Quantifying complexity

Cristina Masoller
Departament de Fisica i Enginyeria Nuclear
Universitat Politecnica de Catalunya, Terrassa, Spain

Osvaldo A. Rosso
Centre for Bioinformatics, Biomarker Discovery and Information-based Medicine
ARC Centre of Excellence in Bioinformatics
School of Electrical Engineering and Computer Science
The University of Newcastle, Australia

Lab Meeting, Terrassa, May 2009
The simple and the complex

- **Complexity?**
- Is a state of affairs that one can easily appreciate when confronted with it; however, is difficult to define it quantitatively.
- We are after a quantity “C” that captures complexity as the entropy “H” captures randomness.
The simple and the complex

$H = 0$

$C = 0$

Mondrian
The simple and the complex

$H = 1$
$C = 0$

Pollock
The simple and the complex

$H \neq 0$

$C \neq 0$

Bosch
### The simple and the complex

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>( H = 0 )</td>
<td>( H \neq 0 )</td>
<td>( H = 1 )</td>
</tr>
<tr>
<td>( C = 0 )</td>
<td>( C \neq 0 )</td>
<td>( C = 0 )</td>
</tr>
</tbody>
</table>
The simple and the complex

<table>
<thead>
<tr>
<th>Complete order</th>
<th>Chaos</th>
<th>Complete disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>H = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H ≠ 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C ≠ 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C = 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Complexity

In between “perfect order” and “complete disorder”, complexity has to do with the existence of a certain degree of structure, temporal and/or spatial correlations hidden in the dynamics.

Statistical complexity is often characterized by the paradoxical situation of a complicated dynamics being generated from relatively simple systems.

One can begin by excluding processes that are not complex:
- periodic motion
- a purely random process

The complexity of a system must reach a maximum between the system’s perfectly ordered and perfectly disordered states.
Motivation

Can complexity be induced by noise?
Outline

- **Introduction:**
  - complexity quantifiers
  - Noise-induced order: coherence and stochastic resonances

- Noise-induced complexity

- Summary, Conclusions & Future Work
Is a useful tool for studying neural computation, analysis of patterns in medical signals and imaging technologies, machine learning, etc.
Measuring structural complexity in brain images

Karl Young and Norbert Schuff


An information theory based formalism for medical image analysis proposed in Young et al. [Young, K., Chen, Y., Kornak, J., Matson G. B., Schuff, N., 2005. Summarizing Complexity in High Dimensions, Phys. Rev. Lett. 94 098701] is described and used to estimate image complexity measures as a means of generating interpretable summary information. An analysis of anatomical brain MRI data exhibiting cortical thinning, currently considered to be a sensitive early biomarker for neurodegenerative diseases, is used to illustrate the method. Though requiring no previous assumptions about the detailed shape of the cortex or other brain structures, the method performed comparably (sensitivity = 0.91) to direct cortical thickness estimation techniques (sensitivity = 0.93) at separating populations in a data set designed specifically to test the cortical thickness estimation algorithms. The results illustrate that the complexity estimation method, though general, is capable of providing interpretable diagnostic information.
Quantifying complexity?

- Tools for detecting and quantifying deterministic *low-dimensional* chaos: metric entropy, Lyapunov characteristic exponents & fractal dimensions.

- For stochastic processes: Shannon’s entropy is a widely use measure of a system’s degree of randomness or unpredictability.

- The statistical complexity is an *information measure*, useful for studying high-dimensional data, generated from complex physical systems with multiple time and/or spatial scales.
Assuming that we know the probability distribution $P = [p_i, i=1,N]$ that fully characterizes a given system, we can define the following information measures

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

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

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

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

where \( I_{\text{max}} = I[P_e] \)

$P_e$ is the equilibrium probability distribution, that maximizes the information measure.

**Example:** if \( I[P] = \) Shannon entropy = \( S_S[P] \)

then \( P_e = [p_i=1/N \text{ for } i=1,N] \)

and \( I_{\text{max}} = \ln(N) \)
Disequilibrium $Q$

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

$$Q[P] = Q_0 D_{P, P_e}$$

where $Q_0$ is a normalization constant such that $0 \leq Q[P] \leq 1$
Distance between \( P \) and \( P_e \)

- Euclidean
- Wootters
- Kullback relative entropy
- Jensen divergence
- etc

\[
D_E[P, P_e] = \|P - P_e\|_E = \sum_i \phi_i - 1 / N
\]

\[
D_W[P, P_e] = \cos^{-1}\left[\sum_i p_i^{1/2} \phi / N^{1/2}\right]
\]

\[
D_K[P, P_e] = K[P | P_e] = I [P_e - I P]
\]

\[
D_J[P, P_e] = \frac{K[P | P_e] + K[P_e | P]}{2}
\]
A family of complexity measures can be defined as:


Where

A = S, T, R (Shannon, Tsallis, Renyi)
B = E, W, K, J (Euclidea, Wootters, Kullback, Jensen)

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


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

Statistical complexity measure $C$
The complexity is a non-trivial function of the entropy: for a given value of $H$, there is a range of possible values of $C$.

The complexity vanishes in the limits $H=0$ and $H=1$. 
Given a time series: \[ X = x_t, \quad t = 1 \ldots M \]

We can define the associate probability distribution based on:

- Histogram of amplitudes
- Binary representation (symbolic dynamics)
- Frequency (Fourier transform)
- Frequency bands (Wavelet transform)
- Ordinal Patterns (attractor representation)
- etc
Binary representation: for each value of r the dynamics was reduced to a binary sequence (0 if $x \leq 1/2$, 0 if $x > 1/2$) and strings of length 12 were considered as states of the system. The probabilities were assigned according to the frequency of occurrence in $2^{22}$ interactions.
The complexity of the Logistic map
Probability distribution from ordinal patterns

Permutation Entropy: A Natural Complexity Measure for Time Series

Christoph Bandt and Bernd Pompe
Institute of Mathematics and Institute of Physics, University of Greifswald, Greifswald, Germany
(Received 19 June 2001; revised manuscript received 20 December 2001; published 11 April 2002)

We introduce complexity parameters for time series based on comparison of neighboring values. The definition directly applies to arbitrary real-world data. For some well-known chaotic dynamical systems it is shown that our complexity behaves similar to Lyapunov exponents, and is particularly useful in the presence of dynamical or observational noise. The advantages of our method are its simplicity, extremely fast calculation, robustness, and invariance with respect to nonlinear monotonous transformations.

Patterns of order D=3, the # of ordinal patterns is 3! = 6
Probability distribution from ordinal patterns

\[ x = (4, 7, 9, 10, 6, 11, 3). \]

We organize the six pairs of neighbors, according to their relative values, finding four pairs for which \( x_t < x_{t+1} \) and two pairs for which \( x_t > x_{t+1} \). So four of six pairs of values are represented by the permutation 01 (\( x_t < x_{t+1} \)) and two of six are represented by 10. We define the permutation entropy of order \( n = 2 \) as a measure of the probabilities of the permutations 01 and 10. So,

\[ H(2) = -(4/6) \log(4/6) - (2/6) \log(2/6) \approx 0.918. \]

As usual, \( \log \) is with base 2, thus \( H \) is given in bit. Next, we compare three consecutive values. \((4, 7, 9)\) and \((7, 9, 10)\) represent the permutation 012 since they are in increasing order. \((9, 10, 6)\) and \((6, 11, 3)\) correspond to the permutation 201 since \( x_{t+2} < x_t < x_{t+1} \), while \((10, 6, 11)\) has the permutation type 102 with \( x_{t+1} < x_t < x_{t+2} \). The permutation entropy of order \( n = 3 \) is \( H(3) = -2(2/5) \log(2/5) - (1/5) \log(1/5) \approx 1.522. \)

\[ \{x_t\} = \{3.4, 5.1, 6.8, 11.5, 1.1, 2.3, \ldots \} \]

\[ (3.4, 5.1, 6.8) \Leftrightarrow (0 \ 1 \ 2) \]

\[ (5.1, 6.8, 11.5) \Leftrightarrow (0 \ 1 \ 2) \]

\[ (6.8, 11.5, 1.1) \Leftrightarrow (2 \ 0 \ 1) \]

\[ (11.5, 1.1, 2.3) \Leftrightarrow (1 \ 2 \ 0) \]

Good statistics when the number of data points: \( M >> D! \)
Noise induced order

- Stochastic resonance
- Coherence resonance
Stochastic resonance

\[ \dot{x} = x - x^3 + A \cos(\omega t) + \sqrt{2D} \xi \]

\[ \langle \xi(t) \xi(0) \rangle = \delta(t) \]

Varying D keeping \( \omega \) fixed

Varying \( \omega \) keeping D fixed

Quantifying SR

Periodic response & SNR

Histogram of residence times
Quantifying SR

Varying $D$ keeping $\omega$ fixed

Varying $\omega$ keeping $D$ fixed
Coherence resonance

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

\[ \frac{dy}{dt} = x + a + D \xi(t). \]

\[ R_p = \frac{\sqrt{\text{Var}(t_p)}}{\langle t_p \rangle} \]

\[ \tau_c = \int_0^\infty C^2(t) dt \]

Quantifying noise-induced complexity
Stochastic resonance

Bandt and Pompe method (ordinal patterns) applied to $M=60000$ consecutive residence times intervals.

Varying $D$ keeping $\omega$ fixed

empty symbols: analysis of the series of residence times intervals, but rearranged the data randomly

Rosso and Masoller, PRE 2009, EPJB 2009
Stochastic resonance

Varying $\omega$ keeping $D$ fixed

Empty symbols: analysis of the series of residence times intervals, but rearranged the data randomly.
Stochastic resonance
Coherence resonance

Bandt and Pompe method applied to $M=60000$ consecutive inter-spike intervals

\[
\varepsilon \frac{dx}{dt} = x - \frac{x^3}{3} - y,
\]
\[
\frac{dy}{dt} = x + a + D \xi(t).
\]

solid symbols: $\varepsilon=0.01$ and $a=1.05$;
empty symbols: $\varepsilon=0.1$ and $a=1.005$
Summary

- We have shown that the information-theory measures, the normalized Shannon entropy, and MPR statistical complexity can be employed to detect and quantify resonant-like behavior in the form of enhanced temporal order induced by the variation in a system parameter or by the variation in the noise level.

- The success of the method is based on an appropriate reconstruction of the attractor and on an appropriate partition of the phase space that results in ordinal patterns having a probability distribution that fully characterizes the temporal correlations in the system.

- It will be interesting to apply the method to experimental data.
FIG. 1. Temporal behavior of the laser intensity for increasing input noise amplitude. From top to bottom: noise $= -60.8$ dBm/MHz (a), $-52.5$ dBm/MHz (b), and $-44.3$ dBm/MHz (c). The horizontal scale is 100 ns/div. The vertical scale is the same for the three plots.

Marino et al, PRL 2002

Giacomelli et al, PRL 2000
Experimental normalized variance of the time between LFF drops as a function of the injection current, for a feedback delay of 6 ns.

Martinez-Avila et al, PRL 2004