ESTIMATING RELATIVE CHANGES IN COMPLEXITY MEASURES

M. Torres and L. Gamero

Universidad Nacional de Entre Ríos,
Facultad de Ingeniería, Bioingeniería,
C.C. 47, Suc. 3, Paraná (3100), E. R. - Argentina.
E-mail: {metorres, lgamero}@alpha.arcride.edu.ar

ABSTRACT

Estimation of complexity is of great interest in nonlinear signal and system analysis. In the present study, complexity measures derived from the Shannon entropy, the Harvda-Charvat-Daróvczy-Tsallis and their corresponding information are presented. The performance of the proposed measures in the presence of changing complexity were evaluated through numerical experiments. An application with real data (electroencephalograms) is presented. The analysis were compared against classical techniques (Ziv and Lempel, Approximate entropy, Lyapunov exponents). The results obtained showed that the entropic approach allows to discern in a similar qualitative fashion the complexity changes, with much less computational cost.

1. INTRODUCTION

In the study of nonlinear dynamical systems we often deal with experimental data where the underlying dynamics is not well known. Most of them present a rich variety of self-oscillating regimes that involve either regular or complex behavior [2].

The algorithm complexity for sequences of finite length was suggested by Ziv and Lempel, and it is related with the number of distinct substrings and the rate of their recurrence along the given sequence [8]. Ziv-Lempel (LZ) complexity can be a finer measure than the Lyapunov exponents for characterizing order. Another measure of complexity (regularity) is the Approximate Entropy (ApEn) [10] that allows complex system classification. It is well known its ability to quantify complexity with a reduced amount of data point although it requires a relative high computational burden. LZ complexity and ApEn are not measures of chaos but they quantify the regularity embedded in the time series.

The renormalized entropy, opposite to other complexity measures, is defined relative to a fixed state and has been used to indicate all transitions from periodic to chaotic behavior as well as between different types of chaos [11]. Basically it is the Kullback Information respect to a state with a given value of effective energy. The main disadvantage of this technique, as most of the usual complexity measures (correlation dimension, Lyapunov exponents), is the large amount of data required for their estimation.

Several notions of entropy have been used to characterize the order degree in ordinary differential or difference equations. The application of quantitative measures for the analysis of such signals has helped to gain better understanding of the system dynamics [11]. The classical Shannon entropy, which comes from information theory, describes the evolution of order. In this study, the Shannon entropy and the more general Harvda-Charvat-Daróvczy-Tsallis [6, 4, 15] (q-entropies) and their corresponding relative informations are presented as complexity measure estimators. We compare the Shannon, (q-entropies) and their information measures against the LZ, ApEn and Lyapunov exponents. Some examples of known nonlinear systems and the analysis of real data are presented.

2. COMPLEXITY MEASURES

In this section we will briefly review the complexity measures considered in this study. For more comprehensive discussions, see e.g. [6, 4, 8, 10, 14].

2.1. Approximated Entropy (ApEn)

The approximated entropy can classify a system given at least 1000 data values in diverse settings both deterministic, chaotic and stochastic processes. The capability to discern changing complexity from such relatively small amount of data holds promise for application of ApEn to a variety of contexts. Given a finite time series $u(1),\ u(2),\ \ldots,\ u(N)$ and a fix positive integer m and r a positive real number we considered the embedding vectors $x(1), \ldots, x(N-m+1)$ in $\mathbb{R}^m$, where $x(j) = [u(j), u(j+1), \ldots,\ u(j+m-1)]^t$. For each i, $1 \leq i \leq N-m-1$ $C_i^m(r) = ($ number of $j \leq N-m+1$ such that $d(x(i), x(j)) \leq r)/(N-m+1)$, where $d(x(i), x(j))$ is the l_∞ norm. Define

$$\text{ApEn}\ (m, r) = lim_{N\to\infty}[\Phi^m(r) - \Phi^{m+1}(r)], \quad (1)$$

where

$$\Phi^m(r) = \sum_{i=1}^{N-m+1} \ln\ C_i^n(r)\ /(N-m+1).$$

2.2. Ziv-Lempel (LZ)

As proposed in [8], the complexity of a finite sequence can be evaluated from the point of view of a simple self delimiting learning machine which, as it scans a given N-digits sequence $\boldsymbol{u}\ =\ u(1), \ldots,\ u(N)$ from left to right, adds a new word to its memory every time it discovers a substring of consecutive digits not previously encountered. The size of the compiled vocabulary and the rate at which new words are encountered along $\boldsymbol{u}$ serve as the basic ingredients in the LZ complexity evaluation.

2.3. Entropies

Given a signal $\boldsymbol{u}$ we can defined its Shannon entropy as [12]:

$$\mathcal{H} = -\sum_{i=1}^{M} p_i\ \ln(p_i), \quad (2)$$

where p_i is the probability that the signal belongs to a considered interval and with the understanding that, $p\ \ln(p) = 0$ if $p = 0$. The entropy $\mathcal{H}$ is a measure of the information needed to locate a system in a certain state j^*; it means that $\mathcal{H}$ is a measure of our ignorance about the system. The Harvda-Charvat-Daróvczy-Tsallis [6, 4, 15] q-entropy, that depends on a single real parameter q, reads as

$$\mathcal{H}_q\ =\ (q-1)^{-1} \sum_{i=1}^{M} [p_i\ -\ (p_i)^q]. \quad (3)$$

2.4. Relative entropies

The entropy of a random variable (rv) is a measure of the uncertainty of the rv; it is a measure of the amount of information required on the overage to describe the rv. The relative entropy is a measure of the distance between two distributions. The relative entropy (or Kullback–Leiber distance) $D(f|g)$ between two probability densities f and g is defined by [3]:

$$D(f|g) = \sum_x f(x)\ \ln \frac{f(x)}{g(x)}, \quad (4)$$

where we use the convention that $y \ln(y) = 0$ if $y = 0$. In the q-entropies case, for $q \neq 1$ and $q \in \mathbb{R}$ the corresponding relative q-entropies are given by [14]:

$$D_q(f|g) = \frac{1}{1-q} \sum_x f(x) \left(1 - \left(\frac{f(x)}{g(x)}\right)^{q-1}\right) \quad (5)$$

The Kullback divergence $KDiv(f|g)$ between two probability densities f and g is defined by [3]:

$$KDiv(f|g) = D(f|g) - D(g|f). \quad (6)$$

3. RESULTS AND DISCUSSION

In order to compare the performance of the complexity measures previously described, we present numerical simulations with nonlinear dynamical systems and electroencephalograms (EEG) signals. The central idea of the entropic analysis is that of

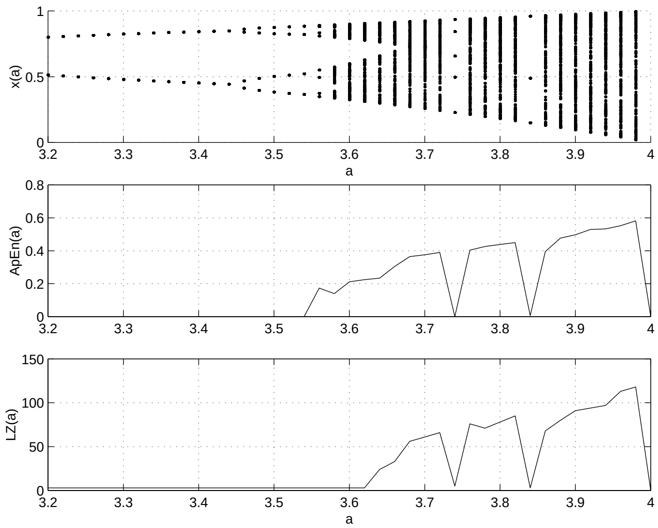

Figure 1: The quadratic map: $x(n+1) = ax(n)(1-x(n))$, evaluated at $N = 1500$ points. The bifurcation diagram of the quadratic map (top), the ApEn (middle) and the LZ (bottom) are plotted as functions of the parameter a.

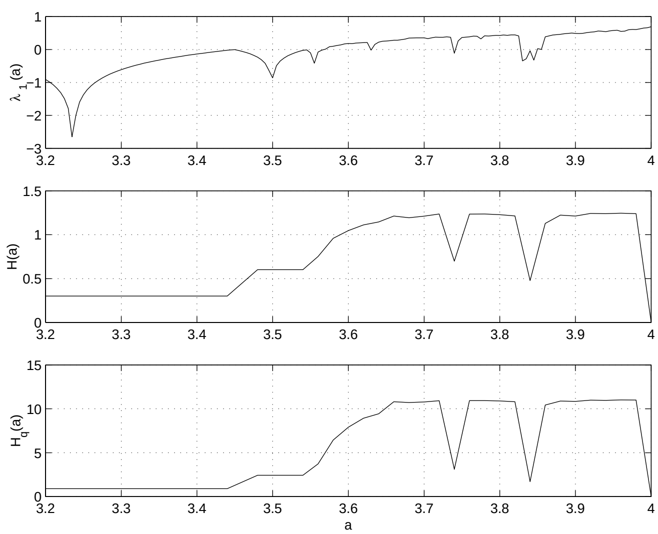

Figure 2: The Lyapunov exponent (top panel), Shannon entropy (middle) and q-entropy (lower plot) of the quadratic map signal corresponding to different a-parameter values.

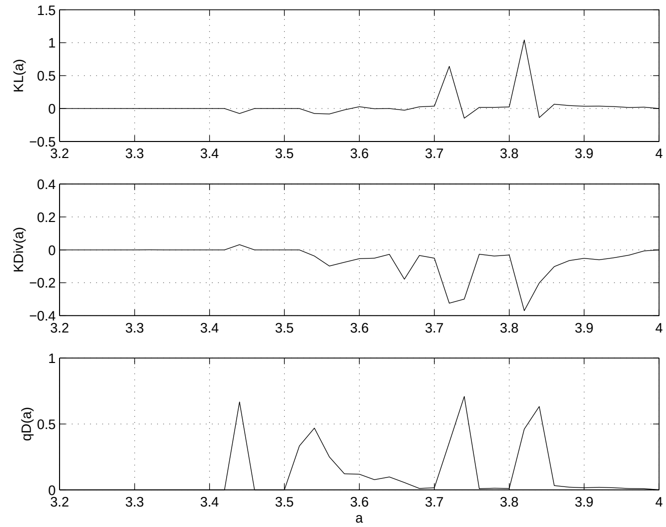

Figure 3: Kullback relative entropy (top), Kullback divergence (middle) and q-divergence (bottom) of the quadratic map signal corresponding to different a-parameter values.

being in a position to generate, from a given nonlinear signal, a suitable probability set. Once the probabilities are determined, the entropies and information measures are computed according to the pertinent definitions. The algorithms have been implemented in Matlab 5.2.

Figure 1 shows at the top panel the quadratic map bifurcation diagram with a parameter $a \in [3.2, 4]$. The middle panel plot the ApEn evaluated over 1500 samples corresponding to each parameter values removing the transient behavior. The LZ (Figure 1 bottom), Lyapunov exponent (Figure 2 top), Shannon entropy (Figure 2 middle) and q-entropy (Figure 2 bottom) were evaluated considering the same signals. In the case of the relative measures (Figure 3) we have considered windows of the same length and consecutive parameter values.

Some physiological systems behave in a nonlinear chaotic fashion and different methods have been developed to estimate the complexity of such behavior. An example of real data corresponding to epileptic EEG signal is presented. Some authors have shown that the variability of the EEG signals does not represent noise but is the signature of an attractor [9, 1]. Iasemidis et al. [7] have

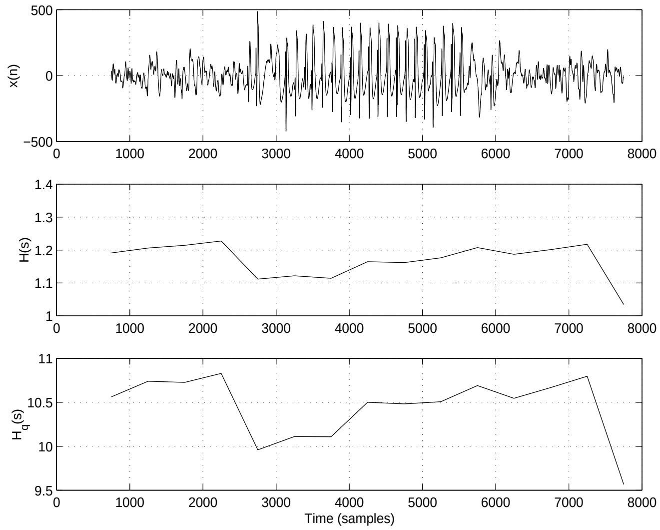

Figure 4: EEG signal (top) corresponding to an epileptic seizure. At the middle and button panels the plots of the entropy and q-entropy evolutions evaluated for adjacent running windows of 500 samples.

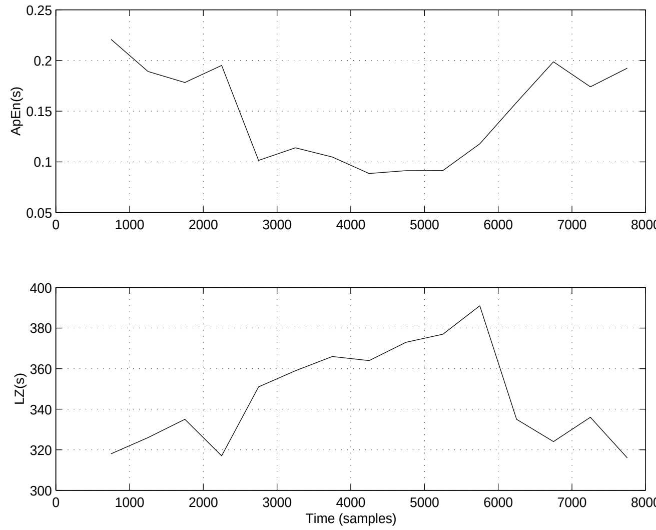

Figure 5: Approximated entropy (top) and Ziv-Lempel complexity (bottom) evolutions corresponding to the EEG signal shown in Figure 4, evaluated for adjacent running windows of 500 samples.

established that, during the preictal period, signals from each electrode exhibit positive values of the first Lyapunov exponent (λ_1), with multiple transient drops; at the onset of the seizure, signals show a drop in λ_1 to its lowest value.

The EEG record here presented correspond to a patient explored with 12 multi lead electrodes. The analysis of interictal and ictal data is usually of visual nature. The signals were amplified and filtered using a 1-40 Hz band-pass filter. A four-pole Butterworth filter was used as an anti-aliasing filter. The EEG was sampled at 256 Hz through a 10 bits A/D converter.

Figure 4 (top) depicts the raw EEG data. The middle and bottom panels plots the entropy and q-entropy evolutions evaluated for adjacent non-overlapped running windows. All the experiments were performed with a window width of 500 samples. The ApEn and Ziv-Lempel complexity (bottom) are presented in Figure 5. The results of evaluating the Kullback relative entropy, Kullback divergence and q-divergence are presented in Figure 6. At the seizure onset that appears in the initial stages as transients in the data, its complexity is higher than basal state as it was expected. As the seizure progresses, the signal appears to be more regular, and thus, the complexity change in a comparative sense. At the end of the seizure, transients are observed again.

It is known the capability of the ApEn to discern changing complexity for relative small amount of data and it was observed in the examples. However, the entropies time evolution, allowed to establish a suitable agreement with the results obtained either with ApEn and LZ measures with a much less computational cost. Our approach allows not only to discern in a similar qualitative fashion the complexity changes, but also, by means of the multiresolution entropy [13, 5], frequency bands of different complexity could be established. Multiresolution analysis allows the possibility of introducing a different, and perhaps more elaborate, information measure, that naturally incorporates all the advantages of wavelet analysis. Application of entropies and information measures in nonlinear signal analysis should be considered as a promising tools in real time applications.

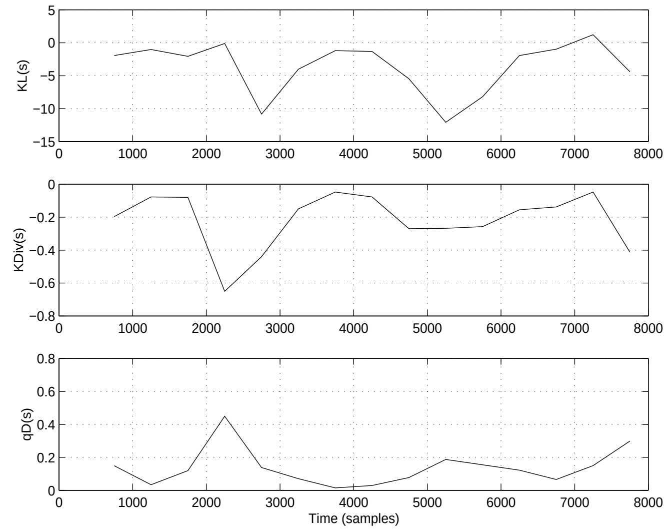

Figure 6: Kullback relative entropy (top panel), Kullback divergence (middle panel) and q-divergence (lower plot) evolutions corresponding to the EEG signal shown in Figure 4.

Acknowledgments

L. Gamero wishes to thank the Facultad de Ingeniería (Universidad de Buenos Aires) for partial financial support.

4. REFERENCES

[1] Basar E. Chaos in Brain Function'. Springer Verlag, Berlin, 1990.

[2] Conant R. C. "Detecting subsystems of a complex system". IEEE SMC, 550–553, 1972.

[3] Cover T. M., and Thomas J. A. Information Theory. John Wiley and Sons, 1991.

[4] Darovczy Z. "Generalized information functions". Inf. Control, 16:36–51, 1970.

[5] Gamero L. G., Plastino A., and Torres M. E. "Wavelet analysis and nonlinear dynamics in a non extensive setting". Physica A, 246:487–509, 1997.

[6] Havrda J.and Charvat F. "Quantification method of classification process: Concept of structural α-entropy". Cybernetica, 3:30–35, 1967.

[7] Iasemidis L. D., Sackellares J. C., Zaveri H. P., and Williams W. J. "Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures". Brain Topogr., 2(3):187–201, 1990.

[8] Lempel A. and Ziv. "On complexity of finite sequences". IEEE Trans. Inf. Theory, 22(1):75–81, 1976.

[9] Meyer-Kress G. and Layne S. C. Dimensionality of human electroencephalogram. In M. F. Shlesinger S. H. Koslow, and A. J. Mandel, editor, Perspectives in Biologycal Dynamics and Theoretical Medicine, number 504. Ann. N.Y. Acad. Sci, New York, USA, 1987.

[10] Pincus S. "Approximate entropy (ApEn) as a complexity measure". Chaos, 5:110–117, 1995.

[11] Saparin P., Witt A., Kurths J., and Anishchenko B. "The renormalized entropy – an appropriate complexity measure?" J. Chaos, Solitons and Fractals, 4(10):1907–1916, 1994.

[12] Shannon C. E. "A mathematical theory of communication". The Bell System Technical Journal, 27(3):379–423, 1948.

[13] Torres M. E., Gamero L. G., and D'Attellis C. E. "A multiresolution entropy approach to detect epileptic form activity in the EEG". In I. Pitas, editor, IEEE Workshop on non linear signal and image processing, volume II, 791–794, 1995.

[14] Torres M. E. El procesamiento de señales ligadas a problemas no lineales. PhD Thesis (in spanish), Universidad Nacional de Rosario, Argentine, 1998.

[15] Tsallis C. "Somme comments on Boltzmann-Gibbs statistical mechanics". Chaos, Solitons and Fractals, 6:539, 1995. and references therein.