# NONLINEAR TIME SERIES ANALYSIS, WITH APPLICATIONS TO MEDICINE

José María Amigó

Centro de Investigación Operativa, Universidad Miguel Hernández, Elche (Spain)

# LECTURE 5

## NONLINEAR METHODS IN MEDICINE II: STUDY CASES

# OUTLINE

# Applications to Medicine

**STUDY CASE 1: Information content in spike trains**[1]

*Peculiarities of the spike trains:*

- Analog (or continuous) signals
- None-or-all signals

They result from very complex interactions $\Rightarrow$ random (point) processes

[1] J.M. Amigó et al., Neural Computation 16 (2004) 717.

**Discretization of time and quantification:**

- If (i) $T =$ duration of the signal, (ii) $Dt =$ duration of the bin,

$$N = \frac{T}{Dt} = \text{ length of the time series}$$

- $Dt$ is the precision of the measurement.
- The quantification $x_1, ..., x_N$ depends on $Dt$.

# 1. Information content in spike trains

**Neuron = Information source.**

- *Probabilities estimates* from $x_1, ..., x_N$ = frequencies in sliding windows of size $L$:

$$\hat{p}(a_1, ..., a_L) = \frac{\#\{n : (x_n, ..., x_{n+L-1}) = (a_1, ..., a_L)\}}{2^L}$$

- *Entropy rate of order* $L$:

$$h(L, Dt) = -\frac{1}{L} \sum \hat{p}(a_1, ..., a_L) \log_2 \ \hat{p}(a_1, ..., a_L)$$

- *Entropy rate (information per symbol)*:

$$h(Dt) = \lim_{L \to \infty} h(L, Dt)$$

# 1. Information content in spike trains

Alternatively, we can estimate $h(Dt)$ via **LZ76 complexity**:

$$\lim_{N\to\infty} \frac{c(x_1^N, Dt)}{Dt} = h(Dt) \quad \text{with probability 1.}$$

# 1. Information content in spike trains

**Experimental work**.

1. *Intracellular periodic current injection in vivo* (frequency $= 2$ Hz, $0.2 \leq i \leq 1.5$ nA, $15.56$ to $47.64$ sec)
2. *Visual stimulation with sinusoidal drifting gratings* ($15.87$ to $23.62$ sec)
3. *Intracellular random current injection in vitro* ($-1.5 \leq i \leq 1.5$ nA, $16.32$ to $35.47$ sec)

**Results on $h(Dt)$.**

- *For periodic current injection in vivo*:

| Coding Frequency | Standard | Complexity |
|---|---|---|
| **100 Hz** | 41.38 | 42.93 |
| **200 Hz** | 59.20 | 60.40 |
| **300 Hz** | 68.42 | 67.00 |

- *For visual stimulation*:

| Coding Frequency | Standard | Complexity |
|---|---|---|
| **100 Hz** | 30.30 | 32.78 |
| **200 Hz** | 47.85 | 50.14 |
| **300 Hz** | 62.55 | 62.11 |

**Results on $h(\Delta t)$:**

- *For random current injection in vitro (slow decay)*:

| Coding Frequency | Standard | Complexity |
|---|---|---|
| **100 Hz** | 52.38 | 53.38 |
| **200 Hz** | 68.69 | 67.23 |
| **300 Hz** | 78.00 | 74.70 |

- *For random current injection in vitro (fast decay):*

| Coding Frequency | Standard | Complexity |
|---|---|---|
| **100 Hz** | 22.31 | 19.00 |
| **200 Hz** | 27.75 | 24.39 |
| **300 Hz** | 31.05 | 26.03 |

# Applications to Medicine

**STUDY CASE 2: Coupling directionality and neuronal signals**.

Let $X = (x_n)$ and $Y = (y_n)$ be two neuronal signals recorded from different brain areas.

- **Question:** *In which direction is information flowing?*

The mutual information,

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

is useless because

$$I(X,Y) = I(Y,X)$$

# 2. Coupling directionality and neuronal signals

Use the idea behind the *Granger causality*: If

(i) $X_\delta = (x_{n+\delta})$, $Y_\delta = (y_{n+\delta})$, and

(ii) information flows from the process $Y$ to $X$ at some later time,

then

$$H(X_\delta \,|X, Y) < H(X_\delta \,|X)$$

for some $\delta > 0$.

**Definition.** *Transfer entropy*

$$\begin{aligned} I^{\delta}_{Y\to X} &\equiv H(X_\delta\,|X) - H(X_\delta\,|X,Y) = I(X_\delta, Y\,|X) \\ &= \sum p(x_\delta, x, y)\log\frac{p(x_\delta, y\,|x)}{p(x_\delta\,|x)\,p(y\,|x)}. \end{aligned}$$

Hence,

$$\boxed{\text{if } Y \rightsquigarrow X, \text{ then } I^{\delta}_{Y\to X} > 0 \text{ for some } \delta > 0.}$$

Similarly one defines

$$I^{\delta}_{X\to Y} \equiv I(Y_\delta, X\,|Y) = \sum p(y_\delta, x, y)\log\frac{p(y_\delta, x\,|y)}{p(y_\delta\,|y)\,p(x\,|y)},$$

$$\boxed{\text{If } X \rightsquigarrow Y, \text{ then } I^{\delta}_{X\to Y} > 0 \text{ for some } \delta > 0.}$$

# 2. Coupling directionality and neuronal signals

In practice one uses averages,

$$I_{Y\to X} = \frac{1}{N}\sum_{\delta=1}^{N} I^{\delta}_{Y\to X},\; I_{X\to Y} = \frac{1}{N}\sum_{\delta=1}^{N} I^{\delta}_{Y\to X},$$

where $N$ is some convenient number of later points.

**Definition.** The *directionality index* between $X$ and $Y$ is

$$D_{XY} = \frac{I_{X\to Y} - I_{Y\to X}}{I_{X\to Y} + I_{Y\to X}} \in [-1, +1]$$

Then

| | | |
|---|---|---|
| $D_{XY} > 0$ | $\Rightarrow$ | $X$ drives $Y$ |
| $D_{XY} = 0$ | $\Rightarrow$ | symmetrical coupling |
| $D_{XY} < 0$ | $\Rightarrow$ | $Y$ drives $X$ |

# 2. Coupling directionality and neuronal signals

If you use ordinal symbolic dynamics, $I_{X\to Y}$ and $I_{Y\to X}$ are called *permutation transfer entropy*.

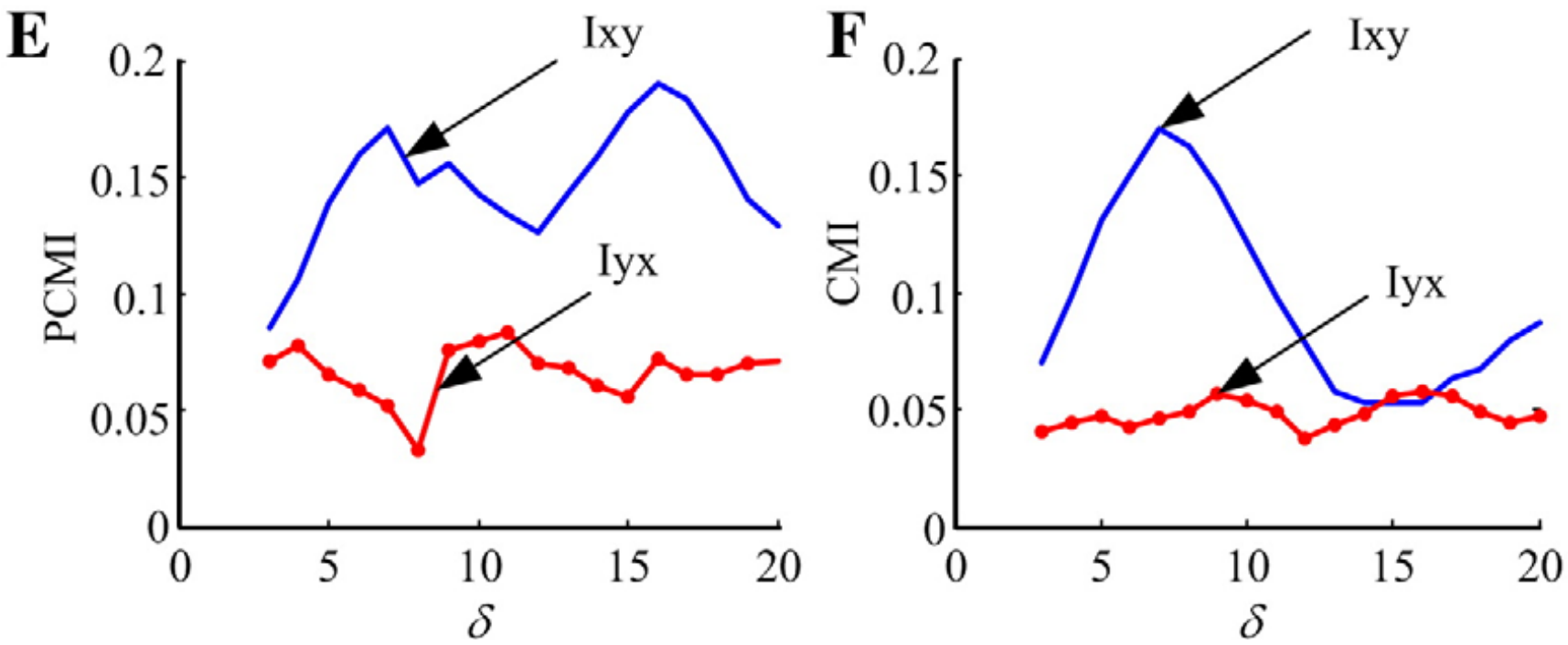
X. Li & G. Ouyang[2] compared *conventional* and *permutation* TE both with numerical models and *intracranial* EEG recorded in the CA1/CA3 hippocampus region of rats.

They confirmed that at the formation of a CA1-CA3 epileptic seizure, the coupling is unidirectional.

[2] X. Li and G. Ouyang, Neuroimage 52 (2010) 497.

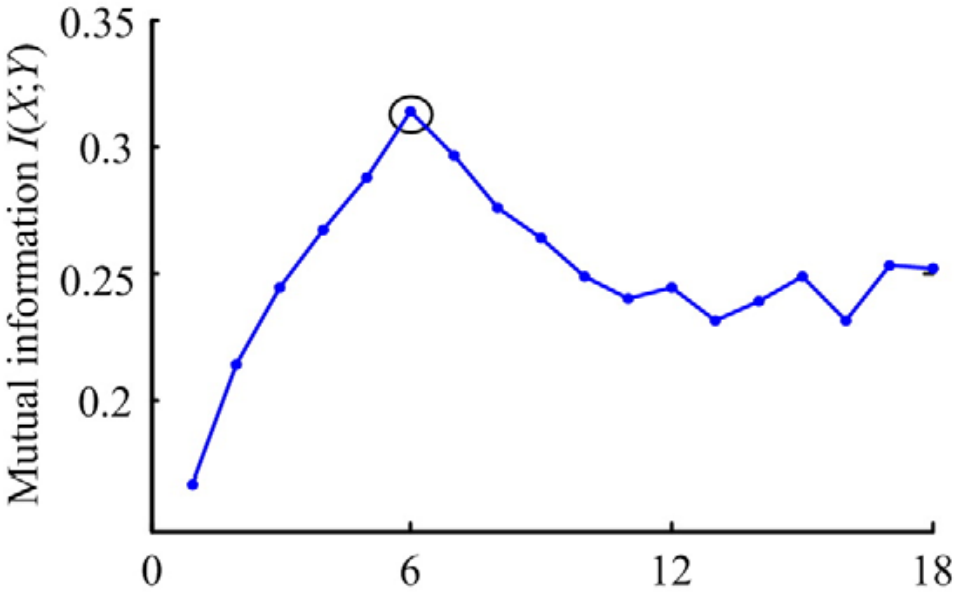
# 2. Coupling directionality and neuronal signals

(X. Li and G. Ouyang)



# 2. Coupling directionality and neuronal signals

The lag $\tau$ was chosen such that $I(X, Y)$ is maximal



# 2. Coupling directionality and neuronal signals

**Advantages** of the permutation transfer entropy[3]:

1. Computationally fast
2. Invariant wrt monotonous transformations
3. Robust against additive and multiplicative noise
4. Does not require long segments of data
5. PTE is superior to conventional TE and Granger causality for identifying the coupling direction between neuronal networks.

[3] X. Li and G. Ouyang, Neuroimage 52 (2010) 497.

# Study case 3

**STUDY CASE 3: Characterization of EEG and ECG**

*Historical background*.

- (**1985**) P.E. Rapp et al, Dynamics of spontaneous neural activity in the simian motor cortex: the dimension of chaotic neurons, *Phys. Lett*. 110, 110
- (**1985**) A. Babloyantz et al., Evidence of chaotic dynamics of brain activity during the sleep cycle, *Phys. Lett A* 111, 152.

**Scope of nonlinear time series analysis**: *Extract information*

But, what information?

# 3. Characterization of EEG and ECG

Very often the analyst only needs to discriminate different kinds of dynamics.

**Examples:**

- *Epilepsy*: normal/abnormal
- *Sleep*: various sleep stages
- *Coma and anaesthesia*: difference depths
- *Mental states and psychiatric disease*
- *Disturbed cognition and dementia*: different degrees
- *Cardiac diseases*: normal/abnormal

# 3. Characterization of EEG and ECG

**Example 1**[4]. *Dynamical change during an epileptic seizure*

- One of the first applications of ordinal patterns and permutation entropy
- Paradigmatic example of application of nonlinear TSA to Medicine

[4] Y. Cao et al., Physical Review E 70 (2004) 046217.

# 3. Characterization of EEG and ECG

(Y. Cao et al.)

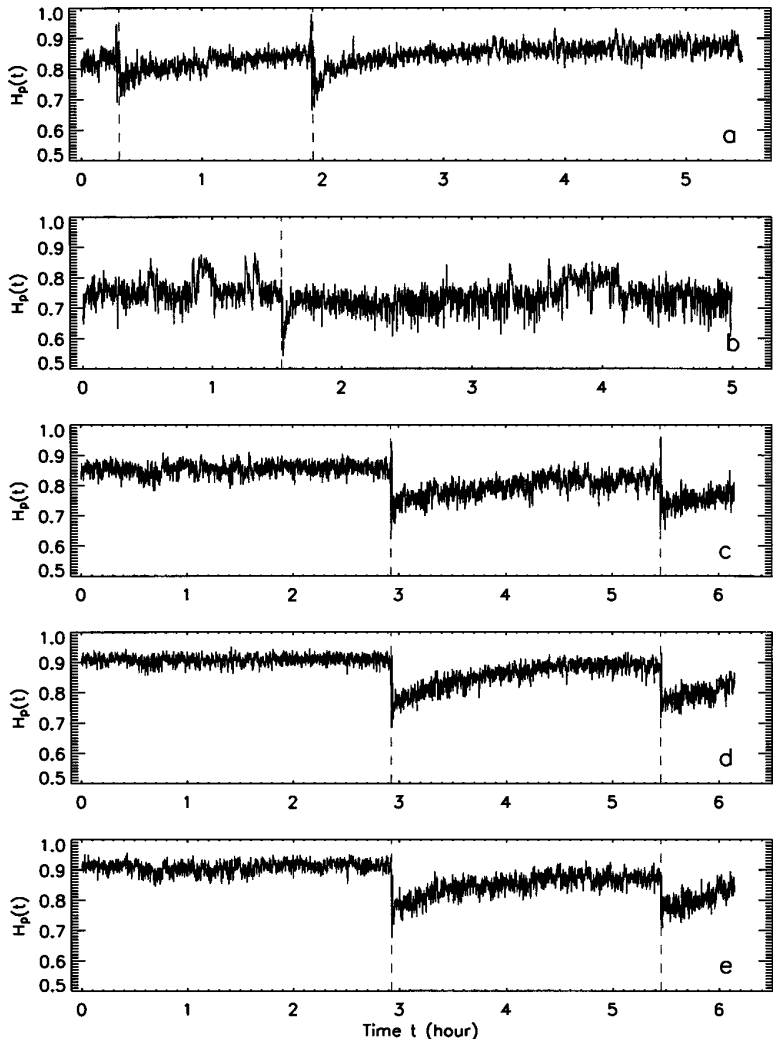
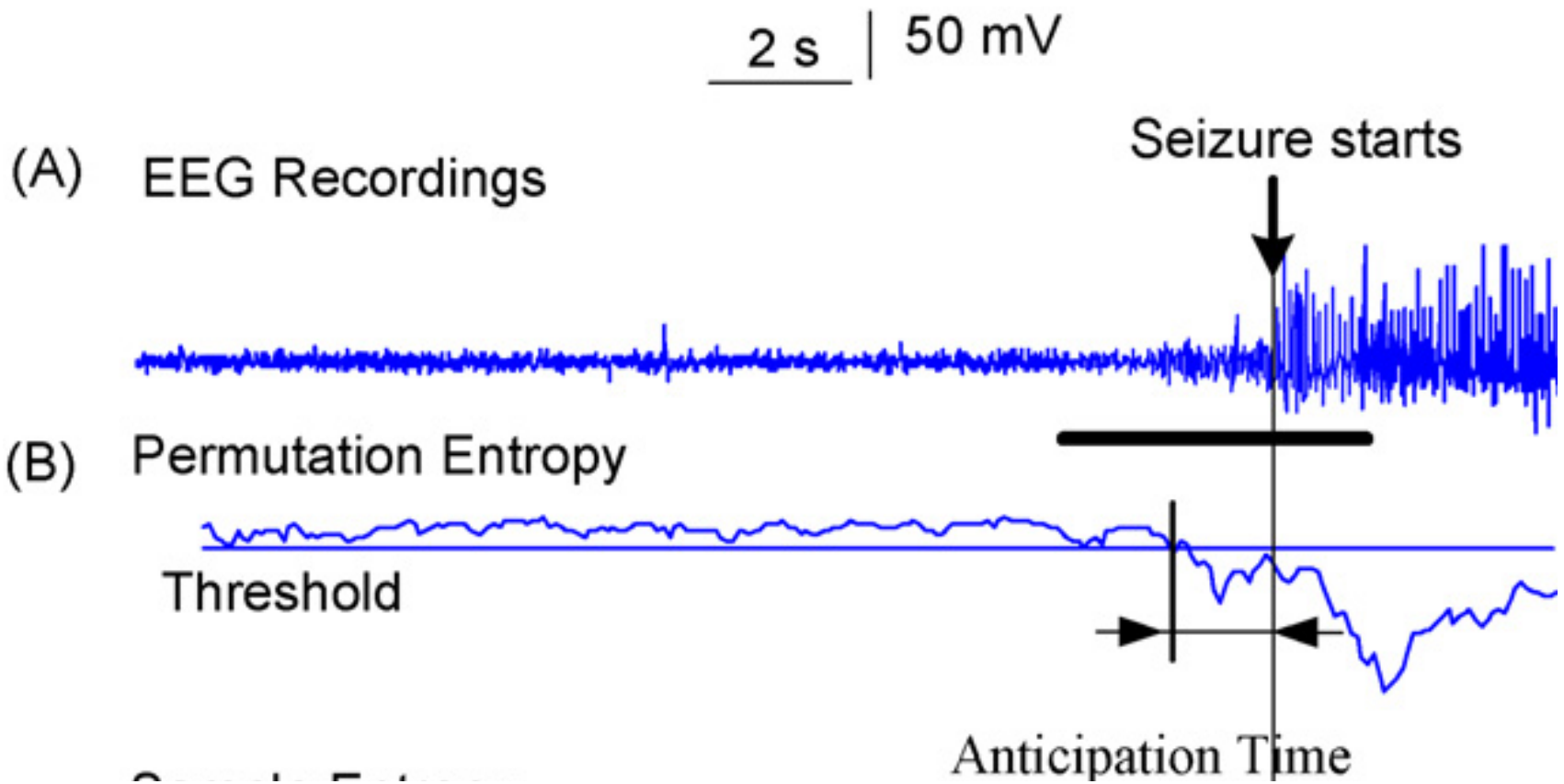



Figure:

**Example 2**[5]. *Prediction of absence seizure*




[5]X. Li and G. Ouyang, Neuroimage 52 (2010) 497.

# 3. Characterization of EEG and ECG

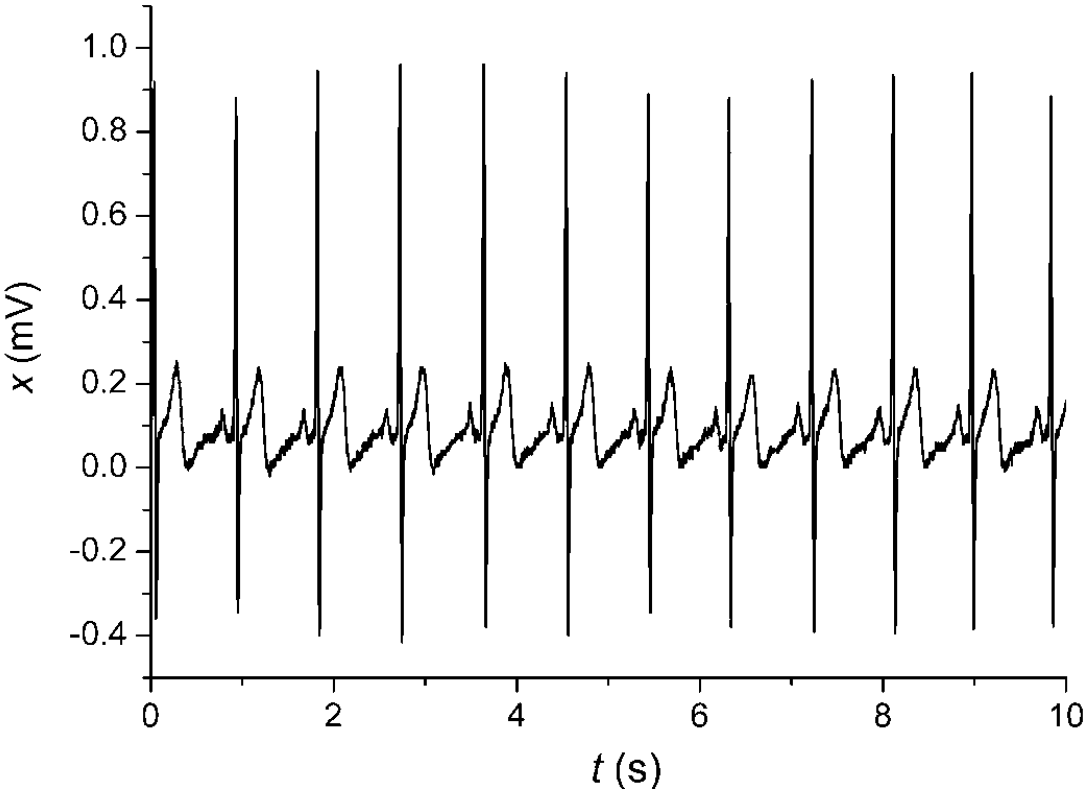
**Example 3**[6]. ECG from patients with congestive heart failure (CHF). U. Parlitz et al. used biomarkers obtained via

- Heart rate variability parameters
- Non-ordinal symbolic dynamics
- Ordinal $3$- and $4$-patterns statistics (with different lags)

for discriminating CHF patients from control groups using beat-to-beat time series $(RR_n)$.

[6] U. Parlitz et al., Comp. Biol. Med. 42 (2012) 319.

# 3. Characterization of EEG and ECG



# 3. Characterization of EEG and ECG

Non-ordinal symbolization used for heart rate variability

- *Binary quantification*:

$$q_n(RR_n) = \begin{cases} 0 & \text{if } |RR_n - RR_{n-1}| < \theta \\ 1 & \text{if } |RR_n - RR_{n-1}| \geq \theta \end{cases}$$

or, to study acceleration and deceleration runs,

$$q_n(RR_n) = \begin{cases} 0 & \text{if } RR_n - RR_{n-1} \geq 0 \\ 1 & \text{if } RR_n - RR_{n-1} < 0 \end{cases}$$

- *Four-symbol quantification*:

$$q_n(RR_n) = \begin{cases} 0 & \text{if } \mu < RR_n \leq (1+a)\mu \\ 1 & \text{if } (1+a)\mu < RR_n < \infty \\ 2 & \text{if } (1-a)\mu < RR_n \leq \mu \\ 3 & \text{if } 0 < RR_n \leq (1-a)\mu \end{cases}$$

# 3. Characterization of EEG and ECG

*The best biomarker of CHF was the mean $\pm$ standard deviation of some ordinal $4$-patterns.*

Other authors are also using *ordinal patterns* as biomarkers to study biomedical time series (G. Graff, K. Keller, G. Ouyang, K. Schindler,...)

# Applications to Medicine

**STUDY CASE 4: Nonlinear analysis of ECG**[7]

*Data acquisition*

- Healthy subject asleep
- Beat rate $\approx 66$ beat/min
- Sampling frequency: 250 Hz ($\Delta t = 0.004$ s)
- Recording time: 3 min ($N = 45000$ data points)

*The full nonlinear analysis involves some 5 basic steps.*

[7] M. Perc, Eur. J. Phys. 26 (2005) 757.

**Step 1: State space reconstruction**

Embedding vectors

$$\mathbf{s}(n) = (s_{n-(m-1)\tau}, ..., s_{n-\tau}, s_n),$$

- False nearest neighbors: $m = 10$.
- Minimum of the mutual information: $\tau = 9$

**Step 2: Noise reduction**

Replace

$$RR_{n+\lfloor m/2 \rfloor \tau} \longleftarrow \frac{1}{|B_\varepsilon(\mathbf{s}(n))|} \sum_{\mathbf{s}(k) \in B_\varepsilon(\mathbf{s}(n))} RR_{n-\lfloor m/2 \rfloor \tau}$$

with $\varepsilon = 2\sigma$ or $\varepsilon = 3\sigma$. Here $\varepsilon = 0.065$.

**Remark.** Some authors interchange steps 1 and 2. In this case, use as a rule

$$1/3 \leq m \cdot \Delta t \cdot \tau \leq 2/3$$

in the noise reduction step.

# 4. Nonlinear analysis of ECG

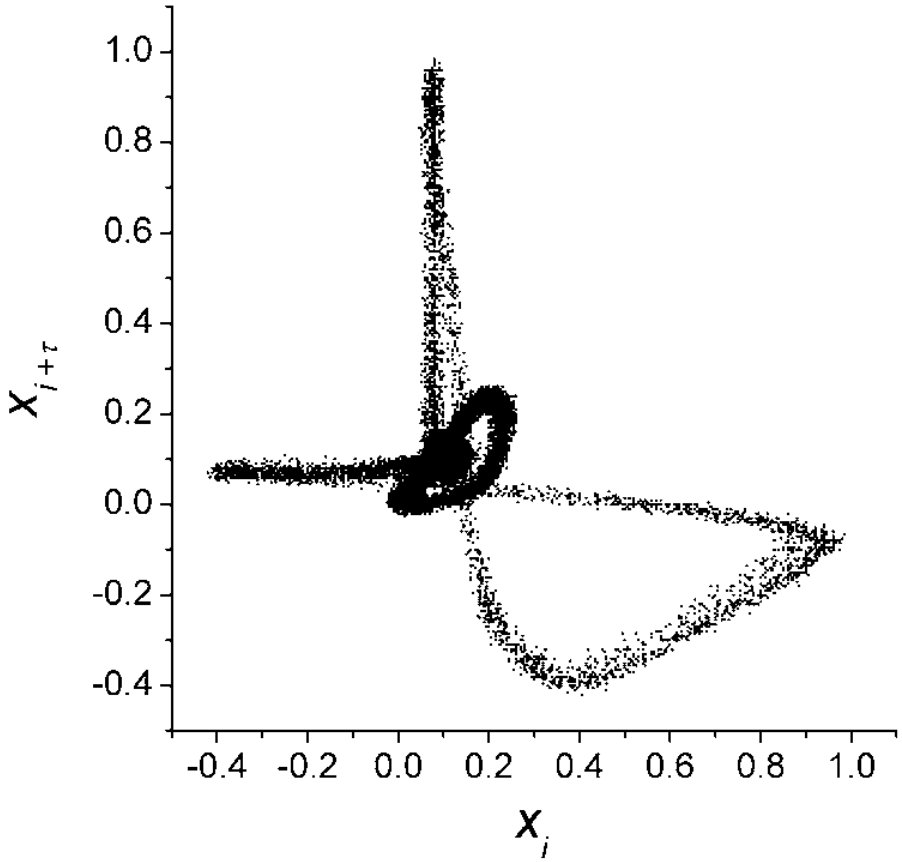



*Figure*. 2D projection of the reconstructed state space with the optimal parameters $m = 10$, $\tau = 9$ before noise reduction (M. Perc).

# 4. Nonlinear analysis of ECG

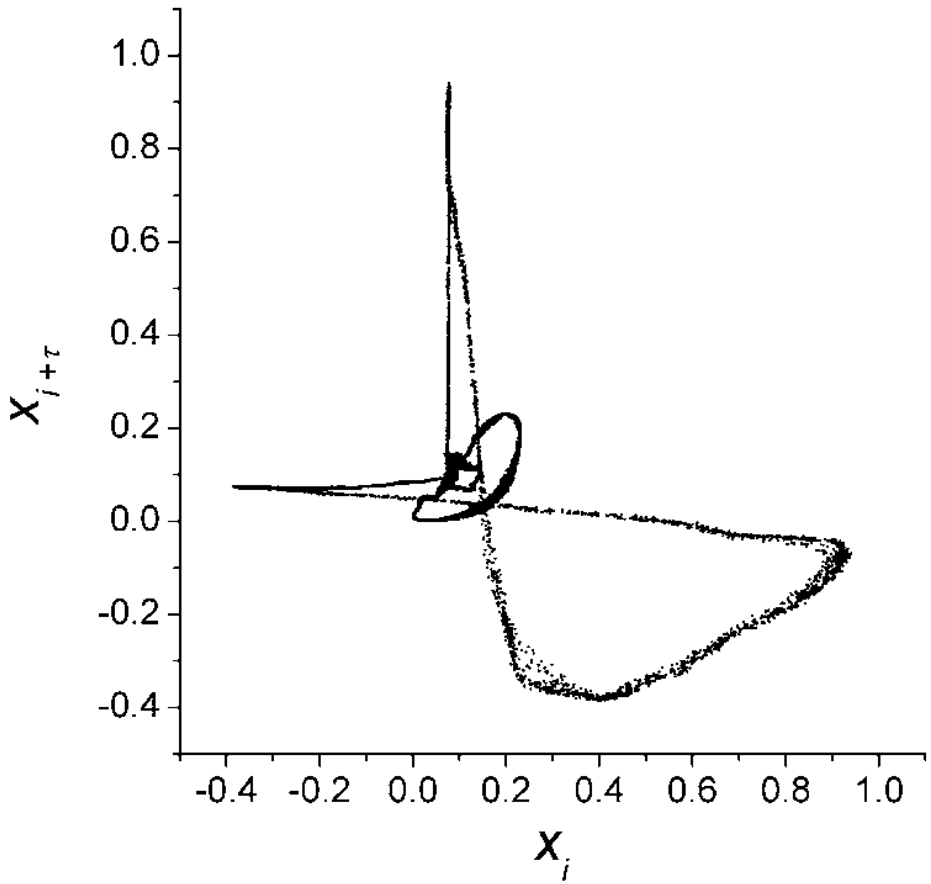



*Figure*. 2D projection of the reconstructed state space with the optimal parameters $m = 10$, $\tau = 9$ after noise reduction (M. Perc).

**Step 3: Determinism test**

*Use the Kaplan-Glass test:*

- Quantize the attractor with a grid of $18 \times 18 \times ... \times 18 \approx 3.6 \times 10^{12}$ boxes
- The average length $\kappa$ of all directional vectors $\mathbf{V}_k$, is $\kappa \approx 0.94$

$\Rightarrow$ *the signal is deterministic*.

**Step 4: Stationarity test**

*Use the cross-prediction error statistic*

- Number of segments: $I = 56$
- Number of points in the segments: $N/I = 800$
- Minimum cross-prediction error: $0.32$
- Maximum cross-prediction error: $0.60$
- Average cross-prediction error: $0.45$

$\Rightarrow$ *the signal is stationary*.

**Step 5: Computation of the attractor invariants**

**Recommendation**: Use well-tested algorithms or off-shelf software

- Attractors dimensions [P. Grassberger, Phys. Lett 97A (1983) 224]
- Lyapunov exponents [M.T. Rosenstein et al., Physica D 65 (1993) 117, H. Kantz, Phys. Lett. A 185 (1994) 77]
- In general:
  - TISEAN project: www.mpipks-dresden.pgg.de/~tisean
  - Mathematica, MatLab, ...

$\Longrightarrow \lambda \approx 0.015 \Longrightarrow$ *the data is slightly chaotic*!

# Conclusions of the course

1. Nature is nonlinear
2. Nonlinear time series analysis is half a science, half an art
    1. The theory is highly sophisticated
    2. The practice requires special skills
3. General recommendations
    - Extract only the information you need
    - Use different approaches and techniques
    - Be aware of the assumptions and approximations
    - Study the dependence on parameters and scaling behavior

# References


1. J.M. Amigó et al. Estimating the entropy rate of spike trains via Lempel-Ziv complexity, Neural Comp. 16 (2004) 717.
2. X. Li et al., Predictability analysis of absence seizures with permutation entropy, Epil. Res. 77 (2007) 70.
3. U. Parlitz et al., Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics, Comp. Biol. Med. 42 (2012) 319
4. M. Perc, Nonlinear time series analysis of the human electrocardiogram, Eur. J. Phys. 26 (2005) 757
5. C.J. Stam, Nonlinear dynamical analysis of EEG and MEG, Clin. Neuro. 116 (2005) 2266.