Analysis of Observed Chaotic Data

By

Mehmet Emre ÇEK

**A Dissertation Submitted to the
Graduate School in Partial Fulfillment of the
Requirements for Degree of**

MASTER OF SCIENCE

**Department: Electrical and Electronics Engineering
Major: Electronics and Communication**

**İzmir Institute of Technology
İzmir, Turkey**

July, 2004

We approve the thesis of **Mehmet Emre ÇEK**

Date of Signature

..... **29. 07. 2004**

Prof. Dr. Ferit Acar SAVACI

Supervisor

Department of Electrical and Electronics Engineering

..... **29.07.2004**

Prof. Dr. Cüneyt GÜZELİŞ

Department of Electrical and Electronics Engineering

Dokuz Eylül University İzmir

..... **29.07.2004**

Assist. Prof. Dr. Mustafa Aziz ALTINKAYA

Department of Electrical and Electronics Engineering

..... **29.07.2004**

Prof. Dr. Ferit Acar SAVACI

Head of Department

Department of Electrical and Electronics Engineering

ACKNOWLEDGEMENT

I am very grateful to my supervisor Prof. Dr. Acar Savacı for his support and guidance throughout my research. His comments and advises directed me to do best.

I would like to thank to Dokuz Eylül University Department of Biophysics and especially to Prof. Dr. Erol Başar for giving permission the EEG signals to be used in the thesis and I also thank to Dr. Murat Özgören for his help during the EEG experiments.

I also offer my special thanks to Dr. Nalan Özkurt due to her contributions during the computer simulations and usage of software packages.

And finally, I wish to express my gratitude to my family, the source of my motivation. To finish this thesis in success would be harder without their continued support and encouragement.

ABSTRACT

In this thesis, analysis of observed chaotic data has been investigated. The purpose of analyzing time series is to make a classification between the signals observed from dynamical systems. The classifiers are the invariants related to the dynamics. The correlation dimension has been used as classifier which has been obtained after phase space reconstruction. Therefore, necessary methods to find the phase space parameters which are time delay and the embedding dimension have been offered. Since observed time series practically are contaminated by noise, the invariants of dynamical system can not be reached without noise reduction. The noise reduction has been performed by the new proposed singular value decomposition based rank estimation method.

Another classification has been realized by analyzing time-frequency characteristics of the signals. The time-frequency distribution has been investigated by wavelet transform since it supplies flexible time-frequency window. Classification in wavelet domain has been performed by wavelet entropy which is expressed by the sum of relative wavelet energies specified in certain frequency bands. Another wavelet based classification has been done by using the wavelet ridges where the energy is relatively maximum in time-frequency domain.

These new proposed analysis methods have been applied to electrical signals taken from healthy human brains and the results have been compared with other studies.

ÖZ

Bu tezde, gözlenmiş kaotik işaretlerin analizi incelenmiştir. Zaman serilerinin analizinin amacı dinamik sistemlerden gözlenen işaretler arasında sınıflandırma yapmaktır. Bu sınıflandırıcılar dinamik sistemlere ilişkin değişmezlerdir. Faz uzayı yapılandırmasıyla elde edilen korrelasyon boyutu sınıflandırıcı olarak kullanılmıştır. Bu nedenle faz uzayı yapılandırmasının parametreleri olan, zaman gecikmesi ve gömme boyutunun, bulunmasına ilişkin yöntemler sunulmuştur. Pratikte gözlenmiş olan işaretlere gürültü bulaştığından dinamik sistemin değişmezleri, gürültü filtrelenmeden bulunamaz. Gürültü filtreleme, yeni önerilen tekil değer ayrıştırması tabanlı rank kestirim yöntemiyle sağlanmıştır.

Bir diğer sınıflandırma, işaretlerin zaman-frekans analizi ile gerçekleştirilmiştir. Zaman-frekans dağılımının bulunmasında, dalgacık dönüşümü değişken zaman-frekans pencereleri sunduğu için tercih edilmiştir. Dalgacık ekseninde sınıflandırma, işaretin belli frekans aralıkları içindeki göreceli dalgacık enerjilerinin toplamı cinsinden ifade edilen dalgacık entropisi ile sağlanmıştır. Diğer bir dalgacık tabanlı sınıflandırma, zaman-frekans ekseninde, enerjinin göreceli olarak maksimum olduğu dalgacık tepeleri kullanılarak yapılmıştır.

Bu önerilen yeni sınıflandırma yöntemleri sağlıklı insanların beyinlerinden alınan elektriksel işaretlere uygulanmış ve sonuçlar varolan diğer çalışmalarla karşılaştırılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADD	Attention Deficit Disorder
AON	Anterior Olfactary Nucleus
ARMA	Auroregressive Moving Average
CNS	Central Nervous System
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyogram
EP	Evoked Potentials
ERP	Event Related Potentials
FNN	False Nearest Neighbour
MCN	Modified Combinatorial Nomenclature
OB	Olfactory Bulb
PC	Prepyriform Cortex
PCA	Principal Component Analysis
REM	Rapid Eye Movement
RWE	Relative Wavelet energy
SMR	Sensory Motor Response
SNR	Signal to Noise Ratio
STFT	Short Time Fourier Transform
SVD	Singular Value Decomposition
WE	Wavelet Entropy
WT	Wavelet Transform

CHAPTER 1

INTRODUCTION

The nonlinear signal processing and analysis of observed data have become an important scientific area. In the introduction of the thesis, the fundamental definitions and concepts about observed chaos described in [1] and the nonlinear time series methods have been given.

A measurable variable of a system, observed continuously in time, may be regarded as an information signal. Sampling this signal, either on a time or event basis and recording the measured values sequentially produces a discrete signal, or time series. For instance, atmospheric pressure may vary continuously throughout the day, but recording the pressure hourly would yield time series.

The process of analyzing time series using mathematical and numerical data transformations or even appropriate graphical displays constitutes a scientific method known as *time-series analysis*. Conventional signal-processing techniques include Fourier transforms, autocorrelation functions and autoregressive data modelling, but these methods generally rely on linear relations and often have been found insufficient in describing the nonlinear structure in chaotic time series [2].

Chaotic time-series analysis, or *nonlinear time-series analysis*, refers to a class of data-analysis techniques employed to provide a richer description of chaotic time series obtained from nonlinear dynamical systems [2]. Chaos is an aperiodic long time behaviour in a deterministic system that exhibits sensitive dependence on initial conditions [2] where aperiodic long time behaviour means that there exist trajectories which do not settle down to fixed points or periodic orbits as time tends to infinity. The deterministic system is the system that has no random or noisy inputs or parameters. Therefore irregular behaviour arises from the system's nonlinearity [3].

Chaos is observed in various types of areas such as; meteorology, aerodynamics and turbulence modelling, chemical reactions (Belousov-Zhabotinskii reaction), nonlinear electronic circuits (Van Der Pol Oscillator, Chua's Circuit), ecology, biology and population evolution, observed time series like electrocardiogram (ECG), electroencephalogram (EEG), and financial data [4].

Aim of this thesis is to apply the nonlinear signal processing methods to see whether classification can be done using the time series. It is expected to obtain this classification by the invariants of dynamical systems. Most of the natural signals exhibit a chaotic behaviour; therefore it is possible to find some numerical invariant values which distinguish the signals having different characteristics. These invariants are fractal dimensions, Lyapunov exponents and entropy values.

In the introduction, the difference between linear and nonlinear signal processing, phase space reconstruction and time series analysis methods have been explained. The results indicate the different characteristics between deterministic and stochastic time series.

Before starting the analysis of observed signal, comparison between linear and nonlinear signal processing techniques has been given in Table 1.1 [1].

Table1.1 Comparison of Linear and Nonlinear signal Processing.

Linear Signal Processing	**Nonlinear Signal Processing**
Obtaining Clean Signal Broadband Noise is separated from narrowband signal. To use matched filter in frequency domain is sufficient for separation.	**Obtaining Clean Signal** Broadband noise is separated from broadband deterministic signal by using blind source separation techniques.
Finding The space Fourier transform is used to turn differential equations into algebraic forms. s(n) is observed; $s(f)=\sum s(n)\exp[j2\pi nf]$ is used.	**Finding the Space** Time lagged variables form phase space in d_E dimensions. (d_E is the embedding dimension, T is the mutual information) $\mathbf{y}(n)=[s(n),s(n+T),s(n+2T)....,s(n+(d_E-1)T)]$
Classification Sharp spectral peaks, that is, resonance frequencies of the system are the classification parameters.	**Classification** Classification parameters are invariants of orbits. These are Lyapunov exponents, fractal dimensions and entropy values.

In Table 1.1 the critical point is that the invariant quantities, which are classification parameters, are independent of initial conditions for both linear and nonlinear signal processing.

Since chaotic signal and noise are both broadband, the difference can not be expressed by Fourier spectrum because of a chaotic signal seems similar to noise in the frequency spectrum. As an example, the Fourier spectrum of Lorenz signal has been illustrated in logarithmic scale where the Lorenz system is described by the following equations,

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \tag{1.1}$$

where σ=16, b = 4 and r = 45.92 [1].

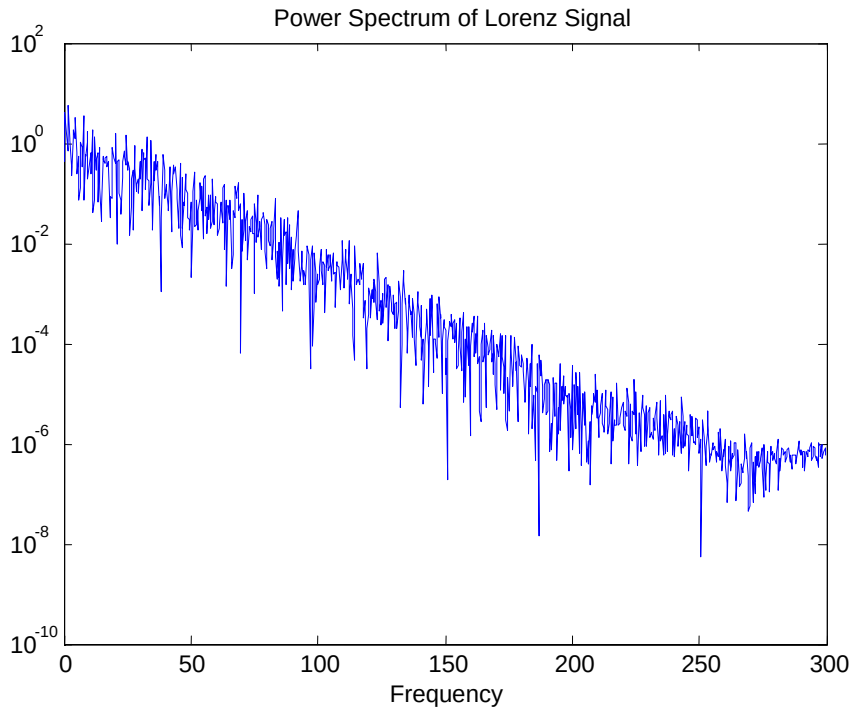

Figure 1.1 Power spectrum of Lorenz signal x(t).

The Fig 1.1 clearly states that to derive an invariant feature from power spectrum for chaotic signals is not possible. Therefore, it will be useful to give the basics of constructing phase portrait for characterizing observed chaotic signals.

1.1 Observed Chaos

Chaos is observed as a feature of **x**(t) arising from nonlinear time evolution rules which are described in the form of differential equations

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{F}(\mathbf{x}(t)) \tag{1.2}$$

with three or more degrees of freedom where $\mathbf{x}(t)=[x_1(t), x_2(t), x_3(t), ..., x_d(t)]$ or with discrete time maps

$$\mathbf{x}(n+1) = \mathbf{F}(\mathbf{x}(n)) \quad n \in \mathrm{Z}. \tag{1.3}$$

The Degree of freedom in the systems which are characterized by ordinary differential equations is expressed by the number of required first order autonomous ordinary differential equations. In dimension two, only the fixed points or limit cycles are possible and chaos can be observed for three or higher dimensions for continuous time dynamical systems described by Eq. 1.2.

The transition from scalar observations to multivariate phase space is explained by embedding theorem which states that if the scalar observations of a vector function **x**(n) which is composed by **F**(**x**(n))=**x**(n+1), is known, the geometric structures of multivariate dynamics can be unfolded from these observations and expressed by new vector ***y***(n) in d_E-dimensional Euclidian space where ***y***(n) is in the form as,

$$\boldsymbol{y}(n)=[s(n), s(n+T),s(n+2T),...,s(n+(d_E-1)T)] \tag{1.4}$$

where T is the common time lag and $s(n)$ is the scalar measurement which is one of the components of vector function **x**(n). As clearly seen in Eq 1.4, the reconstruction of phase space depends on answers of two questions: What time lag and what dimension should be use for reconstruction? If the dimension d_E is not selected necessarily large, there will be intersections of the orbit because of overlapping projections from higher dimensional space to lower dimensional space. According to the embedding theorem [5], if the selected dimension d satisfies the condition $d_E>2d_A$ where d_A is the attractor dimension of the orbit, then d is sufficient to undo all overlaps of the orbit [1]. In some

cases, d can be chosen less than the sufficient dimension d_E which is necessary dimension. Parameter d can be determined by several algorithms. As an example for Lorenz system, fractal dimension d_A is 2.06, therefore the sufficient dimension is at least 5 but there is no intersection in the orbit for dimension 3 and this necessary dimension is enough to unfold the attractor. The algorithms to determine the embedding dimension and time delay have been given in chapter 2.

1.2 Linear Time Series Analysis Methods

With experimental data, the fundamental dynamical variables are not known even the number of them. Another important issue is to decide how frequently to sample the data. Undersampling will miss important dynamics and will cause aliasing in which spurious low frequency components are introduced. Therefore, short record is not preferred. Sampling too frequently taxes the memory limits and slows the calculation for many of the tests. Besides experimental time series will be contaminated by measuring and rounding errors when represent an analogue signal digitally [6].

Most conventional time series analysis methods assume that data come from linear dynamical system with various degrees of freedom and some added noise. The time domain signal is considered as a superposition of sine waves or exponentials which grow or decay [6]. The following methods can be used to characterize the time domain signals.

1.2.1 Stationarity

The scalar time series $x(n) \in R$, exhibit nonstationary behaviour when the mean,

$$\mu = \frac{1}{N}\sum_{n=1}^{N} x(n) \tag{1.5}$$

and the *standard deviation* (root mean square deviation),

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x(n)-\mu)^2} \tag{1.6}$$

change with time. The quantity in the square root in Eq.1.6 is called *variance*, and hence the standard deviation measures the "width" of the distribution. If these quantities are tested for different segments of time series, for example, for first and second halves of the time series, the threshold for comparing the difference will be standard errors, which are expressed as $\sigma/\sqrt{N}$. If the difference of means between two segments is larger than the standard error, it is nonstationarity [6]. Since the standard deviation involves the second order moment of the distribution because of the square in Eq.1.6, it is sensitive to points far from the mean and in some cases it may not converge as the number points increase. *Average deviation* is more robust which is in the form,

$$Average\ deviation = \frac{1}{N}\sum_{n=1}^{N} | x(n) - \mu | \quad . \tag{1.7}$$

Another parameter is skewness which is the third order cumulant of the sequence and the measure of the symmetry of the data about the mean. An estimate of it is given as,

$$skewness = \frac{1}{N}\sum_{n=1}^{N}\left[\frac{x(n)-\mu}{\sigma}\right]^3 \tag{1.8}$$

and an estimate of the fourth order cumulant, *kurtosis*, is

$$kurtosis = \frac{1}{N}\sum_{n=1}^{N}\left[\frac{x(n)-\mu}{\sigma}\right]^4 - 3 \quad . \tag{1.9}$$

Kurtosis measures the peakedness of the distribution relative to a normal (Gaussian) distribution. A distribution with positive kurtosis is called *leptokurtic* or *fat tailed*, and a distribution with negative kurtosis is called *platykurtic* [6].

1.2.2. Probability Density

Another analysis methods is to plot the probability density function of a time series x(t). A sample estimate of the probability density function *P(X)* which is called as the histogram of the data, can be obtained by partitioning the interval from minimum to the maximum of x(t) into some specified numbers which may be $\sqrt{N}$ where N is the data number in the time series [6]. For time series to be *strictly stationary*, *P(X)* must remain constant in time for all x(t) within the statistical uncertainties. Figure 1.2 illustrates the histogram of a realization of Gaussian noise and Lorenz system, respectively.

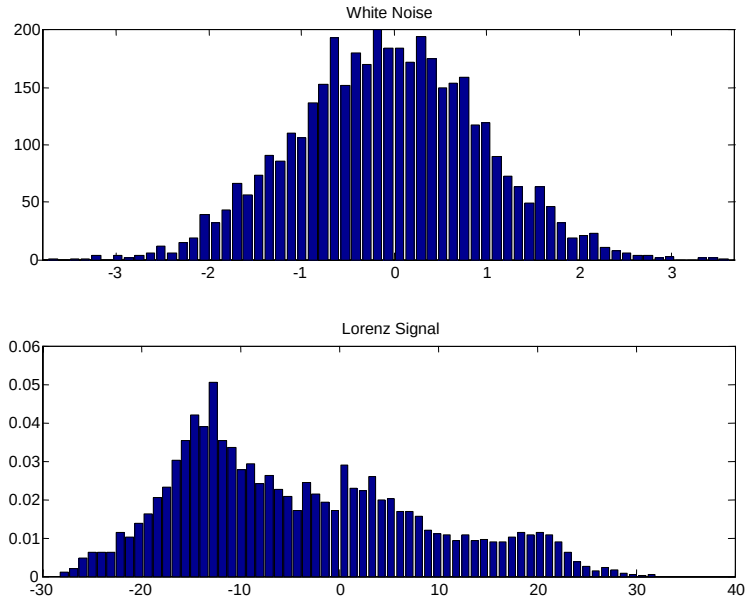

Figure 1.2 Histograms of single realizations of Gaussian noise and Lorenz signals.

1.2.3 Detrending The Data

To reduce the nonstationary of the time series, the data is detrended [6]. If the nonstationarity is in the mean, some smooth function which is a simple least-squares fit such as $f(n)=A+Bn$ is subtracted from the data. The first and second half of the data after detrending will have zero mean. The values A and B can be found as,

$$A = \mu - \frac{N+1}{2} B \quad , \tag{1.10}$$

$$B = \frac{6}{N(N^2-1)} \sum_{n=1}^{N} (2n - N - 1)(x(n) - \mu) \quad . \tag{1.11}$$

Another common method, where all the values of time series are positive, is to compute new time series *y*(n) by taking the *log first differences*,

$$y(n) = \ln(x(n+1)) - \ln(x(n)) = \ln \frac{x(n+1)}{x(n)} \cong \frac{x(n+1) - x(n)}{x(n)} \quad . \tag{1.12}$$

1.2.4. Autocorrelation Coefficient

The correlation coefficient measures how strong a data point is correlated with another point which is *k* time steps away. The formula is in the form,

$$G(k) \cong \frac{\sum_{n=1}^{N-k} (x(n) - \mu)(x(n+k) - \mu)}{\sum_{n=1}^{N-k} (x(n) - \mu)^2} \quad . \tag{1.13}$$

It is the ratio of the autocovariance to the variance of the data. Generally, it falls from the value of 1 at k=0 to zero at large k. The value of k which drops the correlation function to $1/e \approx$ %37 is accepted as *correlation time* τ_c [6]. When x(n) is nearly periodic, the correlation function will be fast decaying where τ_c will be the time for envelope to decay to $1/e$. The correlation function is symmetric about k=0, and the width is considered as $2\tau_c$. It is a measure of how much memory is hold by the system. The reciprocal of the width which is $0.5/\tau_c$, is an estimate of average rate at which predictability is lost. This is also an important quantity because its value is often similar to the largest Lyapunov exponent. In Figure 1.3 below, the autocorrelation function of Lorenz system and its envelope is illustrated. The system has been sampled at 128Hz and the estimated correlation time has been measured as 0.3516sec. The estimated largest Lyapunov exponent λ_1 has been calculated as 1.42 whereas the exact value is 1.54.

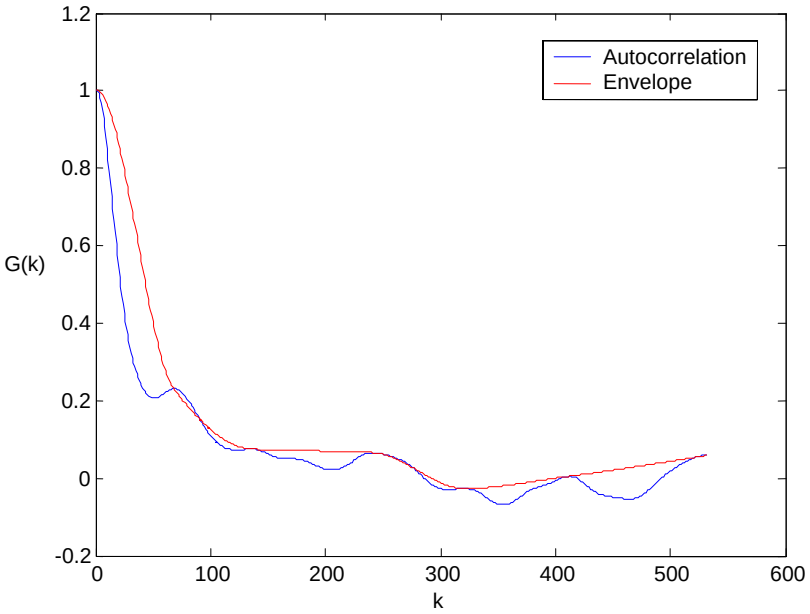

Figure 1.3 Autocorrelation coefficients of Lorenz signal and its envelope.

1.2.5 Hurst Exponent

Hurst exponent enables to set up the level of chaos and fractal dimension of time series. It determines the rate of chaos. If the probability distribution of the system is not normal, a nonparametric analysis is needed. Hurst exponent is such a method in the case of time series. Fractal time series can be distinguished from random time series. If the value of H = 0.5, the time series is normally distributed, if $H_i < H_{i+1}$, where I is the time step, the time series is antipersistent or mean reverting time series [7]. When $H_i > H_{i+1}$, then the time series is persistent or trend reinforcing time series. Hurst exponent can measure how "jagged" a time series is. If the H value is lower, the time series is more jagged. If the H value is higher, the trend is more apparent, and the time series is less jagged. The fractal dimension is calculated as a value D = 2 – H and 0<H<1. Hurst algorithm for obtaining this exponent from a time series with length *N* is as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} \mathrm{x}(i) \qquad , \tag{1.14}$$

$$\boldsymbol{X}(t,N)=\sum_{\boldsymbol{i}=1}^{\boldsymbol{t}}(\mathrm{x}(i)-\mu)\quad t=1,2,...,\mathrm{N}\quad , \tag{1.15}$$

$$\boldsymbol{S}(N)=\sqrt{\frac{1}{N}\sum_{\boldsymbol{i}=1}^{\boldsymbol{N}}(x(i)-\mu)^2}\quad , \tag{1.16}$$

$$R(N)=\max_{1<t<N}X(t,N)-\min_{1<t<N}X(t,N)\quad , \tag{1.17}$$

$$\boldsymbol{H}(\boldsymbol{N})=\log(\boldsymbol{R}(\boldsymbol{N})/(\boldsymbol{S}(\boldsymbol{N})))/\log(\boldsymbol{N})\quad . \tag{1.18}$$

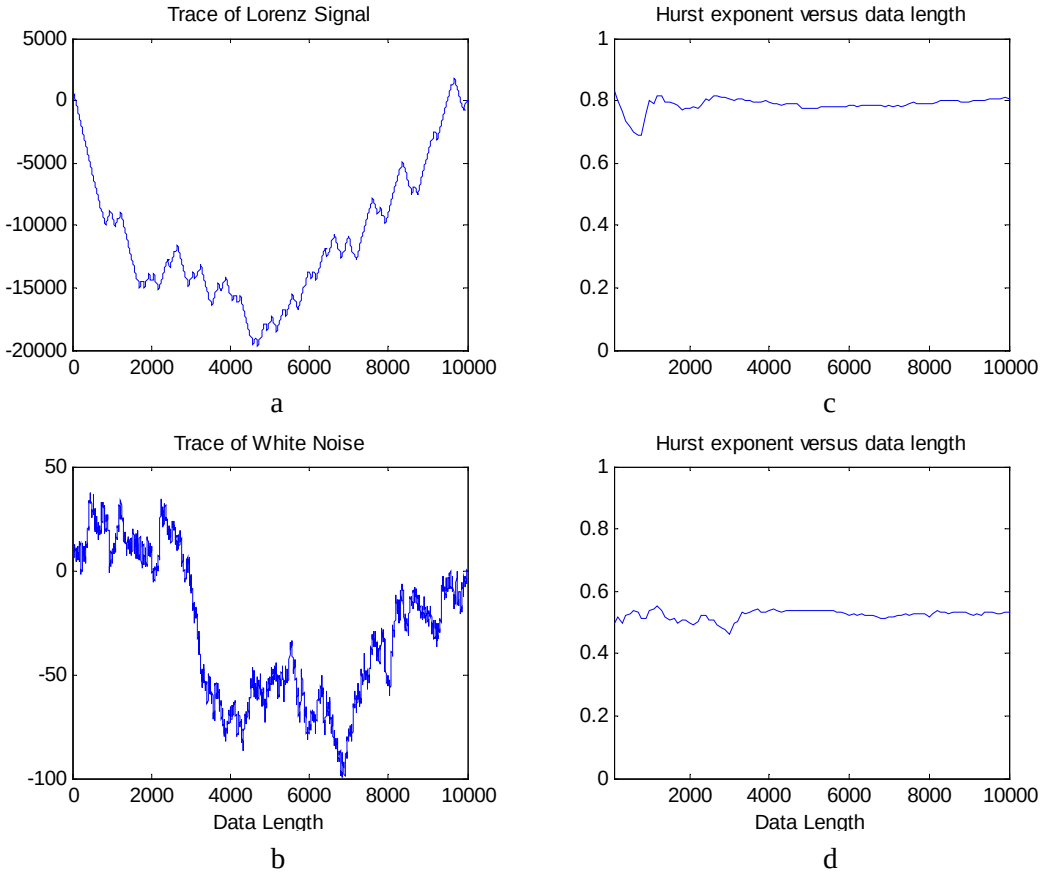

Figure 1.4 a)The trace of Lorenz signal b)The trace white noise c)Hurst exponents of Lorenz signal d)Hurst exponent of white noise.

In Figure 1.4, the Hurst exponent has been calculated as 0.53 for white noise and 0.80 for Lorenz signal. It corresponds to fractal dimensions as 1.47 and 1.20 respectively. The traces of Lorenz signal and noise which have been computed by Eq.

1.15, have been shown in Fig 1.4.a and b. The Hurst exponents have been shown in Fig. 1.4.c and d, respectively.

1.3 Outline of the Thesis

In the following chapter of the thesis, the procedure of obtaining invariants of the dynamical systems and the time delay embedding has been given where these invariants are fractal dimensions and Lyapunov exponents. The Lyapunov exponents reflect the asymptotic behaviour of the system and evaluation in one dimension given in [8]. To obtain the full spectrum of Lyapunov exponents has been given in [1,9,10,11-15]. The largest Lyapunov exponent is more important since the system can be said chaotic if it is positive. Another invariant given in chapter two is fractal dimension which is more frequently used in classification of experimental time series. The fractal dimension can be expressed in different forms as correlation dimension, box counting dimension and Lyapunov dimension. The definitions of these dimensions are given in [1,6,8-10,16]. Furthermore, different fractal dimension definitions have been described in [17] by such as Higuchi, Petrosian and Katz algorithms where the fractal dimension takes values between one and two and express the complexity of the signal between these bounds. In order to choose properly the invariants of the system, time delay embedding [5] has to be applied and the necessary parameters, which are time delay and embedding dimension, have to be determined. Methods for determining time delay and the comparisons with different techniques have been given in [1,10,18]. Embedding dimension is a significant parameter since it gives the dimensionality of the given system where an appropriate selection of embedding dimension provides to be determined the number of Lyapunov exponents, correlation dimension and complexity measure. There are several methods to compute the embedding dimension which are given in [1,10] and in [19], some improvements have been added for more adequate selection of embedding dimension.

In practice the experimental time series are corrupted by noise and this causes false estimation of embedding dimension or invariants of dynamical systems. Therefore, before starting to process an observed signal, noise reduction should be performed. Since there is no foreknowledge about the dynamics of the system, blind source separation techniques are applied to reduce the noise where these methods are principal component analysis and singular value decomposition defined in [20,21]. The important

part of these methods is to make a subspace estimation of full rank noisy data matrix where some techniques for subspace estimation have been given in [22-26]. The proposed subspace estimation method is given in chapter three together with the definition and application of blind source separation methods. In [27], ARMA (Autoregressive Moving Average) have been used as a different method.

In the thesis, observed signals have been taken from different locations of brain while eyes were closed and opened condition. In order to comment on the state changes at the specified regions after physical switching, the structure and physiology of human brain have been informed in [28], so that the accuracy of results can be discussed. In chapter four, in addition to brief information about structure of human brain, electrical activity (EEG) of brain has been given in [29-31] and moreover information about standard 10-20 EEG measurement system has been given in [32,33].

In chapter five, wavelet analysis of nonstationary signals is given. Short time Fourier transform (STFT) is the first approach to time frequency localization but due to constant resolution, the wavelet transform has been needed where the necessity for wavelet transform is explained in [34]. Different time-frequency energy distributions have been explained in [35,36]. Information about wavelet decomposition has been given in chapter five. In this chapter there are three basic applications which have been applied to brain signals where the first one is to discriminate the brain signals into frequency bands using wavelet decomposition so that the correlation dimension of alpha band filtered signal would be computed. The wavelet decomposition has been used in [37,38,39,40] to separate brain signals into natural frequency bands. The second application is to find wavelet entropy which is obtained from relative wavelet energies which has been used as classification parameter. The third application is to detect the wavelet ridges where the wavelet ridges are obtained from wavelet transforms evaluated at ridge points. Evaluation of wavelet ridges with different methods has been explained in [41-43,44]. In the application the ridge distribution of EEG signals for eyes opened and closed conditions have been shown.

In chapter six, correlation dimensions have been used as classification parameters. The application results for spontaneous EEG signals have been given and these results have been compared with other studies given in [45-47]. When state transitions occur in the brain without an input, it is expected to see variations at the invariants of the system. These state changes may occur due to pathological conditions, transitions to different states in sleep. In [47], Freeman has modelled the olfactory

system and correlation dimension of olfactory system for different states such as inhalation and exhalation have been computed. In the application of thesis, correlation dimensions have been computed for eyes opened and eyes closed conditions. This classification parameter gives idea about when and how complexity changes in the brain occur. The computed correlation dimension indicates the signal's degree of chaoticness. Furthermore, measurements of macro-activity of the brain have been expressed as the most adequate method which illustrates the dynamic properties of the integrative brain function where the integrative brain function, defined in [48], states that measurements from brain reflects the total response of neuron populations. According to the superbinding principle given in [48], EEG contains the superposition of the natural frequency bands of the brain where these natural frequency bands are delta, theta, alpha, beta and gamma.

CHAPTER 2

THE INVARIANTS OF DYNAMICAL SYSTEMS AND TIME DELAY EMBEDDING

2.1 Invariants of Dynamical System

In order to identify the systems, the invariant quantities which are not sensitive to initial conditions or small perturbations of an orbit, while individual orbits of the system are exponentially sensitive to such perturbations, are to explored. The invariants of the linear system are the resonant frequencies. For non-linear dynamical systems the invariants such as Lyapunov exponents, fractal dimensions (box counting, Lyapunov, correlation) and entropy values can also be used for classification of an orbit. These classifiers are. The Fractal dimensions are characteristics of the geometry of the attractor and Lyapunov exponents show how orbits on the attractor move apart under the evolution of the dynamics.

2.1.1 Lyapunov Exponent in One Dimension

Consider the two initial points x_0 and $x_0 + \varepsilon$ mapped by the function f: $I \to I$, where $I \subset R$. For n iterations of this map the Lyapunov exponent $\lambda(x_0)$ satisfies the equation

$$\varepsilon e^{n\lambda(x_0)} = f^n(x_0 + \varepsilon) - f^n(x_0). \tag{2.1}$$

If ε is divided and taken the limit as $\varepsilon \to 0$,

$$e^{n\lambda(x_0)} = \left| \frac{df^n(x)}{dx} \right|_{x_0} \tag{2.2}$$

when $n \to \infty$, the definition of Lyapunov exponent is as [8],

$$\lambda(x_0)= \lim_{n\to\infty} \frac{1}{n} \ln \left| \frac{df^n(x)}{dx} \right|_{x_0} \quad . \tag{2.3}$$

From Eq. 2.3, it is clear that $\lambda(x_0)$ is the average exponential stretching of initially nearby points

$$f^n(x_0) = f(f(...f(x_0)....)) \tag{2.4}$$

then,

$$\left.\frac{df^n(x)}{dx}\right|_{x_0} = \left.\frac{df(x)}{dx}\right|_{x_{n-1}} \left.\frac{df(x)}{dx}\right|_{x_{n-2}} .....\left.\frac{df(x)}{dx}\right|_{x_0} = f'(x_{n-1})\, f'(x_{n-2})....f'(x_0) \tag{2.5}$$

where prime denotes the derivative and finally, Eq. 2.3 can be written in the form [8]

$$\lambda(x_0) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln | f'(x_i) | \quad . \tag{2.6}$$

Example 2.1: The tent map given by the formula;

$$\Delta_\mu(x) = 2\mu \begin{cases} x, & \text{if } 0 < x \le \frac{1}{2} \\ 1-x & \text{if } \frac{1}{2} < x \le 1 \end{cases}$$

For μ=1, $|\Delta'(x)| = 2$ for all $x \in [0,1]$ and thus $\lambda(x)$=ln2. This is consistent with the result of the formula,

$f^n(x_0 + \varepsilon) - f^n(x_0) = \Delta x = \varepsilon\, 2^n = \varepsilon\, e^{n\lambda}$. Here it is clear that λ=ln2. For arbitrary μ, it is found as λ=ln2μ.

2.1.2 Fractal Dimension

Fractal dimension is the most widely studied invariant quantity for dynamical systems. Periodic systems have an integer dimension, whereas fractional dimension is the indicator of chaos [45].

There is no unique definition but the basic idea behind the fractal dimension is the number of pair-wise correlation in a ball with radius r about points on attracting set scales as r raised to some power [8]. The dimension derived from a general formula of correlation function which is given in Eq. 2.7,

$$\boldsymbol{C}(\boldsymbol{q},\boldsymbol{r}) = \frac{1}{\boldsymbol{N}}\sum_{\boldsymbol{k}=1}^{\boldsymbol{N}}\left[\frac{1}{\boldsymbol{N}}\sum_{\boldsymbol{n}=1}^{\boldsymbol{N}}\Theta(\boldsymbol{r} - |\, \mathrm{y}(\boldsymbol{n}) - \mathrm{y}(\boldsymbol{k})\,|)\right]^{(\boldsymbol{q}-1)} \tag{2.7}$$

where N is the number of data and the expression in bracket is the ratio of the number of data which are inside the ball with radius r and center $\boldsymbol{y}$(k) and Θ is the Heaviside function

$$\Theta(\mathrm{x}) = \begin{cases} 1 & \mathrm{x} > 0 \\ 0 & \mathrm{x} \le 0 \end{cases} \tag{2.8}$$

By using Eq. 2.7, the dimension D_q can be defined by a limit,

$$\mathrm{D_q} = \lim_{\mathrm{r}\to 0}\frac{\log \mathrm{C(q,r)}}{(\mathrm{q}-1)\log(\mathrm{r})} \tag{2.9}$$

The dimension is accepted as invariant measure for a small neighborhood range because if the r is too small the radius will be insufficient to cover the point on the trajectory and if r is very large, a few *d* dimensional spheres will suffice to cover the whole trajectory. Therefore the range of *r* should be properly chosen such that the dimension D_q is almost constant which the estimated correlation is dimension shown in Fig 2.1.b.

2.1.2.1 Correlation Dimension

For q=2, the *correlation function* C(2,*r*) is found and given in the form [9]:

$$\boldsymbol{C}(2,\boldsymbol{r}) = \frac{2}{\boldsymbol{N}(\boldsymbol{N}-1)}\sum_{\boldsymbol{i}=1}^{\boldsymbol{N}}\sum_{\boldsymbol{j}=\boldsymbol{i}+1}^{\boldsymbol{N}}\Theta(\boldsymbol{r} - \|\, \boldsymbol{y}(\mathrm{i}) - \boldsymbol{y}(\mathrm{j})\,\|) \quad . \tag{2.10}$$

The correlation dimension is more efficient at quantifying the fractal nature of an attractor because it emphasizes regions of state space that contain the most data. For an attractor with uniform measure, it is expected that $D_0=D_2$, otherwise $D_0<D_2$ [6]. In Fig. 2.1.a and Fig. 2.1.b the correlation functions and the local slopes of correlation function which is *correlation dimension*, for Lorenz system have been illustrated respectively.

2.1.2.2 Box Counting Dimension

D_0 is called *capacity* or *box counting* dimension which is defined as the minimum number of spheres with radius r needed to cover all the points in the data set. Then the expression for D_0 can be defined as

$$D_0 = \lim_{r\to 0} \frac{\log N(r)}{\log(1/r)} \tag{2.11}$$

where N(r) is the correlation function for q=0, which yields the relation between N(r) and D_0 as $N(r) \approx r^{-D_0}$ [1].

2.1.2.3 Lyapunov Dimension

The Lyapunov exponents which have been explained in section 2.3 are used as invariants of the dynamical system. The Lyapunov dimension is defined as

$$D_L = K + \frac{\sum_{\alpha=1}^{K} \lambda_\alpha}{|\lambda_{K+1}|} \tag{2.12}$$

where K is the number of nonnegative Lyapunov exponent i.e, $\sum_{\alpha=1}^{K} \lambda_\alpha \geq 0$ and $\sum_{\alpha=1}^{K+1} \lambda_\alpha < 0$ [1]. The advantage of finding Lyapunov dimension D_L is that it is computationally much easier to obtain compared to the correlation dimension and box counting dimension only if the Lyapunov exponents are already known. Secondly, the Lyapunov dimension does not account for seldom visited regions as well as other definitions of fractal dimension [8]. But for the experimental time series, in order to find

Lyapunov dimension, all of the exponents should be calculated and it is no more practical than other methods.

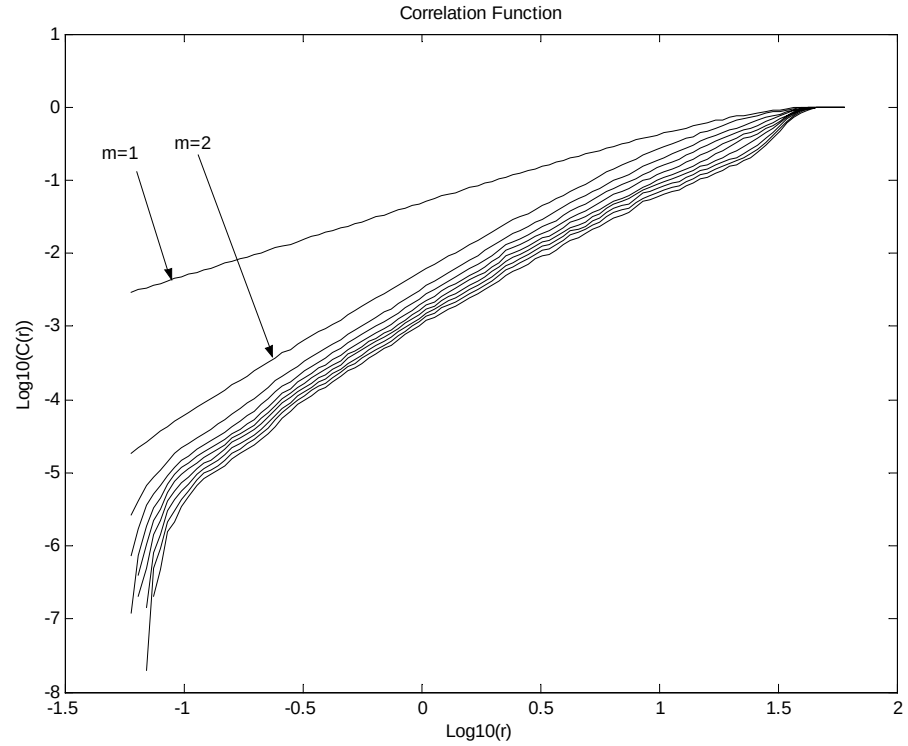

Figure 2.1.a Correlation function for Lorenz system m=1,...10

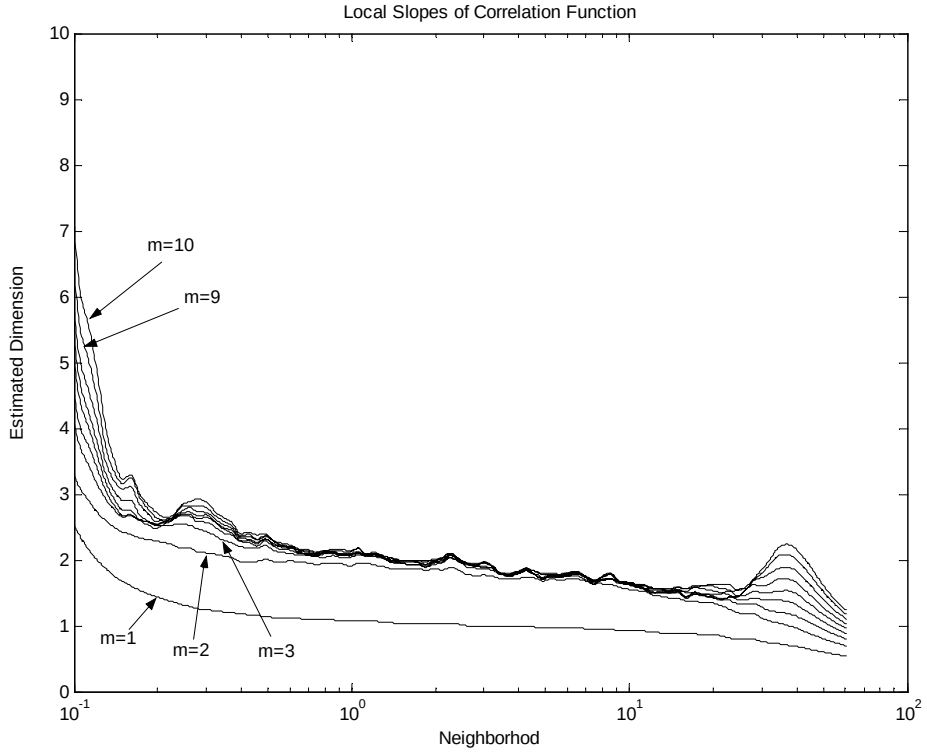

Figure 2.1.b The Correlation Dimension D_2

2.2 Time Delay Embedding Parameters

The necessity of reconstructing phase space has been explained in chapter 1. Two methods are given in [6] and [9] for determining embedding dimension which are *time delay coordinates* and *derivative coordinates* respectively. Derivative coordinates contain higher order derivatives of the observed time series. But since the calculations including derivative are sensitive to noise, it is not very useful for experimental data [16]. Therefore in this thesis *time-delay coordinates* has been used for embedding. The goal of this reconstruction is to transform the scalar measurements into a suitable vector form to find the invariants of the dynamical system. In order to obtain the phase space model, two parameters which are time lag T and embedding dimension d_E have to be determined.

2.2.1 Determining the Time Delay

The first part of the time delay embedding is to select the optimal time intervals T between subsequent samples. Because, if T is chosen too small the samples $s(n)$ and $s(n+T)$ which compose the reconstructed data vector $\boldsymbol{y}(n) \in R^d$, will not be sufficiently independent of each other. In other words, insufficient time will have evolved for the system to have explored of its phase space. If T is chosen very large, the measurements $s(n)$ and $s(n+T)$ will behave as if random numbers were taken from stochastic time series.. Due to these reasons appropriate methods have to be chosen for the optimum time delay T. Although in [18], finding average mutual information is defined as cumbersome, it is reliable and fairly preferable method and explained together with linear autocorrelation function method below.

2.2.1.1 Linear Autocorrelation Function

The first approach might be to choose the time delay value from $s(n)$ using linear autocorrelation function,

$$C(T) = \frac{\frac{1}{N}\sum_{n=1}^{N}\left[(s(n+T) - s_{av})(s(n) - s_{av})\right]}{\frac{1}{N}\sum_{n=1}^{N}\left[s(n) - s_{av}\right]^2} \tag{2.13}$$

where

$$s_{av} = \frac{1}{N}\sum_{n=1}^{N} s(n) \tag{2.14}$$

then the time delay T can be chosen as the smallest time delay where C(T)=0 [10]. In Fig 2.2, first passing through zero is observed at T=540 samples where the sampling rate is 512Hz, i.e, T=1.05 sec, which is quite high for time delay embedding and as will be seen in the sequel, to choose first minimum as the time delay of linear autocorrelation function will completely misleading since linear choice of T is derived from the nonlinear process between $s(n)$ and $s(n+T)$. Hence, average mutual information is more preferable method.

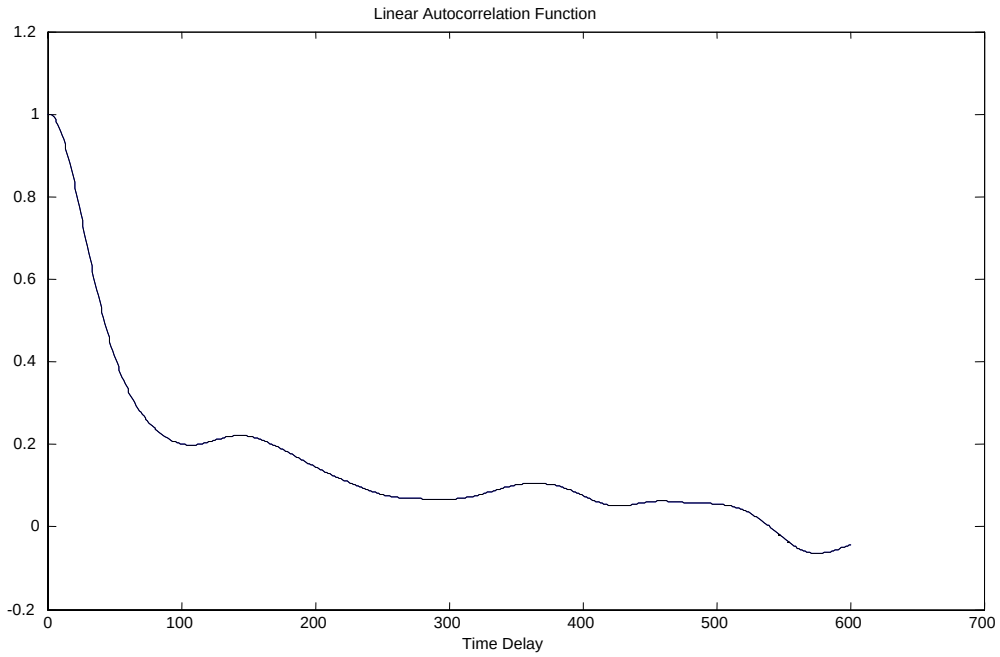

Figure 2.2 Linear autocorrelation function of Lorenz system

2.2.1.2 Average Mutual Information

Shannon's idea of mutual information [49] between two measurements gives the correlation among data points. The connection between two measurements $s(n)$ and $s(n+T)$ express how much these two points are correlated with each other and the criterion in this method is to find the mutual dependence between these two measurements. This relation is expressed as average mutual information which is

$$I(T) = \sum_{s(n),s(n+T)} P(s(n), s(n+T)) \log_2 \left[\frac{P(s(n), s(n+T))}{P(s(n))P(s(n+T))} \right] \tag{2.15}$$

$I(T) \geq 0$. In Eq. 2.15, the expression $P(s(n),s(n+T))$ is the probability of the measurements $s(n)$ and $s(n+T)$ in whole time series. When T becomes large, the chaotic behavior of the signal ($\lambda_{max}>0$) makes the measurements $s(n)$ and $s(n+T)$ to be independent practically and $I(T)$ tends to zero. The first minimum of the $I(T)$ is accepted as the time delay T. In Fig. 2.3, the average mutual information of Lorenz system has been plotted. The system has been sampled at 256 Hz and contains 25600 data points. It is clearly seen that T reaches the first minimum at T=28 samples (54msec).

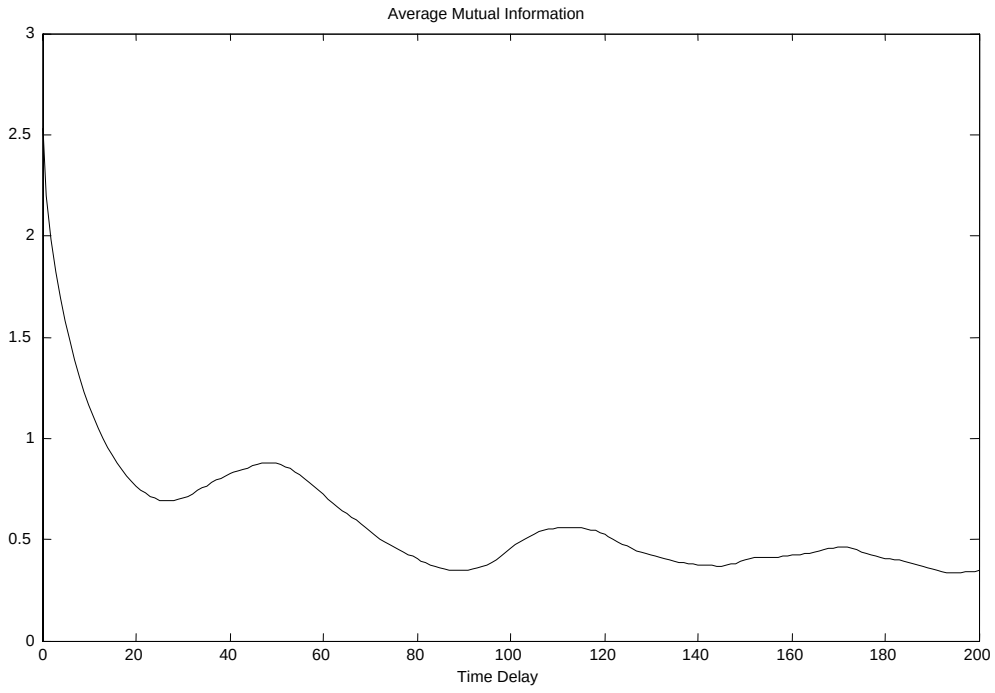

Figure 2.3 Average mutual information for Lorenz system

2.2.2 Determining Embedding Dimension

The dimension, which unfolds the attractor so that there is no remaining overlap in the orbit, is called embedding dimension d_E which is an integer. According to the embedding theorem [5], if the dimension of the attractor is d_A then the attractor will be certainly unfolded at the dimension $d_E > 2d_A$. This condition is not necessary for unfolding but sufficient to ensure that there would be no intersections in the orbit. In the sequel, the methods to determine the embedding dimension have been explained.

2.2.2.1 Eigenvalue Analysis

In this method, the measurement $\boldsymbol{y}(n) \in R^d$ given in Eq. 1.4, is used to obtain the covariance matrix $\mathbf{C} \in R^{d \times d}$. The covariance matrix is obtained as below:

$$\boldsymbol{C} = \frac{1}{\boldsymbol{N}} \sum_{\boldsymbol{n}=1}^{\boldsymbol{N}} [\boldsymbol{y}(\boldsymbol{n}) - \boldsymbol{y}_{\boldsymbol{av}}][\boldsymbol{y}(\boldsymbol{n}) - \boldsymbol{y}_{\boldsymbol{av}}]^T \tag{2.16}$$

where

$$\boldsymbol{y}_{\boldsymbol{av}} = \frac{1}{\boldsymbol{N}} \sum_{\boldsymbol{n}=1}^{\boldsymbol{N}} \boldsymbol{y}(\boldsymbol{n}) \tag{2.17}$$

After eigendecomposition of $\mathbf{C}$, d_E eigenvalues of d is related with the embedding dimension of the system. If the selected dimension d is higher than d_E of the system, then there is d_E-1 nonzero eigenvalues and the redundant eigenvalues are zero. But if there is contamination by noise then the threshold for d_E may be selected in heuristic sense, therefore subspace estimation methods are needed which are the principal component analysis and singular value decomposition and have been explained in chapter 3. The result of this method applied to Lorenz system has been shown in Fig. 2.4.a.

2.2.2.2 Saturation of Correlation Dimension.

Another method is based on finding correlation dimension until large enough values of d. The principle idea of this method is to plot embedding dimension versus correlation dimension D_2. While increasing d, the correlation dimension D_2 approaches

to a constant value where this saturated value is a measure for estimating D_2. The next integer greater than the saturated D_2 value can be accepted as d_E. The disadvantage is that, this method can only be applied to noise free data. Any noise contamination will avoid the saturation of D_2 or cause to be found a false D_2 consequently false embedding dimension d_E. Application of the method for the Lorenz system has been illustrated in Fig 2.4.b in which d_E=3.

2.2.2.3 False Nearest Neighborhood Algorithm (FNN)

The third method for determining d_E obtained by observing intersections of the orbit which is due to projections of high dimensional system to lower dimension. These false crossings cause false neighbors in the orbit. This algorithm plots the percentage of FNN versus each dimension and indicates the necessary dimension to unfold the attractor. The algorithm can be expressed in compact form as following:

i) Start with d=1. For the data point $\boldsymbol{y}(n)$ given in Eq. 1.4, find nearest neighbor $\boldsymbol{y}^{NN}(n)=[s^{NN}(n), s^{NN}(n+T), s^{NN}(n+2T),...,s^{NN}(n+(d-1)T)]$. Repeat this step for all data points,

ii) find Euclidian distance $R_d(k)^2 = \sum_{m=1}^{d}\left[s(k+(m-1)T) - s^{NN}(k+(m-1)T\right]^2$,

iii) nearest neighbors are known but whether which of them are true or false neighbor is not known. To determine it increase d to d+1.

iv) calculate $R_{d+1}(k)^2 = \sum_{m=1}^{d+1}\left[s(k+(m-1)T) - s^{NN}(k+(m-1)T)\right]^2$,

v) find the ratio: $\sqrt{\dfrac{R_{d+1}(k)^2 - R_d(k)^2}{R_d(k)^2}} = \dfrac{|\,s(k+Td) - s^{NN}(k+Td)\,|}{R_d(k)}$,

vi) if this ratio is greater 15 or so $\boldsymbol{y}^{NN}(n)$ is FNN, otherwise it is true nearest neighbor [1]. The dimension is called embedding dimension which drops the percentage of false neighbor to zero.

In [10] and [19], a second criterion is given for false neighbor determination. According to this criterion the nearest neighbor may not be necessarily close. If the nearest neighbor to the given point is false but not close, then the Euclidian distance while passing through d+1 is nearly $2R_A$ where R_A is the mean size of the attractor and expressed as,

$$R_A = \frac{1}{N}\sum_{n=1}^{N}(s(n) - s_{av}) \tag{2.18}$$

where

$$s_{av} = \frac{1}{N}\sum_{n=1}^{N}s(n) \quad . \tag{2.19}$$

Therefore the criterion can be chosen as $\frac{R_{d+1}(n)}{R_A} > 2$.

The advantage of this technique is that it is relatively robust to noise and data corruption. In Fig 2.4.c, the application of the method Lorenz system has been given.

2.2.2.4 True Vector Fields Method

According to this geometrical approach, when a small dimension is used for phase space reconstruction, vector field in any location will not be unique because of the projection from higher dimension. Degree of freedom is determined by $\mathbf{\Delta}(n)$ and $\mathbf{\Delta}^{NN}(n)$ of being parallel where $\mathbf{\Delta}(n) = \mathbf{y}(n+1) - \mathbf{y}(n)$ and, $\mathbf{\Delta}^{NN}(n) = \mathbf{y}^{NN}(n+1) - \mathbf{y}^{NN}(n)$. The percentage variation of parallel vectors versus for each dimension is a measure for degrees of freedom. The steps of algorithm are as follows:

i) find $\frac{\mathbf{\Delta}(n)}{\|\mathbf{\Delta}(n)\|}$ for the given point in phase space,

ii) find $\frac{\mathbf{\Delta}^{NN}(n)}{\|\mathbf{\Delta}^{NN}(n)\|}$,

iii) if the angle between two vectors is too small they can be accepted as parallel.

iv) repeat step i, ii and iii for each point in phase space and derive a percentage. Accept the dimension *d* as necessary dimension which drops this percentage to zero.

This method is similar to the false nearest neighborhood technique. Both of them are geometric and both of them analyze intersections of the orbits. In Fig 2.4.d the results of this algorithm for Lorenz system is plotted which indicates lower performance.

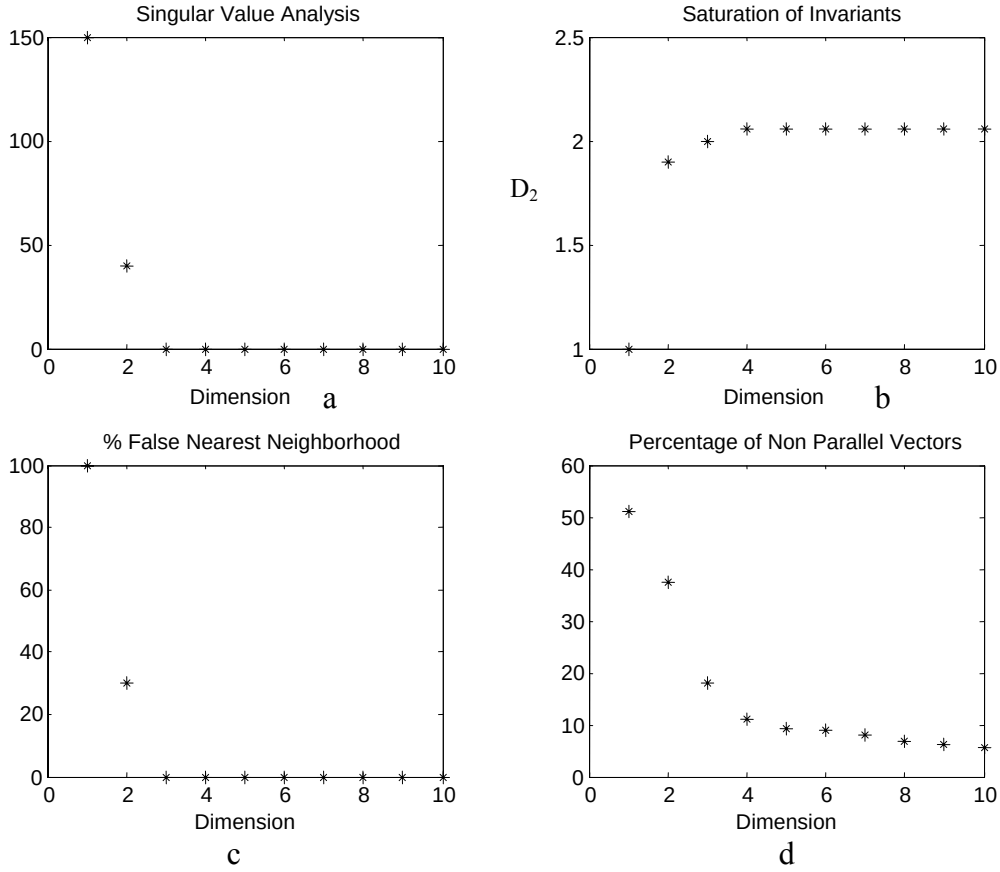

Figure 2.4 The results of embedding dimension methods.

2.3 Lyapunov Exponents from observed time series

The number of Lyapunov exponents is related with the degrees of freedom or equivalently dimensionality of the dynamical system. The most important Lyapunov exponent is maximum Lyapunov exponent λ_{max}, which indicates the chaos if $\lambda_{max}>0$.

The Lyapunov exponents carry the units of the inverse time and give typical time scale for the divergence or convergence of nearby trajectories. If the dynamics of a system are expressed as ordinary differential equations, then it is called *flow data* and while calculating the iterative map correspondence of Lyapunov exponents calculated from flow data, the average time interval τ_{av} should be calculated from Poincare section

of the phase space. The time between two sequential transitions of trajectory in phase space through the Poincare section gives average time interval τ_{av}. Then it can be shown that, Lyapunov exponents computed for the iterative map are related to those of the flow data by $\lambda_{map}=\lambda_{flow}\tau_{av}$ [9].

Table 2.1 Characteristic of attractors for a four dimensional flow.

λ_1	λ_2	λ_3	λ_4	Attractor	Dimension
_	_	_	_	Equilibrium Point	0
0	_	_	_	Limit cycle	1
0	0	_	_	2-Torus	2
0	0	0	_	3-Torus	3
+	0	_	_	Strange (Chaotic)	>2
+	+	0	_	Strange (Hyper chaotic)	>3

When the further analysis of system invariants is required, full spectrum of Lyapunov exponents has to be determined. For experimental time series first step is to put scalar time series into vector form using time delay embedding. Then the stability of an observed point $\boldsymbol{y}(n)$ of the dynamical system $\boldsymbol{y}(k)\rightarrow\mathbf{F}(\boldsymbol{y}(k)) = \boldsymbol{y}(k+1)$ to small perturbations $\Delta(k)$, is expressed by looking at the linearised dynamics and by Taylor series expansion of $\mathbf{F}(\boldsymbol{y}(k)+\Delta(k))$ at $\boldsymbol{y}(k)$ the following is obtained

$$\boldsymbol{y}(k+1) + \Delta(k+1)=\mathbf{F}(\boldsymbol{y}(k) + \Delta(k)) \approx \mathbf{DF}(\boldsymbol{y}(k)) \cdot \Delta(k) + \mathbf{F}(\boldsymbol{y}(k))) \tag{2.20}$$

Hence it yields

$$\Delta(k+1) = \mathbf{DF}(\boldsymbol{y}(k)) \cdot \Delta(k), \tag{2.21}$$

The Jacobian matrix DF is given as

$$\mathrm{DF}(y)_{ij}=\frac{\partial \boldsymbol{F_i}(y)}{\partial y_j} \tag{2.22}$$

while moving from time k to k+L, it can be found that,

$$\Delta(k+L)=\mathbf{DF}(\boldsymbol{y}(k+L-1)) \cdot \mathbf{DF}(\boldsymbol{y}(k+L-2)) \ldots \cdot \mathbf{DF}(\boldsymbol{y}(k)) \cdot \Delta(k) \equiv \mathbf{DF}^{L}(\boldsymbol{y}(k)) \cdot \Delta(k) \quad (2.23)$$

$$\mathbf{DF}^{L}(\boldsymbol{y}(k)) = \prod_{i=1}^{L} \boldsymbol{DF}(\boldsymbol{y}(k+L-i)) \quad (2.24)$$

where $\mathbf{DF}^{L}(\boldsymbol{y}(k))$ is the composition of L Jacobians. If the eigenvalues of $\mathbf{DF}^{L}(\boldsymbol{y}(k))$ behave as $\exp(L\lambda)$ with $\lambda>0$, then the orbit $\boldsymbol{y}(k)$ along which the Jacobians are evaluated, diverges to infinity [1]. When $\boldsymbol{y}(k)$ is a time dependent chaotic orbit then,

$$\|\Delta(k+L)\|^{2}=\Delta^{T}(k) \cdot [\mathbf{DF}^{L}(\boldsymbol{y}(k))]^{T} \cdot \mathbf{DF}^{L}(\boldsymbol{y}(k)) \cdot \Delta(k) \quad (2.25)$$

in the equation 2.25, the key quantity is,

$$[\mathbf{DF}^{L}(\boldsymbol{y}(k))]^{T} \cdot \mathbf{DF}^{L}(\boldsymbol{y}(k))$$

and the Oseledec matrix is given in the form,

$$\mathbf{OSL}(x,L)=([\mathbf{DF}^{L}(\boldsymbol{y}(k))]^{T} \cdot \mathbf{DF}^{L}(\boldsymbol{y}(k)))^{1/2L} \quad (2.26)$$

Then Multiplicative Ergodic Theorem states that the limit of **OSL** matrix exists, i.e. $\lim_{L\to\infty} \mathbf{OSL}(\boldsymbol{x}, \boldsymbol{L}) < \infty$ and independent of almost all points in the basin of the attraction of the attractor.

The logarithm of the eigenvalues of this orthogonal matrix OSL are in non-increasing order and are called Lyapunov exponents of the dynamical system $\mathbf{x}\rightarrow\mathbf{F}(\mathbf{x})$. If one or more of the $\lambda_a>0$, then chaos is observed. The sum of the Lyapunov exponents $\lambda_1+\lambda_2+\ldots+\lambda_d<0$ by the dissipative nature of the systems [1].

The determination of these eigenvalues is given by QR decomposition method when the solution is planned to reached iteratively. In QR decomposition method of the Oseledec matrix of the power 2L is considered as a product of factors which can be represented as $\mathbf{A}(2L) \cdot \mathbf{A}(2L-1),\ldots,\mathbf{A}(1)$ where $\mathbf{A}(j)$ is the component matrix. The component matrix can be written as product of an orthogonal matrix $\mathbf{Q}(j)$ where $\mathbf{Q}(0)=\mathbf{I}$ and upper right triangular matrix $\mathbf{R}(j)$. This can be thought as a polar decomposition of complex number where the matrix **Q** is like the phase factor and the matrix **R** is the magnitude of the complex number. The expression of a matrix in QR form is given as,

$$\mathbf{A}(j) \cdot \mathbf{Q}(j-1) = \mathbf{Q}(j) \cdot \mathbf{R}(j) \quad (2.27)$$

then if iteration starts with j=1,

$$\mathbf{A}(1) = \mathbf{Q}(1) \cdot \mathbf{R}(1),$$
$$\mathbf{A}(2) \cdot \mathbf{Q}(1) = \mathbf{Q}(2) \cdot \mathbf{R}(2),$$
$$\mathbf{A}(2) = \mathbf{Q}(2) \cdot \mathbf{R}(1) \cdot \mathbf{Q}(1)^{T},$$
$$\mathbf{A}(2) \cdot \mathbf{A}(1) = \mathbf{Q}(2) \cdot \mathbf{R}(2) \cdot \mathbf{R}(1)$$
$$\mathbf{A}(3) \cdot \mathbf{A}(2) \cdot \mathbf{A}(1) = \mathbf{Q}(2) \cdot \mathbf{R}(3) \cdot \mathbf{R}(2) \cdot \mathbf{R}(1)$$

and the general statement,

$$\mathbf{A}(2L) \cdot \mathbf{A}(2L-1),...\mathbf{A}(1) = \mathbf{Q}(2L) \cdot \mathbf{R}(2L) \cdot \mathbf{R}(2L-1) ... \cdot \mathbf{R}(1) \qquad (2.28)$$

The Lyapunov exponents are obtained from the product of upper right triangular matrices whose diagonal elements are the Lyapunov exponents in decreasing order

$$\lambda_i = \lim_{L\to\infty} \frac{1}{2L} \sum_{k=1}^{2L} \log(\boldsymbol{R}_i(\boldsymbol{k})) \quad . \qquad (2.29)$$

The dimension for choosing time delay embedding gives the number of Lyapunov exponents. A good choice for number of Lyapunov exponents is given in [50] as making the *d* the next integer greater than or equal to the fractal dimension d_A. This choice will always give at least one negative exponent for a chaotic system [50]. The selected dimension may be higher but in that case there will be at least one spurious Lyapunov exponent. When the order of used polynomial linear map is increased, the spurious exponent is easily distinguishable from true exponents. A second method is to add noise to the system. In that case there will be rapidly drop at spurious exponent but this is not guaranteed that spurious exponent does not exist when there is no rapid drop [50].

CHAPTER 3

NOISE REDUCTION FROM OBSERVED TIME SERIES

Noise limits the ability of extracting the quantitative information from time varying signals. As a characteristic feature, chaotic systems are quite sensitive to the presence of small amount of noise. The simple numerical experiments show that when the white noise is added at even %1 level of the time series, then it is impossible to measure the correlation dimension of the attractor using neighbourhood radius less than 3% of the attractor size [24] defined in Eq. 2.18. Since the chaotic signals have broadband power spectrum, it is difficult to distinguish the white noise from broadband deterministic signals. The main problem is to filter the noise from observed time series under the condition that one has no foreknowledge about the dynamics of the system. This case of noise reduction is called blind source separation. The principal component analysis and singular value decomposition are blind source separation techniques whose basics rely on eigenvalue decomposition and the oriented energy distribution. In this thesis, the method of singular value decomposition has been used to reduce the noise contamination.

3.1 Oriented Energy and Oriented Signal to Signal Ratio Concepts

In a wide variety of systems and signal processing applications, vector sequences are measured and computed [20]. For the analysis of data sequences, scalar data are put into matrix form by using time delay method [5], which is called phase space reconstruction. In this section the purpose is addressed as finding the oriented energy distribution of the matrix $\boldsymbol{M} \in R^{mxn}$ where m is the data length and n is the number of measurement channels or embedding dimension. As an analyzing method, singular value decomposition of matrices formed from observed data has been used to improve the signal parameter estimation and system identification. This method is similar to principal component analysis in which the first few components will account for most of the variation in the original data so that noisy dimension of the data can be reduced. The oriented signal to signal ratio is related to subspace discrimination and that can be used for linear modelling problems if the complexity of the model is the rank of

a certain matrix. Hence, the decision for the complexity reduces to a lower rank of the certain "noise free" matrix [20].

Definition 3.1: Energy of a vector sequence

Consider a sequence of m vectors $\{\boldsymbol{a}_k\}$, $\boldsymbol{a}_k \in R^m$, k=1,…,n and associated with $\boldsymbol{A} \in R^{mxn}$. Its total energy E[$\boldsymbol{A}$] is defined via the Frobenius norm of $\boldsymbol{A}$:

$$E(\boldsymbol{A}) = \| \boldsymbol{A} \|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \quad . \tag{3.1}$$

Definition 3.2: Oriented energy

Let $\boldsymbol{A}$ be a mxn matrix and denotes its n column vectors as $\boldsymbol{a}_k$, k=1,…,n (n>>m), or the indexed vector set $\{\boldsymbol{a}_k\}$ of m vectors $\boldsymbol{a}_k \in R^m$ and for any unit vector $\mathbf{q} \in R^m$ the energy E_q, measured in the direction **q**, is defined as:

$$E_q(A) = \sum_{k=1}^{n} (q^t \cdot a_k)^2 \quad . \tag{3.2}$$

The energy E_q measured in a subspace $Q \subset R^m$, is defined as:

$$\boldsymbol{E}_{\boldsymbol{Q}}(\boldsymbol{A}) = \sum_{\boldsymbol{k}=1}^{\boldsymbol{n}} \| \boldsymbol{P}_{\boldsymbol{Q}}(\boldsymbol{a}_{\boldsymbol{k}}) \|_2^2 \quad . \tag{3.3}$$

where $P_Q(\boldsymbol{a}_k)$ denotes the orthogonal projection of $\boldsymbol{a}_k$ into subspace Q and $\| \cdot \|_2$ denotes the Euclidean norm.

In other words, the oriented energy of a vector sequence $\{\boldsymbol{a}_k\}$, measured in the direction **q** or subspace Q is nothing else than the energy of the signal, orthogonally projected on the vector **q**.

3.2 Blind Source Separation Techniques

3.2.1 Principal Component Analysis (PCA)

The aim of principle component analysis is to reduce the dimensionality of the observed data. Noise reduction by this method corresponds to obtain the subspace related to noise free signal which is realized by finding the eigenvalues of scattering matrix where it is the sum of covariance matrices of each data vector $\mathbf{x} \in R^d$. The eigenvector, which corresponds to highest eigenvalue, gives the direction where most of the energy of the signal is distributed. The algorithm can be given as follow [21],

i) find mean $\mathbf{m} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}_n$ where N is data length,

ii) compute Scatter matrix $\mathbf{S} \in R^{dxd}$ $\mathbf{S} = \sum_{n=1}^{N}(\mathbf{x}_n - \mathbf{m})(\mathbf{x}_n - \mathbf{m})^T$,

iii) find eigenvalues and eigenvectors of matrix **S**, $\mathbf{Se}=\lambda\mathbf{e}$,

iv) obtain new time series by $y_n = \mathbf{e}_1^T \cdot(\mathbf{x}_n-\mathbf{m})$ where n=1,…,N and $\mathbf{e}_1$ is the eigenvector corresponding to the largest eigenvalue. Y corresponds to noise filtered time series.

3.2.2 Singular Value Decomposition (SVD)

The SVD of real matrices is based on the Autonne-Eckart-Young theorem [20], which states that, for any real $\boldsymbol{A} \in R^{mxn}$, there exists a real factorization:

$$\boldsymbol{A} = \boldsymbol{U} \cdot \boldsymbol{S} \cdot \boldsymbol{V}^{\boldsymbol{t}} \tag{3.4}$$

in which the matrices $\boldsymbol{U} \in R^{mxm}$ and $\boldsymbol{V} \in R^{nxn}$ are orthogonal matrices, and the matrix $\boldsymbol{\Sigma} \in R^{mxn}$ is real pseudo diagonal with non-negative diagonal elements.

The diagonal entries σ_i of $\boldsymbol{S}$ are called the singular values and the set of singular values are called singular spectrum of matrix $\boldsymbol{A}$. The columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ are called left and right singular vectors of the matrix $\boldsymbol{A}$, respectively. The space $S^r_U=\text{span}\{u_1,\ldots u_r\}$ is called the r-th left principle subspace. The columns of $\boldsymbol{U}$ and the columns of $\boldsymbol{V}$ are

eigenvectors of $\boldsymbol{AA^t}$ and $\boldsymbol{A^tA}$ respectively [20]. The number of singular values, which are different from zero, equals to the rank of the matrix $\boldsymbol{A}$.

Frobenius norm of matrix $\boldsymbol{A}$ of rank r is defined as:

$$\| \boldsymbol{A} \|_F^2 = \sum_{k=1}^{r} \sigma_k^2 \tag{3.5}$$

where each σ_k is the singular values of $\boldsymbol{A}$. In other words, the total energy in a vector sequence with the associated matrix $\boldsymbol{A}$ is equal to the energy in the singular spectrum [20].

3.3 SVD Based Subspace Estimation Method

The SVD based noise reduction explained in the thesis relies on three basic steps:

i) the observed data y(n) are put into matrix form $\boldsymbol{M}$ as in Eq 3.6,

ii) the dominant singular values of $\boldsymbol{M}$ ($\boldsymbol{M}= \boldsymbol{U \Sigma V^t}$) are estimated,

iii) exact values of dominant singular values are computed,

The scalar observations have been put into the matrix form by the method given in [24]. When the time series of y(n) is partitioned by a rectangular nonoverlapping window function w(n), the data matrix is obtained as:

$$\boldsymbol{M} = \begin{pmatrix} \mathrm{y}(1) & \mathrm{y}(1+T) & \mathrm{y}(1+2T) & \ldots & \mathrm{y}(1+(k-1)T) \\ \mathrm{y}(2) & \mathrm{y}(2+T) & \mathrm{y}(2+2T) & \ldots & \mathrm{y}(2+(k-1)T) \\ \ldots. & \ldots. & \ldots. & & \ldots. \\ \mathrm{y}(p) & \mathrm{y}(p+T) & \mathrm{y}(p+2T) & \ldots & \mathrm{y}(p+(k-1)T) \end{pmatrix} \tag{3.6}$$

where p is the number of reconstruction vectors. The window $w(n)$ has length $T_W = (k-1)T + p$. [23].

There are some critical points that justify the success of the SVD [22]:

i) the information on the underlying model is contained in certain subspaces of the data matrix which can be determined from SVD.

ii) the complexity of the model is given by the approximate rank of data matrix which can be estimated from singular values.
iii) in most applications data are corrupted with additive noise, it is expected that the SVD has a certain noise filtering effect.

3.3.1 Assumptions for Subspace Estimation

Let $\boldsymbol{M}$ be a matrix defined in section 3.3, where $\boldsymbol{M} \in R^{pxq}$ and p>>q. The basic assumption is that the observed data are generated by two unknown matrices; exact data matrix $\boldsymbol{E}$ and additive perturbations $\boldsymbol{N}$ so that

$$\boldsymbol{M}=\boldsymbol{E}+\boldsymbol{N} \tag{3.7}$$

where the matrix $\boldsymbol{N}$ contains the noise, rank ($\boldsymbol{N}$)=rank($\boldsymbol{M}$)=q. The matrix $\boldsymbol{E}$, which contains noise free data is rank deficient (i.e, rank($\mathbf{E}$)=r<q).
Let $\mathbf{V}_m$ be an orthogonal matrix (i.e, $\mathbf{V}^t_m \mathbf{V}_m=\mathbf{I}_q= \mathbf{V}_m\mathbf{V}^t_m$) where $\boldsymbol{V}_m \in R^{pxq}$ and it is partitioned as

$$\boldsymbol{V}_m= [\boldsymbol{V}_{m1} \quad \boldsymbol{V}_{m2}].$$

($\boldsymbol{V}_{m1} \in R^{pxr}$, $\boldsymbol{V}_{m2} \in R^{px(q-r)}$) such that $\mathbf{R}(\boldsymbol{E}^t)=\mathbf{R}(\boldsymbol{V}_{m1})$ and the column space of $\boldsymbol{V}_{e2}$ generate an orthogonal basis for the null space of $\boldsymbol{E}$:

$$\boldsymbol{E}\boldsymbol{V}_{m2}=0 \tag{3.8}$$

For the extraction of $\boldsymbol{E}$ from $\boldsymbol{M}$ it is assumed that the rank of $\boldsymbol{E}$ and its SVD are known and the following assumptions must be satisfied to use SVD as the noise reduction method [22],

1) the exact data should be orthogonal to the noise in the sense that $\boldsymbol{E}^t\boldsymbol{N}=0$,
2) the matrices $\boldsymbol{V}_{m1}$ and $\boldsymbol{V}_{m2}$ must be orthogonal to each other and the matrix $\boldsymbol{N}^t\boldsymbol{N}$ must be a scalar multiple of identity ($\boldsymbol{N}^t\boldsymbol{N}=\sigma^2_{noise}\boldsymbol{I}_q$),
3) the smallest singular value of $\boldsymbol{S}_1$ where $\boldsymbol{S}_1$ is the submatrix of $\boldsymbol{S}$ including to noise free singular values must be larger than the largest singular value of other submatrix $\boldsymbol{S}_2$. The σ_r / σ_{r+1} ratio act as a signal to signal ratio.

3.3.2 Estimation of Clean Data Matrix from Noisy Data Matrix

If the three assumptions given in section 3.3.1 are satisfied, then the exact clean data matrix ***E*** can be estimated from ***M*** according to the relation between singular values of ***E*** and ***M***. There is a gap in the singular spectrum where the smallest value related to signal is larger than the singular value σ_{noise} which is equal to the q-r singular values which illustrate the noise threshold for ***M***. Because of the orthogonality, the relation between exact and observed singular values has been defined in Eq. 3.10,

$$M = (U_{m1} \quad \mathrm{U}_{m2}) \begin{pmatrix} S_{m1} & 0 \\ 0 & \mathrm{S}_{m2} \end{pmatrix} \begin{pmatrix} \mathrm{V}^{t}_{m1} \\ V^{t}_{m2} \end{pmatrix} \quad , \tag{3.9}$$

$$\boldsymbol{S}_{e1} = \sqrt{\boldsymbol{S}^{2}_{m1} - \sigma^{2}_{noise} \boldsymbol{I}_{r}} \quad \text{and} \quad \boldsymbol{S}_{m2} = \sigma_{noise} \boldsymbol{I}_{q-r} \tag{3.10}$$

where $\boldsymbol{S}_{e1}$ is the submatrix related with noise free signal and $\boldsymbol{S}_{m2}$ is the submatrix with related to the noise. If σ_r is the smallest singular value related to $\boldsymbol{S}_{e1}$, then the SNR is defined as:

$$SNR = \left[20 \log \frac{\sigma_{r}}{\sigma_{noise}} \right] \quad . \tag{3.11}$$

3.3.3 Rank Estimation of Noise Free Matrix

In the thesis, the estimation of the rank of the subspace has been obtained by using the proposed method given in [51] and the singular entropy definition given in [52]. The deficient rank has been estimated by using the ratio of the singular entropy increment values of noisy matrix ***M***. Singular entropy is defined in [52] as,

$$\mathrm{E}_{k} = \sum_{i=1}^{k} \Delta \mathrm{E}_{i}, \quad k \leq n \tag{3.12}$$

where E_k is the singular entropy for degree of k and ΔE_i is the increment of singular entropy at orders i,

$$\Delta E_i = -p_i \log[p_i] \tag{3.13}$$

where

$$p_i = \frac{\mu_i}{\sum_{j=1}^{n} \mu_j} \tag{3.14}$$

where μ_i is the singular value of $\boldsymbol{M}$ for i=1,…,n.

3.3.4 Threshold Selection Criteria

For noisy data matrix where $\boldsymbol{M} \in R^{mxn}$, m>>n and $\Delta E_i = \delta_i$, i=1,…,r,r+1…,n

$$\eta(i) \stackrel{\Delta}{=} \delta_i / \delta_{i+1} \tag{3.15}$$

$\eta(i)$ is the ratio index and i=1…n-1. According to this method, if the estimated rank is "r" it is indicated as:

$$\eta(r) = \max_i (\eta(i)) \tag{3.16}$$

3.3.5 Estimates of Noise Variance and Exact Singular Values

Since the rank of matrix $\boldsymbol{E}$ and the noise threshold have been estimated then the noise variance and exact singular values can also be determined. Let μ_i be the singular value of $\boldsymbol{M}$. Then the estimate σ^2_{est} of variance σ^2_{noise} of the elements can be obtained by,

$$\sigma^2_{est} = \frac{\mu^2_{r+1} + \ldots + \mu^2_n}{n - r}. \tag{3.17}$$

The estimate of exact singular values of $\boldsymbol{E}$ can be obtained by the elements of $\boldsymbol{S}_{e1}$,

$$\sigma_i = \sqrt{\mu_i^2 - \sigma_{est}^2} \quad i=1,...,r \quad . \tag{3.18}$$

By defining the relation between $\boldsymbol{U}_{e1}$ where columns of $\boldsymbol{U}_{e1}$ are the left singular vectors of $\boldsymbol{E}$ as given in [22], and $\boldsymbol{U}_{m1}$, the canonical angles are obtained by the following expression,

$$\boldsymbol{U}_{e1}^{t}\boldsymbol{U}_{m1} = \boldsymbol{S}_{e1}\boldsymbol{S}_{m1}^{-1} = \boldsymbol{C} \tag{3.19}$$

where $\boldsymbol{C} \in R^{rxr}$ diagonal matrix whose elements are the canonical angles. The canonical angles are used for obtaining estimated clean data matrix $\boldsymbol{E}_{est}$. Since range spaces are the same $R(\boldsymbol{E}^t) = R(\boldsymbol{V}_{m1})$, the estimated noise free data matrix $\boldsymbol{E}_{est}$ can be expressed as,

$$\boldsymbol{E}_{est} = \boldsymbol{U}_{m1}(\boldsymbol{S}_{e1}\ \boldsymbol{C})\boldsymbol{V}_{m1}^{t}. \tag{3.20}$$

The estimated noise filtered time series $y_{est}(n)$ is smoothed in order to reduce the remaining small ripples by the given formula,

$$\hat{y}_{est}(n) = \frac{1}{2m+1}\sum_{j=-m}^{m} y_{est}(n+j) \quad . \tag{3.21}$$

where m is the number of neighbour points of the point $y_{est}(n)$.

3.4 Application

The time series from Lorenz system has been used for simulations which is defined by differential equations,

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \tag{3.22}$$

where σ=16,b=4 and r=45,92 and the system has 6000 data point length, 256 Hz sampling rate and 15.67dB SNR. Fig 3.1 clearly indicates the singular spectrum in which the singular values related with cleaned data and noise have been illustrated for the chosen time window.

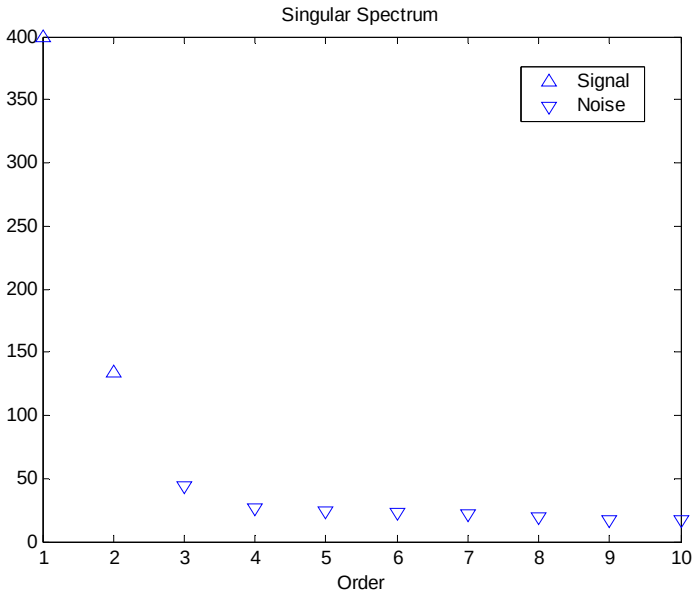

Figure 3.1 Singular spectrum for Lorenz system

In Fig 3.1, first two singular values have been detected as related with signal and the rest are related with noise. The time domain representation of signal has been obtained in Fig 3.2. The correlation dimension results have been used as the performance criterion, and as shown in Fig 3.3 the filtered signal has constant local slope for small range of neighbourhood which corresponds the estimated dimension 2.24 whereas the original dimension is 2.06. The method approximates the exact value which can never be obtained since it is possible to construct the space E. The phase space reconstruction results indicate how much noise has been filtered which are as illustrated in Fig 3.4.

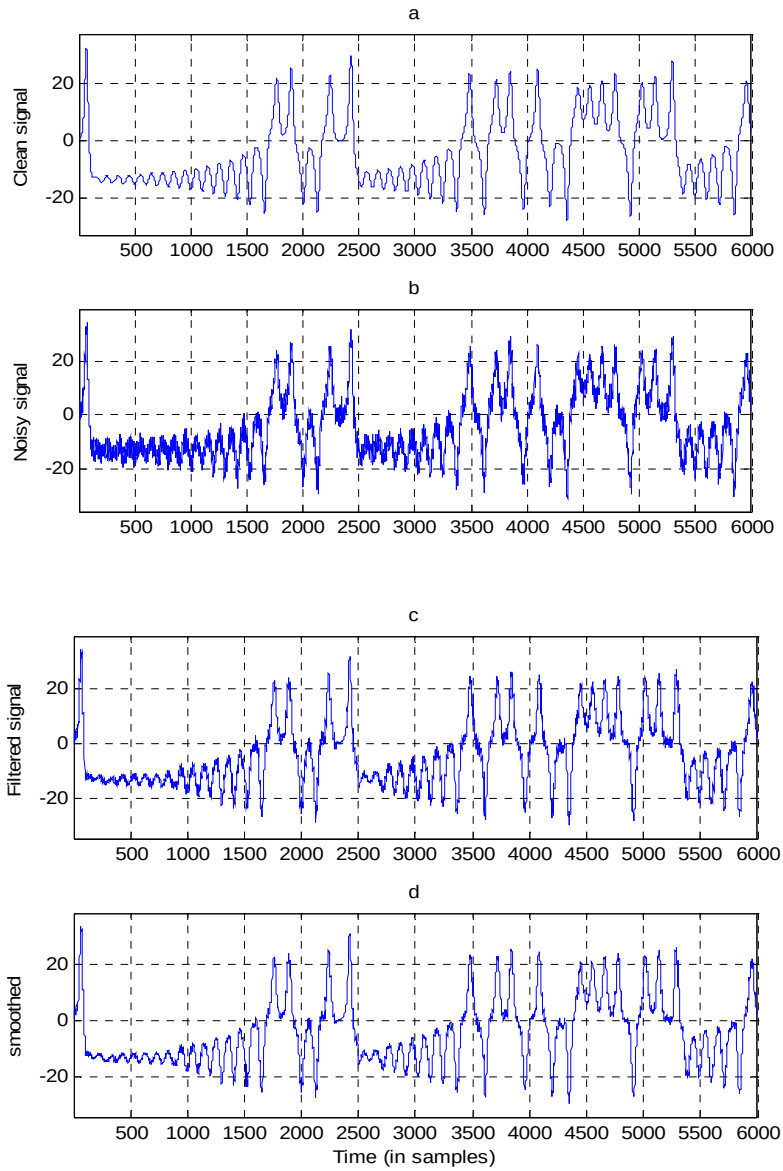

Figure 3.2 Time domain illustration of Lorenz signal. a) Original, b) Noisy, c)Noise Filtered, d)Smoothed signal after filtering.

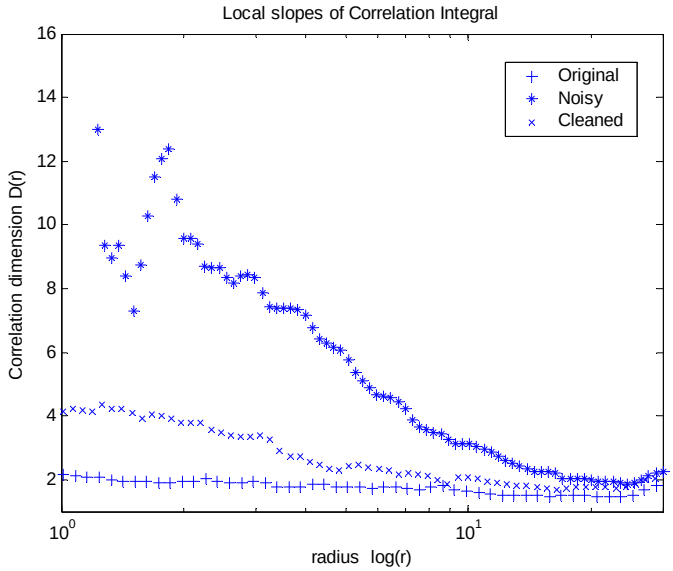

Figure 3.3 Local slopes of correlation function versus neighbourhood radius r.

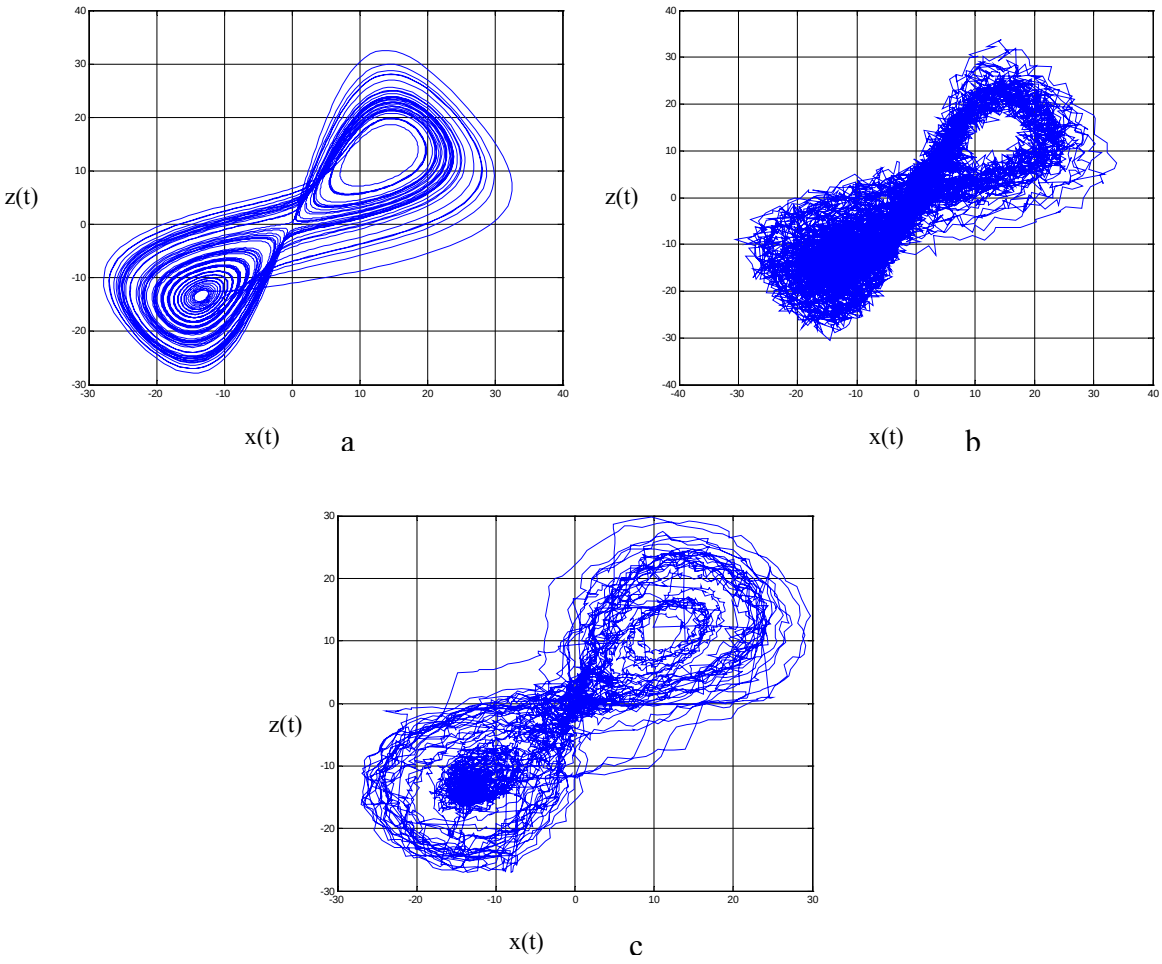

Figure 3.4 Phase spaces for a) Clean, b) Noisy, c) Filtered signal

Another critical parameter in SVD based noise reduction is to choose the order of embedding. The defined SNR [22] of the system has been expressed in Eq. 3.11 and when it is computed for each embedding dimension and then the SNR value decreases while order of embedding increases which limits to choose higher order values. Fig. 3.5 indicates the variation of SNR estimation by embedding order.

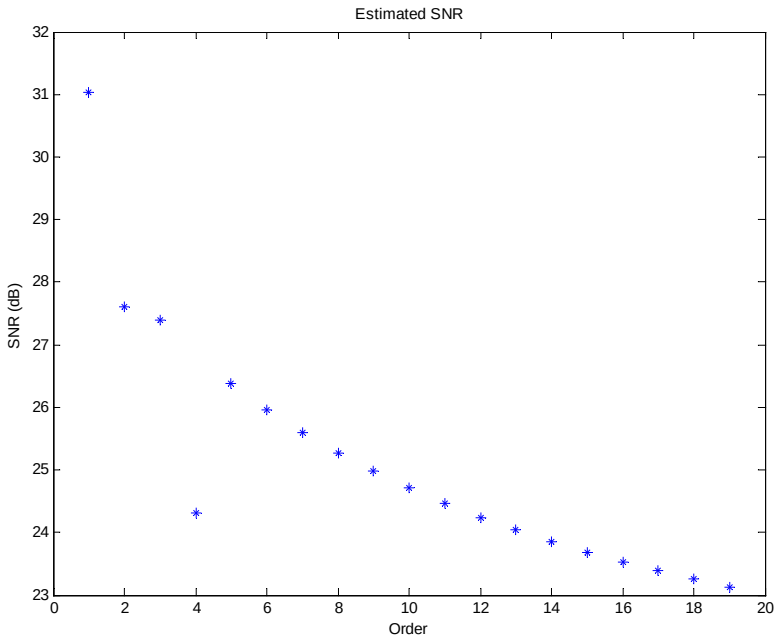

Figure 3.5 Estimated SNR variation versus embedding order.

In Fig. 3.6, blue signs indicate clean; red indicates noisy signals from Poincare section and black signs noise filtered section. For order 12, reconstructed trajectory has diverged from original trajectory more then order of 5. To choose high embedding orders cause the reconstructed phase space to diverge from the original space. The reconstruction performance can be visualized by Poincare sections shown in Fig. 3.6 and the Poincare sections of reconstructed signal have been shown by orders of 5 and 12, respectively. According to the embedding theorem, it is known that embedding dimension should be $d_E > 2d_A$. This can be considered as the lower limit for choosing the

embedding order. If these limitations are taken into account, one can suggest of choosing embedding order as $2d_A$ of the system.

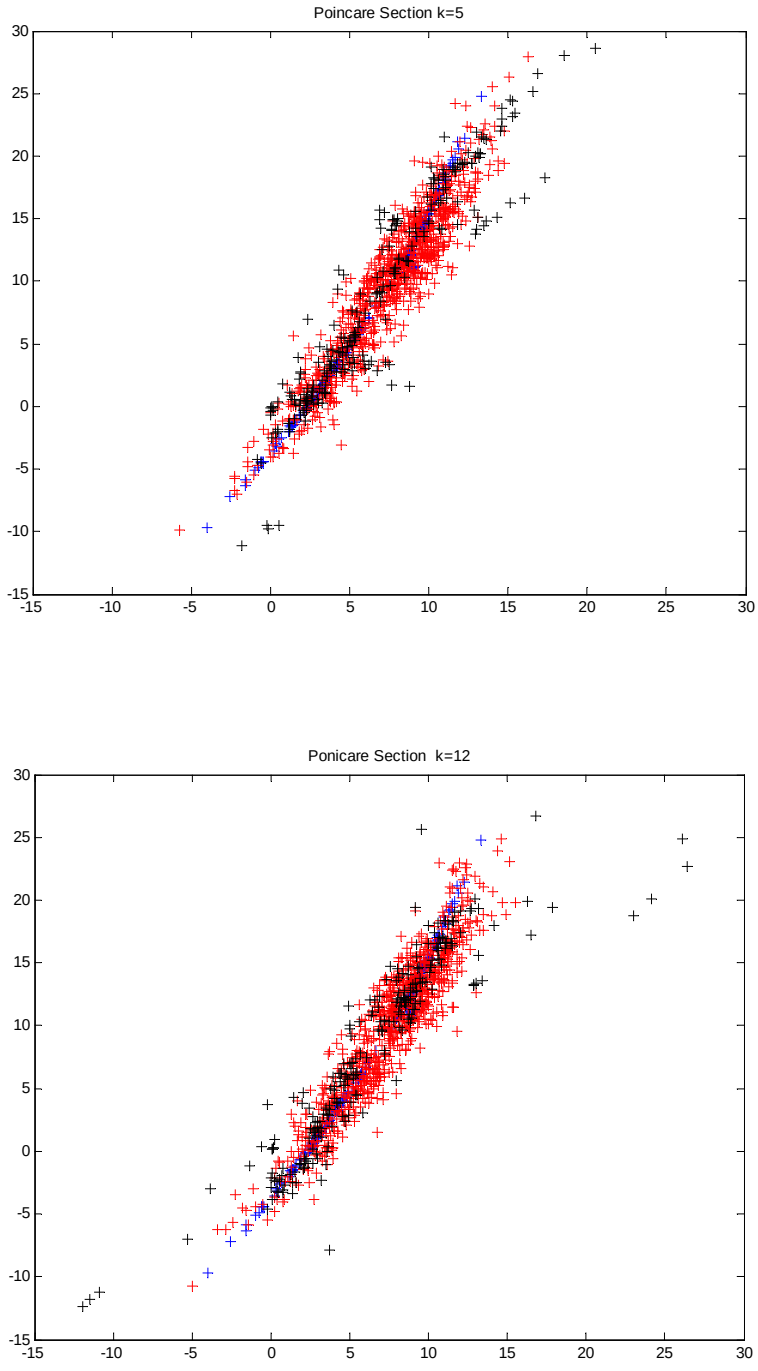

Figure 3.6 Poincare sections for embedding dimensions of 5 and 12 respectively.

CHAPTER 4

PHYSIOLOGICAL INFORMATION IN THE BRAIN

Since the observed signals taken from the brain has been used in latter chapters, it is necessary to give brief information about some part of the brain structure, electrical activity and the natural frequency bands of the brain with their functionality.

Brain is the most critical organ in human body that is responsible for whole mental and physical activities during life and it produces electrical oscillations while performing any task. The electrical activity of the brain is observed by electrical signals taken from different spatial locations of the brain which compose the Cerebral Cortex and are called Cerebral cortex lobes. Hence, Cerebral Cortex and its lobes are point of view in order to have idea about meaning of the electrical activity changes.

4.1 Structures of Cerebral Cortex and Functions of Cerebral Cortex Lobes

The Cerebral Cortex is responsible by critical functions which are sorted as determining intelligence, personality, interpretation of sensory impulses, motor functions, planning and organization and touch sensation. The cerebral cortex is the outer portion (1.5mm to 5mm) of the cerebrum which is located in the anterior portion of the forebrain. It is divided into two hemispheres that are connected to each other by a thick band of fibres called corpus callosum. The hemispheres are divided into lobes [28]:

Cerebral Cortex Lobes:

i) Frontal lobes are responsible of motor functions, planning reasoning, judgment, impulse control and memory. It is at the anterior portion of the cerebral cortex.

ii) Insula is associated with visceral functions and integrates autonomic information. It is located within the cerebral cortex, beneath the frontal, parietal and temporal opercula.

iii) Occipital lobes controls vision and colour recognition. It is the most caudal portion of the cerebral cortex.

iv) Parietal lobes are responsible for cognition, information processing, pain and touch sensation, spatial orientation, speech and visual perception. It is anterior to the occipital lobes and posterior to the frontal lobes.

v) Temporal Lobes are related with emotional responses, hearing memory and speech. It is anterior to occipital lobes and lateral to fissure of sylvius [28].

4.2 The Electroencephalogram (EEG)

As an historical evolution, in 1929, German psychiatrist Hans Berger indicated the existence of the electrical activities inside the brain by using galvanometer and electrodes located to the scalp. In 1930, he showed that the electrical activity varies by closing and opening the eyes and called these electrical waves "Electroencephalography" [53]. More technical description will be given in the sequel.

When the neurons of the human brain process information, they do so by changing the flow of electrical currents across their membranes. These changing currents generate electric and magnetic fields that can be recorded from the surface of the scalp. The electric fields are measured by attaching small electrodes to the scalp [29].

The potentials between different electrodes are then amplified and recorded as the "electroencephalogram" (EEG), which means writing out of the electrical activity of the brain (that which is inside the head) [29]. It is a completely noninvasive procedure that can be applied repeatedly in patients, normal adults, and children with virtually no risks or limitations. Local current flows are produced when brain cells (neurons) are activated. However, only electrical activity generated by large populations of neurons concurrently active can be recorded on the head surface. The small electrical signals detected by the scalp electrodes are amplified thousands of times, then displayed on paper or stored in computer memory [30].

4.2.1 EEG Frequency Bands

The basic EEG rhythms are summarized briefly as follows, with regard to their typical distribution on the scalp, subject states, tasks, physiological correlates, and the effects of training [31].

4.2.1.1 Delta (0.5-4 Hz):

Distribution: generally broad or diffused, may be bilateral, widespread.

Subjective feeling states: deep, dreamless sleep, non-REM sleep, trance, unconscious.
Associated tasks & behaviors: lethargic, not moving, not attentive.
Physiological correlates: not moving, low-level of arousal.
Effects of Training: can induce drowsiness, trance and deeply relaxed states.

4.2.1.2 Theta (4-8 Hz):

Distribution: usually regional, may involve many lobes, can be lateralized or diffuse;
Subjective feeling states: intuitive, creative, recall, fantasy, imagery, creative, dreamlike, switching thoughts, drowsy.
Associated tasks & behaviors: creative, intuitive; but may also be distracted, unfocused
Physiological correlates: healing, integration of mind/body.
Effects of Training: if enhanced, can induce drifting, trancelike state if suppressed, can improve concentration, ability to focus attention.

4.2.1.3 Alpha (8-12 Hz):

Distribution: regional, usually involves entire lobe; strong occipital with eyes closed.
Subjective feeling states: relaxed, not agitated, but not drowsy; tranquil, conscious.
Associated tasks & behaviours: meditation, no action.
Physiological correlates: relaxed, healing.
Effects of Training: can produce relaxation.

Sub-band low alpha: 8-10: inner-awareness of self, mind/body integration, balance.
Sub-band high alpha: 10-12: centring, healing, mind/body connection.

4.2.1.4 Beta (above 12 Hz) :

The beta band has a relatively large range, and has been defined as anything above the alpha band.

i) Low Beta (12-15 Hz), formerly "SMR":

Distribution: localized by side and by lobe (frontal, occipital, etc)
Subjective feeling states: relaxed yet focused, integrated

Associated tasks & behaviours: low SMR can reflect "ADD", lack of focused attention
Physiological correlates: is inhibited by motion; restraining body may increase SMR
Effects of Training: increasing SMR can produce relaxed focus, improved attentive abilities, may remediate ADD.

ii) Midrange Beta (15-18 Hz):

Distribution: localized, over various areas. It may be focused on one electrode.

Subjective feeling states: thinking, aware of self & surroundings
Associated tasks & behaviours: mental activity
Physiological correlates: alert, active, but not agitated
Effects of Training: can increase mental ability, focus, alertness, IQ

iii) High Beta (above 18 Hz):

Distribution: localized, may be much focused.
Subjective feeling states: alertness, agitation
Associated tasks & behaviours: mental activity, e.g. math, planning, etc.
Physiological correlates: general activation of mind & body functions.
Effects of Training: can induce alertness, but may also produce agitation, etc.

4.2.1.5 Gamma (40 Hz):

Distribution: Much localized
Subjective feeling states: thinking; integrated thought
Associated tasks & behaviours: high-level information processing, binding
Physiological correlates: associated with information-rich task processing
Effects of Training: not known [31].

4.2.2 Experimental Measurement Information about EEG

In the thesis EEG measurements have been performed from 9 healthy subjects (5 male, 4 female) with an average age of 27,4 (minimum 21, maximum 44). The EEG records were taken for eyes closed and eyes opened case with duration of 100 seconds each other. International 10-20 system was used for placements and measurements of the electrodes. The measurement procedure is given below:

The International 10-20 System: of Electrode Placement is the most widely used method to describe the location of scalp electrodes. The 10-20 system is based on the relationship between the location of an electrode and the underlying area of cerebral cortex. Each site has a letter (to identify the lobe) and a number or another letter to identify the hemisphere location.

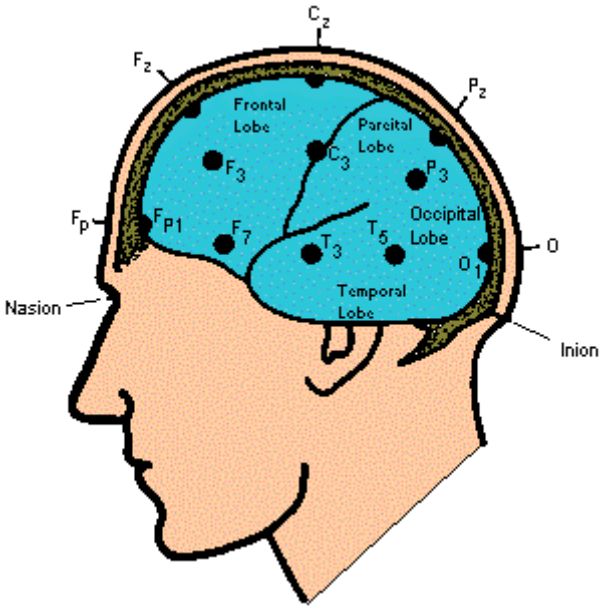

Figure 4.1 Cerebral cortex lobes and electrode displacement

The letters used in figure2.1 are: "F" - Frontal lobe, "T" - Temporal lobe , "C" - Central lobe , "P" - Parietal lobe, "O" - Occipital lobe. (Note: There is no central lobe in the cerebral cortex. "C" is just used for identification purposes only.)
Even numbers (2, 4, 6, 8) refer to the right hemisphere and odd numbers (1, 3, 5, 7) refer to the left hemisphere. "Z" refers to an electrode placed on the mid line. The smaller the number, the closer the position to the mid line. "Fp" stands for front polar. "Nasion" is the point between the forehead and nose. "Inion" is the bump at the back of the skull. The "10" and ""20" (10-20 system) refer to the 10% and 20% inter electrode distance [32].

Why are percentages used?

The skull may be different from subject to subject. A paediatric or adolescent may be smaller than an adult, and different adults have slightly different size head (and brains). The percentage relationship remains the same for the location of the internal brain lobes [33].

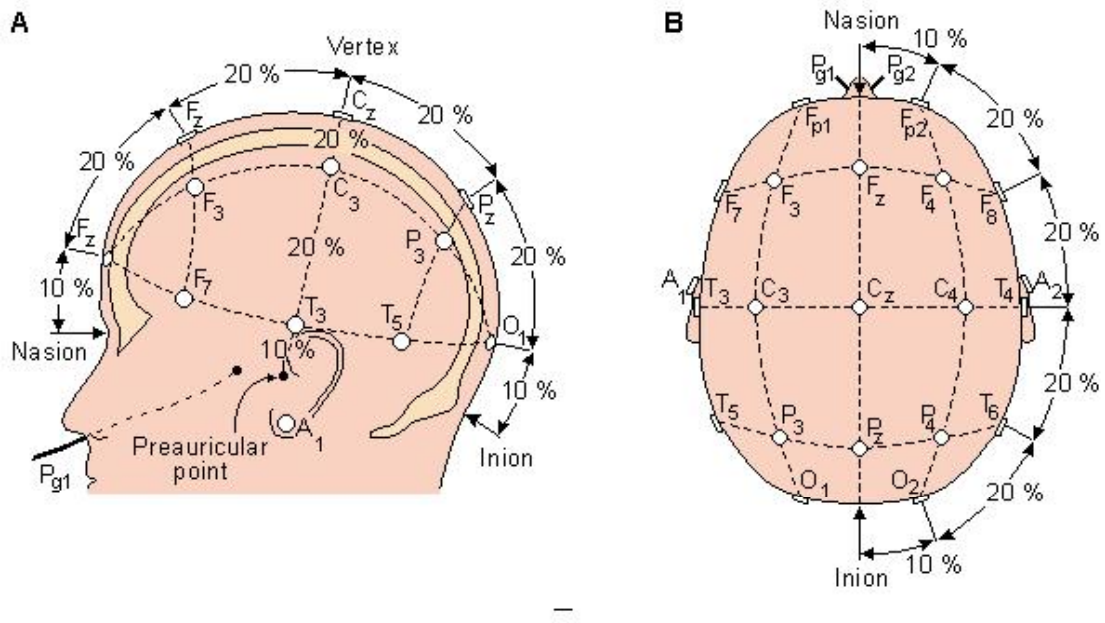

Figure 4.2 Spatial distributions of electrodes

When recording a more detailed EEG with more electrodes, extra electrodes are added utilizing the spaces in-between the existing 10-20 system. This new electrode-naming-system is more complicated giving rise to the Modified Combinatorial Nomenclature (MCN). This MCN system uses 1, 3, 5, 7, 9 for the left hemisphere which represents 10%, 20%, 30%, 40%, and 50% of the inion-to-nasion distance respectively. 2, 4, 6, 8, 10 are used to represent the right hemisphere. The introduction of extra letters allows the naming of extra electrode sites [32].

4.2.3 EEG in Medical Research

EEG signals can be used in various types of research areas due to the frequency variation based meaning of the state transitions. These are given below [53],

i) neurology: Discovering the brain pathology together with EMG, echocardiogram and neurological controls.

ii) neurosurgery: Discovering the abnormal pathological tissues like malignant tumors.

iii) anesthesia: To adjust the anesthesia level of the patient under anesthesia.

iv) pediatrics: Identify the visual and auditory problems of new born babies by using averaged evoked potentials.

v) psychiatry: Discovering the existence of any organic brain disorder in order to identify the mental abnormality more certainly.

vi) criminology: Discriminating the EEG patterns which belong to guilty person from innocent person.

CHAPTER 5

TIME-FREQUENCY ANALYSIS OF OBSERVED SIGNALS

In many applications of signal processing, one is interested in the frequency content of a signal locally in time. That is, the frequency evolves over time. Such signals are called *non-stationary*. For a non-stationary signal, $f(t)$, the standard Fourier Transform is not useful for analyzing the signal since the information which is localized in time such as spikes and high frequency bursts cannot be easily detected from the Fourier Transform.

5.1 Short Time Fourier Transform

Time localization can be achieved by first windowing the signal so as to cut off only a well-localized slice of $f(t)$ and then taking its Fourier Transform. This gives rise to the **Short Time Fourier Transform**, (STFT) or Windowed Fourier Transform. A function $f \in L^2(R)$ is used to represent a continuous signal and its Fourier transform $\hat{f}$ is,

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} f(t)\,dt \tag{5.1}$$

where $f \in L^2(R)$ means f has finite energy [54],

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt < \infty \tag{5.2}$$

and $f \in L^1(R)$ is,

$$\int_{-\infty}^{\infty} |f(t)|\, dt < \infty \quad . \tag{5.3}$$

The formula 5.1 gives the spectral information *f* the signal but it does not give any idea about local observation of *f*. To achieve this, a time window is needed.
A function $g \in L^2(R)$ can be window function if it satisfies the following requirement [54],

$$t \cdot g(t) \in L^2(R). \tag{5.4}$$

5.1.1 Time Frequency Windows and Uncertainty Principle

It is supposed that any window function $g \in L^2(R)$ has a region in the time frequency plane (t,ω), where the domain centre of the window function is,

$$t_m = \int_{-\infty}^{\infty} t \,|\, g(t) \,|^2 \, dt \tag{5.5}$$

and time domain radius of the window function is,

$$\Delta_t = \int_{-\infty}^{\infty} (t - t_m)^2 \,|\, g(t) \,|^2 \, dt \quad . \tag{5.6}$$

The centre frequency of the window function *g* is defined as,

$$\omega_m = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega \,|\, g(\omega) \,|^2 \, d\omega \tag{5.7}$$

and the frequency radius of the window function is,

$$\Delta_\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\omega - \omega_m)^2 \,|\, g(\omega(|^2 \, d\omega \quad . \tag{5.8}$$

The time and frequency widths are $2\Delta_t$ and $2\Delta_\omega$ respectively, since the window function is real and symmetric. By using these formulas one can determine the window region as,

$$[\, t_m + b - \Delta_t ,\ t_m + b + \Delta_t] \times [\, \omega + \omega - \Delta_\omega ,\ \omega + \omega + \Delta_\omega] \ . \qquad (5.9)$$

The area of the time frequency window is

$$4\Delta_t\Delta_\omega \qquad . \qquad (5.10)$$

The area found in Eq. 5.10 constitutes Heisenberg Boxes and the width of the time frequency window remains unchanged for localizing signals in both high and low frequencies. According to the Heisenberg Uncertainty principle [55]

$$\Delta_t\Delta_\omega \geq \frac{1}{2} \qquad (5.11)$$

then the Eq. 5.11 becomes $\Delta_t\Delta_\omega = \frac{1}{2}$ if the window function *g* is Gaussian [55].

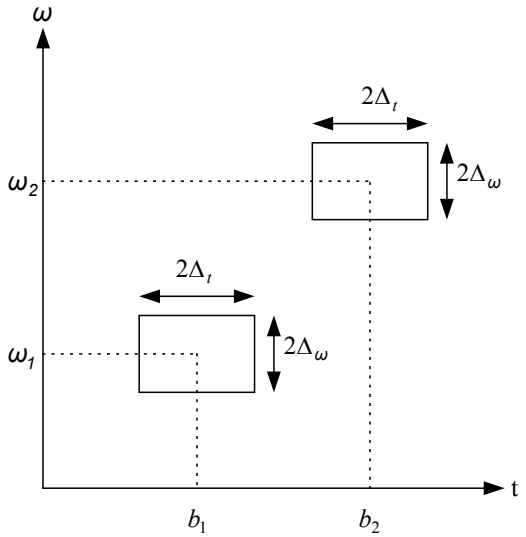

Figure 5.1 Representation of time frequency windows in STFT, $b_1 < b_2$ and $\omega_1 < \omega_2$.

5.1.2 Short Time Fourier Transform

A real and symmetric window $g(t) = g(-t)$ which satisfies the condition given in Eq. 5.4, is translated by b and modulated by ω:

$$g_{b,\omega}(t) = e^{j\omega t} g(t-b) \quad . \tag{5.12}$$

The window function g is normalized $\|g(.)\|_2 = 1$ so that $\|g_{b,\omega}(.)\|_2 = 1$ for any $(b,\omega) \in R^2$. The windowed Fourier transform of $f \in L^2(R)$ is,

$$Sf(b,\omega) = \int_{-\infty}^{\infty} f(t) g(t-b) e^{-j\omega t} dt \quad . \tag{5.13}$$

The transform given in 5.13 is called *Short Time Fourier Transform* since the term $g(t-b)$ localises the Fourier integral in the neighbourhood of $t = b$. Various window function types are given in table 5.1.

Table 5.1 Window function types for STFT [55].

Name	$g(t)$	Δ_ω
Rectangle	1	0.89
Hamming	$0.54+0.46\cos(2\pi t)$	1.36
Gaussian	$\exp(-18t^2)$	1.55
Hanning	$\cos^2(\pi t)$	1.44
Blackman	$0.42+0.5\cos(2\pi t)+0.08\cos(4\pi t)$	1.68

Frequency parameters of windows shown in table 5.1 are restricted [-0.5 0.5]. These windows are normalized so that $g(0) = 1$ but $\| g(.) \| \neq 1$.

5.1.3 Gabor Transform

The Gabor transform is a windowed Fourier transform with any Gaussian function g_α as the window function. The optimal window for time localization is achieved by using Gaussian function,

$$g_\alpha(t) = \frac{1}{2\sqrt{\pi\alpha}} e^{-\frac{t^2}{4\alpha}} \tag{5.14}$$

where $\alpha>0$ fixed. The optimal window is Gaussian because it gives minimum window size (i.e highest resolution in both time and frequency) that fits to the Uncertainty Principle given in 5.1.1. The Gabor transform of an $f \in L^2(R)$ is defined by

$$(G_b^\alpha f)(\omega) = \int_{-\infty}^{\infty} f(t) g(t-b) e^{-j\omega t} dt \quad . \tag{5.15}$$

The width of the window is determined by the positive α and is found by,

$$\Delta_{g_\alpha} = \frac{1}{\| g_\alpha(.) \|_2} \left\{ \int_{-\infty}^{\infty} t^2 g_\alpha^2(t) dt \right\}^{1/2} \tag{5.16}$$

Since g_α is even, the width of the function is $2\Delta_{g_\alpha}$ and for each $\alpha>0$,

$$\Delta_{g_\alpha} = \sqrt{\alpha} \, . \tag{5.17}$$

That is, the width of the function is $2\sqrt{\alpha}$. The time frequency region is then obtained as

$$[\, b - \sqrt{\alpha} \, , b + \sqrt{\alpha} \,] \times [\omega - \frac{1}{2\sqrt{\alpha}} \, , \omega + \frac{1}{2\sqrt{\alpha}}] \quad . \tag{5.18}$$

It is found that $\Delta_t = \sqrt{\alpha}$ and by the similar way $\Delta_\omega = \frac{1}{2\sqrt{\alpha}}$ and the result $\Delta_t \Delta_\omega = \frac{1}{2}$ which is the smallest value of Eq.5.11.

5.2 Wavelet Transform

In section 5.1, it was observed that time frequency window of the short time Fourier transforms is constant for all frequency and time segments. This limits to obtain various resolutions at different time frequency ranges. Therefore, more flexible basis functions are needed which are added dilation parameter in addition to translation parameter so that window size can change at different frequencies.

5.2.1 Integral Wavelet Transform

Wavelets ψ are the small waves that fast decays to zero that is,

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad . \tag{5.19}$$

This is the reason why they are called *wavelets*. A wavelet function ψ is called basic wavelet if it satisfies the admissibility condition,

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad . \tag{5.20}$$

The integral wavelet transform of a function f is,

$$W_{\psi} f(b,a) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi^*\left(\frac{t-b}{a}\right) dt \qquad f \in L^2(R) \tag{5.21}$$

The function f can be reconstructed from its wavelet transform if the admissibility condition is satisfied. The reconstruction formula is,

$$f(t) = \frac{1}{C_{\psi}} \int_{0}^{\infty} \int_{-\infty}^{\infty} W_{\psi} f(b,a) \, f(t) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) db \frac{da}{a^2} \tag{5.22}$$

where $a,b \in R$ are scale and translation parameters respectively. The time frequency window in continuous wavelet transform is expressed as

$$\left[b + at_m - a\Delta_\psi , b + at_m - a\Delta_\psi\right] x \left[\frac{\omega_m}{a} - \frac{1}{a}\Delta_{\hat{\psi}} , \frac{\omega_m}{a} + \frac{1}{a}\Delta_{\hat{\psi}}\right] \quad . \qquad (5.23)$$

From Eq.5.23, it is seen that the time width of the window function is $2a\Delta_\psi$ and the frequency width is $\frac{2}{a}\Delta_{\hat{\psi}}$. The area of the window is found $4\Delta_\psi \Delta_{\hat{\psi}}$, it is independent of the scale parameter a, but the time and frequency width of the window function changes with scale. The time frequency windows for wavelet transform is shown in Fig. 5.2

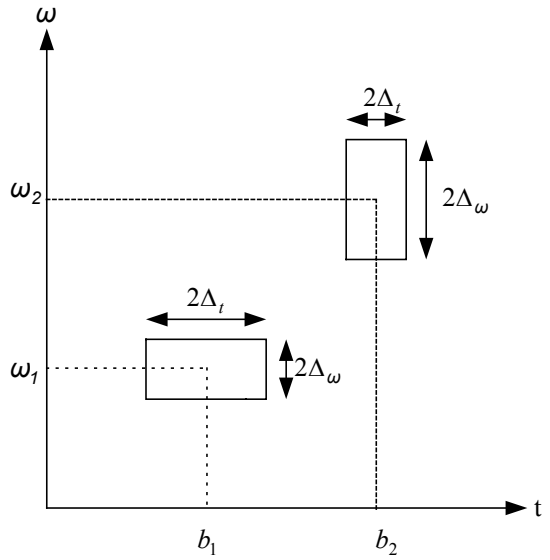

Figure 5.2 Time frequency windows in wavelet transform, $b_1 < b_2$ and $\omega_1 < \omega_2$.

5.2.2 Wavelet Decompositions

In wavelet analysis, waves can be partitioned into consecutive "*octaves*" or namely frequency bands. For computational efficiency, the powers of 2 are used frequency partitioning, then the mother wavelet is ψ considered as

$$\psi(2^j t-k) \quad j,k \in Z. \qquad (5.24)$$

If a function $\psi \in L^2R$ has unit length, then $\psi_{j,k}$ defined by,

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^{j/2}t - k) \quad j,k \in Z \tag{5.25}$$

also has unit length, that is $\| \psi_{j,k}(.) \|_2 = \| \psi(.) \|_2 = 1$. Every function $f \in L^2(R)$ can be written as

$$f(t) = \sum_{j,k=\infty}^{\infty} c_{j,k}\psi_{jk}(t) \tag{5.26}$$

and the expression for f given in Eq. 5.26 is called *wavelet series*. The *wavelet coefficients* are shown as

$$c_{j,k} = (W_\psi f)(\frac{k}{2^j}, \frac{1}{2^j}) \quad . \tag{5.27}$$

For any wavelet $\Psi_{j,k}$ where $j,k \in Z$, let W_j denotes the closure of the linear span of $\Psi_{j,k}$, then W_j is defined as [54],

$$W_j \overset{\Delta}{=} clos_{L^2(R)} < \psi_{j,k} : k \in Z > \quad . \tag{5.28}$$

Then the $L^2(R)$ can be decomposed as a direct sum of the spaces W_j,

$$L^2(R) = \sum_{j \in Z}^{\bullet} W_j = \ldots \dot{+} W_{-1} \dot{+} W_0 \dot{+} W_1 \ldots \tag{5.29}$$

where the term $\dot{+}$ denotes the direct sum. According to Eq 5.29, it can be said that every function $f \in L^2(R)$ has a unique decomposition [54],

$$f(t) = \ldots + g_{-1}(t) + g_0(t) + g_1(t) + \ldots \tag{5.30}$$

where $g_j \in W_j$, $\forall\, j \in Z$. If Ψ is an orthogonal wavelet, then the subspaces W_j of $L^2(R)$ are mutually orthogonal,

$$\langle g_j, g_m \rangle = 0, \quad j \neq m \qquad \text{where } g_j \in W_j \text{ and } g_m \in W_m \ . \tag{5.31}$$

In this case, the notation will be $W_j \perp W_m$ $j \neq m$ and consequently the direct sum becomes orthogonal sum [54],

$$L^2(R) = \bigoplus_{j \in Z} W_j \overset{\Delta}{=} \ldots \oplus W_{-1} \oplus W_0 \oplus W_1 \ldots \ . \tag{5.32}$$

The decomposition in Eq. 5.32 is usually called "orthogonal decomposition". Any wavelet generates a direct sum decomposition given in Eq. 5.29 of $L^2(R)$. For each $j \in Z$ consider the closed subspace,

$$V_j = \ldots \dot{+} W_{j-2} \dot{+} W_{j-1} \tag{5.33}$$

and these subspaces have the following properties [54]:

i) $\ldots \subset V_{-1} \subset V_0 \subset V_1 \ldots$

ii) $\operatorname{clos}_{L^2}\left(\bigcup_{j \in Z} V_j \right) = L^2(R)$

iii) $\bigcap_{j \in Z} V_j = \{\theta\}$

iv) $V_{j+1} = V_j + W_j$, $j \in Z$ and

v) $f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}$, $j \in Z$.

The subspace given in Eq. 5.33 is generated by a function, called *scaling function*, $\phi \in L^2(R)$ and W_j is generated by some wavelet $\Psi \in L^2(R)$. The scaling function supplies complement of information corresponding to the wavelet transform of a function f, $Wf(b,a)$, for $a > a_0$ when $Wf(b,a)$ is known only for $a < a_0$. The relation between scaling function and the wavelet function is given in terms of their Fourier transform, which is defined as [55],

$$|\phi\ (\omega)|^2 = \int_1^{\infty} |\hat{\psi}(a\omega)|^2\ \frac{da}{a} \tag{5.34}$$

The multiresolution decomposition separates the signal into detail coefficients at different scales and the remaining part being a coarser representation of the signal called approximate coefficients which are denoted by D_j and A_j for each level *j*, respectively. The detail coefficients "D" correspond to high pass filtered part of the signal and approximate coefficients "A" correspond to low pass filtered part [38]. The representation of wavelet decomposition of signal is shown in Fig 5.3 schematically and wavelet decomposition of EEG signal has been illustrated in Fig. 5.4.

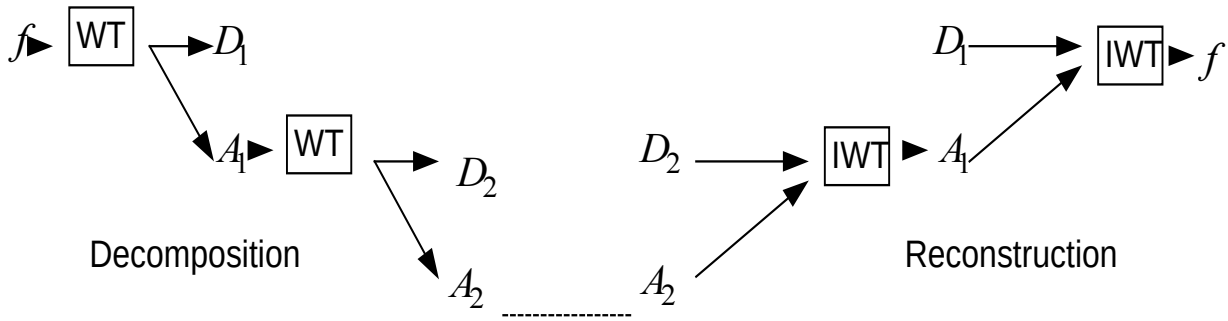

Figure 5.3 Wavelet decomposition and reconstruction scheme of signal *f*.

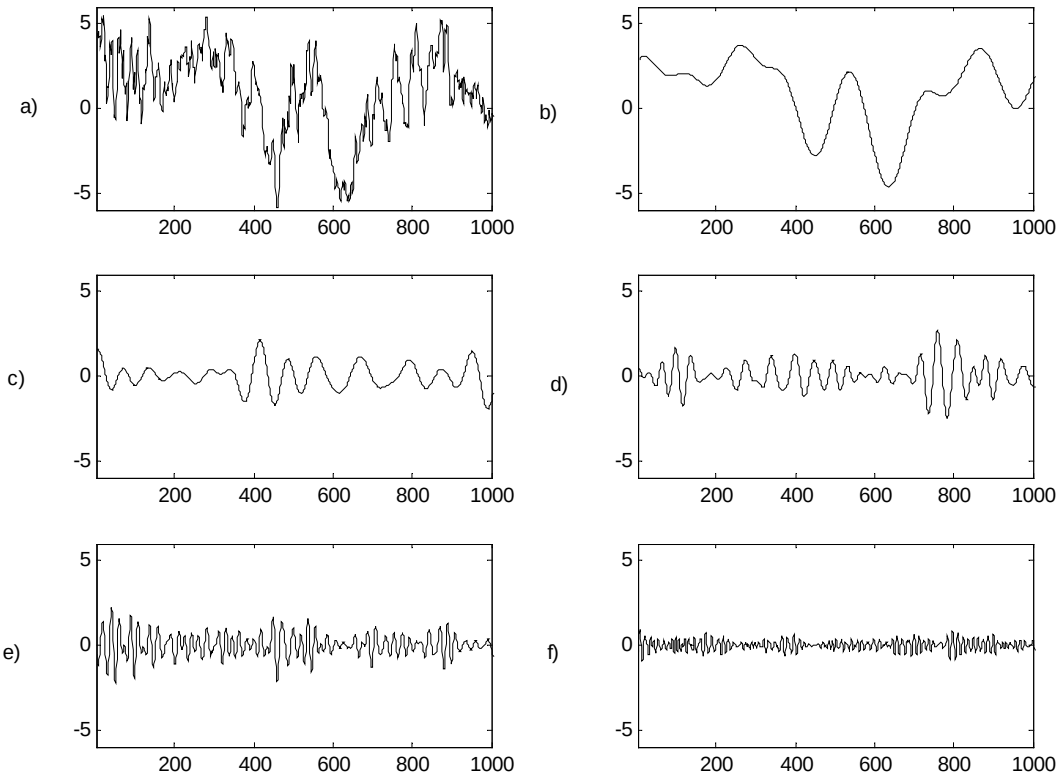

Figure 5.4 Wavelet decomposition of an EEG signal, a) Original signal, b) Delta band, c) Theta band, d) Alpha band e) Beta band f) Gamma band

5.3 Time Frequency Energy Distributions

The energy distribution of a signal in time-frequency domain is represented by *spectrogram* or *scalogram* obtained by short time Fourier transforms and wavelet transforms respectively.

5.3.1 Spectrogram

Spectrogram of a signal $f(t)$ localised by window function $g(t)$ is the square modulus of STFT and it is defined as

$$|\mathrm{S}f(\mathrm{b},\omega)|^2 = \left| \int_{-\infty}^{\infty} f(\mathrm{t})\mathrm{g}(\mathrm{t}-\mathrm{b})\mathrm{e}^{-\mathrm{j}\omega t}\mathrm{dt} \right|^2 \quad . \tag{5.35}$$

It measures the energy of f in time frequency neighbourhood of (b,ω). This energy distribution has constant time frequency window, hence the resolution is limited. The total energy found by spectrogram can be shown as [35],

$$E_T = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Sf(b,\omega)|^2 \, dtd\omega \quad . \tag{5.36}$$

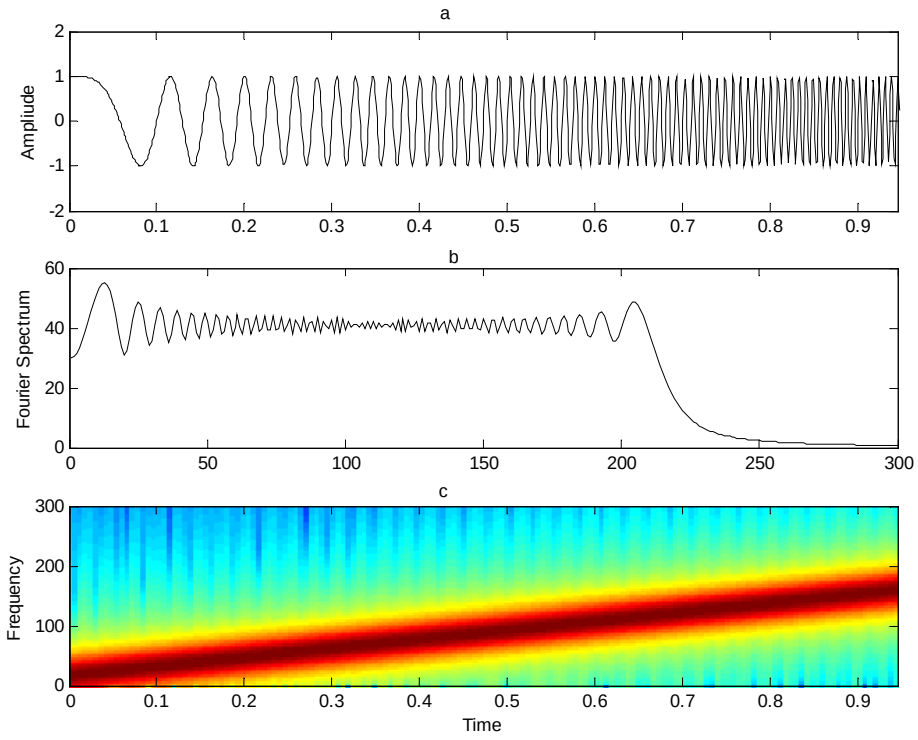

Figure 5.5 Spectrogram of linear chirp signal. a) Time domain, b) Fourier spectrum c) Spectrogram

In Fig. 5.5, the spectrogram of a linear chirp signal, which is expressed as, $s(t) = e^{jat^2}$ is given with a sampling frequency of 1kHz.

5.3.2 Scalogram

Scalogram is the square modulus of the wavelet transform and defined by,

$$\left|\mathrm{W}f(\mathrm{b},\mathrm{a})\right|^2 = \left|\int_{-\infty}^{\infty} f(\mathrm{t})\frac{1}{\sqrt{a}}\psi\left(\frac{\mathrm{t}-b}{a}\right)\mathrm{dt}\right|^2 \tag{5.37}$$

and the total energy found by scalogram is denoted as [35],

$$\mathrm{E_T} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\mathrm{W}(b,a)\right|^2 \frac{\mathrm{d}b\mathrm{d}a}{a^2} \quad . \tag{5.38}$$

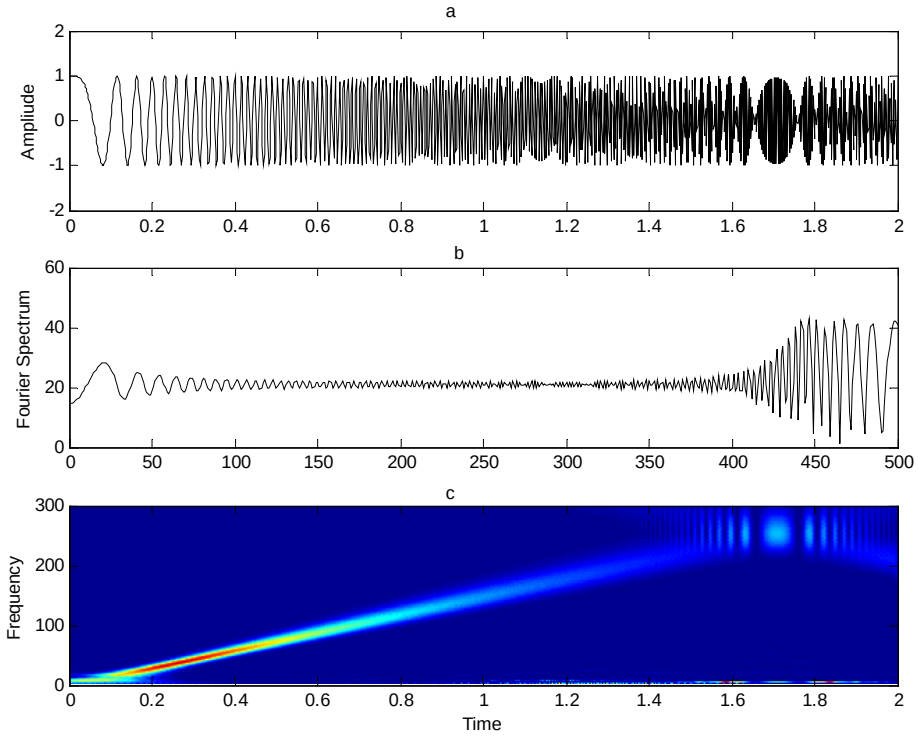

Figure 5.6 Scalogram of linear chirp signal, a) Time domain signal b) Fourier spectrum c) Scalogram

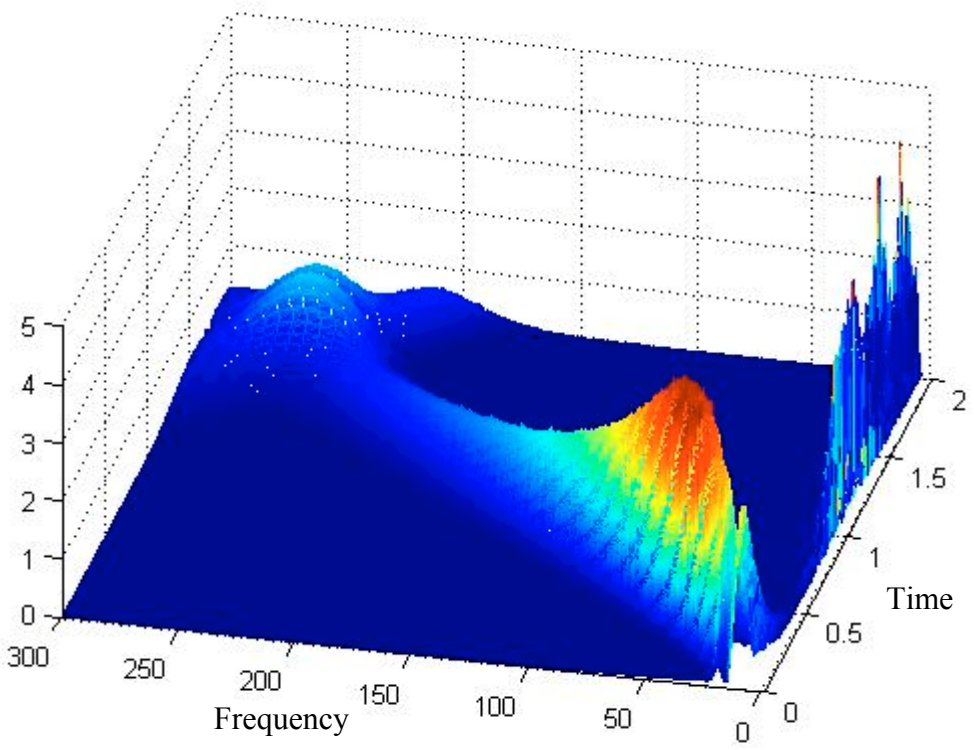

Figure 5.7 Wavelet transform of a linear chirp signal.

Scalogram has more sensitive time frequency resolution than spectrogram due to flexible window function. In Fig 5.6 and 5.7, the scalogram and wavelet transform of linear chirp signal has been shown where the sampling frequency is 512 Hz.

5.4 Classification with Wavelets

The methods containing wavelet transforms can be used to classify the nonstationary signals. They are preferred when the classifications are needed for short time recording applications since nonlinear classifiers such as correlation dimension D_2 and maximum Lyapunov exponent λ_{max} require long time evolution of trajectory. The discriminating parameters obtained from wavelet transforms are wavelet ridges and the wavelet entropies [37].

5.4.1 Wavelet Entropy

The wavelet entropy is obtained from relative wavelet energies, which provide information about the relative energy associated with different frequency bands in time

domain signal [37]. The wavelet analysis is performed by taking dilation parameter $a=2^{-j}$ and the translation parameter $b=2^{-j}k$ of mother wavelet with $j,k\in Z$.

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^{j/2}t - k) \quad . \tag{5.39}$$

Then, the energy of the detail signal for each resolution level is given by square modulus of wavelet coefficients by the following form [37],

$$E_j = \sum_{k=1}^{K} \left| C_j(k) \right|^2 \tag{5.40}$$

where K is data length, the resolution level j=-1,...,-N and the wavelet coefficients $C_j(k)$ contain the information of signal f between the frequencies $2^{j-1}\omega_s \le |\omega| \le 2^j\omega_s$ where $\omega_s=2\pi f_s$ and f_s is the sampling frequency. The energy at each sampled time will be

$$E(k) = \sum_{j=-N}^{-1} \left| C_j(k) \right|^2 \quad . \tag{5.41}$$

The total energy can be obtained by

$$E_{tot} = \sum_{j=-N}^{-1} \sum_{k=1}^{K} \left| C_j(k) \right|^2 = \sum_{j=-N}^{-1} E_j \quad . \tag{5.42}$$

Then the normalized values which represent the relative wavelet energy [37],

$$p_j = \frac{E_j}{E_{tot}} \quad . \tag{5.43}$$

In Eq 5.43, p_j can be considered as time-scale density. This parameter enables to detect a variation in a certain frequency band. The expression wavelet entropy is obtained from the relative wavelet energy shown in Eq 5.43. Wavelet Entropy "WE" appears as a measure of the degree of order/disorder of the signal. It can be shown as,

$$WE = -\sum_{j=-N}^{-1} p_j \ln(p_j) \tag{5.44}$$

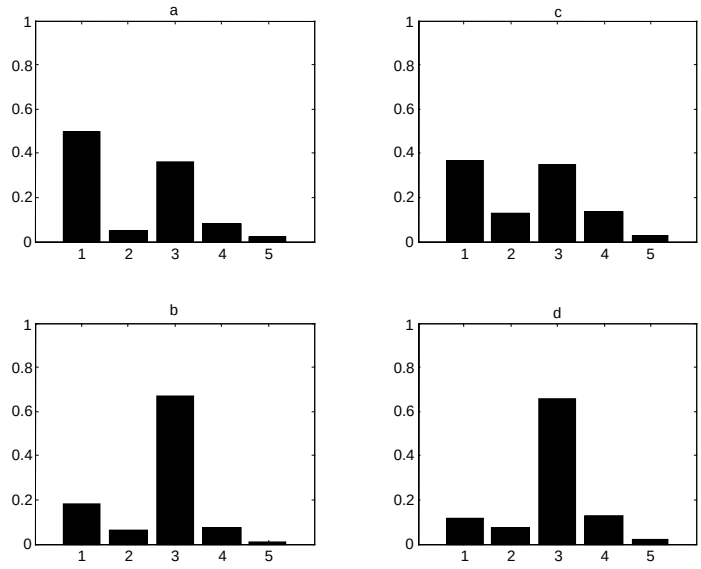

Figure 5.8. Relative wavelet energies (RWE) of delta, theta, alpha, beta and gamma respectively a) Channel O1-eyes opened b) Channel O1-eyes closed c) Channel O2-eyes opened d) Channel O2 eyes closed.

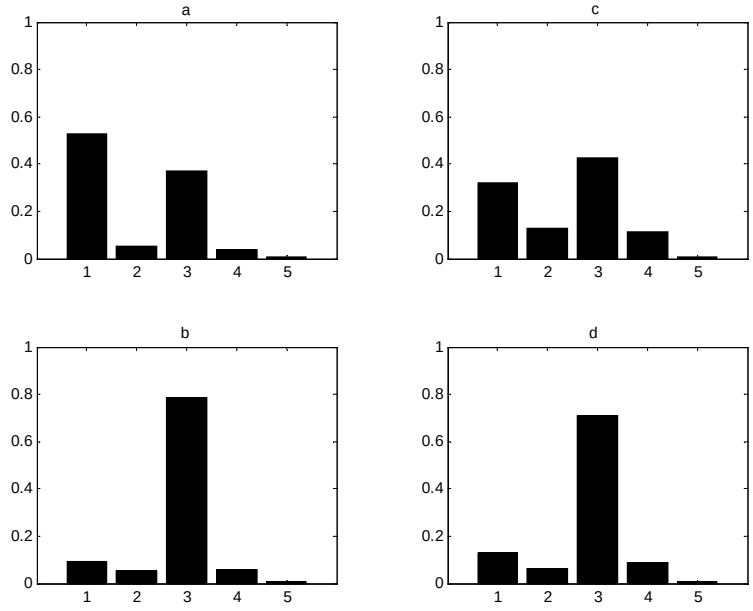

Figure 5.9 Relative wavelet energies (RWE) of delta, theta, alpha, beta and gamma respectively a) Channel P3-eyes opened b) Channel P3-eyes closed c) Channel P4-eyes opened d) Channel P4 eyes closed.

In Fig. 5.8 and 5.9, the relative wavelet energies of delta, theta, alpha, beta and gamma bands has been illustrated for different channels when eyes are opened and closed respectively. The wavelet entropy result which is given in Fig 5.10 indicates that when the eyes are closed, then the wavelet entropy decreases. The wavelet entropy is a measure of the degree of order/disorder of the signal and an ordered signal can be thought to have narrowband spectrum therefore wavelet decomposition of such a signal will be dominant in one resolution level whereas a signal generated from random process has wavelet resolution in all frequency bands and will represent a disordered behaviour. Then it can be said that signals taken from closed eyes exhibit an ordered activity and have lower wavelet entropy. The results in Fig. 5.10 illustrate the wavelet entropy variation for opened and closed eyes condition and confirms this idea.

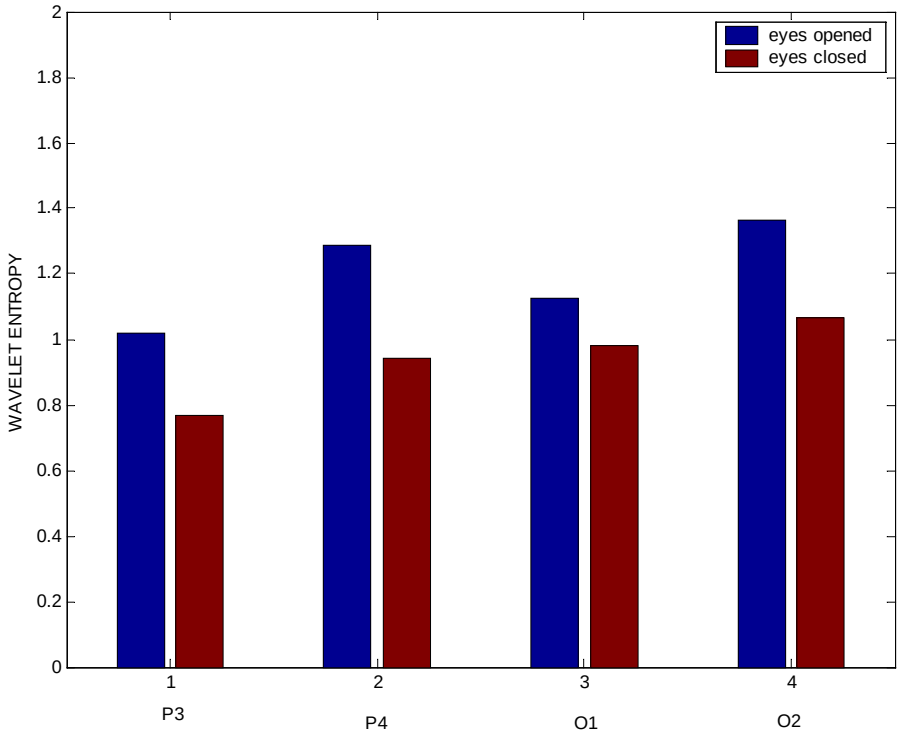

Figure 5.10 Wavelet Entropy results for four channel P3,P4,O1 and O2 respectively.

5.4.2 Wavelet Ridges

To find the wavelet ridges of wavelet transformed signal can be used as a second classification parameter. The skeleton of the wavelet transform of *s*(t) is the wavelet transform evaluated on the ridge [41]. The ridges are determined by different methods, which can distinguish as phase, and modulus based methods.

5.4.2.1 The Stationary Phase Method

In this method the time derivative of the phase of the wavelet transform is given as,

$$\left.\frac{\partial \Phi_{a,b}}{\partial b}\right|_{a(b)} = \frac{1}{a}\varphi'_{\psi}(0) \tag{5.45}$$

indicates the ridges in time-frequency domain where

$$\Phi_{\boldsymbol{a},\boldsymbol{b}}(t) \stackrel{\Delta}{=} \varphi(t) - \varphi_{\psi}\left(\frac{t-b}{a}\right) \tag{5.46}$$

and $\varphi(t)$ is phase component of complex signal $Z_s(t)$ defined as,

$$Z_s(t) \stackrel{\Delta}{=} s(t) + j\,\tilde{s}(t) = A(t)\exp(\mathrm{j}\,\varphi(\mathrm{t})) \tag{5.47}$$

where $\tilde{s}(t)$ is the Hilbert transform of real signal $s(t)$ given in [49]. The term $\varphi_{\psi}\left(\frac{t-b}{a}\right)$ is described as the phase component of asymptotic mother wavelet which is given by [42],

$$\psi(t) = A_{\psi}(t)e^{\varphi_{\psi}(t)} \tag{5.48}$$

The points given by Eq. 5.45 are called stationary phase points, so this method is called as stationary phase method.

5.4.2.2 Carmona Method

Carmona method is a modulus based ridge extraction method. Extracting the ridge from the modulus of the transform is more robust than the phase, because the extraction from the modulus does not involve differentiation of phase [41]. In Carmona method, the ridges are obtained by the solution of an optimization problem. The

minimization of penalty function is equivalent to maximization of the modulus of the transform [41]. The ridge is searched as a(b)=φ(b) which minimizes the following penalty function,

$$F(\varphi(b)) = -\int \left|W_s(b,a)\right|^2 db + \int \left[\lambda\varphi'(b)^2 + \mu\varphi''(b)^2\right] db \qquad (5.46)$$

And in Eq. 5.46, the first term maximizes energy density along the ridge and the second term contains first and the second derivative of the ridge function φ(b) is concerned with the smoothness of the ridge [42]. This problem can be solved by simulated annealing algorithm, which is given in [41].

5.4.2.3 Simple Method

In this method, the maximum components of the modulus of the wavelet transform is searched when the mother wavelet is assumed to have a peak at t=0 [42] and the most of the energy of the signal will be localized around the ridge. Therefore, simplest way of determining wavelet ridge is to compute the scalogram where,

$$\left.\frac{\partial Ps(a,b;\psi)}{\partial a}\right|_{a=a(b)} = 0 \quad \text{and} \quad \left.\frac{\partial^2 P_s(a,b,\psi)}{\partial a^2}\right|_{a=a(b)} < 0 \qquad (5.47)$$

and $P_s(a,b,\psi)$ is the square of the modulus of the wavelet transform coefficients. The simple method works well when there is no contamination by noise. In noisy situations there will be spurious local maximum and this will prevent to detect the ridges successfully. In the thesis, since the aim is to exhibit the different patterns under different conditions, simple method has been applied due to fast execution. In Fig 5.11 the ridges of linear chirp signal have been plotted. Fig. 5.12 indicates the wavelet transform of the mean of ten subjects where at time t=100 seconds eyes of the ten subjects are closed and an abrupt increase of magnitude of wavelet coefficients at about 10 Hz can be clearly seen in the second half of recording.

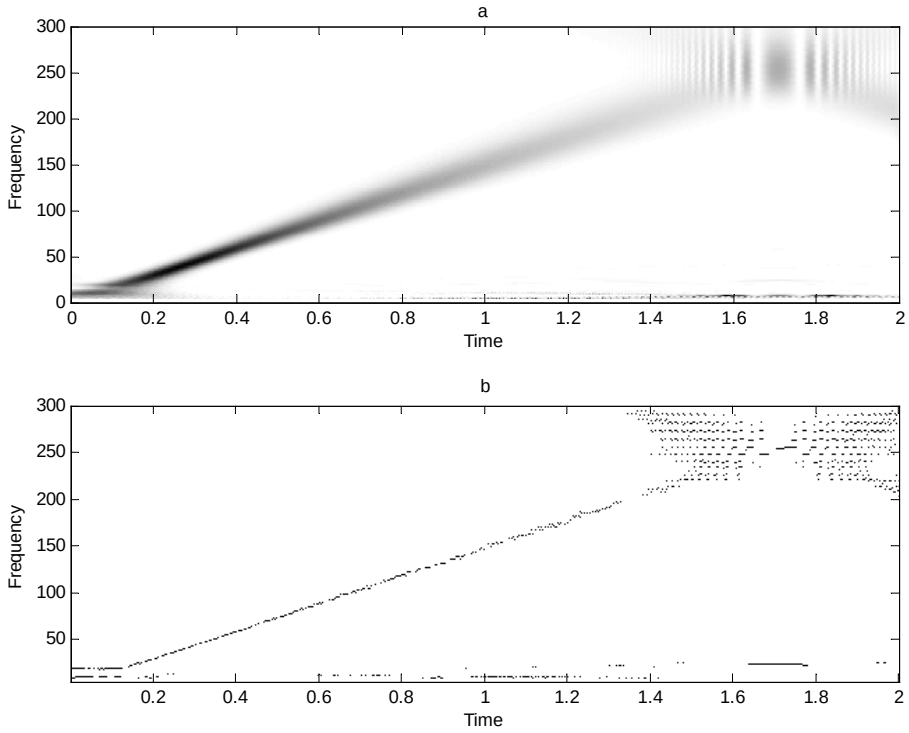

Figure 5.11 a) Wavelet transform of linear chirp signal b) Wavelet ridges

In Fig. 5.13, the measurements have been done while eyes are opened until t=100 sec and at the second half, measurements have been done while eyes are closed. It can be observed that after t=100 sec the ridges are concentrated near to 10Hz, whereas there is weak signal density near 10 Hz between the interval t=0-100 sec.

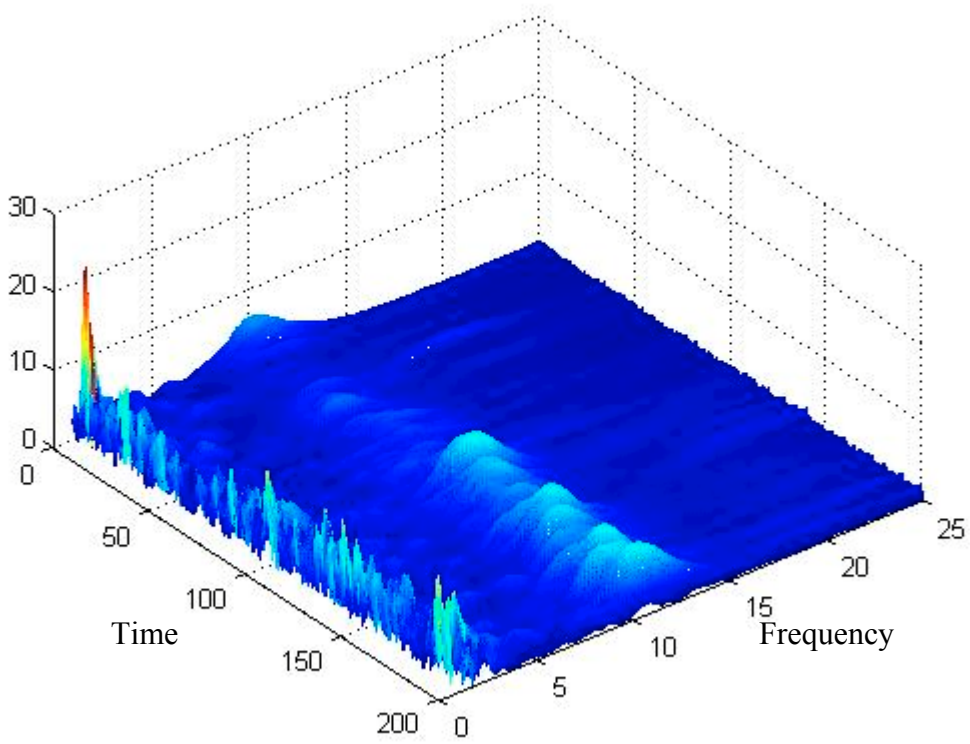

Figure 5.12 Wavelet transform of spontaneous EEG data

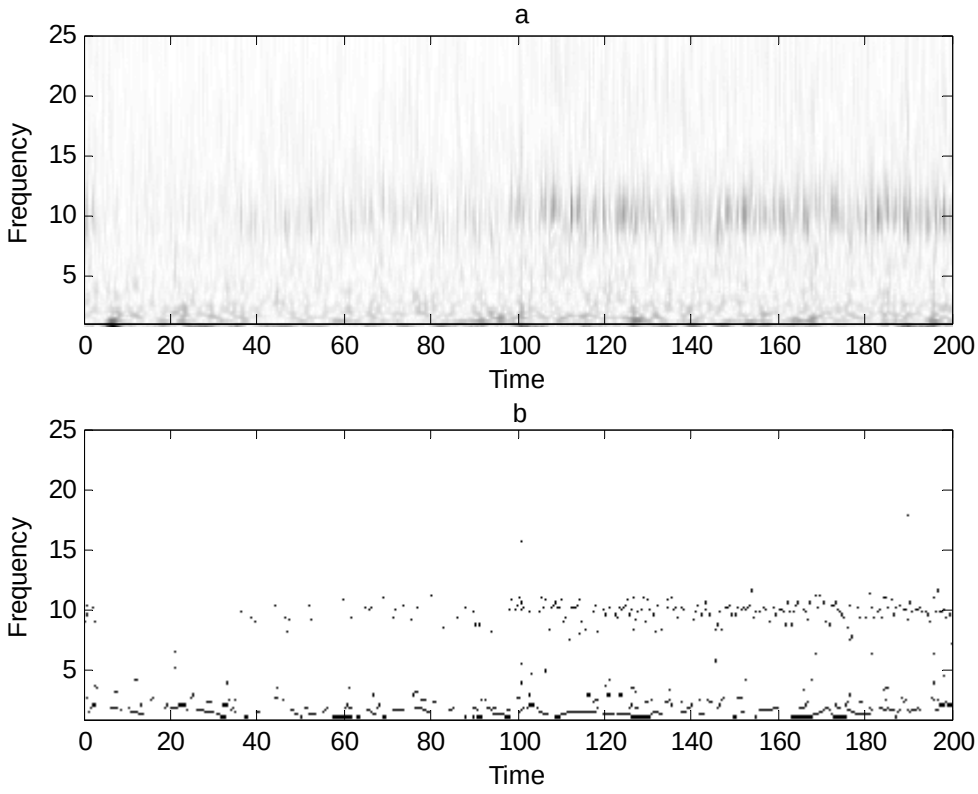

Figure 5.13 a) Scalogram of spontaneous EEG Signal b) Wavelet ridges

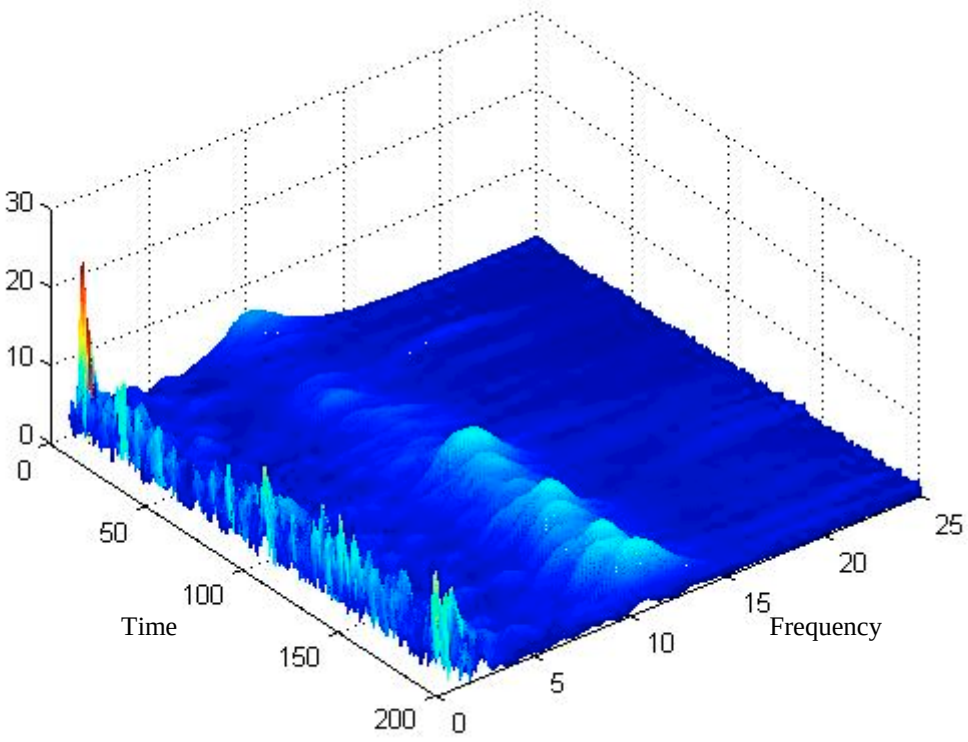

Figure 5.12 Wavelet transform of spontaneous EEG data

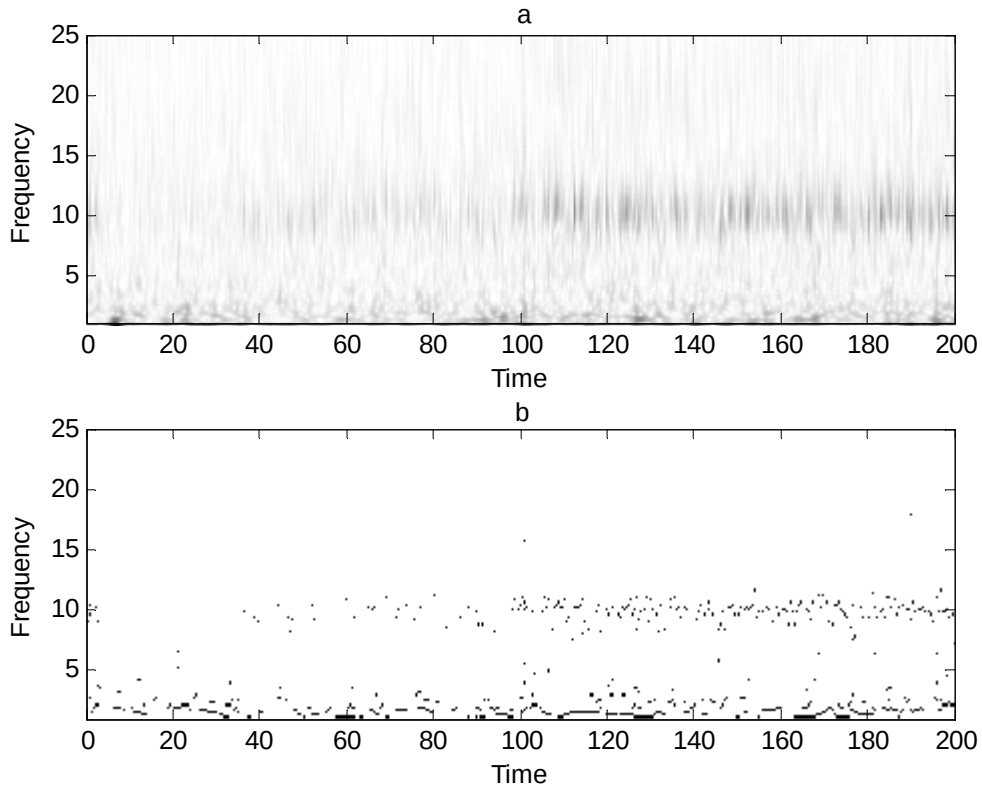

Figure 5.13 a) Scalogram of spontaneous EEG Signal b) Wavelet ridges

CHAPTER 6

CHAOS IN BRAIN

Chaos analysis is a critical point in brain researches. The electroencephalography (EEG) defined in section 4.2, is used for chaos analysis of the brain. The main goal is to classify the EEG signals under different states and conditions. According to the theory of chaos, it is reasonable to assume that the neuronal network underlying an EEG is a complex dynamical system [56] and besides in [6], neural networks are said to represent broadband high dimensional chaos. In order to make a classification, the invariants of the system have to be found where correlation dimension (D_2), is an adequate choice as a discriminating parameter. It is used to distinguish random activity from deterministic activity. To be able to find D_2 data length should be sufficiently long, so that attractor structure can be observed in the phase space. Since evoked potentials (EP) and event related potentials (ERP) are based on an applied stimulus and the short time range before and after the stimulus is treated, data length is not sufficient for chaos analysis. That is why correlation dimensions have been calculated for spontaneous EEG data in application part of this thesis.

6.1 Correlation Dimension Parameters

In Chapter 2, the methods for obtaining the necessary parameters which are time delay and embedding dimension in order to obtain correlation dimension have been given. However, in practical applications, it has been seen that correlation dimension can have various values depending on some other parameters such as sampling frequency, data length and upper limit of correlation dimension. Therefore, it is necessary to take into account these additional parameters so that correlation dimension can be used as an invariant measure.

6.1.1 Sampling Frequency

Improper choice of sampling frequency may be the source of an error. Although time domain signals are continuous in nature, they are recorded discretely. The standard

criterion is the Nyquist rate which says that the lowest sampling frequency must be at least two times greater than the bandwidth of the signal. In applications, EEG is also digitized into discrete points and information between neighbouring sampling points is lost. To avoid this, one may consider using very high sampling frequency but in that case computation will take long time for long recordings and memory will be insufficient. Hence, the sampling frequency should be decreased but there is need to obtain optimum value to overcome information losing problem due to select an improper sampling frequency can affect the estimates of the dimensional complexity. In order to determine sampling frequency, in [56], an application is given about the variation of correlation dimension versus different sampling frequencies and resolutions, which are obtained from analogue to digital converters. The results of this work are given in Table 6.1.

Table 6.1 Correlation Dimensions versus Sampling Rate and Resolution [56].

Sampling Frequency (Hz) / Resolution (bit)	128	256	512	1024
8	7.003±0.60	6.608±0.42	6.425±0.42	6.314±0.40
12	7.002±0.49	6.699±0.44	6.502±0.44	6.412±0.40
16	6.997±0.48	6.680±0.45	6.485±0.42	6.392±0.39

The sampling frequencies are increased by factor two because it is purposed that the midpoint between two samples would be included. Then the results are compared with highest sampling rate because it is nearest to the waveform. For the frequencies 256 and 128 Hz, the correlation dimension is relatively higher, that is, one observes a false complexity because of insufficient sampling rate. After 512 Hz, there is no significant difference among the correlation dimensions, so the optimum value can be selected as 512 Hz because it corrupts less location in memory of computer compared with higher sampling rates.

6.1.2 Time Delay *T*

In section 2.2.1, suitable algorithms for determining *T* have been given. But in any case a check for stationarity of this parameter is strongly recommended while EEG

signals are under consideration. . Therefore, the saturation of system invariants methods should also be applied for each time lag *T*. The value, after which D_2 as a function of *T* becomes constant, is chosen as time delay parameter.

6.1.3 Data Length

In order to get a good evaluation of D_2, very long time recordings are needed. But, in practice it is not possible when the EEG data do not remain stationary for long data length. Especially, for high dimensional systems, short time series yield wrong estimate for correlation dimension underestimate the correlation dimension. As an example, for alpha waves, correlation dimension varies from 2.6 to 6.6 [45]. Therefore, to be able to accept the data length as sufficient, the substantial increase in time series should not change the correlation dimension. That is, there should be no effect of data length in the invariants of the system.

6.1.4 Upper Limit of Correlation Dimension

When analyzing REM sleep and different stages of waves during the sleep, they look like highly random signals and one can find values D_2=9.7 and D_2=8.9 [45]. When dimension reaches to such high values, it is suspicious that these dimensions indicate very highly deterministic chaos. The reason of suspicion is that even pure noise signal becomes to saturate at high embedding dimensions and gives correlation dimension of order nine or ten. To rely on these values, one should be sure that the observed data is strongly noise free. Therefore it is recommended that any value of D_2>8 is considered as unreliable [45].

When these precautions are taken into account, the correlation dimensions can be reliably evaluated by digitizing the EEG signal.

6.2 Applications of EEG

Different workgroups had experimental studies on EEG [45]. These studies were about with analysis of EEG taken from patient, sleeping or awake subjects. Most commonly correlation dimension has been used as classification parameter. Exhibiting

chaotic structure of EEG signals and the dynamical characteristics of brain have been the source of these studies.

6.2.1 Classification of Pathological EEG Signals

When the EEG signals are observed at pathological conditions, it has been seen that there exist an explosion of oscillations during the period of seizure and this different pattern causes the correlation dimension to be found different from healthy condition. The studies are generally focused on Epilepsy and Schizophrenic patients. In the study of Saermark [45] the correlation dimension has been measured as dropping from 11 to 7, after epileptic attack started. At another study performed by Babloyantz [45], D_2 has been found in interval of 3,7-5,4 for the case of Creutzfeld-Jakob Disease.

The EEG gives information about the synchronization among the stages of neural stages. Low amplitude EEG reflects relatively desynchronized states, whereas high amplitude waves are the indicator of synchrony between neural masses [45]. During the normal states of the brain activity, dimension of chaotic attractors decreases as the amplitude of the waves increases. But for the pathological conditions one can not claim the same situation. Creutzfeld-Jakob disease is characterized by higher amplitude and higher dimension than epileptic seizure. To compare the dimensions of healthy and patient subjects can be said more feasible to make discrimination [45].

In the EEG, another parameter which changes the dimension is stationarity. When a nonstationary EEG rithym passes through stationary state, periodicity of the signal will increase and the stationary characteristic will cause to decrease the complexity of signal. Hence, one can say that the dimension of the signal observed in stationary time interval is smaller than nonstationary time interval.

6.2.2 Observations on Dynamics of Olfactory System

In [47], a sample model has been constructed and developed to describe the neural dynamics responsible for odor recognition and discrimination. The olfactory system has been chosen to observe the neural dynamics since it has the simplest structure among other systems such as auditory and visual systems. Therefore the basic phenomenon has been olfaction and the background activity has been spontaneous EEG. Freeman made experiments on rabbits and according to the neurophysiologic

results, from measurements of receptor unit activity, the electro olfactogram shows that, receptor cells which are sensitive to particular odorants are clustered nonuniformly in density in the mucosa and their patterns of olfaction differ for differing odorants. In order to observe the self characteristics of olfaction, the bulb and prepyriform cortex which are the parts of olfactory system have been isolated from other neuronal parts of brain except circulation and it has been seen that these structures are "silent" except when they are electrically or chemically stimulated and after stimuli the state falls "silent" again. These stimuli are considered as perturbations and according to these results the state of dynamic structure is said to be stable if the system returns to old state after perturbation. If the spontaneous activity is steady and nonoscillating, the system is said to be at equilibrium.

But when the bulb and the prepyriform cortex are connected to each other, they can not stay at equilibrium even if they are not connected to the rest of the brain. A bifurcation takes place when system undergoes a major state transition in its dynamics, for example transition from sleep to waking or from normal to seizure activity corresponds to bifurcation. Then it can be said that the control of the system dynamics is shifted from point attractor to chaotic attractor.

The Haussdorf dimension of the background EEG is observed between 4 and 7. This indicates that the background activity reflects not noise but chaos and it has relatively smaller degrees of freedom than noise. The EEG olfactory system has been modelled by nonlinear ordinary differential equations and these equations had been used to separately model the olfactory bulb (OB), the anterior olfactory nucleus (AON) and prepyriform cortex (PC).

When a small change in gains between bulb and AON is entered then a degenerative state with Haussdorf dimension 2, which reflects electrically induced olfactory epileptic seizure and it corresponds to transition from chaotic attractor to point attractor. The chaotic attractor of the seizure state is a two torus whereas the background activity is hyper chaotic and has a much higher dimension.

The smallest dimension is observed at the seizure state and the dimension increases at inhalation, exhalation and motivation, waking rest, deep anesthesia states respectively.

The role of chaos in odor recognition can be observed during inhalation and exhalation. The inhalation for different odor can be assumed as an input to the system and this

causes the transition from disordered state to ordered state so that the dimension of the signal decreases. Under this condition chaos can be considered in “off” state.

If the odor is novel and the system does not already have a global activity pattern corresponding to the odor, then instead of producing one of its previously learned activity patterns, the system falls into a high level chaotic state rather than into the basin for the background odor. This enables the system to conserve the previously learned activity patterns.

6.2.3 Classifications of EEG Using Olfactory System

Freeman designed a dynamical model of olfactory system which is the simplest part of the cerebral cortex, to simulate the chaotic activity. The microscopic characteristics of this activity, when it is observed at a level of single neurons, are shown by spontaneous pulse trains. The macroscopic characteristics are aperiodic and unpredictable oscillations with amplitude histograms near to Gaussian, autocorrelation functions approaching to zero rapidly.

The aim of building a model is to construct nonlinear ordinary differential equations under the known constraints of the anatomy and physiology and to find aperiodic solutions that simulate the statistics, spectra and visually displayed patterns of EEGs [46].

The potential difference between two points in the brain is the sum of contributions from many populations depending on their geometries and their distances from the recording sites. The individual neurons can generate chaotic activity with the time scales in milliseconds and distances in microns. These remarks state that chaotic dynamics can be expected at several hierarchical levels of brain function and it can be distinguished by the time and distance scales [46]. EEG can be considered as the composition of macroscopic variables. EEG activities are generated by interconnected neurons over broad areas of cortex. The Olfactory system is the best known part of the cortex.

While investigating an unknown dynamical system, first critical point is to consider the model as an autonomous. So that the system will be closed to external inputs and intrinsic construction can be discovered. The second point is to make the system stationary. That is, it has constant statistical properties and there are no internal changes from one region to another.

Some of behavioural states of the brain, such as awake, sleeping, resting and aroused are major determinants of cortical properties through neuromodulators. The transition between these states is often abrupt but any state may last in seconds to minutes and this is enough to assume stationarity [46].

The sample model olfactory system has three main parts. These are the olfactory bulb (OB) which receives input from axons of sensory receptors and transmits in parallel to the anterior olfactory nucleus (AON) and preyriform cortex (PC). The AON feeds back to the OB, and the PC feeds back to both AON and OB. The PC is main output pathway to the limbic and motor systems. When the three parts are separated from each other and receive no external stimulation, they go to rest and stay there. This rest state manifests a point attractor. Under certain condition of input, each part can enter into a stable oscillation at its characteristic frequency. When it is perturbed by electrical or chemical stimuli, each part returns to its oscillations that manifests a limit cycle attractor. This biosystem can not maintain a chaotic activity unless three parts are interactively connected. If the OB does not receive sensory input, it maintains steady state aperiodic activity called basal or spontaneous. After an electrical stimulus is applied, it will return to spontaneous pattern.

This olfactory model that simulates the spontaneous EEG activity suggests that there are three main conditions that lead to chaos. First, the characteristic frequencies of the three parts are incommensurate. The second, there are delay in the long feedback paths and the long delays introduce low frequencies in delta and theta ranges, and it makes a feed back of output with out of phase with input. Third, the feedback is inhibitory on some paths which causes to occur negative Lyapunov exponents and excitatory on other paths which causes to occur one or more positive Lyapunov exponents. Hence the system can not go to equilibrium or limit cycle on a torus, it maintains robust aperiodic activity. The "t*wist and fold*" flow of activity within the system suggests a basic similarity between dynamics of the olfactory system and some class of models like Lorenz, Rössler and Chua attractors.

Epilepsy Notes: There are various types of epilepsy with different degrees of complexity but the general characteristics are,

i) an abrupt and global onset of broad spectrum, high amplitude, aperiodic activity,
ii) loss of responsiveness to stimuli
iii) abrupt and global termination of the epileptic activity.

The loss of behaviour and the lack of responsiveness to stimuli indicate that the cortex has entered into an autonomous state during an epileptic seizure [46]. Within a few seconds a spike and wave appears repeating at a rate of nearly 3 Hz which indicates that the equilibrium is unstable. The point is repellor and not an attractor. These spikes seem like periodic but when it is plotted versus PC and OB, then it is seen that trajectory does not close itself in a limit cycle or chaotic attractor of epileptic form.

When the results of study on epilepsy are associated with dimension analysis it is observed that the attractor of spontaneous chaotic state is embedded in 6 dimensions. The dynamics of the olfactory system appears to be less complex during the seizure than the spontaneous state. The measurements were made for 5 sec with 5000 data points. The result of correlation dimensions of epileptic EEG for biological model is 2.52 and 3.76 for simulated model. The D_2 result of simulated spontaneous EEG is 5.92, biological spontaneous EEG is 5.46. The D_2 results taken from inhalation during 5 sec records of simulated EEG is 3.77, for experimental EEG the results varies between 4.10 and 4.88. These findings suggest that the seizure state is less complex than the spontaneous state and the burst state (inhalation) is also less complex than spontaneous activity [46].

6.2.4 EEG Studies about classifications of Sleep Stages

EEG signals are favourable for pattern classifications on sleep since there are different sleep stages which have individually different complexities. The sleep experiments require long EEG recordings therefore D_2 is also suitable for using as classification parameter of different sleep stages.

In a study of Babloyantz [45], the sleep stage 2 and stage 4 has been observed and the correlation dimensions were found as $D_2 = 5.03 \pm 0.07$ and $D_2 = 5.05 \pm 0.1$ for stage 2 and $D_2 = 4.08 \pm 0.05$ and $D_2 = 4.37 \pm 0.1$ for stage 4. According to the results, it can be said that complexity of stage 4 is relatively smaller than stage 2 sleep.

In the study of Başar and Röschke, the dimensionality of "Slow Wave Sleep" (SWS) were investigated where the EEG signals have been taken from cats' hippocampus and it was observed that when cat goes to waking state from SWS state, frequency spectrum shifts from deltha activity to theta activity and the correlation dimension increases from 4 to 5. This transition corresponds to a bifurcation, which describes different states of Central Nervous System (CNS) [45].

6.3 Experimental Application on Spontaneous EEG Data

In this thesis, EEG signals taken from 9 healthy subjects (5 male and 4 female) were analyzed. The records have been taken from eyes opened and eyes closed condition and the correlation dimensions of these two states were evaluated. The record time was 200 sec (100 sec for eyes opened, 100 sec for eyes closed) so that trajectory was sufficient to constitute the phase space. This enables the correlation dimension to be evaluated without an underestimation. The measurements have been taken from scalps and used standard 10-20 measurement system given in chapter 4.2.2. In the tables 6.2, 6.3,6.4 and 6.5, there are symbols "ga" and "gk" which represent *eyes opened* and *eyes closed* respectively.

6.3.1 Application Results

The result has been illustrated from eight electrodes F3, F4, C3, C4, P3, P4, O1 and O2 where the capital F represents frontal lobe, C central lobe, P parietal lobe and Occipital lobe. There are four tables representing eyes opened, eyes closed, eyes opened alpha band filtered and eyes closed alpha band filtered dimensions of the signals. Inside the tables, there are some empty boxes since there are distortions at some signals due to exceeding voltage range and cause unexpected D_2 values. The results for eyes opened and closed are illustrated in table 6.2 and 6.4 respectively. In Fig 6.1 and 6.2, the phase space reconstructions of 5 sec trajectories from eyes opened and closed signals and in Fig 6.3 and 6.4 the phase space reconstructions of alpha band filtered eyes opened and closed signals have been shown, respectively. It is clear that when the alpha band has been filtered, the structure of attractor has become more visible especially for closed eyes condition.

Table 6.2 Correlation dimensions for eyes opened, sbxx means subject number

Ga	Sb01	Sb02	Sb03	Sb04	Sb05	Sb06	Sb07	Sb09	Sb10	Mean	St.dev
F3	6,30	7,01	6,60	6,64	7,52	6,95	6,72	6,12	6,12	6,68	0,44
F4	5,35	6,89	6,52	7,28	7,60	7,10	7,13	5,66	5,66	6,58	0,82
C3	6,64	7,21	6,78	7,02	7,11	6,62	7,60	6,52	6,52	6,91	0,35
C4	6,29	6,70	6,64	6,42	7,14	6,44	6,68	6,58	6,58	6,53	0,35
P3	6,59	7,10	6,38	6,08	7,55	7,33	6,79	5,71	5,71	6,65	0,60
P4		6,85	5,60	5,94	6,60	6,60		5,72	5,72	6,17	0,50
O1	6,62	7,28	5,80	5,98	7,62	6,89	7,59	6,87	6,87	6,73	0,70
O2	6,39	7,32	6,05	6,02	7,23	6,95	7,11	7,42	7,42	6,75	0,57

These results have been compared with alpha band (8-16 Hz) filtered EEG signals' correlation dimensions which are given in table 6.3 and 6.5 for eyes opened and closed, respectively.

Table 6.3 Correlation dimensions for eyes opened alpha band filtered.

Ga	Sb01	Sb02	Sb03	Sb04	Sb05	Sb06	Sb07	Sb09	Sb10	Mean	St.dev
F3	5,80	5,46	5,46	5,86	5,71	5,50	5,73	5,51	5,31	5,59	0,19
F4	5,33	5,40	5,61	5,99	5,60	5,48	5,53	5,48	5,68	5,57	0,19
C3	6,05	5,82	5,60	5,57	5,37	5,54	5,35	5,60	5,30	5,58	0,24
C4	5,80	5,31	5,66	5,25	5,57	5,55	5,50	5,85	5,49	5,55	0,20
P3	5,70	5,60	5,01	5,49	5,71	5,55	6,05	5,95	5,79	5,65	0,30
P4		5,57	5,66	5,00	5,58	5,75		5,50	5,39	5,49	0,25
O1	5,44	5,56	5,24	5,36	5,91	6,08	5,76	5,80	5,95	5,68	0,29
O2	5,40	5,69	5,69	5,25	5,48	6,83	6,12	5,56	5,96	5,66	0,28

Table 6.4 Correlation dimensions for eyes closed condition.

Gk	Sb01	Sb02	Sb03	Sb04	Sb05	Sb06	Sb07	Sb09	Sb10	Mean	St.dev
F3	7,02	6,35	7,22	7,09	7,75	6,83	7,02	7,37	6,19	6,98	0,48
F4	6,60	6,50	5,4	6,84	6,65	6,63	7,21	5,15	5,82	6,31	0,69
C3	7,25	6,45	6,61	7,59	4,58	6,83	7,95	7,44	6,40	6,79	0,99
C4	6,80	6,30	7,34	7,08	6,36	7,25	7,14	7,72	5,90	6,88	0,58
P3	6,61	7,18	6,65	7,10	7,42	7,18	7,14	6,86	5,80	6,88	0,48
P4		6,55	5,26	6,62	6,38	7,10		7,92	5,04	6,41	1,00
O1	6,82	6,75	6,90	5,93	7,31	6,80	6,95	6,82	5,82	6,68	0,48
O2	6,53	6,82	6,77	6,82	6,71	6,65	6,79	7,43	5,85	6,71	0,41

Table 6.5 Correlation dimensions for eyes closed alpha band filtered signals.

Gk	Sb01	Sb02	Sb03	Sb04	Sb05	Sb06	Sb07	Sb09	Sb10	Mean	St.dev
F3	5,30	5,59	5,66	5,51	5,59	5,63	5,75	5,72	5,57	5,59	0,13
F4	5,32	6,01	4,30	5,62	5,70	5,42	5,74	5,16	5,09	5,37	0,50
C3	5,15	5,71	5,69	5,57	4,52	5,33	5,80	5,46	5,49	5,41	0,39
C4	5,46	5,98	6,12	5,35	5,49	5,41	5,95	5,65	5,23	5,63	0,32
P3	5,60	5,58	5,56	5,42	5,67	5,65	5,42	5,48	5,22	5,51	0,14
P4		5,68	4,35	5,48	5,55	5,60		5,80	5,25	5,39	0,49
O1	5,68	5,60	5,39	5,52	5,86	5,51	5,44	5,73	5,10	5,54	0,22
O2	5,61	6,40	5,22	5,25	5,40	5,22	5,48	5,49	5,11	5,46	0,39

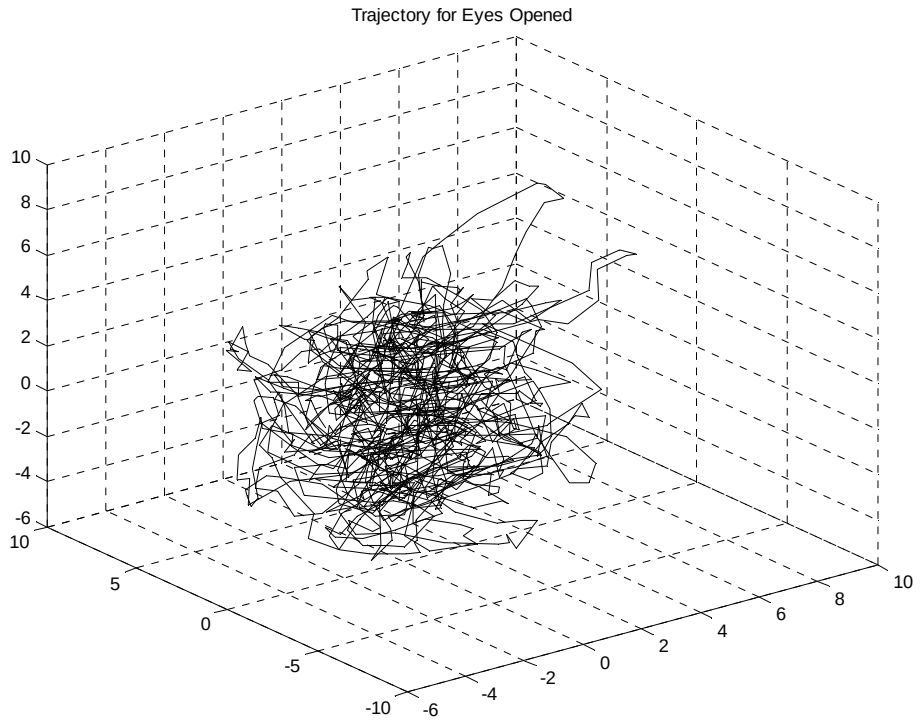

Figure 6.1 Trajectory of 5 sec. data taken from electrode O1 of mean of 9 subjects for eyes opened condition with time lag $T = 30$

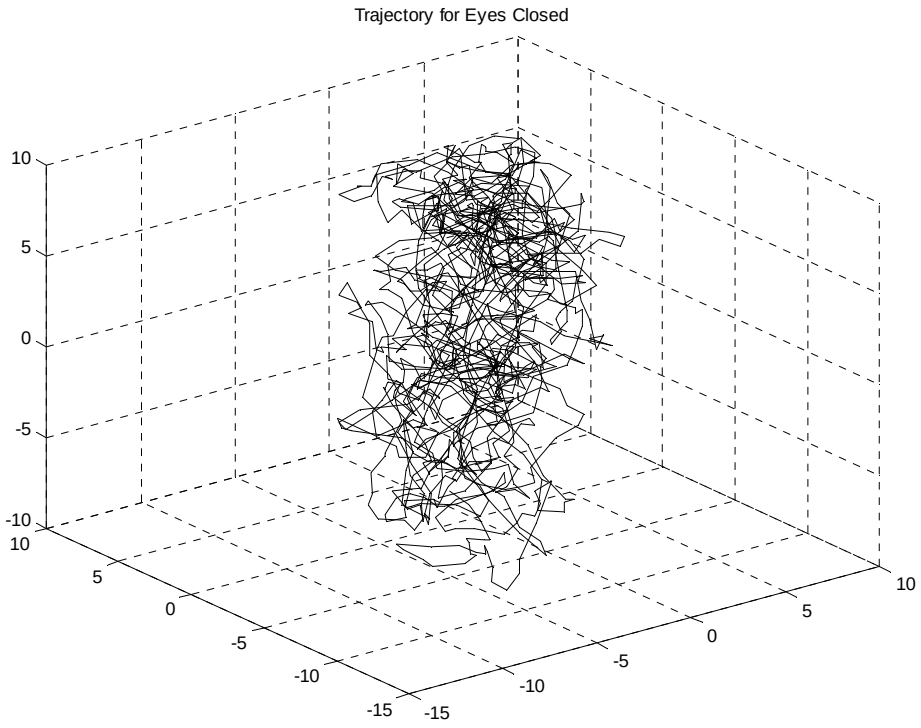

Figure 6.2 Trajectory of 5 sec. data taken from electrode O1 of mean of 9 subjects for eyes closed condition with time lag $T = 30$

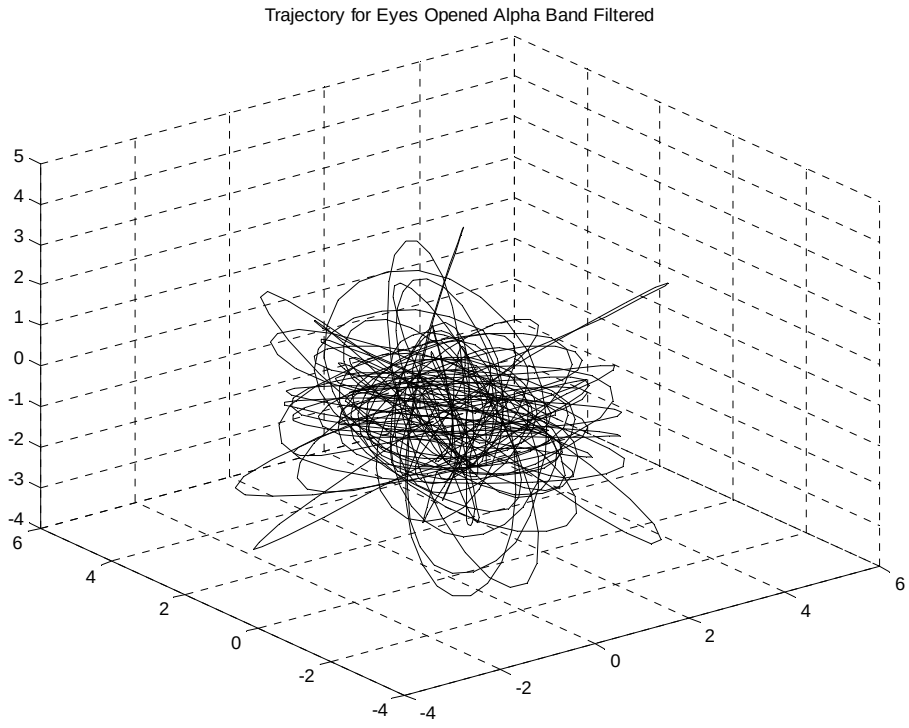

Figure 6.3 The trajectory of 5 second eyes opened alpha band filtered response, $T = 30$.

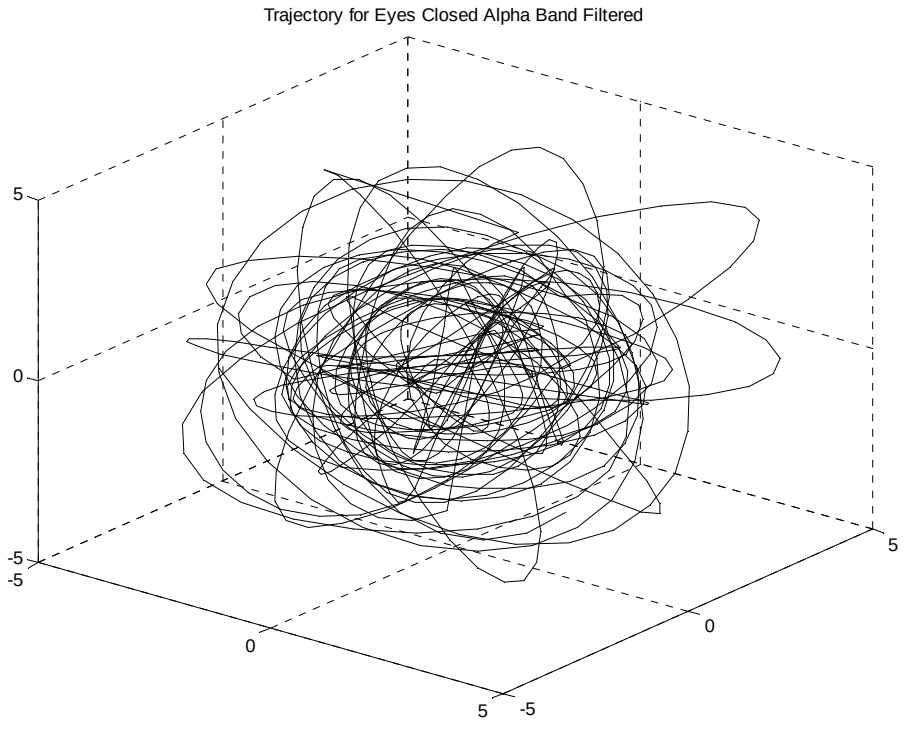

Figure 6.4 The trajectory of 5 second eyes closed alpha band filtered response, $T = 30$.

6.3.2 Comparison of the results

In these observations there are two critical situations that are used while classifying the dimensions. The first is conditions of opened and closed eyes and the second is correlation dimensions of alpha band filtered signals. To close the eyes causes an explosion of oscillations in parietal and occipital lobes of the brain since these lobes are sensitive to visual activities. Therefore, investigating D_2 for the eyes opened and closed conditions is an important comparison study for EEG signals and to investigate the dimensions for parietal and occipital lobes are more meaningful. Since the frequency spectrum is more concentrated near 10 Hz when the eyes have been closed, to investigate the dimension variation for alpha band filtered signals is expected to give sharper variation that is why correlation dimensions of specific frequency band are evaluated. The mean values have been used as comparison because comparing individual dimension results do not give a clue for classification. Indeed, it is seen that for electrode O1, an average D_2=6,73 while eyes closed case D_2=6,68 and similarly for electrode O2, D_2=6,75 whereas D_2=6,71 when eyes are closed. One can say that the results are near to each other but when the alpha band results are compared, it is observed that for O1, D_2=5,68 for eyes opened and D_2=5,54 for eyes closed and similarly for O2, D_2=5,66 for eyes opened and D_2=5,46 for eyes closed. It is quite clear that there is a sharper difference among the signals eyes opened and eyes closed.

This result is not surprising because, when eyes closed the stationarity of the signal increase, it becomes more periodic and it causes the complexity to decrease and these results reflect to the dimensions.

CHAPTER 7

CONCLUSION

In this thesis, the necessary methods for analyzing observed chaotic data have been given. Since the chaotic signals have broadband power spectrum, they can not be analyzed only by finding resonant frequencies where they are used as invariants of linear systems. Classification has been realized by computing the correlation dimensions and wavelet entropies of EEG signals which exhibit high dimensional chaotic behavior.

The phase space reconstruction is the critical part of the analysis and there exist several methods to determine embedding dimension and the time delay which are the parameters of the reconstruction and they affect the results of invariant measures directly.

Due to any noise contamination causes to estimate divergent invariant measures, the noise reduction from chaotic signals has a vital importance. Since there is no fore-knowledge about the signal, the blind source separation techniques have been used. The common methods are principle component analysis (PCA) and singular value decomposition (SVD). The noise is eliminated by detecting the size of the subspace related with noise free signal. SVD based method has been proposed to estimate the rank of the diagonal singular value submatrix which is related to noise free signal. After the threshold has been determined, the lower singular values have been truncated and effect of noise due to orthogonality has been eliminated on the remaining singular values. Furthermore, as a result of proposed rank estimation method it has been observed that the proper value of the embedding order of noisy matrix is $2D_2+1$ which confirms the condition given in embedding theorem. When these procedures have been completed, time series is ready to be processed for classification.

Wavelet transform is quite preferable to observe time-frequency distribution of signals. The energy of a signal distributed in time frequency plane is represented by scalogram so that the stationarity of signal can be determined. The classification using wavelet analysis has been performed by wavelet ridges and wavelet entropy where the wavelet entropy represents the energy density in a certain frequency band and it is based on relative wavelet energies. Wavelet ridge is the wavelet transform of a signal

evaluated at the ridge points where the ridge points are determined by phase based or modulus based methods. The results of simple method which is one of the modulus based methods have been represented for EEG signals for open and closed eyes conditions. It is observed that when the eyes are closed the ridge patterns are gathered near to 10 Hz, (i.e., alpha band) whereas there is more dispersed patterns for opened eyes case. The variation of pattern distribution is a good classification tool for experimental signals. Another application, wavelet entropy gives the energy density distribution for frequency localized signals after the result of wavelet decomposition. It is observed that the relative wavelet energy of alpha band has a significant increase when the eyes are closed. The entropy is lower for eyes closed than eyes opened. This shows that when the eyes are closed, the complexity of signal decreases, alpha band becomes the dominant part of the total signal whereas the signal is more distributed among the other frequency bands for eyes opened.

Another classification parameter used in the thesis has been correlation dimension which is most widely used classification parameter in EEG research. Dimensionality of brain has been evaluated for all channels and two cases of eyes closed and opened. The results indicate that when eyes are closed, the correlation dimension of signals observed from parietal and occipital regions decrease relatively. When the alpha band is individually observed, clearer attractor structure is seen. But this complexity decrease is not as much as other state transitions such as transition to pathological state from healthy state or transition between different sleep states because the characteristic of spontaneous EEG for opened and closed eyes does not significantly change like the other conditions. The correlation dimension difference gives a clue about state transitions (i.e., bifurcations) in the brain and at which conditions complexity changes. To make chaos analysis for event related potentials(ERP) or evoked potentials (EP) which are stimulus based observations, is not possible because in order to determine correlation dimension properly, long time recordings are needed so that the trajectory will suffice to observe unchanged invariant measure by the time. Therefore, chaos analysis on EEG signals in the thesis has been made for spontaneous activity that contains the experiments of opened and closed eye conditions and for pathological and sleeps EEG signals.

The remaining applications on analysis of experimental chaotic time series is to develop models to increase the prediction time and to study on specified methods of discriminating interfered chaotic time series by using support vector machines.

REFERENCES

[1] H.I.Abarbanel, "*Analysis of Observed Chaotic Data*", Springer,1996.

[2] http://www-chaos.engr.utk.edu/res/ctsa.html

[3] S.H. Strogatz, "*Nonlinear Dynamics and Chaos*", Addison-Wesley, 1994.

[4] U. Gürsoy, "*Kaotik Sistemlerde Lyapunov Karakteristik Üstelleri*", Graduate Thesis-İTÜ,1996.

[5] F. Takens, "*Detecting Strange Attractors in Turbulence*", Dynamical systems and Turbulence, Springer, p.366, 1981.

[6] J.C. Sprott, "*Chaos and Time Series Analysis*", Oxford University Press,2003.

[7] http://www.iqnet.cz/dostal/CHA2.htm

[8] S. Rasband, "*Chaotic Dynamics of Nonlinear Systems*", John Wiley Sons, 1989.

[9] H. Kantz & T. Schreiber, "*Nonlinear Time Series Analysis*" Cambridge University Press 1997.

[10] H.I. Abarbanel, R. Brown,J.J. Sidorowich, Lev. Sh. Tsimring, "The analysis of Observed Chaotic Data in Physical Systems", *Rev. Mod. Phys*, Vol.65,No:3, October 1993.

[11] M. Sano & Y. Sawada,"Measurement of The Lyapunov Spectrum from a Chaotic Time Series", *Physical Review Letters*, Vol.55, No:10, pp.1082-1085.

[12] J.A. Vastano, E. J. Kostelich, "Comparison of Algorithms for Determining Lyapunov Exponents from Experimental Data ", *Entropies and Dimensions*, Springer 1986, pp.100-107.

[13] K. Geist, U. Parlitz & W. Lauterborn, "Comparison of Different Methods for Computing Lyapunov Exponents", *Progress of Theoretical Physics*, Vol. 83, No.5, May 1990.

[14] J. Eckmann & S. Kamphorst, D. Ruelle, S. Ciliberto, "Lyapunov Exponents from Time Series", *Physical Review A*, Vol.34, No.6, 1986.

[15] A. Wolf, J. B. Swift, H. L. Swinney & J. A. Vastano, "Determiming Lyapunov Exponents from Time Series ", *Physica 16D*, pp.285-317, 1985.

[16] U. Parlitz, "Nonlinear Time Series Analysis", *Georg-August Universitat*.1995

[17] R. Esteller, G. Vachtsevanos, J. Echauz, B. Litt, "A Comparison of Fractal Dimension Algorithms Using Synthetic and Experimental Data", *IEEE*

Transactions on Circuits & Systems, vol 48(2), February 2001, pg 177-183.

[18] W. Liebert, H.G.Schuster, "Proper Choice of The Time Delay for The Analysis of Chaotic Time Series", *Physics Letters A*, Vol:142, pp.107-111, December 1989.

[19] T.W. Frison, H.I. Abarbanel, "Identification and quantification of Nonstationary Chaotic Behaviour", *IEEE, Acoustics, Speech & Signal Processing*, 1997, Vol.3,pp.2393-2396.

[20] B. De Moor, "*Mathematical Concepts and Techniques for Modelling of Static and Dynamic Systems*", Phd. Thesis, Katholieke Universiteit, 1988.

[21] R. O. Duda, P.E. Hart, D.G. Stork, "*Pattern Classification*", John Wiley & Sons, 2001.

[22] B. De Moor, **"**The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices", *IEEE Transactions on Signal Processing*, Vol. 41, No.9, pp.2826-2838, September 1993.

[23] B. Pilgram**,** W. Schappacher, **"**Estimation of the Dominant Singular Values for SVD Based Noise Reduction Methods". *Int. Journal of Bifurcation& Chaos* Vol.8, No.3, pp.571-580, 1998

[24] E. Kostelich, T. Schreiber, "Noise Reduction in Chaotic Time Series Data: A Survey of Common Methods", *Physical Review E*, Vol. 48, No.3, pp. 1752-1763 September 1993.

[25] "Estimating the Number of Sinusoids in Additive Noise" , *IEEE Transactions Acoustics, Speech and Signal Processing*, Vol. 36, No.12, December 1988.

[26] K. Shin, K. Hammond, P. R. White, "Iterative SVD Method for Noise Reduction of Low Dimensional Chaotic Time Series", *Mechanical Systems and Signal Processing*, Vol.13, No.1, 1999.

[27] S. A. Billings, K. L. Lee, "A Smoothing Algorithm for Nonlinear Time Series", *International Journal of Bifurcation & Chaos*, Vol.14, No.3, pp. 1037-1051, 2004

[28] http://biology.about.com/library/organs/brain/blbrain.htm

[29] http://www.sbg.ac.at/psy/people/schabus/tutorials/eeg%20intro/eeg_intro.html

[30] http://www.u.arizona.edu/~jallen/EEG.HTM

[31] http://www.altered-states.co.nz/cgi-local/reload.cgi?// www.google.com.tr/search?q=EEG+eyes+closed+open+concepts&hl=tr&lr=&ie

=UTF-8&oe=UTF-8&start=10&sa=N^/brainmaster/freqs.htm

[32] http://www.brainmaster.com/generalinfo/electrodeuse/eegbands/1020/1020.html

[33] http://www.qeeg.com/1020System.html

[34] O. Rioul, M. Vetterli, "Wavelets and Signal Processing", *IEEE Magazine*, October 1991.

[35] P. Flandrin, P. Gonçalves, "From Wavelets to Time Scale Energy Distributions", *Recent Advances in Wavelet Analysis*, Academic Press, pp.309-334, 1994.

[36] L. Cohen, "Time Frequency Distributions- A Review", *Proceedings of IEEE*, Vol.77, No.7, July 1989.

[37] O. Rosso, S. Blanco, Juliana Yordanova, V. Kolev, A. Figliola, M. Schürmann, E. Başar, "Wavelet Entropy: A New Tool for Analysis of Short Duration Brain Electrical Signals", *Journal of Neuroscience Methods*, No.105, pp.65-75, 2001.

[38] R. Q. Quiroga, O.W. Sakowitz, E. Başar, M. Schürman, "Wavelet transform in The analysis of The Frequency Composition of Evoked Potentials", *Brain Research Protocols*, No.8, pp.16-24, 2001.

[39] O. W. Sarkowitz, R. Q. Quiroga, M. Schürmann, E. Başar, "Bisensory Stimulation Increases Gamma Responses Over Multiple Cortical Regions", *Cognitive Brain Research* (11) pp.267-279, 2001.

[40] Y. U. Khan, J. Gotman, "Wavelet Based Automatic Seizure Detection Intracerebral Electroencephalogram", *Clinical Neurophysiology* 114 (2003), pp. 898-908.

[41] M. I. Todorovska, "*Estimation of Instantaneous Frequency Signals Using The Continuous Wavelet Transform*", University of Southern California, December 2001.

[42] N. Özkurt, F. A. Savacı, "Determination of Wavelet Ridges of Nonstationary Signals by Singular Value Decomposition Method", Accepted in *IEEE Circuits and Systems*.

[43] N. Özkurt, F. A. Savacı, "The Reconstruction of Nonstationary Signals in Wavelet Domain Based on Singular Value Decomposition" *NDES, 2003* pp.181-185.

[44] N. Delprat, B. Escudie, P. Guillemain, R. K. Martiner, P. Tchamitchian, B. Torresani, "Asymptotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequencies", *IEEE Transactions on Information Theory*, Vol. 38

No.2, March 1992.

[45] E. Başar, "*Chaos in Brain Function*" , Springer, 1990.

[46] W. J. Freeman, "Tutorial on NeuroBiology: From Single Neuron to Brain Chaos", *International Journal of Bifurcation and Chaos*, Vol.2 No.3, pp.451-482, 1992.

[47] C. A. Skarda, W. J. Freeman, "How Brain Make Chaos in Order to Make Sense of The World", *Behavioural and Brain sciences*, Cambridge University Press, 1987.

[48] E. Başar, "Macrodynamics of Electrical Activity in The Whole Brain" *International Journal of Bifurcation and Chaos*, Vol. 14, No. 2 (2004) 363-381.

[49] R.G. Gallager, "*Information Theory & Reliable Communication*", John Wiley & Sons, 1968.

[50] P. Bryant, R. Brown, H.I. Abarbanel, "Lyapunov Exponents from Observed Time Series", *Physical Review Letters*, Vol.65 No:13, pp.1523-1526, 1991

[51] M. E. Çek, F. A. Savacı, "Rank Kestirim Yöntemi Kullanarak SVD Tabanlı Gürültü Filtreleme", *Sinyal İşleme Uygulama Konferansı* (SIU), s.320-324, 2004

[52] Wen-Xian Yang, P.W. Tse, "Development of an Advanced Noise Reduction Method for Vibration Analysis Based on Singular Value Decomposition", *NDT&E International*, No.36, pp.419-432, 2003.

[53] E. Yazgan, M. Korürek, "*Tıp Elektroniği*", İTÜ Matbaası, 1996.

[54] C. K. Chui, **"***Wavelet Analysis and Its Applications*", Academic Press, 1992.

[55] S. Mallat, "*A wavelet Tour of Signal Processing*", Academic Press, 1998.

[56] H. Jing, & M. Takigawa, "Low sampling Rate Induces High Correlation Dimension on Electroencephalograms From Healthy Subjects", *Psychiatry and Clinical Neurosciences*, No.54, pp.407-412, 2000.