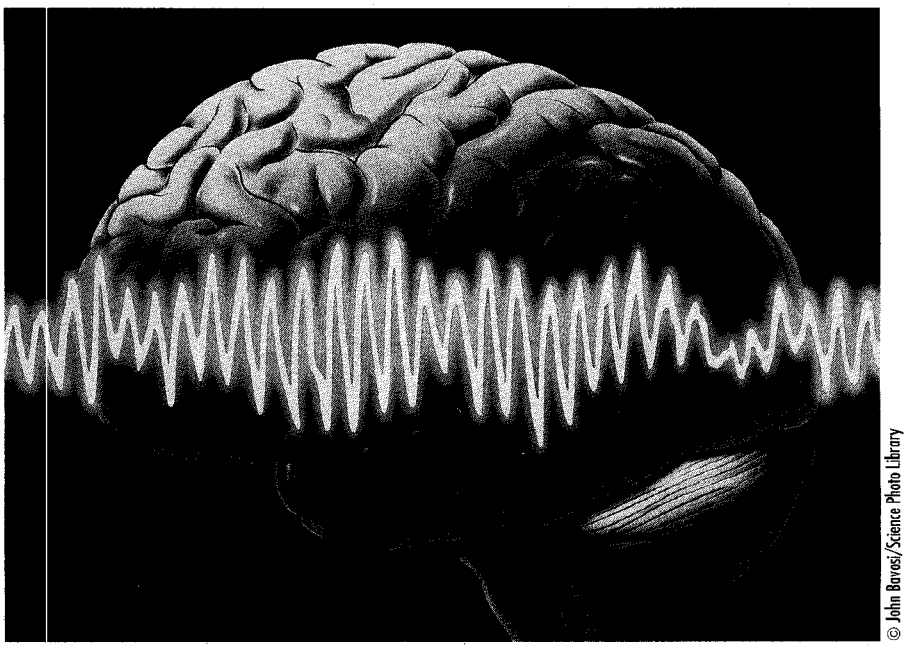




# Estimating Regularity in Epileptic Seizure Time-Series Data

## A Complexity-Measure Approach

**N. Radhakrishnan**
Department of Electrical Communication Engineering,
Indian Institute of Science
**B.N. Gangadhar**
Department of Psychiatry,
National Institute of
Mental Health and Neurosciences (India)

In the study of complex systems one maybe faced with experimental data collected on the variables without knowing much about the way in which these variables interact. The parts of such a complex system interact in a norisimple way [1]. Wolfram *[2]* found that there exists a complex dynamic state which may be even more complex than chaotic motions. Complex systems have: been found to have an underlying deterministic model and exhibit chaotic dynamics in some cases, epilepsy in particular [3]. The use of quantitative measures for the analysis of these systems has helped gain better insight into system dynamics. In this article we apply Ziv-Lempel (LZ) complexity and approximate entropy (ApEn) as measures to quantify the regularity in the various epochs of epileptic seizure time-series data.

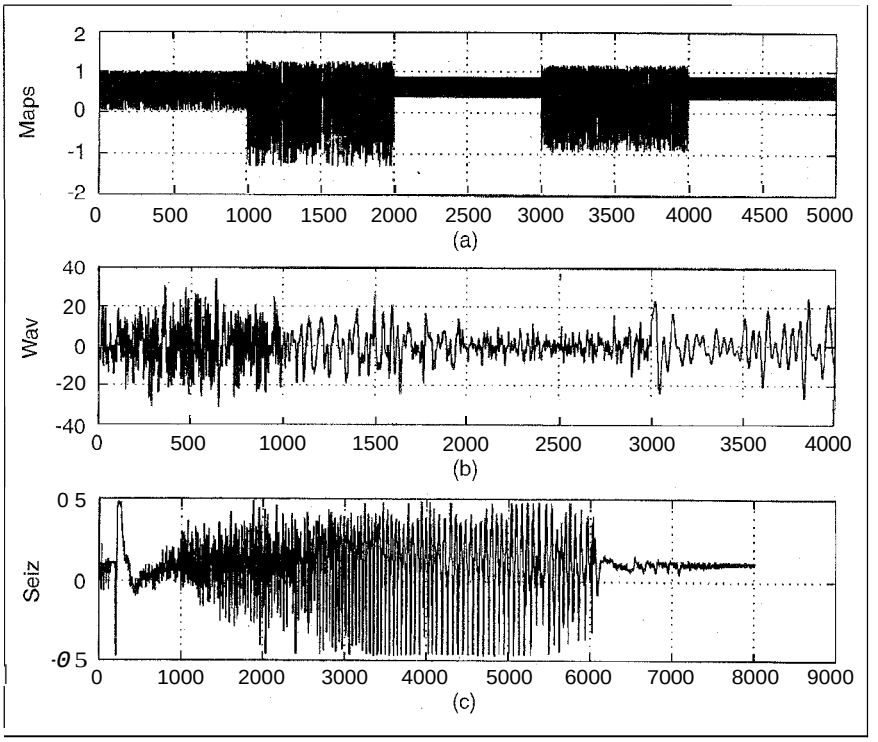
### Analyzing Complex Systems

Nonlinear dynamics has been one of the most popular approaches for analyzing complex systems. However, application to neuronal processes is under criticism where the EEG signal may be considered to have a simpler stochastic description, and chaotic dynamic measures can turn out to be spurious and unnecessary. The EEG signal has been found to be both nonstationary and high dimensional, and the calculation of quantities such as dimensions is not strictly accurate and has meaning only in a comparative sense [4]. Validity of surrogate data testing for low-dimensional chaos has also been questioned [5].

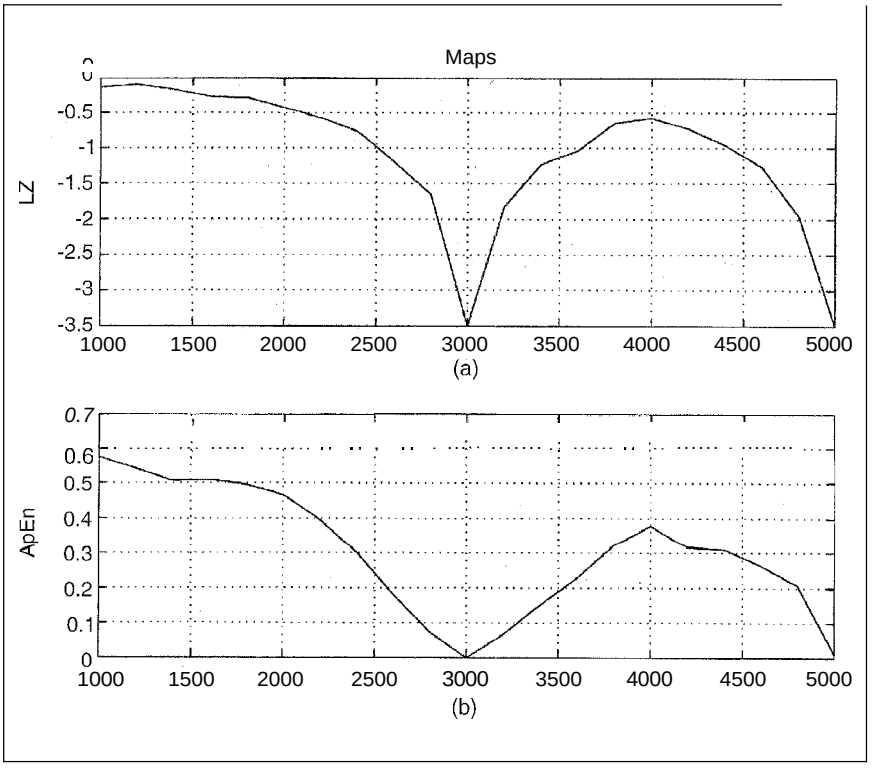
Some popular invariants that have been used to characterize the regularity of the systems have been the correlation dimension (D2) and the Lyapunov exponent (LE). Correlation dimension gives us an idea about the minimal dimension that the state space should have in order to reconstruct the strange attractor. The traditional Grassberger-Procaccia algorithm [6] assumes the time series to be stationary and noise free. However, these assumptions are not generally true. The stationarity constraints imposed on the time-series data results in the reduction of the scaling region, making the calculation of D2 error prone. The estimation of LEs, which quantifies the long-term average rate of exponential growth of small perturbations to initial conditions [7], is very sensitive to noise. The numerical effort required to extract its spectrum is quite large.

Since computational complexity of these algorithms is also high, one should be very clear whether it is *necessary* to apply these algorithms to the problem con-

 0739-5175/98/$10.00©1998 




**1. (a) Time series made of standard maps in the order mentioned in the article; (b) time series made of alpha, beta, theta, and delta waves in the order mentioned in the article; (c) time series of an epileptic seizure (y-axis: magnitude, x - axis: samples).**




**2. (a) The variation of Ziv-Lempel complexity for the time series made of well-known maps (y-axis: magnitude of the LZ complexity, x-axis: samples); (b) the variation of the approximate entropy for the time series made of well-known maps (y-axis: magnitude of the ApEn measure, x-axis: samples). The x-axis starts from 1000 because we need 1000 samples for the first window to find the complexity.**

cerned or is there an *alternate* easier way. New methods for EEG analysis hold promise, and limited progress has been made in finding new methods for diagnosis [4]. The method of symbolic dynamics, an algebraic approach that originated from abstract topological theory of dynamical systems, has been successfully applied to one-dimensional maps [8, 9]. It was also carried out successfully for the multicomponent Belousov-Zhabotinskiis reaction [10].

In this article, we do not to argue for the existence of *chaos* or estimate the regularity by finding the invariants of the epileptic seizure time-series data, but use alternate *complexity measures* to quantify the regularity embedded in the time series. Epileptic seizures represent a pathological state of brain activity, characterized by synchronous discharge of large groups of neurons. In particular, we have studied electroconvulsive (ECT) therapy-induced seizures. We are concerned with generalized epilepsy, where the EEG activity during the seizure is found to switch into an apparent oscillating mode, with a succession of more or less regular and extremely coherent waves. We have used two mathematically well-established measures, namely LZ complexity and ApEn, as relative indices to quantify the regularity of this time-series data. Other measures of complexity can also be applied for analysis [2, 11].

The algorithmic complexity, $c(n)$, for sequences of finite length was suggested by Ziv and Lempel [12]. It is related to the number of distinct substrings and the rate of their recurrence along the given sequence. $c(n)$ reflects the order that is retained in the sequence. In this work, we have coded the time series as a binary sequence and then evaluated its LZ complexity. It has been shown that LZ complexity can be a finer measure than the LE for characterizing order [13].

The ApEn can classify complex systems given at least 1000 data values in diverse settings, both deterministic chaotic and stochastic processes. The capability to discern changing complexity from such a relatively small amount of data holds promise for application of ApEn to a variety of contexts [14]. We shall consider three time series data: (1) the artificial time series data made of well-known maps with tunable parameters; (2) the time series made up of a sequence of normal waves; alpha, beta, theta, delta; and (3) the EEG time-series data obtained

from patients who received ECT. We compute the variation of the above-mentioned complexity measures as we run through the series, thus establishing the effectiveness of these measures. The seizure data under consideration consists of transients in the initial stages, so we expect the series to be more complex in this stage. Regularity sets in the series as the seizure progresses, which indicates a reduction in complexity in a comparative sense. Transients are observed again at the end of the series.

## Algorithmic Complexity — An Overview

The probabilistic interpretation of distinguishing sequences may turn out to be singularly unhelpful, failing to give any information about the order embedded in the sequence. One such example is given below. The first ideas in the field of algorithmic complexity were introduced by Kolmogorov [15] and Chaitin [16]. Algorithmic complexity is defined as the length in bits of the shortest algorithm required by a computer to produce the given string. The shortest algorithms are referred to as *minimal programs*. The complexity of a string is thus the length, in bits, of the minimal program necessary to produce that string.

In the case of a random string, the algorithmic complexity is found to be equal to the length of the string. In other words, a random string cannot be compressed to a more compact form; any attempt to do so would result in information loss. It can also be seen that a system fails to characterize a sequence appropriately if the sequence has as much information as the system itself. Consider the two binary sequences **(A),** (B) of length $n = 16$, where (A) has been chosen purposely to be a periodic pattern and (B) is chosen seemingly random:

(A) 0101010101010101

(B) 0001101001000101

The probability of the occurrence of string **(A)** and string (B) are the same.

## Ziv-Lempel Complexity (LZ)

We first introduce some basic definitions:

A: Alphabet of symbols (for a binary sequence we have two symbols, namely 0 & 1)

s: finite-length sequences formed by A, $s = s1\ s2..sn$ where $si \in A$

$v(s)$: vocabulary of sequence $s$; it the set of all substrings of $s$

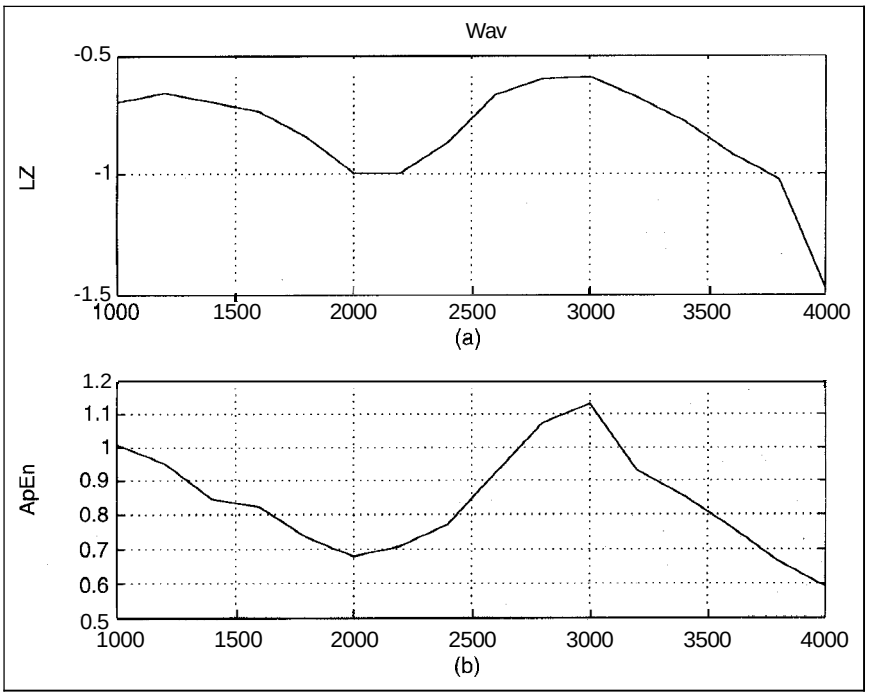
$s\pi$: number of elements in the set $s$ minus one

i.e, if $A = \{0, 1\}$ and $s = 010$ then $v(s) = \{0, 1, 01, 10, 010\}$ and $v(s\pi) = (0, 1, 01\}$.

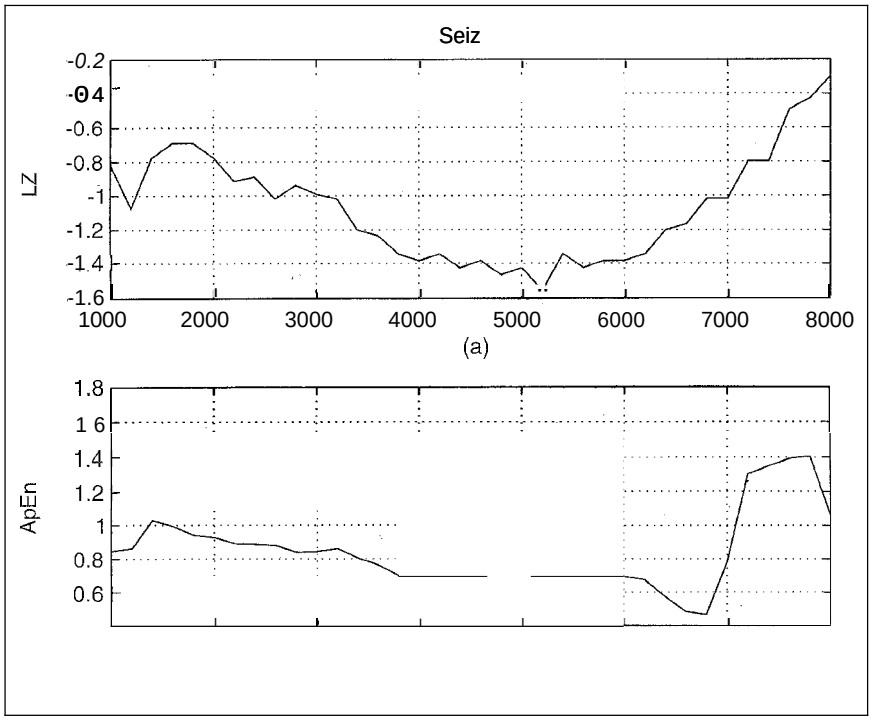
An absolute measure for complexity is believed to be nonexistent. However, complexity of a finite sequence is evaluated from the point of view of a simple self-delimiting learning machine that, as it scans a sequence from left to right, adds a new word to its memory every time it discovers a substring of consecutive digits not previously encountered. The first symbol is always inserted and the procedure is iterated. The size of the compiled *vocabulary* and the rate at which new words are encountered along $s$ serve as the basic ingredient in the proposed evaluation of the complexity of s.

Let us assume that a given string, s1 $s2...sn$, has been reconstructed by the program up to the digit $sr$, and $sr$ has been newly inserted. The string up to s r will be denoted by $s = s1\ s2. . .$ sr., where the dot (after sr) denotes that $sr$ is newly inserted in order to check whether the rest of the string $sr+1 . . . sn$ can be reconstructed by simple copying. First, we take $q = sr+1$ and see whether $p = sq$ is reproducible from s. If the answer is "no" then we insert $q = sr+1$ into the sequence followed by a dot. Thus, it could not be obtained by copying. If the answer is "yes," then no new symbol is needed and we proceed with $q = sr+1\ sr+2$ and carry out the same procedure. The LZ complexity is the number of dots (plus one if the string is not terminated by a dot). We shall not go into the details of the mathematical proofs of this measure. but we do consider the tw

**It has been shown that LZ complexity can be a finer measure than the LE for characterizing order or disorder.**




**3. (a) The variation of Ziv-Lempel complexity for the time series made up of normally occurring brain waves (y-axis: magnitude of the LZ complexity, x-axis samples); (b) the variation of the approximate entropy for the time series made up of normally occurring brain waves (y-axis: magnitude of the ApEn measure, x-axis: samples).**




**4. (a) The variation of Ziv-Lempel complexity for the seizure time-series data (y-axis: magnitude of the LZ complexity, x-axis: samples); (b) the variation of the approximate entropy for the seizure time-series data (y-axis: magnitude of the ApEn measure, x-axis: samples).**

strings mentioned above and find their LZ complexity.

Consider sequence (A): 0101010101010101

0

$s = 0; q = 1; sq = 01;$

$v(sq\pi) = \{0\}, q \notin v(sq\pi)$

$s = 01; q = 0; sq = 010;$

$v(sq\pi) = \{0,1\}; q \in v(sq\pi)$

Proceeding this way, we get the parsed sequence as 0.1.01010101010101, $c(s)$ = 2+1 = 3

Consider sequence (B): 0001101001000101

0

$s = 0; q = 0; sq = 00;$

$v(sq\pi) = \{0\}; q \in v(sq\pi)$

$s = 0; q = 00; sq = 000;$

$v(sq\pi) = \{0,00\}; q \in v(sq\pi)$

$s = 0; q = 001; sq = 0001;$

$v(sq\pi) = \{0,1,00,000\}; q \notin v(sq\pi)$

Proceeding this way, we get the parsed sequence as 0.001.10.100.1000.101, $c(s)$ = 5+1 = 6

We state here that only relative values of $c(s)$ are meaningful. We compare $c(s)$ of astring withrespecttoarandomstring of the same length. For a random string the complexity measure $b(n)$ is defined as:

$$b(n) = (hn)/(\log_k n)$$

where

$h$ = the normalized source entropy.

$k$ = number of elements in the alphabet A.

$$h = -\sum_{i=1}^{k} pi \log(pi)/\log k$$

let

$$\gamma = \lim_{n\to\infty} c(s)/b(n)$$

i.e., we have normalized $c(s)$ with respect to $b(n)$, where $b(n)$ gives the asymptotic behavior of $c(s)$. Consider a sequence generated from a random source having $k$ states, where $pi$ is the probability of finding the state, $i$. If the probability $pi < 1/k$, then we expect its complexity to be less than that of a random string with $pi = 1/k$. It should be noted that the source entropy is maximum at $pi = 1/k$. This implies that one can find out whether the deviation of "$\gamma$" from unity is either due to the fact the source entropy differs from unity or is due to pattern formation in the sequence.

## Approximate Entropy Measure (ApEn)

Approximate entropy measure [14] can classify complex systems. Its ability to quantify with a limited amount of data points and to distinguish between multiple periodic systems is to be noted. It has been found [14] that invariant measures can arise from deterministic as well as stochastic settings, and it is in generally not valid to infer the presence of an underlying deterministic system from the convergence algorithms designed to encapsulate properties of invariant measures. The ApEn has also been used to quantify sequential irregularity applied to both finite- and infinite-length sequences, and thus identify maximally irregular sequences [17]. Below we discuss the estimation of the ApEn in brief; for a detailed discussion please see [14].

Fix $m$, a positive integer, and $r$, a positive real number. If the given time series is of the form $u(1),u(2),\ldots,u(N)$, we form vectors $x(1),\ldots,x(N-m+1)$ in $R^m$ where

$$x(i) = [u(i),u(i+1),\ldots,u(i+m-1)].$$

This procedure is known as *embedding.*

We define for each i,1 $\le i \le N-m+1$

$$C_i^m(r) = (\text{number o f j such that } d[x(i),x(j)] \le r)/N-m+1)$$

We define $d[x(i),x(j)]$ for vectors $x(i)$ and $x(j)$. We follow [18], by defining

$$d[x(i),x(j)] = \max_{k=1,2,\ldots,m} (|u(i+k-1)-u(j+k-1)|)$$

Define

$$\phi^m(r) = 1/(N-m+1)\sum_{i=1}^{N-m+1} \log C_i^m(r)$$

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

$$\text{ApEn}(m,r,N) = \phi^m(r)-\phi^{m+1}(r)$$

## Methods and Materials

Electroconvulsive therapy was administered to chronically depressed patients after obtaining proper consent. Treatments were given under intravenous anesthesia using thiopentone (3 mg/kg), succinylcholine (0.5mg/kg), and atropine (1.2 mg). Patients were ventilated with 100% oxygen after the injections and throughout the seizures, until they resumed spontaneous respiration. Stimulus

was applied either bifi-onto-temporally (BL) or unilaterally (UL) on the right side. The ECT instrument delivered 800 mA constant current pulses of 1.5 ms with alternating polarity, at a rate of 12.5 Hz.

### EEG Data Acquisition

The EEG was recorded on two channels, from left and right frontal regions (F3 and F4), referenced to ipsilateral mastoids, with a ground on the forehead. The electrode impedance was < 10 kΩ. The EEG was amplified by a factor of 1000 with analog filter settings of 1.6 Hz (high pass), 35 Hz (low pass), and 50 Hz (notch). Square wave pulses of 100 μV were used for calibration. The EEG was acquired for 200 s following the end of stimulation. The EEG was digitized using a 12-bit A/D converter at the rate of 128 Hz per channel, using Labtech Notebook, and the data were stored for off-line analysis. The data were reaiewed and artifact-free segments were chosen from each channel (right and left). Each segment was low-pass filtered (30 Hz), using an FIR linear-phase digital filter of 80th order. All simulations were run on a Sparc Ultra Workstation and the graphs were plotted with the Matlab program.

## Simulation Results

We have Considered three time-series data and have quantified the regularity in them with overlapping segments. The window length was fixed at 1000 samples, with an overlap of 200 samples, and the parameters for the ApEn were $m$ =2, $r$ = 0.1SD). We coded the series about the mean by a binary sequence in the calculation of LZ complexity. The binary sequence consisted of two states, zero and one (choice of higher number of states was also possible). The plots for each of the cases under consideration are self explanatory.

### Case 1

Logistic map is given by:

$$x(n+1) = \alpha x(n)(1-x(n))$$

Henon map is given by:

$$x(n+1) = \beta y(n) + 1 - 1.4x(n)^2$$
$$y(n+1) = 0.3\beta x(n)$$

where $\alpha, \beta$ are the map parameters. We shall denote the logistic map with parameter $\alpha$ as $l(\alpha)$ and the henon map with parameter $\beta$ as $h(\beta)$.

In Fig. 1(a), we have considered a time series of 5000 samples formed by standard maps. The series consists of $l(3.95)$, $h(1.0)$, $l(3.5)$, $h(0.8)$, Z(3.56) of 1000 samples each in that order. $l(3.95)$, $h(1.0)$, $h(0.8)$ are chaotic series, whereas $l(3.5)$ is of period 4 and $l(3.56)$ is of period 8. The hierarchy of the decreasing regularity is in the order of $l(3.5)$, $l(3.56)$, $h(1.0)$, $h(0.8)$, $l(3.95)$. In Fig. 2(a), we see the variation of the LZ complexity. In Fig. 2(b), we see the variation of the ApEn as we run the window through the series. Both of these figures are in accordance with the above-mentioned hierarchy.

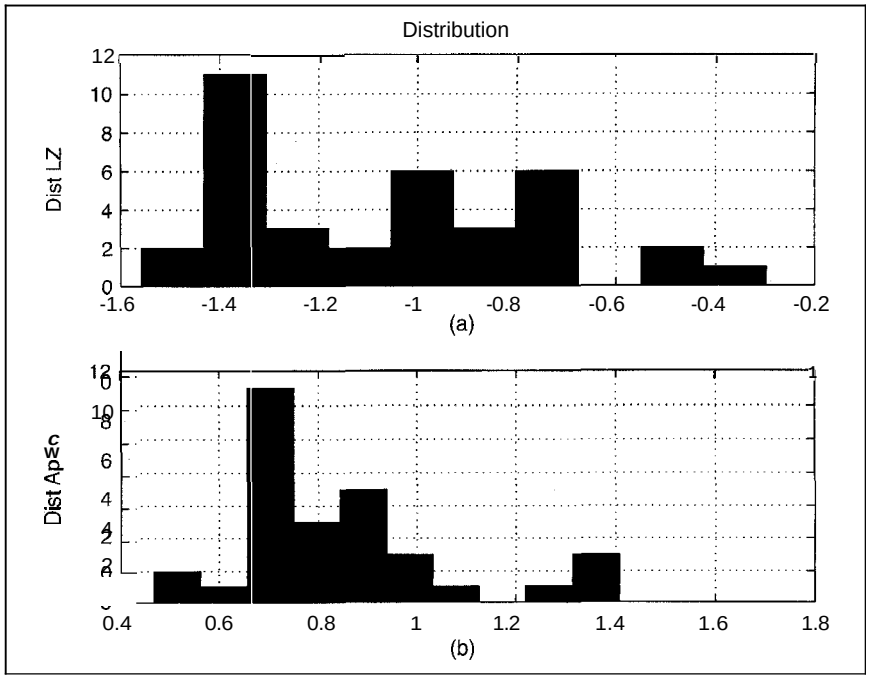
**In the case of a random string, the algorithmic complexity is found to be equal to the length of the string.**

### Case 2

In Fig.(1b), we consider a series of 4000 samples composed of alpha, theta, beta, and delta waves of 1000 samples each, in that order. Alpha waves have frequency range of 8-15 Hz and they occur over the occipital lobes in the awake, mentally relaxed state, with the eyes closed. Beta waves have a range 1.5-30**Hz** and they occur over the parietal and frontal lobes. Theta waves have a range of 4-7 Hz and they occur in adults during sleep and in children. Delta waves contain all the EEG activity below 4 Hz and occur during deep sleep, in premature babies, and in infants. The hierarchy of decreasing regularity is in the order delta, theta, alpha, beta. We can see the variation LZ complexity in Fig. 3(a) and of the ApEn in Fig. 3b as we move through the series. This is in accordance with the noted hierarchy.

### Case 3

We have used the complexity measures to quantify the regularity embedded in the seizure time-series data as the seizure progresses. The epileptic seizure time-series data has transients at the beginning and at the end. The seizure becomes more regular and coherent in the middle part. The seizure data considered is shown in Fig. 1(c). In Fig. 4(a), we see the value of LZ complexity falling gradually as the seizure progresses. The same is




*5.* (a) Distribution **of** the **LZ** complexity values shown in Fig. **4(a).** (b) Distribution **of** the **ApEn** values obtained in **Fig. 4(b).**

demonstrated by the ApEn in Fig. 4(b). We see the complexity more or less plateaus in the mid-seizure range. We have also plotted the distribution of the LZ measures in Fig. 5(a) and ApEn in Fig. 5(b). The frequency distribution gives us an idea about the relative occurence of the complexity values and the segment range where these values occur.

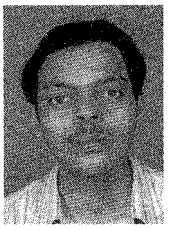
## Conclusion

We have demonstrated the potential of complexity measures such as LZ and ApEn in quantifying the regularity at different epochs of epileptic seizure time-series data. It is clearly shown that these measures have high values at the beginning and the end of the seizure, and that they decrease during mid-seizure. In fact, we observe in the histogram plot that the frequency of the complexity measure in mid-seizure is quite prominent. This gives us an idea about the epoch where we can find more regular patterns. These measures can also be used as relative indices (comparing across state), rather than absolute indices, by using a larger number of subjects to obtain statistical validity in comparing across conditions.

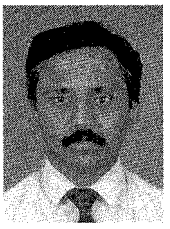
The analysis of time series obtained from complex systems, such as the brain, by the above measures provides an alternative easy way to quantify the regularity with finite-length segments (of the order of 1000 samples). The same can be inferred by calculating the D2 and LE, but the algorithms used to estimate these invariants are susceptible to error due to the finite sample size and are also highly sensitive to noise. The computational complexity of these algorithms is also high. We have also applied these measures across the various states of epilepsy (the details of which will be discussed in a later article).


## Acknowledgments

We would like to thank Prof. D. Narayanadutt, Prof. A.G. Ramakrishnan, and Prof. N.V. Joshi for their comments and suggestions as well as Prof. Joe Mietus for clarifying certain doubts related to approximate entropy. We would also like to thank the Supercomputer Education Research Center (SERC) for providing the computing facilities.


*N. Radhakrishnan* is a student pursuing a master's degree in the Department of Electrical Communication Engineering at the Indian Institute of Science. He is interested in the analysis of biomedical signals and complex systems.


*B.N. Gangadhar* is an Additional Professor in the Department of Psychiatry of the National Institute of Mental Health and Neurosciences (India). His research interests include estimation of the EEG signal variables obtained during electroconvulsive therapy.

**Address for Correspondence:** N. Radhakrishnan, Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560 012, India. E-mail: radha@protocol.ece.iisc.ernet.in.


## References


1. **Conant RC:** Detecting subsystems of a complex system. *IEEE SMC* 550 - 553, 1972.

2. **Wolfram** S: Computation theory of cellular automata. *Commu'n. Math.* Phys, 96: 15-57, 1984.

3. **Baloyantz A, Dextexshe A:** Low dimensional chaos in an instance of epilepsy. *Proc. Natl. Acad of Sci* (USA) 83: 3513-3517,1986.

4. **Layne SP, Mayer-Kress G, & Holzfuss J:** Problems associated with the analysis of EEG data, in Dimensions *and Entropies in Chaotic* Systems, *Springer* NY. 1986.

5. **Pradhan et al:** Validity of complexity measures to EEG. *Intl. Jrnl of bifurcation &* chaos 17: 173 - 186, 1997.

6. **Grassberger P, Procaccia I:** Characterisation of strange attractors. Phy. *Rev. Lett* 50: 346 - 349, 1983.

7. **Wolf et al:** Determining lyapunov exponent from a time series. *Physica* 16D: 285 -317,1985.

8. **Hao BL:** Symbolic dynamics and systematics of period windows. *Physica* 140 A: 85 -95,1986.

9. **Metroplis N, Stein ML & Stein PR:** On finite limit sets for transformation on the unit interval. *Journal Combinatorial Theory* A15: 25 - 43, 1973.

10. **Wolf et al:** Multiplicity in a chemical reaction with one dimensional dynamics. Phys. *Rev. Lett* 49: 245 - 248, 1982.

11. **G'Allesandro et al:** Hierarchial approach to Complexity with Applications to Dynamical Systems. Phys. *Rev. Lett* 64: 1609 - 1612, 1990.

12. **Lempel A and Ziv:** On Complexity of finite sequences. *IEEE Trans. Information 13 Theory* 22: 75 - 88, 1976.

13. **K. Kaspar, H.G. Schnster:** Easily calculable complexity of spatio-temporal patterns, *Phys.Rev A* (36): 843 - 848, 1987.

14. **Pincus SM:** Approximate entropy as a measure of system complexity. *Proc. Natl. Acad* (USA) *88:* 3017 - 3026, 1991.

15. **Kolmogorov AN:** Three approaches to the quantitative definition of information. *Information. Trans* 1: *3* - 11, 1965.

16. **Chaitin GJ:** Randomness and Mathematical proof. Scientific American 232: 47 - 52, 1975.

17. **Pincus S M et al:** Not all (possibly) "random sequences " are created equal. *Proc. Natl. Acad.* Sci (USA), 94: 3513 - 3518, 1997.

18. **Takens F:** Atas do 13. Col. brasiliero de Mathematicas (Rio deJanerio).