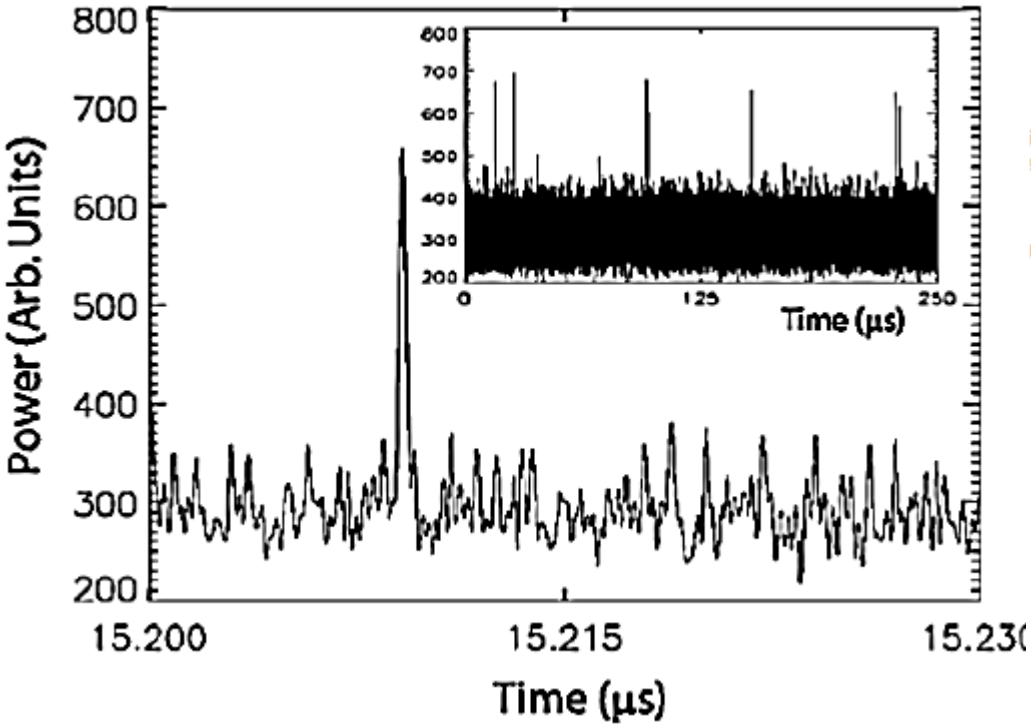


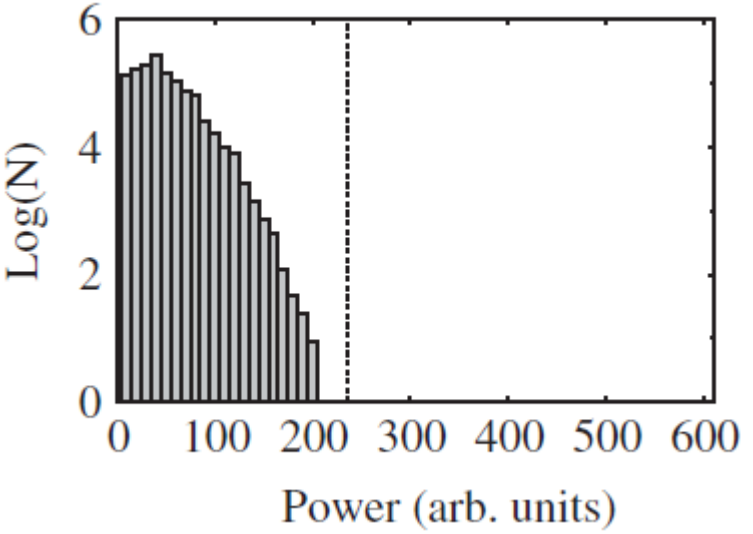
- **K<3:** the distribution produces fewer and less extreme outliers than the normal distribution. An example is the uniform distribution.
- **K=3:** Normal Gaussian
- **K>3:** the tail approaches zero more slowly than a Gaussian, and therefore produces more outliers than the normal distribution. An example is the Laplace distribution.

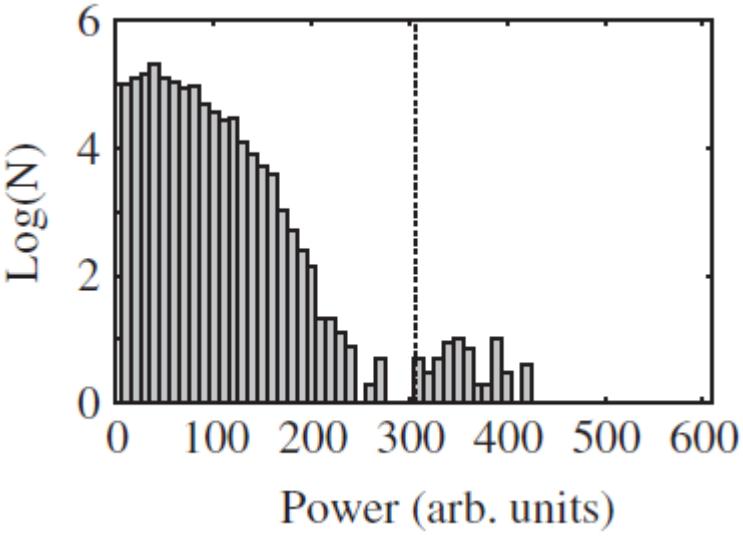
Press WH et al. Numerical recipes: the art of scientific computing (Cambridge University Press)

Tailed distribution? “Extreme values”

Example: optical rogue waves



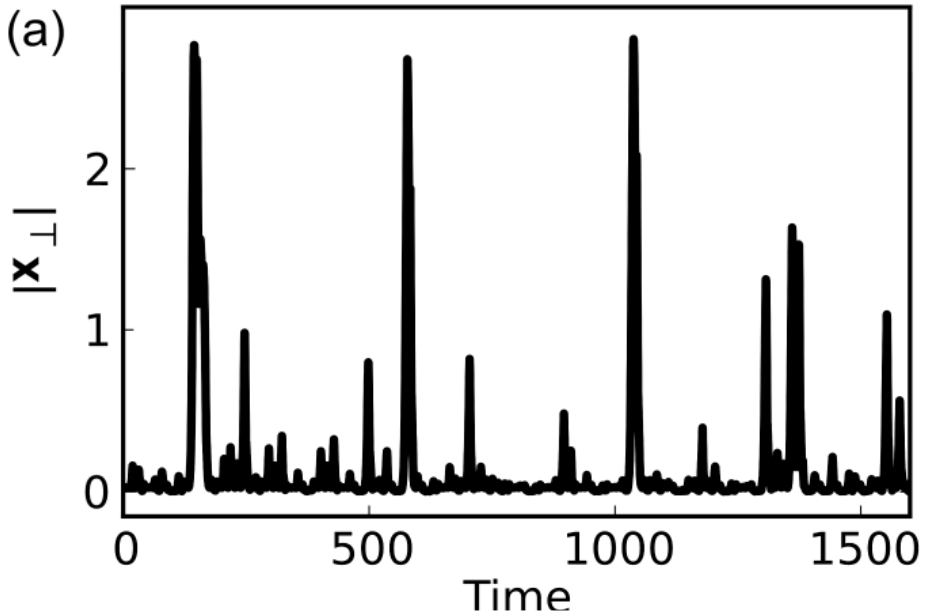


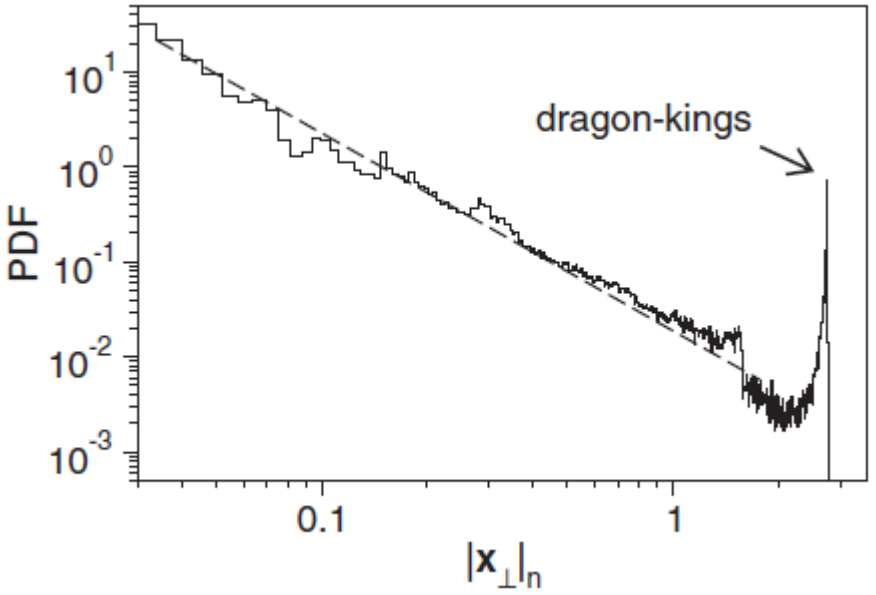


Rogue wave if the pulse height is: $I > \langle I \rangle + 8\sigma$

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

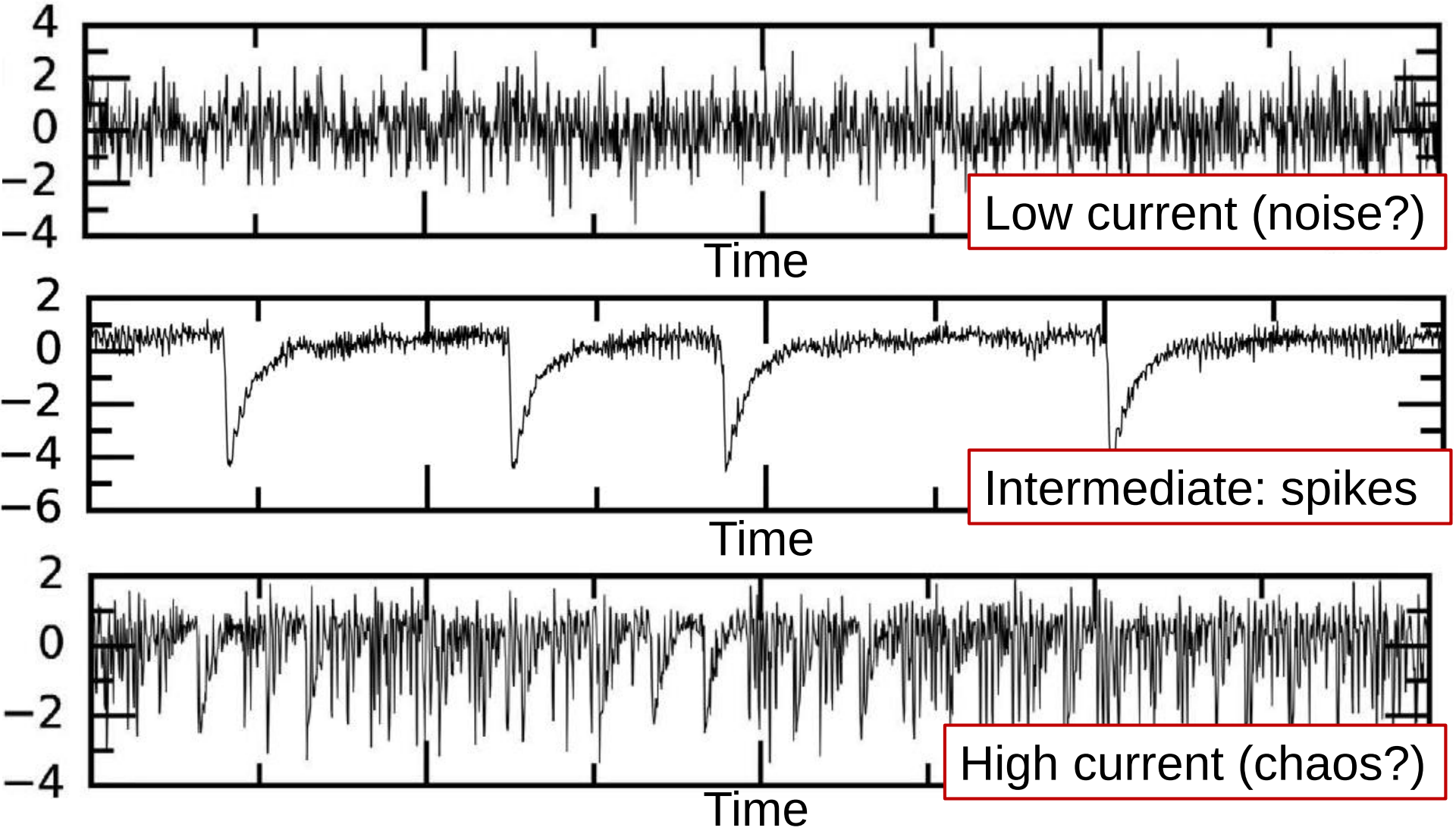
“Dragon kings”: extreme values or “outliers”?





Hugo L. D. de S. Cavalcante et al, Phys. Rev. Lett. 111, 198701 (2013)

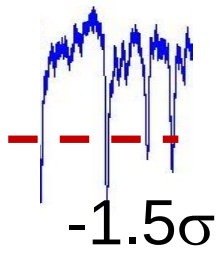
Example: intensity emitted by a diode laser with feedback, as the pump current increases (video seen in previous class).

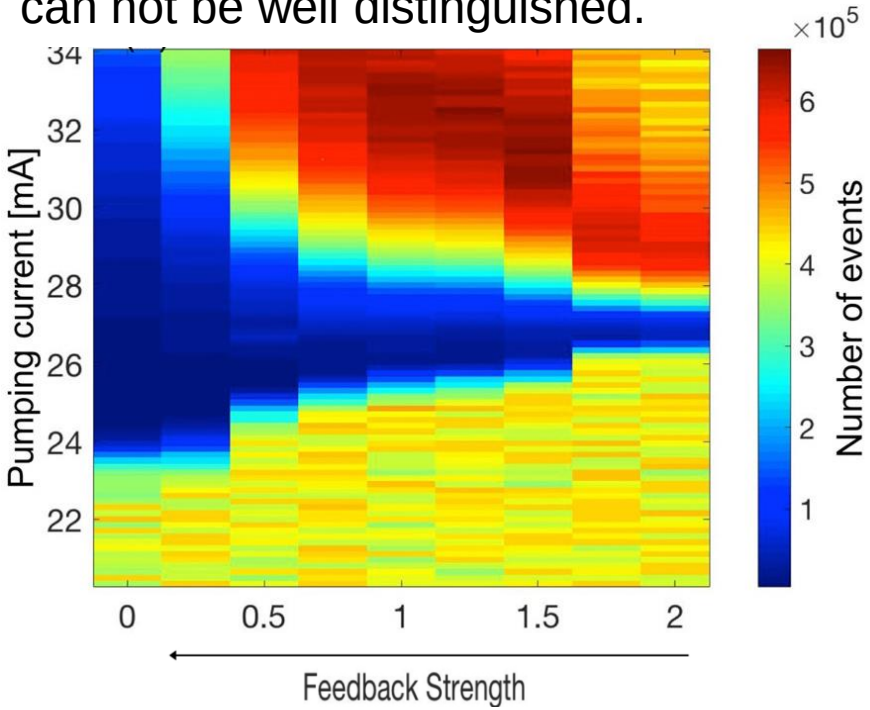


Can we distinguish *quantitatively* the three regimes?

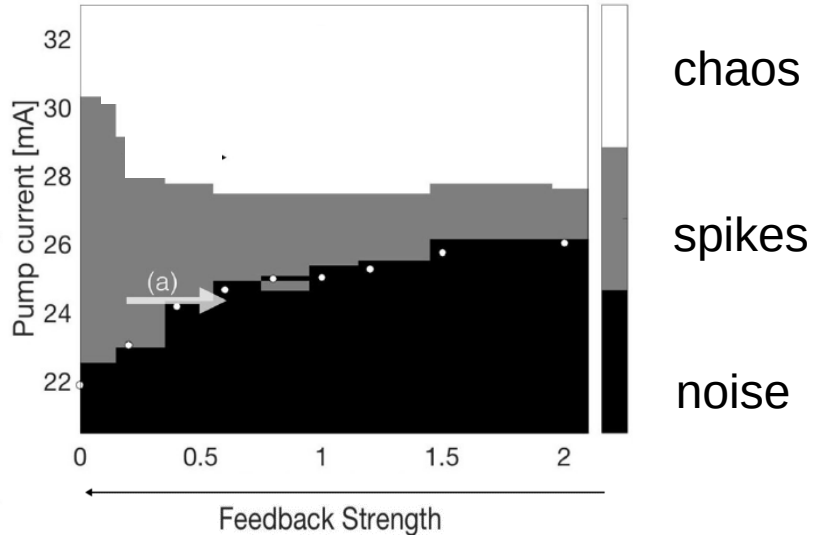
We recorded a large number of time series varying two experimental parameters: feedback strength and laser current.

First method: **Count** the number of “spikes events” (# of times the intensity falls below a threshold).
Problem: chaos and noise can not be well distinguished.





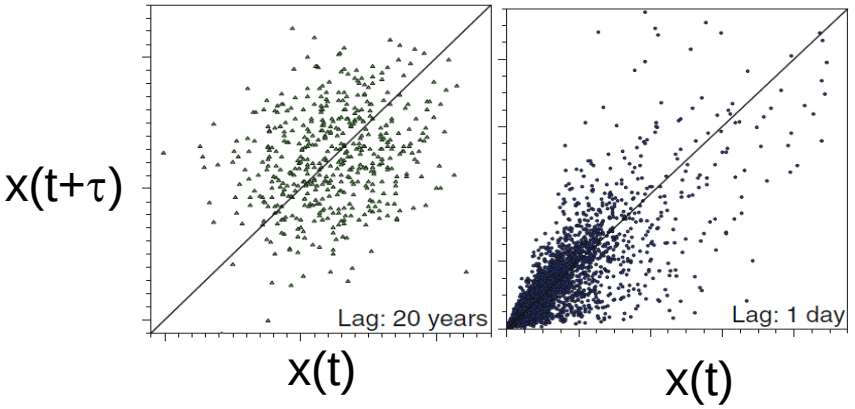
Second method: Analyze the distribution of intensity values
If K in 3-3.3 (Gaussian dist.) ⇒ Noise
Else
If σ increases with the pump current (keeping feedback constant) ⇒ Spikes
Else ⇒ Chaos



[Panozzo et al, Chaos 27, 114315 (2017)]

Autocorrelation function (ACF)

For two geophysical time series



- The return map allows us to see if x(t) and x(t+τ) are "*correlated*".
- How to **quantify**?

- The autocorrelation function $C(\tau) = \dfrac{\langle [x(t)-\mu][x(t+\tau)-\mu] \rangle}{\sigma^2}$
- By definition: *C*(0)=1

- **C(τ)=0** indicates that x(t) and x(t+τ) are **uncorrelated**.
- **C(τ)>0** indicates **persistence**: large values tend to follow large ones, and small values tend to follow small ones, on average more of the time than if the time series were uncorrelated.
- **C(τ)<0** indicates **anti-persistence**: large values tend to follow small ones and small values large ones.

Autocorrelation function (ACF)

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

$$C(\tau) = \frac{1}{\sigma_x^2} \frac{1}{N - \tau} \sum_{t=1}^{N-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x})$$

- After processing the time series to remove its mean value and to normalize to σ [$x_i=(x_i-\mu)/\sigma$]: $\mu=0$, $\sigma=1$

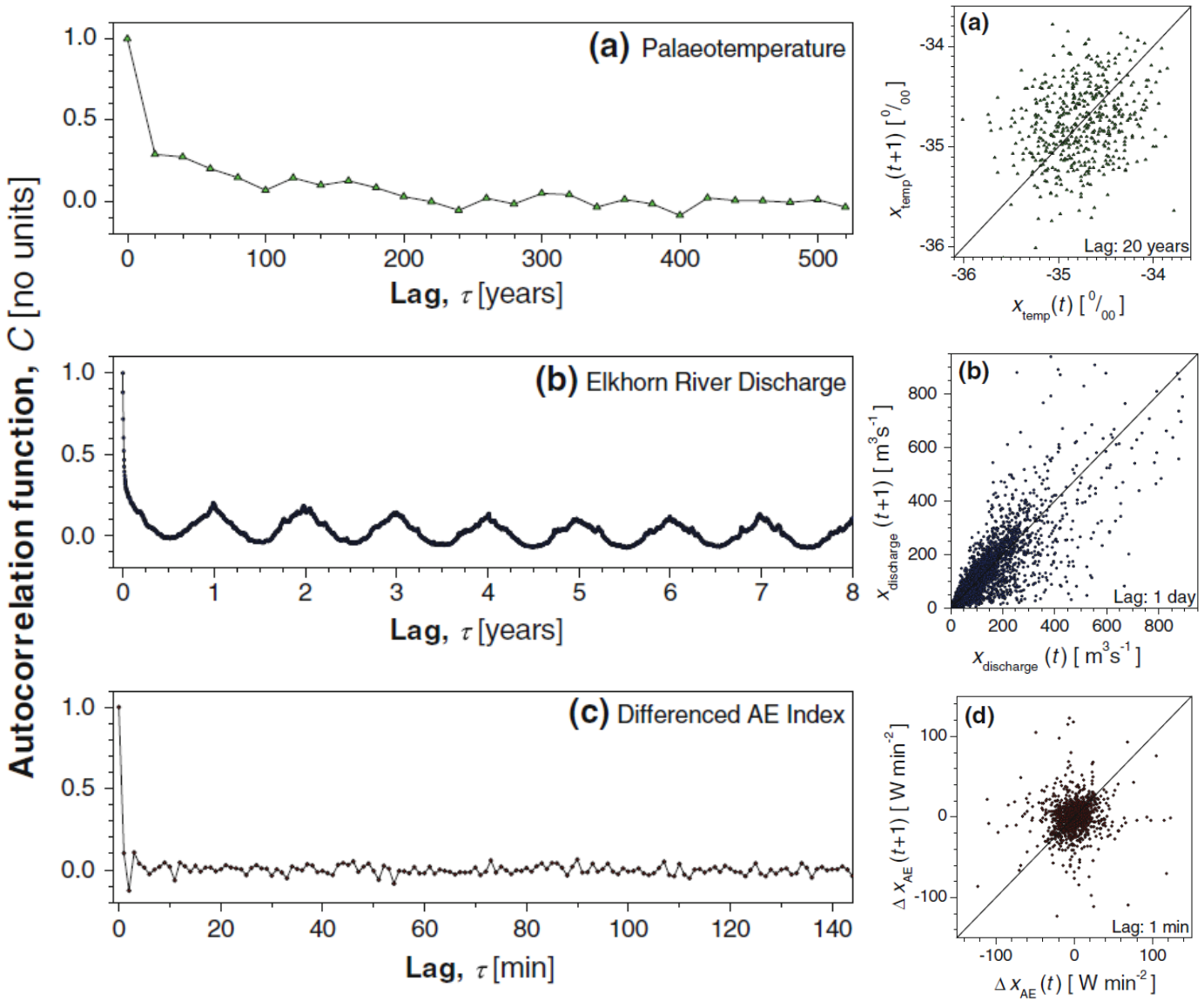
$$C(\tau) = \langle x(t)x(t+\tau) \rangle$$

- For a stationary process μ and σ are constant in time

$$\Rightarrow C(\tau) = C(-\tau)$$

- Correlogram: plot of $C(\tau)$ vs. τ.
- **Matlab**: **corr(A)**. A is a matrix with two columns: x(t) and x(t+τ). Returns: $\begin{bmatrix} 1 & C(\tau) \\ C(\tau) & 1 \end{bmatrix}$

Back to the three examples of geophysical time series



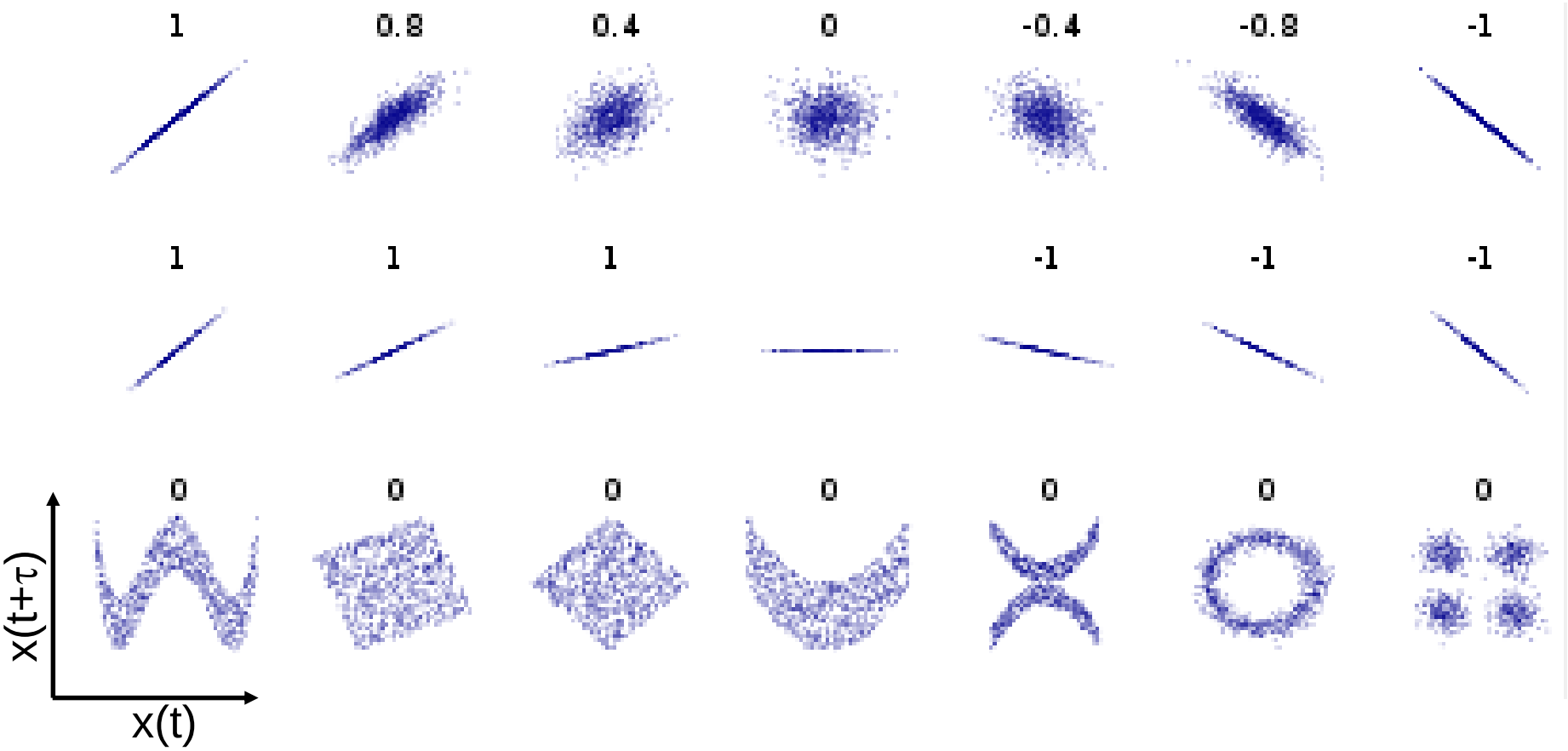
Slow decay: **long-range** correlations.

Rapid decay: **short-range** correlations.

$$|C(\tau)| \leq \kappa_0 \exp(-\kappa\tau)$$

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

Problem with the ACF: it only detects linear correlations between two data points ⇒ it is important to analyze nonlinear correlations and higher order correlations.



Source: wikipedia

Hands-on exercise 2: ACF

3.1 Generate a time series with the map: $x(i+1) = 1 - 2\,x(i)^2$

and calculate the ACF.

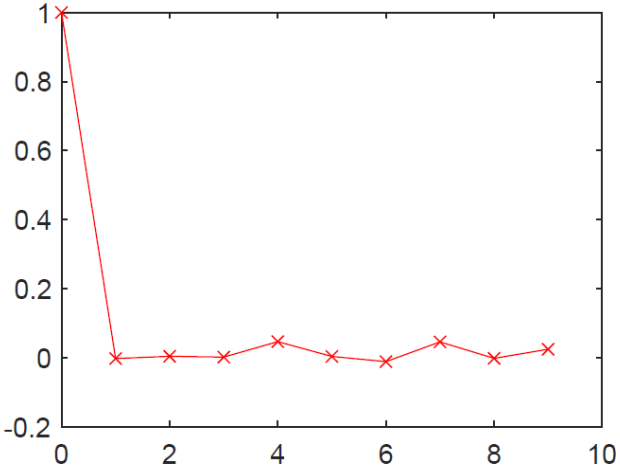
3.3 Plot the return map *x*(i+1) vs. *x*(i). Are *x*(i+1) and *x*(i) correlated?

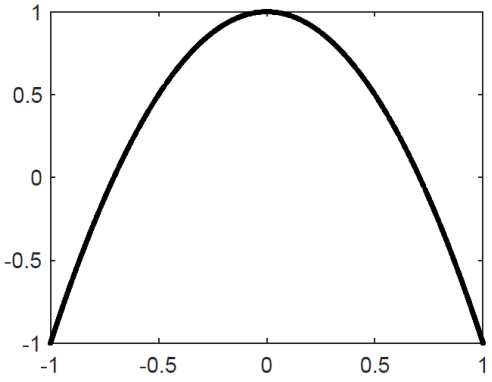
3.4 Generate a time series with an **autoregressive process** of order 1, AR(1), and calculate the ACF.

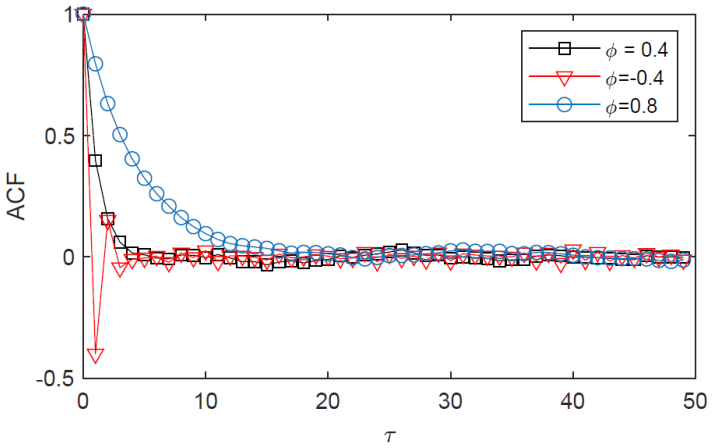
$$x(i+1) = \phi\, x(i) + \varepsilon(i)$$

Gaussian white noise (μ=0, σ=1)

$-1 < \phi < 1$



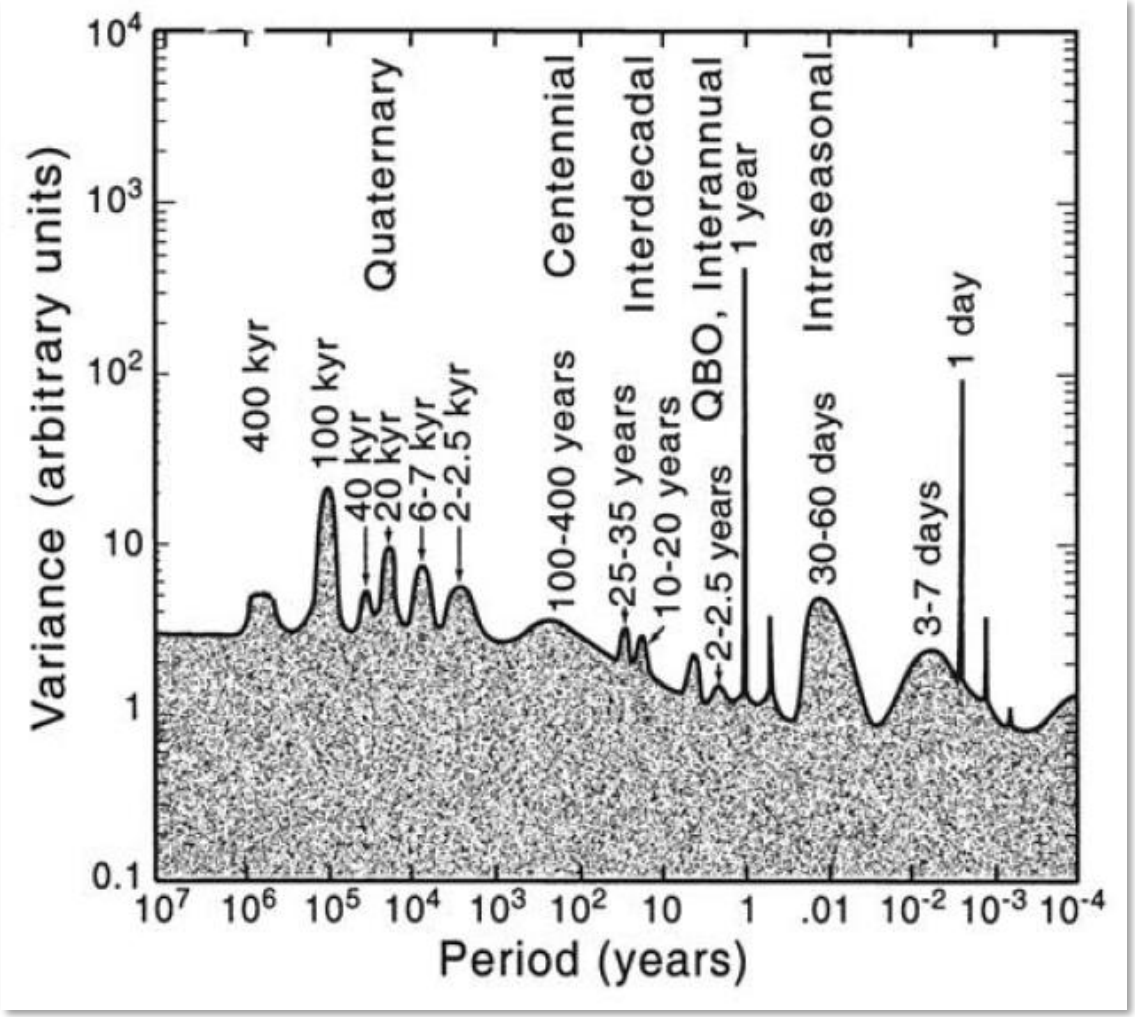




Fourier analysis

Describes a time series in terms of a linear superposition of oscillatory components.

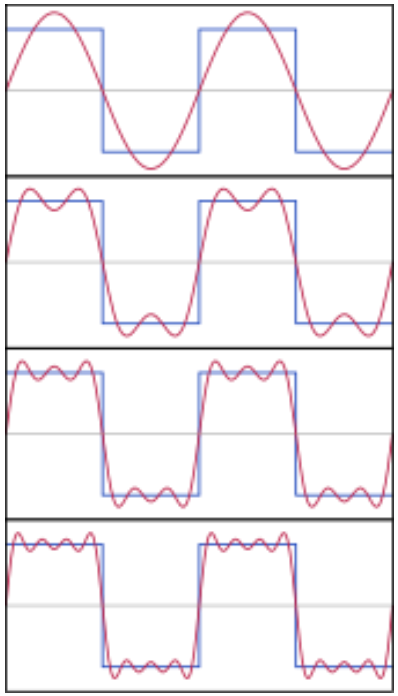
Example: climatic time series have oscillations with a wide range of time scales.



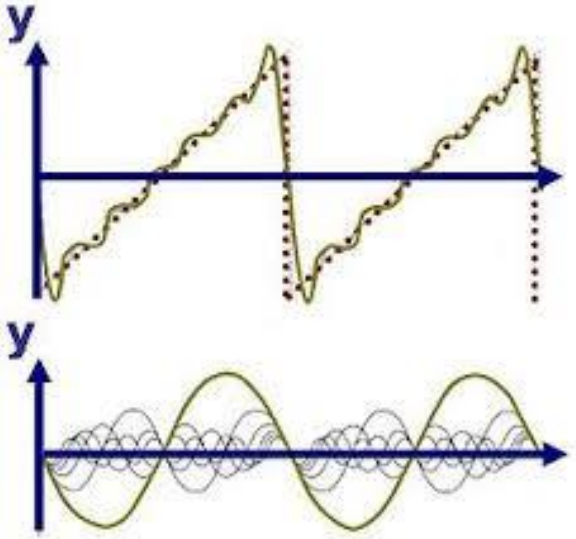
Adapted from M. Ghil (2002).

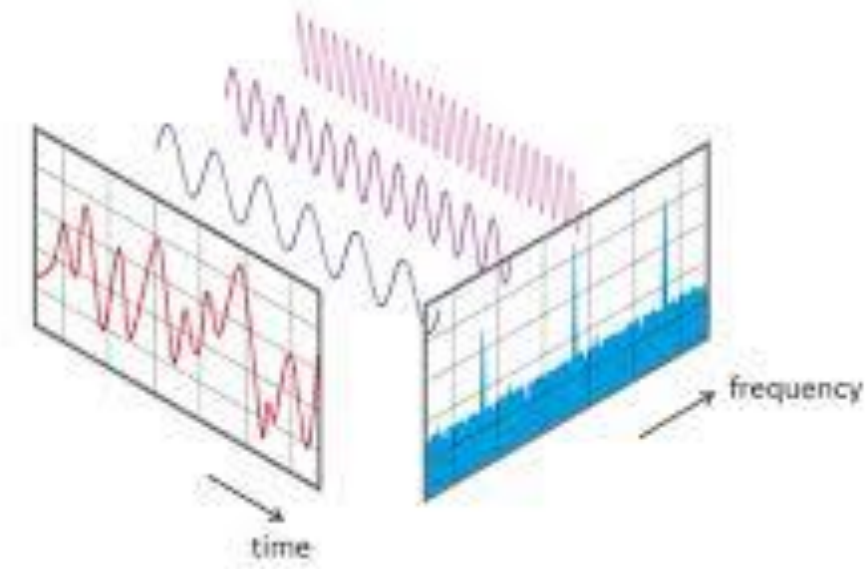
Examples

Sum of the first four terms of the Fourier series of a square wave.



The sum of many terms of the Fourier series gives a triangular signal.





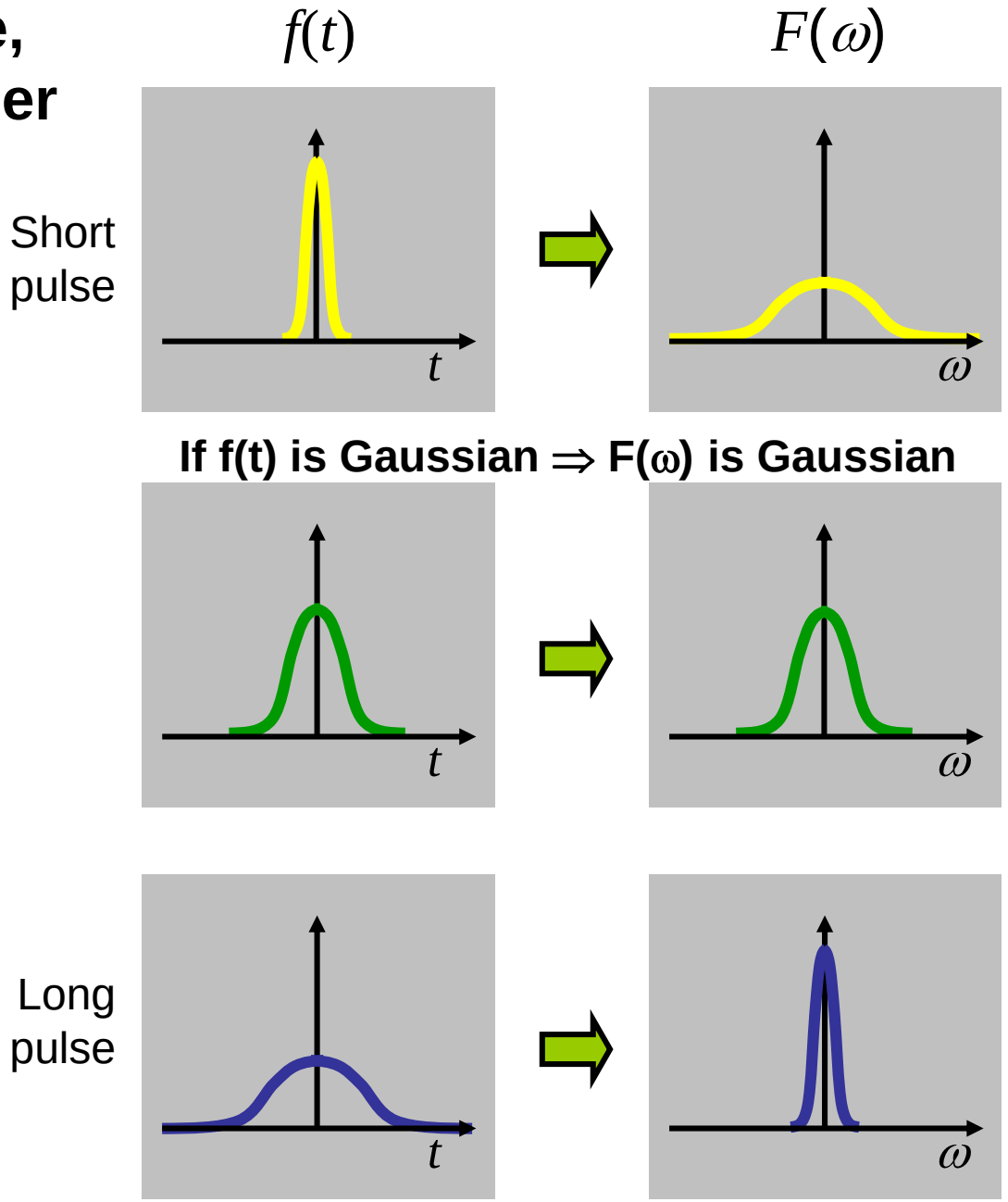
The Fourier spectrum of a sinusoidal signal is the "Delta" function.

Which is the Fourier spectrum of a short pulse (localized in time)?

The shorter the pulse, the broader the Fourier spectrum.

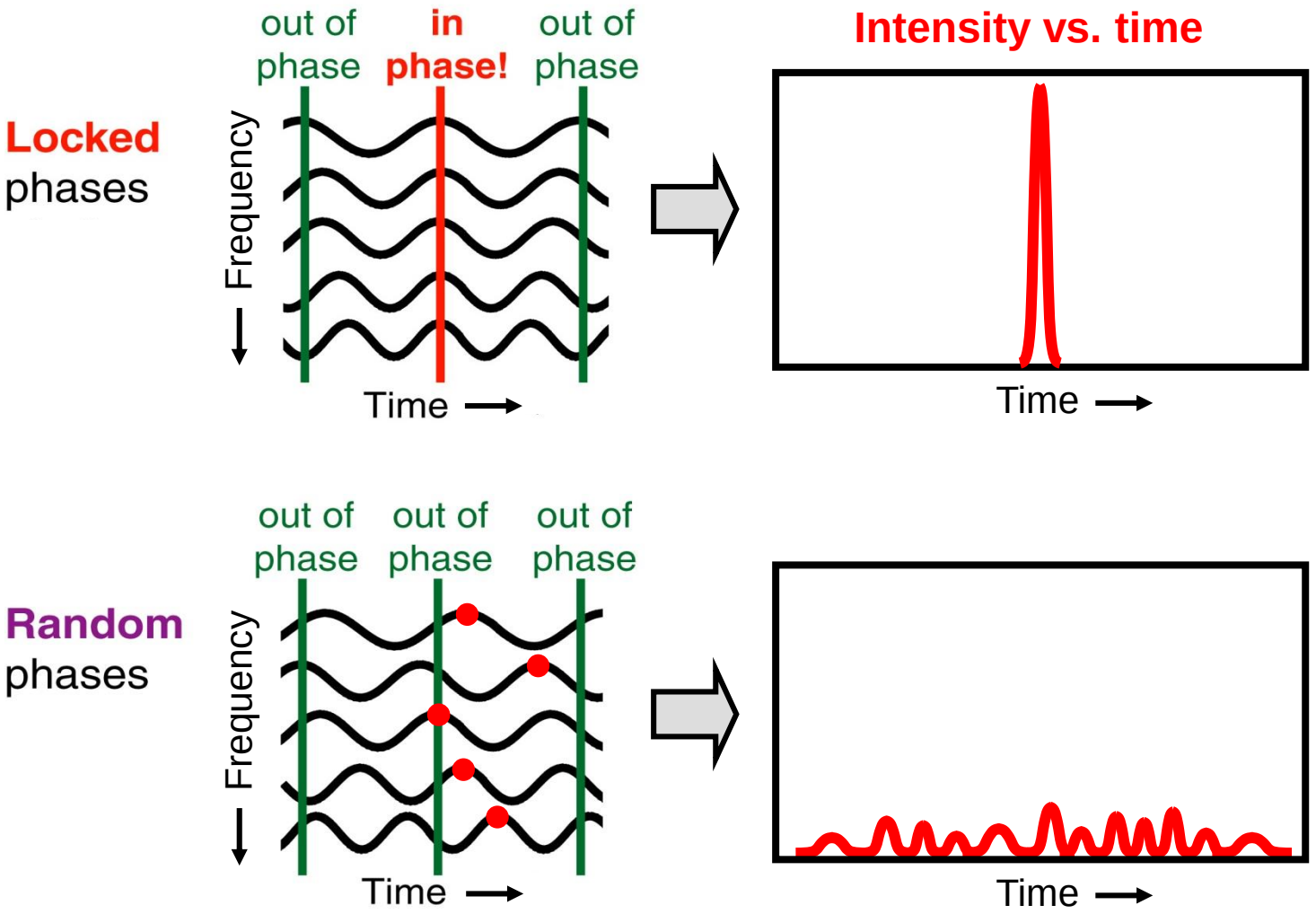
One cannot simultaneously sharply localize a signal in both, the time domain and the frequency domain.

(This is the essence of the **Uncertainty Principle** of Quantum Mechanics)



Source: Prof. R. Trebino lectures, Georgia Tech, USA

Superposition of oscillations with different frequencies: depending on their relative phases, the result is a short pulse or noisy fluctuations



Source: Prof. R. Trebino lectures, Georgia Tech, USA

Discrete Fourier Transform

- $x = \{x_0, x_1, \ldots x_{N-1}\}$ is described as a superposition of waves.

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k \cdot e^{i\frac{2\pi}{N}kn} \qquad X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{i2\pi}{N}kn}$$

- The DFT of x is the set of complex numbers

  $X = \{X_0, X_1, \ldots X_{N-1}\}$

  associated to the frequencies $f_k = k/(N\Delta)$ with Δ = sampling time (time interval between x_i and x_{i+1}) and N = # of data points.

- **Important property**: If x is a real signal: $X_K = (X_{N-k})^*$

- Another important property: The Fourier transform is a **linear** transformation. If $\{X_K\}$ is the DFT of $x = \{x_0, x_1, \ldots\}$ and $\{Y_K\}$ is the DFT of $y = \{y_0, y_1, \ldots\}$, then the DFT of $(ax + by)$ is $\{aX_K + bY_K\}$.

Power spectral density (PSD)

- DFT is the set of complex numbers, $X = \{X_0, X_1, \ldots X_{N-1}\}$

- The PSD is a set of real numbers that give the "strength" of each frequency component:

  $$\text{PSD} = \{|X_0|^2, |X_1|^2, \ldots, |X_{N/2}|^2\}$$

- *Wiener-Khinchin theorem*: if the time series is stationary the PSD is the Fourier transform of the autocorrelation function.

$$S(f) = \sum_{k=-\infty}^{\infty} r_{xx}[k] e^{-i(2\pi f)k}$$

$$r_{xx}[k] = \mathrm{E}\left[x[n]^* x[n-k]\right]$$

Fast Fourier Transform (FFT)

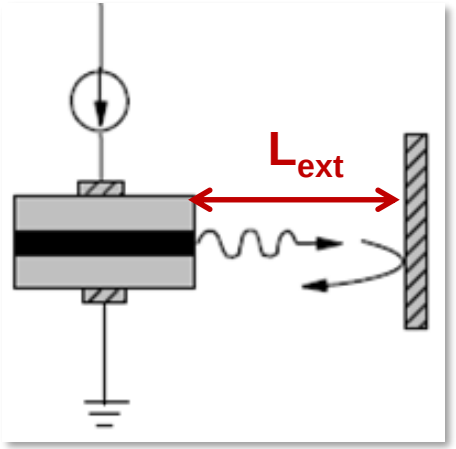
- Fast Fourier Transform (FFT) algorithm applied to x= $\{x_0, x_1, \ldots x_{N-1}\}$ returns $X = \{X_0, X_1, \ldots X_{N-1}\}$
- It is faster if $N=2^m$.
- The order of the components of X depends on the implementation.
- Matlab implementation:

$f_k = k/(N\Delta) \Rightarrow df = 1/(N\Delta)$

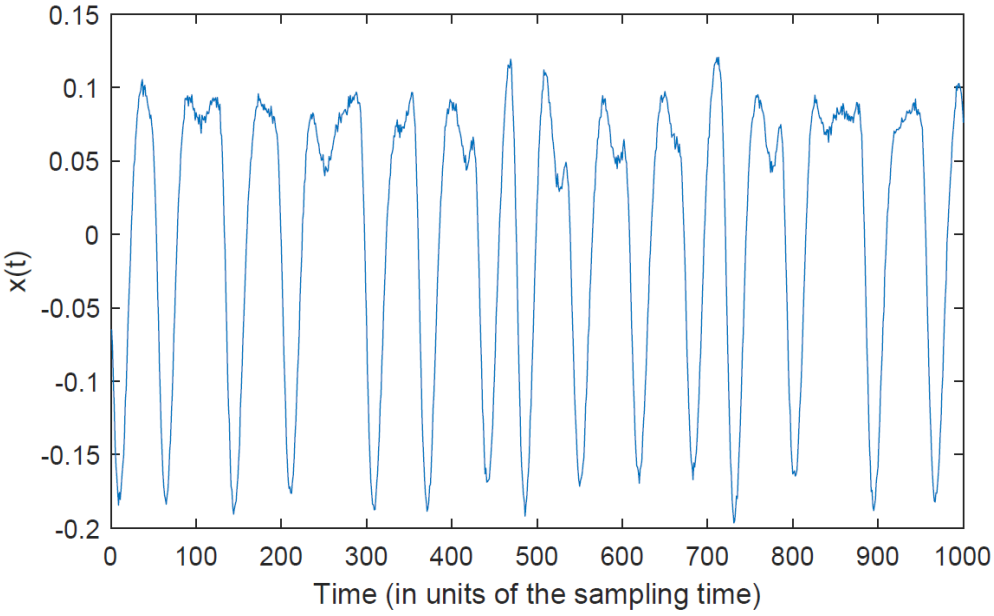
x is a real signal: $X_K = (X_{N-k})^*$
$\Rightarrow$ disregard second half.
q(0): dc value –we also disregard.

```
x=datos(:,1);
n = length(x)
dt=1.
df=1./(dt*n);
f=df.*[1:1:(n/2)-1];
z=fft(x);
q=z.*conj(z);
q=q(2:n/2);
q=q./max(q);
semilogy(f,q)
```

Hands-on exercise 3: Fourier spectrum of the intensity of the laser with optical feedback.

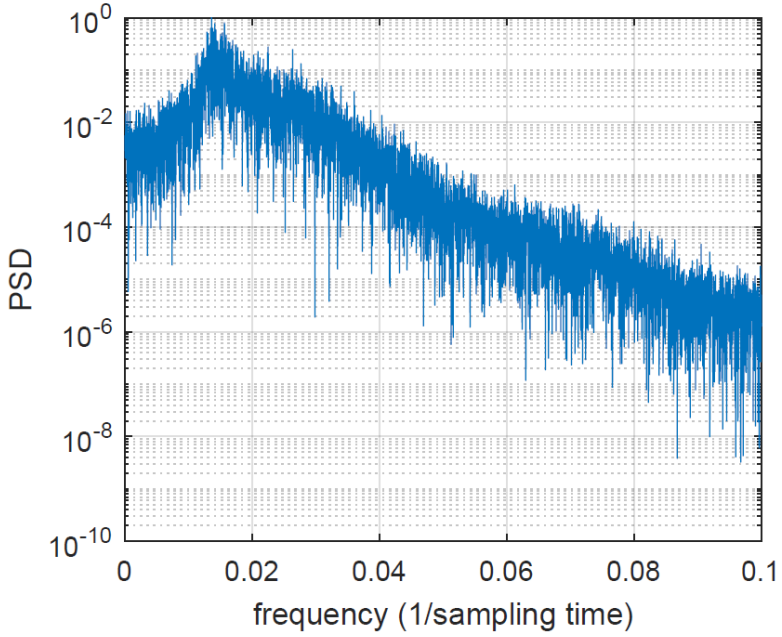


- Plot the time series & estimate the mean inter-spike-interval.
- Calculate the FFT. Which is the frequency of the main peak?
- Verify f_{max} = 1/<ISI>



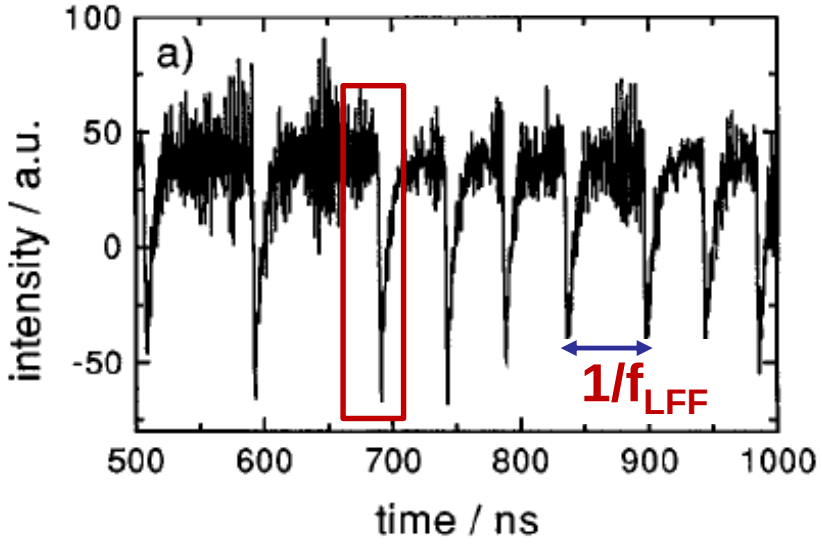
<ISI>≈1000/14

Lower part of the spectrum

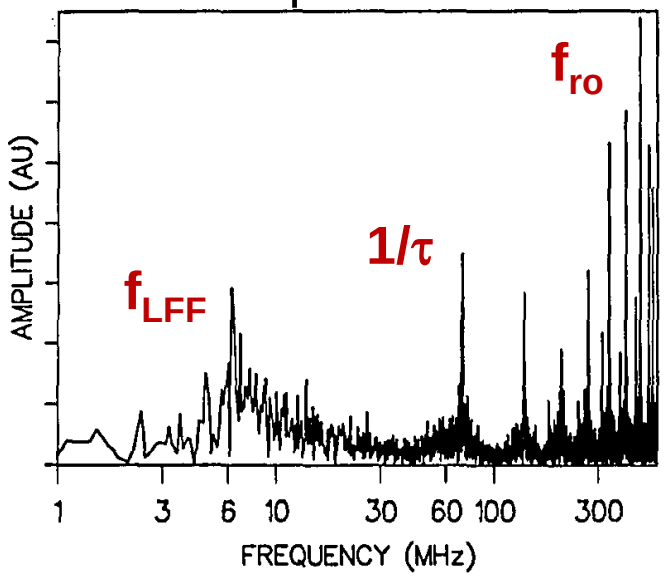


f_{max} ≈0.014

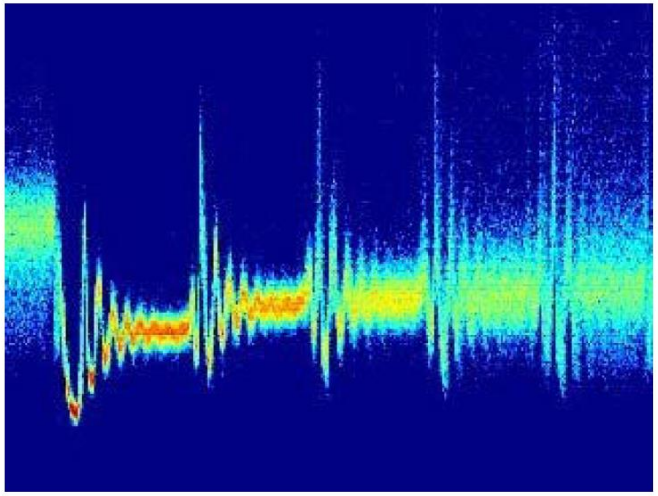
Warning! The importance of the sampling time: some examples of time series with higher temporal resolution



Fourier spectra



The laser intensity after a drop: steps of duration τ with relaxation oscillations



Time

$$\tau = \frac{2L_{ext}}{c}$$

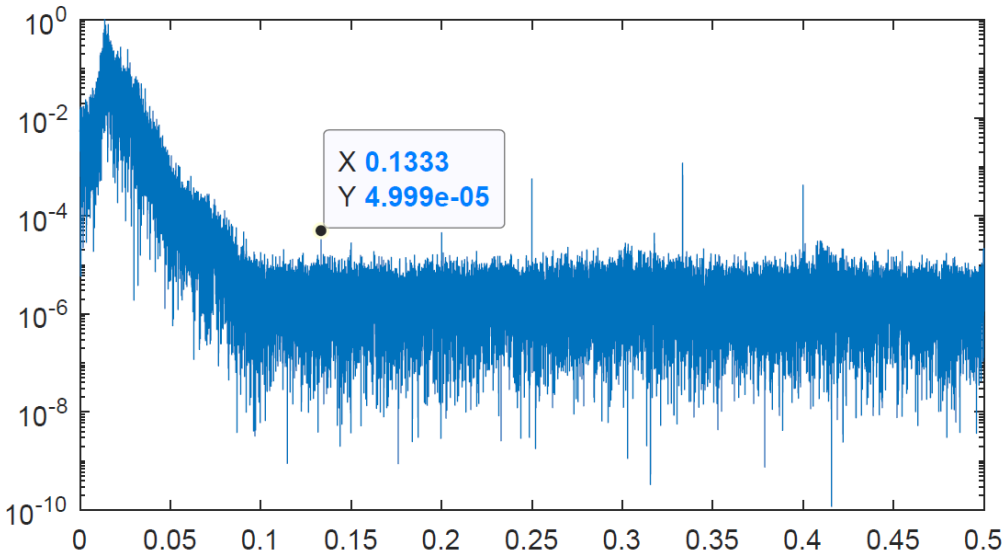
If L_{ext} = 1 m
$\Rightarrow \tau$ = 6.7 ns

Three main frequencies:
- relaxation oscillations f_{ro}
- the “steps”: external cavity frequency $1/\tau$
- the “drops”: f_{LFF}

M. Sciamanna (PhD Thesis 2004).
Langley et al, Opt. Lett. 19, 2137 (1994).

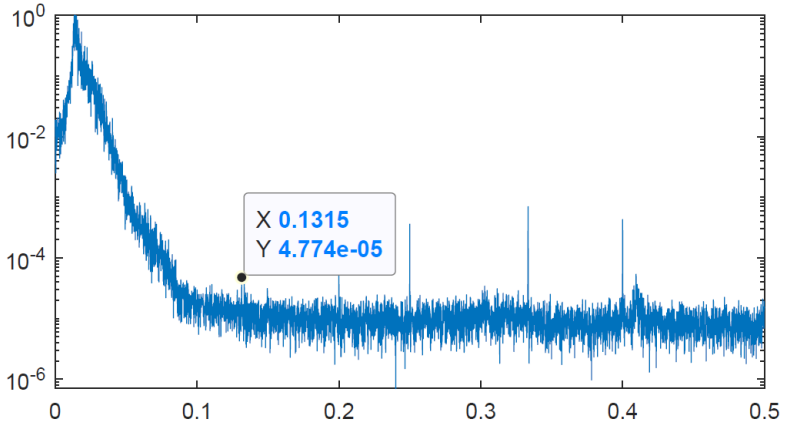
Can we infer the length of the external cavity?

$$\tau = \frac{2L_{ext}}{c}$$



Frequency in units of 1/dt.

Dividing the time-series in 5 non-overlapping segments and averaging the PSD in each segment.



Let's assume dt=1 ns

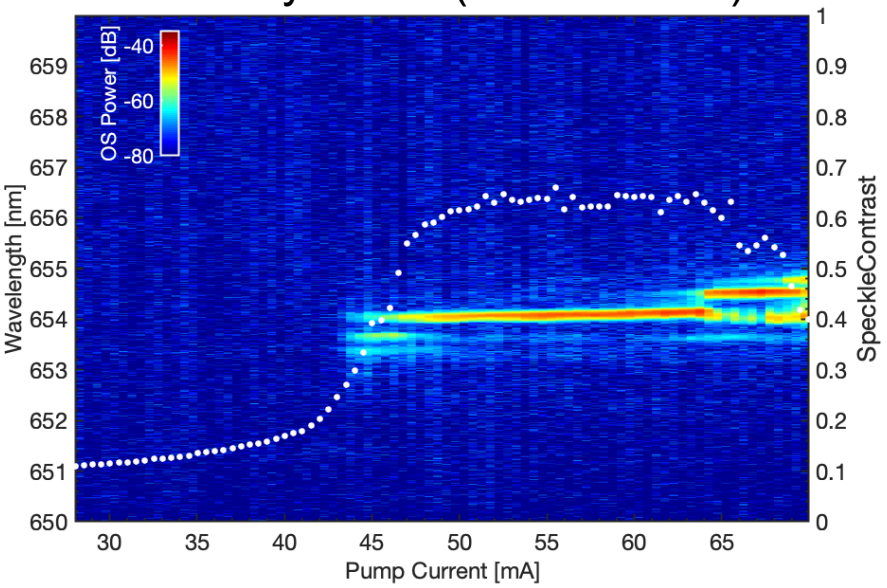
$\Rightarrow$ f is measured in GHz.

f = 0.1333 GHz (= 133 MHz)

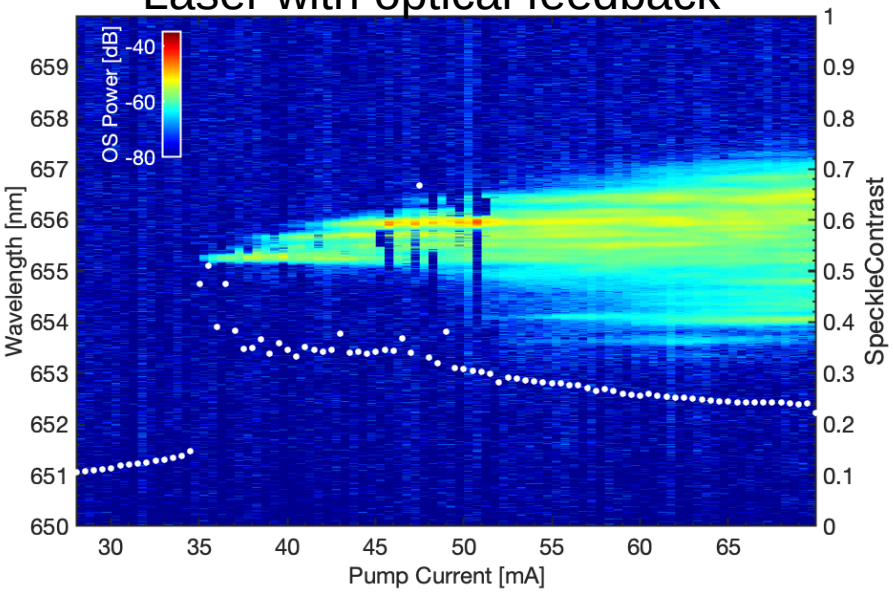
τ = 1/f = 7.5 ns. L_{ext} = c τ/2 = 1.125 m

Spectrogram: Fourier spectrum as a function of time, or as a function of a control parameter.

“Solitary” laser (no feedback)



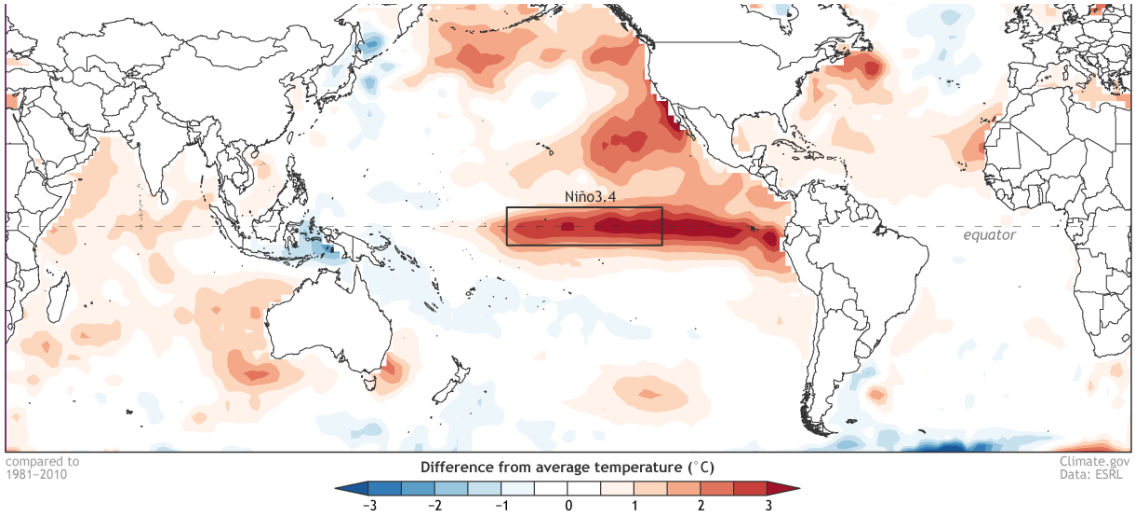
Laser with optical feedback



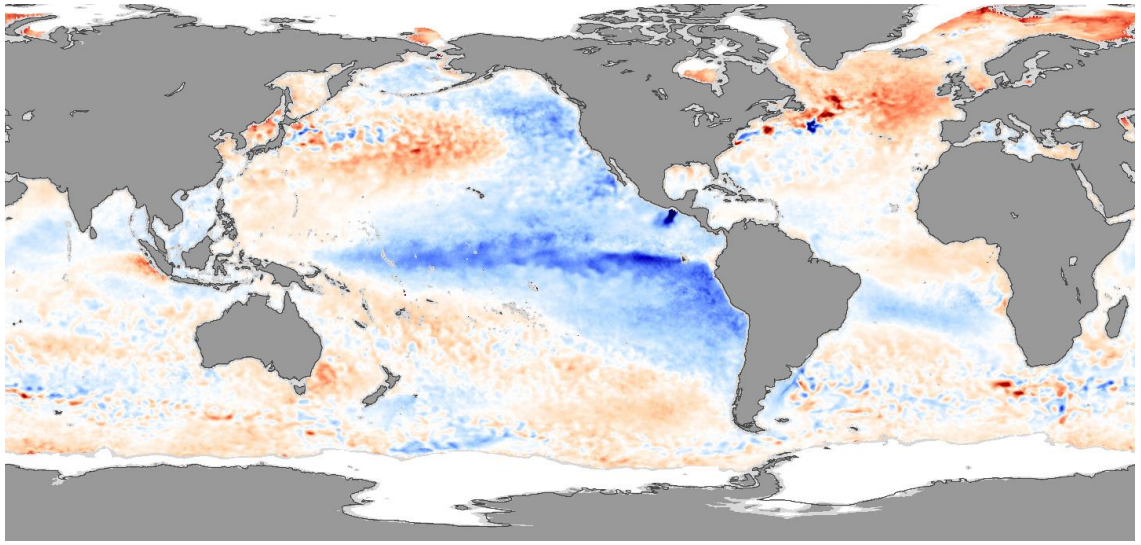
Hands-on exercise 4: analysis of the index Nino3.4

- What is it?
- ENSO: El Niño / La Niña-Southern Oscillation.
- Nino 3.4: sea surface temperature (SST) anomaly, averaged in a geographical rectangular region, and monthly averaged.

SST anomaly during a El Niño event (Nov. 2015)



SST anomaly during a La Niña event (Nov. 2007)



Hands-on exercise 4: analysis of el Niño index 3.4

- ENSO is the most important climatic phenomena on the planet.
- It occurs with ≈ 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.
- Download data from

https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino34.long.anom.data

In the period 1870-2020: time series with 1800 data points.

Hands-on exercise 4: analysis of el Niño index 3.4

4.1 Plot the time series.

4.2 Plot the return map with τ=1 month, 1 year, 4 years.

4.3 Plot the distribution of data values.

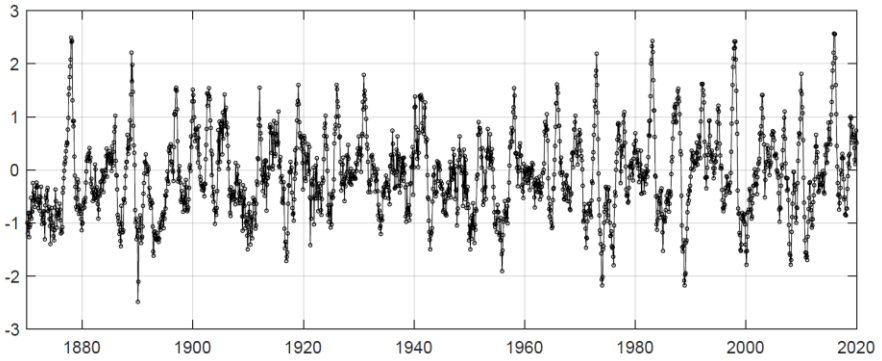
4.4 Calculate μ, σ, S, and K.

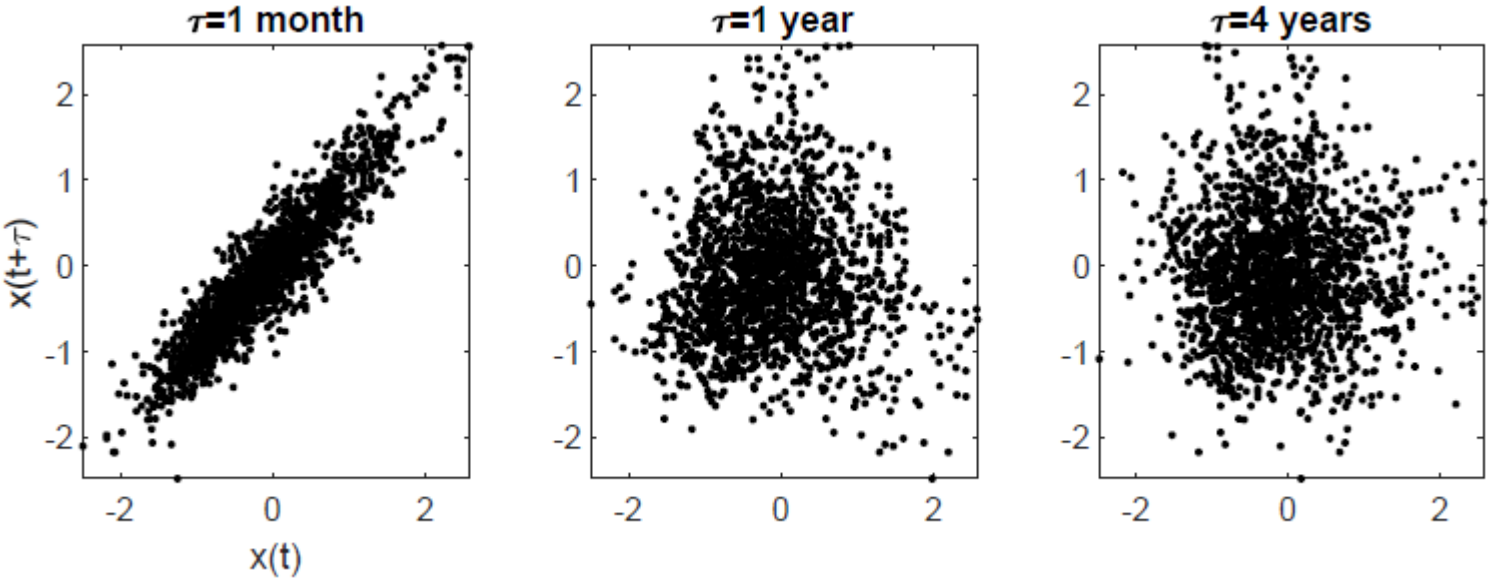
4.5 Calculate the ACF.

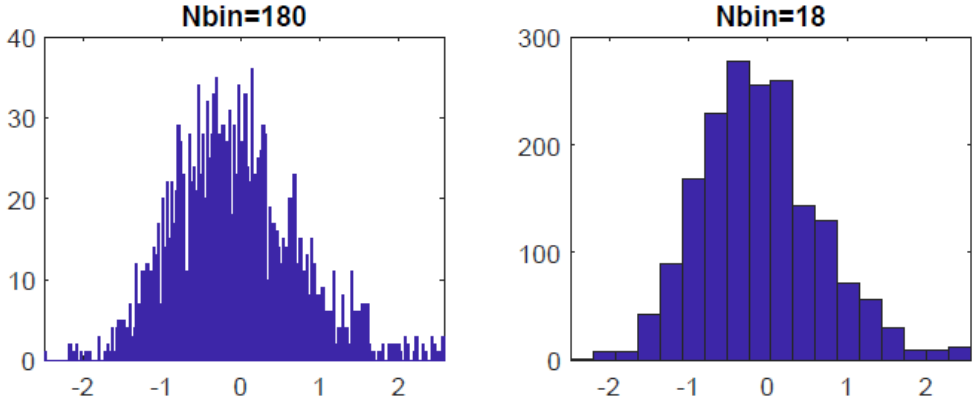
4.6 Calculate the PSD.

4.7 Compare with the ACF & PSD of Gaussian white noise.

μ = -0.0987
σ = 0.7713
S = 0.4547
K = 3.4132



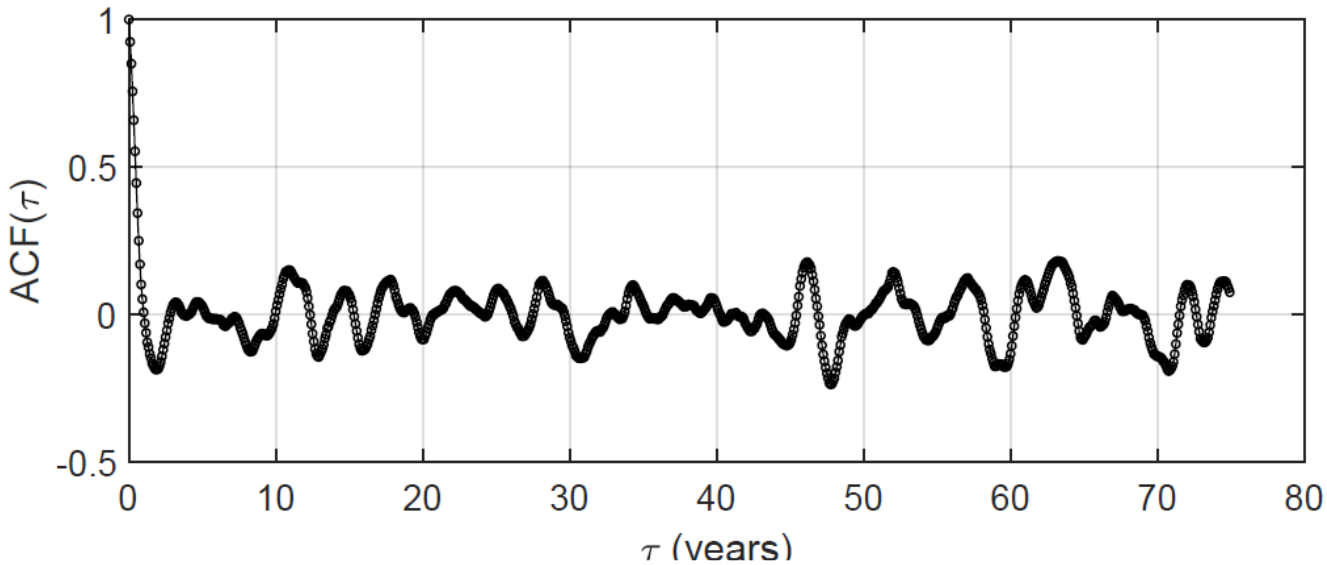


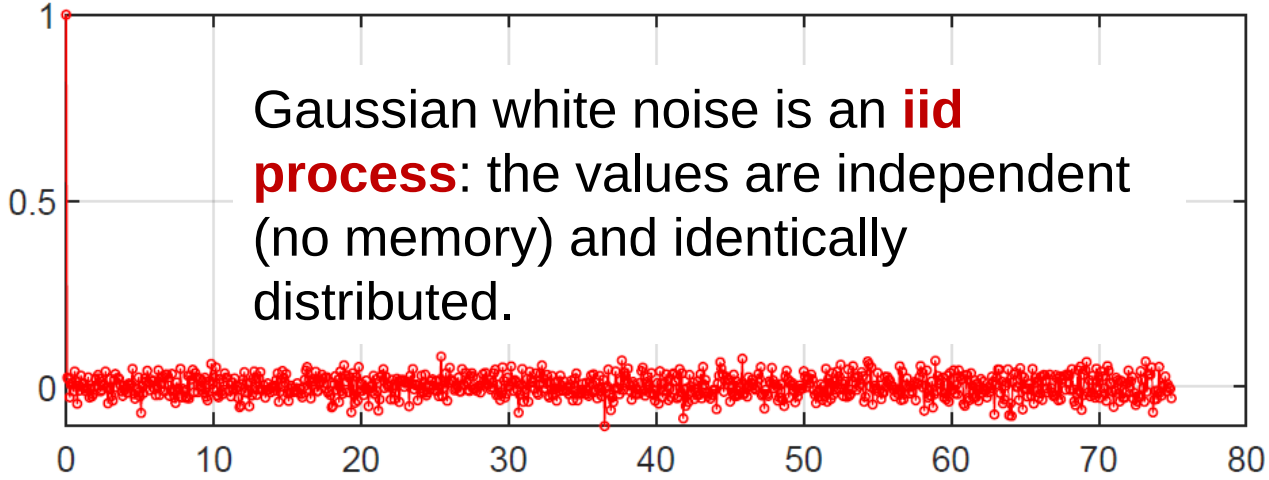


ACF El Niño3.4 index

$$C(\tau) = \langle x(t)x(t-\tau) \rangle = \frac{1}{N-\tau} \sum_{t=\tau}^{N} x(t)x(t-\tau)$$

(μ=0, σ=1)



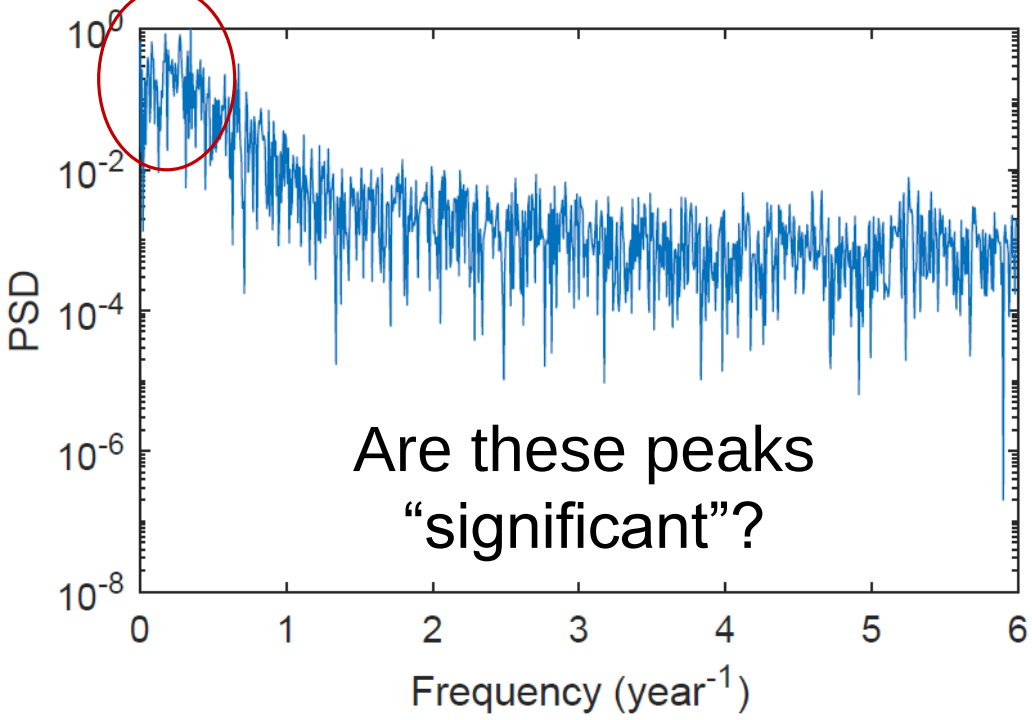


Gaussian white noise is an **iid process**: the values are independent (no memory) and identically distributed.

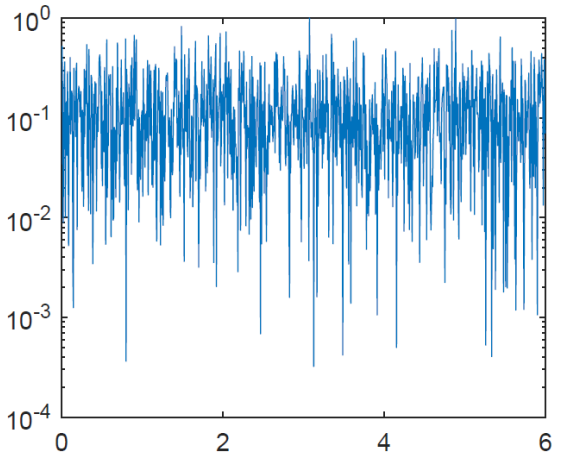
PSD of El Niño3.4 index

```
X=nino3.4(:,1);
z=fft(x);
q=z.*conj(z);
ndats = length(x)
dt=1./12; % (SAMPLING
TIME IN UNITS OF YEARS)
df=1./(ndats*dt);
f=df.*[1:(ndats/2)-1];
q=q(2:ndats/2);
q=q./max(q);
semilogy(f,q)
```

(Main peak T=2.7 years)

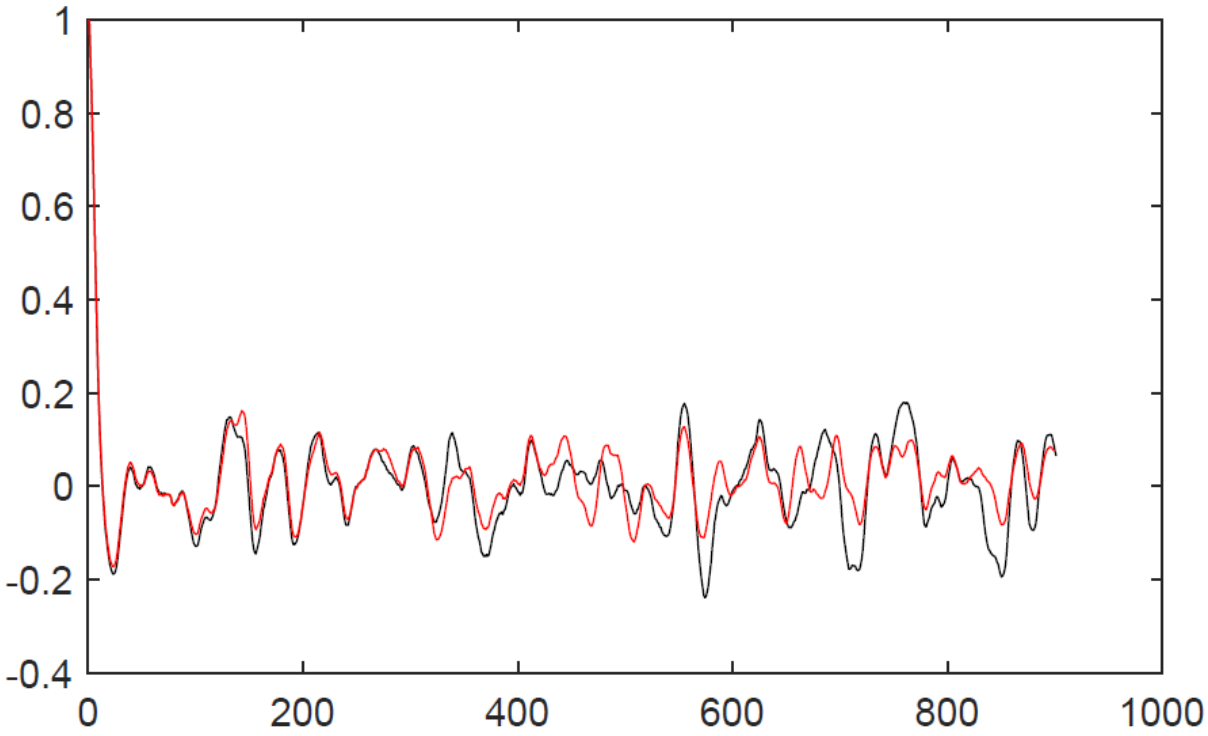


For Gaussian white noise: $C(\tau)=\delta(\tau)$ $\Rightarrow$ PSD is flat (all frequencies have the same contribution).

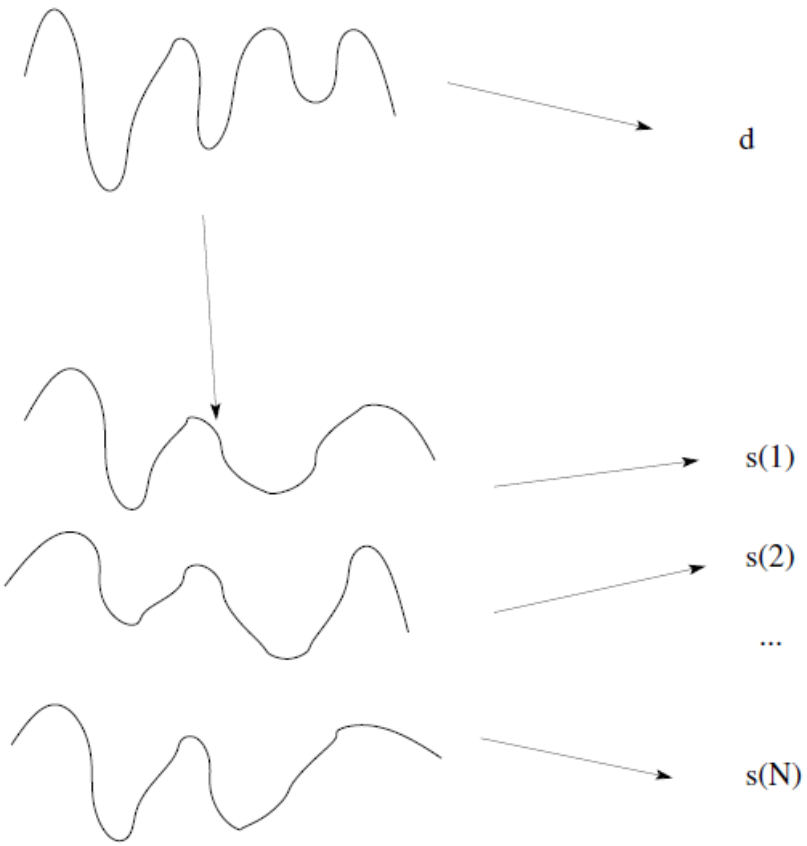


Verification of the *Wiener-Khinchin theorem*: if the time series is stationary PSD is the Fourier transform of ACF, and the inverse Fourier transform of PSD is ACF.

```
plot(acf,'k')
hold
z=fft(x);
q=z.*conj(z);
w=ifft(q);
w=w(1:ndats/2);
plot(w/max(w),'r')
```



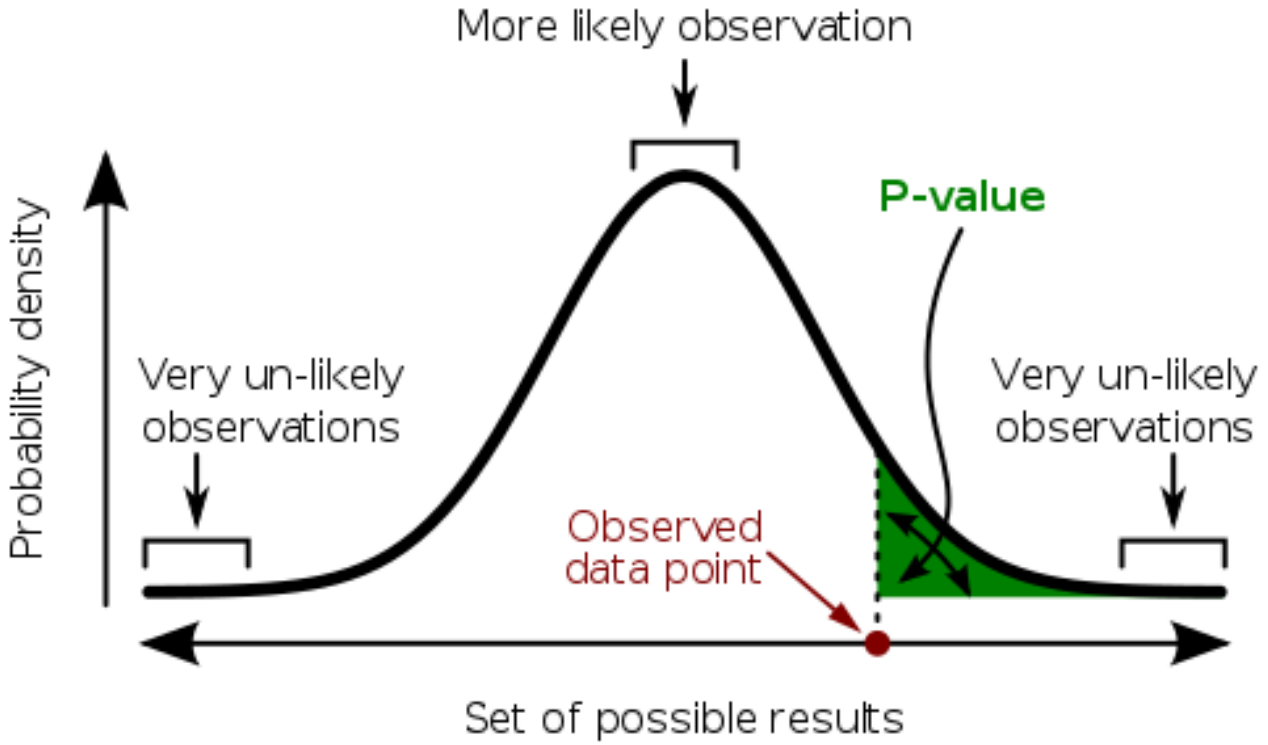
Significance analysis: the method of surrogate data



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.

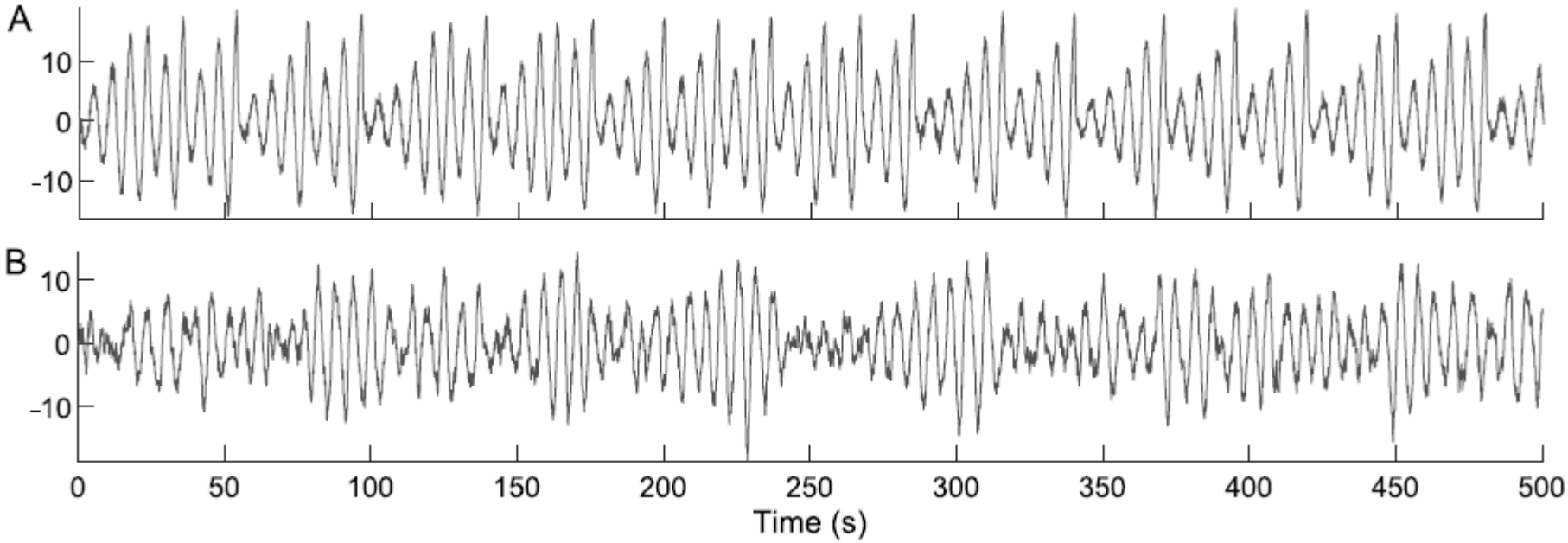
p value



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

Altman and Krzywinski, 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

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

A proper surrogate test detects nonlinearity in A (reject NH) but not in B (fail to reject NH).

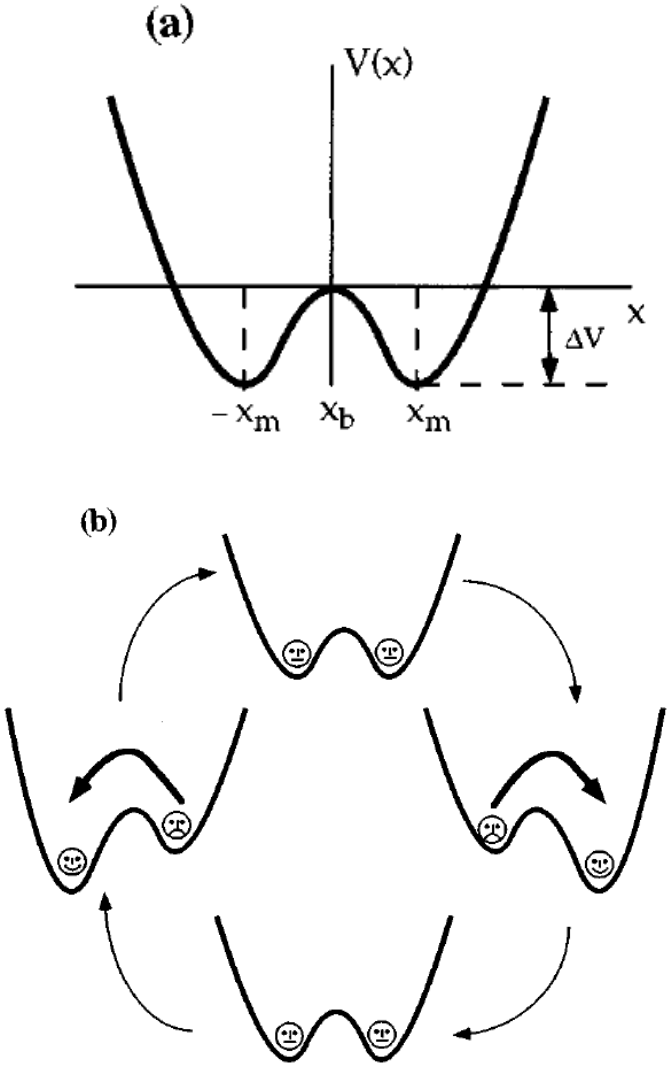
G. Lancaster et al, "Surrogate data for hypothesis testing of physical systems", Physics Reports 748, 1 (2018).

Examples of applications:

- Stochastic resonance in bistable systems

- Coherence resonance in excitable systems

Stochastic resonance in a bistable system: a classical example of the nontrivial interplay of noise and nonlinearity.



In the presence of noise a particle with switch back and forth (stochastic motion).
The average "residence time" in each well (Kramer's time) is $\propto e^{\Delta V/D}$ with D being the strength of the noise.

An external periodic force moves the potential back and forth.

An optimal level of noise can "synchronize" the switching of the particle with the tilting of the potential.

Gammaitoni et al, Reviews of Modern Physics 70, 223 (1998).

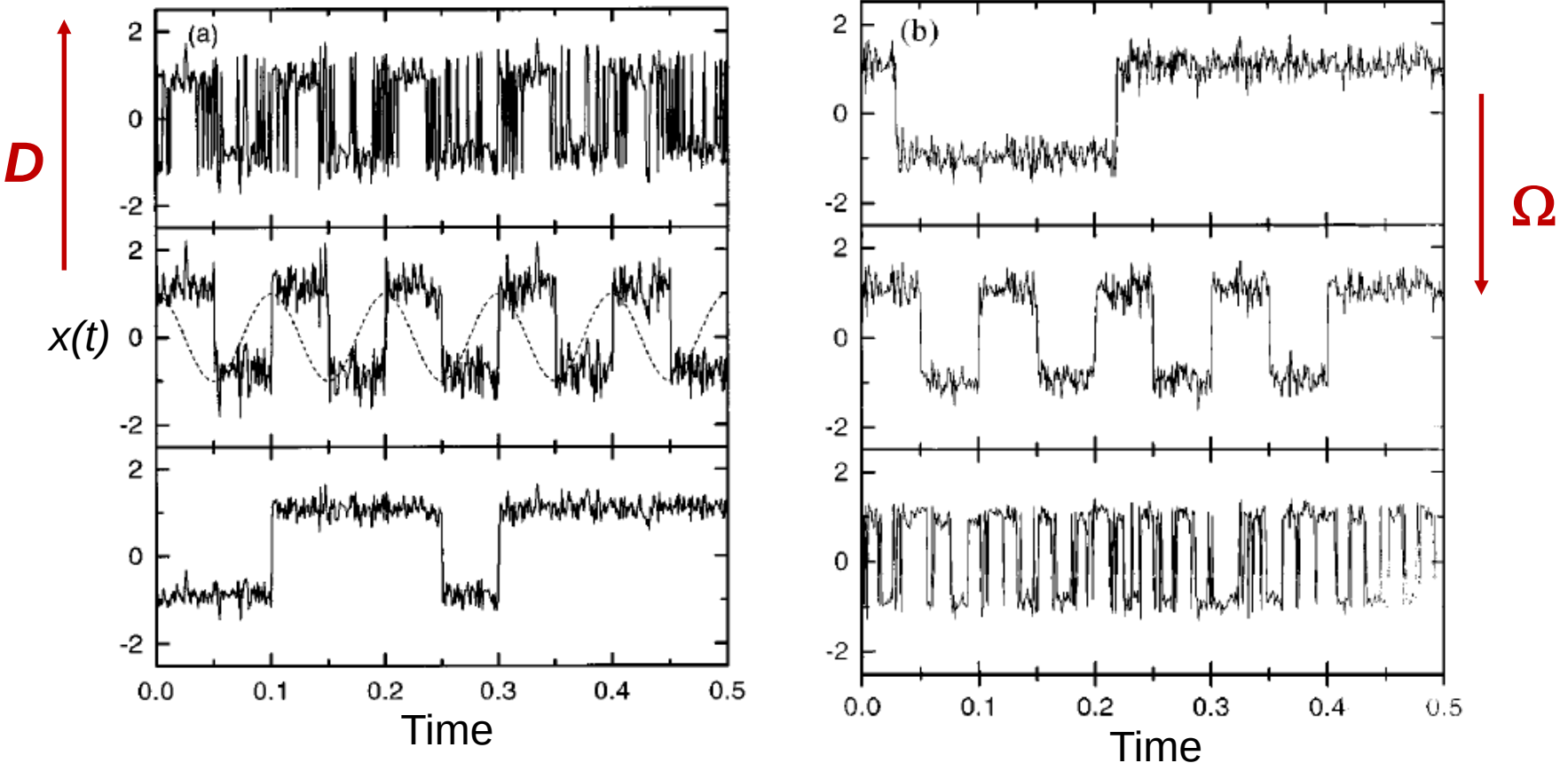
Bistable system with sinusoidal forcing and noise

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

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

Varying *D*; Ω constant

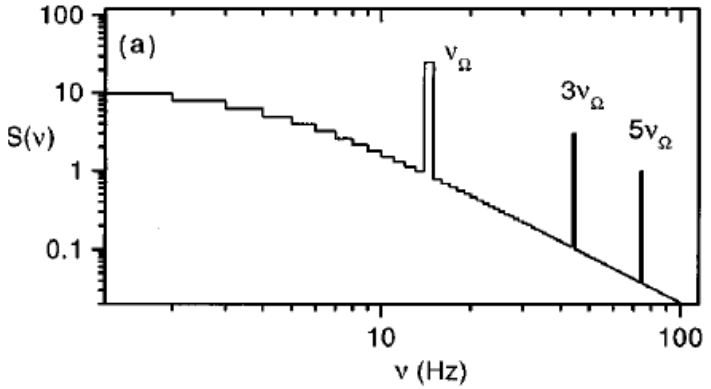
Varying Ω; *D* constant



Gammaitoni et al, Reviews of Modern Physics 70, 223 (1998).

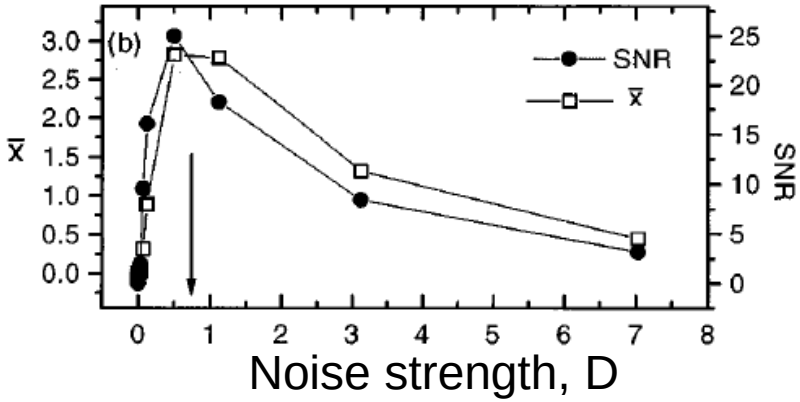
Quantification of stochastic resonance with spectral analysis

- *Phase-averaged* power spectral density: average over many realizations of the noise and average over the input initial phase φ.



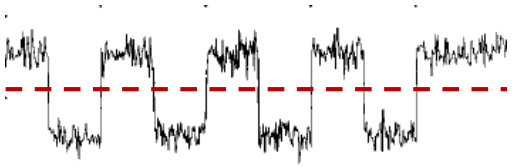
- Signal to noise ratio at Ω

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



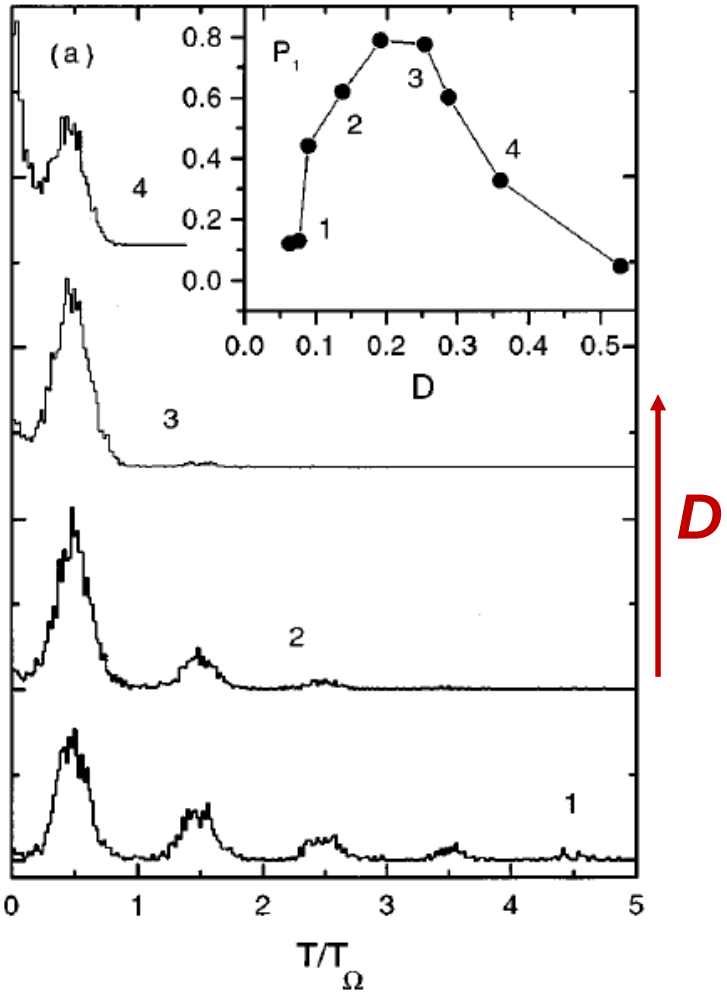
Gammaitoni et al, Reviews of Modern Physics 70, 223 (1998).

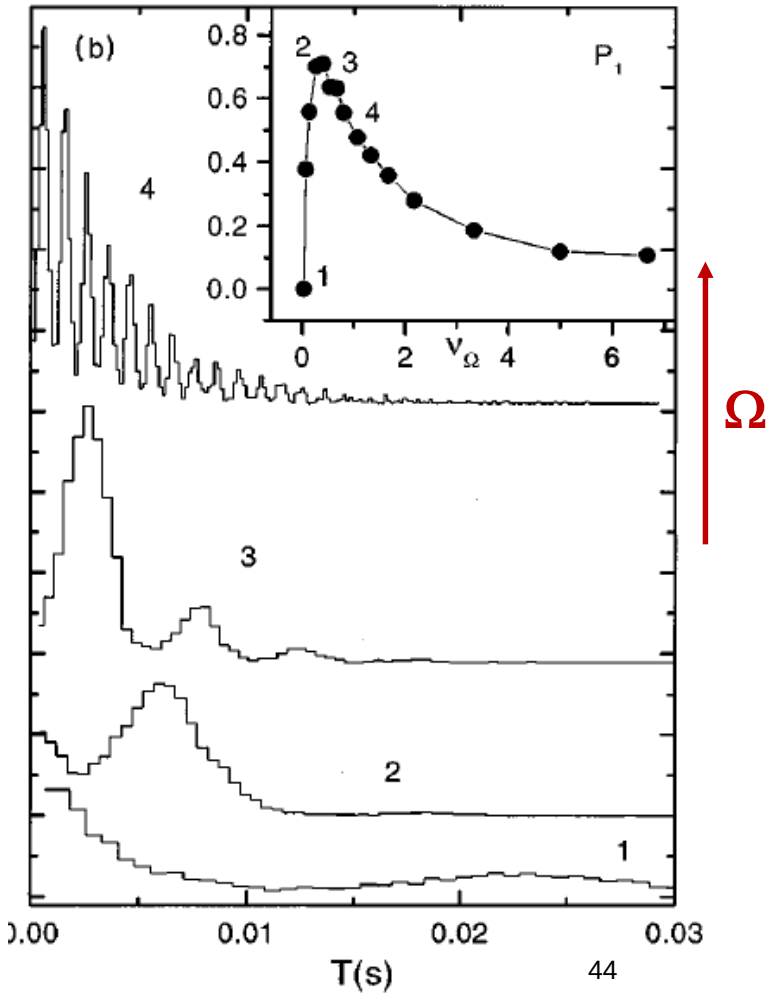
Quantification of stochastic resonance with the *distribution of residence times*



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

Number of switching times

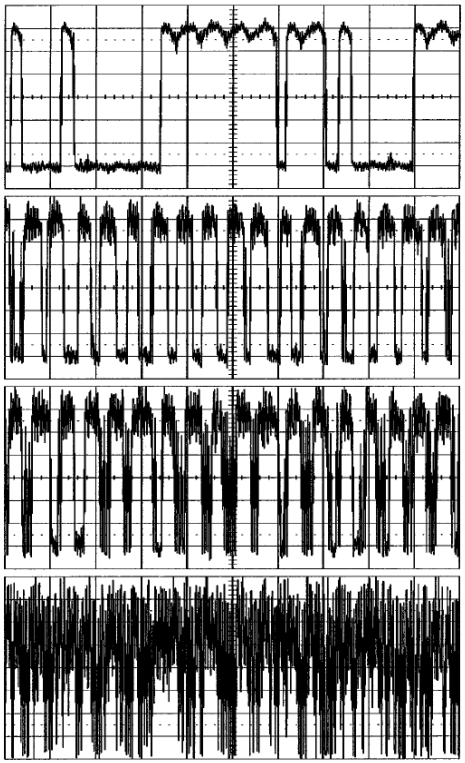




Gammaitoni et al, Reviews of Modern Physics 70, 223 (1998).

The SR phenomenon has been observed experimentally in many nonlinear systems. Two examples from in lasers:

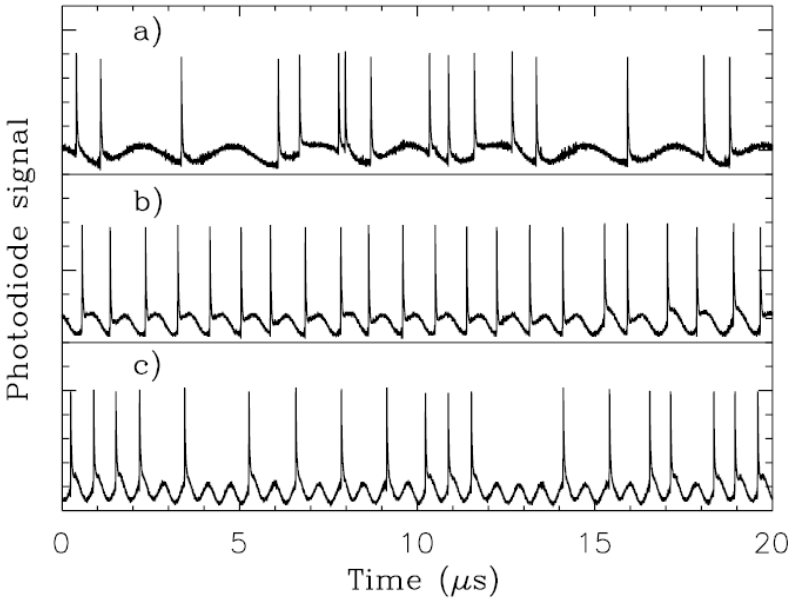
Temporal behavior of the polarized laser intensity in the presence of a current modulation, increasing the input noise from top to bottom.



Time (10 μs/division)

Giacomelli et al, Phys. Rev. Lett. 82, 675 (1999)

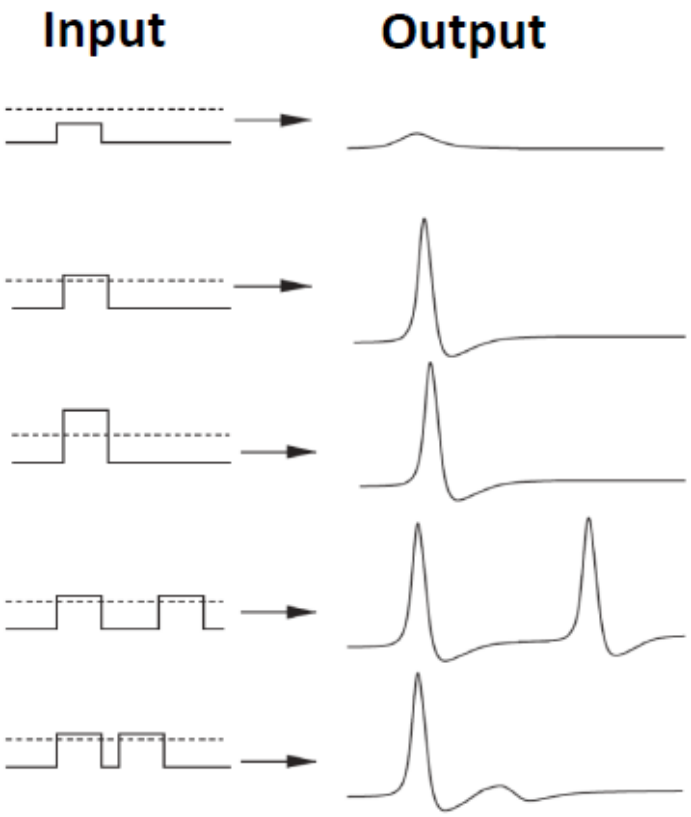
Time traces of the laser output intensity for a fixed noise level and increasing forcing frequency 0.4 MHz (a), 1.1 MHz (b), and 1.8 MHz



Marino et al, Phys. Rev. Lett. 88, 040601 (2002)

Excitable systems

- High nonlinear response to a perturbation:
  - Below a threshold, small response.
  - Above the threshold, rapid and large response.
- Refractory period: cannot support another event until a certain amount of time has passed



- Examples: neuronal spikes and cardiac beats.

B. Lindner et. al. Phys. Rep., 392:321, 2004.

Effect of noise in an excitable system: coherence resonance.

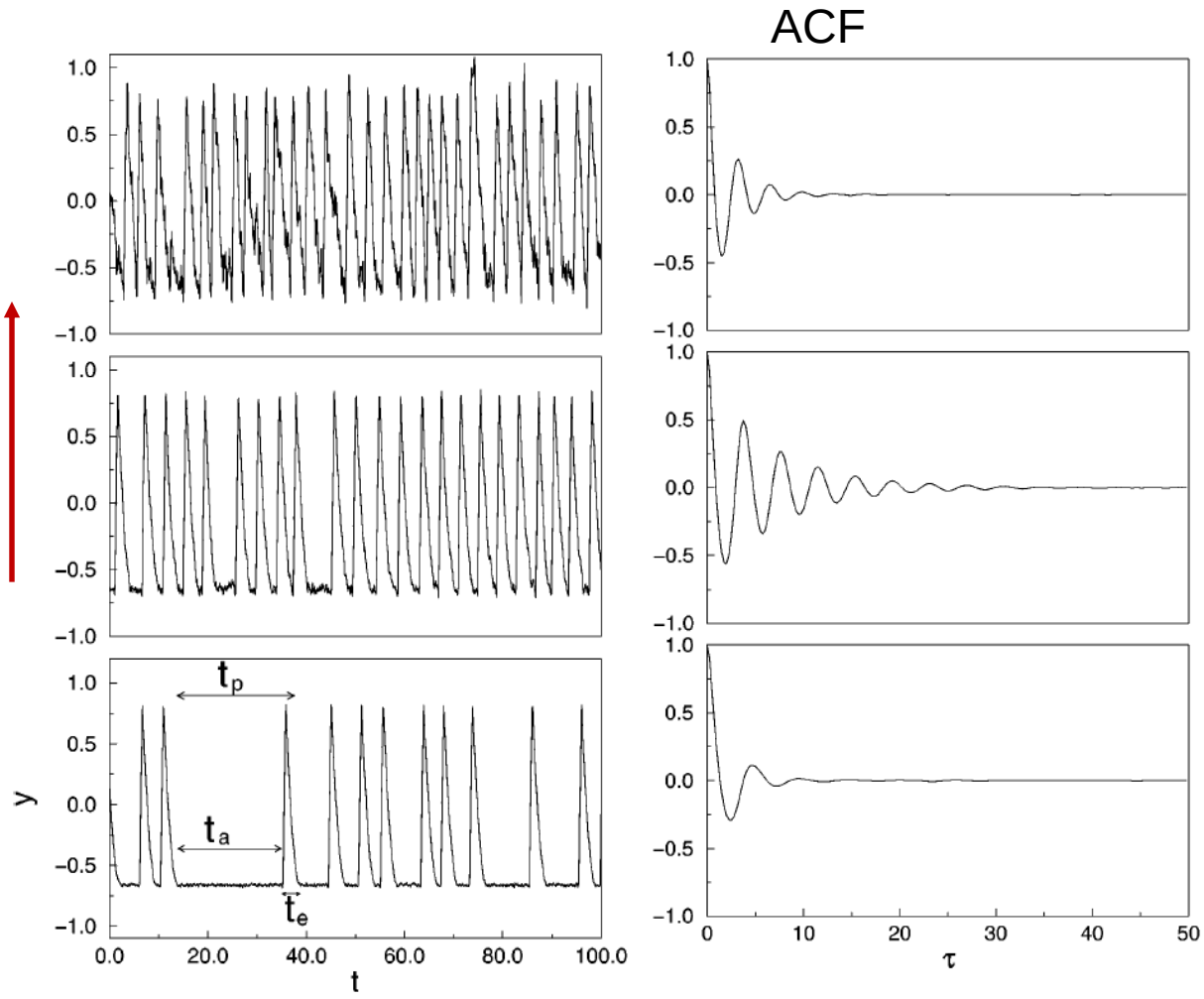
The system shows **maximum regularity (minimum disorder)** for an optimal level of noise.

Fitz Hugh–Nagumo model: classical example of an excitable system.

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

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

Excitable behavior for ε=0.01 and a =1.05.

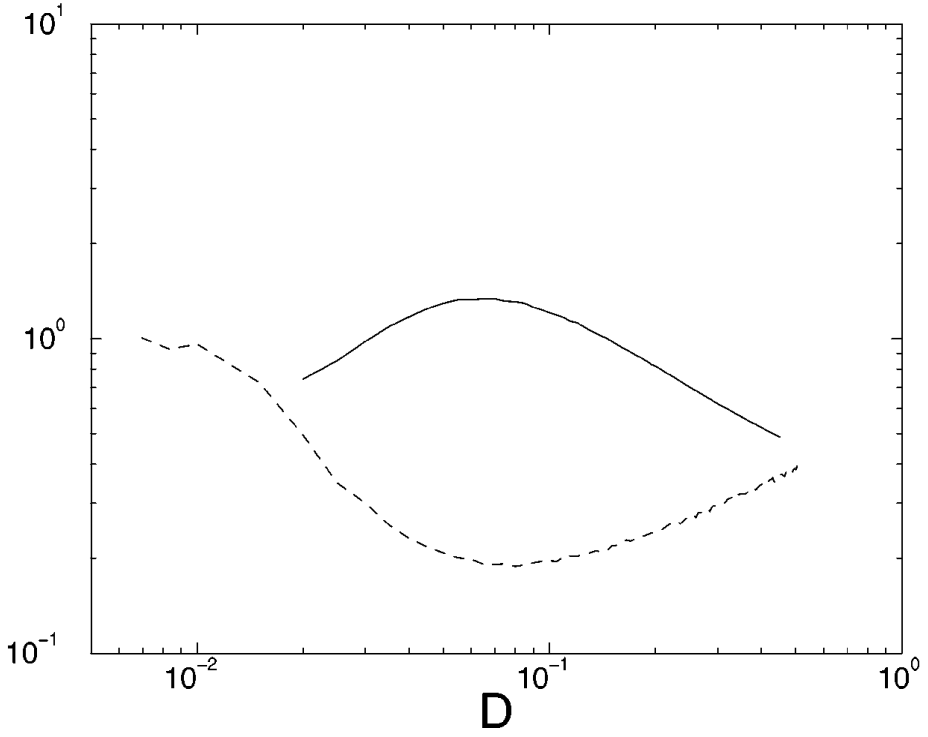


Pikovsky and Kurths, Phys. Rev. Lett. 78, 775 (1997)

How to quantify coherence resonance?

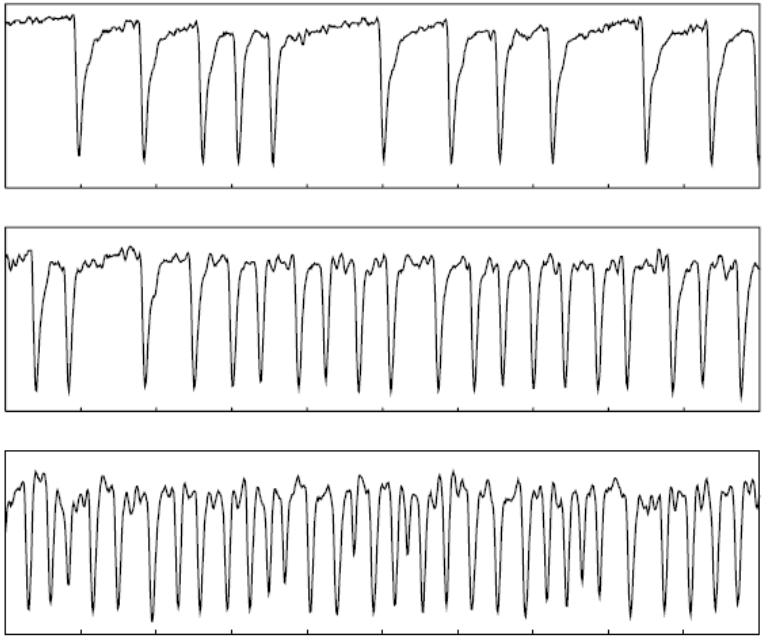
- The correlation time $\tau_c = \int_0^\infty C^2(t)\,dt$
- Coefficient of variation of the distribution of inter-spike-intervals $C_v = \sigma/\mu$

τ_c solid
C_v dashed

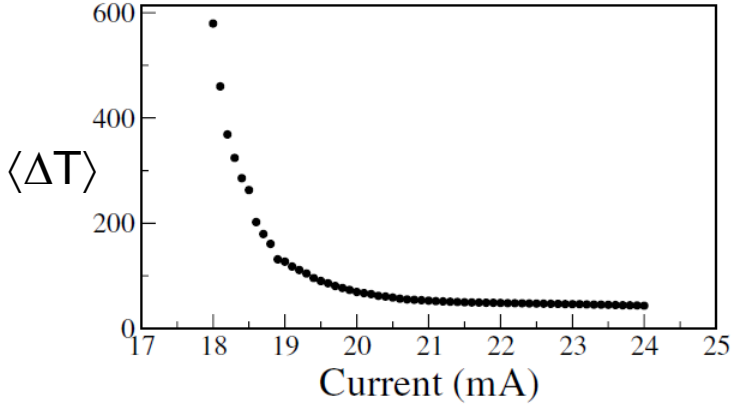


Pikovsky and Kurths, Phys. Rev. Lett. 78, 775 (1997)

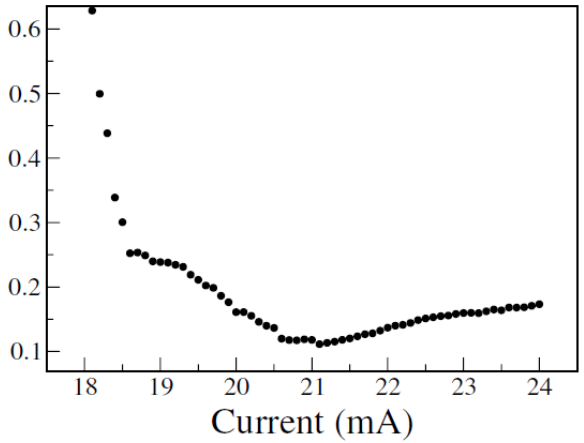
Coherence resonance has been observed in many excitable systems, including lasers.



Temporal behavior of the laser intensity for increasing input noise amplitude. The horizontal scale is 100 ns/div. The vertical scale is the same for the three plots.



$C_v = \sigma / \langle \Delta T \rangle$



Martinez Avila et al,
Phys. Rev. Lett. 93, 144101 (2004).

Tredicce et al, Phys. Rev. Lett. 84, 3298 (2000).

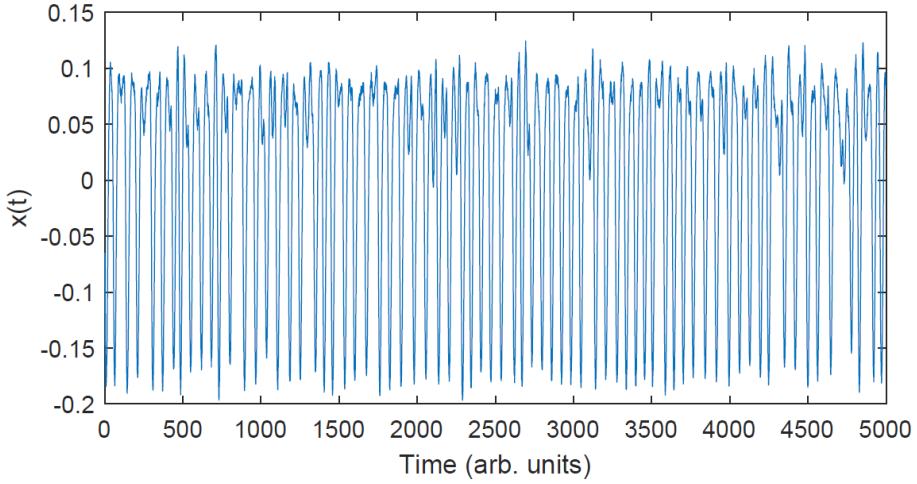
Methods of time series analysis

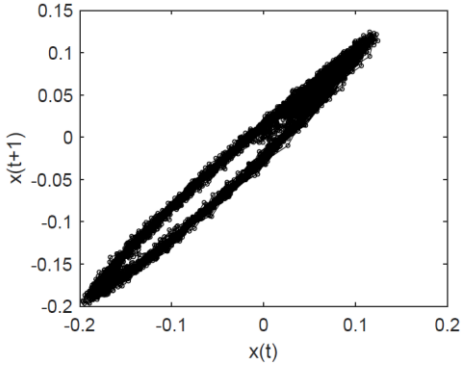
- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction, Lyapunov exponents, and fractal dimension
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
-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

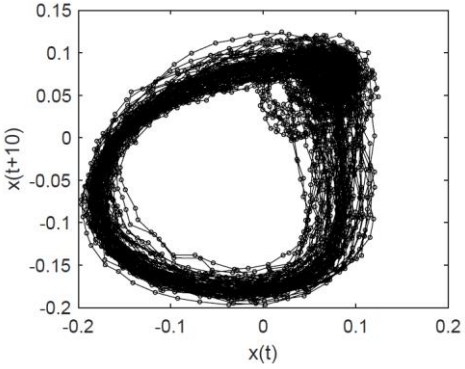
Why we want to reconstruct the phase space of a system from an observed (scalar) time series?

- Real systems are in general high-dimensional and we can only measure a few (hopefully relevant) variables.
- Models are too complex and have many parameters: reconstructing the phase space may allow to understand the effect of different parameters.

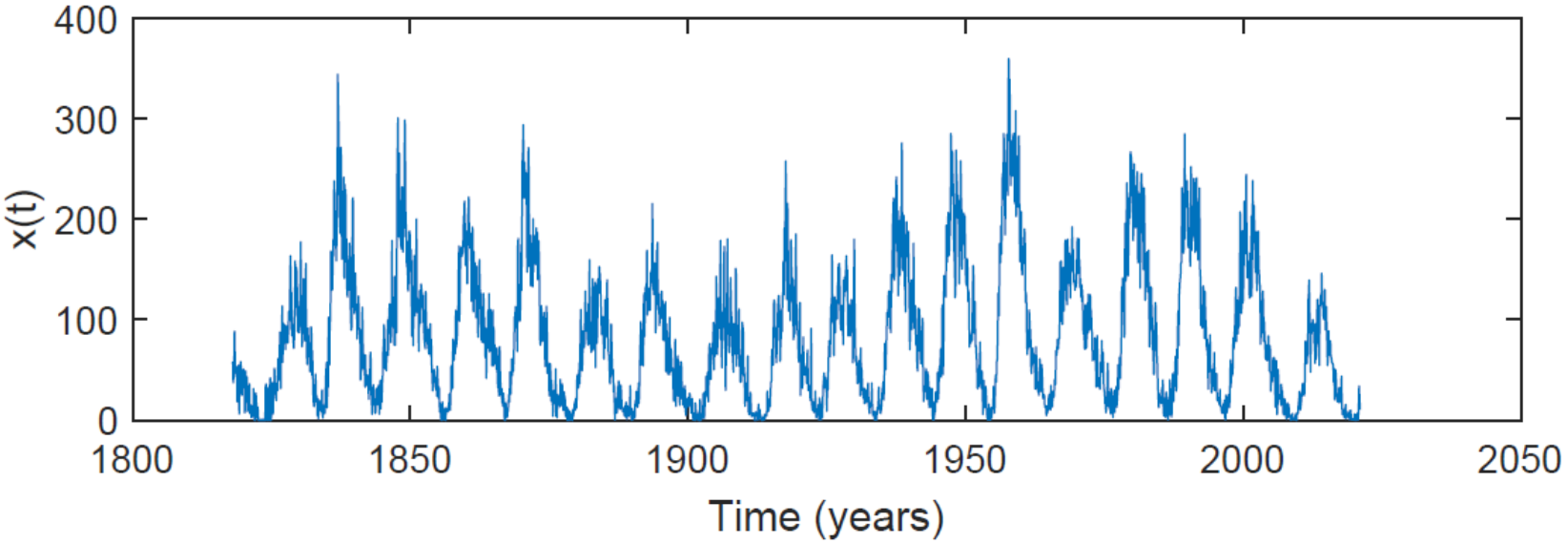
Example: the intensity emitted by a diode laser with optical feedback

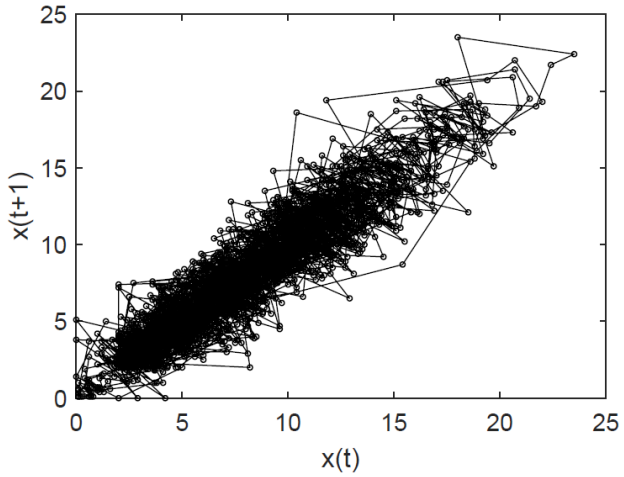


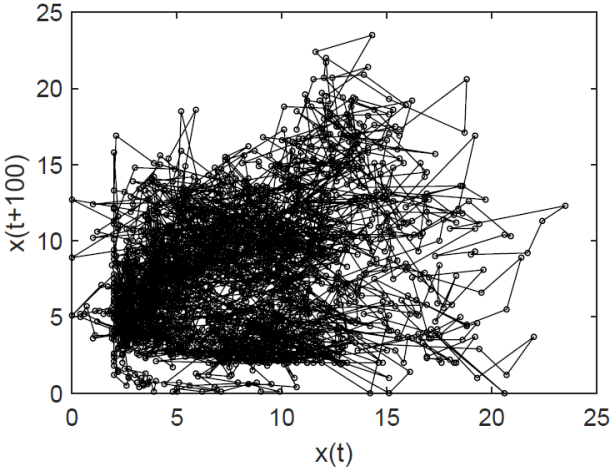




A popular time series: monthly mean total sunspot number







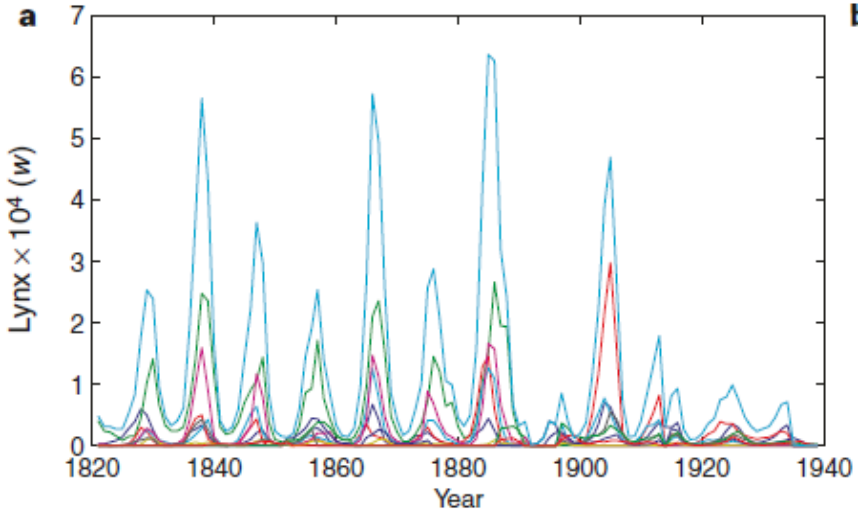
http://sidc.oma.be/silso/infosnmtot

Complex dynamics and phase synchronization in spatially extended ecological systems

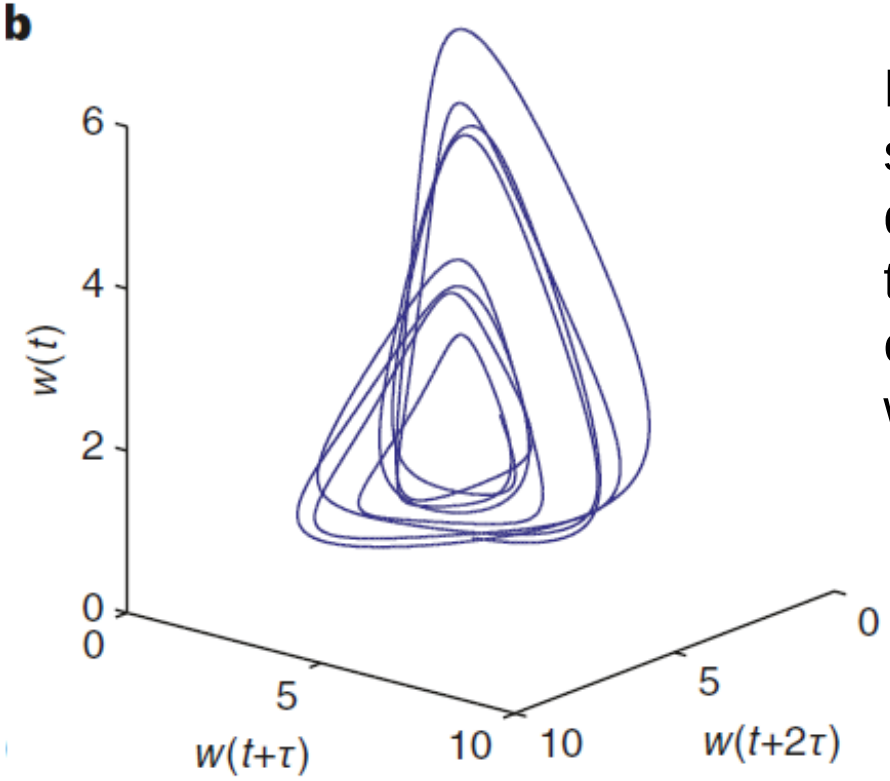
Bernd Blasius, Amit Huppert & Lewi Stone

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



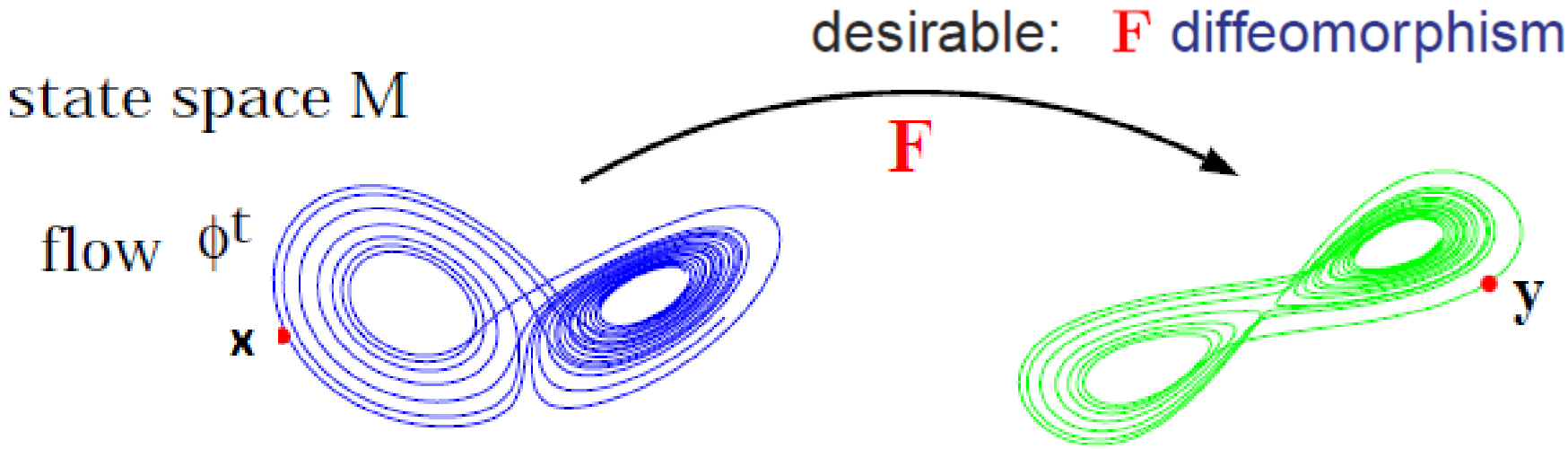
Reconstructed attractor obtained by spatially averaging all regional lynx data and embedding the resulting time-series w(t) using lagged coordinates, w(t) versus w(t+τ) versus w(t+6τ) after filtering and interpolating.

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

Observed time series S = {s(1), s(2), … s(t) … }

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

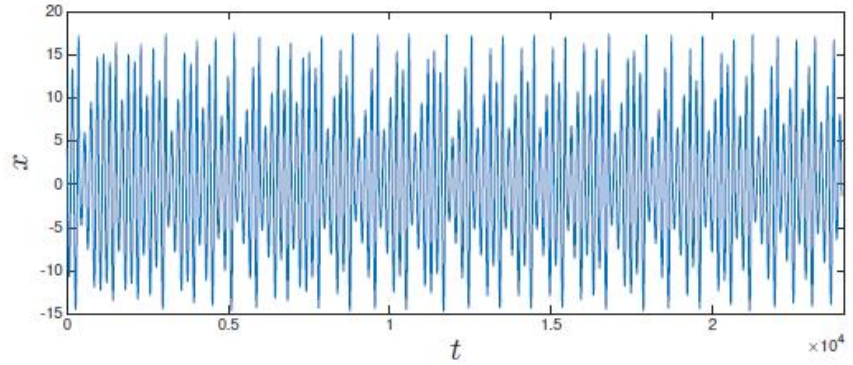
$$y(t) = \left(s(t), s(t+\tau), \ldots, s(t+(d-1)\tau)\right)$$



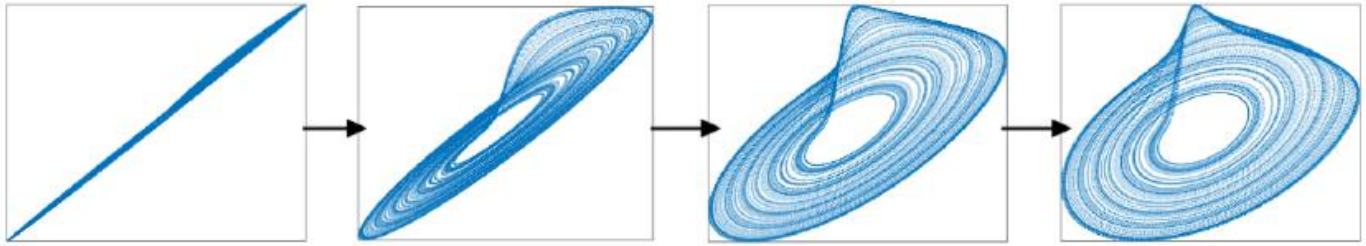
Adapted from U. Parlitz (Gottingen)

F.Takens, Lecture Notes in Mathematics 366 (1981)
Sauer et al., J. Stat. Phys. 65 (1991) 579

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 chose the embedding parameters (lag τ, dimension ***d***)

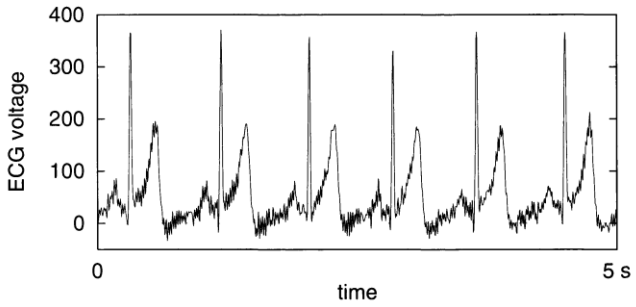
Bradley and Kantz, CHAOS 25, 097610 (2015)

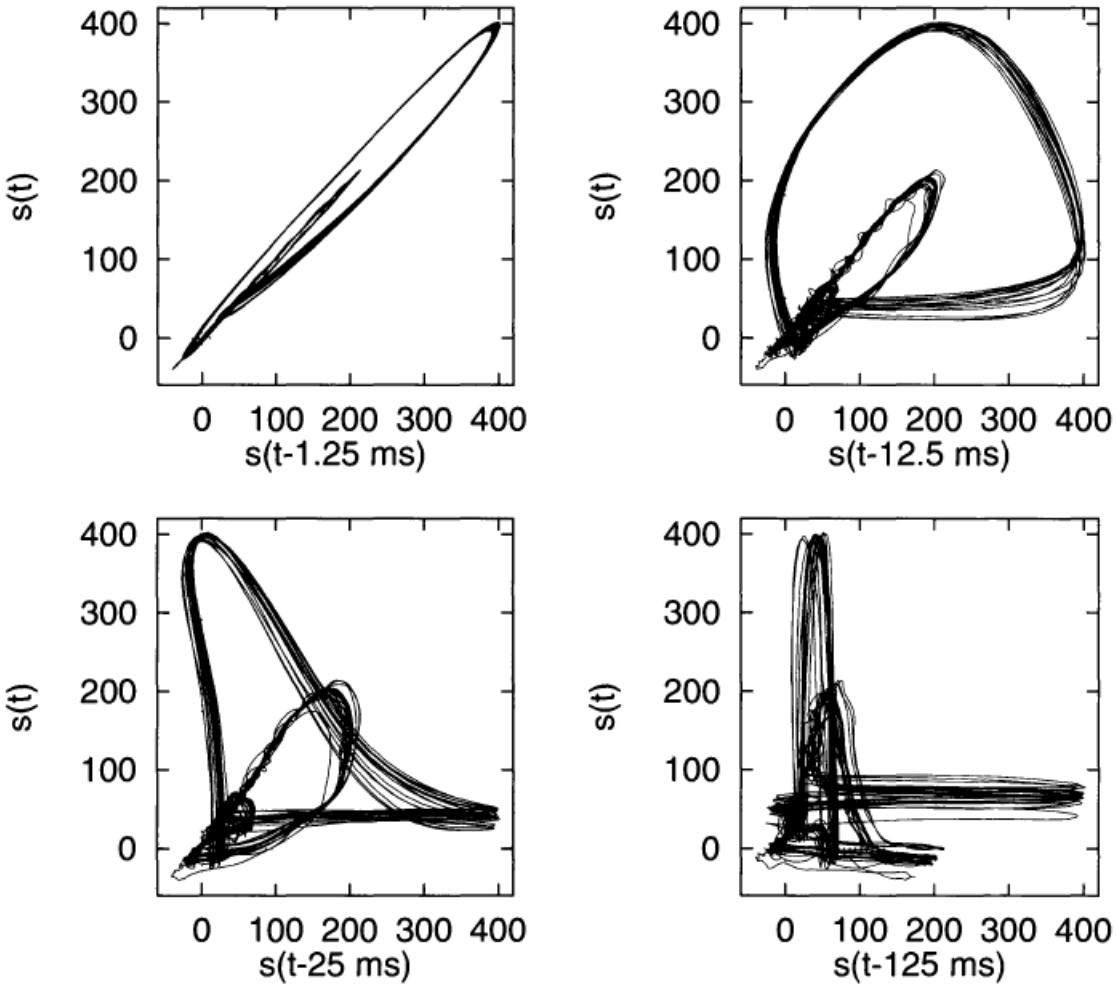
How to chose the lag τ and the dimension *d*

- τ 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.**

Example: 2D representation of a human ECG signal

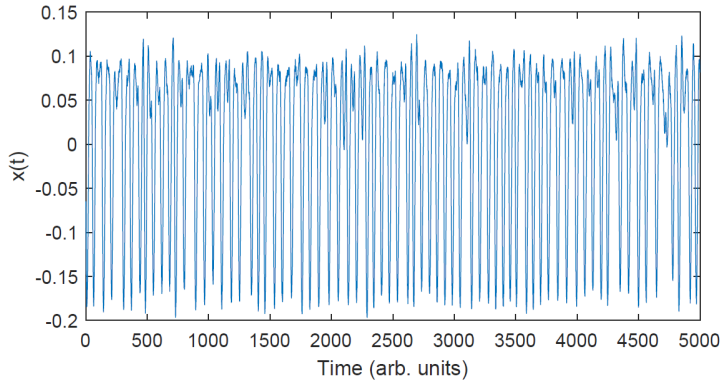


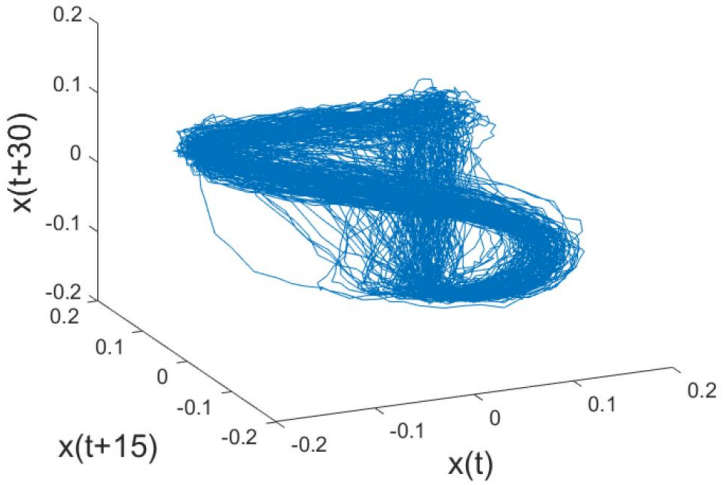


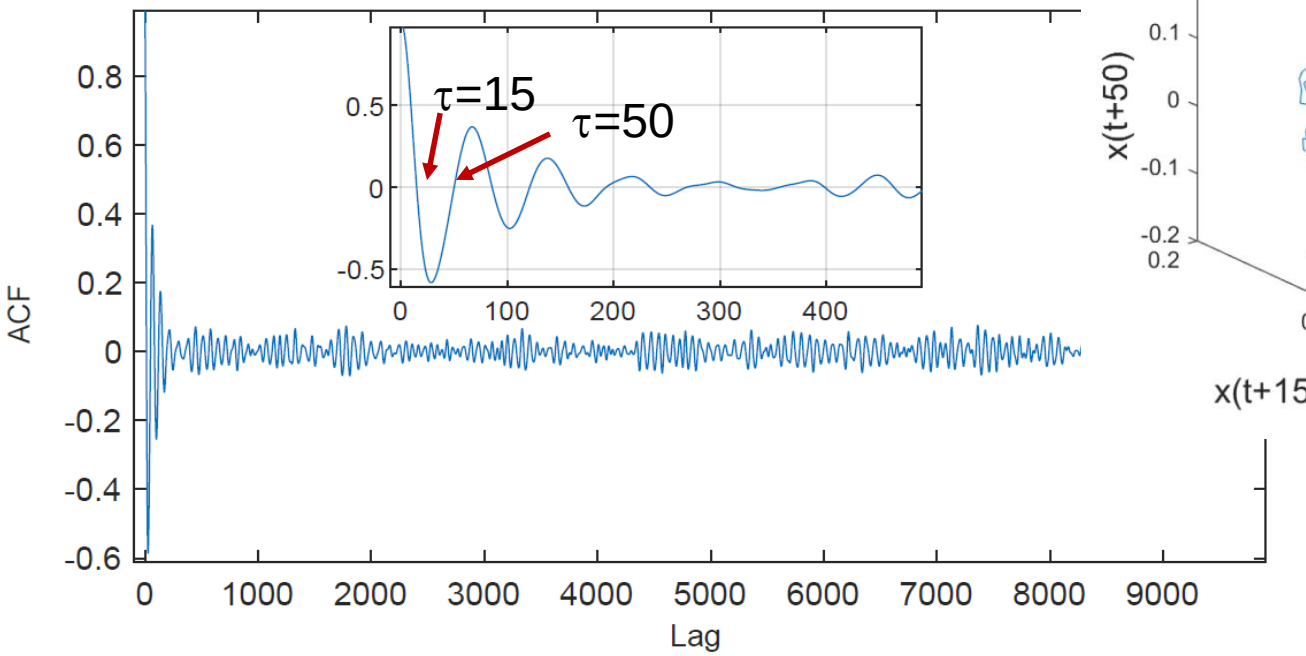
H. Kantz and T. Schreiber, Cambridge University Press 2003

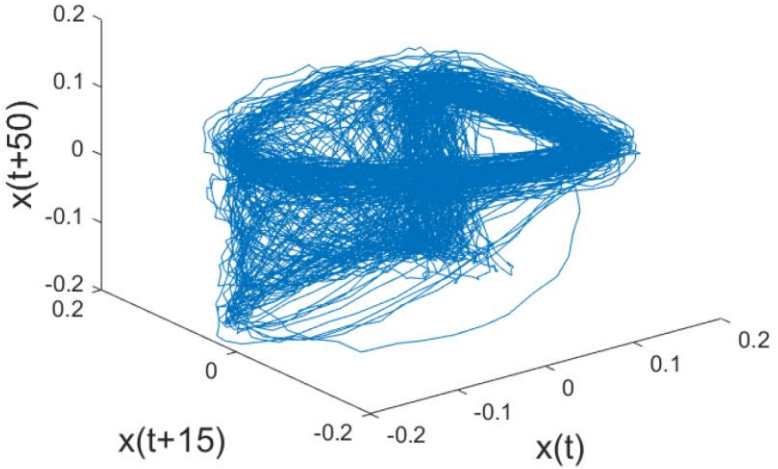
Hands-on exercise 5:

5.1 Reconstruct the attractor from the laser intensity time series



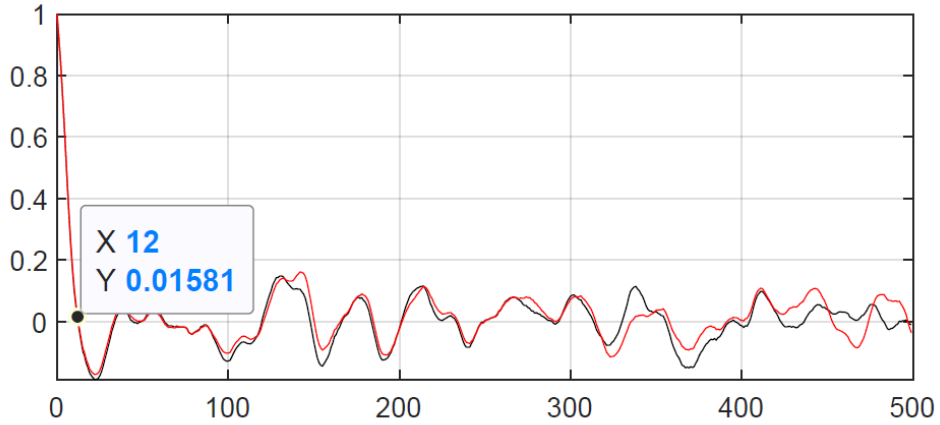


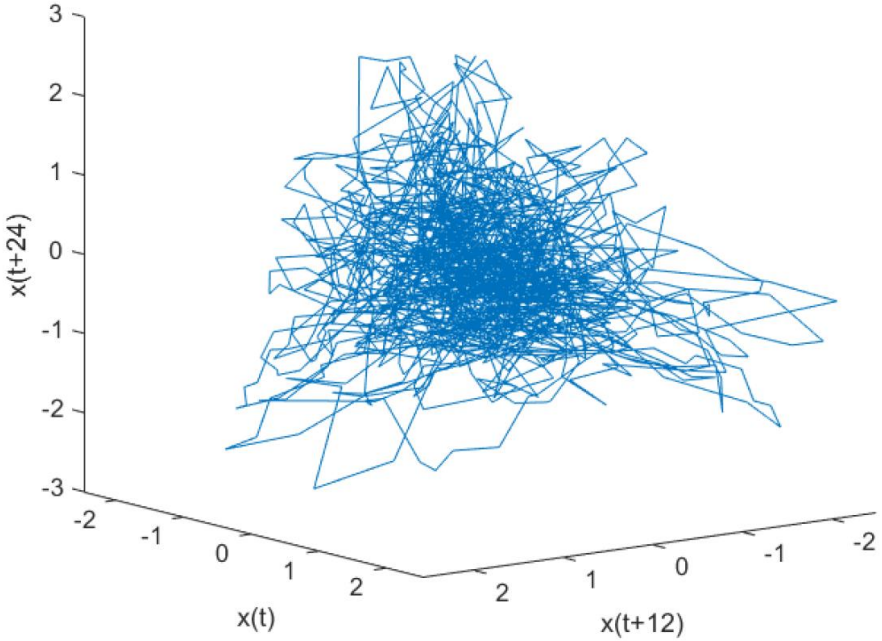




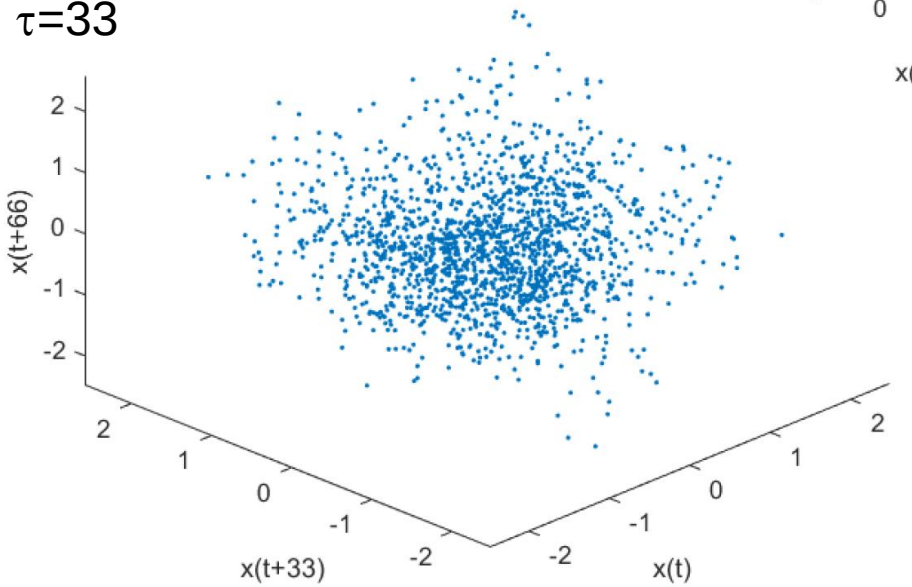
Hands-on exercise 5:

5.2 Reconstruct the attractor of El Niño index





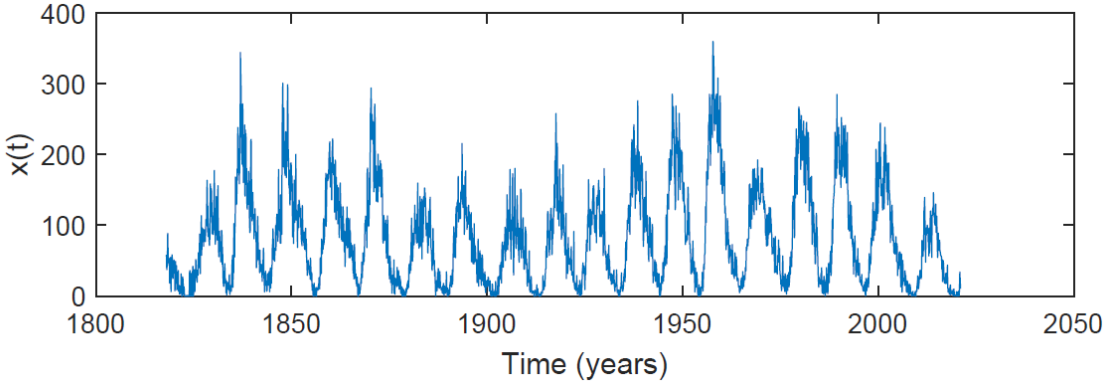
Next 0: τ=33

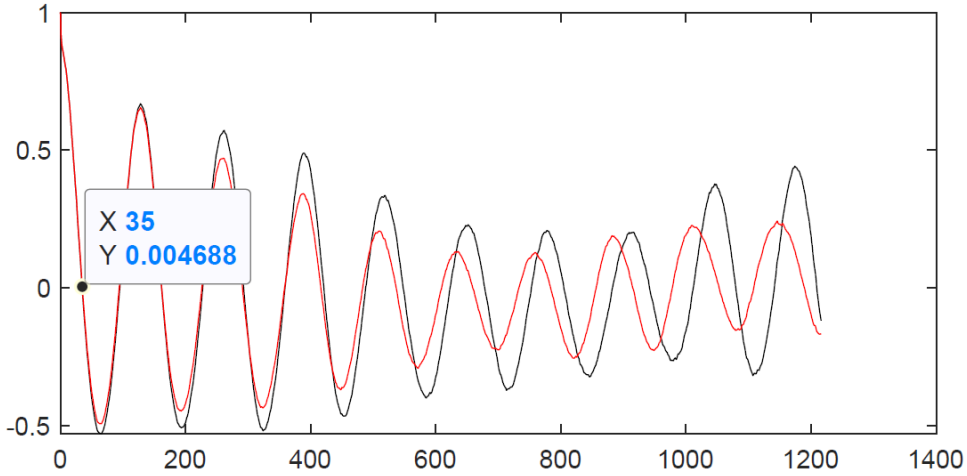


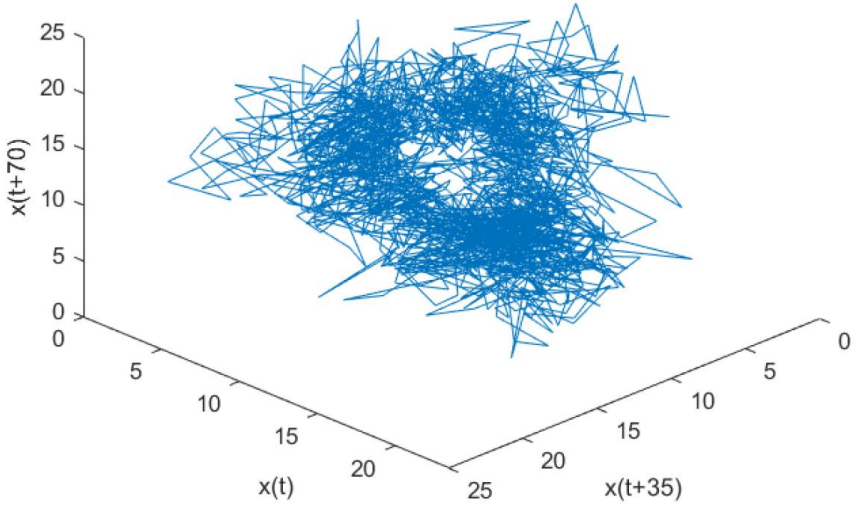
Hands-on exercise 5:

5.3 Reconstruct the attractor of the monthly mean total sunspot number.

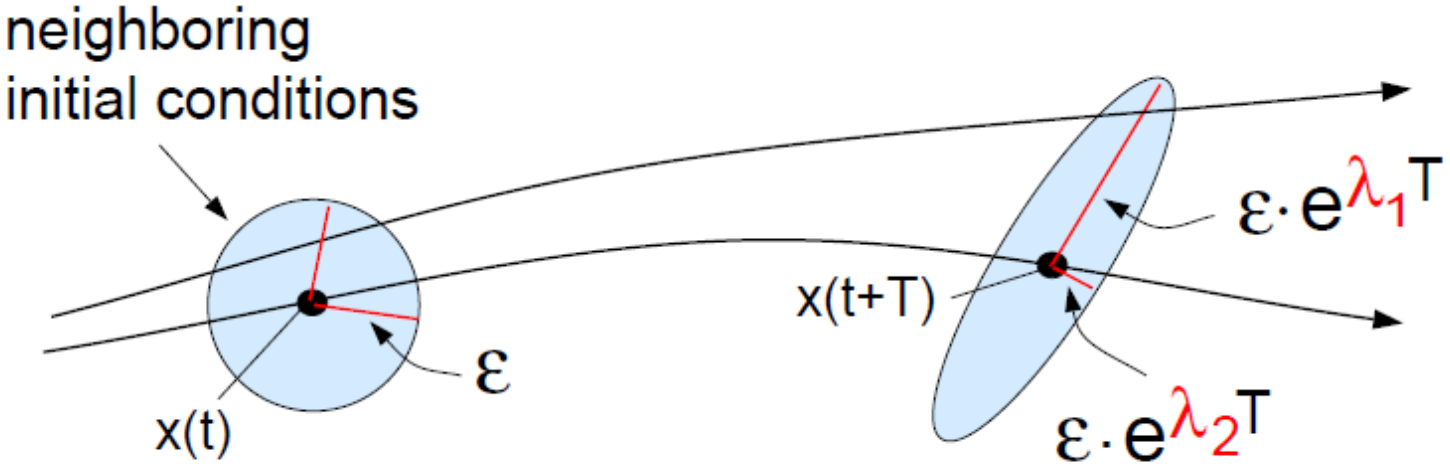
- Down load data set from http://sidc.oma.be/silso/infosnmtot
- Plot and analyze the 4th column of SN_m_tot_V2.0.txt







Lyapunov exponents: measure how fast neighboring trajectories diverge.



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

Adapted from U. Parlitz (Gottingen)

Steps to compute the maximum LE

- Initial distance $\delta_{\mathrm{I}} = |s_i - s_j|$
- Final distance $\delta_{\mathrm{F}} = |s_{i+T} - s_{j+T}|$
- Local *exponential* grow $\lambda^*_{\mathrm{local}} = \frac{1}{T}\log(\delta_{\mathrm{F}}/\delta_{\mathrm{I}})$
- The rate of grow is averaged over the attractor, which gives λ_{max}

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

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

On the interpretation of the maximum Lyapunov exponent: a word of warning!

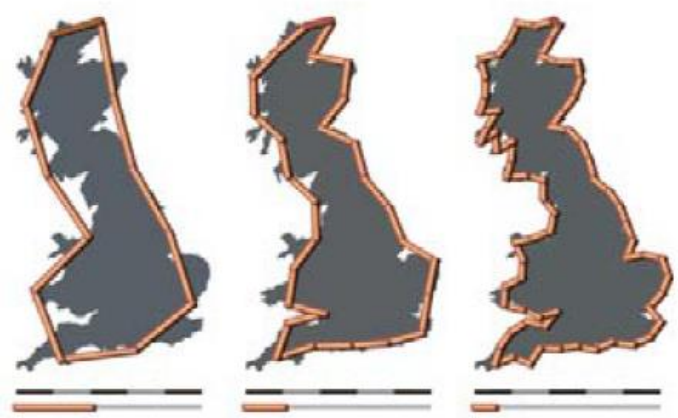
- The algorithm returns λ in the fastest expansion direction.
- The algorithm always returns a positive number!
- This is a main problem when computing the LE of noisy data.

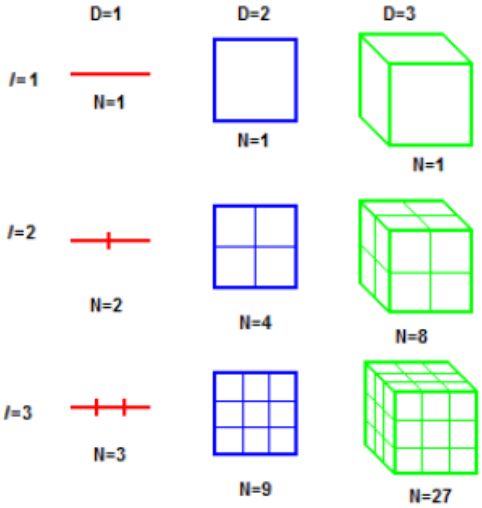
Every time series analysis algorithm returns a number of any time series. But is it useful?

Further reading:
F. Mitschke and M. Damming, Chaos vs. noise in experimental data, Int. J. Bif. Chaos 3, 693 (1993)

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.

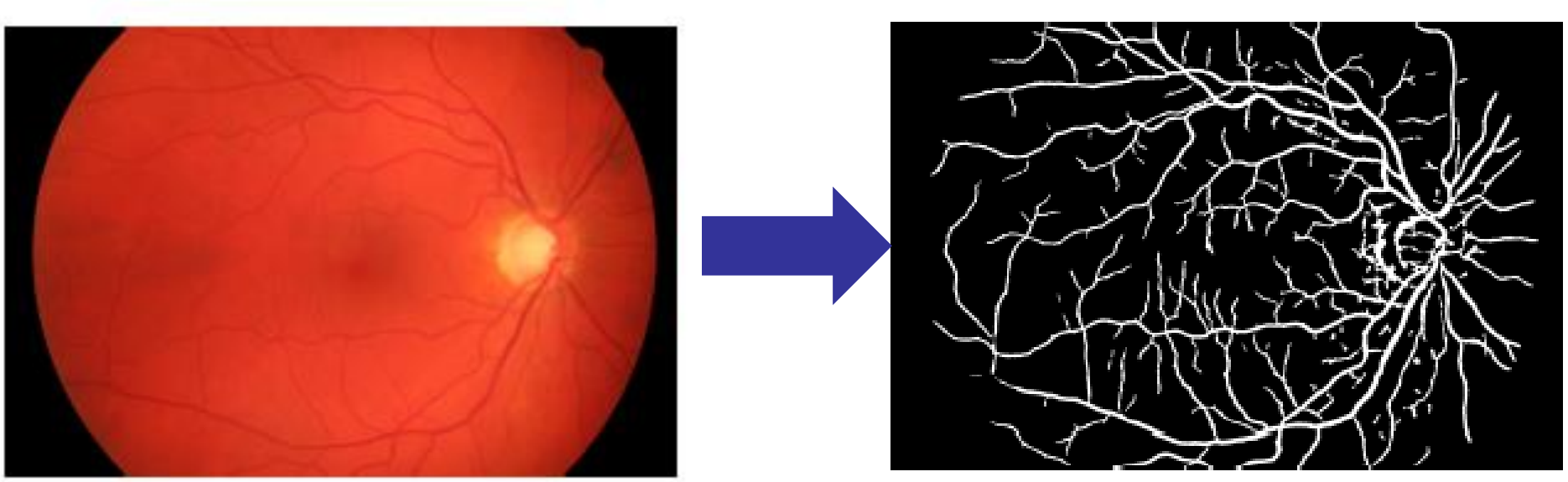




- Fractal dimension: $N \propto \epsilon^{-D} \rightarrow D_0 = \lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}}$

Source: Wikipedia

Application of fractal analysis



The fractal dimension of the blood vessels in the normal human retina is about 1.7 while it tends to increase with the level of diabetic retinopathy.

P. Amil et al., PLoS ONE 14, e0220132 (2019).

Grassberger-Procaccia correlation dimension algorithm

- Another very popular method for detecting chaos in real-world data.
- Fractal dimension (box counting dimension): $D_0 = \lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}}$
- **Problem**: for time-series analysis, D_0 does not distinguish between frequently and unfrequently visited boxes.
- An alternative: the *correlation dimension*, based on calculating the number of pairs of points with distance between them < ε.

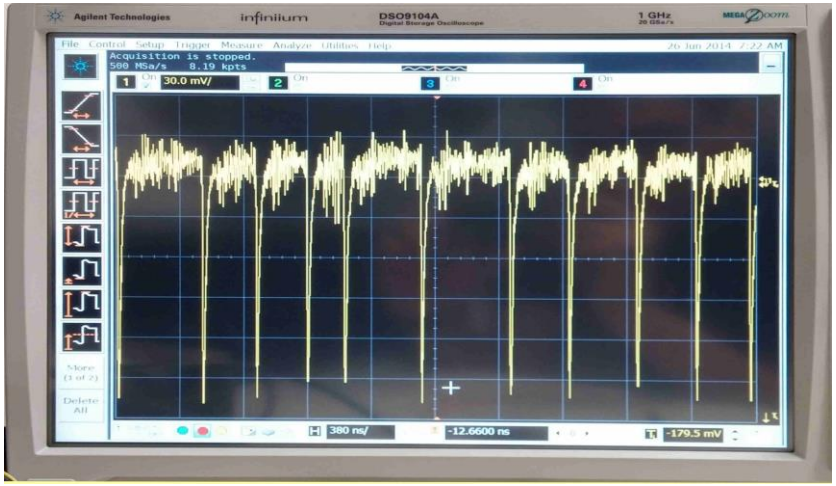
Further reading:
P. Grassberger and I. Procaccia, "Measuring the Strangeness of Strange Attractors". Physica D vol. 9, pp.189, 1983.
L. S. Liebovitch and T. Toth, "A fast algorithm to determine fractal dimensions by box counting," Physics Letters A, vol. 141, pp. 386, 1989.

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 dimension
- 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

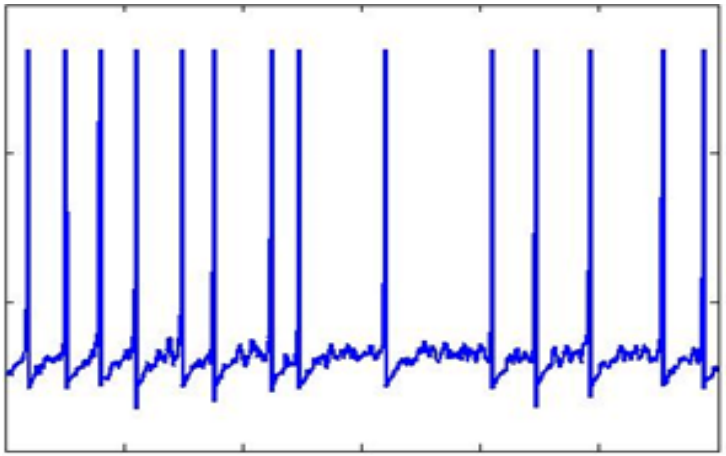
Can lasers mimic real neurons?

Laser spikes



Time (μs)

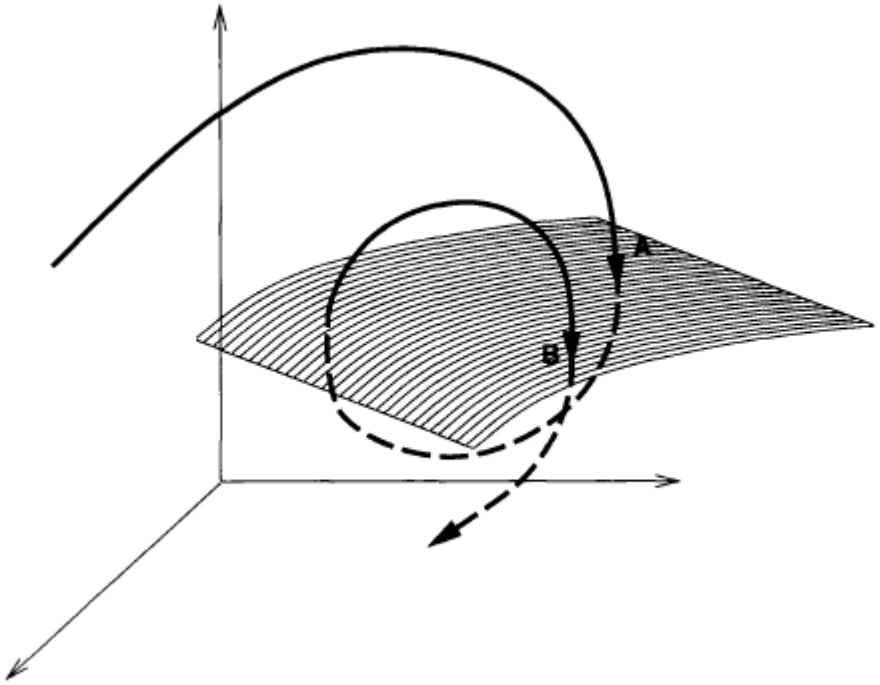
Neuronal spikes



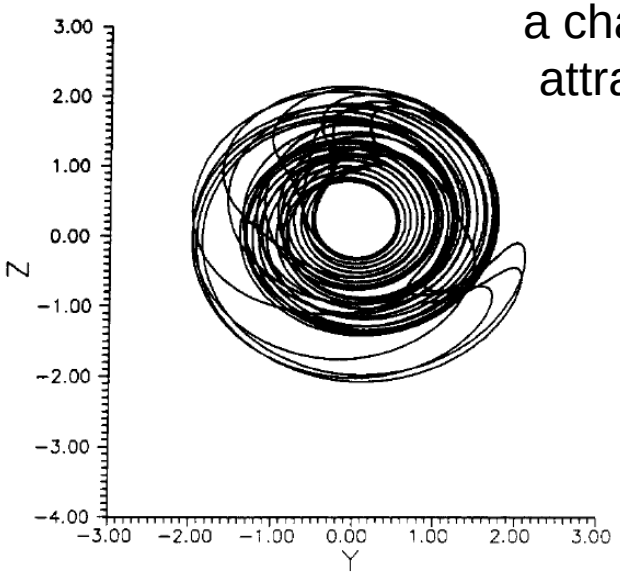
Time (ms)

- Are there statistical similarities?
- A popular technique: define spike times via “threshold crossings” and analyze the statistical properties of the inter-spike-intervals (ISIs). Data compression!
- Results should be robust to small variations of the threshold.

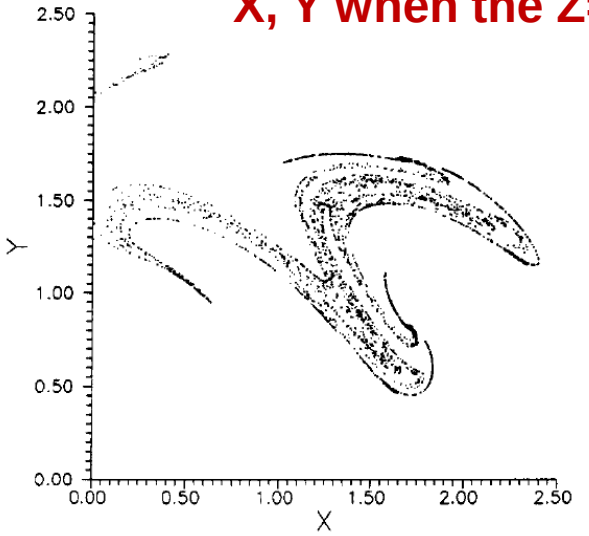
Another way to define events for “compressing” the data: the Poincare section



Example of a chaotic attractor



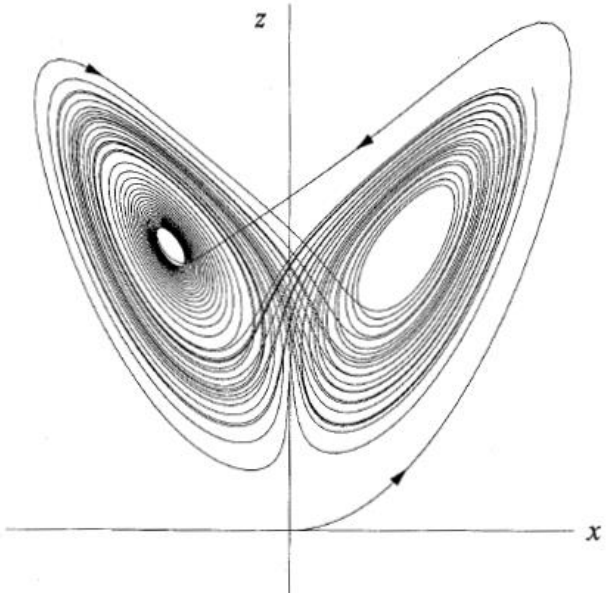
X, Y when the Z=0

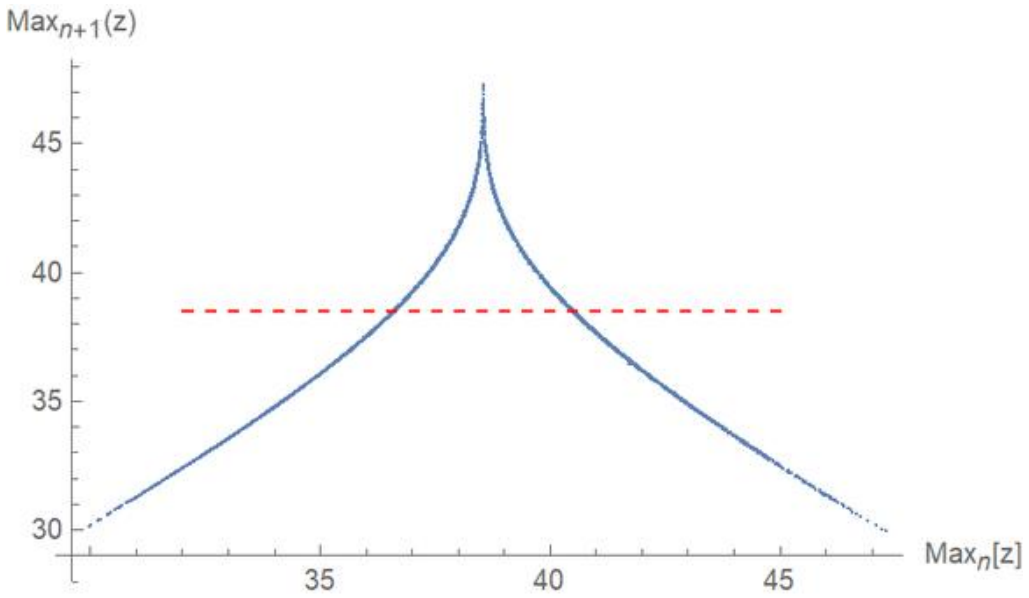


H. Kantz and T. Schreiber, Cambridge University Press 2003
C. Masoller et al, Phys. Lett. A 167, 185 (1992).

A third way to define events: analyze only the extreme values

Lorenz created a Poincaré plot by plotting the relative maximum value of the z value, against the previous one. He found that when the maximum z value is above a certain value, the system will switch to the other lobe.





E. N. Lorenz, "Deterministic nonperiodic flow". Journal of the Atmospheric Sciences 20, 130 (1963).

Back to the comparison of neuronal and optical spikes: inspection of the distribution of inter-spike-intervals

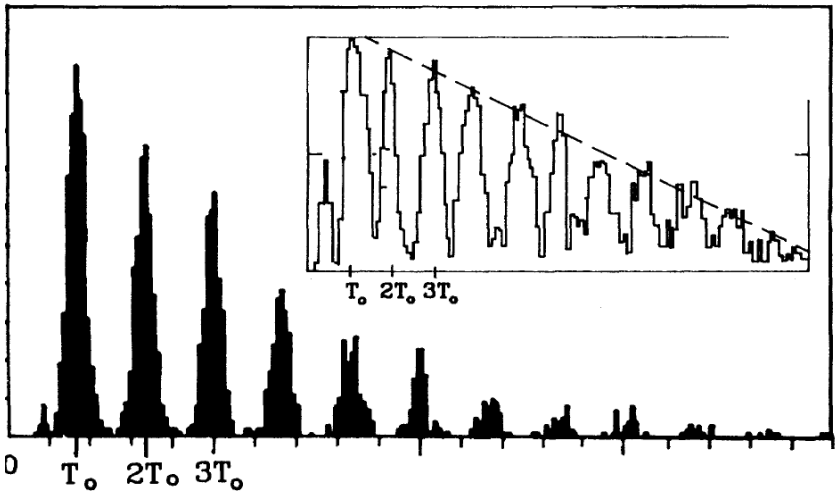
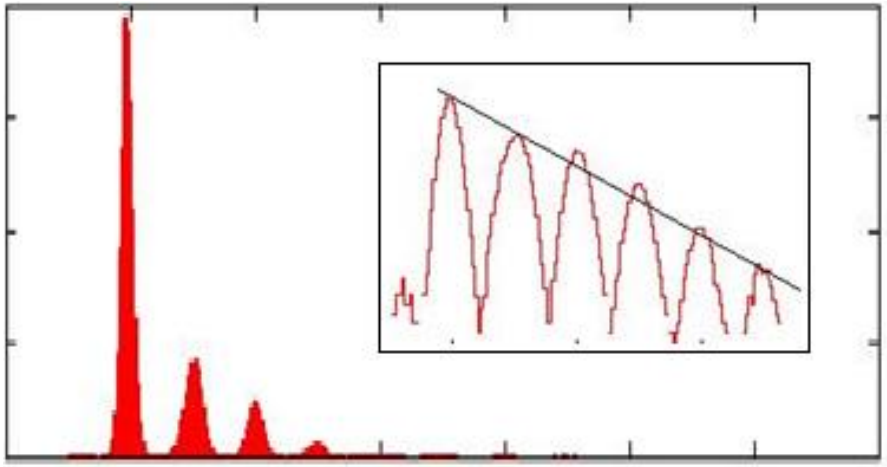


FIG. 1. (a) An experimental ISIH obtained from a single auditory nerve fiber of a squirrel monkey with a sinusoidal 80-dB sound-pressure-level stimulus of period $T_0 = 1.66$ ms applied at the ear. Note the modes at integer multiples of T_0. Inset:



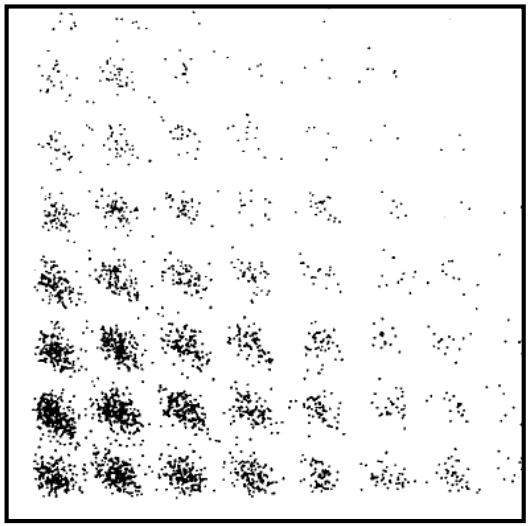
Experimental data when the laser current is modulated with a sinusoidal signal of period T_0.

A. Longtin et al Phys. Rev. Lett. (1991)

A. Aragoneses et al Optics Express (2014)

Return maps of the inter-spike-intervals

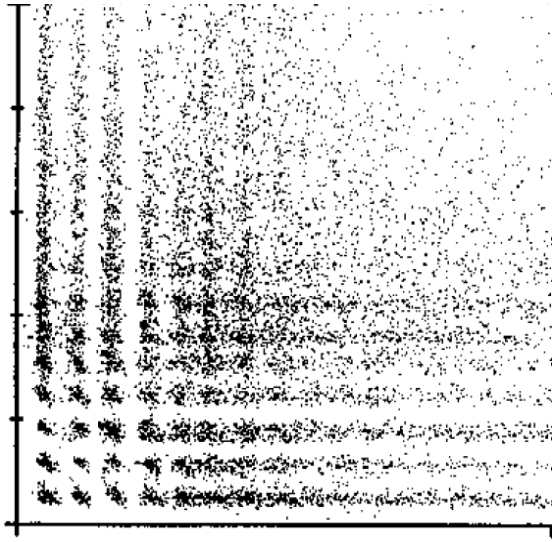
Neuron



ΔT_{i+1}

ΔT_i

Laser



HOW TO INDENTIFY SIMILAR TEMPORAL ORDER?
MORE/LESS EXPRESSED PATTERNS?

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

M. Giudici et al Phys. Rev. E (1997)
A. Aragoneses et al Optics Express (2014)

Symbolic analysis

- The time series $\{x_1, x_2, x_3, \ldots\}$ is transformed (using an appropriated **rule**) into a sequence of symbols $\{s_1, s_2, \ldots\}$
- Symbols are 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 $\mathbf{2^D}$
- More thresholds allow for more letters in the "alphabet" (and more words in the dictionary). Example:

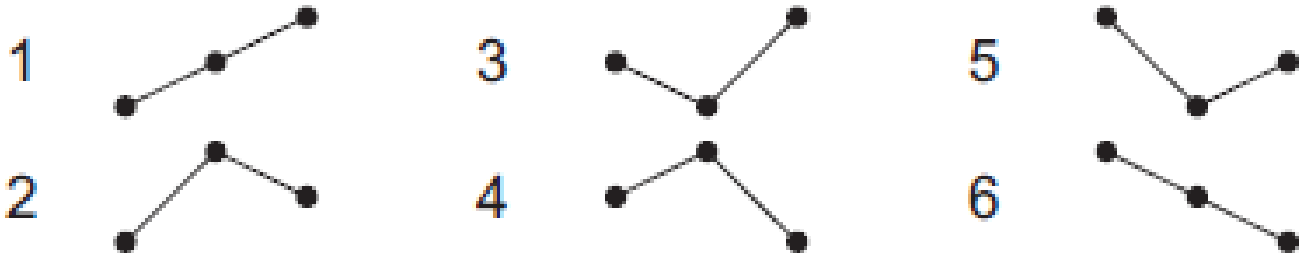
  if $x_i > x_{th1} \Rightarrow s_i = 0$;

  else if $x_i < x_{th2} \Rightarrow s_i = 2$;

  else $(x_{th2} < x_i < x_{th1}) \Rightarrow s_i = 1$.

Ordinal analysis: threshold-less method to define symbols

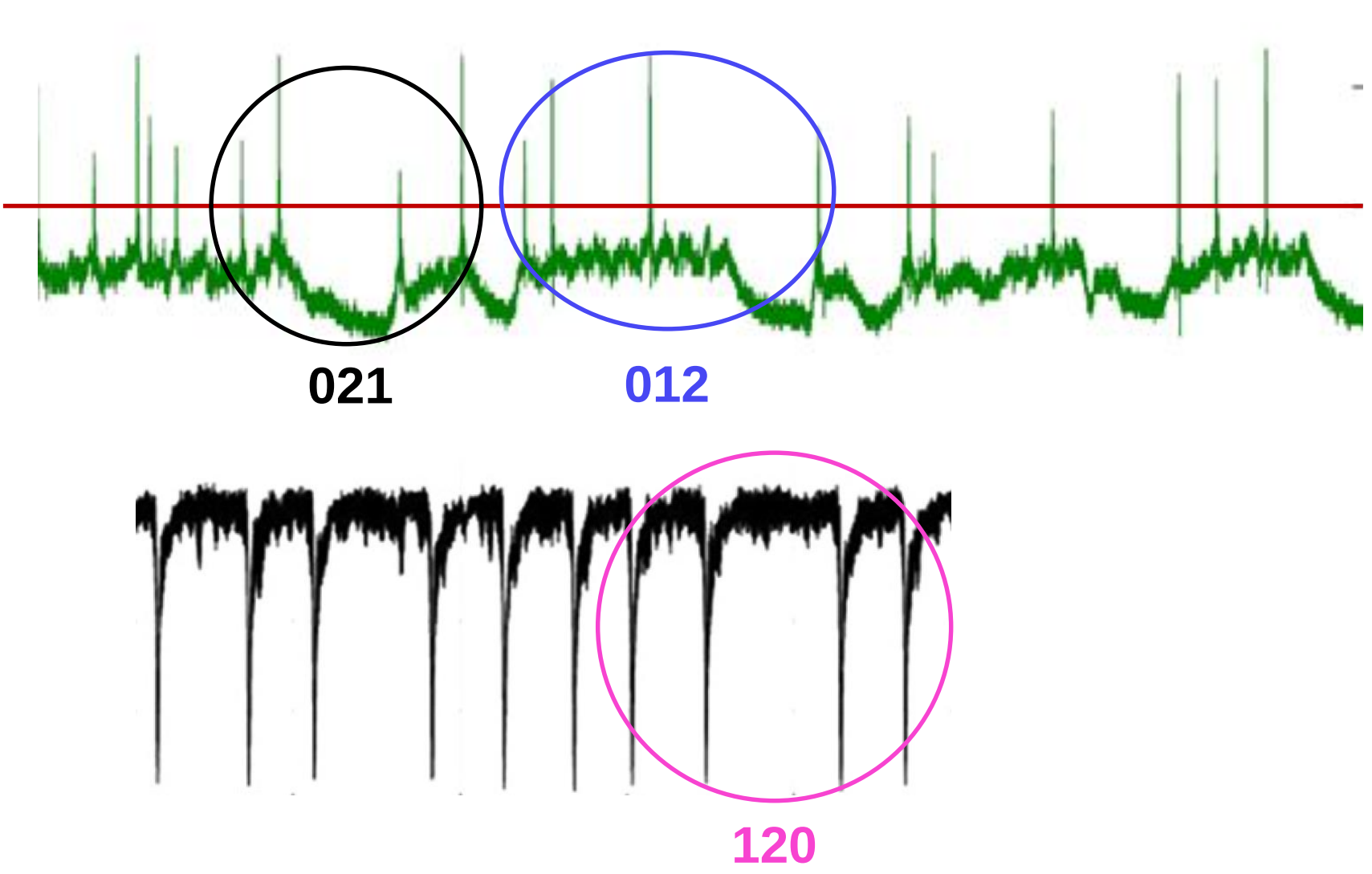
- Consider a time series $x(t)=\{\ldots x_i, x_{i+1}, x_{i+2}, \ldots\}$
- Which are the possible order relations among three consecutive data points?



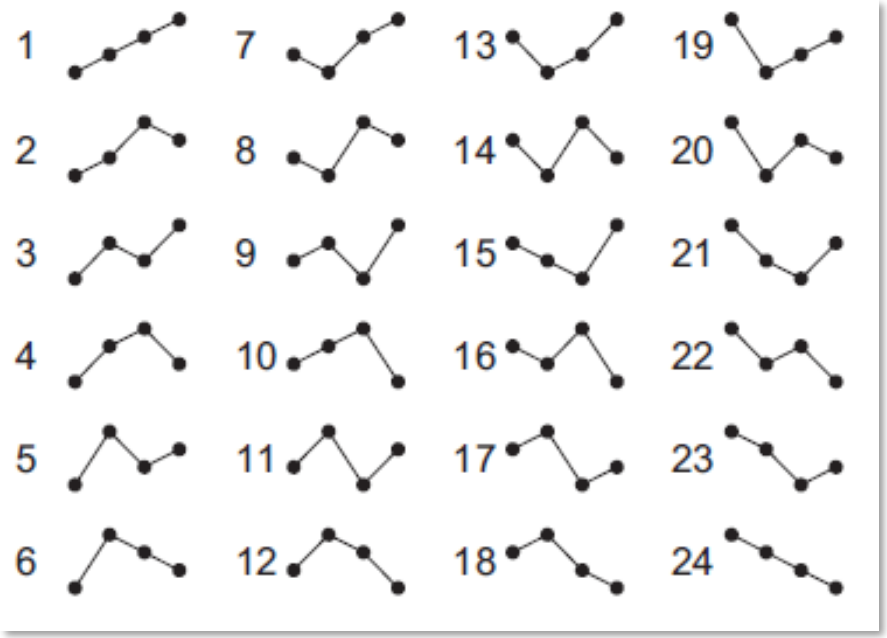
- Count how many times each “ordinal pattern” appears.
- Advantages: allows to identify temporal structures & is robust to noise.
- Drawback: information about actual data values is lost.

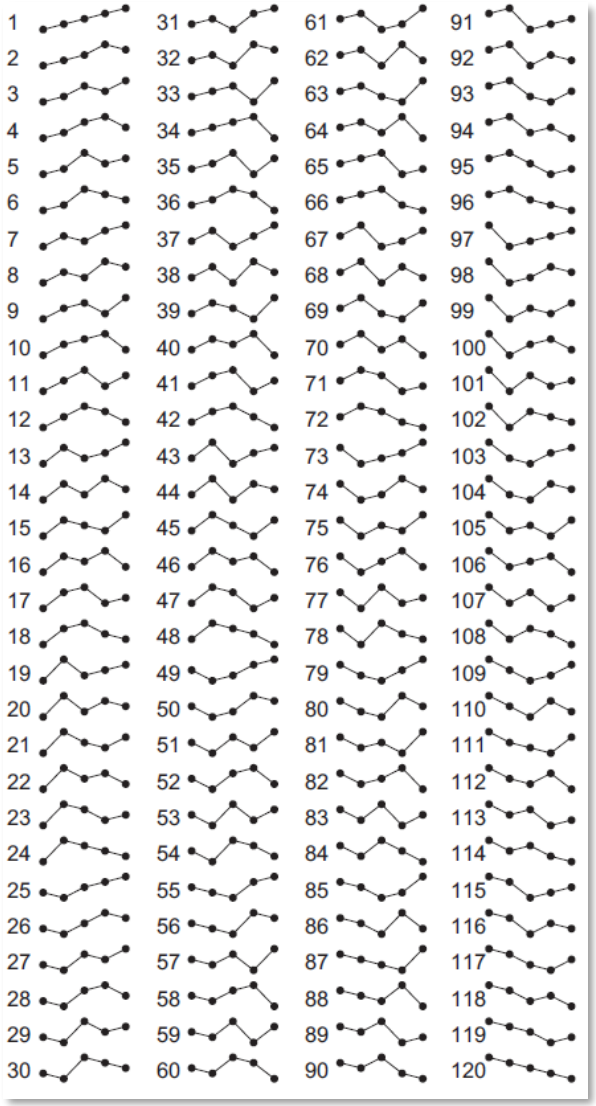
Bandt and Pompe, Phys. Rev. Lett. 88, 174102 (2002)

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

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

Comparison between the two rules to define symbols

Threshold transformation:

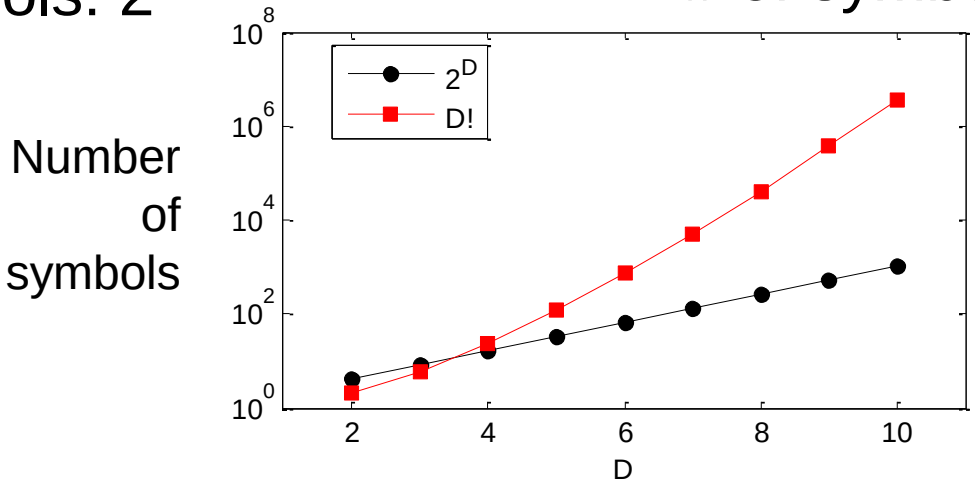
if $x_i > x_{th} \Rightarrow s_i = 0$; else $s_i = 1$

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

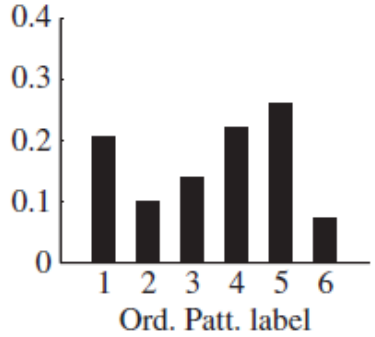
Ordinal transformation:

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
- # of symbols: D!



Are the *D*! ordinal patterns equally probable?



- **Null hypothesis**:

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

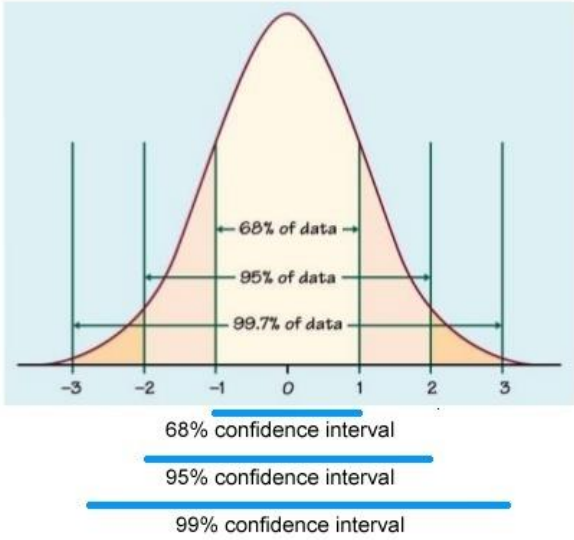
- If at least one probability **is not** in the interval $p \pm 3\sigma$ with $\sigma = \sqrt{p(1-p)/N}$

  and *N* the number of ordinal patterns:

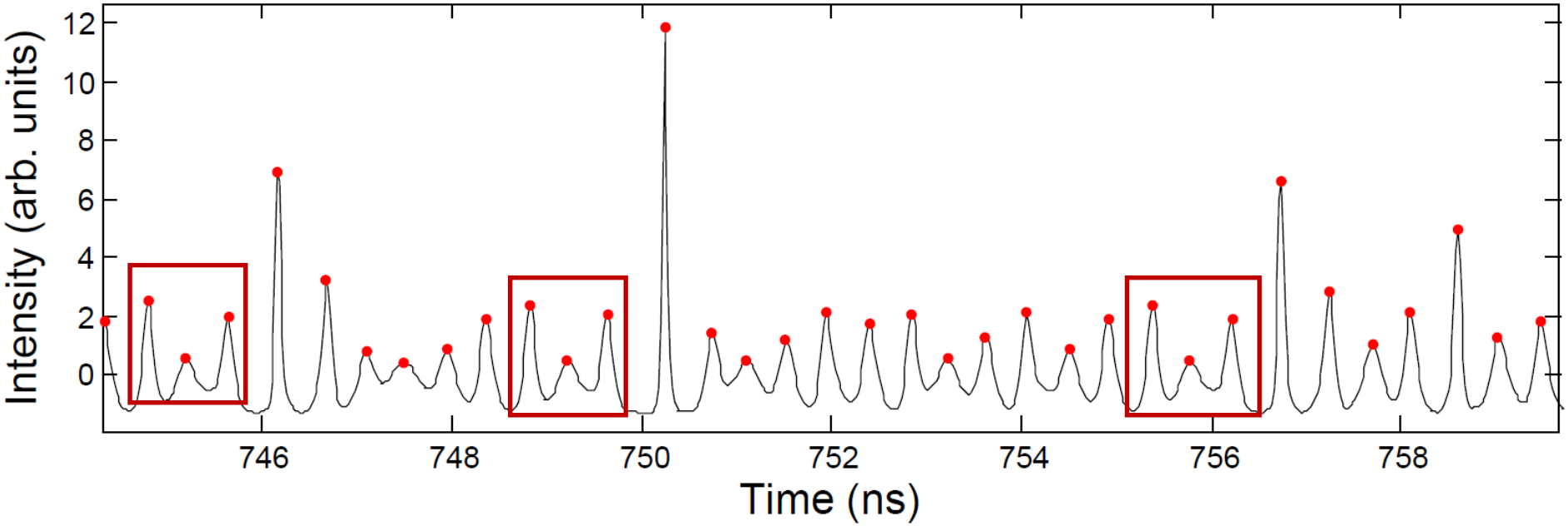
  We **reject** the NH with 99.74% confidence level.

- Else

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



Example: intensity pulses emitted by a chaotic laser



N. Martinez Alvarez et al, Eur. Phys. J. Spec. Top. 226, 1971 (2017).

How to detect longer temporal correlations?

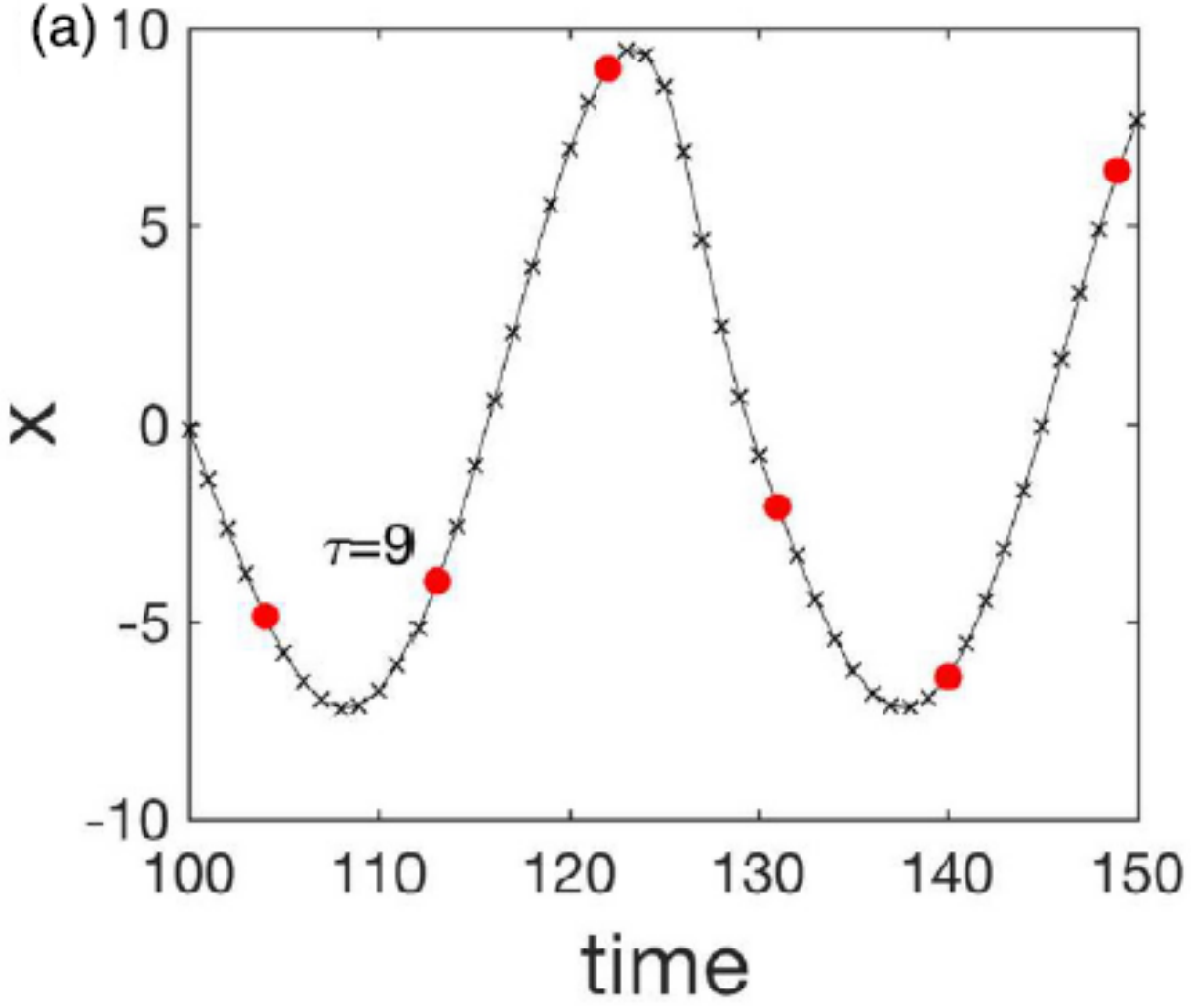
$$[\ldots x(t), x(t+1), x(t+2), x(t+3), x(t+4), x(t+5)\ldots]$$

- Problem: number of patterns increases as D!.

- Solution: a lag τ allows considering long time-scales without having to use words of many letters

$$[\ldots x(t), x(t+2), x(t+4),\ldots]$$

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



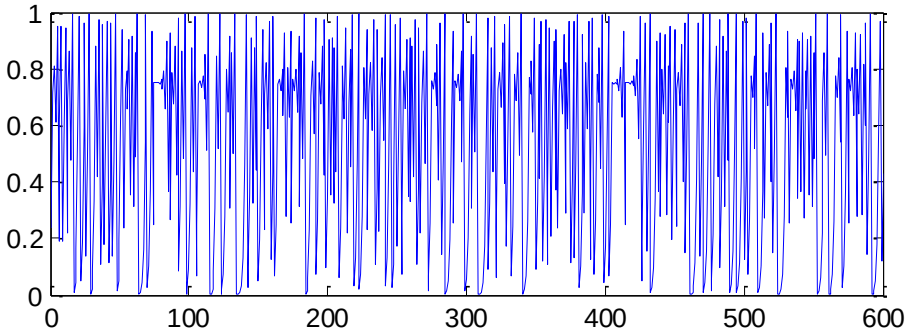
What to do if two values are exactly equal?
Which is the pattern?

Several possible solutions, a simple one is to add a very small amount of noise: $x(t) = x(t) + \xi$.

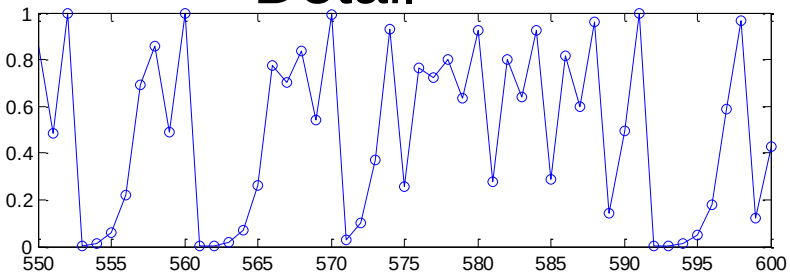
Y. Zou, R.V. Donner, N. Marwan et al. / Physics Reports 787 (2019) 1–97

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

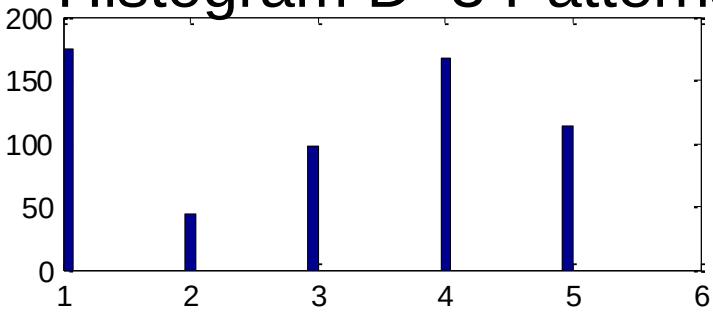
Time series



Detail

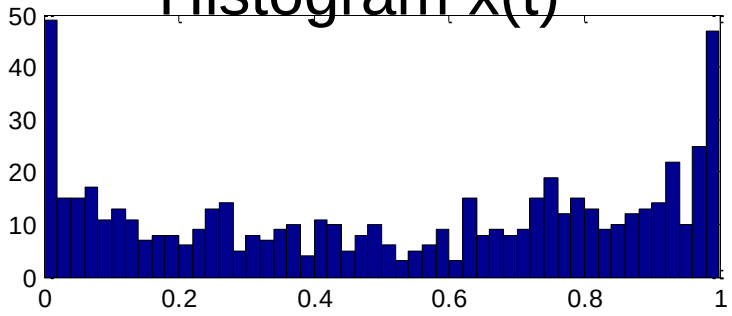


Histogram D=3 Patterns



↑ **forbidden**

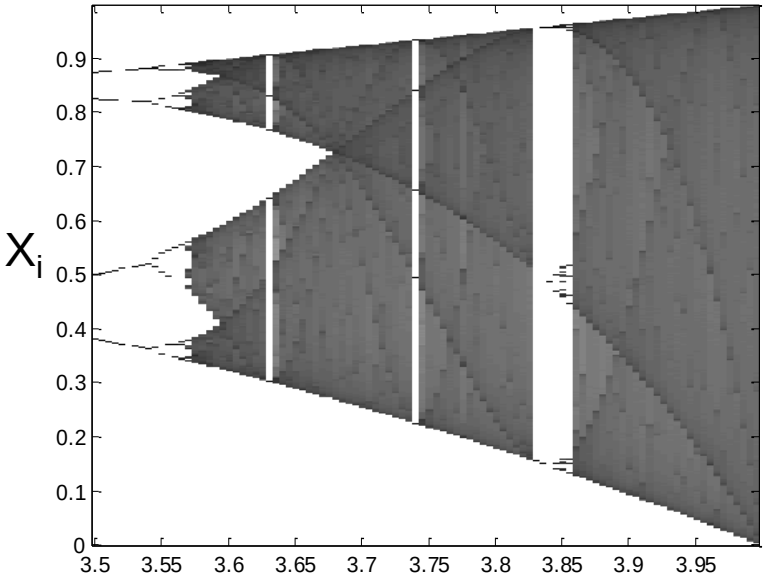
Histogram x(t)



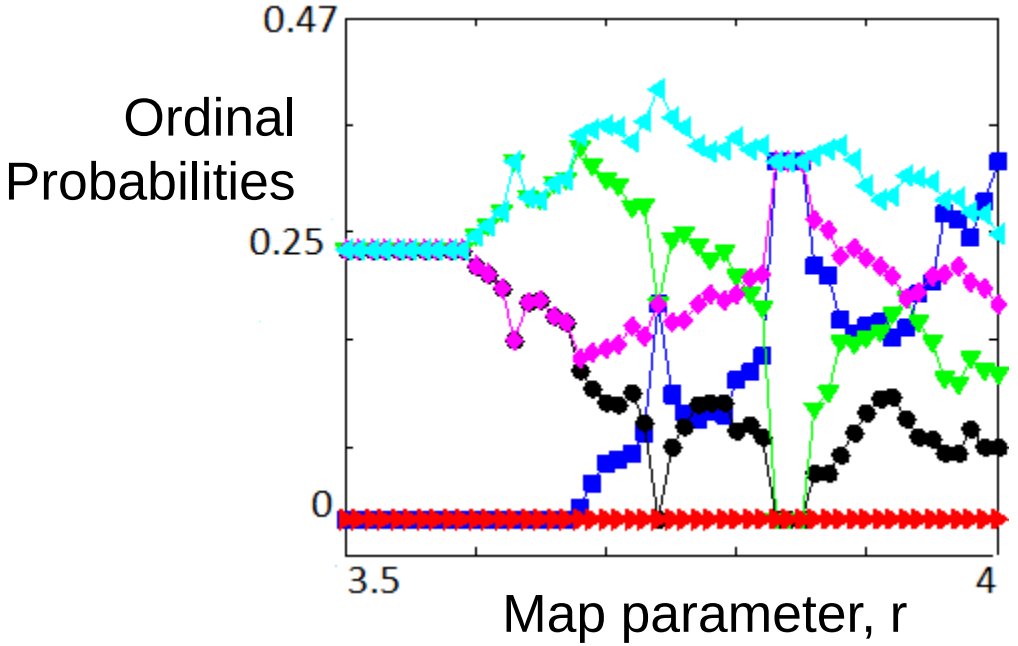
Ordinal analysis yields information about more and less expressed patterns in the data

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

Normal bifurcation diagram



Ordinal bifurcation diagram

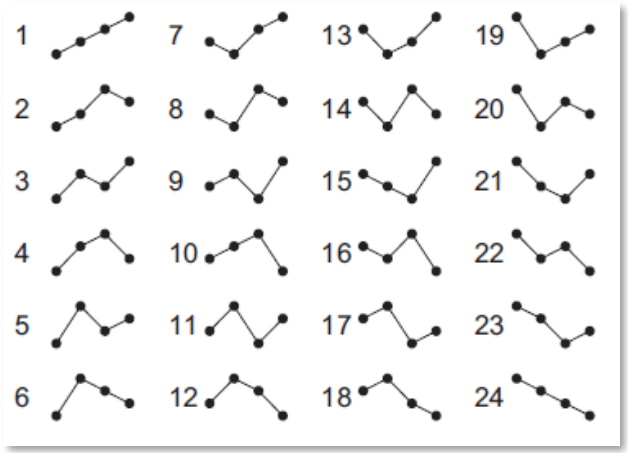


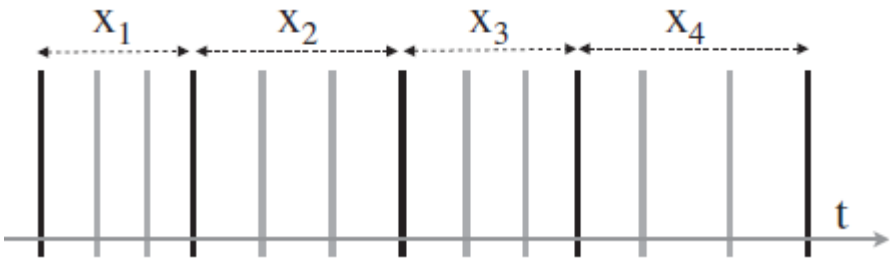
012 021 102 120 201 210

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

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)]





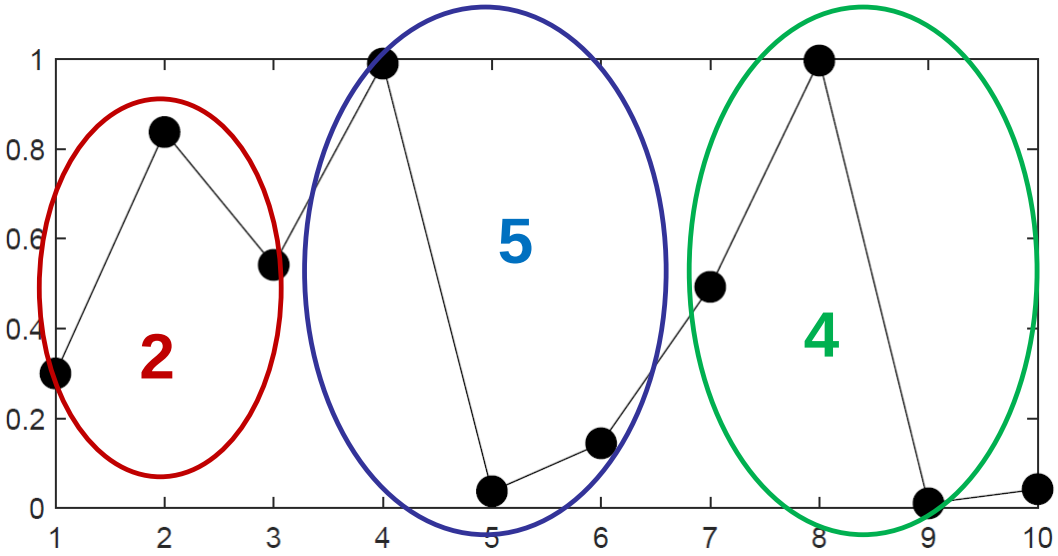
World length (wl): 4
Lag = 3 (skip 2 points)
Result:

indcs= 3

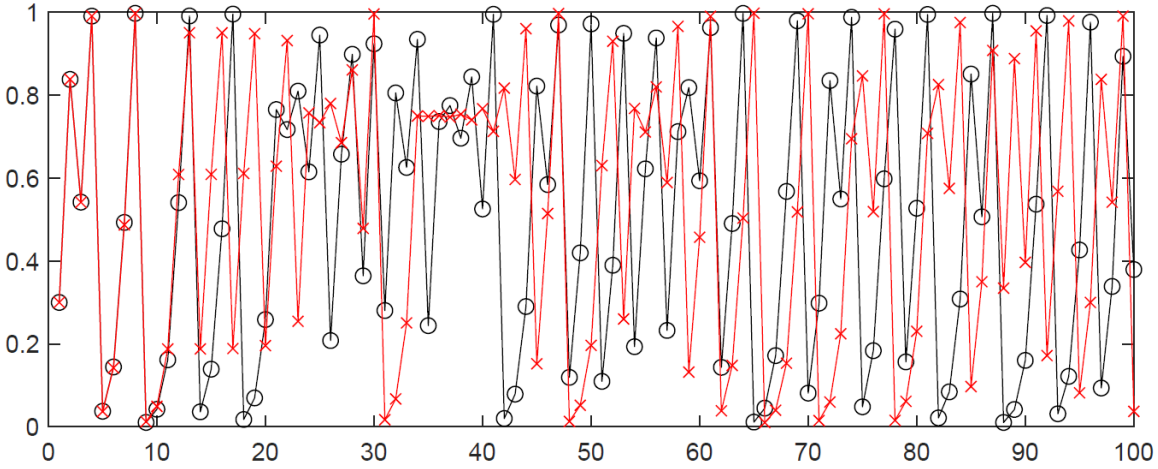
```
function indcs = perm_indices(ts, wl, lag) ;
m = length(ts)-(wl-1)*lag;
 indcs = zeros(m,1) ;
for i = 1:wl-1 ;
   st = ts(1+(i-1)*lag : m+(i-1)*lag) ;
   for j = i:wl-1 ;
      indcs = indcs+(st > ts(1+j*lag : m+j*lag));
   end
   indcs = indcs*(wl-i) ;
end
indcs=indcs + 1 ;
```

Hands-on exercise 6: ordinal analysis of the logistic map

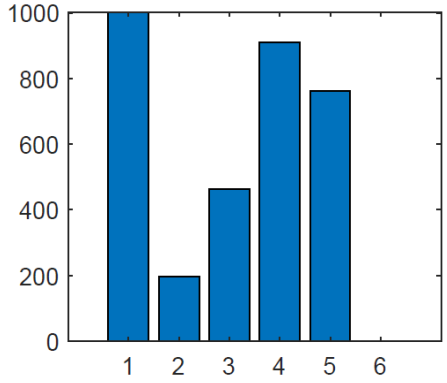
- Test that the program works with some examples.
- Calculate the probabilities of the 6 ordinal patterns of length D=3 for the logistic map with r=3.99.
- Calculate the ordinal bifurcation diagram: r in (3.5,4) with D=3.
- For r=3.99 generate two trajectories starting from very similar initial conditions and calculate the sequence of ordinal patterns and their distribution.

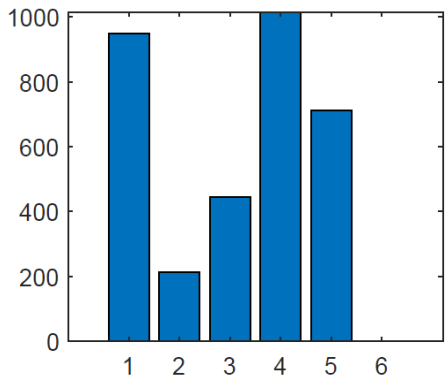


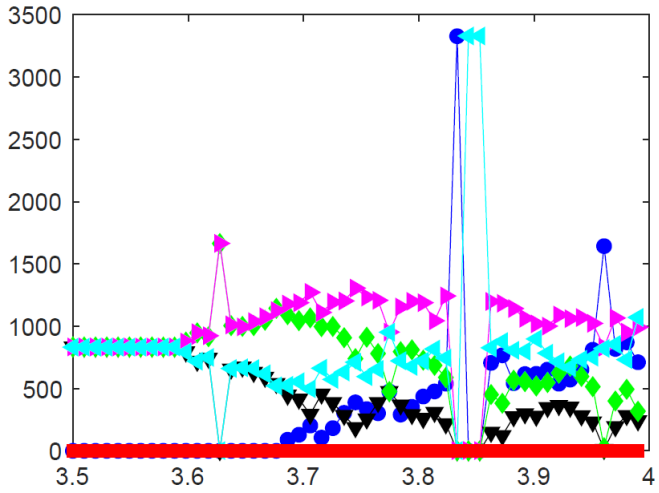
r=3.99, x(1)-y(1)=1e-4

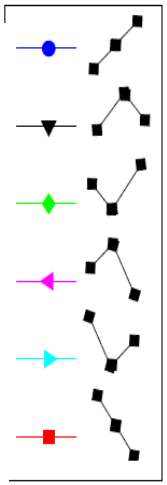


In the periodic windows the probabilities depend on the initial conditions.







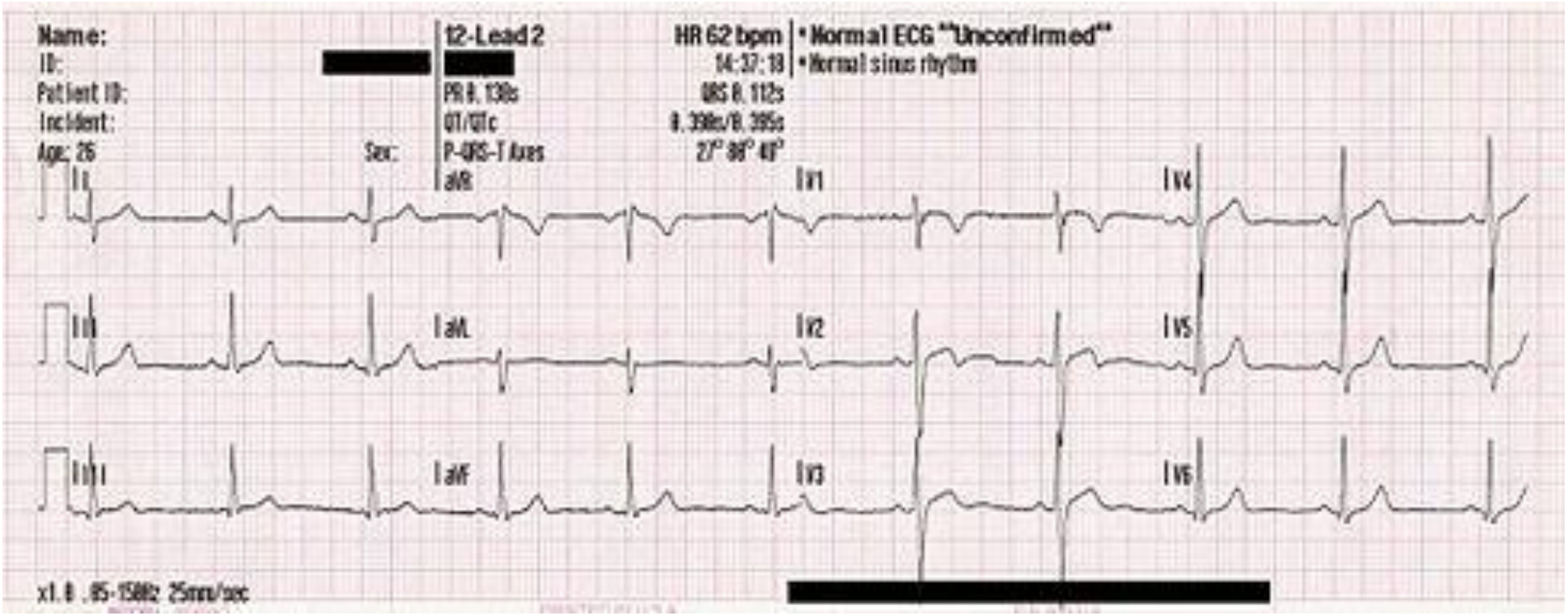


Examples of application

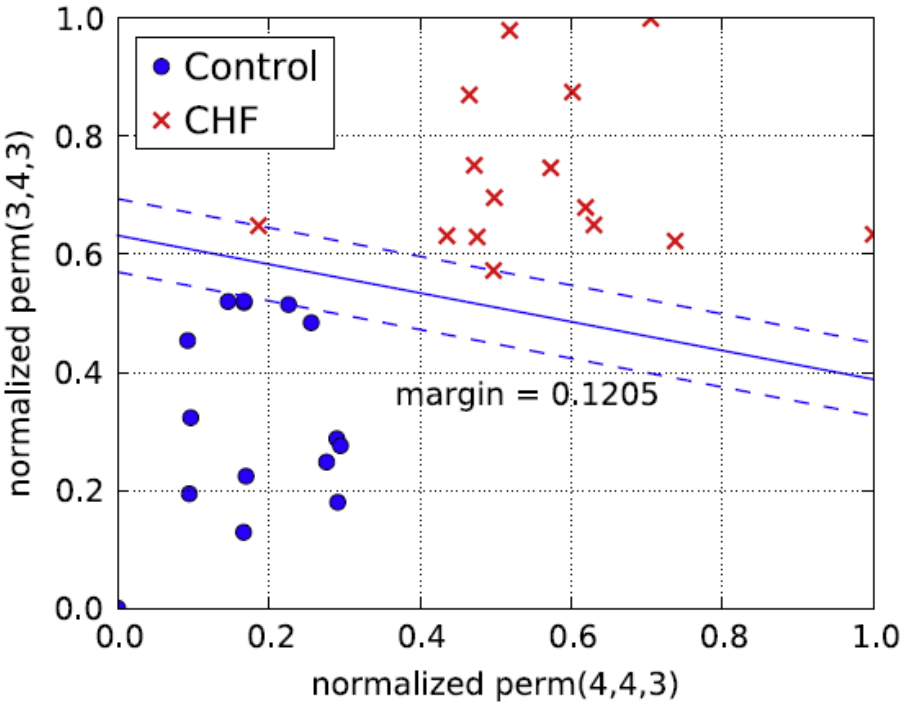
-cardiac signals

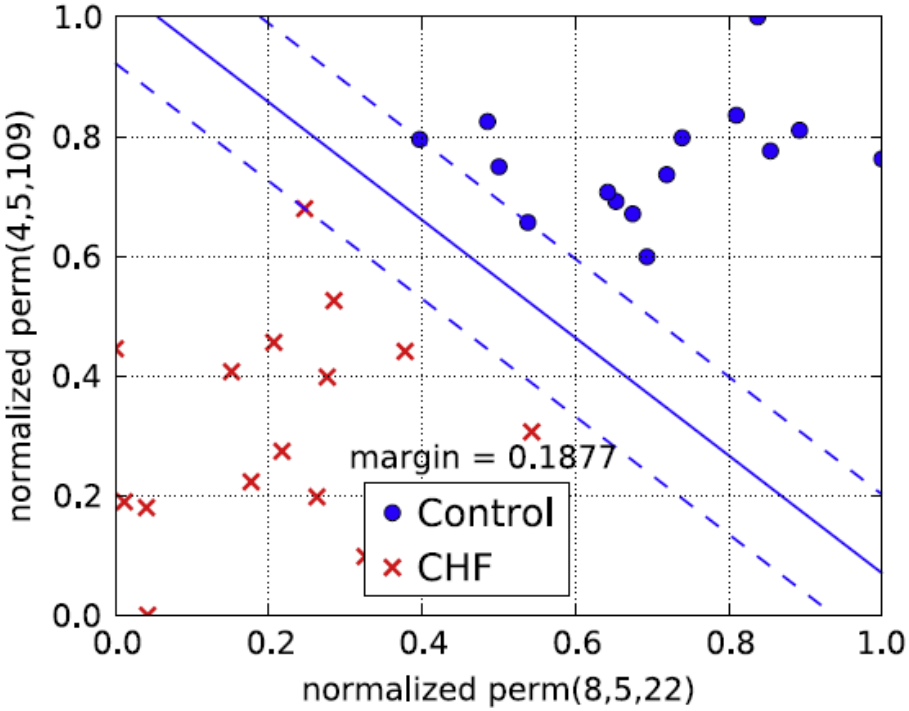
-lasers & neurons

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



Classifying ECG signals according to ordinal probabilities

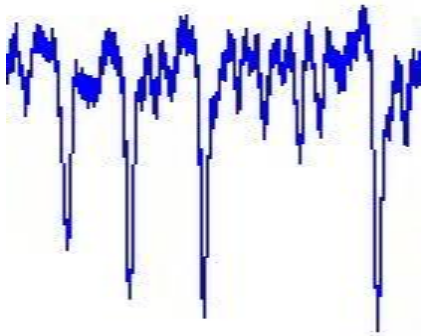




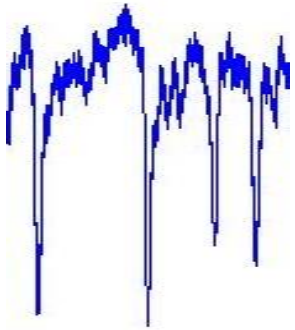
- Analysis of raw data (statistics of ordinal patterns is almost unaffected by a few extreme values)
- The probabilities are normalized with respect to the smallest and the largest value occurring in the data set.

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

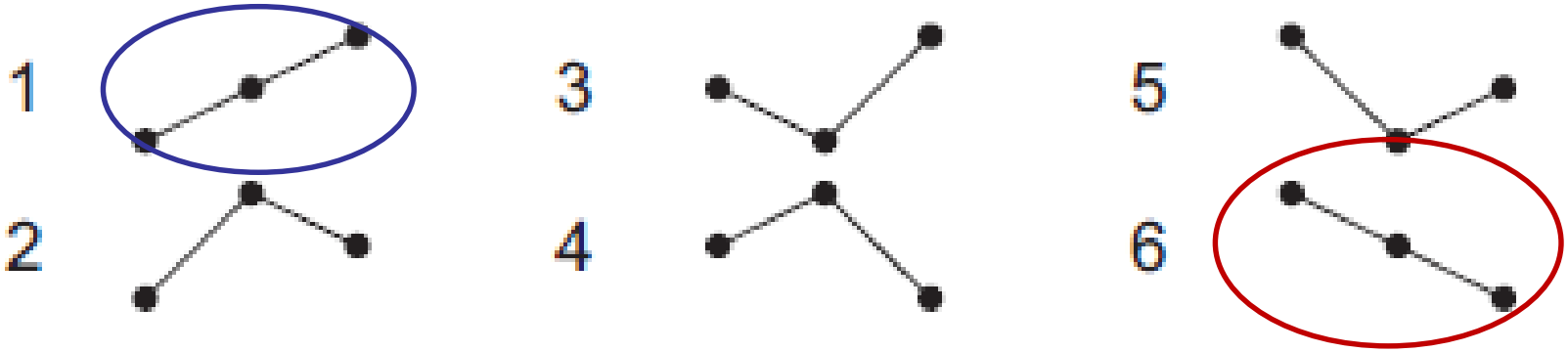
Laser spikes: analysis of inter-spike intervals



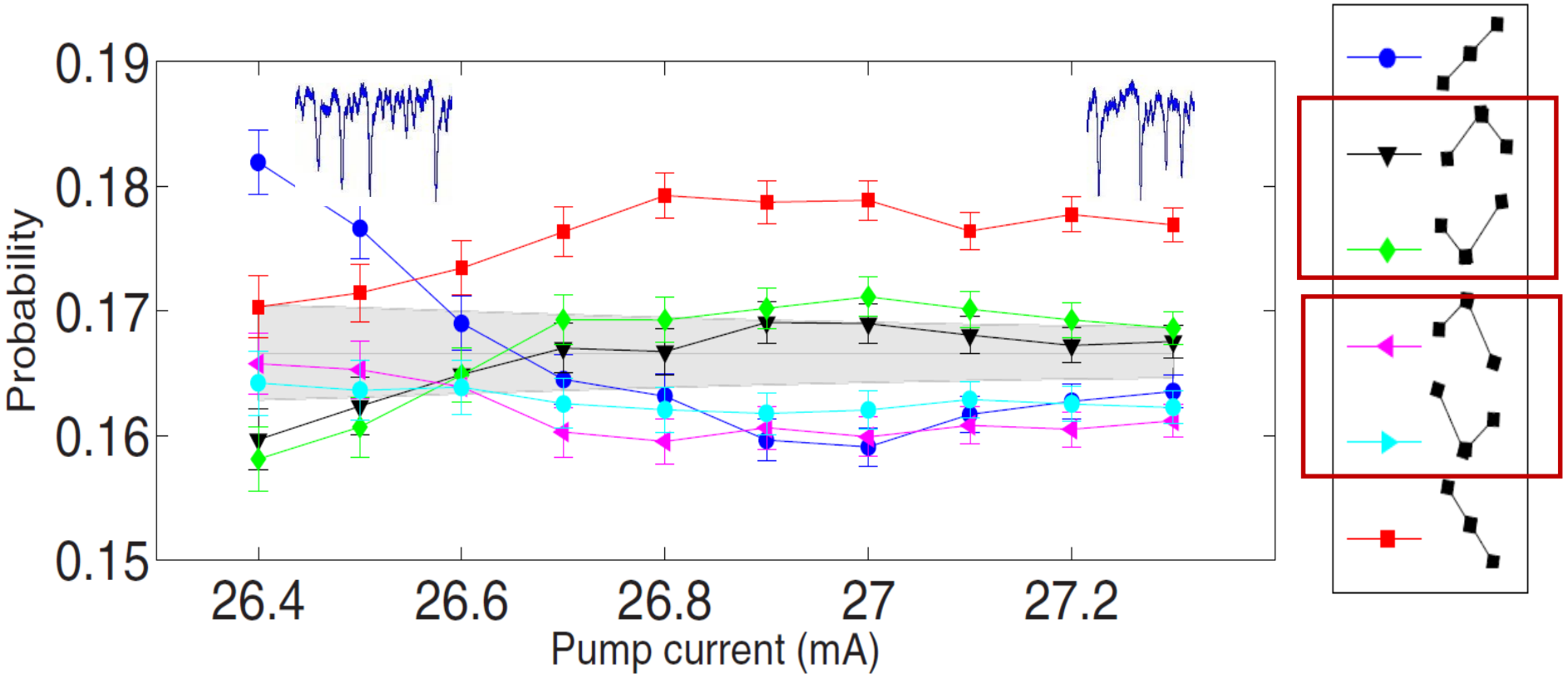
⇒012



⇒210



More probable pattern varies with the laser pump current

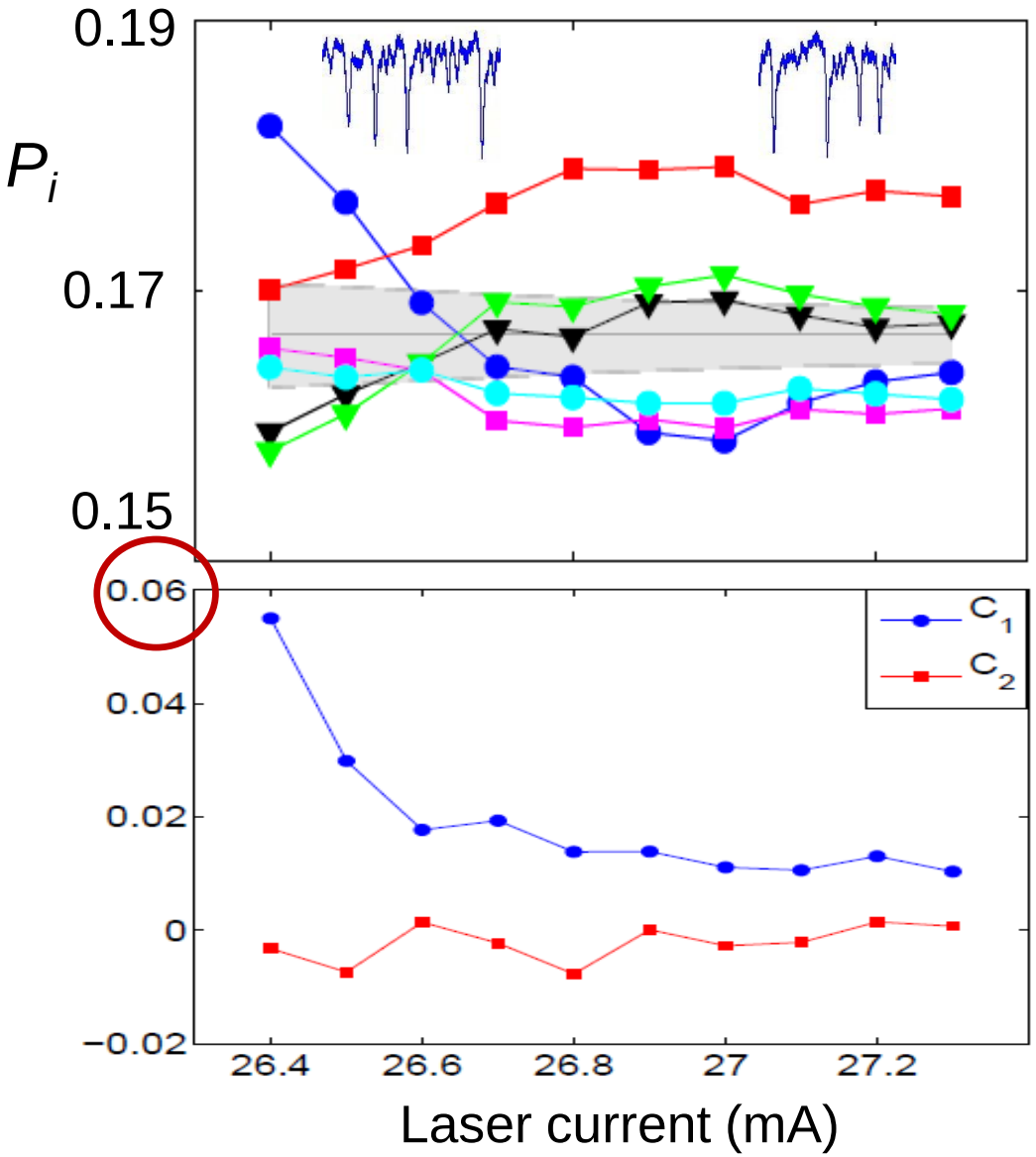


75,000 – 880,000 spikes

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

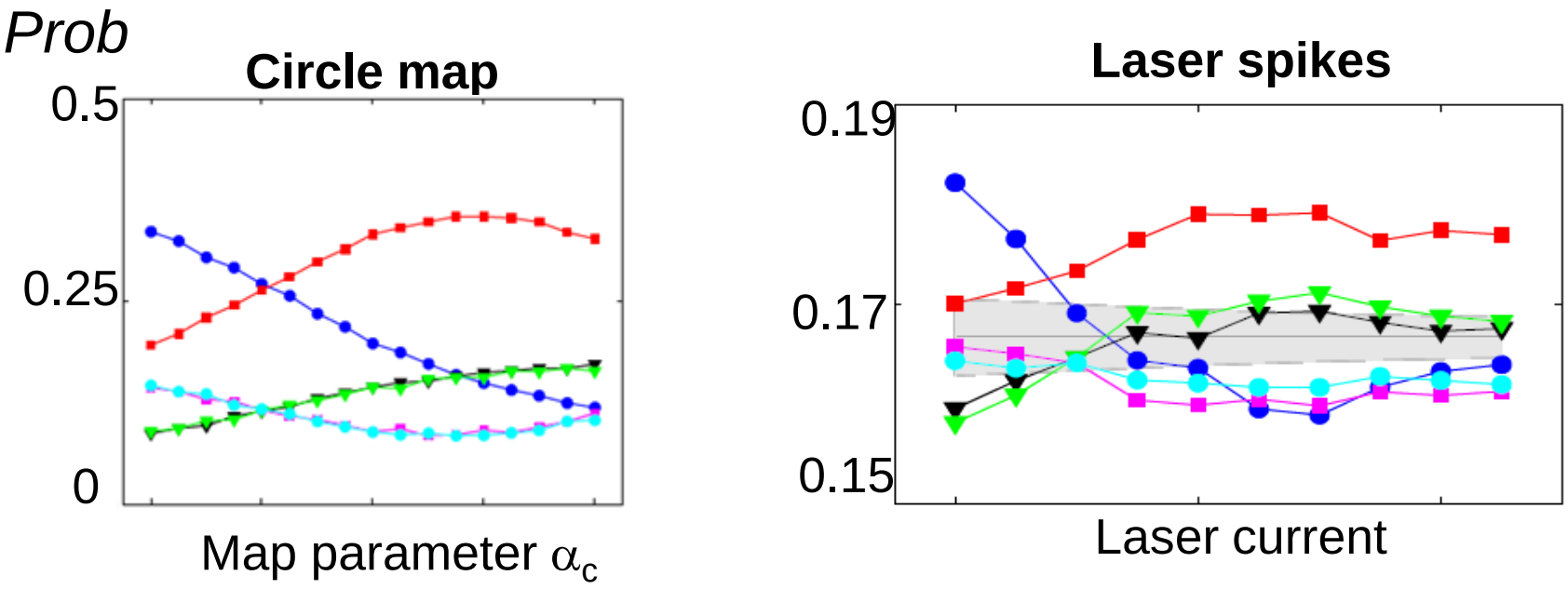
$$\varphi_{i+1} = \varphi_i + \rho + \frac{K}{2\pi}\left[\sin(2\pi\varphi_i) + \alpha_c \sin(4\pi\varphi_i)\right] + D\zeta$$

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

ρ = natural frequency / forcing frequency

K = forcing amplitude

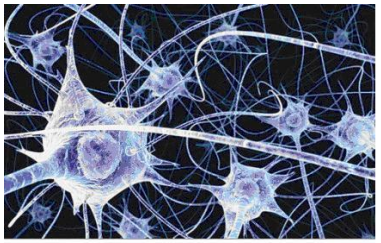
D = noise strength



Same “clusters” & same hierarchical structure

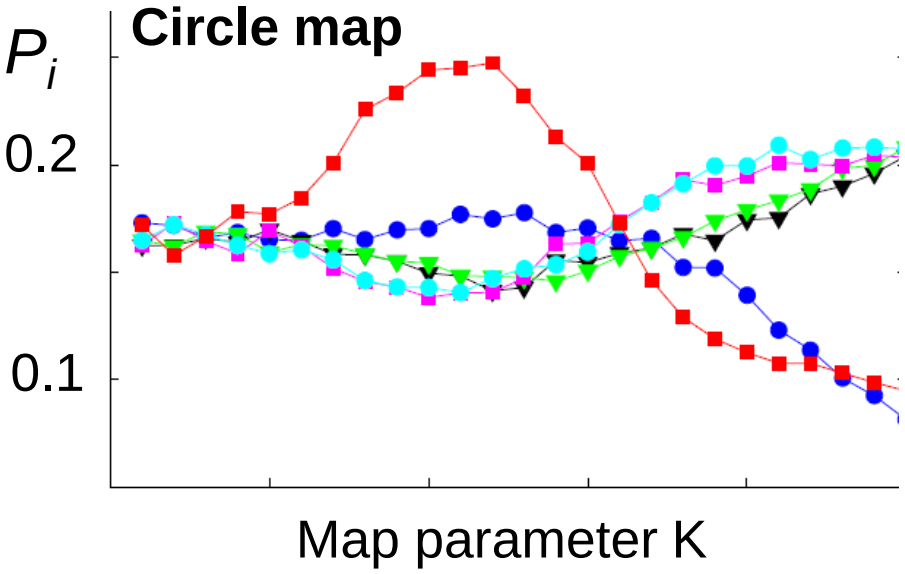
A. Aragoneses et al, Sci. Rep. 4, 4696 (2014)

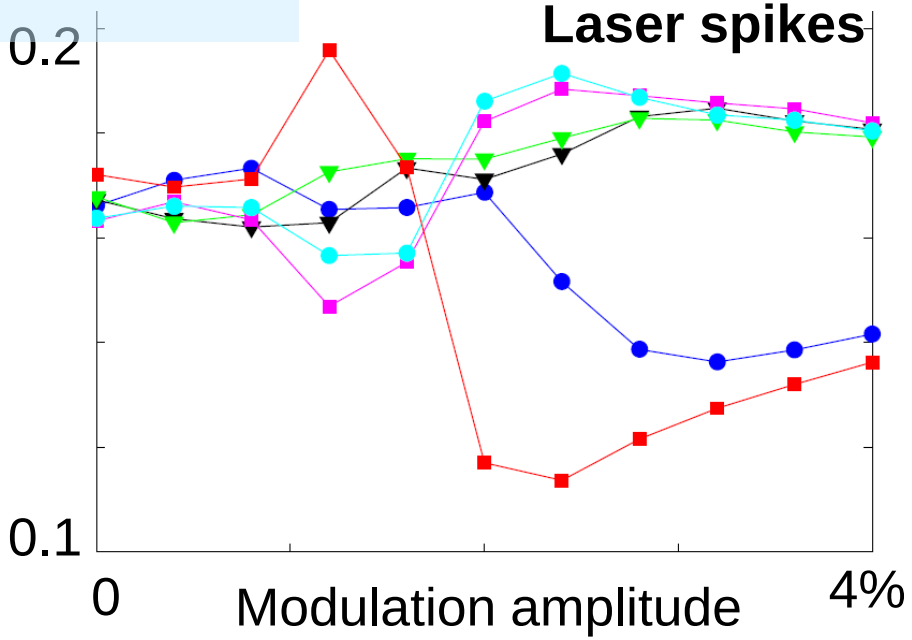
Connection with neurons: the circle map describes many excitable systems



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

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





Comparing with synthetic neuronal spikes: good agreement

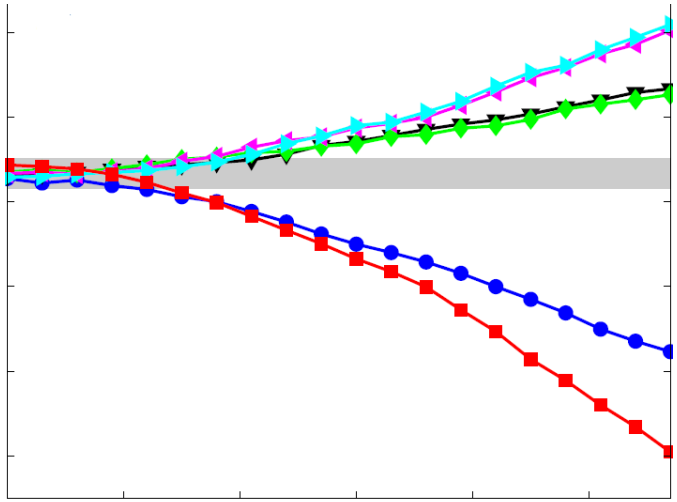
Neuron model with Gaussian white noise and **weak** sinusoidal input: spikes are noise-induced

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

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

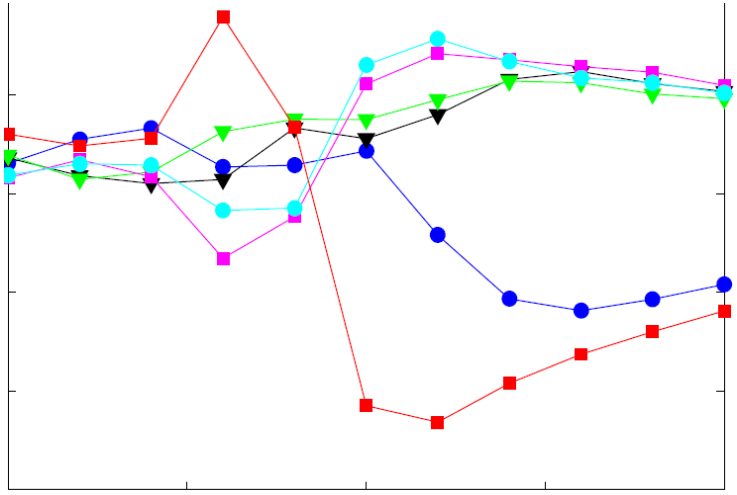
Synthetic spikes

P_i



Modulation amplitude

Empirical laser data



Modulation amplitude

J. M. Aparicio-Reinoso et al, Phys. Rev. E 94, 032218 (2016)

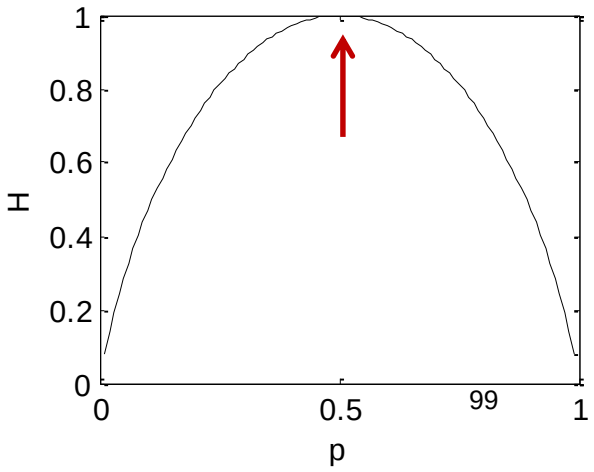
How to quantify unpredictability and complexity?

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: $$H = -\sum_{i=1}^{N} p_i \ln p_i$$
- Interpretation: "*quantity of **surprise** one should feel upon reading the result of a measurement*" Faser and Swinney (1986)
- Simple example: a random variable takes values 0 or 1 with probabilities: p(0) = p, p(1) = 1 − p.
- H = −p ln(p) − (1 − p) ln(1 − p).
  ⇒ p=0.5: Maximum **unpredictability.**

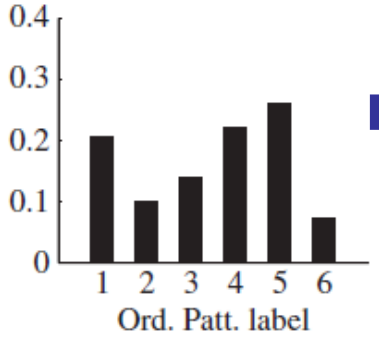


Permutation entropy

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

Time series

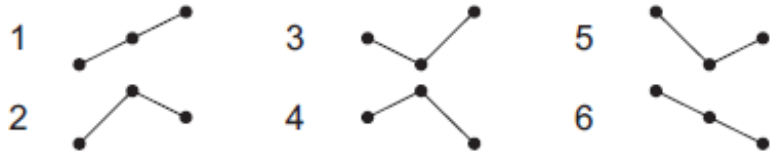
Ordinal probabilities



Permutation entropy

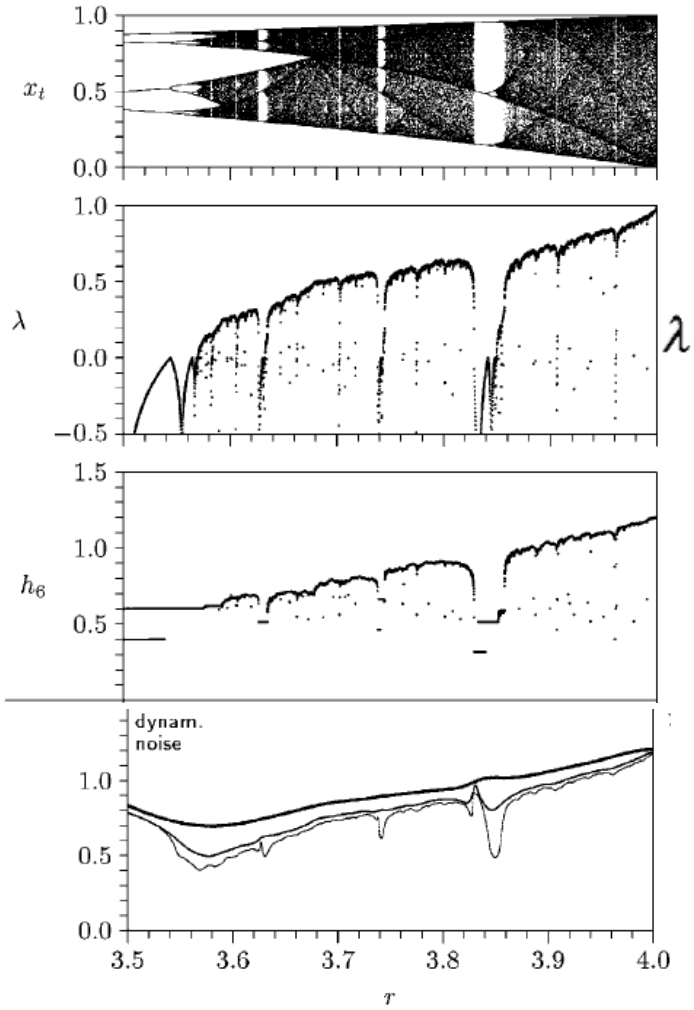
$$H = -\sum_{i=1}^{N} p_i \ln p_i$$

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



Permutation entropy (PE) of the Logistic map

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



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

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

Entropy per symbol:

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

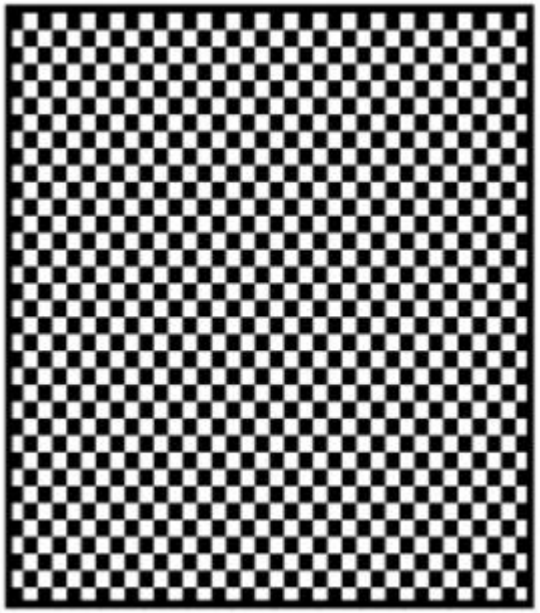
Robust to noise

Bandt and Pompe
Phys. Rev. Lett. 2002

The entropy measures the degree of 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

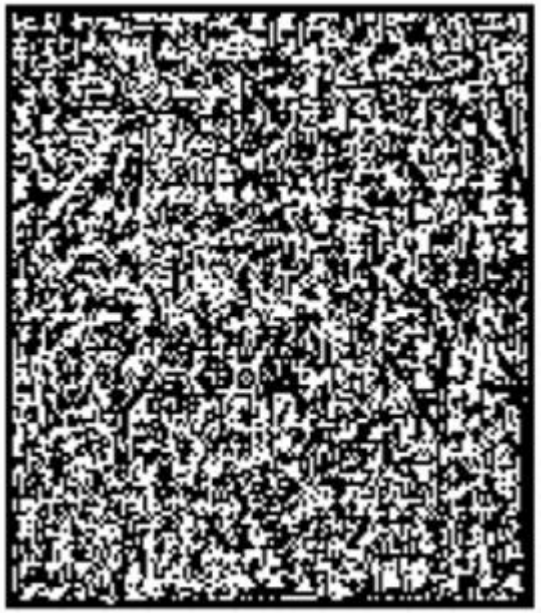


H = 0
C = 0

Chaos

H ≠ 0
C ≠ 0

Disorder



H = 1
C = 0

Source: O. A. Rosso

“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.”

Feldman, McTague and Crutchfield, Chaos 2008

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

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

Normalization

$$H[P] = \frac{I[P]}{I_{\max}} \qquad 0 \le H[P] \le 1$$

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

Example: if *I[P]* = 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

$$Q[P] = Q_0 D\left[P, P_e\right]$$

where Q_o is a normalization constant such that $0 \le Q[P] \le 1$

Distance between two distributions P and P_e

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

Kullback–Leibler divergence (relative entropy) $$D_{KL}[P, P_e] = \sum_i p_i \log \frac{p_i}{p_{i,e}}$$

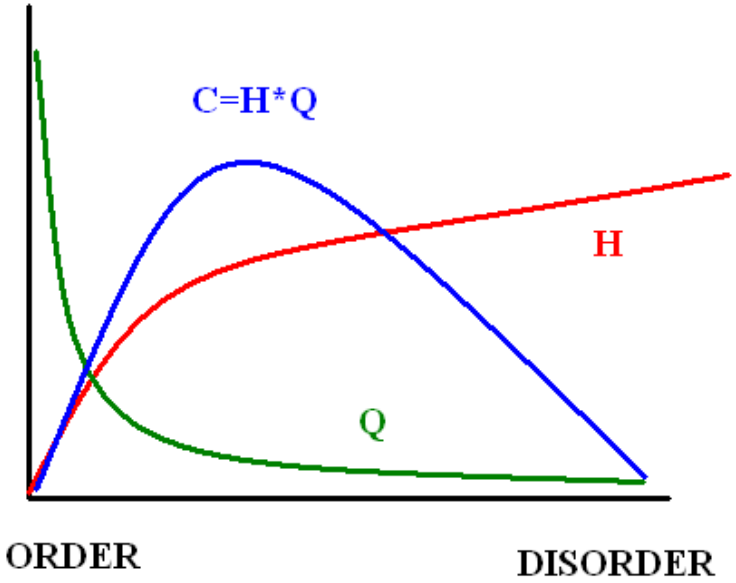
Jensen divergence $$D_J[P, P_e] = \frac{D_{KL}\left[P \middle| P_e\right] + D_{KL}\left[P_e \middle| P\right]}{2}$$

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)

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

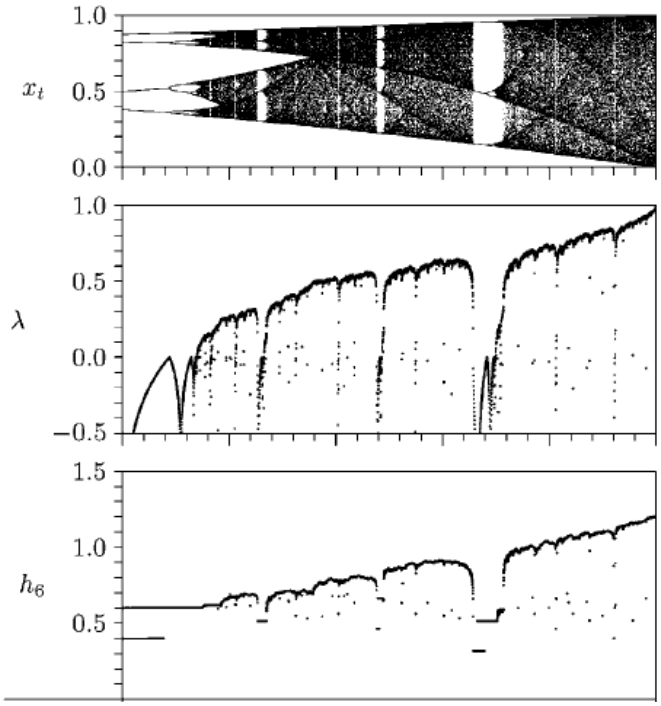
Lopez-Ruiz, Mancini & Calbet, Phys. Lett. A (1995).
Anteneodo & Plastino, Phys. Lett. A (1996).

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

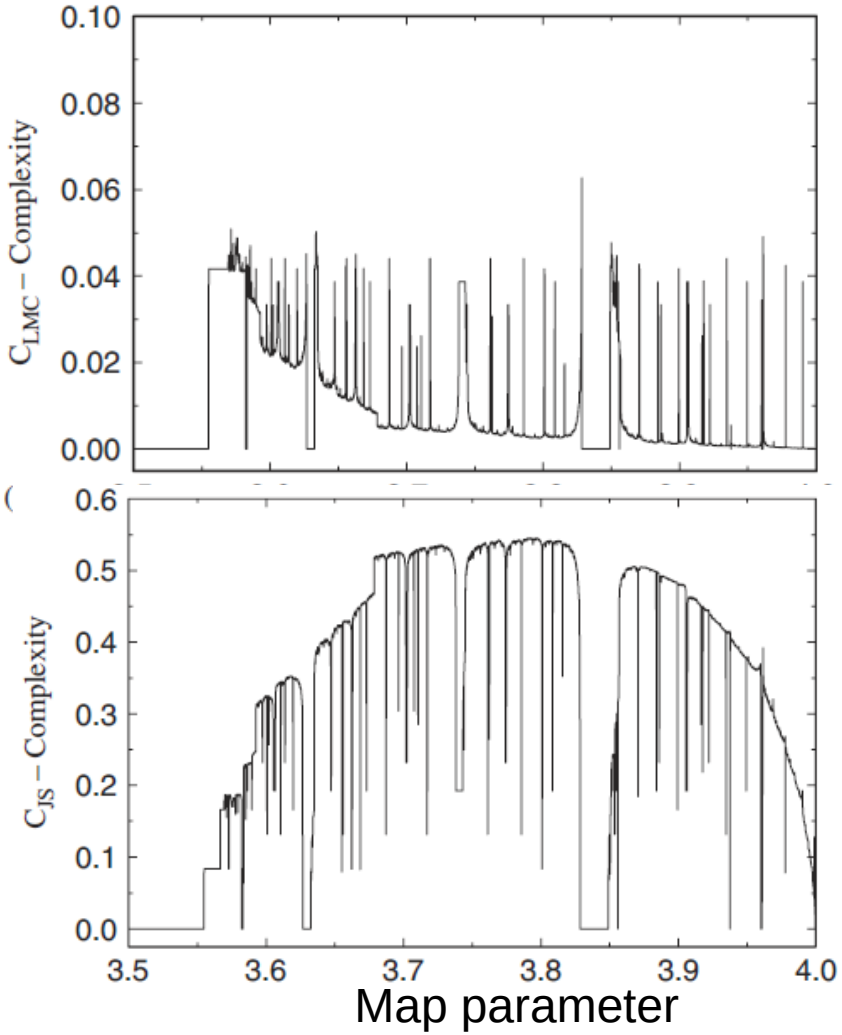
Martín, Plastino & Rosso, Phys. Lett. A (2003).

The complexity of the Logistic Map

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



Map parameter



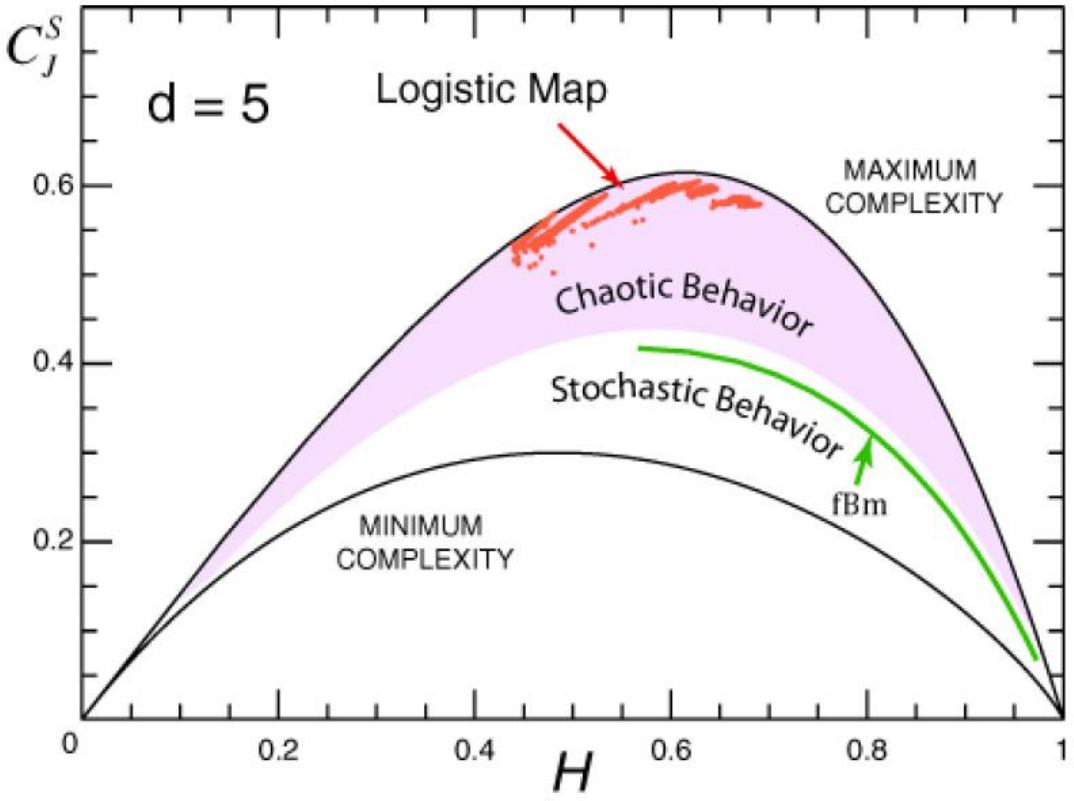
Map parameter

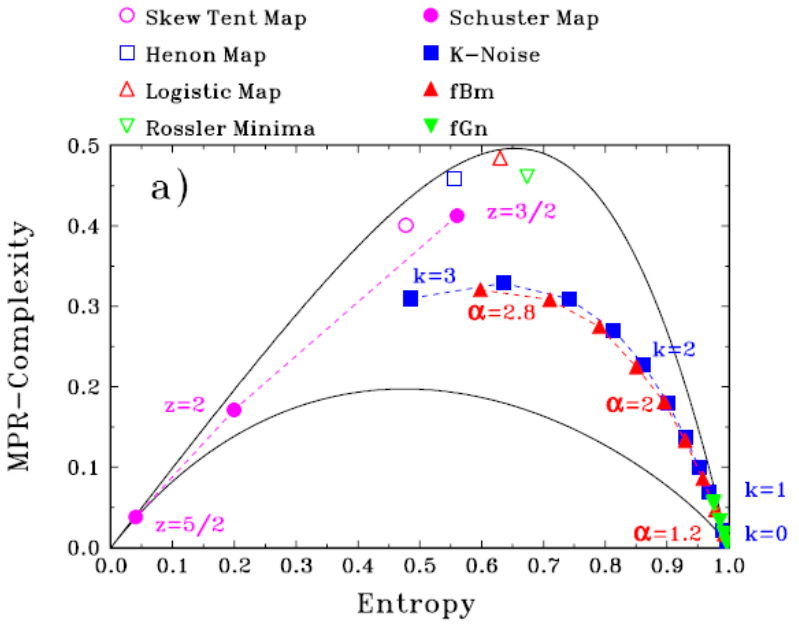
Euclidian distance

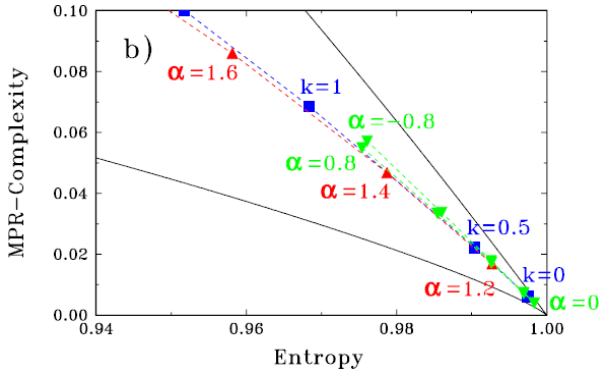
Jensen distance

Martín, Plastino, & Rosso, Physica A 2006

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

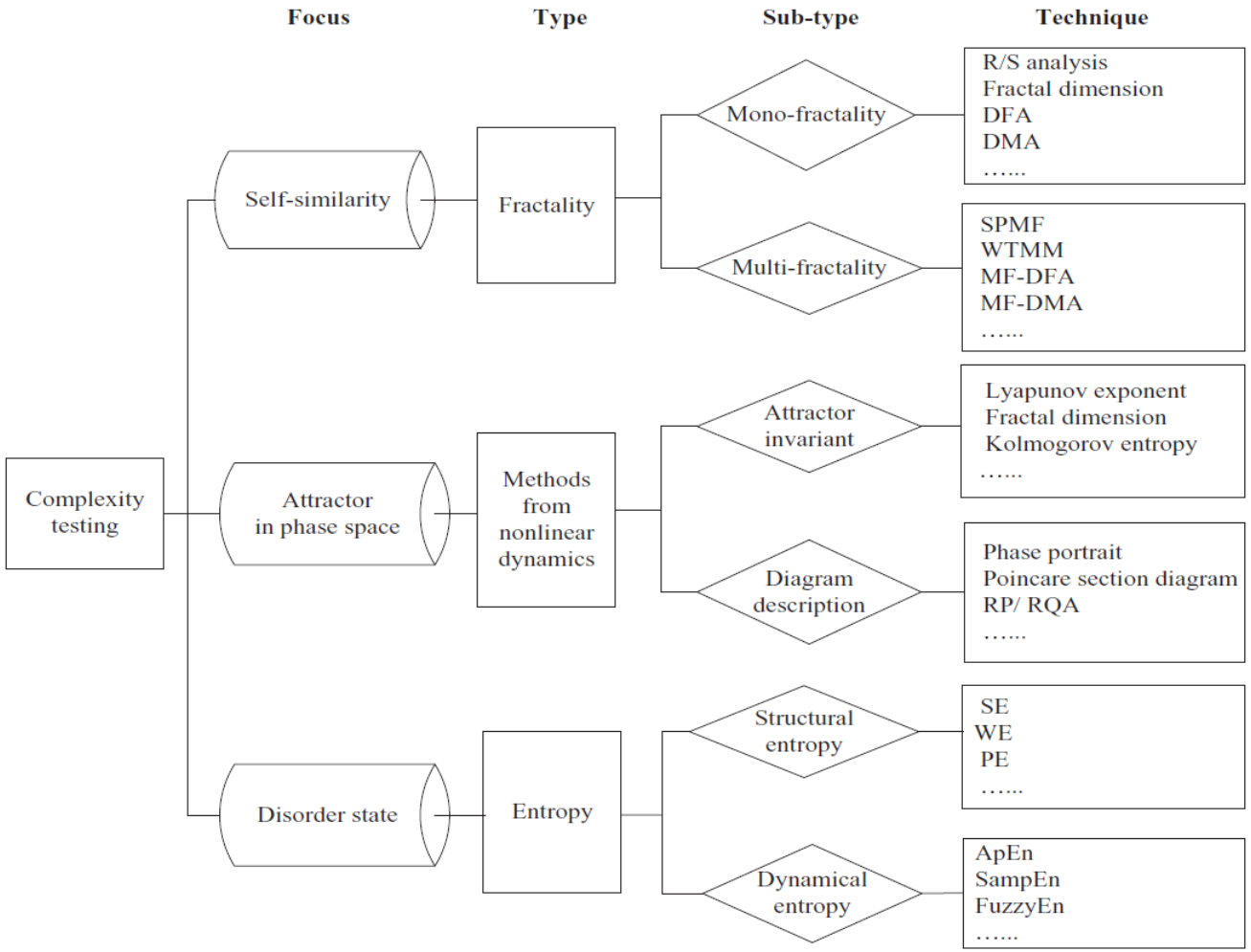






O. A. Rosso et al, Phys. Rev. Lett. 99, 154102 (2007)

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

Kolmogorov complexity

- The complexity of an object is a measure of the computability resources needed to specify the object.

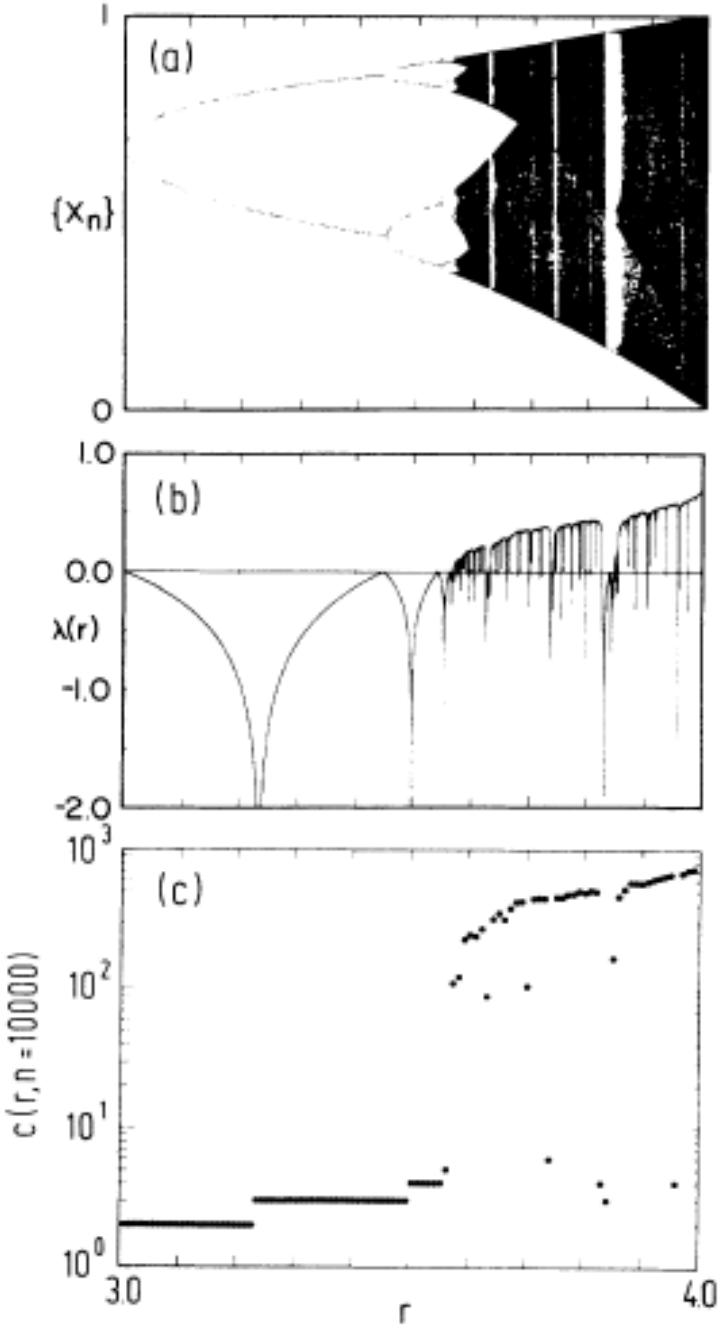
  Example: Let's consider 2 strings of 32 letters:

  abababababababababababababababab

  4c1j5b2p0cv4w1x8rx2y39umgw5q85s7

- The first string has a short description: "ab 16 times".
- The second has no obvious description: complex or random?
- The ***Lempel & Zip complexity*** is an estimation of the Kolmogorov complexity.

Lempel & Zip complexity of the Logistic Map



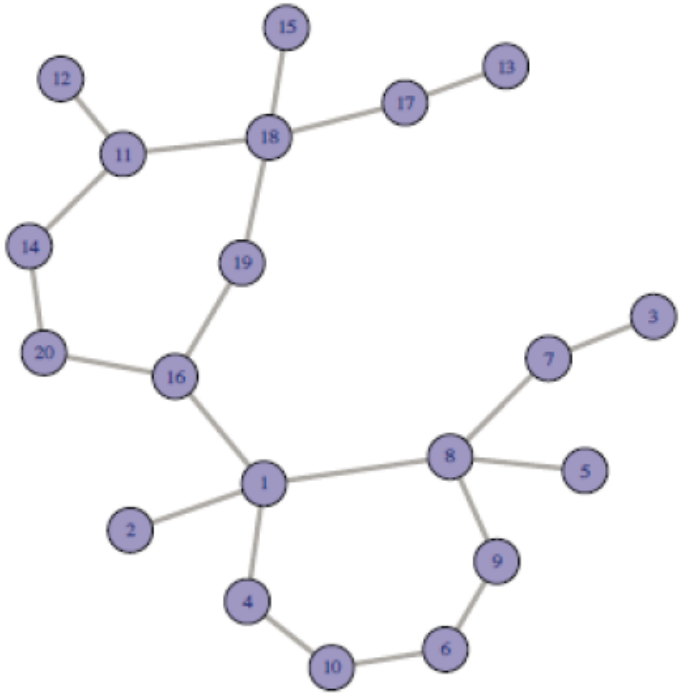
Kaspar and Schuster, Phys Rev. A 1987

Methods of time-series analysis

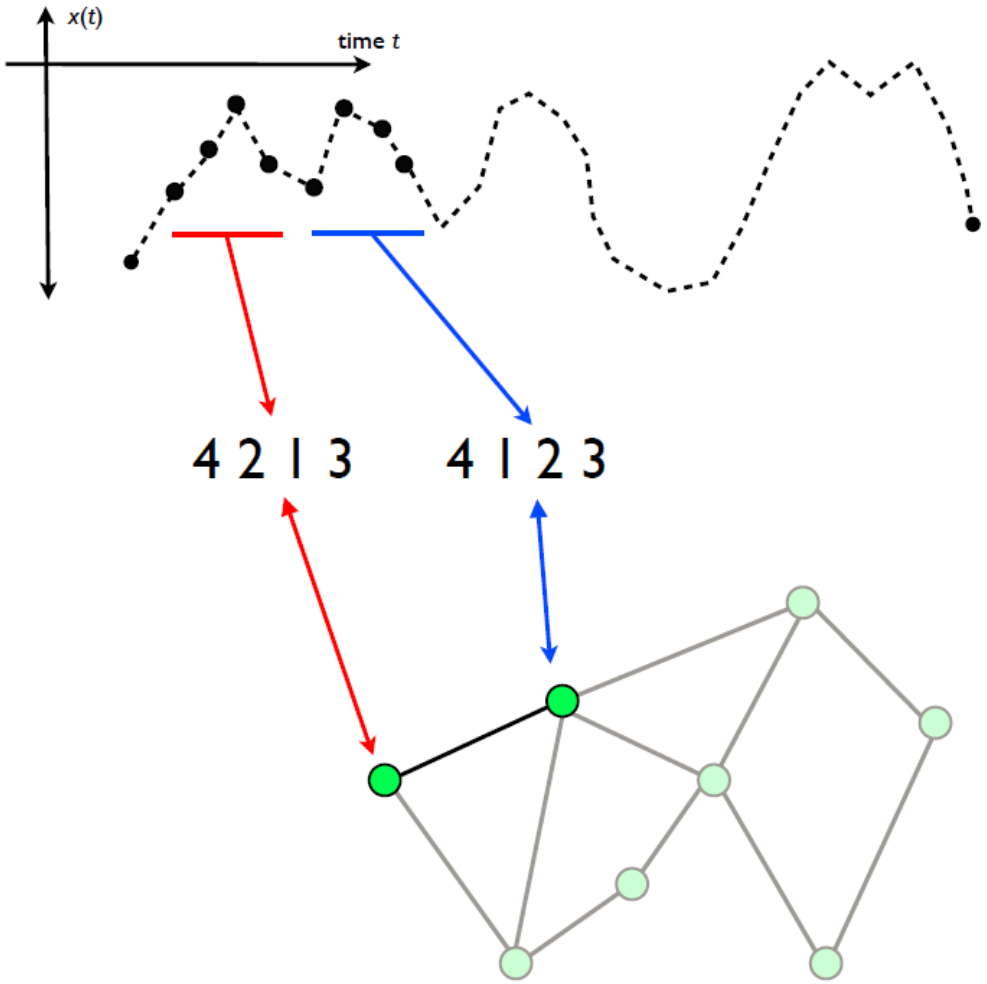
- Return maps
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Distribution of data values
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
-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

  $(\sum_i p_i=1)$

$\Rightarrow$ **Weighted and directed network**

Adapted from M. Small (The University of Western Australia)

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)

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

- Asymmetry coefficient: normalized difference of transition probabilities, P(‘01’→ ‘10’) - P(‘10’→ ’01’), etc.

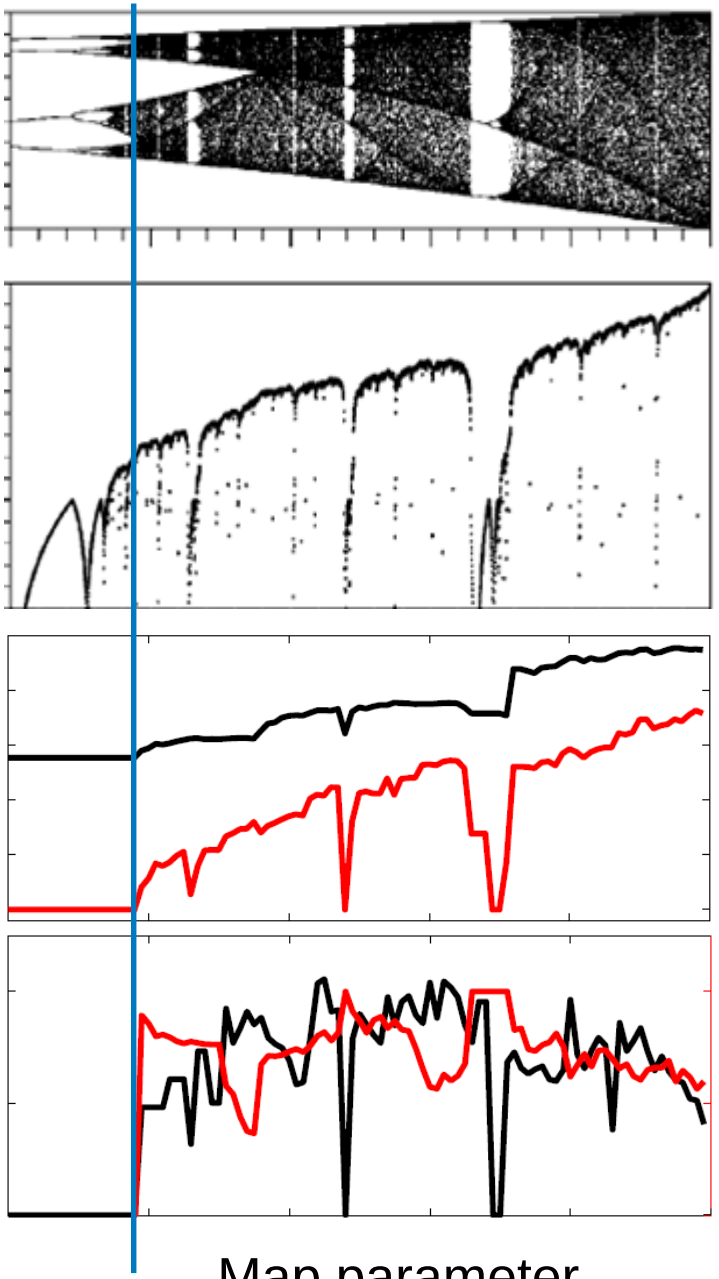
$$a_c = \frac{\sum_i \sum_{j \neq i} \left| w_{ij} - w_{ji} \right|}{\sum_i \sum_{j \neq i} \left(w_{ij} + w_{ji} \right)}$$

(0 in a fully symmetric network;
1 in a fully directed network)

A first test with the Logistic map

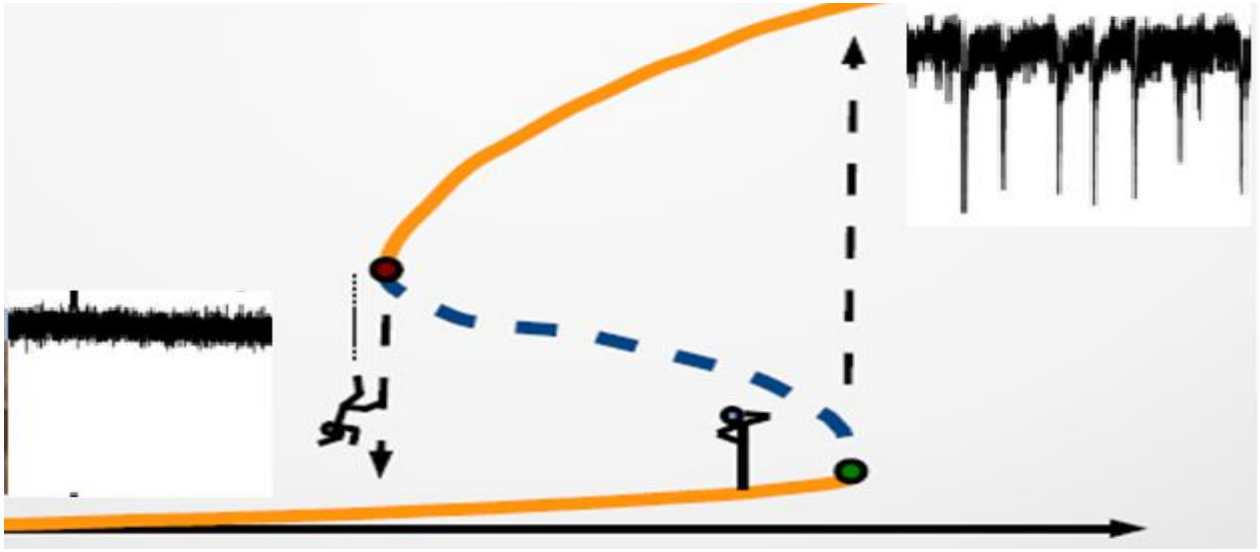
D=4

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



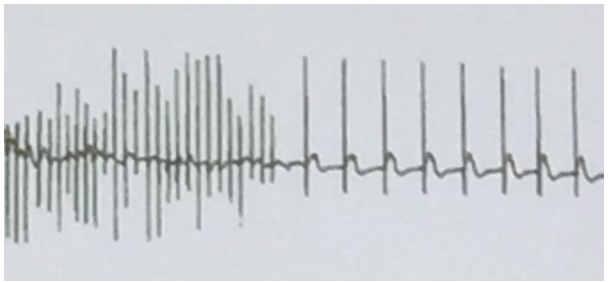
C. Masoller et al, NJP (2015)

Approaching a “tipping point”



Control parameter

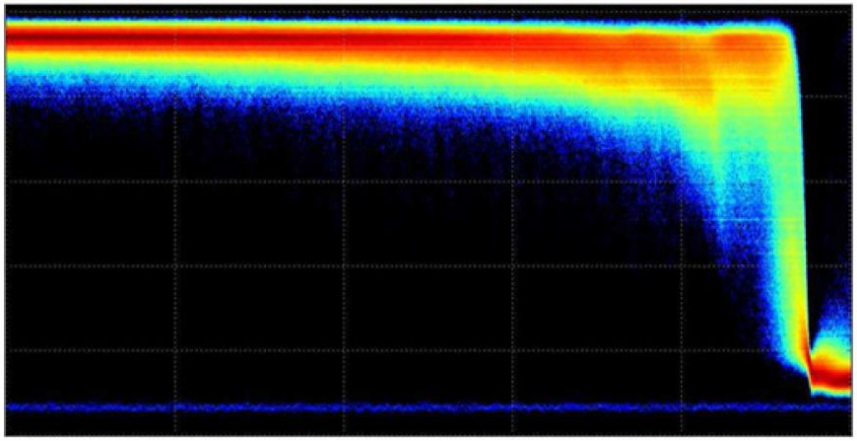
Can we use the ordinal network method to detect an *early warning signal* of a transition to a different dynamical regime?



Yes! Two examples: optical signals and brain signals

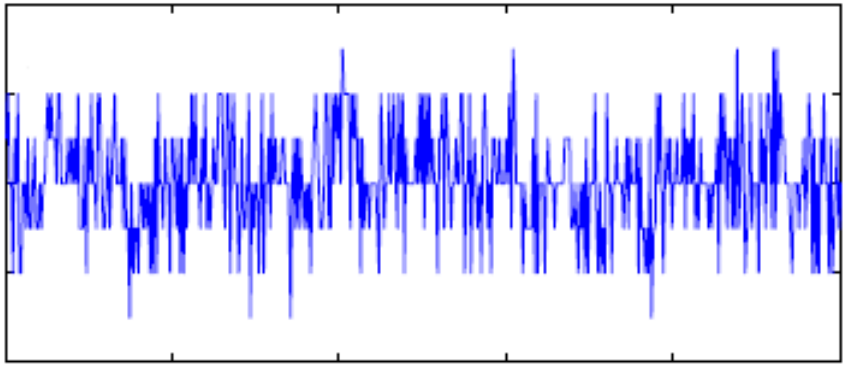
Apply the ordinal network method to laser data

As the laser current increases



Time

Intensity @ constant current

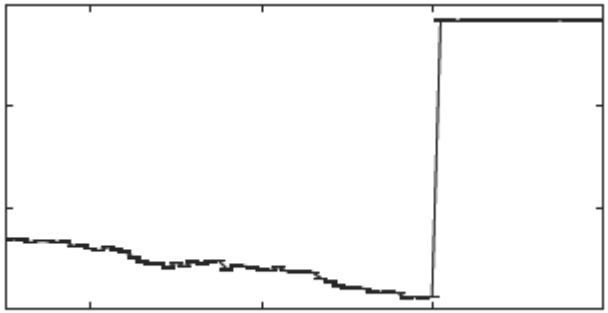


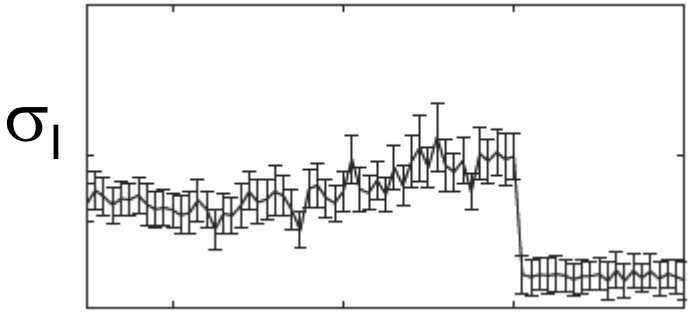
Time

- 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

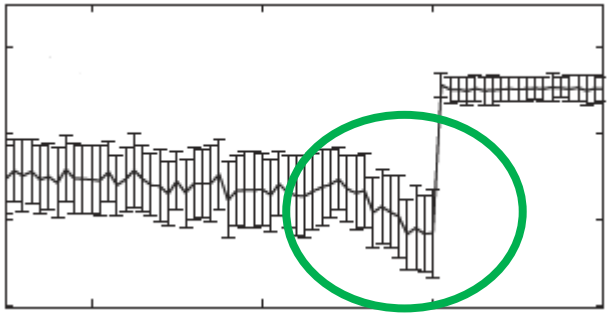




L=1000
100 windows

⇒ **No warning**

Node entropy s_n (D=3)

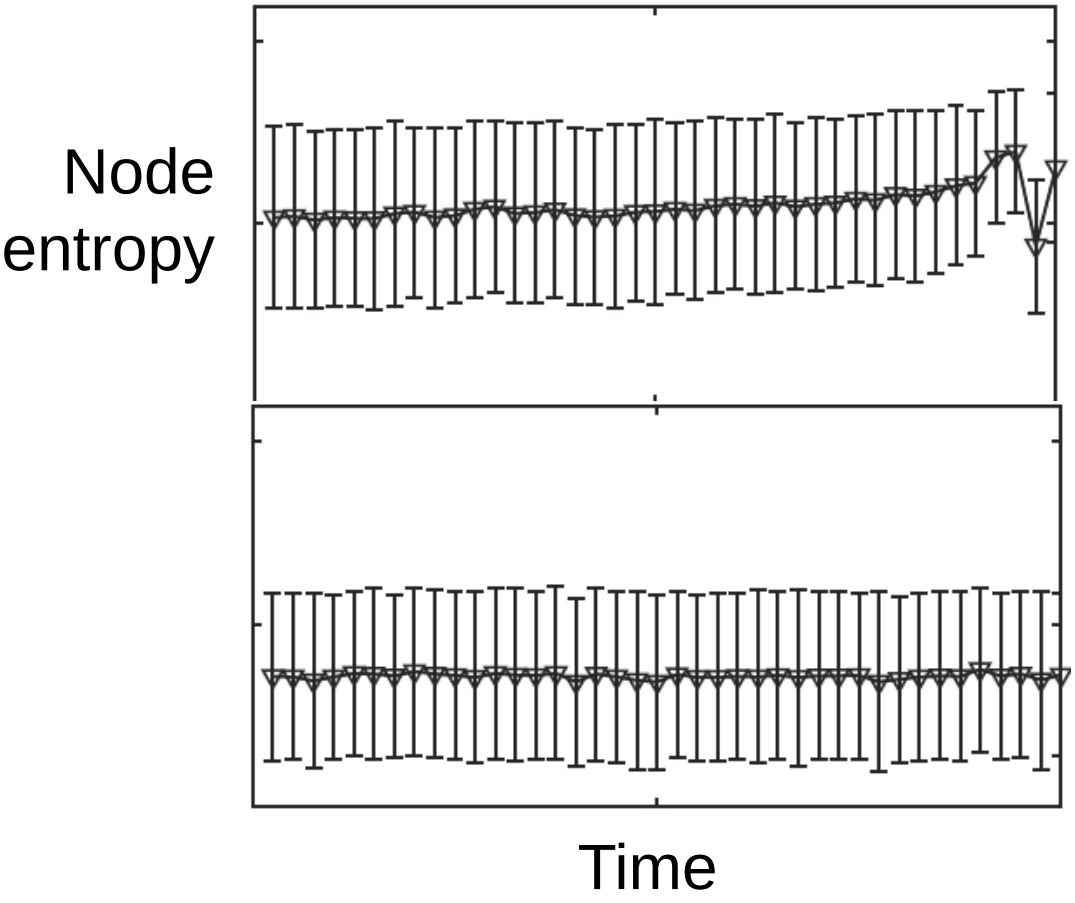


Early warning

⇒ Deterministic mechanisms must be involved.

C. Masoller et al, NJP (2015)

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.

C. Masoller et al, NJP (2015)

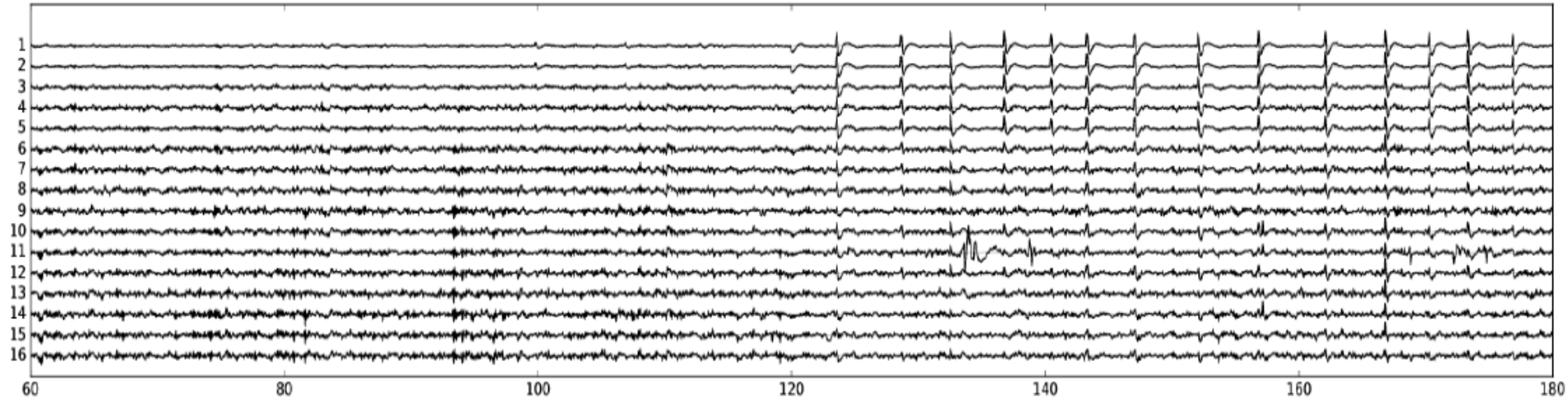
Second application of the ordinal network method: distinguishing *eyes closed* and *eyes open* brain states

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

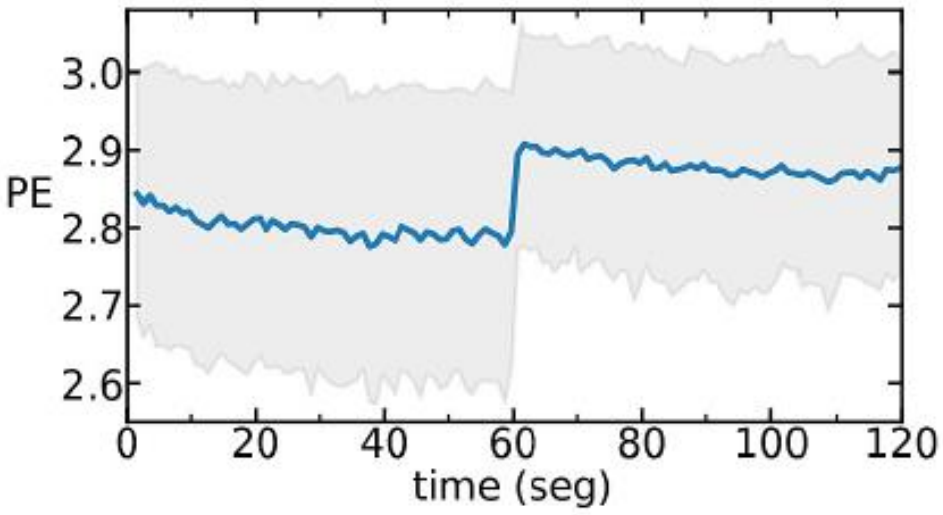
Eye closed

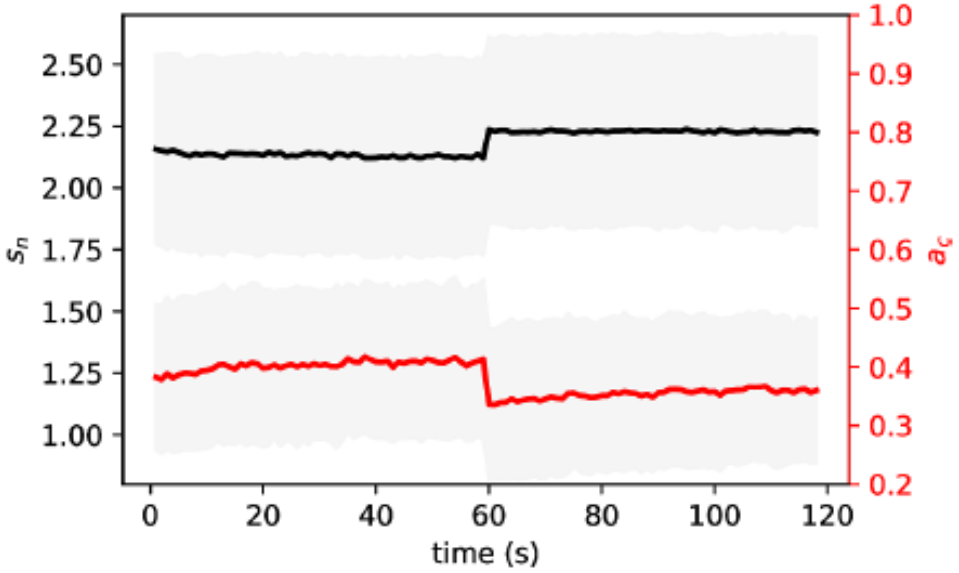
Eye open



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

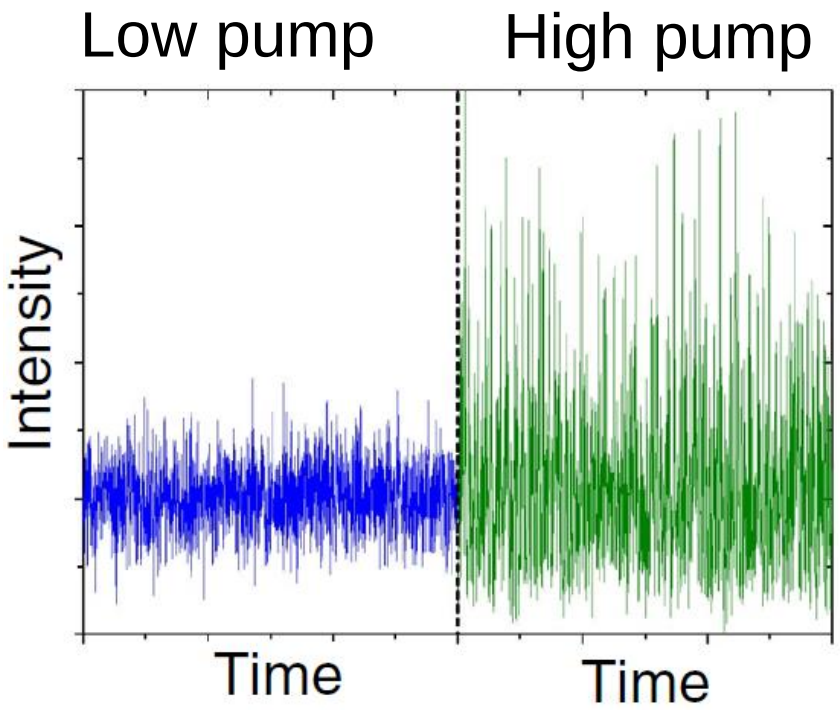


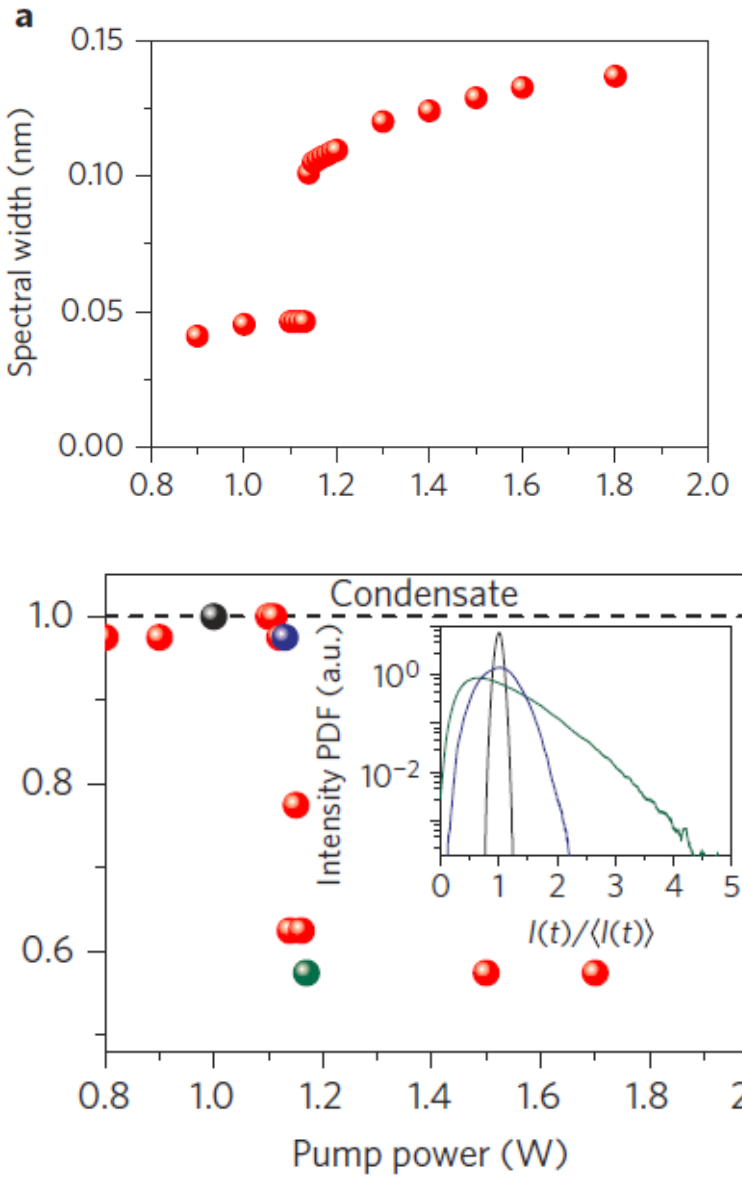


Time window = 1 s
(160 data points)

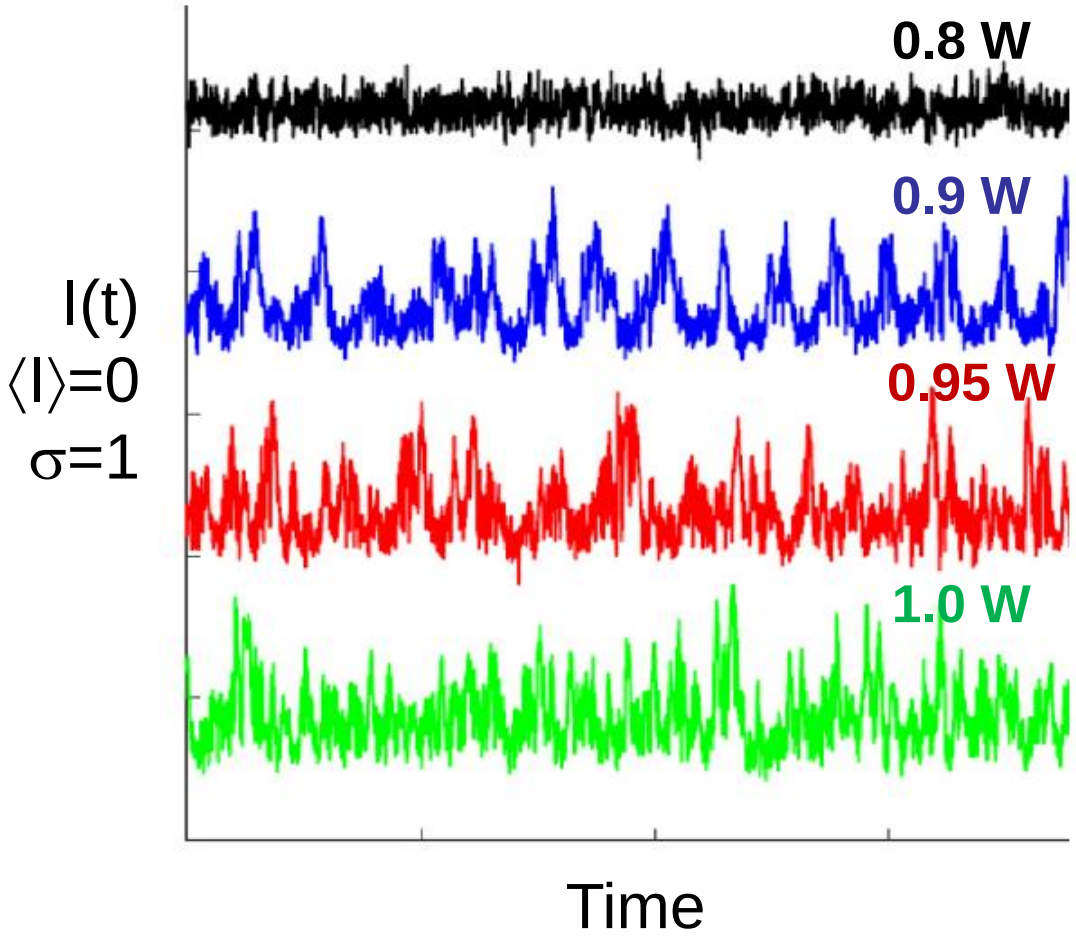
C. Quintero-Quiroz et al, “*Differentiating resting brain states using ordinal symbolic analysis*”, Chaos 28, 106307 (2018).

Example of application: Laminar → Turbulence transition in a fiber laser as the pump (control parameter) increases



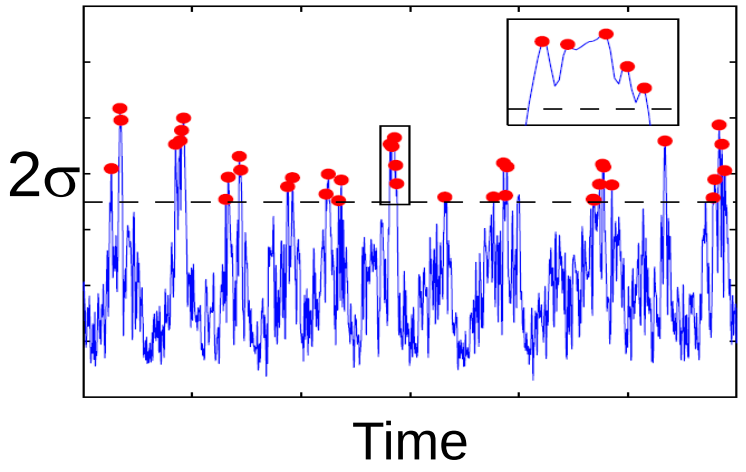


E. G. Turitsyna et al Nat. Phot. 7, 783 (2013)



Nonlinear temporal correlations?

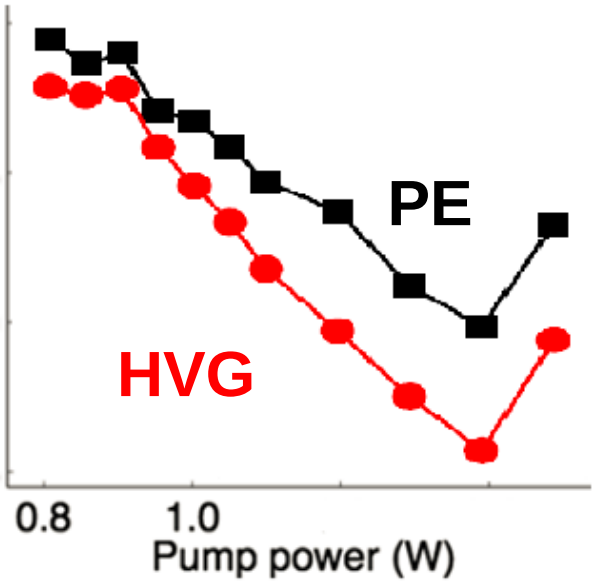
Raw and **thresholded** data



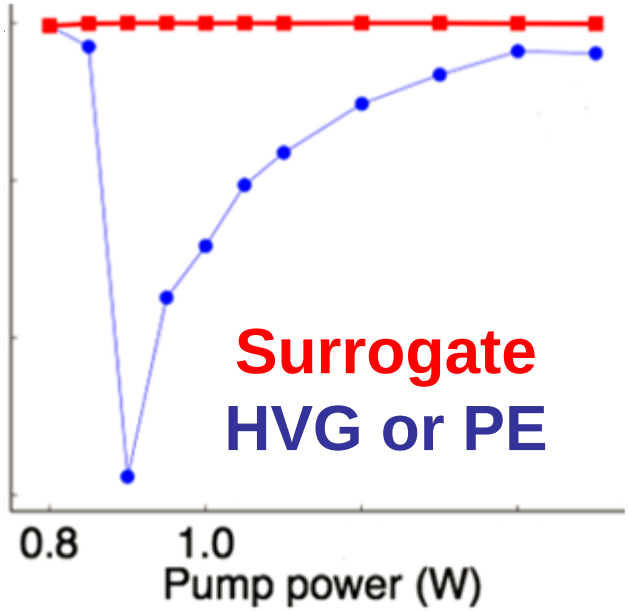
L. Carpi and C. Masoller, "*Persistence and stochastic periodicity in the intensity dynamics of a fiber laser during the transition to optical turbulence*", Phys. Rev. A **97**, 023842 (2018).

Comparison of different ways to compute the entropy

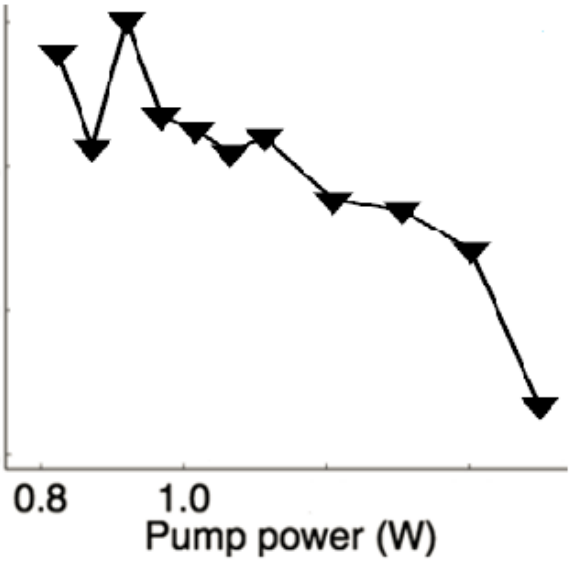
PE/HVG from "raw" data



"thresholded" data



Histogram of "raw" values

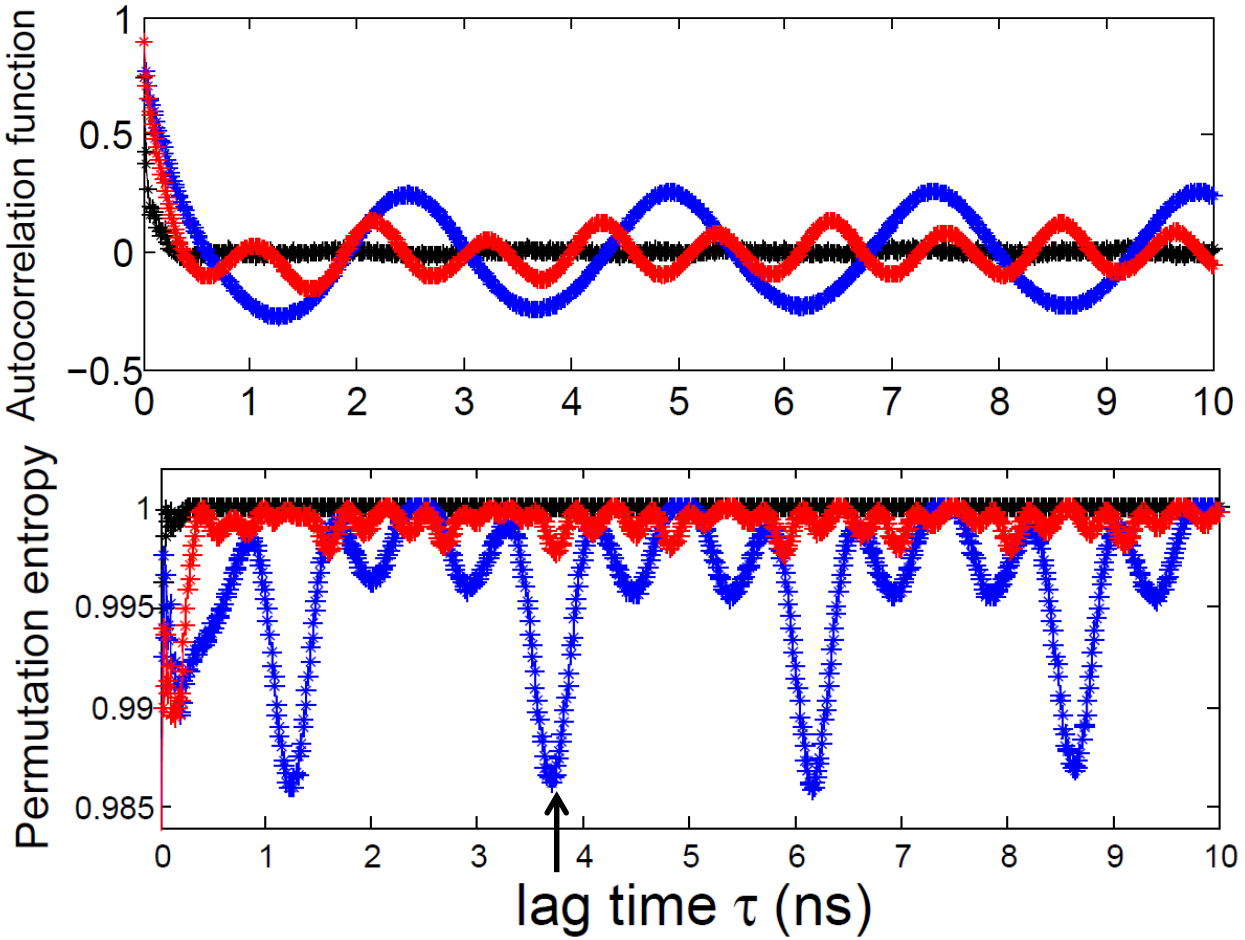


(the abrupt transition is robust with respect to the selection of the threshold)

Aragoneses et al, PRL (2016)

Changing τ identifies ``hidden'' periodicity

$$\{I_i, I_{i+\tau}, I_{i+2\tau}, \ldots\}$$



$$C(\tau) = \frac{\left\langle \left[I_i - \langle I \rangle \right] \left[I_{i-\tau} - \langle I \rangle \right] \right\rangle}{\sigma^2}$$

Permutation entropy

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

Aragoneses et al, PRL (2016)

Space-time representation of a time series

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

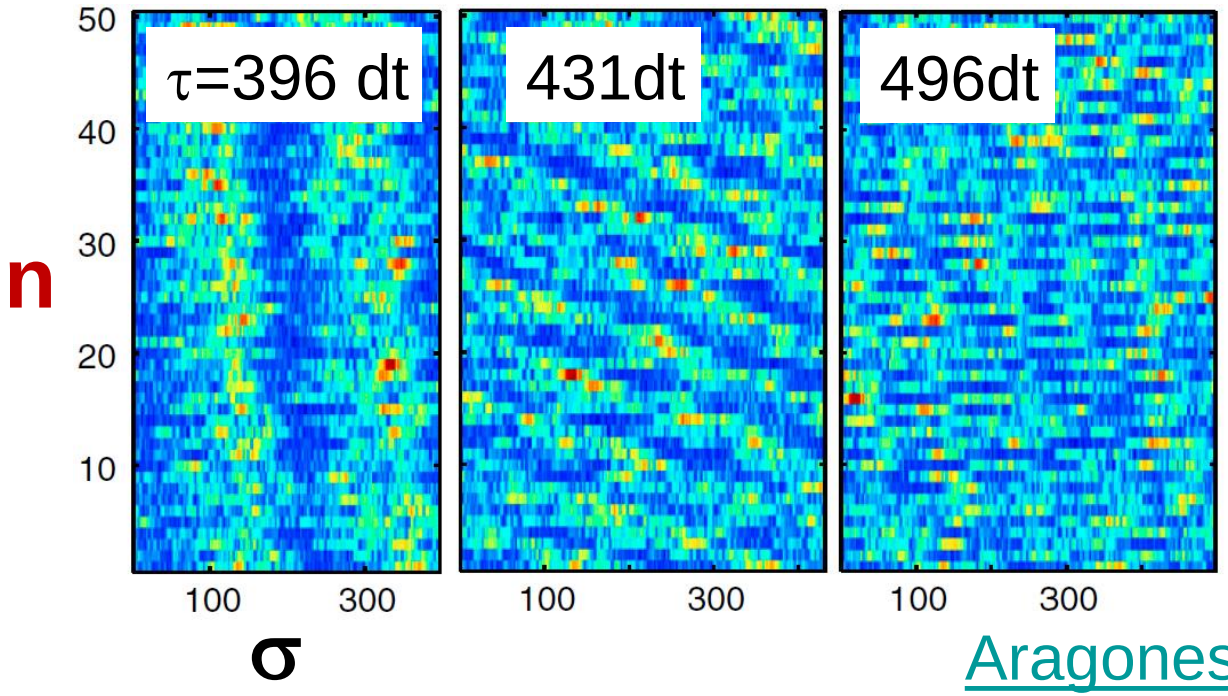
$\{I_1, I_2, \ldots I_\tau, I_{\tau+1}, \ldots\}$ ➡

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ I_{2\tau+1} & I_{2\tau+2} & \cdots & I_{3\tau} \\ I_{\tau+1} & I_{\tau+2} & \cdots & I_{2\tau} \\ I_1 & I_2 & \cdots & I_\tau \end{bmatrix}$$

Color scale: I_i

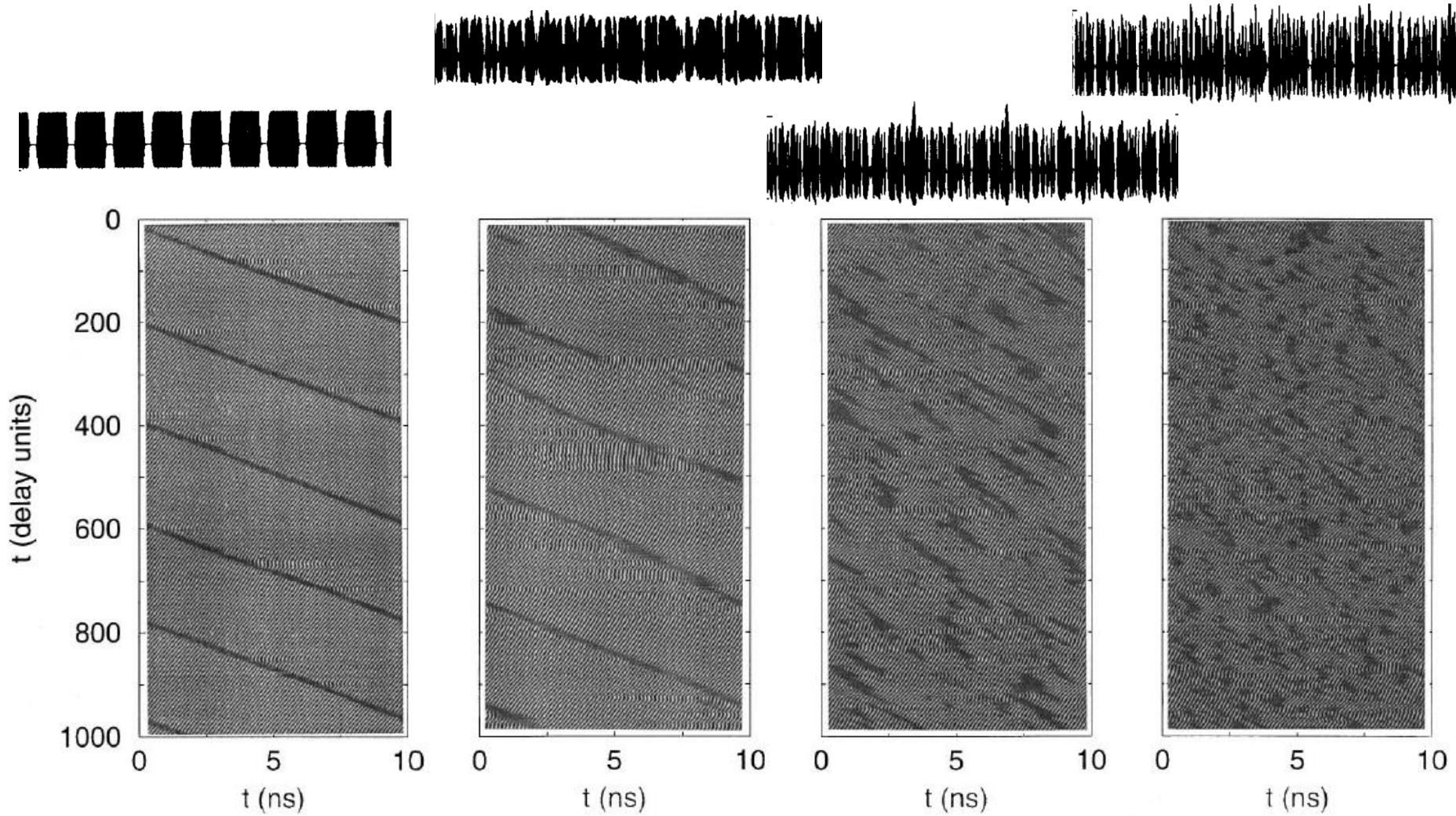
↑ n

σ →



Aragoneses et al, PRL (2016)

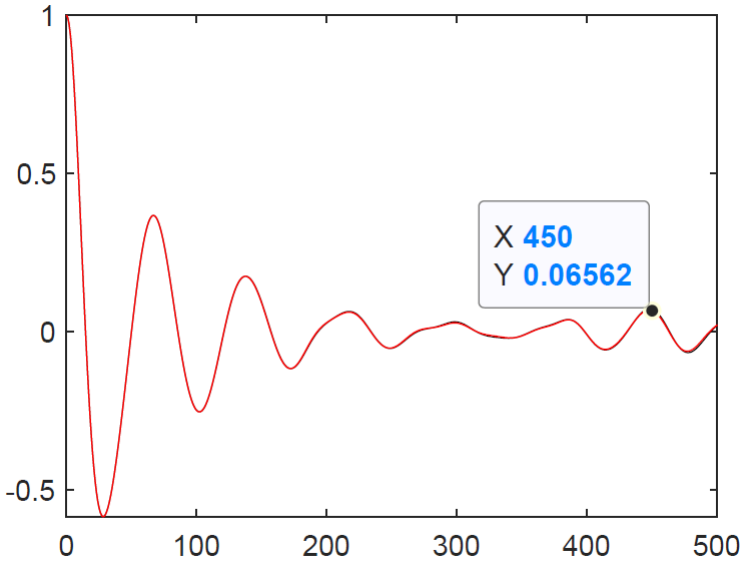
Another example: spatio-temporal representation of the intensity of a laser with optical feedback



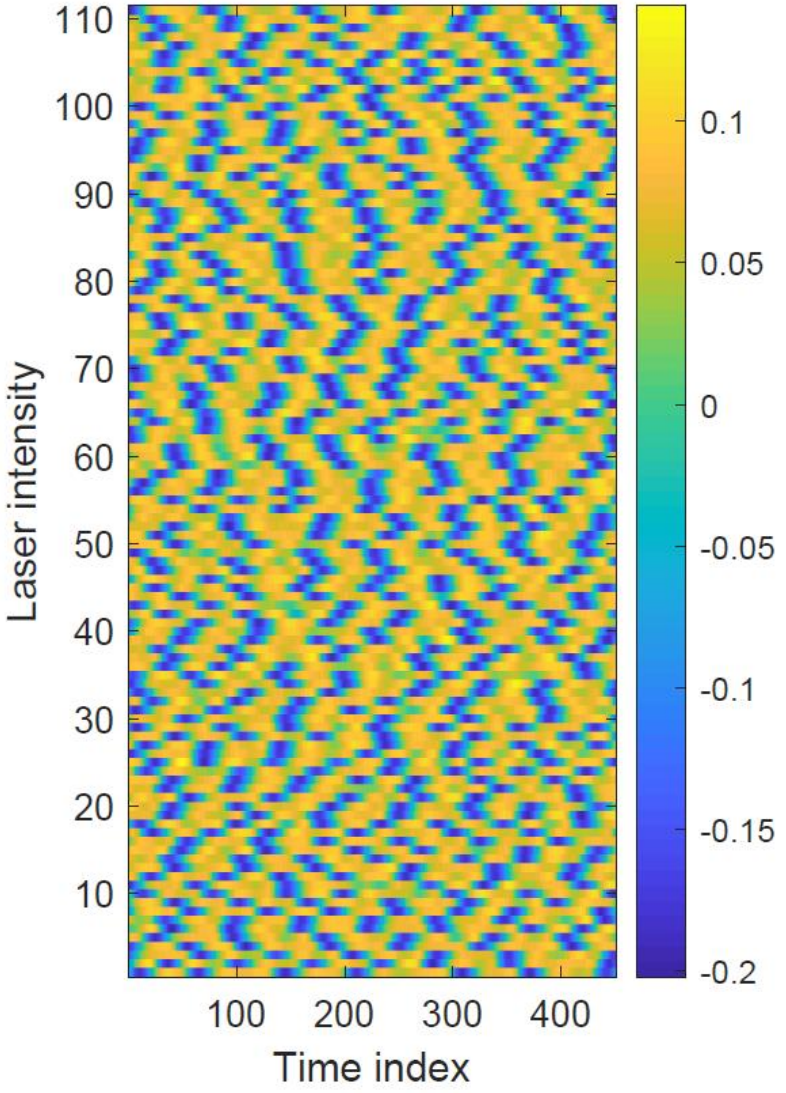
C. Masoller, Chaos 7, 455 (1997)

delay time: 10 ns

Hands-on exercise 7: space-time representation of the laser intensity



Matlab command: imagesc

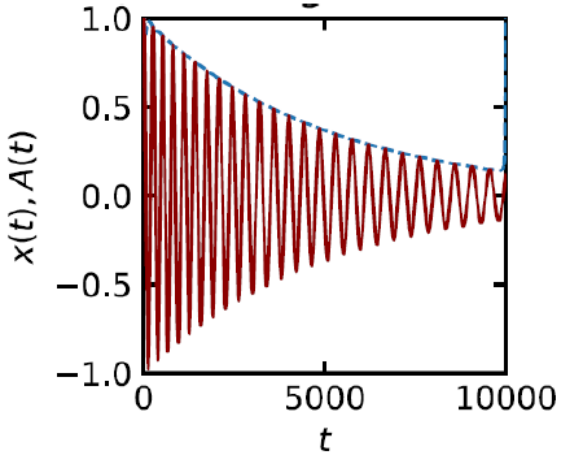


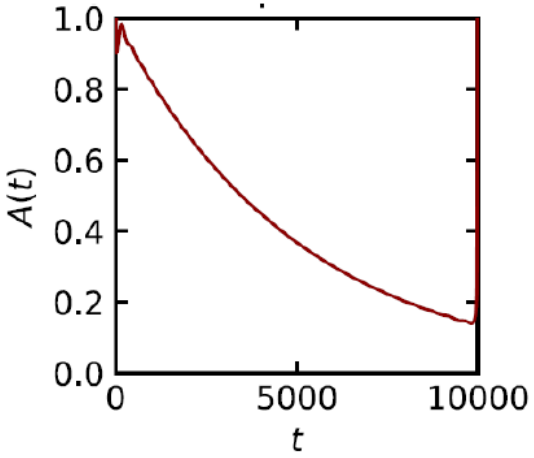
Methods of time-series analysis

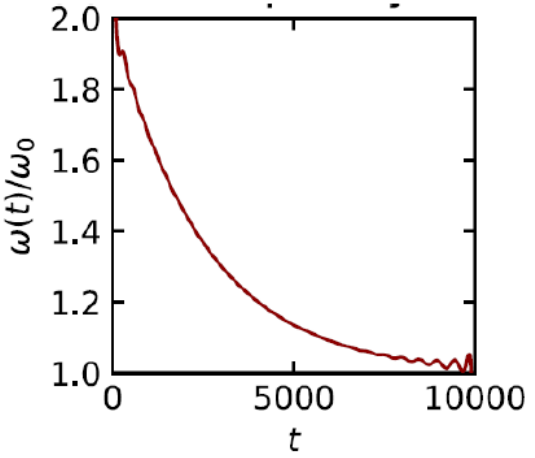
- Return maps
- Distribution of data values
- Correlation and Fourier analysis
- Stochastic models and surrogates
- Attractor reconstruction: Lyapunov exponents and fractal dimensions
- Symbolic methods
- Information theory measures: entropy and complexity
- Network representation of a time-series
-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

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

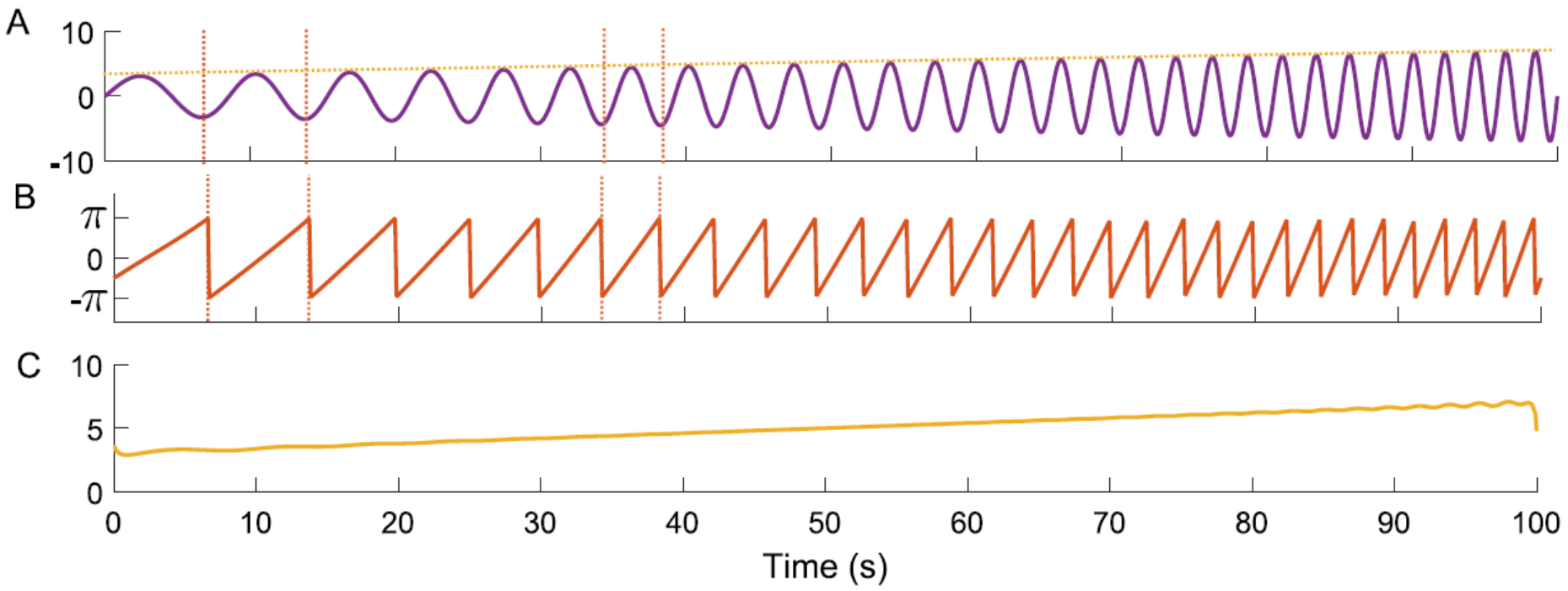
$$x(t) = e^{-\alpha t}\cos[\omega_0(1 + e^{-2\alpha t})]$$







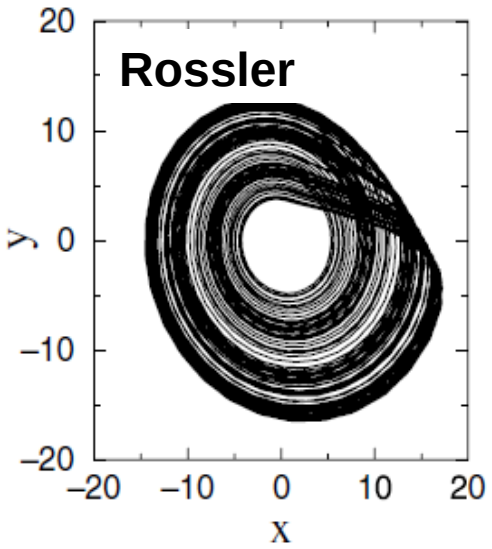
Example: sine wave with increasing amplitude and frequency

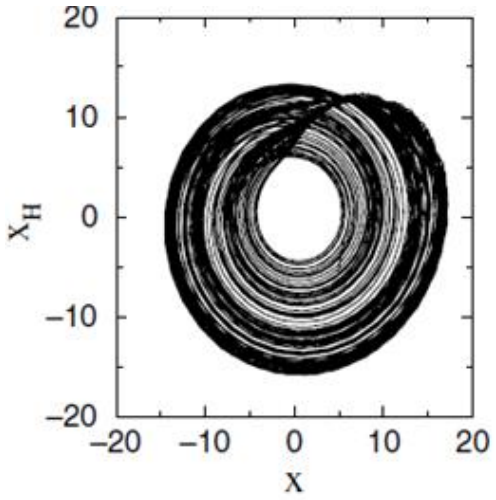


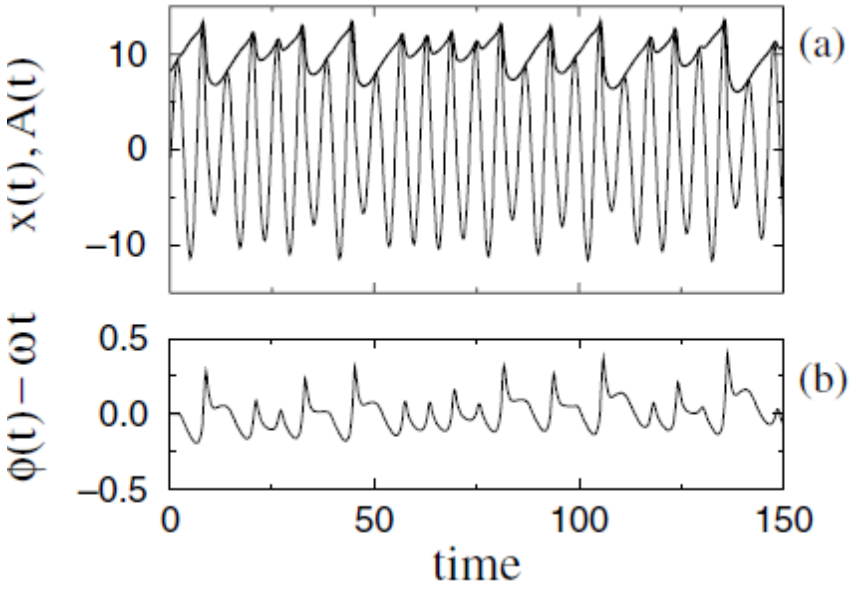
- (A) The original signal. (B) The instantaneous phase extracted using the Hilbert transform. (C) The instantaneous amplitude.
- A = C cos(B).

G. Lancaster et al, Physics Reports 748 (2018) 1–60

Second example







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

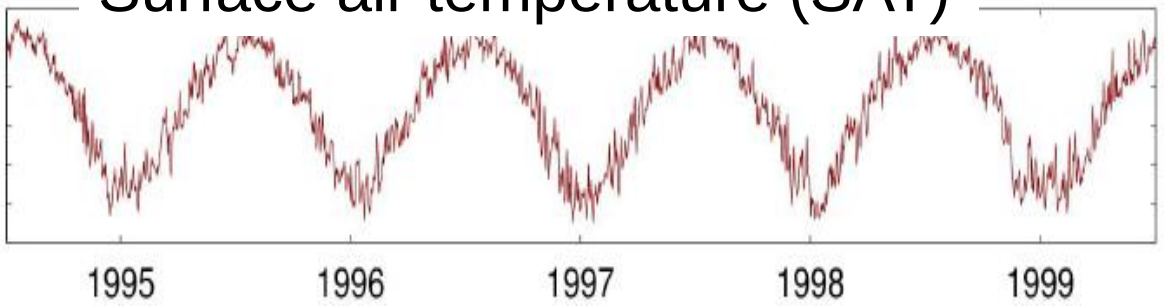
A. Pikovsky, M. Rosenblum and J. Kurths 2001

Third example

Surface air temperature (SAT)

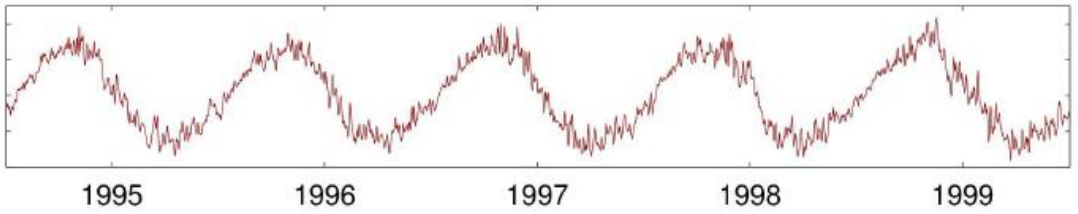
Normalization: $\mu=0$, $\sigma=1$

x

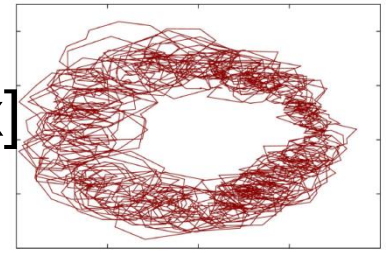


- HT[sin(ωt)]=cos(ωt)

HT[x]



y=HT[x]



x

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

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

Zappala, Barreiro and Masoller, Entropy (2016)

Hilbert transform

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

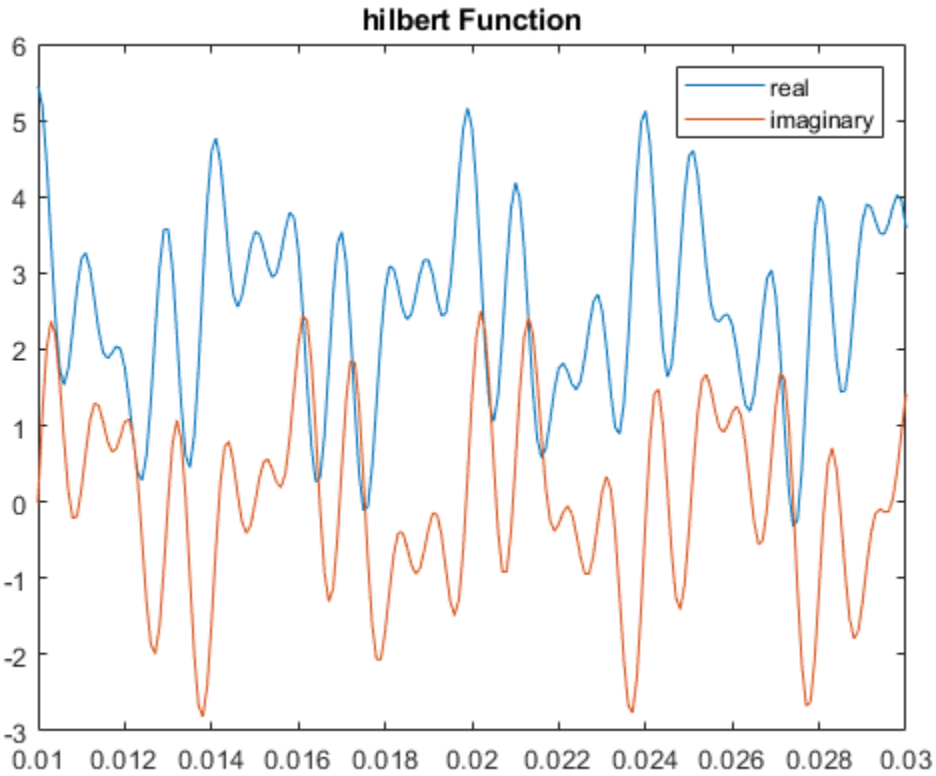
$$\zeta(t) = x(t) + iy(t) = a(t)e^{i\varphi(t)}$$

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

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')
```

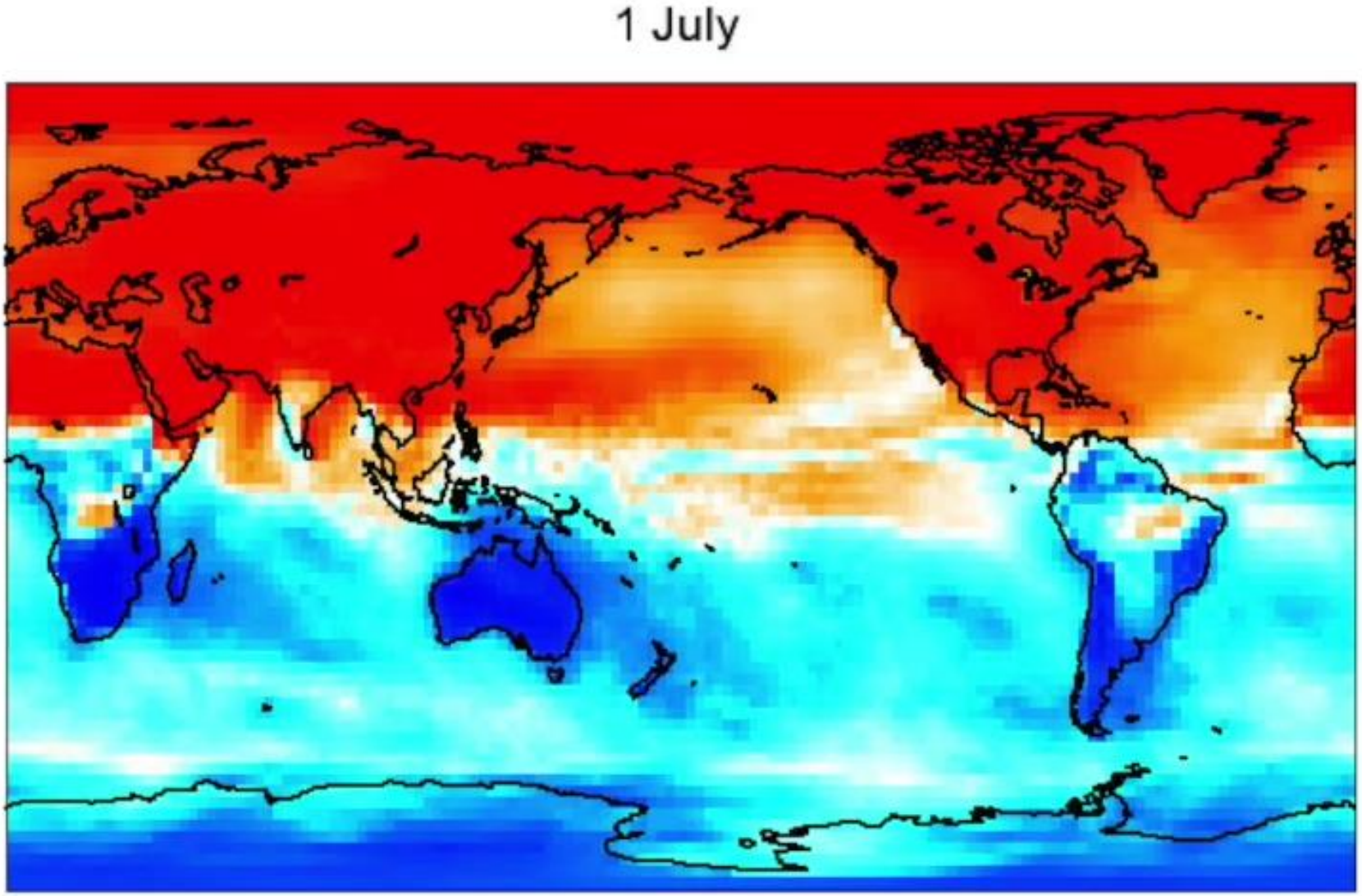


The sampling rate must be chosen in order to have at least 20 points per characteristic period of oscillation.

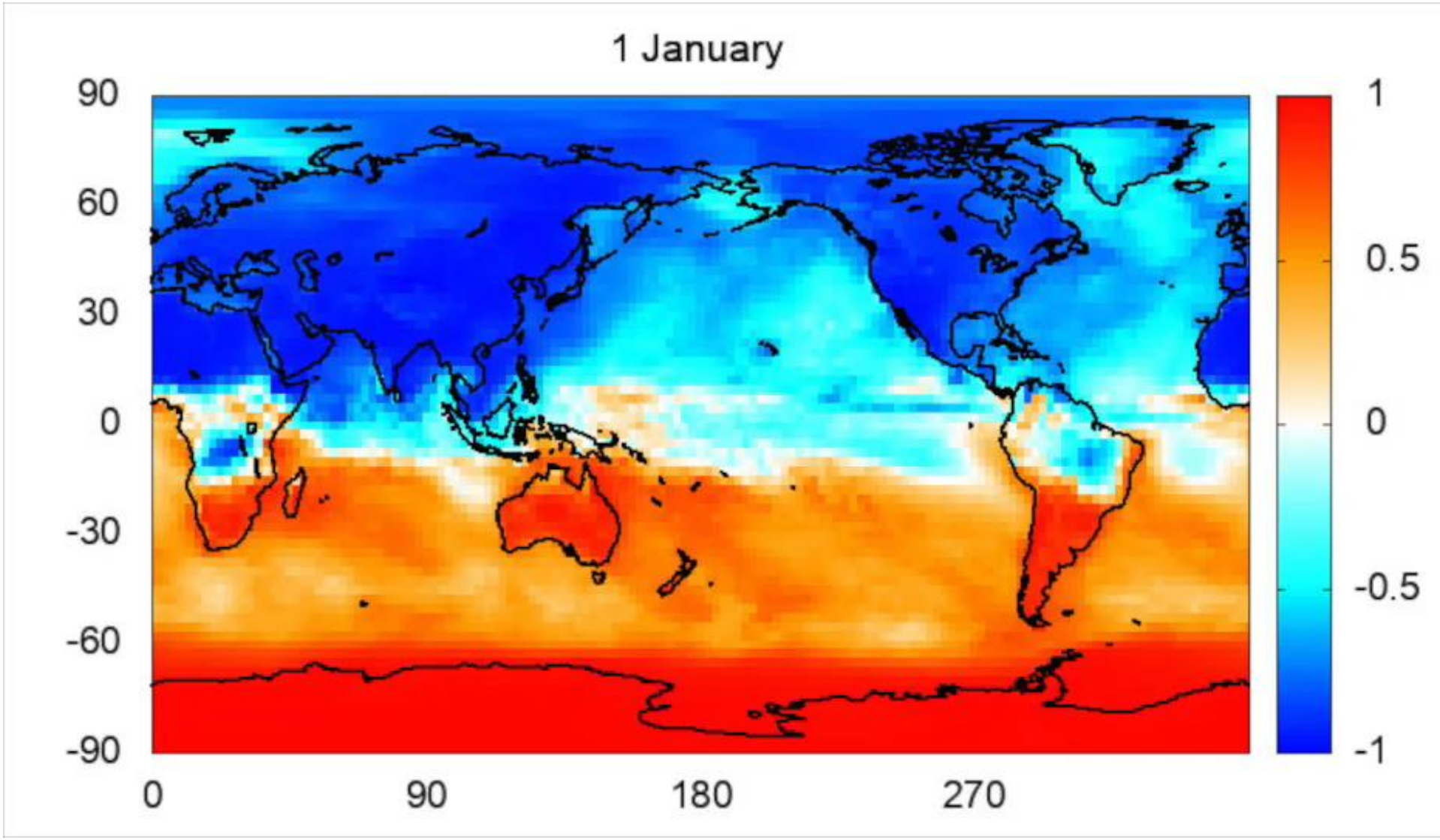
Application to climate data

- 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.

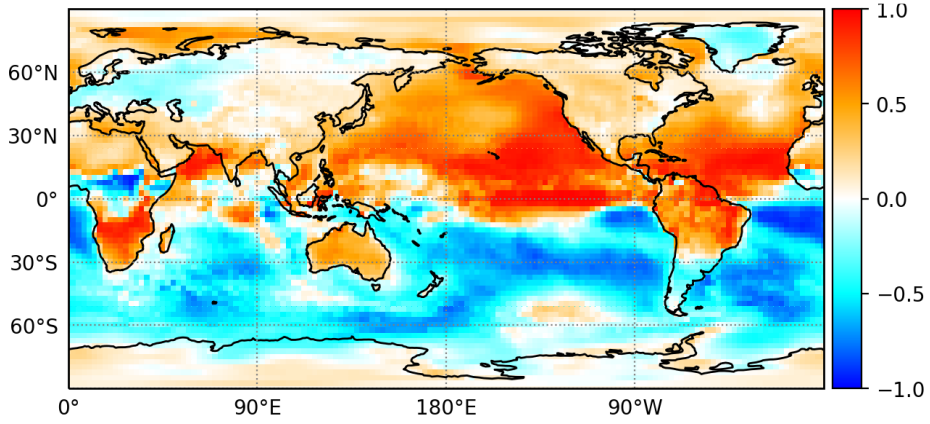
Cosine of Hilbert phase



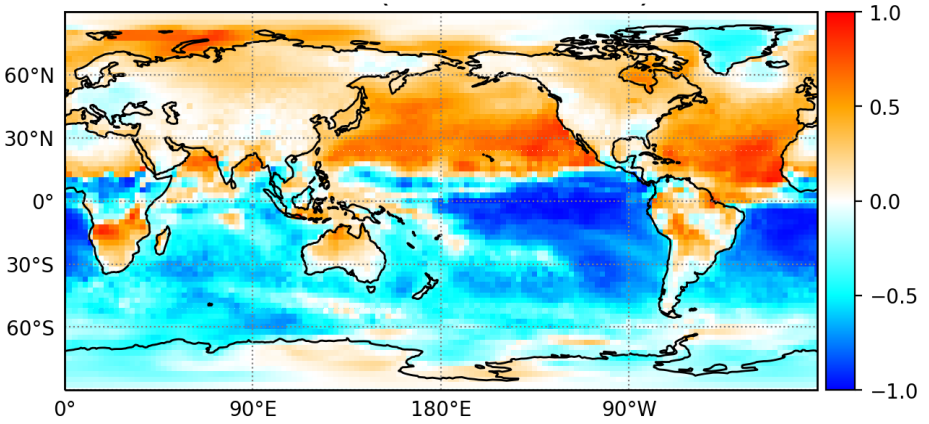
How the seasons evolve?
Temporal evolution of the cosine of the Hilbert phase



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 2.5^0 x $2.5^0 \Rightarrow$ 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, σ_a, σ_ω

Relative decadal variations

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

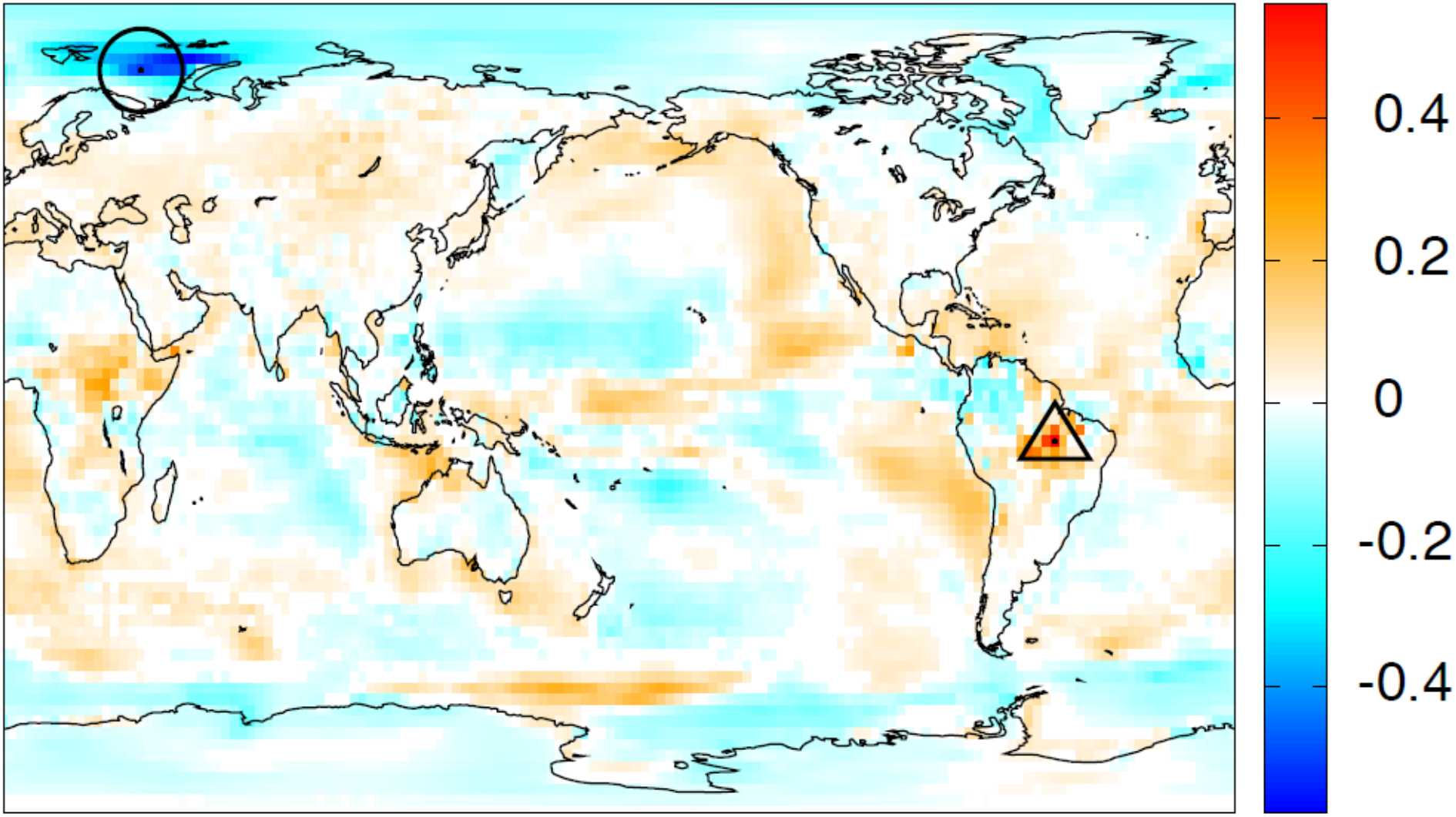
$$\frac{\Delta a}{\langle a \rangle_{2016-1979}}$$

Relative variation is considered significant if:

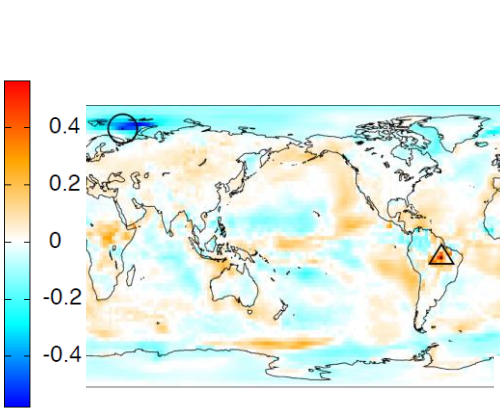
$$\frac{\Delta a}{\langle a \rangle} \geq \langle . \rangle_s + 2\sigma_s \quad \text{or} \quad \frac{\Delta a}{\langle a \rangle} \leq \langle . \rangle_s - 2\sigma_s$$

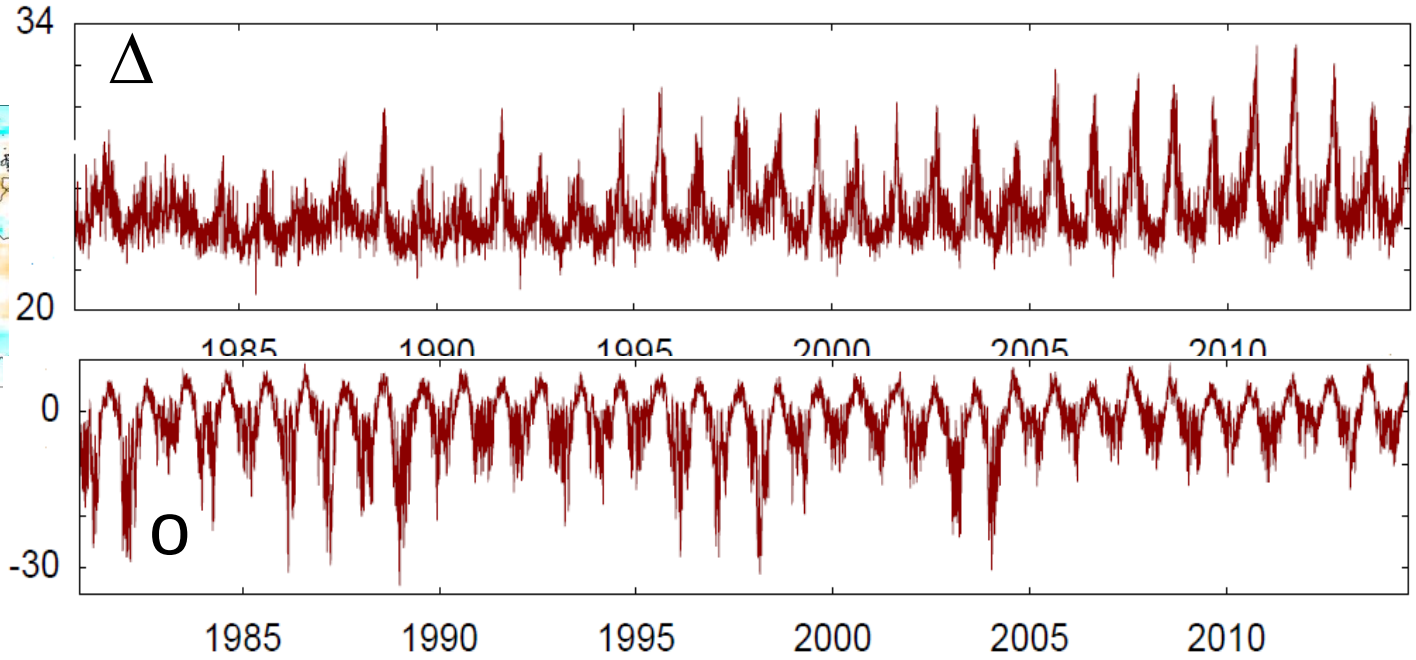
100 “block” surrogates

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)


0.4
0.2
0
-0.2
-0.4

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.

D. A. Zappala et al., Earth Syst. Dynam. **9**, 383 (2018)

And many more time series analysis methods

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

Take home messages

- Linear tools (correlation, Fourier, distribution of datapoints) and nonlinear tools (symbolic analysis, networks, spatiotemporal representations, etc.) are useful for investigating complex signals.

- Different tools 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…"

Bradley and Kantz, CHAOS 25, 097610 (2015)