Permutation Entropy
&
Statistical Complexity

in the analysis of biomedical and economic signals

R.Riera – Dep. Física – PUC-Rio

INCT- Sistemas Complexos
Reunião anual – CBPF 22-24 abril 2013

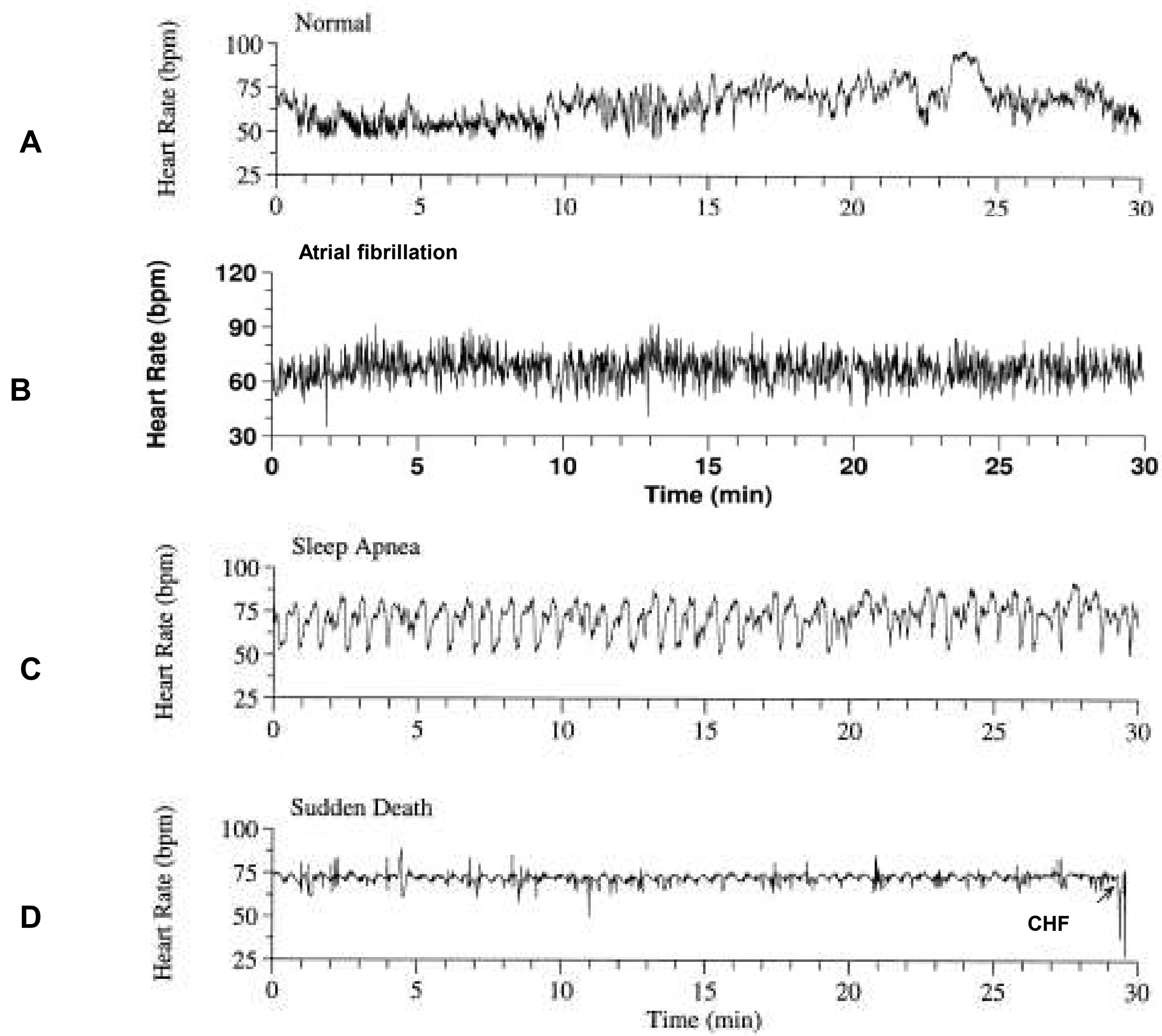
Findings

- **Biomedical**: pathological recordings are less complex than normal ones (diseases destroy intrinsic meaningful structures)

- **Ecological** : the flow rate of river systems changes its complexity due to human intervention

- **Economic** : market inefficiency enhances complexity

Representative physiological fluctuations.



Representative heart rate recordings of 30 minute-time series from a subject: (**A**) healthy, (**B**) with atrial fibrillation (AF) - which produces an erratic heart rate, (**C**) with obstructive sleep apnea, and (**D**) with congestive heart failure (CHF).

Goldberger A L et al, Circulation ,**101**:e215-e220 (2000) & PNA **99**:2466 (2002).

Empirical analysis

- The healthy heartbeat time series exhibits visual patches suggesting the existence of memory effects

- Their breakdown in disease is associated with the emergence of uncorrelated randomness (B) or regularity (C)

- A healthy system needs to exhibit processes which run on several time scales to be able to respond to unpredictable stimuli and stresses.

- Multi-scaling and complexity degrade with aging and disease reduces the adaptability of the organism.

Entropy measures – I and II

Consider the PDF P(x) of a variable X that assumes values x belonging to a finite alphabet $\mathbb{A}$ with N classes: $P \equiv \{p_i \,;\, i=1, \ldots, N\}$

Shannon Entropy (quantifies randomness at a global level):

$$S[P] \equiv -\sum_{x \in \mathbb{A}} P(x) \ln P(x) \qquad H[P] = S[P] / S_{\max} \ \textit{with} \ S_{\max} = \ln N$$

Consider the m-dimensional sequence vectors $u^{(m)}(i) \equiv \{x_i , x_{i+1} ,\ldots, x_{i+m-1} \}$

Sample Entropy (quantifies the regularity of finite length time series):

$SampEn\ (m, r, N)$ reflects the likelihood that vector sequences that are close to each other within a fixed threshold value r, remain close when more data points are known.

Approximate Entropy (ApEn) & Sample Entropy (SampEn)

Consider the distance between two vectors as the maximum of the absolute differences between their components and fix a threshold value r for determining when these vectors are close to each other. ApEn reflects the likelihood that sequences that are close to each other, i.e., within r, for m consecutive data points remain close when one more data point is known. Mathematically, ApEn is computed as follows: Let $\{X_i\} = \{x_1, \ldots, x_i, \ldots, x_N\}$ represent a time series of length N. Consider the m-length vectors: $u_m(i) = \{x_i, x_{i+1}, \ldots, x_{i+m-1}\}$. Let $n_{im}(r)$ represent the number of vectors $u_m(j)$ within r of $u_m(i)$. $C_i^m(r) = n_{im}(r)/(N-m+1)$ is the probability that any vector $u_m(j)$ is within r of $u_m(i)$. Define, $\Phi^m(r) = 1/(N-m+1) \sum_{i=1}^{N-m+1} \ln C_i^m(r)$. ApEn is defined as $\mathrm{ApEn}(m, r) = \lim_{N\to\infty} \Phi^m(r) - \Phi^{m+1}(r)$. For finite N, it is estimated by the statistics, $\mathrm{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r)$. Lower values of ApEn reflect more regular time series while higher values are associated with less predictable (more complex) time series.

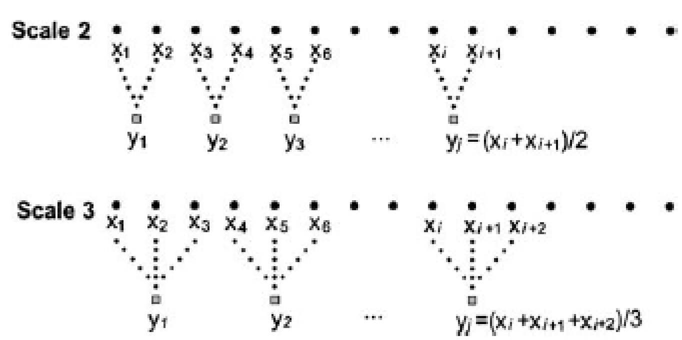
SampEn is defined as in the above procedure, but avoiding self matching.

Multi-scale Sample Entropy (MSE) method

Given a time series, $\{x_1,\ldots,x_i,\ldots,x_N\}$, we first construct consecutive coarse-grained time series by averaging a successively increasing number of data points in non-overlapping windows . Each element of the coarse-grained time series, $y_j^{(\tau)}$, is calculated according to the equation:

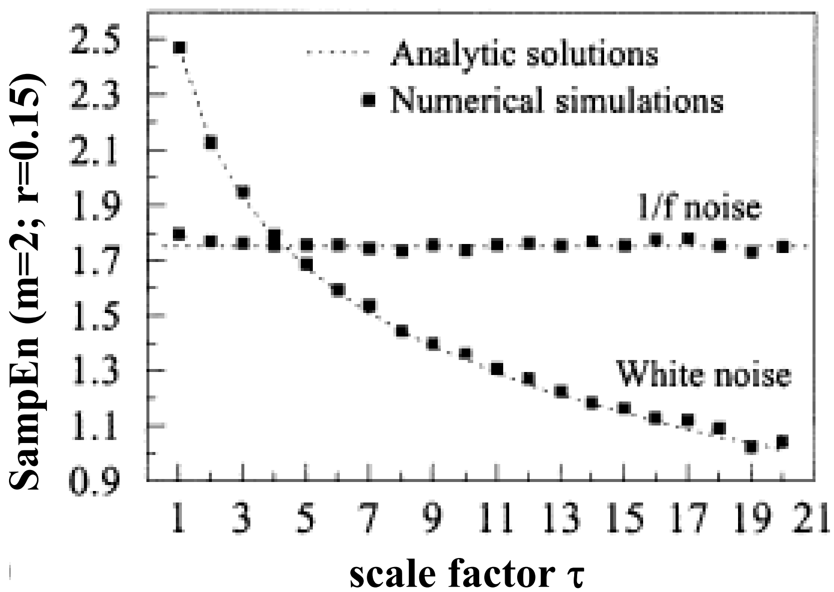
$$y_j^{(\tau)} = 1/\tau \sum_{i=(j-1)\tau+1}^{j\tau} x_i \ ,$$

where τ represents the scale factor and $1 \leqslant j \leqslant N/\tau$. The length of each coarse-grained time series is N/τ. For scale 1, the coarse-grained time series is simply the original



Schematic illustration of the coarse-graining procedure for scales 2 and 3.

Control datasets



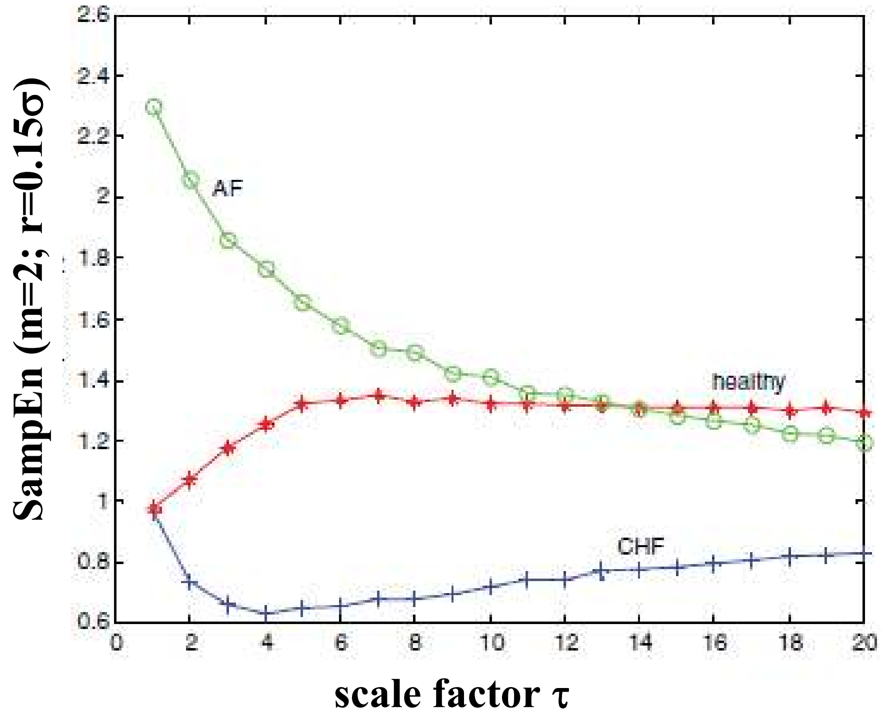
MSE analysis of standard Gaussian white noise and 1/f noise

Uncorrelated random signals are highly unpredictable, but, at a global level, they admit very simple description and therefore , are not really complex.

At small-scale factors, the white noise has entropy larger than the 1/f noise, but at larger scales, the long-ranged noise produces larger entropies

→ The MSE method provides a suitable measure of complexity for stochastic processes.

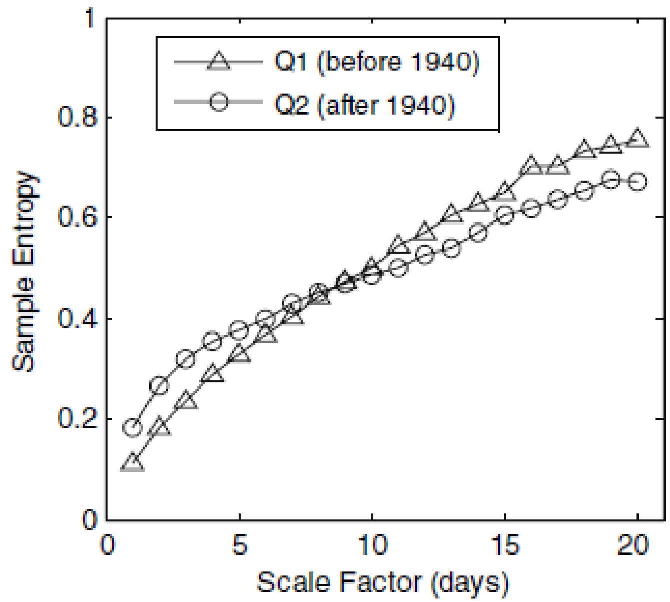
Analysis of cardiac data using MSE



Plot of the multi-scale (averaged) signature for healthy hearts (stars), with atrial fibrillation (circles) and congestive heart failure (crosses).
Thuraisingham R A and Gottwald G A, Physica A **366**, 323 (2006)

MSE analysis of river flow

Similar to a biological system, a river system may be categorized into healthy or pathologic, depending on the conditions in the watershed.



- Sample entropy increases, indicating increasing complexity over larger time scales due to long-range correlations in stream flow.

- Fluctuations of the Mississippi river (MR) flow after 1940 become rougher (more random) than before .

- MR system may lose its complexity due to human intervention

Entropy based complexity measures: Multi-scale Sample Entropy

- Sample entropy is bigger for random systems in short scales

- Multi-scale sample entropy measures how order/disorder is kept along time scales: large fluctuations at large timescales means larger entropy

- MSE is able to quantify the degree of complexity

Entropy measures - III

Causal information may be incorporated into the evaluation of the PDF if a symbol Π is assigned to a trajectory's portion.

Consider the n-dimensional sequence vectors $u^{(n)}(i) \equiv \{x_i , x_{i+1} ,\dots, x_{i+n-1} \}$.

The n real values of the vector are permutated in increasing order:

$$x_{i+k1-1} < x_{i+k2-1} < \dots < x_{i+kn-1} \quad with \quad 1 \leq k1, k2, \dots kn \leq n$$

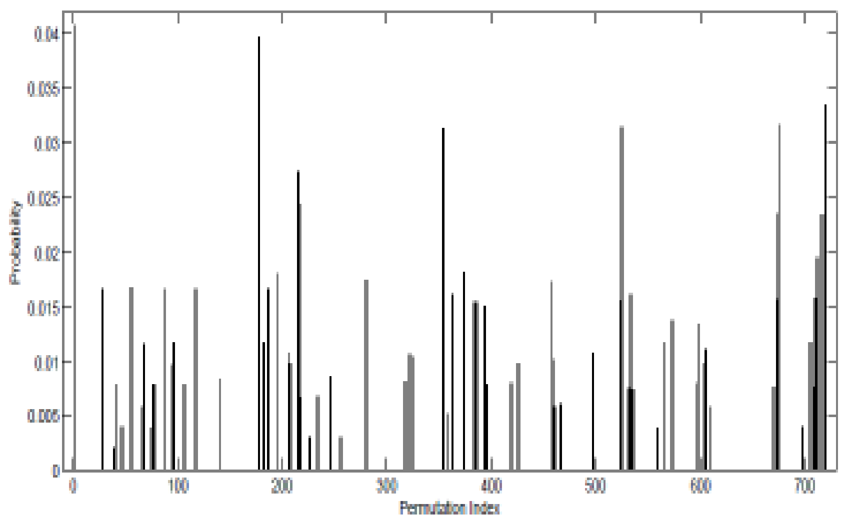
As a result, any sequence vector is mapped onto a vector $\Pi \equiv [k1, k2, \dots kn]$ which is one of the n! permutations of the n integers.

Consider the relative frequency p(Π) of the permutation pattern Π.

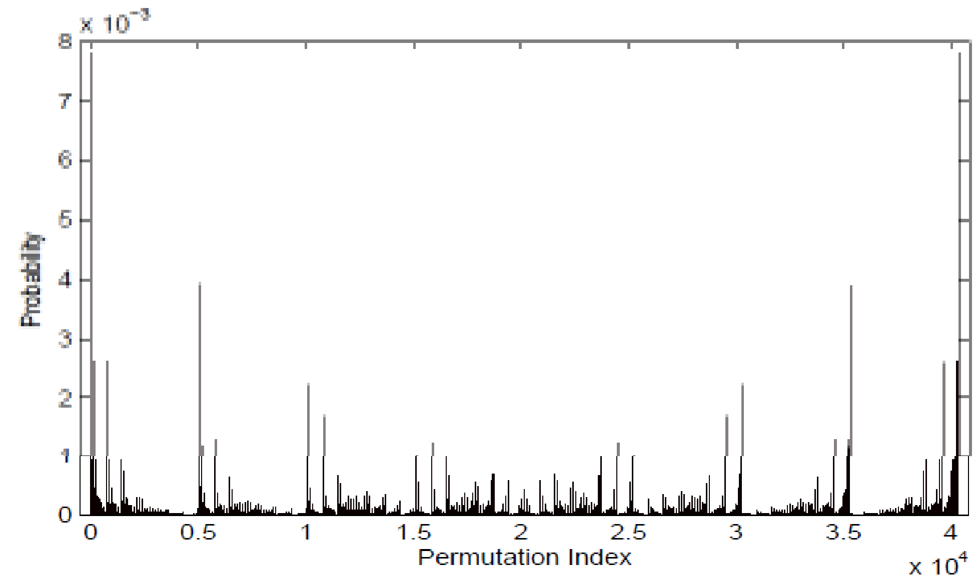
Permutation Entropy : $$S[P] \equiv -\sum_{\Pi \in \mathbb{A}} P(\Pi) \ln P(\Pi)$$

$$H[P] = S[P] / S_{\max} \quad with \quad S_{\max} = \ln n!$$

Distinguishing deterministic and stochastic dynamics



Permutation histogram with $n = 6$ for the fully chaotic logistic map



Permutation histogram with $n = 8$ for Brownian Motion

- Deterministic processes display forbidden permutation patterns

- Stochastic processes display higher permutation entropy values than the deterministic ones.

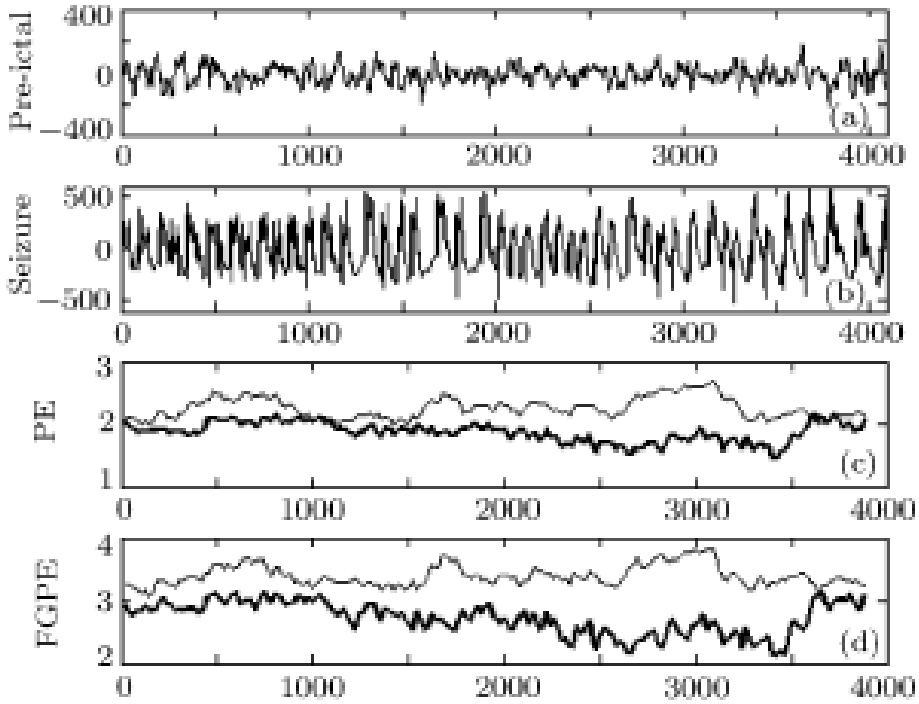
Fine-grained Permutation Entropy (FGPE)

Permutation Entropy focuses only on the order of the elements in the time series, disregarding the magnitude of the difference between neighboring elements.

A factor q is introduced to quantify the major difference between them and it is appended to the permutation pattern as an additional element.

In the calculation of FGPE, one considers the relative frequency of this new defined permutation pattern . In this way, the FGPE distinguish two sequences according to the magnitude of the fluctuations.

Detection of dynamical changes in epileptic data using PE and FGPE



- Accurate detection of transitions from normal to pathological states may improve diagnosis

- The FGPE improves the performance for detecting dynamical changes in time series

Electroencephalography (EEG) signal in normal state (a) and seizure state (b);
PE(c) and FGPE (d) for order n=4 in moving windows of length 200:
normal state (thin line) and seizure state (bold line).
Xiao-Feng L and Yue W, Chinese Physics 18, 1674 (2009)

Entropy based complexity measure: Permutatin Entropy

- Permutation Entropy (PE) effectively discriminate between deterministic chaos and random noise

- The PDF of ordinal patterns are invariant with respect to non-linear monotonous transformations
(a desirable property when dealing with experimental data)

- PE is suitable to capture the structure transitions between healthy and pathological states

Statistical Complexity

Jensen-Shannon Divergence: $J[P|U] = H\left[\frac{P+U}{2}\right] - \frac{H[P]}{2} - \frac{H[U]}{2}$

where U is the uniform distribution, taken as a reference

Disequilibrium: $Q[P] = J[P|U] / J_{\max}$ *with* $J_{\max} = \ln 2$

Statistical Complexity: $C[P] = H[P] * Q[P]$

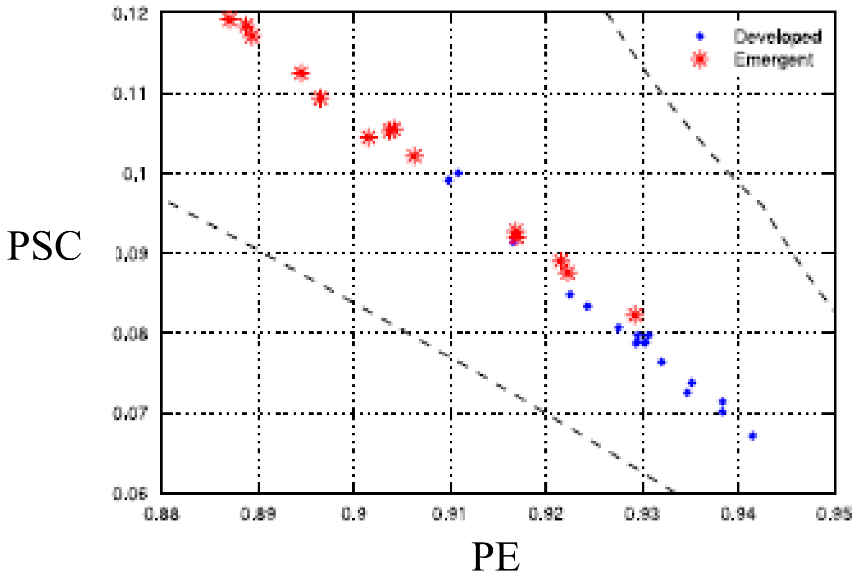
- The C[P] measure ascribes null value of complexity for both regular and completely random series. Therefore, systems exhibiting a behavior in-between these two extremes, although encompassing intermediate content of information, are more complex.

Permutation Statistical Complexity (PSC)

An important consequence of handling the permutation probability P(Π) compared to standard sample distribution P(x) is that one improves the performance of the information-based quantifiers, capturing not only randomness but also the memory structures.

In this case, while the H[P] measures the overall diversity of the observed ordinal patterns, the statistical complexity measure encompass additional information of the ordinal pattern distribution, revealing new characteristics of the data flow dynamics within permutation order n.

PE - PSC Plane



- The complexity-entropy plane can discriminate emergent and develop markets
- inefficiency enhances complexity

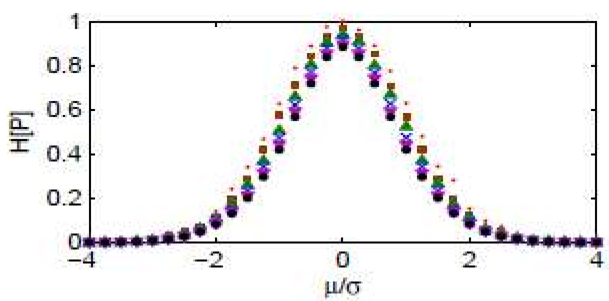
Localization of world stock indices daily time series for permutation order *n=4*.
Zunino L et al, Physica A 389, 1891 (2010)

Representative stochastic processes

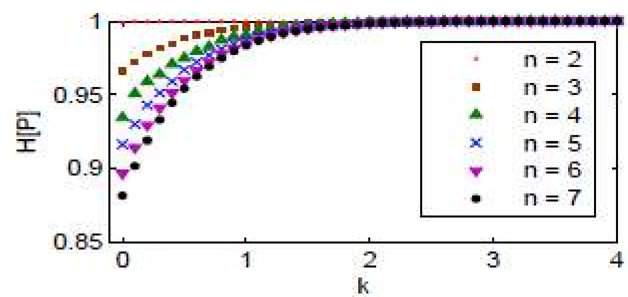
Consider stochastic processes described by: $dx = \mu(x)dt + \sigma(x)dW$

class of drifting processes: μ= cte and σ=cte

class of mean-reverting processes: $\mu(x) = -k(x - x^*)$ and $\sigma = cte$



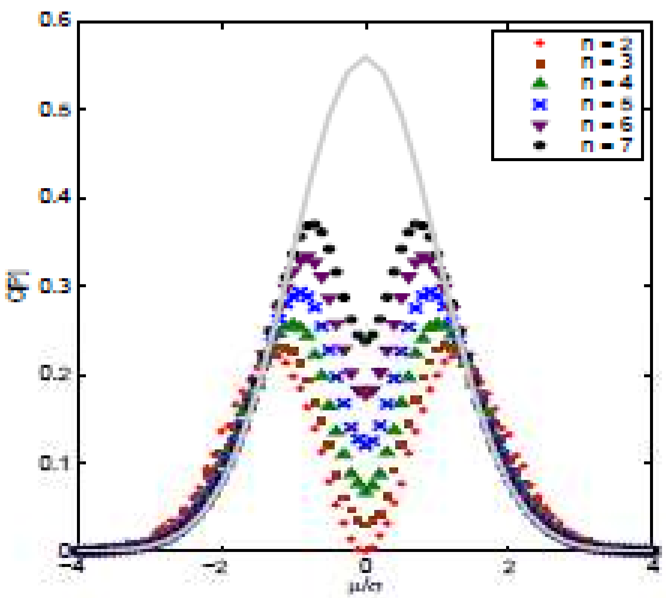
(drifting)



(reverting)

- PE discriminates real data according to the nature (drift or reverting) of the deterministic forces underlying complexity : for each order $n \geq 3$, the class of mean reverting process furnishes PE values larger than the class of drift processes

Multi-scale PSC curve for drift processes



PSC curve for several orders *n* and the limiting shape for arbitrary large scales. Ribeiro A and Riera R, Physica A (submitted, 2013)

Brownian motion is a scale invariant process at a global level, and therefore, are rather complex.

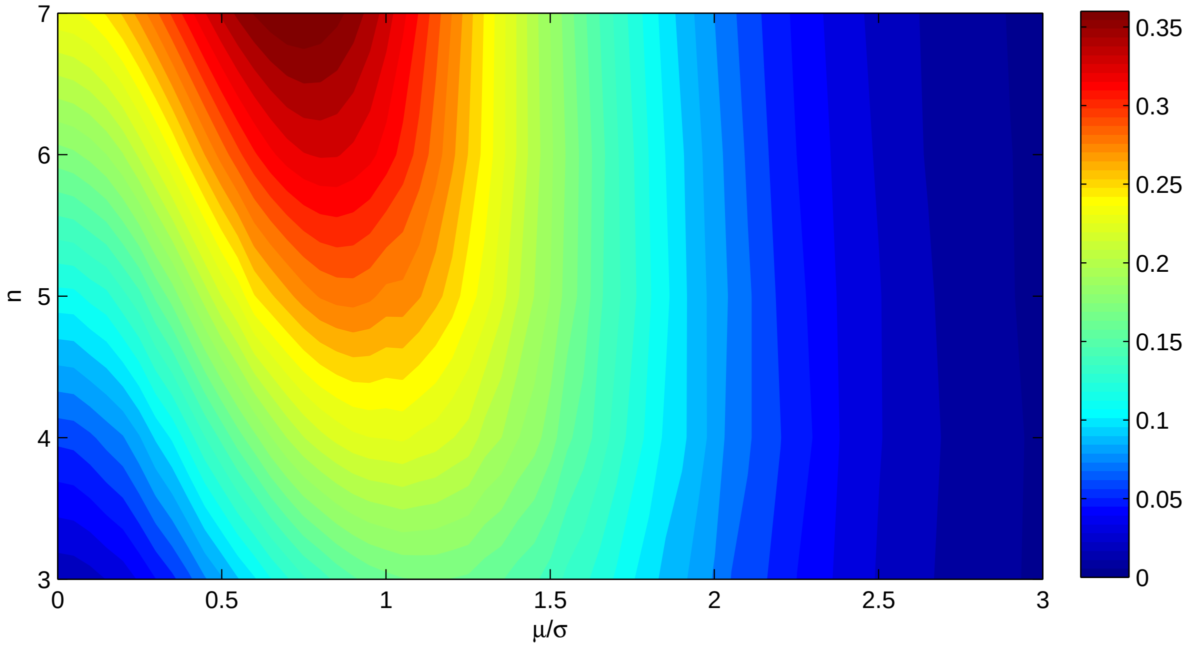
At small-scale factors, the Brown noise has PSC values smaller than the noise with drift, but at larger scales, it produces larger complexity.

→ global trends are associated with complexity loss, indicating forecast opportunities

Multi-scale PSC

- The complexity measure PSC depends crucially on the permutation order under consideration

- similar systems show different signatures of their own complexity dependency on *n,* giving rise to crossovers



Color map of PSC values for drift processes
Ribeiro A and Riera R, Physica A (submitted, 2013)

Conclusions

- New informational structures are captured by PSC and MSE as time scales increases

- The degree of complexity depends crucially on the time scales τ (or order n) under consideration

- Some time series may look complex in short time intervals whereas the actual signal is not → empirical analysis based on low-order complexity measures should be taken with caution