CHAPTER 3

FEATURE EXTRACTION TYPE AND ANALYSIS

3.1 INTRODUCTION

Feature extraction plays a major role in signal processing. Here the EEG signal is acquired on different cognitive conditions and the necessary statistical parameters are calculated. To begin with before going for nonlinear analysis, a study of linear parameters are analyzed. The following are the parameters such as Mean, Variance, Std-deviation, Median, Energy, Kurtosis, Skewness, Pearson Correlation, Histogram analysis by considering the single channel system. The objective of this study is to know the features of linear and significance of nonlinear features under cognitive state analysis with the single channel systems and the approach changes by considering bi-variants and multivariate systems.

There are several stages to process the data from any EEG experiment. They are the pre-processing steps which can be applied to the data to improve the detection of activation of events. The prerequisite for data analysis is the Signal acquisition process. It includes an experimental setup involving a predefined paradigm.

3.1.1 EEG Data Acquisition and Preprocessing

Electrical signals detected along the scalp by an EEG, they are originated from non-cerebral origin are called artifacts. EEG data is almost always contaminated enormous noise by such artifacts. The amplitude of
artifacts can be quite large relative to the size of amplitude of the cortical signals of interest. This is one of the reasons why it takes considerable experience to correctly interpret EEGs clinically.

3.1.2 Experimental Set-up

![Figure 3.1 EEG-Acquisition system](image1)

![Figure 3.2 Layout in Lab chart software](image2)
The EEG data was acquired in the R and D department of The Art of Living Foundation using AD Instruments Power Lab equipment and Lab Chart software.

Initially, the subject was asked to lie down and the cap is applied.

The subject was instructed to be in the normal state of mind for about 2 minutes.

Then, the subject was asked to perform different tasks containing baseline task, figure rotation, letter task, math task and visual counting based on the instructions set for all the tasks.

The same procedure was repeated for all the subjects taken for several number of days.

Figure 3.3 Experimental Setup for Signal-Acquisition
3.1.3  **Detail on the Protocol Design on Different Tasks (SET-1)**

The following five tasks were performed by the subjects:

a) **Baseline Task:** Preceding the stimulus program, the subject was asked not to perform a specific mental task, but to relax as much as possible, make as few movements as possible, and think of nothing in particular. This task is considered a baseline task for alpha wave production and was used as a control measure of the EEG.

b) **Letter Task:** The subject was asked to rehearse the English alphabets a to z continuously in mind without moving their lips or vocalizing them.

c) **Math Task:** The subject was asked to perform nontrivial multiplication problems, such as 89 multiplied by 67, and to solve them without vocalizing or making any physical movements. The problems were designed so that they could not be solved in the time allowed.

d) **Geometric Figure Rotation:** The subject was shown images of three-dimensional figures and asked to visualize them rotating about an axis.

e) **Visual Counting:** The subject was asked to imagine an image of black Roman numerals on a white background, and asked to visualize similar numerals being written onto a blackboard, one after another, sequentially in ascending order, the previous numeral being erased before the next being written.
3.2 EVALUATION OF LINEAR FEATURE EXTRACTION ANALYSIS METHODS

Feature extraction plays a major role in signal processing. Here the EEG signal is taken on the basis of different cognitive conditions and the necessary statistical parameters are calculated. To begin with first for the nonlinear analysis, a study of linear parameters are also analyzed. The following are the parameters such as Mean, variance, std deviation, median, energy, kurtosis, skewness, Pearson Correlation, Histogram analysis by considering the single channel system. The objective of this study is to know the features of linear and significance of nonlinear features under cognitive state analysis with the single channel systems and the approach changes by considering bi-variants and multivariate systems.

3.2.1 Mean

The mean is calculated after the removal of baseline wondering effect. The mean of the random process like EEG is defined as the average value of the all its sample function is given by the probability density function is defined as,

\[ E\{ \} \text{prob}\{ d \} d \]  \hspace{1cm} (3.1)

\[ \frac{1}{N} \sum_{i=1}^{N} x_i \]  \hspace{1cm} (3.2)
Figure 3.4  The mean is calculated after the removal of baseline wondering effect

From the mean of the data the difference in peak values are noted. Based on the activities it changes and out of all the tasks for multiplicative task it shows highest in amplitude. Shows that the person is active saying his beta activity is more.

### 3.2.2 Median

The median is one of statistical distribution parameter. D(x) is the value x such for a symmetric distribution; it is therefore equal to the mean. Given the statistical median of the random sample is defined by

\[ Y_{x} \begin{cases} \frac{N+1}{2} & \text{if N is odd} \\ (Y(N/2) + Y1(N/2))/2 & \text{if N is even} \end{cases} \]

(3.3)

And commonly denoted \( \mu_{1/2} \). Here Y and Y1 are two different signals. The efficiency of the median, measured as the ratio of the variance of
the mean to the variance of the median, depends on the sample size \( N = 2n+1 \) as

\[ \frac{\text{mean}}{\text{variance}} \approx 0.637 \]

which tends to a value of \( 2/3 \) approximately 0.637 as \( N \).

**Figure 3.5 Median**

3.2.3 **Variance**

The square of the standard deviation is given by .

\[ V = \sigma^2 \]  \hspace{1cm} (3.4)

= Is the standard deviation.
Figure 3.6 Variance Representation of different task

At the time of Base line the beta frequency of EEG have low variance however in Letter composition task and Multiplicative task shows high variance (97.819).

3.2.4 Kurtosis

Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution, that is datasets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than sharp peak. A uniform distribution is the extreme case. The kurtosis for a real signal x(n) is defined as
3.2.5 Skewness

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point. The skewness is defined for a signal as

\[ \text{Skewness} = \frac{E[x(n)]^3}{\mu^3} \]  \hspace{1cm} (3.5)

where \( \mu \) and \( \sigma \) are the mean and standard deviation respectively and \( E \) denotes statistical expectation. If the distribution is more to the right of the mean point the skewness is negative and vice versa. For a symmetric distribution such as Gaussian, the skewness is zero.
3.2.6 Pearson-Auto Correlation

Pearson Autocorrelation gives information on the correlations in time present in the signal. It is the standard autocorrelation for the given range of delays. The delay $t$ can be an array of positive integers or a single integer.

In addition the Cumulative Pearson Autocorrelation is computed for the same range of delays. Also, if the autocorrelation for delays up to the maximum given delay crosses the $1/e$ or zero level the delay of decorrelation or zero-autocorrelation, respectively, is assigned a value. The cumulative Pearson autocorrelation and the delay of decorrelation and zero autocorrelation can then be simply assigned to the respective measures if these are selected with the same set of delay values.
Figure 3.9 Pearson Autocorrelation of subjects

<table>
<thead>
<tr>
<th>TASKS</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>-0.18</td>
<td>-0.13</td>
<td>0.033</td>
<td>-0.05</td>
</tr>
<tr>
<td>T2</td>
<td>-0.08</td>
<td>-0.11</td>
<td>-0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td>T3</td>
<td>-0.15</td>
<td>-0.18</td>
<td>-0.12</td>
<td>-0.1</td>
</tr>
<tr>
<td>T4</td>
<td>-0.16</td>
<td>-0.19</td>
<td>-0.13</td>
<td>-0.11</td>
</tr>
<tr>
<td>T5</td>
<td>-0.25</td>
<td>-0.3</td>
<td>-0.26</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

Figure 3.10 Pears Autocorrelation vs delay
Pearson autocorrelation describes the correlation between values of the time series at different times, as a function of the two times or of the time lag. It gives a value between [-1,1] with 1 indicating perfect correlation, 0 is no correlation and -1 indicating negative correlation.

3.2.7 Histogram Analysis

A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous (quantitative variable) and was first introduced by Karl Pearson. To construct a histogram, the first step is to "bin" the range of values; that is, divide the entire range of values into a series of intervals and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent, and are usually equal size.

Figure 3.11(a) Histogram analysis
Figure 3.11(b) Histogram analysis

Figure 3.11(c) Histogram analysis
Histograms give a rough sense of the density of the underlying distribution of the data, and often for density estimation: estimating the probability density function of the underlying variable. The total area of a histogram used for probability density is always normalized to 1. If the lengths of the intervals on the x-axis are all 1, then a histogram is identical to a relative frequency plot.

3.3 NONLINEAR INVARIANT ANALYSIS UNDER COGNITIVE ACTIVATION

3.3.1 Motivation

The application of medical signal processing for effective noninvasive diagnosis has led to a huge growth in the application of digital signal processing techniques for solving related medical problems. Many useful phenomena in nature is due to the presence of nonlinearity. The theory of nonlinear dynamical systems, also called as 'chaos theory', has now progressed to a stage, where it is possible to study self-organization and pattern formation in the complex neuronal networks of the brain. According to Kellert (1994), chaos theory is "the qualitative study of unstable aperiodic behavior in deterministic dynamical systems". The brain is a highly complex network and important organ of a human body. The neurons interact with the local as well as the remote ones in a very complicated way (Kalpana et al. 2015).

These interactions evolve as the spatiotemporal electromagnetic field of the brain and are recorded as Electroencephalogram (EEG). It has been proven that the EEG recordings exhibit chaotic nature from experiments such as the EEG models proposed by Skarda & Freeman (1987) or Wright & Liley (1996) for chaotic dynamics to meet requirements in Kalpana et al. (2015) neurobiology. EEG data are very useful for many branches of the
neurosciences, and, many advanced experiments in cognitive neuro science have shown that EEG and evoked potentials are strongly correlated with specific cognitive tasks (Kalpana et al. 2015). Many pathologic states have been examined as well, ranging from toxic states, seizures (Pijn et al. 1997), and psychiatric disorders to Alzheimer's (Jaesung et al. 2001), Parkinson's and Creutzfeldt Jakob's disease.

3.3.2 Nonlinear Parameters and Overview

With the advent of nonlinear methods, most particularly complexity indices such as approximate entropy and sample entropy, has allowed psychologists to examine the behavior dynamics. In general, the history behind the use of nonlinear modeling in the study of human psychology. Initially, the applications of nonlinear methods derived from sources like engineering and physics is the EEG data series is chaotic in nature? The use of common methods to discover structure in behavior time series, e.g. mood ratings, obtained from individuals.

In studying neurobiological signals, it has always been a challenge how to gain information from them. It is important to find what is happening in the supposed frequency and time-related components of those signals. The results of the combined time-frequency domain analysis are very challenging.

There are procedures which can give us only time information and those which give only frequency information, but the best methods for non-stationary signals are time-frequency procedures. These can be the most useful methods applied for analyzing EEG signal an unusual problem in biomedical signal analysis is to devise techniques to understand and, more importantly, to the prediction of clinical onset in advance, epileptic seizures that affect about 1% of the population in industrialized countries.
Epileptic seizures are characterized electrographically by sudden changes at a time changes in power spectral density and increases in the EEG wave rhythmicity. The changes obtained due to brain activity, whether local or global, can be monitored via electrodes on the scalp electroencephalogram, EEG. The recordings gave a window, perhaps the only practically accessible window at present, through which the dynamics of epilepsy can be (Liu et al. 2003) investigated.

An approach that is gaining increasing attention is the application to this problem of techniques from nonlinear dynamics and chaos, originally developed for the study of low-dimensional, nonlinear, deterministic systems. The nonlinear Measures are helpful for characterization of low-dimensional chaotic systems, of EEG signals The most common tests are the Approximate Entropy, and the Lyapunov exponents, correlation dimension, have been various reasons that epileptic seizures can also be predicted and detected up to several minutes or hours before their clinical studies and diagnosis and also for sleep data analysis for knowing the cognitive activities during all the stages of sleep cycle.

3.3.3 Approximate Entropy

Approximate entropy (ApEn) is defined as the "logarithmic likelihood that runs of patterns of data that are close to each other will remain close on next incremental comparisons. ApEn is a statistical metric that which quantifies the complexity or irregularity of signals that are both deterministic and stochastic in nature (Flores et al. 2013). It reflects the rate of new pattern generation and is thus related to the concept of entropy. This method was first proposed by Pincus (1991). One of the main significance of ApEn is that it is very useful for small duration datasets, that may also be polluted by noise and interference because it is not sensitive to them. It has two important user-specified parameters: a run length m and tolerance window r small values of
ApEn imply a greater likelihood that similar measurements will follow certain patterns of measurements.

If the time-series is highly irregular, the occurrence of similar patterns in the future is less likely.

The algorithm is as follows:

1) Form a vector \( \mathbf{x} \) defined by \( x_m = \{ x_1, x_2, \ldots, x_m \} \) where the original data are \( x_1, x_2, \ldots, x_m \) and \( m \) is the total number of data points.

2) Set the distance between \( x \) and \( y \) by \( d(x, y) \), define as the maximum absolute difference between their respective scalar elements.

3) For \( d(x, y) \), find the number of \( j \) so that \( d(x, y) < \epsilon \).

4) For each \( \epsilon \), compute the natural algorithm and average it over \( \epsilon \).

5) Increase the dimension to, repeat steps 1 to 4 and find \( \epsilon \).

6) The number of data point \( N \) defined when the data as \( N \) length is \( N \) and it is denoted.
3.3.4 Results

Table 3.1 Result of approximate entropy for various cognitive tasks performed by subjects

<table>
<thead>
<tr>
<th>ApEn</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.215</td>
<td>0.201</td>
<td>0.201</td>
<td>0.201</td>
<td>0.319</td>
</tr>
<tr>
<td>S2</td>
<td>0.482</td>
<td>0.451</td>
<td>0.451</td>
<td>0.451</td>
<td>0.451</td>
</tr>
<tr>
<td>S3</td>
<td>0.083</td>
<td>0.077</td>
<td>0.231</td>
<td>0.077</td>
<td>0.077</td>
</tr>
<tr>
<td>S4</td>
<td>0.22</td>
<td>0.231</td>
<td>0.231</td>
<td>0.231</td>
<td>0.231</td>
</tr>
</tbody>
</table>

Figure 3.12 Approximate entropy of subjects

m and r are critical in determining the outcome of ApEn; no of guidelines exist for optimization of their values. By principle, the accuracy and confidence of the entropy estimate improve as the number of matches of length m and m+1 increases. The number of events can be taking increasing values by choosing small m (short templates) and large value of r (wide
tolerance). However, there are consequences for deciding the criteria that are too relaxed (Pincus 1991). For smaller r values, one usually achieves weak conditional probability estimates, while for larger values of r Figure 3.13. The EEG waves on different cognitive states.

![Figure 3.13 Conditional probability estimates](image)

To avoid a significant contribution of noise in an ApEn calculation, one must choose r bigger than most of the noise (Pincus 1991). It is possible by selecting the appropriate tolerance threshold value r.

For this study, m is set to 2 and r is set to 15% of the standard deviation of each time series.

These values are chosen here (Kalpana et al. 2015) are on the basis of prior studies indicating good statistical validity for ApEn within these variable ranges (Pincus & Goldberger 1994).
3.3.5 Correlation Dimension

Correlation dimension describes the dimensionality of the underlying process concerning its geometrical reconstruction in phase space. It is used for the detection of chaotic behavior in dynamical systems (Sprott & Rowlands 2001). Since in principle D2 converges to finite values for deterministic systems and does not converge in the case of a random signal, D2 is a useful parameter for evaluating the deterministic or noisy inherent nature of a system. Grassberger & Procaccia (1983) algorithm (GPA) computes correlation dimension based on the following approximation: The probability that two points of the set are in the same cell of dimension r is approximately equal to the probability that two points of the set are separated by a distance less than or equal to r.

Thus \( C(r) \) is approximately given by:

\[
CD \lim_{r \to 0} \frac{\log C(r)}{\log(r)}
\]  

(3.6)

where \( N \) is the number of data points and the Heaviside function \( H \) is defined as:

\[
C(r) \frac{1}{(N_{\min}) (N_{\min} - 1)} \frac{N}{N} \left( r \mid X_X \quad X_y \right)
\]

(3.7)

where \( XX \) and \( Xy \) are the points of the trajectory in the phase space, for \( N \) number of data points or the radial distance around each reference point \( Xi \), the Heaviside function \( min \) the average correlation time.

For example, if we have a set of random points on the real number line between 0 and 1, the correlation dimension will be \( = 1 \), while if they are
distributed on say, a triangle embedded in three-dimensional space (or m-dimensional space), the correlation dimension will be $= 2$. This is what we would intuitively expect from a measure of dimension. The real utility of the correlation dimension is in determining the (possibly fractional) dimensions of fractal objects. The correlation dimension has a advantage of being directly and quickly calculated, of being less noisy when only a small number of points is available and is often in agreement with other calculations of dimension. If the number of points is sufficiently large, and evenly distributed, a log-log graph of the correlation integral versus will yield an estimate of $D$.

This idea can be qualitatively understood by realizing that for higher-dimensional objects, there will be more ways for points to be close to each other, and so the number of pairs close to each other will rise more rapidly for higher dimensions. Correlation Integral, a measure originally proposed by Grass Berger and Procaccia, which has become one of the most popular nonlinear dynamics based tools in the analysis of EEG/ECoG data.

A previous study has demonstrated that decreases in correlation dimension are predictive of seizure onset. The driving force behind a large number of studies on dimension analysis of EEG/ECoG is that epileptic seizures are regarded as emergent states with reduced dimensionality compared to non-epileptic activity.

This concept finds support in the observation that neuronal hyper synchrony underlies seizures; a phenomenon during which the number of independent variables required to describe. Thus, measures that detect reduced dimensionality of EEG may allow for prediction of seizures.
Table 3.2 Result of correlation dimension for various cognitive tasks performed by subjects

<table>
<thead>
<tr>
<th>Correlation Dimension</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.73</td>
<td>0.758</td>
<td>0.781</td>
<td>0.758</td>
<td>0.718</td>
</tr>
<tr>
<td>S2</td>
<td>0.736</td>
<td>0.719</td>
<td>0.777</td>
<td>0.752</td>
<td>0.725</td>
</tr>
<tr>
<td>S3</td>
<td>0.789</td>
<td>0.769</td>
<td>0.982</td>
<td>0.769</td>
<td>0.774</td>
</tr>
<tr>
<td>S4</td>
<td>0.785</td>
<td>0.717</td>
<td>1.1</td>
<td>0.718</td>
<td>0.936</td>
</tr>
</tbody>
</table>

Figure 3.14 Graphical representation of correlation dimension for various cognitive tasks performed by subjects

Correlation dimension is a sensitive parameter in the analysis of electrical brain activity. The technique can be used to distinguish between (deterministic) chaotic and truly random behavior. It quantifies the variability in a time series.
As can be inferred from Table 3.3 there are slight differences in the correlation dimension of the subjects between the tasks. Nevertheless the table indicates that the letter composition activity (T3) is consistently higher despite the relatively small changes. The correlation dimension of most mental activities appeared to be higher compared to the relaxation condition. Hence it is concluded that cognitive and mental activity is associated with a higher correlation dimension in the EEG.

Table 3.3 Experimental Results

<table>
<thead>
<tr>
<th>Bicorrelation at t=1</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>-0.0195</td>
<td>0.0025</td>
<td>0.034</td>
<td>0.0541</td>
<td>0.178</td>
</tr>
<tr>
<td>S2</td>
<td>-0.070</td>
<td>-0.327</td>
<td>-0.134</td>
<td>-0.145</td>
<td>0.087</td>
</tr>
<tr>
<td>S3</td>
<td>-0.003</td>
<td>-0.089</td>
<td>-0.018</td>
<td>-0.079</td>
<td>0.019</td>
</tr>
<tr>
<td>S4</td>
<td>-0.112</td>
<td>-0.199</td>
<td>-0.073</td>
<td>-0.135</td>
<td>0.047</td>
</tr>
</tbody>
</table>

3.3.6 Bi-Correlation

Bicorrelation is a nonlinear third-order correlation measure. It is used as a statistic for the test of linearity or nonlinearity called the Hinrich test in the time domain. There is no predefined range of expected values for bicorrelation and cumulative bicorrelation, other than [-1,1]. If the EEG time series shows a higher value of bicorrelation it is said to be linear. A decrease in the value of bicorrelation corresponds to the EEG time series being nonlinear in nature. From Figure 3.15 the activity of imagining an object rotating (T2) has negative correlation and visual counting activity (T5) has positive correlation thus showing that T2 is more complex and random and T5 is more predictable.
3.3.7 Hurst Exponent

The Hurst exponent is a measure that has been widely used to evaluate the self-similarity and correlation properties of fractional Brownian noise, the time-series produced by a fractional (fractal) Gaussian process. It is used to evaluate the presence or absence of long-range dependence and its degree in a time series. The Hurst exponent is the measure of the smoothness of a fractal time series based on the asymptotic behavior of the rescaled range of the process. However, non-stationerities are often present in physiological data and may compromise the ability of some methods to measure self-similarity. The Hurst exponent $H$ is defined as:

\[
H = \lim_{t \to \infty} \frac{\log R(t)}{\log t}
\] (3.8)
where $T$ is the duration of the sample of data and $R/S$ the corresponding value of rescaled range. The above expression is obtained from the Hurst's generalized equation of time series that is also valid for Brownian motion. A Hurst exponent, $H$, between 0 to 0.5 is said to correspond to a mean reverting process (anti-persistent), $H=0.5$ corresponds to Geometric Brownian Motion (Random Walk), while $H >= 0.5$ corresponds to a process which is trending (persistent). $H$ is related to the fractal dimension $D$ given by: $H = E + 1 - D$ (3) where $E$ is the Euclidean dimension.

Figure 3.16 Hurst exponent graph

3.3.8 Phase-Space Plot

We construct the EEG attractors of all five kinds of mental activities of 4 subjects and find that EEG attractors of various patterns have similar characteristics. As can be seen from Figure 3.17, the attractors' track often rotates in an extremely complex way but there is still internal structure when the attractors is magnified. The attractors of resting, composition of a letter and visualizing a 3-dimensional object being rotating about an axis.
often distribute in a small ellipse region, while the point in the attractors of mental arithmetic and visualizing numbers being

![EEG attractors](image)

**Figure 3.17** EEG attractors of five kinds of cognitive tasks of a subject in the order of Relaxation, Visualizing a 3D figure rotating, Mental composition of a letter, Arithmetic and Visualizing numbers being written or erased on a blackboard

Construct the EEG attractors of all five kinds of mental activities of 4 subjects and find that EEG attractors of various patterns have similar characteristics. As can be seen from Figure 3.17, the attractors' track often rotates in an extremely complex way but there is still internal structure when the attractors is magnified. The attractors of resting, composition of a letter and visualizing a 3-dimensional object being rotating about an axis often distribute in a small ellipse region, while the point in the attractors of mental arithmetic and visualizing numbers being written centralize nearby the 45 degree line and there is a large distributing range along the 45 degree line. This is because during rational computation such as mathematics or
imagination, the value of the adjacent sampling points of EEG signals are close, and the amplitude values of the whole EEG signals are great.

3.3.9 **Significance of Multichannel Data Analysis**

The execution of the task is usually not because of any single stimulus; rather there is involvement of multiple stimuli. For example, if we consider the speech activity, it includes the motor activity of the lip and the tongue. The memory region is activated for want of words which involves the activity in the parietal lobe of the brain. The voice modulations are dependent on the hearing capability. The activity in the auditory region of the brain is also involved. This clearly indicates the involvement of multiple stimuli. The stimuli can be categorized into main activity stimulus, which in the above-discussed case would be speech and the stimuli responsible for accompanying activities would be memory and auditory activities.

The analysis of multiple stimuli is thus a challenging area of research since it gives details of the accommodative character and efficient execution of various parallel activities by the human brain. Even though multiple regions are responsible for the execution of the primary task, the features of only the main activity are extracted for activity pattern classification and comparison pattern procedures.

3.3.10 **Non-Linear Measurement and Analysis under Diseased Conditions**

The brain is a highly complex and vital organ of a human body whose neurons interact with the local as well as the remote ones in a very complicated way these interactions evolve as the Spatio-temporal electromagnetic field of the brain, and are recorded as Electroencephalogram (EEG). It has been established that EEG recordings exhibit chaotic behavior. The
theory of nonlinear dynamical systems, also called 'chaos theory', has now progressed to a stage. Here, it is possible to study self-organization and pattern formation in the complex neuronal networks of the brain. According to Kellert, chaos theory is “the qualitative study of unstable a periodic behavior in deterministic dynamical systems. Some of the common parameters are, approximate entropy, Hurst-exponent, Bi-correlation.

3.3.11 Approximate Entropy (ApEn)

In statistics, approximate entropy (ApEn) is a technique used to quantify the amount of regularity and the unpredictability of fluctuations over time-series data.

\[
\text{ApEn} = \frac{\log C \times \log C_j}{(3.9)}
\]

The presence of repetitive patterns of fluctuation in a time series renders it more predictable than a time series in which such patterns are absent. ApEn reflects the likelihood that similar patterns of observations will not be followed by additional similar observations. A time series containing many repetitive patterns has a relatively small ApEn; a less predictable process has a higher ApEn. Epilepsy from the focal region has the highest approximate entropy; hence it is less predictable than the EEG from healthy subjects as indicated by the low value of ApEn is 0.032 for O-data-set much greater than 0.002 for N-data-set.

3.3.12 Bi-Correlation

Bi-Correlation is a statistical measure that indicates the extent to which two or more variables fluctuate together. A positive correlation indicates the extent to which those variables increase or decrease in parallel; a negative correlation indicates the extent to which one variable increases as the
other decreases. When the fluctuation of one variable reliably predicts a similar fluctuation in another variable, there’s often a tendency to think that means that the change in one causes the change in the other. However, correlation does not imply causation. There may be, for example, an unknown factor that influences both variables similarly data set shows the highest bi-correlation value, indicating their variables are increasing over the entire duration of recording of EEG. Bi-correlation, or three-point autocorrelation, or higher order correlation, is the joint moment of three variables formed from the time series and two delays \( t \) and \( s \). A simplified scenario for the delays is implemented, \( s = 2t \),

Formulae for calculating Bi-correlation

\[
E [x(i), x(i+t), x(i+2t)]
\]  \hspace{1cm} (3.10)

So the bi-correlation is where the mean value is estimated by the sample average. In this way, the bi-correlation is a function of a single delay \( t \).

![Figure 3.18 Bi-correlation and ApEn comparison](image)

Figure 3.18 Bi-correlation and ApEn comparison
Figure 3.19 Bi-correlation values

Figure 3.20 Bi-correlation for depression data set
Even in the segmented time-series analysis O-data set shows the higher values of change, EEG from a healthy subject has low bi-correlation and low approximate entropy values. Observe the comparison between variations in bi-correlation and ApEn values, a small change in ApEn causes a huge change in magnitude of bi-correlation values.

![Figure 3.21 Time-Series plot of Epilepsy from Focal Region and Seizure plot](image)

3.3.13 Hurst Exponent

The Hurst exponent is used as a measure of long-term memory of time series. It relates to the autocorrelations of the time series, and the rate at which these decrease as the lag between pairs of values increases. Studies involving the Hurst exponent were originally developed in hydrology for the practical matter of determining optimum dam sizing for the Nile River’s volatile rain and drought conditions that had been observed over a long period of time. The name "Hurst exponent", or "Hurst coefficient", derives from
Harold Edwin Hurst, who was the lead researcher in these studies; the use of the standard notation H for the coefficient relates to his name also.

Figure 3.22 Hurst exponent for epilepsy from focal region

Figure 3.23 Hurst exponent for seizures
The Hurst exponent is used as a measure of long-term memory of time series. It relates to the autocorrelations of the time series, and the rate at which these decrease as the lag between pairs of values increases. Studies involving the Hurst exponent were originally developed in hydrology for the practical matter of determining optimum dam sizing for the Nile River’s volatile rain and drought conditions that had been observed over a long period of time. The name "Hurst exponent", or "Hurst coefficient", derives from Harold Edwin Hurst, who was the lead researcher in these studies; the use of the standard notation H for the coefficient relates to his name also.

**The expression for calculating Hurst exponent**

The Hurst exponent H is defined as:

\[
H = \frac{\log \frac{\log \left( \frac{R}{S} \right)}{\log T}}{\log \left( \frac{R}{S} \right)},
\]

where T is the duration of the sample of data and R/S the corresponding value of rescaled range. The Hurst exponent is referred to as the "index of dependence" or "index of long-range dependence". It quantifies the relative tendency of a time series either to regress strongly to the mean or to cluster in a direction. A value H in the range 0.5–1 indicates a time series with long-term positive autocorrelation, meaning both that a high value in the series will probably be followed by another high value and that the values a long time into the future will also tend to be high.

A value in the range 0-0.5 indicates a time series with long-term switching between high and low values in adjacent pairs, meaning that a single high value will probably be followed by a low value and that the value after that will tend to be high, with this tendency to switch between high and
low values lasting a long time into the future. A value of $H = 0.5$ can indicate a completely uncorrelated series, but in fact it is the value applicable to series for which the autocorrelations at small time lags can be positive or negative but where the absolute values of the autocorrelations decay exponentially quickly to zero.

### 3.4 SUMMARY

Both bi-correlation and Pearson autocorrelation are measures of testing linear or nonlinear correlation in the time series. From the results, the authors think that a higher value corresponds to linear nature and a decrease in value to nonlinear nature of the EEG time series of various mental states. The self-similarity parameter (Kalpana et al. 2015) is higher when the subject is performing tasks which require greater consciousness than the rest owing to a smoother time series without noise. The phase space plots show unique patterns for each consciousness state useful in computing correlation dimension.

The above analyses indicate that different cognitive activities have profound nonlinear dynamic characteristics. Some differences are difficult to perceive, and the linear and nonlinear quantitative parameters of different individuals have great differences. Hence it is a critical problem to find a widely applicable criterion, which needs to be explored for a long time.