Nonlinear, Multidimensional Transformations and their Applications to Signal Processing

Thesis Submitted to the Department of Studies in Mathematics
of University of Mysore in fulfillment of the requirements for the Degree of

Doctor of Philosophy

by

Sajini Anand P S

under the guidance of

Dr. Prabhakar G Vaidya

School of Natural Sciences and Engineering
National Institute of Advanced Studies (NIAS)
Indian Institute of Science (IISc) Campus
Bangalore 560012, India

JULY 2014

NATIONAL INSTITUTE OF ADVANCED STUDIES
Indian Institute of Science campus, Bangalore-560 012, India

DECLARATION

I, Sajini Anand P S, hereby certify that the thesis entitled, **'Nonlinear, Multidimensional Transformations and their Applications to Signal Processing'** is the result of research work done by me under the supervision and guidance of Dr. Prabhakar G. Vaidya at School of Natural Sciences and Engineering, National Institute of Advanced Studies (NIAS), Bangalore. I am submitting this thesis for possible award of Doctor of Philosophy (Ph. D) degree in Mathematics of the University of Mysore.

I further certify that this thesis has not been submitted by me for award of any other degree/diploma of this or any other university.

(Signature of the Doctoral candidate)

Sajini Anand P S

Date : 10/July/2014

NATIONAL INSTITUTE OF ADVANCED STUDIES

(Signature of Guide)

Dr. Prabhakar G Vaidya
NIAS, Bangalore

Date : 28/June/2014

National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore - 560 012

Counter signed by

Dr.V S Ramamurthy
Director, NIAS,
Bangalore

Date : 10.July.2014

Prof. V. S. RAMAMURTHY
DIRECTOR
National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore - 560 012

Telephone: +91 80-2218 5000 Fax: +91 80-2218 5028 Email: niasoff@nias.iisc.ernet.in

NATIONAL INSTITUTE OF ADVANCED STUDIES
Indian Institute of Science campus, Bangalore-560 012, India

DR. PRABHAKAR G. VAIDYA E-mail: pgvaidya@gmail.com

Professor, School of Natural Sciences and Engineering

CERTIFICATE

The research embodied in this thesis entitled "Nonlinear Multidimensional Transformations and their Applications to Signal Processing" was conducted by Sajini Anand P S in the School of Natural Sciences and Engineering, National Institute of Advanced Studies, Bangalore under the supervision and guidance of Dr. Prabhakar G. Vaidya. It is hereby certified that the thesis submitted is a bonade record of research done by the candidate in fulfillment of the requirements for Ph.D. degree, and that the thesis has not previously formed the basis for the award to the candidate of any degree, diploma or any other similar title of any other university or society.

(Signature of Guide)

Dr. Prabhakar G Vaidya
NIAS, Bangalore

Date : 29th June 2014

National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore - 560 012

List of publications

International Peer-reviewed Journal Publications

1. Prabhakar G Vaidya, Sajini Anand P S and Nithin Nagaraj, *A Nonlinear Generalization of Singular Value Decomposition and its Applications to Mathematical Modeling and Chaotic Cryptanalysis* (Springer), Acta applicandae mathematicae Vol. 112.2 (2010) pp. 205–221. ISSN: 0167-8019 (print version), ISSN: 1572-9036 (electronic version).

2. Prabhakar G. Vaidya and Sajini Anand P S, *Estimating the Dimension of a Manifold and finding local charts on it by using Nonlinear Singular Value Decomposition*, Topology Proceedings, Vol. 43 (2014) pp. 1–15. ISSN: 0146-4124 (Print version), ISSN: 2331-1290 (electronic version).

National and International Conference Papers

1. Prabhakar G Vaidya and Sajini Anand P S, *Estimating the Dimension of a Manifold using Nonlinear Singular Value Decomposition*, 27^{th} Summer Conference on Topology and its Applications, July 25–28, 2012, Minnesota State University, Mankato, Minnesota, USA.

2. Prabhakar G Vaidya and Sajini Anand P S, *Application of Topology for ECG Analysis*, 27^{th} Summer Conference on Topology and its Applications, July 25–28, 2012, Minnesota State University, Mankato, Minnesota, USA.

3. Sajini Anand P S and Narayana Dutta, *Topological structure of Epilepsy EEG signals*, International Conference on Networks in Biology, Social Science and Engineering, July 12–14, 2012, IISc, Bangalore.

4. Sajini Anand P S, Nithin Nagaraj and Prabhakar G Vaidya, *Imprecise Synchronization– A Study*, International Conference on Nonlinear Dynamical Systems and Turbulence , July 17– 22, 2008 IISc, Mathematics Initiative (IMI), Indian Institute of Science, Bangalore.

5. Prabhakar G Vaidya, Kishor Bhat and Sajini Anand P S, *Laughter signature using a special nonlinear filter*, International conference on Perspectives in Nonlinear Dynamics PNLD 2010, July26-29, IISc, Bangalore.

6. Prabhakar G Vaidya, Nithin Nagaraj and Sajini Anand P S, *Topological Derivatives and other Embeddings for Ocean Floor Tsunami Data*, International Conference on Tsunami and Non-linear Waves, Saha Institute of Nuclear Physics, Salt Lake, Calcutta, India, March 6–10, 2006.

7. Prabhakar G Vaidya, Sajini Anand P S, *Rapid Chaotic Synchronization by Intermittent Driving Signals.*, National Conference in Nonlinear Systems and Dynamics, NCNSD Proceedings (2006) pp. 99–101.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Prof. Prabhakar G. Vaidya who gave me an opportunity to pursue a PhD and guided me towards exciting research questions in the field of nonlinear dynamics and mathematical modeling that forms the basis of this thesis. I am very thankful to Dr. Nithin Nagaraj for his guidance and support throughout the PhD research. I thank Prof. Tim Poston and Prof. Ramachandra for many useful mathematics discussions we had at NIAS.

Thanks are due to my Doctoral Committee members Dr. Amit Apte (TIFR–CAM/ ICTS), Dr. Chadrasekhar jog (ME dept., IISc.) and Prof. Narayan Ramani (NIAS) for their patience and feedback on my thesis work. My gratitude to Prof. B.V Sreekanton, Prof. Sangeetha Menon, Prof. K. Kasturirangan and Prof. V. S. Ramamurthy for their support and encouragement during my PhD programme. I am grateful to NIAS for allowing me to use all the facilities and I thank DST for the financial support for my Ph. D research. I am thankful to IISc and TIFR–CAM for enabling me to avail library books, various courses, workshops, tutorials, conferences and other valuable resources.

I am deeply grateful to Ms. Sandhya for the dealing with all the formalities of University of Mysore. Thanks to Ms. Mariammal, Ms. Hamsa, Ms. Vijaylakshmi, Ms. Stella, Mr. Devaraju, Mr. Sasi and Mr. Aithal for their co-operation. I am grateful to Rajani, Kishor, Meera, Nalini, Asha, Sreeja, Anu, Jafar and Lakshmi for all the personal and professional help during my Ph. D work. I thank Nimisha for all the help with the proof reading of the thesis. Special thanks to my colleagues and all my good friends at NIAS for the lively multidisciplinary discussions we had at NIAS.

I wish to thank my Amma and Achan, Ammachi and Achachan, my siblings Unni and Vishnu, and my extended family for their affection, care and patience. I am grateful to my husband Saneesh and my little daughter Nivedita for the happiness and peace they bring into my life. I dedicate my thesis to my Amma and Achan, who struggled in their life to bring us up and motivated us to excel in studies.

ABSTRACT

Modeling a system based on time series is a complicated problem in general, especially when the time series is nonlinear and chaotic. The goal of the thesis is to introduce a method of prediction and modeling that exploits the property of recurrence in dynamical systems. A time series is said to be recurrent if keeps on visiting a particular neighborhood in the state space. The thesis demonstrates that the inherent redundancy structure of a well known topological technique known as delay embedding can be coupled with recurrence property to develop a new method of prediction. The modeling procedure empirically finds the recurrence neighborhoods from the signal, which are then subdivided into various equivalence classes based on their recurrence timings. A set of affine maps are then derived across these equivalence classes. This gives is a possibility of simplifying the dynamics in terms of affine transformations in small neighborhoods. The delay-embedding (done in a dimension much higher than the inherent dimension of the dynamics) is used as a scaffolding to analyze the global structure of the system. A projection to a lower dimension was followed to take care of the fundamental issues related to high dimensional models that describe a low dimensional dynamics. Local analysis of the system was done in the low dimensional projected space. A topological conjugacy of the recurrence neighborhoods in both the lower and the higher dimensional spaces are demonstrated.

The proposed model uses a nonlinear generalization of a well known linear algebra technique named Singular Value Decomposition (SVD) for data analysis. The method of nonlinear SVD and its uses (i) to determine nonlinearity in a time series and (ii) to empirically arrive at an upper bound for the dimension of a manifold where the data resides are demonstrated. The proposed method of prediction and modeling was used for the analysis of (i) data generated by the Duffing oscillator and (ii) an Electrocardiogram (ECG) record. It is shown that the entire nonlinear structure can be deduced from one or few overlapping neighborhoods for these data. A method of stability analysis by studying the properties of affine maps specific

to the neighborhoods are demonstrated for both these data. The thesis gives a theoretical justification for a well known experimental observation that the heart rate variability– a variability in beat-to-beat intervals of the heart is a necessity for healthy functioning of the heart. The relevance and contribution of the introduced method for biomedical signal processing is justified by using it successfully for analyzing a set of multi-channel physiological data.

CHAPTER 1
INTRODUCTION

Modeling a system based on an observed time series is a difficult problem general as the measurements in addition to being noisy contain only partial information about the system. Key goals of research in various fields ranging from theoretical mathematics to nonlinear system analysis and atmospheric– ocean modeling to biomedical signal processing focus on this issue. The foundation of this thesis is an exploratory research for finding new, effective methods for modeling and prediction of biomedical signals. The main goal of the research reported in the thesis is to propose a new method of modeling and prediction with a dynamical system perspective for a special class of signals that exhibit the property of recurrence– the repetition of patterns in the time series. Further the thesis discusses a few novel analysis techniques that are developed for processing signals generated by nonlinear systems. In what follows, the overview of the thesis– the purpose and the significance of the study with respect to the background literature, key ideas and areas that are explored and objectives of the thesis, new ideas explored in the research, methodology of research, analysis of results and its significance, and the contribution of the thesis to the existing body of knowledge are discussed.

1.1 Overview

Most of the dynamical systems in nature are fundamentally nonlinear. Some of these systems were modeled and studied in a linear paradigm before the developments in nonlinear dynamics. The field of nonlinear dynamics is interdisciplinary in nature as it analyzes the complex behaviors of non–linear systems in various disciplines ranging from theoretical mathematics to biological entities and engineering models. Nonlinear systems exhibit a range of complex phenomena that have given useful insights to the field of time series analysis, signal processing and mathematical modeling of realtime systems [2, 3]. Emergence of new methods for analysis and modeling based on a dynamical systems perspective is one of the significant achieve-

ments of the nonlinear dynamics in the past few decades [4, 5]. These methods have found numerous applications in different disciplines spanning from atmospheric–ocean models to biomedical analysis.

The main area of research the thesis focusses on is the modeling of dynamical systems based on a measured signal or a time series generated by the system. This type of modeling is perceived as a complicated problem especially when the data under study is nonlinear, chaotic and noisy. Most of the physiological signals in the area of biomedical signal processing are considered nonlinear. Depending upon the objective of study, many linear methods are still popular to process these signals. Representing a signal as linear combination of set of orthogonal functions is a widely used methodology in signal processing. Two of them are: (i) Fourier series based methods use a set of prefixed basis functions, and (ii) class of methods based on Singular Value Decomposition (SVD) or Karhunen Loeve Transform that do not have a prefixed set of basis functions but finds the best basis for the signal empirically such that it satisfies the law of parsimony.

SVD and some of its related variations (Independent Component Analysis, Principal Component Analysis and Blind Source Separation) are also widely used to process data [6, 7, 8, 9]. SVD is an established method in linear algebra for matrix decomposition and it has been used for finding the dimension of a linear system as it gives statistically independent set of variables which could span the state–space [10, 11]. Since SVD based methods are inherently designed to identify linear dependence or correlation, they work extremely well when the signals are generated by linear systems. A method based on SVD is proposed by Broomhead *et al.* for detecting underlying nonlinearity from a time series in a qualitative matter [12]. But this method fails to distinguish a chaotic time series from its surrogates– the stochastic counterparts of the data which has the same power spectrum [13, 14]. Additionally, there are many novel algorithms that extend regression methods to estimate of model parameters based on the best nonlinear polynomial fit. The goal of such algorithms are to find the parameters of the appropriate models from the data based on various least square techniques and most of them are found to be successful in a large number of engineering problems that deal with time series

analysis and signal processing [10, 15].

Other than the engineering methods, there are exciting developments in Chaos theory and Topology in the past few decades that are useful to understand the dynamics of nonlinear signals. An important property of dynamical systems was observed by Packard et. al. in 1980 when they could reconstruct a state–space of the system from a time series generated it [16]. This opened up a possibility of associating geometrical and topological structures with an observed time series of a system whose state-space is unknown otherwise [16, 17]. One goal of modeling a system based on time series is to develop equations for the system from the time series [18, 19] for prediction or a description of the dynamics [20, 21]. These type of modeling methods generally fall into two categories: (i) Global methods that find equations valid for the entire statespace [22] and (ii) Atlas methods that develop local charts for small neighborhoods of the statespace [23].

The global embedding of a dynamical system in general requires a dimension higher than the intrinsic dimension of the system. According to Whitney's theorem a d–dimensional manifold can almost always be embedded into the Euclidean space R^N given $N > 2d$ [24]. Later, Takens in 1981 proved a theorem according to which the state space of a system can be reconstructed in a high dimension using time–delayed copies of a measured variable and this method is known as delay embedding [25]. According to this theorem given a time series that belonged to system defined on a d–dimensional manifold, the state–space reconstructed using delay vectors in R^N such that $N \geq 2d + 1$ is an embedding. This theorem implied a one–one similarity between the reconstructed space in the delay embedding and the original state space of the system. Delay embedding of data in a higher dimensional space has became a standard procedure for analyzing experimental times series in nonlinear dynamical system studies since then [5].

In general, for global embedding of a d–dimensional manifold the minimum dimension required for embedding is $2d + 1$. But note that, if the aim is to use Atlas methods to find local charts that cover the statespace, it might be possible to represent the system in just d–dimensions. Since Takens theorem doesn't specify any upper limit on the delay embedding dimension, nonlinear systems are often

embedded into high dimensions for global analysis. Also in practical cases where the dimension of the system and the dynamics that generated the data is unknown, embedding of the data into a dimension much higher than inherent dimension is a usual practice. But there are some fundamental issues that causes spurious instabilities and ambiguities in models when the system is represented in a dimension higher than the required [26]. Other possibilities reported are (i) increase in model dimension that expands the effect of noise (ii) use of higher order polynomials in the model equation that increases the number of coefficients to be estimated (iii) some orbits of the model statespace that goes outside the region containing the time series – thus making the model no longer connected with the object under study [27]. In addition to the issue of stability, there is also a practical reason why one prefers a lower dimensional model. A model in a lower dimensional space is more economical in the sense it requires lesser amount of data compared to its high dimensional counterparts.

Nonlinear chaotic signals exhibit many properties including non-periodicity and non-stationarity. Chaotic signals also mimic random signals despite their deterministic origin. Recurrence plots and Poincare maps are some classic techniques that are used to analyze these types of signals [28, 29, 4]. The behaviour of Recurrence (the property that a typical trajectory of the system keeps on visiting the neighborhood of a particular state in the state space) is a characteristic of nonlinear chaotic systems [30]. Identifying recurrence patterns and analyzing them has become a prominent method for nonlinear data analysis [31, 32] and these methods have found applications in numerous fields as financial time series analysis and physiological signal processing [33, 34, 35, 36]. The key objective of the thesis is to propose a new method of modeling and prediction for a special class of nonlinear signals that exhibits the property of recurrence.

A potential application of the proposed model is the analysis of biomedical signals that exhibit the property of recurrence. Most of the cardiovascular signals signals as electro cardiogram (ECG) , blood pressure (BP) and respiration records belong to this category. Real time monitoring of these signals are important in clinical research and diagnosis of diseases. It is reported that corruption and loss of

these signals are common in the hospital settings [37]. A trained personal or a doctor may be able to deal with some distortion of signals, loss of signals, some of the noises that are present in the signal due to the cognitive abilities of human brain. It is a challenging task to develop an algorithm that can perform some cognitive tasks based on the contextual information as prediction the gaps or loss, identifying noises in the physiological data. There is a need to find methods that can deal with the corruption and loss of signals for the purpose of diagnosis and prediction. There are few methods available in literature (algorithms that were based on neural networks, kalman filters, adaptive filters and their combinations [38, 39] and collections of methods based on linear regression, pattern matching, average substitution, principal component analysis and wavelet decompositions [40, 41, 42, 43]) to address these issues [44]. Neural network based methods are reported as the current best solution and many of them can give reasonable predictions with intense training sessions [45, 46]. But the main drawbacks of neural network based methods are the following: (i) they are computationally very extensive (ii) most of them demand a long training time (iii) prediction is limited to the patterns or behaviors that are available during the training time otherwise that behavior is not learned or completely lost during prediction. A novel method that can make use of the recurrence property of the biomedical signals for prediction and modeling is demonstrated in the thesis. The redundancy structure of the topological technique delay embedding is exploited to exploited to reduce the computational load.

1.2 Objectives of the Thesis

Having identified the areas to be explored, the key objectives of the thesis are listed as follows.

1. Propose a nonlinear generalization of the method of singular value decomposition (SVD) for detecting and quantifying the nonlinearity in data.

2. Extend the method of nonlinear SVD to find the dimension of the manifold where data reside.

3. Propose a method of modeling and prediction for a time series that exhibit the

property of recurrence. explore the possibilities for the model for (i) Duffing Oscillator generated chaotic data and (ii) healthy Electrocardiogram (ECG) measurement.

4. Demonstrate a new method of stability analysis based on affine transformations specific to the recurrence neighborhood and verify the method by empirically determining the floquet coefficients from Duffing Oscillator generated data. Use the method of stability analysis for ECG data.

5. Extend the method of prediction based on recurrence neighborhoods for simultaneous multichannel physiological measurements for predicting one physiological signal from another signal.

6. Demonstrate the application of the methods for a multichannel data set of 100 records (that include various ECG channels, continuous invasive blood pressure channels, intra-cranial pressure, central venous pressure, respiration and raw fingertip plethysmogram outputs) of ICU patients.

1.3 Research Methodology

The research embodied in the thesis explores many ideas of Topology, Nonlinear Dynamics and Chaos theory. The proposed method of modeling demonstrates that the best way to exploit the recurrence property of a signal for prediction is to couple it with the delay embedding procedure. The key ideas on which the model is built are as follows:

Redundancy Structure of Delay Embedding

There is a special structure to the delay embedding matrix as it contains lot of redundant information. Delay vectors that constitute the embedding matrix are interconnected to each other. Due to this property, intermediate vectors can be written as a sum of few transformed end vectors. This relation is true even when the delay vectors belong to a nonlinear time series. One can exploit the special structure that delay embedding form in R^N for the purpose of prediction. The aim

of the proposed model of the thesis is to exploit this property of the delay embedding to reduce the computational load that is inevitable in nonlinear data analysis.

Affine Maps in the Neighborhoods of a Manifold of a Dynamical System

Consider an important property of the manifolds that represents a differential equation. A Differential equation can be seen as model that represents a given dynamic process. It generates maps known as one–parameter group of diffeomorphisms at every time of the evolution [47]. Hence given an initial condition, its evolution after a fixed time can be specificized by a unique map. This unique map, even if it is nonlinear has an affine form in a small neighborhood around it. This follows from the property that a nonlinear real analytic map has a valid Taylor series expansion at every point on the manifold. Further an affine map can be reduced to linear maps with respect to its rest points. Since a differential equation that model a given dynamic process generate these maps, there is an option to simplify the dynamics in terms of affine approximations in small neighborhoods everywhere on the manifold. This leads to a possibility of using a set of overlapping neighborhoods with specific affine transformations for each of them to represent the local dynamics.

Recurrent Timings reveal Dynamics

It is well known fact that Symbolic Sequences reveal information about the initial condition of the nonlinear chaotic systems [5]. Infinite precision of an initial condition in a chaotic system requires infinite length of symbolic sequence. In practice the longer the Symbolic Sequence, the more specific it renders information about the initial condition. A finite symbolic sequence can be seen as address of a partition of the statespace or a neighborhood of the manifold where the initial condition belongs to. The proposed model makes an assumption that recurrence timings, similar to the symbolic sequence contains information about the system under study.

Proposed Method of Prediction and Modeling

Given a time series that exhibits the property of recurrence the proposed method of modeling is as follows:

- A delay embedding of the data is done in a higher dimensional space R^N for the global representation of dynamics. Embedding is necessary to exploit the redundancy property of the delay embedding to reduce the computational load. The embedding dimension N and the Recurrent timings (the array that record the time of recurrence for the entire data) are empirically found.

- A Recurrent neighborhood is defined for a recurrent point in a high dimensional space R^N

  In order to take care the issues of over embedding, a practical solution is proposed by using Broomhead's theorem in Topology that Finite Impulse Response (FIR) filters preserve all the information one wants to extract by embedding techniques [48, 5]. Since such a filtered time series of a system preserves the embedding, a projection from the high dimensional space to a lower dimensional space is used. The lower dimensional model is economic in the sense data requirement, further it takes care of the ambiguity generated by high dimensional models [26, 27].

- A Recurrent neighborhood is defined in the lower dimensional space R^d using an FIR filter based projection and the local analysis of the system was done in R^d.

  The set of maps known as one–parameter group of diffeomorphisms exists at every time of the evolution on the manifold of a differential equation [47]. Since nonlinear real analytic maps has a valid Taylor series expansion at every point on the manifold, there is an option to simplify the dynamics in terms of affine approximations in small neighborhoods everywhere on the manifold. Hence a set of overlapping neighborhoods with specific affine transformations for each of them can be used to represent the local dynamics on the manifold.

  Recalling the Hartman-Grobman theorem that the behaviour of a nonlinear

system near an equilibrium point the system is qualitatively similar to the behaviour of a linear system [49], one can propose a generalization of the theorem. The proposed model relies on a hypothesis that the dynamics can be simplified in terms of affine transformations in small neighborhoods everywhere (and specifically as linear transformations in small neighborhoods around the rest points, as the affine map reduces to a linear map at the rest point).

- Neighborhood vectors in the lower dimensional space are collected into a set of equivalence classes based on their recurrent times (the time delay between two adjacent recurrences). For the local analysis of dynamics in R^d, the set of affine maps which are functions of only the recurrent times are derived across these equivalence classes. Since these maps are functions of the recurrent time alone, once the recurrent timings are known, one can generate an evolution of vectors in the lower dimensional space.

- A topological conjugacy between the recurrence neighborhoods reconstructed in both the lower and the higher dimensional spaces are demonstrated. Hence the evolution in the higher dimension space can be predicted from the evolution of the vectors in the lower dimensional space.

These possibilities are explored for a numerically data generated by Duffing oscillator first, and then for ECG data and other multichannel cardiovascular signals. Thesis shows that an extremely good prediction can be obtained by using a small amount of initial data and only information about recurrence intervals. Once the maps are determined, one could start at any point in the future and hence the initial data at that point is eventually unnecessary. The key inferences that came out by studying the model were:

1. Recurrence timings reveal lot of the information about the system dynamics that can be used for prediction and modeling.

2. The local dynamics of the system can be approximated by a set of few over lapping neighborhoods with specific affine transformations for each of them

3. By exploiting the redundancy structure of delay embedding and the conjugacy of the neighborhoods in lower and higher dimension the entire nonlinear structure can be approximated by one neighborhood (for a healthy ECG data and ECGI, ECGII, ECGIII, ECGV, AVR, ABP channels of multichannel records) or a few overlapping neighborhoods (for Duffing oscillator generated data)

1.4 Structure and Contents of Thesis

One of goals of the thesis was to introduce a new modeling method for nonlinear signals focusing on signal processing applications. Hence a few novel techniques were developed for processing nonlinear data. The proposed method of modeling uses these techniques for data processing. A nonlinear generalization of SVD is proposed for finding nonlinearity in a time series both in a qualitative and in a quantitative manner. This is a novel method of finding nonlinearity from data using an extension of SVD technique and the topological method of delay embedding without ignoring the dynamical perspectives. Thesis demonstrates the method of the technique of nonlinear SVD to retrieve nonlinearity from data generated by chaotic dynamical systems: Logistic map, Henon map, Van der Pol oscillator and Duffing oscillator. The recovery of parameters are shown in the following scenarios: (i) data generated by nonlinear maps and flows (ii) comparison of the method for both noisy and noise-free nonlinear data (iii) surrogate data analysis for both the noisy and noise-free cases, and also discuss two particular applications of the method: (i) Mathematical Modeling and (ii) Chaotic Cryptanalysis.

The method of nonlinear SVD is further extended to find the dimension of the system based on the time series. If the data on a manifold embedded in some R^N is available a method to compute the dimension of a manifold is discussed in the thesis. Two specific cases are discussed: For the simple case where the manifold is in the form of a lower dimensional affine subspace, the standard SVD is used to (i) calculate the dimension of the manifold and (ii) to get the equations which define the subspace. For the general case of manifolds, the nonlinear SVD is used (i) to search for an upper bound for the dimension of the manifold and (ii) to find the equations for the local charts of the manifold. This introduced method is highly

useful in the context of the Takens' embedding- a technique that is used through out this thesis for data analysis. Finding a good estimate of the underlying dimension of an embedded data is a requirement while modeling a system based on local charts. Thesis provides examples of this type of modeling.

The key goal of the thesis is to demonstrate the proposed method of modeling and prediction for the analysis of human electrocardiogram (ECG) data. First, the possibilities of the model are explored for a numerically generated data by an ideal dynamical system for a proper understanding and demonstration of the proposed method. Duffing oscillator generated data was specifically selected for this purpose as it is a well-studied system in nonlinear dynamics and it exhibit the property of recurrence. Then, the method is demonstrated for the analysis of the ECG record. For both the cases of the Duffing data and the ECG data, thesis shows that the entire nonlinear structure can be deduced from one or few overlapping neighborhoods.

Thesis also introduces a method for the stability analysis of the system based on data by studying the affine maps specific to the neighborhoods. This method was verified using the data generated by the Duffing oscillator as (i) it is a well studied system in nonlinear dynamics and (ii) for the stability analysis, floquet exponents of the system can be calculated and verified experimentally as their sum is known to be equal to trace of the coefficient matrix of the system [50]. Thesis also demonstrates that representing a system in a dimension higher than the inherent dynamics of the system leads to wrong inferences about its stability. For the case of ECG data, since we do not know the exact dimension of the dynamics, a new method of nonlinear singular value decomposition was used to find both the dimension of the manifold and the local dynamics. The dimension of the manifold is equal to the dimension of the system due to the conjugacy property. For the ECG data, an inference about the stability was made by studying the properties of affine maps specific to the neighborhoods. The thesis tries to give a theoretical justification of a well known experimental observation that the heart rate variability– the variability in beat-to-beat intervals of the heart, is necessary for a healthy functioning of the heart. Heart rate variability implies a stable dynamics where as a uniform heartbeat could result in instability based on the properties of the affine maps.

While analyzing the ECG signal using the proposed model, it was observed that there is lot of information available on the RR intervals that can be used for the purpose of modeling and prediction. The motivation for this study was the Physionet challenge 2010 which was about predicting a short segment of missing data from one of the channels in a multichannel physiological signals. Thesis demonstrates an effective way to model various ECG and BP signals by exploiting the recurrence property. For some specific class of signals like various ECG signals and BP, the cross predictions were excellent implying a generalized synchronization that exists between them. Analysis results showed that the predictions gives a good fit to the actual target for some of the signals of the data set under study. It is inferred that the model could potentially reflect the changes in the cardiovascular system efficiently for specific signals (ECGI, ECGII, ECGIII, ECGV, AVR, ABP). An excellent quality of reconstruction was obtained for all the ECG and BP signals. The PLETH, RESP, CVP scores were not as good as the ECG and BP scores. One reason for the bad prediction could be that either there is no connection between the channels (ECG Vs PLETH, RESP, CVP signals) or the connection was highly non-linear or the model developed was inefficient to extract the connection. RESP, CVP signals have identical recurrence patterns with respect to each other but a different recurrent timing with respect to the ECG signals. For the numerical simulation, recurrent neighborhoods and the recurrent timings were calculated with respect to the ECG /BP signals. There is an opportunity to improve the reconstruction of RESP, CVP signals by finding the recurrent neighborhoods and the recurrent timings with respect to them instead of ECG signal which we plan as a direction for future research.

The proposed new model is critically analyzed with respect to its ability to contribute in this crucial area of biomedical research. An online data set of 300 ICU patients published by Physionet as part of the challenge 2010 was used to evaluate the relevance of the model [51]. The challenge was to find a short stretch of missing data in one of the channels of a multichannel data, using the data available in all channels. The multichannel recordings were measurements of various physiological functions. Given that some of these signals had a lot of mutual information as they

represented the activities of the same system, there was a possibility of some relationship between different channels that are simultaneously recorded. The proposed new model was successful to exploit the redundancy and mutual relationships across channels for prediction. The prediction results are excellent for the signals that had identical recurrent patterns as demonstrated in the thesis.

For the prediction of missing data 300 data sets in the data analysis section, the proposed method of prediction worked extremely well (in a range of scores 0.90 to 0.99 for most of cases) for cardiovascular signals (various ECG signals and some blood pressure (ABP, ICP) signals) that had identical recurrent timings. This can be attributed to the fact that they carried a lot of mutual information by basically representing activities of the same system. Also there was a possibility of some relationship between different channels that are simultaneously recorded from the same person. The model could efficiently exploit the redundancy and mutual relationships across some channels to predict or reconstruct the lost or corrupt as discussed in the last chapter of the thesis. But prediction were not that efficient for some other signals: plethmosgraphy (scores ranging around the value 0.5 with a highest score approaching 0.99) and CVP (score ranging around 0.5 with a highest score approaching 0.91) and Respiration (wide range of scores from 0 to 0.9 with 2 data sets scoring up to 0.92).

1.5 Organization of Thesis Chapters

First section of the thesis (Chapters 2 and 3) covers few main techniques along with its theoretical and algorithmic details, that are used for processing data throughout the thesis. Next section (Chapters 4–7) covers the proposed method of modeling and prediction, stability analysis and prediction results for a chosen data set. Thesis consists of the following chapters and their brief description are given below:

- **Chapter 1: Introduction** – This is an introduction to the thesis. It gives an overview of the thesis– the purpose and the significance of the study with respect to the background literature, key goals and objectives of the thesis,

new ideas explored in the research, methodology of research, structure and contents of the thesis.

- **Chapter 2: Nonlinear generalization of Singular Value Decomposition (SVD)**– This chapter introduces the method of nonlinear SVD for both the qualitative detection and quantitative determination of nonlinearity in a time series. Two applications of the method (i) Mathematical Modeling and (ii) Chaotic Cryptanalysis are demonstrated.

- **Chapter 3: Nonlinear SVD to estimate the Dimension of the Manifold and finding Local Charts**– This chapter propose a method to arrive at an upper bound for the dimension of a manifold which is embedded in some R^N using nonlinear SVD. This method is useful in the context of the Takens' embedding which is used for analysis of data in throughout this thesis. A good estimate of the underlying dimension of an embedded data is required for developing a model based on local charts. Next two chapters provide an example of such an application.

- **Chapter 4: Prediction based on the Recurrence Neighborhoods** – This chapter introduce the new method for prediction that exploits the property of recurrence in a dynamical system. It demonstrates the possibility of predicting the entire state space based on the information available on few small neighborhoods alone using data generated by Duffing oscillator under chaos.

- **Chapter 5: A Topological Structure of ECG signals** – This chapter demonstrates the proposed method of modeling and prediction for a healthy ECG data. For the special case of ECG, the entire data was predicted based on the information available on one small neighborhoods by appropriately choosing the embedding dimension.

- **Chapter 6: Stability Analysis based on Affine maps specific to Recurrent Neighborhoods**– This chapter demonstrates a new method of stability analysis using affine maps specific to the recurrence neighborhood. It

also contains a discussion about the possible theoretical foundations of an important experimental observation regarding heart rate variability.

- **Chapter 7: Prediction across Recurrent Neighborhoods: A Study on Multichannel Data**– This chapter extends the proposed method of modeling and prediction for multichannel physiological data and contains a study of the proposed algorithms on a set of 100 data sets.

- **Chapter 8: Conclusions and Future research directions** – This concluding chapter contains the significance of results, contribution of the thesis and lists some open problems and possible future research directions.

CHAPTER 2
Nonlinear Generalization of Singular Value Decomposition

Singular Value Decomposition (SVD) is a powerful tool in linear algebra and has been extensively applied to Signal Processing, Statistical Analysis and Mathematical Modeling. The goal of this chapter is to introduce a nonlinear generalization of SVD for finding nonlinearity in a time series in a qualitative and quantitative manner. It demonstrates how the method can retrieve nonlinearity from data generated by discrete dynamical systems: Logistic map and Henon map. Recovery of parameters are successfully demonstrated for real and surrogate data for both the noisy and noise–free cases. The method is extended to identify the system parameters for continuous systems: Van der Pol oscillator and Duffing oscillator. Two specific applications of the method (i) Mathematical Modeling and (ii) Chaotic Cryptanalysis are discussed.

(Note: Research embodied in this chapter is published in a peer reviewed journal and [52] is the reference for that publication)

2.1 Introduction

Singular Value Decomposition (SVD) is a standard technique for matrix decomposition in linear algebra and it has found numerous applications in various fields such as signal processing, statistical analysis, biomedical engineering, genetics, mathematical and statistical modeling and graph theory [6, 7, 8, 9, 53]. Historically SVD has been used for finding the dimension of a linear system as it gives statistically independent set of variables which could span the state space [10, 11]. Since methods like SVD are inherently designed to identify linear dependence or correlation, they work extremely well when the signals are generated by linear systems. Though SVD is fundamentally a linear technique, a method based on SVD known as singular spectrum analysis was introduced by Broomhead for a qualitative detection of underlying nonlinearity [12]. An abrupt decrease in the profile of the singular spectrum was shown as an indication of lower dimensional determinism or chaos. But this

method fails to distinguish a chaotic time series from its surrogates; the stochastic counterparts of the data which has the same power spectrum [13, 14]. Later, a new method of quadratic scaling of singular values was proposed that can distinguish between the data series and its surrogates [54]. Quadratic scaling improves the significance of decreasing singular values thus highlighting the deterministic or stochastic features.

Chaotic signals resembles noise though they are generated by lower dimensional deterministic dynamical systems. Noise on the other hand is stochastic in nature and very less is known about its origin. Noise is defined mainly based on its statistical properties as variance, mean, correlation and entropy; where as chaotic signals are defined based on its dynamic properties, initial conditions and generating equation. Chaotic signals generally have smaller degrees of freedom compared to noise. An algorithm that could determine correlation dimension from an irregular time series was proposed by Grassberger–Procaccia (known as GP algorithm) [55]. A finite correlation dimension is considered as an indication of underlying determinism. The GP algorithm has a drawback as it fails when the data is noisy and is unable to distinguish stochastic processes with power–law power–spectra from chaos [56, 57]. A distribution of correlation coefficients is also reported as a qualitative method to distinguish chaos from noise because the spectrum is flat for noise but gradually decays for chaos. But this method fails to distinguish between correlated noise and chaos [58].

Additionally, there are many novel algorithms that extend regression methods to estimate model parameters based on the best nonlinear polynomial fit. The goal of such algorithms are to find the parameters of the appropriate models from the data based on various least square techniques and most of them are found to be successful in a large number of engineering problems. One method which is an alternative to linear regression, known as ACE (Alternating Conditional Expectation) algorithm is used to identify the optimal transformations to the data such that a linear regression model can be applied to the transformed data [59]. Some of the other robust algorithms available are: Fast Orthogonal Search (FOS) method that can obtain correct model parameters irrespective of the model selection [60], one of its

variation known as Optimal Parameter Search (OPS) algorithm [61], various least square techniques: Least Square [10], Total Least Square (TLS) [62] and Minimizing the Hyper surface Distance (MHD) method [63].

From a dynamical system perspective, embedding based techniques [24, 16] are useful as they can give qualitative information of the dynamics from the experimental time series. The data generated by nonlinear chaotic systems live in a lower dimensional manifold, unlike the data from stochastic systems that tend to spread over the whole phase space. This property of nonlinear chaotic systems can be exploited for modeling and prediction of nonlinear chaotic signals as we will see in the following chapters. Takens suggested a reconstruction based on delay vectors of the time series and gave a firm theoretical foundation for both delay and derivative embedding [25]. Broomhead and King proposed an alternative method 'SVD based embedding' as SVD can get the best orthogonal basis to represent the data [12]. In this context Porta *et al.* proposed a local nonlinear prediction method that used non–overlapping hypercubes of the embedded phase space for the purpose of prediction of biomedical signals [64, 65].

Since the goal of the thesis is to introduce a new modeling method for nonlinear signals focusing on signal processing applications, few new techniques are developed for processing the data generated by nonlinear system. The proposed method of modeling that is described in the following chapters uses these techniques for data processing. The objective of this chapter is to introduce a nonlinear generalization of SVD and demonstrate how it can be used for detecting and quantifying nonlinearity in a time series. Next chapter extends the method to find the local dimension of the manifold were the data resides. The proposed method of nonlinear SVD for the case of discrete dynamical systems is similar to the nonlinear Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) regression proposed by Lu *et al.* [61] and Marmarelis [66]. An advantage of this method over others is that it is not limited to just polynomial regression. Any deterministic nonlinearity present in the data could potentially be recovered by choosing the appropriate functions for regression. Nonlinear SVD method could be considered as functional regression in an extended phase space. This chapter further demonstrates that nonlinear SVD coupled with

the method of finding accurate derivatives from data [67] can be used to identify the nonlinearity from in a time series generated by continuous dynamical systems.

2.2 Singular Value Decomposition (SVD)

Singular Value Decomposition is a generalization of the eigen decomposition of square matrices, for the decomposition of rectangular matrices. SVD decomposes a rectangular matrix into three simple matrices: two orthogonal matrices and one diagonal matrix. In general, SVD theorem [11] can be stated as: Any $m \times n$ real matrix A of dimension $m \geq n$ can be factored into three matrices: U (orthogonal, $m \times m$ matrix), W (diagonal, $m \times n$ matrix) and V (orthogonal $n \times n$ matrix) such that,

$$A = UWV^T \tag{2.1}$$

For a complex matrix A, U and V are unitary matrices and W is a real matrix. V^T denotes the transpose of V. The diagonal elements of W matrix are known as the singular values of A.

This decomposition works well with matrices that are either singular or else numerically very close to singular. Hence SVD is used to calculate pseudo–inverses when the natural inverse of the matrix does not exist [10]. SVD and pseudo–inverses are used in statistics for solving least square problems. Data compression using SVD is one of the standard applications in image processing [68].

2.3 Method of Nonlinear Singular Value Decomposition

This section explains the method of nonlinear SVD using a delay embedding matrix. Consider a finite time series $\{x_1, x_2, x_3, \ldots, x_m\}$ generated by a system. Delay embedding is a method of reconstruction of the state space using delayed data segments of the time series known as embedding vectors [25]. An embedding matrix E (which represents an embedding of the data in R^N) can be created using

time delayed vectors as follows.

$$\begin{aligned} Y^1 &= (x_1\ x_2 \ldots x_N\ ) \\ Y^2 &= (x_2\ x_3 \ldots x_{N+1}\ ) \\ &\ldots \\ Y^i &= (x_i\ x_{i+1} \ldots x_{N+i-1}\ ) \\ &\ldots \end{aligned}$$

Let Y^i represents a typical embedding vector of length N in R^N. The collection of vectors $\{Y^i\}$ in R^N represents the delay embedding of the given time series. The corresponding embedding matrix E (of dimension $p \times N$) that represents p embedding vectors in R^N is,

$$E = \begin{bmatrix} x_1 & x_2 & \ldots & X_N \\ x_2 & x_3 & \ldots & x_{N+1} \\ & & \ldots & \\ x_p & x_{p+1} & \ldots & x_{N+p-1} \end{bmatrix} \tag{2.2}$$

The procedure of nonlinear SVD is to extend the embedding matrix E by adding nonlinear functions of the columns of the E matrix. Let F be the extended embedding matrix and $f_1, f_2 \ldots f_K$ represents the additional columns.

$$F = [E\ :\ f_1\ f_2\ \ldots f_K\ ] \tag{2.3}$$

In general, a nonlinear function refers to the square, cube, any other higher powers of the elements in a column, the element–wise product of two or more columns or any other kind of nonlinearity such as exponentiation and trigonometric functions of the entries in a column of the E matrix. If there is a nonlinear relationship between these column vectors, it could be interpreted as a linear relationship between the original columns vectors and the added nonlinear columns.

Assuming that E is extended by adding K nonlinear columns, the dimension of F matrix would be $p \times q$ where $q = N + K$. This extended embedding matrix F can

be considered as a higher dimensional linear system. Now the SVD procedure can find the linear relation between the embedding vectors and corresponding nonlinear columns, thus recovering any nonlinear relationship that is inherent in the data. Let the SVD of F be,

$$F = UWV^T \tag{2.4}$$

Post multiplying with V on both sides,

$$FV = UW \tag{2.5}$$

Expanding Eq. 2.5

$$F \begin{bmatrix} V^{<1>} & \dots & V^{<q>} \end{bmatrix} = \begin{bmatrix} U^{<1>} & \dots & U^{<q>} \end{bmatrix} \begin{bmatrix} W_{1,1} & 0 & \dots & 0 \\ 0 & W_{2,2} & \dots & 0 \\ \dots & & \dots & \dots \\ 0 & 0 & \dots & W_{q,q} \end{bmatrix} \tag{2.6}$$

Using partitions of V and W and expanding along the last column of V and W such that,

$$\begin{bmatrix} F^{<1>} & \dots & F^{<q>} \end{bmatrix} . \begin{bmatrix} V_{1,q} \\ V_{2,q} \\ \dots \\ V_{q,q} \end{bmatrix} = \begin{bmatrix} U^{<1>} & \dots & U^{<q>} \end{bmatrix} . \begin{bmatrix} 0 \\ 0 \\ \dots \\ W_{q,q} \end{bmatrix} \tag{2.7}$$

Note that $F^{<i>}$ represents the i^{th} column of F matrix and $W_{i,j}$ is the element on i^{th} row and j^{th} column of W matrix. Expanding Eq. 2.7,

$$\begin{aligned} V_{1,q}F^{<1>} + V_{2,q}F^{<2>} + \ldots + V_{q,q}F^{<q>} &= 0.U^{<1>} + \ldots + W_{q,q}U^{<q>} \\ &= W_{q,q}U^{<q>} \end{aligned}$$

If the q^{th} singular value of W is zero, $W_{q,q} = 0$ then,

$$V_{1,q}F^{<1>} + V_{2,q}F^{<2>} + \ldots + V_{q,q}F^{<q>} = 0 \tag{2.8}$$

Eq. 2.8 states that the columns of F including those formed by the nonlinear functions $\{f_1, f_2 \ldots f_K\}$ now span a linear vector space. This equation is true for all rows of F and hence by exploiting this relation, the dependance between the linear and nonlinear columns of the F matrix can be recovered.

2.4 Nonlinear SVD for Identifying Nonlinearity in Data

The key assumption behind the method of SVD is that the underlying equation is a function of the delay vectors of the data. If the data is nonlinear, standard SVD does not give a zero singular value even when it is noise–free. Then the nonlinear SVD procedure is to try different nonlinear functions $f_1, f_2 \ldots f_K$ for extending the embedding matrix as shown in Section 2.3. The selection of nonlinear functions depends on the domain of application. The criterion is to try different choices of f_i's for the F matrix and select the F which gives at least one 'zero' or 'nearly zero' singular value. Even if the correct nonlinear functions are chosen, a zero singular value will be obtained only when the data is noise–free. When the data is noisy, the singular value $W_{q,q} \neq 0$ even if the chosen F matrix has the right set of nonlinear functions. Hence the procedure is to try different nonlinear functions and choose the F that gives the lowest value of $W_{p,p}$. For the purpose of numerical analysis the ratio $\frac{W_{q,q}}{W_{1,1}}$ (the ratio of q^{th} singular value to the 1^{st} singular value) was set to be below some preset criterion. As the noise level in the data increases, chances are higher that the method of nonlinear SVD fails. Therefore our confidence in the estimated model equation goes down with the increase in noise.

The choice of appropriate nonlinear functions is crucial for finding the nonlinearity in data. The selection of the required nonlinear functions are dependant on the domain of application. This chapter illustrates a choice of quadratic, cubic polynomials and sinusoidal functions. These three sets of functions are chosen as they are solutions of low order differential equations with constant coefficients. In practice,

solutions of differential equations with variable coefficients (e.g. Bessel functions) or nonlinear low dimensional equations are good candidates for this purpose.

2.5 Numerical Analysis using Data generated by Logistic Map

The section shows how the method of nonlinear SVD can be used to empirically detect the nonlinearity in the data and determine the equation which generated it. Numerical analysis is done using data generated by Logistic map. It is demonstrated that a choice of quadratic functions for the extended matrix of nonlinear SVD procedure retrieves the exact equation from the time series.

2.5.1 Recovering Nonlinearity

Consider a discrete dynamical system, the Logistic map $F(X) = \lambda X(1 - X)$ where $0 \leq X < 1$ and $\lambda = 4$. Let $\{x_1, x_2, \ldots, x_m\}$ be a finite time series generated by the map such that,

$$x_{i+1} = \lambda x_i(1 - x_i) \tag{2.9}$$

The objective is to identify the form of this equation including the parameter values from the time series. For the purpose of demonstration consider a delay embedding of the numerical data in to R^3 using p embedding vectors as shown by the following matrix:

$$E = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \\ & \cdots & \\ x_p & x_{p+1} & x_{p+2} \end{bmatrix}$$

The dimension of E matrix is $p \times 3$. After SVD operation on the embedding matrix E, it was observed that none of the singular values of E go to zero indicating no linear dependance present in the data vectors. Next step is to extend the E matrix by adding nonlinear columns. If a quadratic function is chosen as a candidate F

would be,

$$\begin{aligned} F &= [E \ : \ f_1 \ ] \\ &= [E^{<1>} \ E^{<2>} \ E^{<3>} \ f_1 \ ] \\ &= [E^{<1>} \ E^{<2>} \ E^{<3>} \ {E^{<1>}}^2 \ ] \end{aligned}$$

A typical row of F can be seen as [X_n X_{n+1} X_{n+2} ${X_n}^2$]. SVD of F gives a zero singular value $W_{4,4} = 0$ indicating a linear dependence between the column [X_n X_{n+1} X_{n+2} ${X_n}^2$]. Expanding Eq. 2.8 for this particular numerical example,

$$V_{1,4}(X_n) + V_{2,4}(X_{n+1}) + V_{3,4}(X_{n+2}) + V_{4,4}(X_n^2) = 0 \tag{2.10}$$

Substituting the numerical values in Eq. 2.10,

$$\begin{aligned} -0.696311X_n + 0.174078X_{n+1} + 0X_{n+2} + 0.696311X_n^2 &= 0. \\ 4X_n - X_{n+1} - 4{X_n}^2 &= 0 \\ 4X_n(1 - X_n) &= X_{n+1} \end{aligned}$$

Hence the exact equation is recovered from the data. For the numerical example discussed above, the initial condition that generated the trajectory was 0.02 and the parameter value $\lambda = 4$. The numerical results can be reproduced for any initial condition (which is not equal to zero) and for parameter values $0 < \lambda \leq 4$.

2.5.2 Singular Value Spectrum for Standard SVD and Nonlinear SVD

This section discusses the properties of singular value spectrum of standard SVD and nonlinear SVD after quadratic scaling of singular values. Quadratic scaling is known to improve the significance of decreasing singular values thus highlighting the deterministic or stochastic features [54]. Fig. 2.1 shows the quadratically scaled spectrum of singular values for the case of standard SVD ($n^2.\sigma_n$) and nonlinear SVD ($n^2.\rho_n$), where σ_n is the n^{th} singular value generated by the standard SVD on the embedding matrix E. The dimension of the embedding matrix selected was 21×21 for this particular example. Standard SVD operation gave 21 non-zero singular

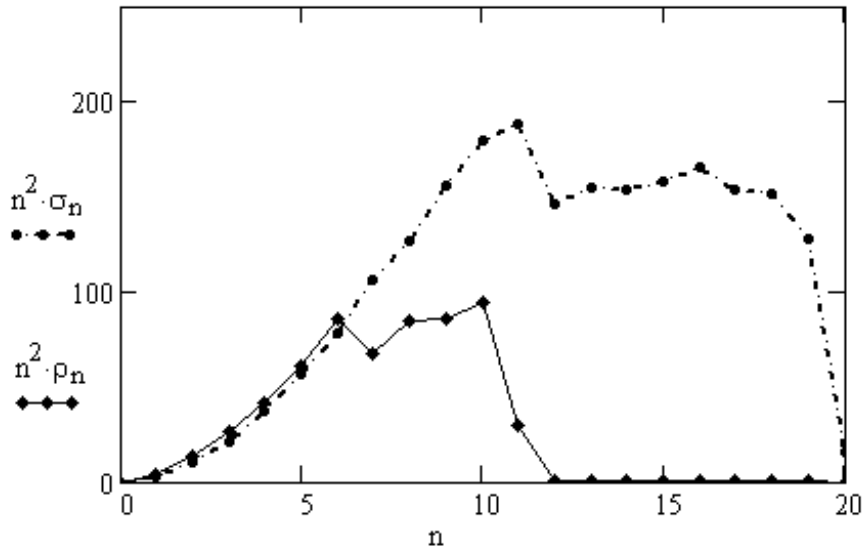

Figure 2.1: Scaled spectrum of singular values for the case of standard SVD $(n^2.\sigma_n)$ Vs n (where σ_n is the n^{th} singular value of E matrix) and nonlinear SVD $(n^2.\rho_n)$ Vs n (where ρ_n is the n^{th} singular value of F matrix).

values: note that the profile gradually increases and slowly comes down.

For the nonlinear SVD operation, the dimension of the embedding matrix F was kept same as that of E i.e. 21×21; for which the last 10 columns of E were replaced by the squares of the first 10 columns. Nonlinear SVD gave 21 singular values, out of which the last 10 were zero. Observe that the nonlinear SVD profile in Fig. 2.1 is significantly different from that of conventional SVD case as the former is 'flat' towards the end. There is a significant qualitative change in the spectrum as the profile of nonlinear SVD drops to zero rapidly compared to the case of standard SVD. Existence of a zero singular value indicates that the columns of corresponding embedding matrix can be expanded in the form of a linear equation as Eq. 2.8. This implies that the underlying nonlinear equation can be recovered as shown in Section 2.3.

The time series for the simulation was generated by the Logistic map under chaos: $X_{n+1} = \lambda X_n(1 - X_n)$ where $0 < X_n < 1$ and $\lambda = 4$. The embedding matrix

E was created as explained in Section 2.3 and the extended embedding matrix F was generated by replacing the last 10 columns of E by the squares of the first 10 columns.

2.5.3 Comparison of the Method on Data and its Surrogates

This section discusses whether nonlinear SVD can differentiate between the chaotic data and its surrogates. Surrogates are the non–deterministic counterparts of the data which are generated by randomizing the phases of the Fourier Transform of the data; thereby maintaining the same probability distribution and power spectrum as the chaotic data [69, 70].

The time series $\{X\}$ for the simulation was generated by the Logistic map under chaos: $X_{n+1} = \lambda X_n(1 - X_n)$ where $0 \leq X_n \leq 1$ and $\lambda = 4$. A set of surrogate data series $S_1, S_2, \ldots, S_i$ were generated from $\{X\}$ such that $\{S_i\}$ and $\{X\}$ had the same power spectrum. Fig. 2.2 (i) shows the quadratically scaled spectra of singular values for the Logistic data and two of its surrogates S_1, S_2 using the standard SVD procedure under noise–free conditions. Fig. 2.2 (ii) shows the spectra for the case of nonlinear SVD. It can be observed from Fig. 2.2 that the standard SVD does not distinguish between the data and the surrogates, but the nonlinear SVD clearly distinguishes them. The qualitative change in the profile of the spectrum happened due to the zero singular values in the case of nonlinear SVD procedure.

2.5.4 Comparison of the Method for different choices of Nonlinear Functions

Depending upon the domain of application of Logistic map data which is very popular in population dynamics studies, a quadratic nonlinearity is a decent candidate for the nonlinear function. But assume that one wants to see how the profile of singular spectrum changes if the type of the nonlinear functions of the embedding matrix F varied. If cubic columns are selected in the F matrix instead of quadratic columns for the case of logistic map data, spectrum looks as in Fig. 2.3. This figure shows that both the spectrums (cubic and quadratic case) looks qualitatively similar, but the exact relationship cannot be retrieved quantitatively in the cubic case as none of the singular values have exact zero vales. Hence according

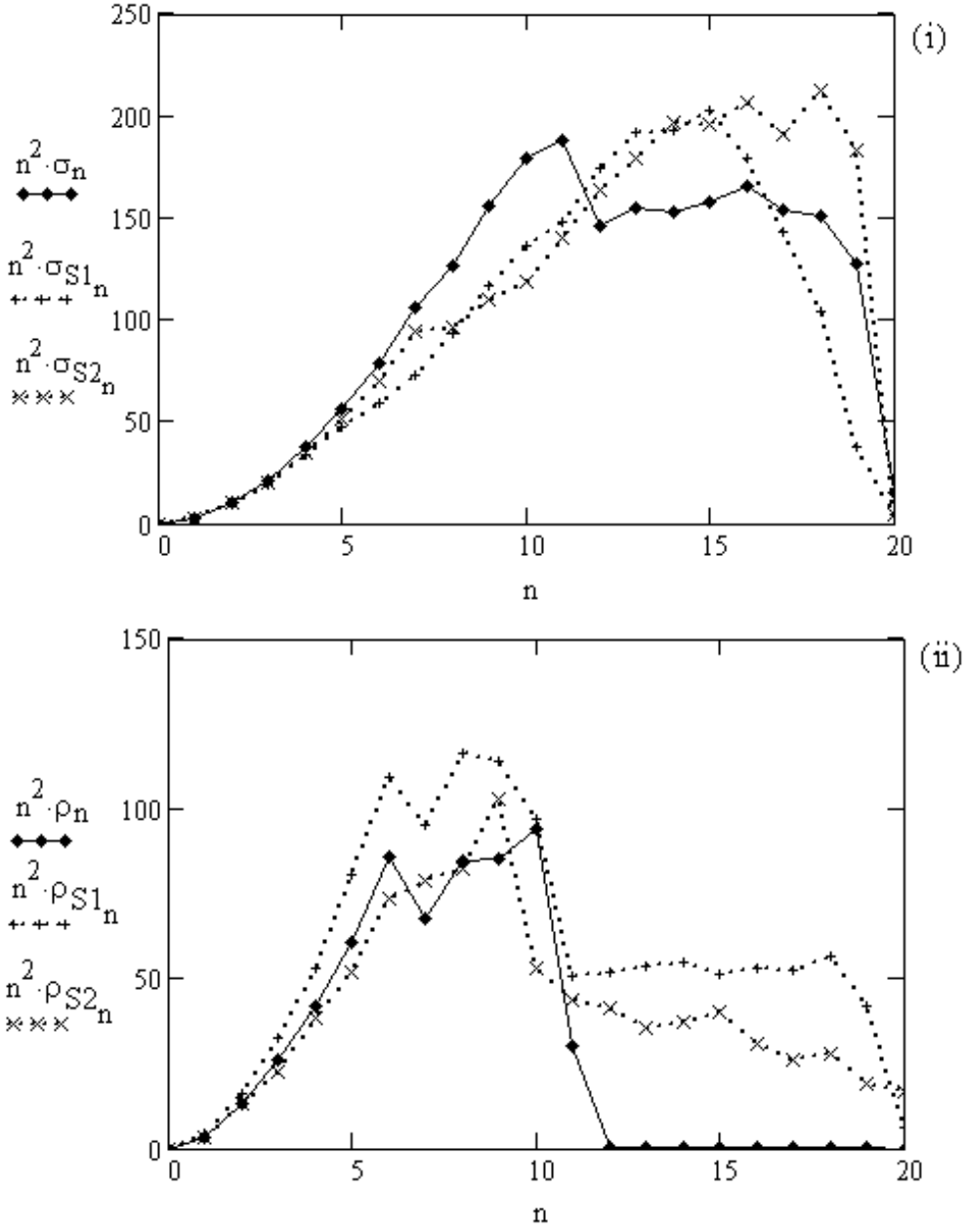

Figure 2.2: Scaled spectrum of Singular Values for standard SVD and nonlinear SVD for Logistic map generated data and its Surrogates: (i) spectrum for the data $(n^2.\sigma_n)$ and its surrogates $(n^2.\sigma_{S1_n})$ and $(n^2.\sigma_{S2_n})$ using standard SVD, (ii) spectrum for the data $(n^2.\rho_n)$ and its surrogates $(n^2.\rho_{S1_n})$ and $(n^2.\rho_{S2_n})$ using nonlinear SVD.

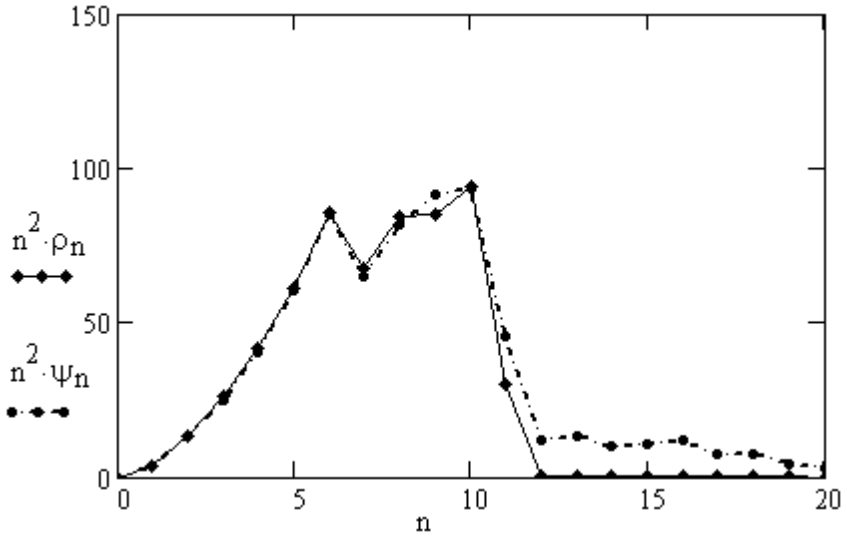

Figure 2.3: Scaled singular value spectra of nonlinear SVD for different choices of nonlinear functions on data generated by Logistic map: $n^2.\rho_n$ versus n (where ρ_n is the n^{th} singular value of F with quadratic columns) and $n^2.\psi_n$ versus n (where ψ_n is the n^{th} singular value of F with cubic columns)

to the proposed method, one has to try different choices of embedding matrices and select the one which gives at least one nearly zero singular value.

2.6 Recovering Nonlinearity: Higher order Maps

2.6.1 Data generated by Henon Map

This section demonstrates how to recover the nonlinear equation when the data is generated from a higher order map of the following form,

$$X_{n+2} = f(X_n, X_{n+1}) \tag{2.11}$$

The Henon map, falls in this category.

$$X_{n+1} = c - aX_n^2 + Y_n \tag{2.12}$$
$$Y_{n+1} = bX_n \tag{2.13}$$

Consider the X data generated by this map for parameter values $a = 1.4$ $b = 0.3$ and $c = 1$. The objective of the analysis is to check if one can recover the nonlinear equation from the X data. An embedding matrix F with quadratic nonlinear functions, such that F had the form [1 X_{n+2} X_{n+1} X_n ${X_n}^2$ X_{n+1}^2 $(X_{n+1}X_n)$] was set up this purpose. Since the nonlinear SVD procedure gave a zero singular value for this matrix, the columns were expanded using Eq. 2.8 to get the following equation,

$$\begin{aligned} 0.496904 - 0.6956656X_{n+1}^2 + 0.1490712X_n - 0.496904X_{n+2} &= 0 \\ 1 - 1.4X_{n+1}^2 + 0.3X_n - X_{n+2} &= 0 \\ 1 - 1.4X_{n+1}^2 + 0.3X_n &= X_{n+2} \end{aligned}$$

The equation $X_{n+2} = 1 - 1.4X_{n+1}^2 + 0.3X_n$ is equivalent to Eq. 2.12 and 2.13 for parameter values $a = 1.4$ $b = 0.3$ and $c = 1$. Hence the proposed method is successful in retrieving the form of the equation and its the parameters from the data.

2.6.2 Even Iterates of Data generated by Logistic Map

Consider the case of data generated by the Logistic map where all the odd iterates of the time series are suppressed or removed. The goal is to find the parameters λ from the available data. In this case one could assume that the underlying equation is of the form,

$$X_{n+2} = g(X_n) \tag{2.14}$$

Since the alternate iterates are suppressed in the data the underlying nonlinearity is quartic or fourth order of the form $g(x) = \alpha x + \beta x^2 + \gamma x^3 + \delta x^4$ where $\alpha, \beta, \gamma, \delta$ are functions of the parameter λ of the Logistic map. An embedding matrix

of the form $F = [\ X_n\ X_{n+2}\ {X_n}^2\ {X_n}^3\ {X_n}^4]$ can be used to retrieve quartic nonlinearity. For this specific case the retrieved equation corresponding to zero singular value was,

$$0 = 14.761X_n - X_{n+2} - 71.473{X_n}^2 + 113.423{X_n}^3 - 56.712{X_n}^4 \quad (2.15)$$

$$X_{n+2} = 14.761X_n - 71.473{X_n}^2 + 113.423{X_n}^3 - 56.712{X_n}^4 \quad (2.16)$$

Comparing Eq. 2.16 to the second iterate map of Logistic equation,

$$X_{n+2} = (\lambda^2)X_n - (\lambda^2 + \lambda^3){X_n}^2 + (2\lambda^3){X_n}^3 - (\lambda^3){X_n}^4 \quad (2.17)$$

One can recover the exact parameter value $\lambda = 3.842$ that was used for the numerical simulation to generate the data, by comparing Eq. 2.16 and Eq. 2.17.

2.7 Numerical Analysis for Noisy Data

This section contains numerical analysis results for the noisy data. The method of nonlinear SVD works reasonably well to identify the underlying deterministic equation from a noisy data, provided the noise is below a threshold. A criterion is developed to determine the upper bound of acceptable noise, beyond which the confidence in the proposed method goes down.

Assume that data $\{X_n\}$ is contaminated by some additive noise $\{P_n\}$ which is either gaussian or uniform, $\{\widehat{X}_n\} = \{X_n\} + \{P_n\}$. The procedure explained in Section 2.3 can be used to create F matrix with added nonlinear columns from the noisy data $\{\widehat{X}_n\}$. For the numerical examples discussed in this section, this upper bound was empirically determined using the singular value spectrum by nonlinear SVD as a guide. The ratio $\frac{W_{q,q}}{W_{1,1}}$ mentioned in Section 2.4 was used for this purpose and the upper bound for the ratio was set as 10^{-6}. In practice this spectral criterion translates into a criterion for the Peak Signal to Noise Ratio (PSNR) of the signal. PSNR is defined as $20\log_{10}(\max(\text{Signal})/\sqrt{MSE})$ where MSE represents the Mean Squared Error– the average of the square of the noise present in the signal. Note that for the case of standard SVD on the noisy data, the ratio $\frac{W_{q,q}}{W_{1,1}}$ was in the range $(10^{-1}, 10^{-3})$ for different noise levels for Logistic and Henon Map data discussed in

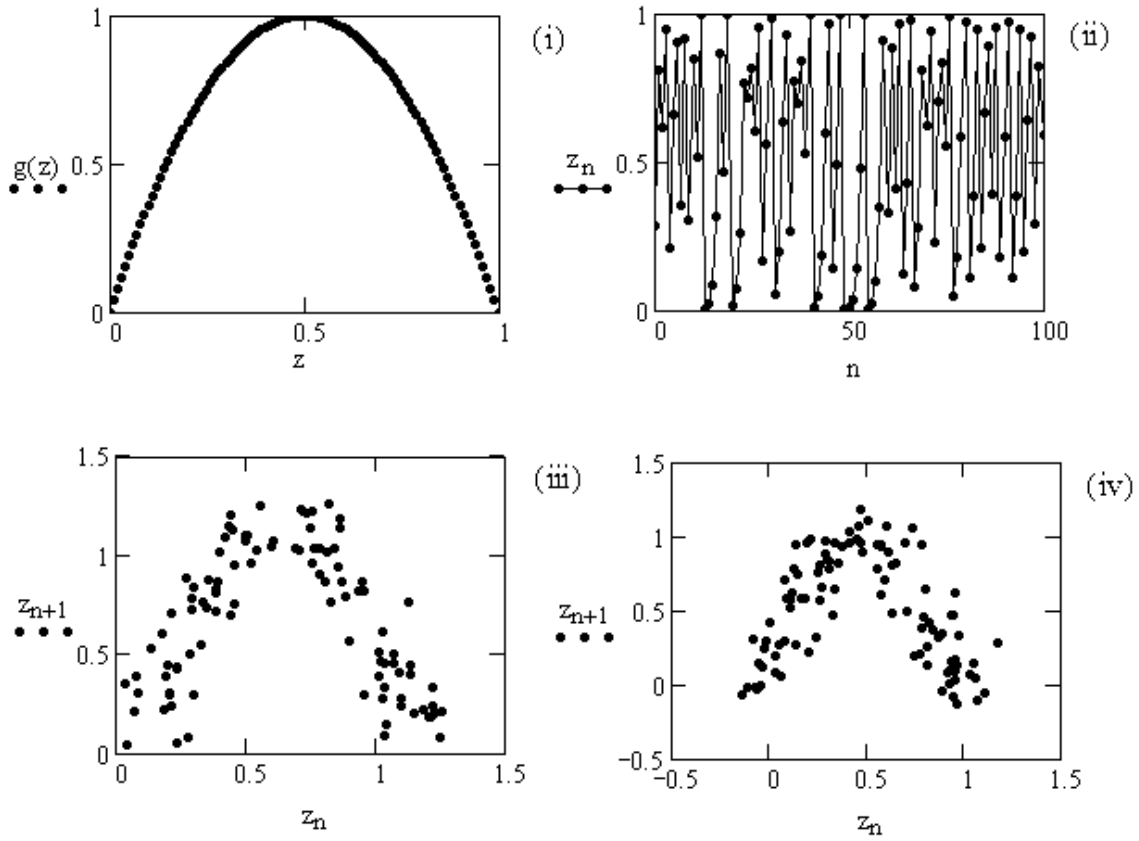

Figure 2.4: Phase space of Logistic map (i) and a chaotic time series generated from it (ii). Phase space reconstructed from noisy data (z_{n+1} Vs z_n) for noisy data (iii) with additive uniform noise N(0,1) with noise level 28.089 % (iv) with additive gaussian noise N(0,1) with noise level 28.123 %

next two sections.

2.7.1 Retrieving Logistic Map Parameters from Noisy Data

This section shows how to retrieve the nonlinear equation from a data generated by Logistic map which is contaminated by noise. Fig. 2.4 (i) and 2.4 (ii) shows the phase space of Logistic map and a chaotic time series generated from it. Fig. 2.4 (iii) and 2.4 (iv) show the state space created from the noisy data for both the cases of uniform and gaussian noises. Retrieved parameters for both the cases of uniform and gaussian additive noises are given in Table. 2.1. It lists

the estimated values of parameter $\hat{a}, \hat{b}$ for the data generated by Logistic family of maps: $X_{n+1} = a_1 X_n - a_2 {X_n}^2$ where $0 \leq X_n \leq 1$ for the values of $a = 4$ and $b = 4$ for different Peak Signal to Noise Ratio (PSNR) using the nonlinear SVD method. When the preset criterion $\frac{W_{q,q}}{W_{1,1}}$ was below 10^{-6}, an assumption was made that the singular values smaller than this can be considered zero and the underlying equation is extracted as explained in the noise-free case. For the numerical examples cited below the spectral criterion was satisfied for PSNR above 28 for gaussian noise and 33 for uniform noise.

Table 2.1: Estimated values of parameters $\hat{a_1}, \hat{a_2}$ for the Logistic map: $X_{n+1} = a_1 X_n - a_2 {X_n}^2$ where $0 \leq X_n \leq 1$ in the presence of noise using nonlinear SVD.

Parameters	Noise (Uniform Distribution)			Noise (Gaussian Distribution)		
	$\hat{a_1}$	$\hat{a_2}$	PSNR	$\hat{a_1}$	$\hat{a_2}$	PSNR
	4.036	3.994	44.231	4.099	4.087	43.450
$a_1 = 4$	4.142	4.020	38.730	4.106	4.091	41.839
$a_2 = 4$	3.875	3.847	37.796	4.247	4.156	37.095
	4.235	4.151	36.462	4.299	4.148	34.523
	4.389	4.250	33.609	4.156	3.995	32.856
	4.480	4.278	30.711	4.030	3.824	28.817
	4.429	3.850	24.631	3.804	2.562	24.151

Table 2.2: Estimated values of parameters $\hat{a}, \hat{b}, \hat{c}$ for the Henon map $X_{n+1} = c - aX_n^2 + Y_n$; $Y_{n+1} = bX_n$ in the presence of noise using nonlinear SVD.

Parameters	Noise (Uniform Distribution)				Noise (Gaussian Distribution)			
	$\hat{a}$	$\hat{b}$	$\hat{c}$	PSNR	$\hat{a}$	$\hat{b}$	$\hat{c}$	PSNR
	1.389	0.289	1.003	46.863	1.400	0.299	1.003	44.515
a = 1.4	1.381	0.292	1.010	40.764	1.398	0.296	1.005	40.930
b = 0.3	1.357	0.276	1.009	36.178	1.400	0.316	0.988	35.931
c = 1.0	1.371	0.262	1.020	32.975	1.437	0.325	1.008	34.113
	1.367	0.275	1.067	30.493	1.396	0.327	1.040	31.336
	1.296	0.268	1.034	26.549	1.520	0.340	1.080	28.325
	1.326	0.156	1.148	20.835	1.251	0.291	0.954	22.714

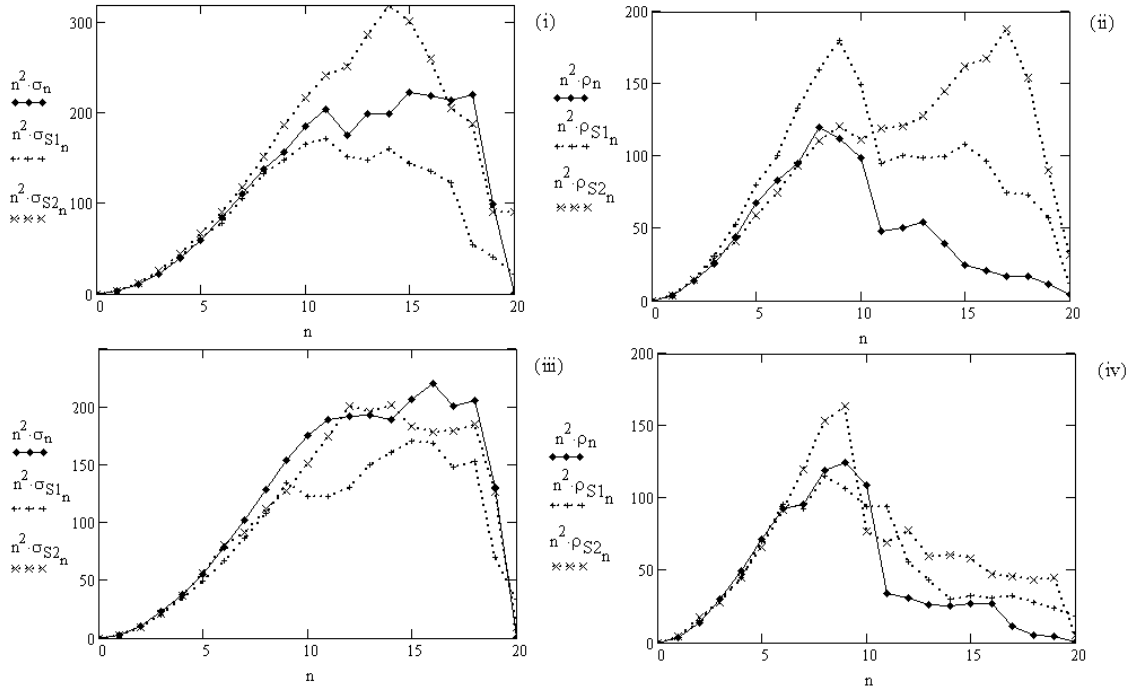

Figure 2.5: Singular value spectrum for noisy data and its surrogates for Logistic data: (i) for standard SVD $(n^2.\sigma_n)$ and its surrogates $(n^2.\sigma_{S1_n})$ and $(n^2.\sigma_{S2_n})$ and (ii) for nonlinear SVD $(n^2.\rho_n)$ and its surrogates $(n^2.\rho_{S1_n})$ and $(n^2.\rho_{S2_n})$ for the case of additive gaussian noise N(0,1) with noise level 27.484%. Similar spectra for (iii) standard SVD and (iv) nonlinear SVD spectra for the case of additive uniform noise N(0,1) with noise level 28.182%.

2.7.2 Retrieving Henon Map Parameters from Noisy Data

The estimated values of parameters of Henon map from a noisy data for both the cases of uniform and gaussian additive noises are listed in Table. 2.2. It shows the estimated parameters $\hat{a}, \hat{b}, \hat{c}$ for the Henon map $X_{n+1} = c - aX_n^2 + Y_n;\ Y_{n+1} = bX_n$ in the presence of noise using nonlinear SVD.

2.7.3 Singular Value Spectra for Noisy Data and its Surrogates

This section discusses the analysis results for the case of standard SVD and nonlinear SVD for noisy data. Fig. 2.5 shows the quadratically scaled singular spectra by standard SVD on the noisy logistic data and its surrogates along with similar spectrum by nonlinear SVD for the noisy data and its surrogates. Surrogate

data spectra are added for a comparative analysis. Fig. 2.5 (i) is the spectra of standard SVD and 2.5 (ii) is the spectra for nonlinear SVD for the case of gaussian noise in the range $(0, 1)$ added to the logistic data with a noise level 27.484%. Fig. 2.5 (iii) is the spectrum of standard SVD and 2.5 (iv) is the spectrum for nonlinear SVD for the case of uniform noise in the range $(0, 1)$ added to the same data with noise level 28.182%. Singular value spectra for two surrogates $S1$ and $S2$ are added for a comparative analysis in all the 4 cases of Fig. 2.5. Noise level is defined as the ratio of the maximum noise value to the maximum signal value. The size of the embedding matrix F was kept constant for finding the spectra for both the standard and nonlinear SVD operation.

It is clear from Fig. 2.5 that the method of nonlinear SVD is able to distinguish the original data from its surrogates under the presence of noise, provided the noise level is low. But note that if the noise level goes up the method fails to distinguish between data and surrogates.

2.8 Discussion on Linear Vs Nonlinear Models

At this point, there is a need to distinguish between two possibilities that could happen: (i) the underlying dynamics is linear but the data is contaminated with noise (ii) the underlying dynamics is nonlinear with or without noise present in the data. In both of these cases, standard SVD will give non zero singular values. For these two cases, one could explore simple nonlinear alternatives if the ratio of singular values falls below the preset criterion confirming the second possibility. But if the goal is to detect the nonlinearity but not to quantify it, one could use the method suggested by Porta *et al.* [64]. They have discussed the case of the logistic model with parameter value $\lambda = 3.7$ under various noise levels. It was shown that an error function dips much further with actual data than with its surrogates for their model. A comparison of this method with various other methods are also discussed [65].

A second order linear system $X_n = a.X_{n-1} + b.X_{n-2}$ was simulated with parameter values $a = -1.5$ and $b = -1$. For the data generated by this system in the absence of noise the linear SVD had a sharper fall off than the nonlinear SVD.

In the case of surrogate data neither linear SVD nor nonlinear SVD had a sharp fall off. In such cases the linear model would be chosen based on the grounds of parsimony. In the presence of small amount of noise the same situation continues. However when the noise is beyond a certain value neither linear nor nonlinear SVD will show a substantial qualitative difference with the surrogate data to have any degree of confidence in either of the models.

In the case of nonlinear chaotic data, one can observe from the simulations results discussed so far that the method of nonlinear SVD works reasonably well even in the presence of noise, provided the noise is below some threshold. As the noise level increases the ratio of $\frac{W_{q,q}}{W_{1,1}}$ crosses the preset threshold 10^{-6}, the confidence in the model gets reduced.

2.9 Extension of the method of Nonlinear SVD to Flows

This section demonstrates the utility of nonlinear SVD for recovering a nonlinear relationship in time series generated by continuous dynamical systems or flows. The first section gives an application of the method for mathematical modeling using the Van der Pol Oscillator. The second section demonstrates an application of the method for chaotic cryptanalysis using the Duffing oscillator.

2.9.1 Application to Mathematical Modeling

The goal of this section is to explore the possibility of nonlinear SVD for the modeling a differential equation based on experimental data. The key assumptions behind this model are: (i) the sampling is frequent (ii) noise level is very low. It is demonstrated here how the equation can be retrieved using nonlinear SVD. A method of finding accurate derivative from data [67] was used in the numerical simulation.

Assume that the time series is generated by a lower dimensional deterministic dynamical system whose exact equations are unknown. This section will illustrate the method and the issues involved by finding the unknown differential equation from the time series. The selected time series in this case– Z displayed in Fig. 2.6 (ii) is relatively short, which is clear from the state space picture depicted in Fig.

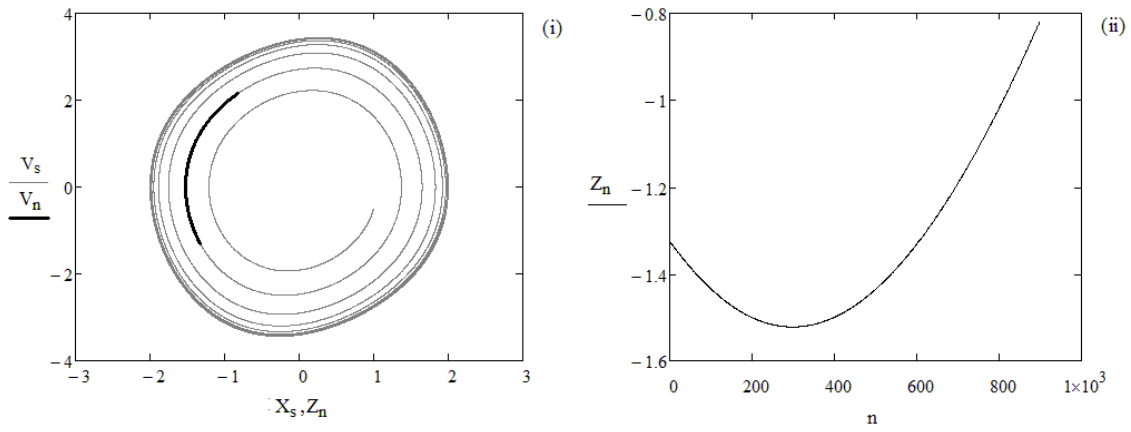

Figure 2.6: (i) Phase space of Van der Pol oscillator with a darkened portion highlighting the short data segment of data Z used for analysis and (ii) is the time series Z.

2.6 (i). In practice, one might be able to do better if a longer time series is available as some noise reduction methods might work better it. Finding a procedure for a shorter time series is a tougher challenge.

Steps of the numerical procedure to identify the system are listed below. Assume that data measurements Z. Calculate n derivatives of Z represented as $(Y_2, Y_3, \ldots Y_{n+1})$ using the method mentioned in [67]. Y_n is the $n-1^{th}$ derivative of Z, hence Y_1 (the 0^{th} derivative of Z) is same as Z; and these notations are followed for the rest of the section.

Step. 1 Given a highly sampled data Z, calculate its derivatives using the method [67].

For the numerical analysis Z data (accurate up to 10^{-16}) frequently sampled (with a step size 0.001) with a length 1000 is chosen (displayed in Fig. 2.6 (ii)).

Step. 2 Make an assumption that data is generated by a differential equation of the following form.

$$\begin{aligned} \frac{d}{dt}X_1 &= X_2. \\ \frac{d}{dt}X_2 &= G(X_1, X_2). \end{aligned}$$

where G is a relatively simple multinomial function that has to be determined

empirically.

Step. 3 Construct a candidate function $G1$ for G (based on the domain on application where data belongs to).

Let $G1$ consisting of linear and quadratic terms of the derivatives Y_1 and Y_2 be,

$$G1 = c_1Y_1 + c_2Y_2 + c_3Y_1^2 + c_4Y_1.Y_2 + c_5Y_2^2. \tag{2.18}$$

Step. 4 Retrieve the parameters $c_1, c_2, \ldots, c_5$ of Eq. 2.18 by choosing an embedding matrix of the form, $F = [Y_3\ Y_1\ Y_2\ {Y_1}^2\ {Y_2}^2\ (Y_1Y_2)]$. Find the singular values of F and check if the preset criterion $\frac{W_{q,q}}{W_{1,1}} \leq 10^{-6}$ is satisfied.

Singular values of F matrix were [201.33, 162.24, 83.95, 3.95, 0.44, 0.03] and and the ratio $\frac{W_{6,6}}{W_{1,1}} = 1.5 \times 10^{-4}$ did not satisfy the preset criterion.

Step. 5 Since the preset criterion is not satisfied, the function $G1$ is improved by adding cubic functions. Let the new function $G2$ be,

$$G1 = c_1Y_1+c_2Y_2+c_3Y_1^2+c_4Y_2^2+c_5Y_1.Y_2+c_6{Y_1}^3+c_7{Y_2}^3+c_8{Y_1}^2.Y_2+c_9{Y_2}^2.Y_1 \tag{2.19}$$

Step. 6 To retrieve the parameters $c_1, c_2, \ldots, c_9$ of Eq. 2.19 by choosing an embedding matrix of the form, $F = [Y_3\ Y_1\ Y_2\ {Y_1}^2\ {Y_2}^2\ (Y_1Y_2)\ {Y_1}^3\ {Y_2}^3\ ({Y_1}^2Y_2)\ ({Y_2}^2Y_1)]$. Find the singular values of F and check if the preset criterion $\frac{W_{q,q}}{W_{1,1}} \leq 10^{-6}$ is satisfied.

The singular values of F were (181.0635, 115.3297, 44.2941, 2.1510, 22.1215, 0.1768, 0.0247, 0.0044, 0.0003, 10^{-11}) and the ratio $\frac{W_{10,10}}{W_{1,1}}$ satisfied the preset criterion.

Step. 7 Retrieve the coefficient array for this case assuming the last singular value $W_{10,10} = 0$ (expanding the corresponding equation for Eq. 2.8).

The retrieved coefficient array was [1, 2.895, −0.237, 0, 0, 0, 0, 0, 0.237, 0].

Rewriting Eq. 2.19 in terms of the coefficients,

$$0 = Y_3 - 0.237Y_2 + 0.237Y_2Y_1^2 + 2.895Y_1. \quad (2.20)$$

$$Y_3 = 0.237Y_2 - 0.237Y_2Y_1^2 - 2.895Y_1. \quad (2.21)$$

$$Y_3 = 0.237Y_2(1 - Y_1^2) - 2.895Y_1. \quad (2.22)$$

Since Y_2 and Y_3 are the 1^{st} and 2^{nd} derivatives of Y_1, rewriting the Eq. 2.22,

$$\frac{d}{dt}Y_1 = Y_2.$$
$$\frac{d}{dt}Y_2 = cY_2(1 - Y_1^2) - kY_1.$$

This is the Van der Pol equation with parameter values $k = 2.895$ and $c = 0.237$. The estimated values $k = 2.895$ and $c = 0.237$ from the data series are in agreement with the values used in the Van der Pol to generate the data series. The form of the equation and the parameter values are exactly predicted using the proposed method as the data was noise free. In addition to the data, information about the sampling interval is necessary to calculate the derivatives using the method mentioned in [67]. Fig. 2.6 (i) shows the Phase space of the Van der Pol oscillator which was used to generate the Z data of length 1000 (shown in Fig. 2.6 (ii)) used for analysis.

2.9.2 Application to Chaotic Cryptanalysis

This section explains how the method of nonlinear SVD can be used for cryptanalysis. The key assumptions behind the model are the same as those explained in the previous section.

Consider the following narrative of cryptanalysis. Assume that a sender A is using a communication channel for sending message to a receiver B. For using Chaotic Cryptography methods there is a set of shared information across A and B known as super key (the equations and the parameter values that are used for generating data on which the message is encrypted). Assume that C an eavesdropper who does not have access to the super key but wants to retrieve it by listening to

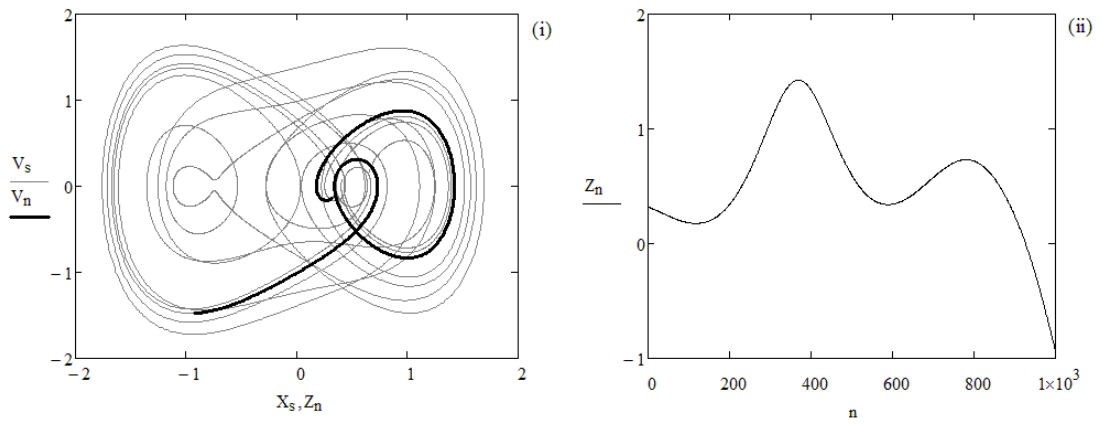

Figure 2.7: (i) Phase space of the Duffing oscillator under chaos with a darkened section highlighting the short data segment Z, and (ii) the message Z sent across the communication channel by A, that was used for cryptanalysis by C.

the messages across the communication channel.

Assume that the sender A is generating data from Duffing equation under chaos,

$$\frac{d}{dt}X_1 = X_2 \tag{2.23}$$

$$\frac{d}{dt}X_2 = -kX_1 - cX_2 - \delta {X_1}^2 + Acos(\omega t + \Phi) \tag{2.24}$$

Parameter values chosen by A are $k = 0.01$, $c = 0.04496$, $\delta = 1$, $A = 1.02$, $\omega = 0.44964$ and $\Phi = 0$. A has generated a data segment Z (accurate up to 10^{-16}) of length 1000 using the Duffing equation by Runge–Kutta method from an initial condition $(0.656, 0.172)$ with sampling step size 0.01 and was send to team B across the channel as a message. Phase space of the Duffing oscillator under chaos is shown in Fig. 2.7 (i) and the message Z sent to B is shown in Fig. 2.7 (ii).

The eavesdropper C who is listening to the channel is left with the data segment Z. The objective of C for cryptanalysis is to find the form of the equation used by A along with the parameter values. Assuming that (i) sampling frequency of the message is known to C by some other means and (ii) noise level in the message is very low, the methods used by C to fulfil the objective are given below.

Step. 1 Make an assumption that data Z is generated by a differential equation of the following form.

$$\frac{d}{dt}X_1 = X_2. \tag{2.25}$$

$$\frac{d}{dt}X_2 = G(X_1, X_2) + A.cos(\omega t + \Phi). \tag{2.26}$$

where G is a low order polynomial function.

Step. 2 Calculate n derivatives of $Z : Y_2, Y_3, \ldots Y_{n+1}$, where Y_n is the $n-1^{th}$ derivative of Z and $Z = Y_1$ using the method [67].

Step. 3 Construct a candidate function $G1$ consisting of linear, quadratic and cubic terms of Y_1 and Y_2 and sinusoidal functions.

$G1 = c_1Y_1 + c_2Y_2 + c_3Y_1^2 + c_4Y_1.Y_2 + c_5Y_2^2 + c_6{Y_1}^3 + c_7{Y_2}^3 + c_8({Y_1}^2.Y_2) + c_9({Y_2}^2.Y_1)] + c_{10}sin(\omega t) + c_{11}cos(\omega t)$.

Step. 4 Retrieve the parameters $c_1, c_2, \ldots, c_{11}$ by choosing an embedding matrix of the form,
$F = [Y_3,\ Y_1,\ Y_2,\ {Y_1}^2,\ {Y_2}^2,\ (Y_1Y_2),\ {Y_1}^3,\ {Y_2}^3,\ ({Y_1}^2Y_2),\ ({Y_2}^2Y_1), \sin(\omega t),\ cos(\omega t)]$

Step. 5 Find the singular values of F and check if the preset criterion $\frac{W_{q,q}}{W_{1,1}} \leq 10^{-6}$ is satisfied. Retrieve the coefficient array.

SVD operation of F gives a set of singular values (46.1294, 28.0900, 20.9784, 18.8628, 12.5940, 6.8899, 4.8714, 4.2835, 2.0742, 0.8881, 0.5682, 0). According to Eq. 2.8 a relationship can be recovered across the columns of F matrix since the last singular value is exactly zero.

[1, 0.01, 0.04496, 0, 0, 0, 1, 0, 0, 0, −0.98958, 0.24723] is the recovered coefficient array.

Step. 6 Retrieve the equation using the coefficient array.

$$Y_3 + 0.01Y_1 + 0.0449Y_2 + Y_1^3 - (\sqrt{0.989^2 + 0.247})cos(\omega t) = 0 \quad (2.27)$$

$$Y_3 + 0.01Y_1 + 0.0449Y_2 + Y_1^3 - 1.02cos(\omega t) = 0 \quad (2.28)$$

Since Y_2 and Y_3 are the 1^{st} and 2^{nd} derivatives of Y_1, Eq. 2.28 can be rewritten to get the Duffing equation with parameter values $k = 0.01$, $c = 0.04496$, $\delta = 1$, $A = 1.02$, $\omega = 0.44964$ and $\Phi = 0$ as,

$$\frac{d^2Y_1}{dt^2} + c\frac{dY_1}{dt} + kY_1 + \delta Y_1^3 = Acos(\omega t) \quad (2.29)$$

whose state space representation is,

$$\frac{dY_1}{dt} = Y_2 \quad (2.30)$$

$$\frac{dY_2}{dt} = -cY_2 - kY_1 - \delta Y_1^3 + Acos(\omega t) \quad (2.31)$$

Thus C was successful in cryptanalysis as the estimated parameter values and the form of the equation are in exact agreement with the Duffing equation used by A and B for communication.

2.10 Conclusions

A nonlinear extension of the singular value decomposition (SVD) technique is demonstrated in this chapter. In principle, the method can work with any type of nonlinearity. The method of nonlinear SVD is for detecting and quantifying nonlinearity in a time series is explained using data generated by discrete dynamical systems: the Logistic map and the Henon map and continuous dynamical systems: the Van der Pol oscillator and the Duffing oscillator. The proposed method works in the presence of noise, provided the noise level is not high. Recovery of parameters is demonstrated for both the noise–free and noisy cases. Two specific applications of the method are discussed. Next chapter discusses an extension of the nonlinear SVD method for finding the dimension of the manifold where data resides.

CHAPTER 3
Nonlinear SVD for estimating the Dimension of a Manifold and finding Local Charts

The goal of this chapter is to propose a new method based on nonlinear Singular Value Decomposition (SVD) to arrive at an upper bound for the dimension of a manifold which is embedded in some R^N. Assume that some data on the manifold is available and also there exists at least one small neighborhood with sufficient number of data points. Given these conditions, this chapter shows a method to compute the dimension of a manifold. Two specific cases are discussed in this chapter. For the simple case where the manifold is in the form of a lower dimensional affine subspace, standard SVD is used to (i) calculate the dimension of the manifold and (ii) to get the equations which define the subspace. For the general case of manifolds, nonlinear SVD is used (i) to search for an upper bound for the dimension of the manifold and (ii) to find the equations for the local charts of the manifold. The proposed method is highly useful in the context of delay embedding– a technique that is used through out this thesis for data analysis. Finding a good estimate of the underlying dimension of an embedded data is a requirement while modeling a system based on local charts. Following chapters of the thesis provides examples of this type of modeling.

(Note: Research embodied in this chapter is published and [71] is the reference for that publication)

3.1 Introduction

Embedding based techniques are extremely useful in analyzing data as they can give qualitative information of the dynamics of the system that generated the data. Data generated by nonlinear chaotic systems usually live in a lower dimensional manifold, unlike the data from stochastic systems that tend to spread over the whole phase space. This property of nonlinear chaotic systems can be exploited for modeling and prediction of nonlinear chaotic signals.

Embedding is defined as a smooth map Φ from a manifold M to a space U such that its image $\Phi(M) \subset U$ is a smooth manifold of U and Φ is a diffeomorphism between M and $\Phi(M)$ [5]. The existence theorem for embeddings in Euclidean spaces was given by Whitney [24]. He proved that a smooth d–dimensional manifold can be embedded in R^n if $n > 2d$. In 1980's Packard *et al.* reported that phase portraits similar to the underlying dynamical system could be reconstructed from measurements [16]. Later Takens gave a firm theoretical foundation for the reconstruction techniques based on measurements known as delay embedding which soon became a popular method for analyzing dynamical systems based on time series measurements [25].

Broomhead and King proposed an alternative method 'SVD–based embedding' to get the best orthogonal basis for the data [12]. SVD–based embeddings have some advantages as the set of linearly independent orthogonal vectors given by SVD is used to form a local basis of the embedded manifold $\Phi(M)$. SVD–based embedding can be used for noise reduction by partitioning the embedding space and rejecting the out of band noise– the less significant partition of the embedding space. In this context of modeling realtime signals, different local nonlinear prediction(LNP) methods were proposed that used embedding based techniques. Porta *et al.* proposed an LNP method of prediction that used delay embedding of data in a higher dimensional phase space [64, 65]. The phase space was subdivided into non–overlapping hypercubes and the prediction was based on the behavior of the median of the values of the past samples that belonged to the same hypercube.

One goal of modeling a system based on time series is to develop equations for the system from the time series [18, 19] for prediction or a description of the dynamics [20, 21]. These type of modeling methods generally fall into two categories: Global methods that find equations valid for the entire statespace, and Atlas methods that develop local charts for small neighborhoods of the statespace [22, 23]. In general the dimension required for the global embedding of a dynamical system is higher than the intrinsic dimension of the system. According to Whitney and Takens, global embedding of a d–dimensional manifold requires a minimum dimension $2d+1$ [24, 25]. But if the aim is to use Atlas methods to find local charts that cover

the statespace, it might be possible to represent the system in just d–dimensions. This chapter shows a method to compute the dimension of a manifold and to find equations for its charts using the method of nonlinear SVD explained in the previous chapter. The main application of this method is in the context of an analysis of a dynamical system from observed data that is embedded in some R^N. Chapter. 5 on ECG data analysis uses the results of this chapter in this context.

Consider an example of a geometrical object. Let the assumptions on the data available about this object are: (i) the manifold has already been embedded in some R^N and the data about its co-ordinates is available, (ii) there exists at least one small neighborhood with sufficient number of data points. For the case of dynamical systems, this requires the existence of a property known as Recurrence [30]. A point in the state space is said to be recurrent if a time series generated by the system keeps on visiting the neighborhood of that point. The method demonstrated in the chapter is useful for data generated by dynamical systems that exhibit the property of recurrence.

Two specific cases are discussed in this chapter. For the simple first case, when the manifold is in the form of a lower dimensional affine subspace, the technique of SVD can be used to (i) calculate the dimension of the manifold and (ii) to get the equations which define the subspace as shown in Section 3.2. Section 3.2.1 discusses a numerical example of finding the dimension a subspace embedded in R^4. For the general second case, the method of nonlinear SVD is used to (i) to search for an upper bound for the dimension of the manifold and (ii) to find the equations for the local charts of the manifold as shown in Section 3.3. Section 3.3.1 shows a numerical example of finding the equation for a local neighborhood on a manifold using the co-ordinate data of möbius strip. Chapter concludes in Section 3.4 and it contains a brief discussion about how the proposed method would be useful in the context of delay embedding.

3.2 The case of an M dimensional Subspace embedded in R^N where $N > M$

Consider a manifold that is embedded in R^N, and assume that the coordinates of a large number of points which belong to this manifold are available as a numerical data. This section discusses a numerical technique to investigate whether the data belongs to a manifold which might have a dimension less than N. This section looks at a simpler case when the manifold is an affine subspace of R^N. Next section will look at the general case of an M dimensional manifold embedded in R^N. In any case, given the data, it would always be worthwhile to first check if an affine subspace is an adequate description of the manifold under consideration.

The first step to check for the possibility of the existence of an affine subspace is to locate the centroid of the data and reset it as a new origin. Trivially if the data belonged to an affine subspace this centroid would also belong to the same subspace. Further, if the centroid is set as the new origin, one would get a linear vector subspace if and only if the original data belonged to an affine subspace. This reduces the investigation to check if the modified data belongs to a linear subspace. In practice the method of SVD is quite suited to do this task.

Recalling the SVD theorem mentioned in Section 2.2 of Chapter 2, a matrix B of size $(P \times N)$ can be decomposed into three matrices as

$$B = UWV^T \tag{3.1}$$

where U, a column orthogonal matrix of size $(P \times N)$; W, a diagonal square matrix of size $(N \times N)$ and V, an orthogonal square matrix of size $(N \times N)$. V^T represents transpose of the matrix V and the diagonal entries of W are called the singular values of B.

To use this theorem, construct a matrix D of size $(P \times N)$ whose rows represent various data points and the columns correspond to various coordinates in R^N. For practical cases, the total number of points $P \geq N$ (usually $P \gg N$), as having more data is better for dealing with the noise in the data. Thus the entry in the i^{th} row and j^{th} column would be the value of the j^{th} co-ordinate of the i^{th} data point. If

$x_1, x_2, \ldots x_N$ are the variables representing the N coordinates of R^N, $D_{i,j}$ would be the value of x_j at i^{th} observation point.

Find the centroid of all the data points represented by D. The n^{th} coordinate of the centroid is given by,

$$d_n = \frac{1}{P}\sum_{p=1}^{P} D_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{3.2}$$

Set the centroid as origin by creating a new matrix A such that,

$$A_{p,n} = D_{p,n} - d_n \qquad \text{for all p} \tag{3.3}$$

Using the result of SVD given in Eq. 3.1 two theorems can be proven. For these theorems, the matrices W and V are computed from the data matrix using standard numerical procedures [10].

Theorem 3.1. *If the SVD of the matrix A defined above gives Q singular values which are zero then there exists an $N-Q$ dimensional affine subspace on which the data resides and the equations defining the subspace are given as follows,*

$$\sum_{n=1}^{N} V_{n,q}(Z_n - d_n) = 0; \qquad \text{for q} = (\text{N} - \text{Q} + 1) \ldots \text{N} \tag{3.4}$$

where, V is the orthogonal matrix we got after SVD. Z represents the co-ordinates of R^N and d_n is the centroid of the data points.

Theorem 3.2. *If the matrix A defined above, represents a collection of points which lie on a linear subspace of dimension $N-Q$ which is defined by Q independent linear algebraic homogenous equations, then the SVD of A will have at least Q singular values which are zero.*

Proof of the Theorem 3.1 is given in this chapter. (Refer Appendix A for the proof of Theorem 3.2. Theorem 3.2 is not needed to explain the results discussed in this chapter.)

Proof of Theorem 3.1

SVD of matrix A,

$$A = UWV^T \tag{3.5}$$

By post multiplying Eq. 3.5 by V,

$$AV = UW \tag{3.6}$$

Upon transposing,

$$V^T A^T = WU^T \tag{3.7}$$

Note: $W^T = W$ as W is a diagonal square matrix. Expanding Eq. 3.7 for any column p,

$$\sum_{n=1}^{N} V_{m,n}^T A_{n,p}^T = \sum_{n=1}^{N} W_{m,n} U_{n,p}^T \tag{3.8}$$

Define $Y_n(p) = A_{n,p}^T + d_n$ for all p, given that $Y(p)$ is the original data vector at p^{th} data point. Therefore by using Eq. 3.8, all the data vectors (from the original data, before the correction for centroid was applied) will fit the following equation if Q singular values are zero:

$$\sum_{n=1}^{N} V_{n,q}(Z_n - d_n) = 0 \qquad \text{for q} = (\text{N} - \text{Q} + 1) \ldots \text{N} \tag{3.9}$$

3.2.1 Numerical Example of a Linear Subspace

This section explains the results of Theorem 3.1 using a numerical example. Consider a set of 100 samples each of two random variables $g1$ and $g2$. Both were selected using a program to find uniform random numbers in the interval $[0, 1]$. From each of these pairs, coordinates for points in R^4 were selected by using linear transformations. These coordinates became the column entries of the A matrix. The first column of this matrix, representing a coordinate $S1$ consisted of $0.2g1 + 0.4g2$; second column $S2$ consisted of $0.3g1 + 0.5g2$; third column $S3$ consisted of $0.5g1 + 0.6g2$; and the last column $S4$ consisted of $0.7g1 + 0.1g2$. Singular values of

the data matrix A (after the removal of the column means d_n) were,

$$\begin{bmatrix} 3.29739 \\ 1.32848 \\ 0 \\ 0 \end{bmatrix}$$

The mean vector d was,

$$\begin{bmatrix} 0.2925 \\ 0.39093 \\ 0.53994 \\ 0.40198 \end{bmatrix}$$

The V matrix was,

$$\begin{bmatrix} -0.35101 & -0.33214 & -0.87549 & 0 \\ -0.46934 & -0.34316 & 0.31836 & -0.74874 \\ -0.64884 & -0.21017 & 0.33987 & 0.64756 \\ -0.48531 & 0.85308 & -0.12906 & -0.14165 \end{bmatrix}$$

Since the last 2 singular values were zeros, expanding Eq. 3.9 using the last 2 columns of this V matrix,

$$\sum_{n=1}^{4} V_{n,3}(Z_n - d_n) = 0 \tag{3.10}$$

$$\sum_{n=1}^{4} V_{n,4}(Z_n - d_n) = 0 \tag{3.11}$$

Expanding Eq. 3.10 and 3.11 in terms of the coordinates $(S1, S2, S3, S4)$,
$-0.74874(S2 - 0.39093) + 0.64756(S3 - 0.53994) - 0.14165(S4 - 0.40198) = 0$
and
$-0.87549(S1-0.2925)+0.31836(S2-0.39093)+0.33987(S3-0.53994)-0.12906(S4-0.40198) = 0$
Re-write these equations to get the coordinates $S1$ and $S2$ in terms of $S3$ and $S4$

as follows,

$$\begin{aligned} S1 &= 0.7027S3 - 0.07862S4 - 0.05531. \\ S2 &= 0.86487S3 - 0.18918S4 - 0.00001. \end{aligned} \tag{3.12}$$

Consider a new space R^2 given by co-ordinates $(T1, T2)$ where $T1 = S3$, $T2 = S4$. The following diffeomorphisms across R^4 and R^2 define the local chart co-ordinates:

$$\begin{bmatrix} S1 \\ S2 \\ S3 \\ S4 \end{bmatrix} = \begin{bmatrix} 0.7027 & -0.07862 \\ 0.86487 & -0.18918 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} T1 \\ T2 \end{bmatrix} + \begin{bmatrix} -0.05531 \\ -0.00001 \\ 0 \\ 0 \end{bmatrix} \tag{3.13}$$

and the inverse is,

$$\begin{bmatrix} T1 \\ T2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} S1 \\ S2 \\ S3 \\ S4 \end{bmatrix} \tag{3.14}$$

Thus $(T1, T2)$ define local coordinates of a two dimensional manifold which contains the data. In this simple case, they also turn out to be the global co-ordinates. This creates an upper bound on the dimension of the manifold as 2. This is only an upper bound because there may be a submanifold which contains the data.

3.3 Finding Nonlinear Equations for Charts on a Manifold using Nonlinear SVD

The method discussed in Section 3.2, using the conventional method of SVD may not work for many of the practical cases for the reason that the data may not be confined to a linear subspace. For such situations, the nonlinear SVD procedure discussed in Chapter 2 can be used as shown in this section. An assumption is made that a neighborhood in the manifold of the system might confine to an affine

subspace in the $N + K$ dimensional space.

Recall that every row of the original data matrix D consisted of the observed values of the N co-ordinates. According to the nonlinear SVD procedure, augment D matrix by K additional columns of nonlinear functions of the N variables [52]. If the k^{th} candidate function is denoted by $f_k(x_1, x_2, \ldots x_N)$, the augmented matrix E can be represented as,

$$\begin{aligned} E_{p,n} &= D_{p,n}; \qquad \text{for n} = 1, 2 \ldots \text{N} \\ E_{p,N+k} &= f_k(D_{p,1}, D_{p,2}, \ldots D_{p,N}); \qquad \text{for k} = 1, 2 \ldots \text{K} \end{aligned}$$

This is equivalent to embedding the N dimensional system in a higher dimension $N + K$. Though the system is nonlinear in N dimensions, it happens many times that the system may be linear in the higher dimension $N + K$, if the choice of candidate functions is successful. The physical situation of the problem might suggest some specific choice of trial functions and hence the choice of functions are dependent on the domain of application. Often the trial functions are polynomial combinations of the variables.

Following the same idea used in section 3.2 for constructing the matrix A from the matrix D, construct a matrix F from the matrix E by removing the centroid. In addition to the mean vector d using Eq. 3.2, define a mean vector g for the additional columns as,

$$g_k = \frac{1}{P} \sum_{p=1}^{P} E_{p,N+k} \qquad \text{for k} = 1, 2 \ldots \text{K} \tag{3.15}$$

The modified F matrix is given by,

$$F_{p,n} = E_{p,n} - d_n; \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{3.16}$$

$$F_{p,N+k} = E_{p,N+k} - g_k; \qquad \text{for k} = 1, 2 \ldots \text{K} \tag{3.17}$$

If SVD of the F matrix gives Q singular values that are zero, one can get Q equations using a slightly modified formula in the extended $N + K$ dimensional

space as,

$$\sum_{n=1}^{N+K} V_{n,q} F_{n,q}^{T} = 0; \qquad \text{for q} = (\text{N} - \text{Q} + 1) \ldots \text{N} \tag{3.18}$$

where V is an orthogonal matrix that has dimension $(N + K) \times (N + K)$. Splitting the F matrix, this equation can be rewritten as,

$$\sum_{n=1}^{N} V_{n,q}(Z_n - d_n) + \sum_{k=N+1}^{N+K} V_{k,q}(f_k - g_k) = 0 \quad \text{for q} = (\text{N} - \text{Q} + 1) \ldots \text{N} \tag{3.19}$$

The data vectors that constitute the matrix D will obey Eq. 3.19 at all the points if the singular values are exactly zero. As in the linear case, if Q such equations exists for the set Eq. 3.19, one can conclude that the manifold is of at the most $N - Q$ dimension.

3.3.1 Numerical example: Möbius Strip embedded in R^3

This section explains the results of Section 3.3 using a numerical example of möbius strip. The conventional method of SVD will not work for a practical case as the data may not be confined to a linear subspace. For such situations, the nonlinear SVD procedure can be used to find the coordinates of the local chart. Consider the data generated on a möbius strip that is represented by the following parametrization,

$$\begin{aligned} x(u,v) &= (1 + v.cos\frac{u}{2})cos(u) \\ y(u,v) &= (1 + v.cos\frac{u}{2})sin(u) \\ z(u,v) &= v.sin(u) \end{aligned} \tag{3.20}$$

where $0 \leq u < 2\pi$ and $-0.3 \leq v \leq 0.3$. The parameter u runs around the strip while v moves from one edge to the other. The möbius strip simulated in R^3 is shown in Fig. 3.1 and its projection in R^2 can be seen in Fig. 3.2.

Assume that the data is generated the form of co-ordinates (x, y, z) in R^3 using the set of equations given by Eq. 3.20. One need to find a small neighborhood in

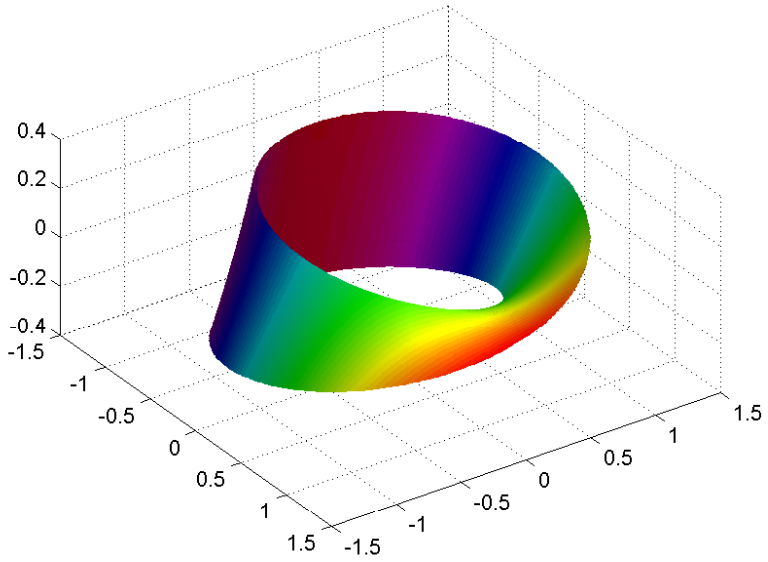

Figure 3.1: Möbius Strip in R^3 plotted using (x, y, z) variables generated using Eq. 3.20.

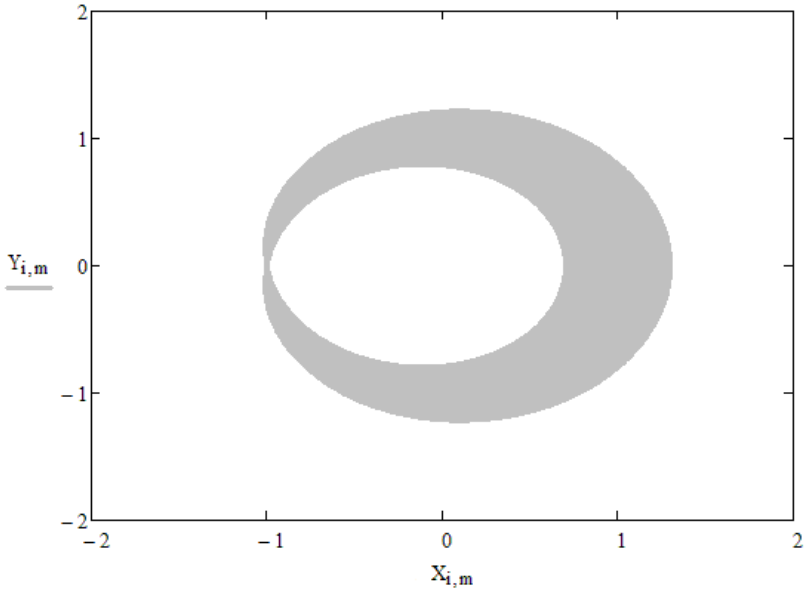

Figure 3.2: Möbius Strip in R^2 plotted using (x, y) variables generated using Eq. 3.20.

the manifold where sufficient number of data points are available from these data.

Identification of a Recurrence Neighborhood from a Time Series

A point in state space is said to be recurrent if the time series generated by the system keeps on visiting the neighborhood of the point [30]. One can identify a neighborhood for some recurrent point from a time series as follows: Start with an initial reference point and find its evolution in the state space. For the neighborhood of the reference point, the Euler metric– L_2 norm of the difference vector (the distance of the reference point with respect to all other vectors) is defined to be less than some threshold value Δ. For the numerical simulation explained below $(0.7, 0, 0)$ was chosen as the reference point and the threshold $\Delta = 0.032$ (refer section 4.9 of Chapter 4 for an algorithm to find a recurrence neighborhood).

Collect P data vectors belonging to the neighborhood of the m*ö*bius strip in R^3 into a matrix D, thus D has a dimension $P \times 3$. For the numerical simulation explained in this section $P = 43$. Create the matrix A from matrix D by removing the column averages of the data points from D as shown in Eq. 3.3. The singular values of the A matrix that represented the selected neighborhood were,

$$\begin{bmatrix} 0.352 \\ 0.01 \\ 2.286 \times 10^{-5} \end{bmatrix}$$

The low third value is encouraging but to improve the accuracy, the coordinates were embedded in R^9 (from R^3) by using nonlinear functions of the x, y, z co-ordinates of the data points (for this particular demonstration the nonlinear functions were limited to the quadratics). Hence, the extended data matrix E was created from D matrix by augmenting the trial functions x^2, y^2, z^2, xy, xz, yz to the D matrix. Now the E matrix can be denoted as,

$$E = \begin{bmatrix} x & y & z & x^2 & y^2 & z^2 & xy & xz & yz \end{bmatrix}.$$

Note that E has dimension $P \times 9$ where P is the number of data points in the

neighborhood. Next step is to generate the matrix F from E by removing the mean vector d and g from E as explained in Eq. 3.16 and 3.17.

The mean vector d was,

$$\begin{bmatrix} 0.694 \\ 0.086 \\ -0.018 \end{bmatrix}$$

The mean vector g was,

$$\begin{bmatrix} 0.481 \\ 0.01 \\ 4.667 \times 10^{-4} \\ 0.059 \\ -0.013 \\ -2.172 \times 10^{-3} \end{bmatrix}$$

Singular values of F matrix were,

$$\begin{bmatrix} 0.43287 \\ 0.024 \\ 4.02788 \times 10^{-4} \\ 8.31511 \times 10^{-6} \\ 2.43449 \times 10^{-8} \\ 1.28177 \times 10^{-10} \\ 2.19759 \times 10^{-12} \\ 9.08467 \times 10^{-14} \\ 0 \end{bmatrix}$$

Since the 9^{th} singular value is zero, the following relationship between the coordinates can be developed using the Eq. 3.19,

$$\sum_{n=1}^{3} V_{n,9}(Z_n - d_n) + \sum_{k=4}^{9} V_{k,9}(f_k - g_k) = 0 \tag{3.21}$$

Expanding this, an equation for the local neighborhood on manifold can be

obtained as,

$0.05(x-d_1)+(1.185\times10^{-4})(y-d_2)+(4.378\times10^{-4})(z-d_3)+0.334(x^2-g_1)+0.319(y^2-g_2)-0.855(z^2-g_3)+(5.139\times10^{-5})(xy-g_4)-(4.046\times10^{-4})(xz-g_5)-0.231(yz-g_6)=0$

Substituting the mean values for d and g, the exact equation for the manifold is, $0.05x+(1.185\times10^{-4})y+(4.378\times10^{-4})z+0.334x^2+0.319y^2-0.855z^2+(5.139\times10^{-5})xy-(4.046\times10^{-4})xz-0.231yz-0.198=0$.

This equation implies that the manifold is at the most 2 dimensional. One can choose y and z coordinates and create a chart that goes from (y,z) to (x,y,z). For the equation mentioned above, at every point (y,z) one can get a quadratic equation of the form $(\alpha x+\beta x^2+\gamma)=0$ for the neighborhood on the manifold where,

$$\begin{aligned}\alpha &= 0.05+(5.139\times10^{-5})y-(4.046\times10^{-4})z\\ \beta &= 0.334\end{aligned} \tag{3.22}$$

$\gamma=(1.185\times10^{-4})y+(4.378\times10^{-4})z+0.319y^2-0.855z^2-0.231yz-0.198.$

Solution of this quadratic equation gives the value of the x coordinate. Fig. 3.3 (i) shows the x coordinate of a data segment in the neighborhood along with its prediction using the chart co-ordinates (y,z). Fig. 3.3 (ii) shows the error between predicted and original x data. (Refer Appendix. B for a verification of the procedure of nonlinear SVD to find explicit nonlinear equations for charts on the manifold).

3.4 Conclusions

This chapter demonstrated how to find the dimension of a manifold when it is in the form of a lower dimensional linear subspace using the SVD method. The technique of SVD was used to calculate the dimension of the manifold and further to get the linear equations which define the linear subspace where data resides. For the general case of manifolds the method of nonlinear SVD was used to find the dimension of the manifold and the equation of the local chart.

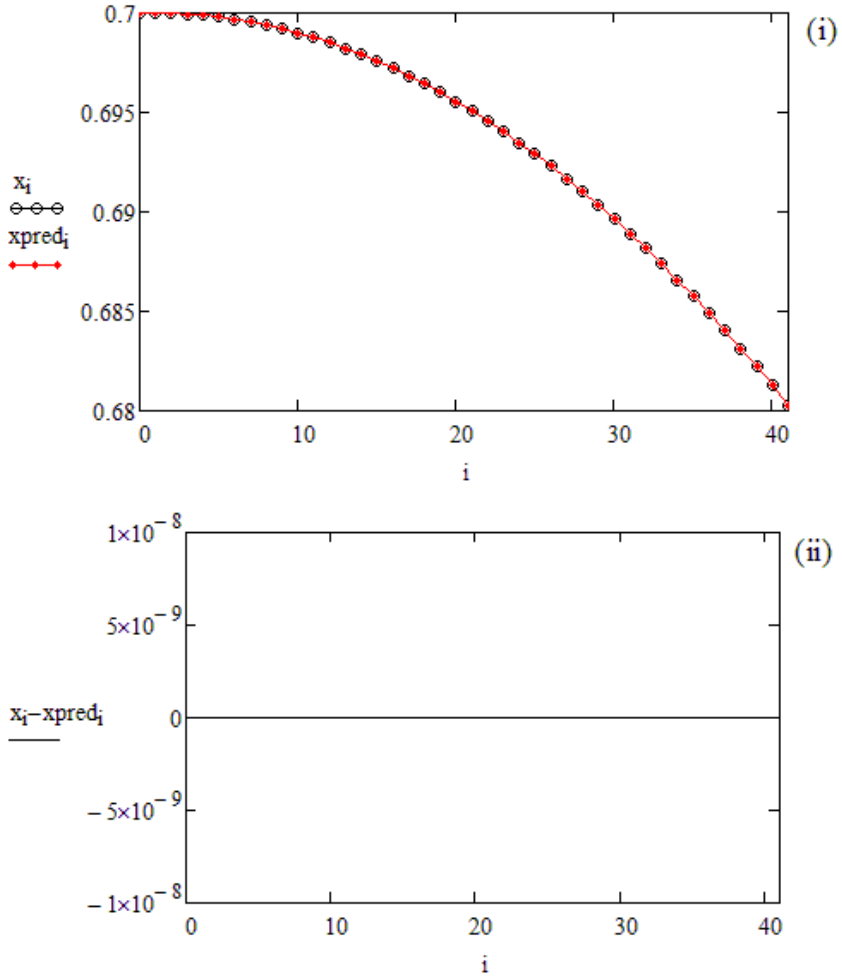

Figure 3.3: Prediction of x coordinate of a data segment in the neighborhood using the chart co-ordinates (y, z): (i) shows x data and its prediction $xpred$ (ii) shows the error between the x and $xpred$.

Embedding a data in a higher dimensional space is a standard technique in nonlinear dynamics to gain some insights about the system that generated the data [16, 5]. Embedding theorems [25, 24] gives the conditions under which a dynamical system can be reconstructed from a time series generated by it. This type of reconstruction is diffeomorphic to the original dynamics and it preserves the properties of the dynamical system and it does not change under smooth coordinate changes. To be sure of the embedding, the minimum embedding dimension is

$(2d+1)$ where d is the dimension of the manifold on which the dynamics resides. It is a common practice to embed the data in to space whose dimension is higher than the actual required value. However, there are a few problems associated with high dimensional model that represent lower dimensional dynamics [26]. When the data is actually generated by a nonlinear dynamical system of dimension d, it would occupy a d dimensional submanifold in a high dimensional embedded space R^N ($d << N$). Then one could find a large number of competing dynamical systems which agree on the d dimensions but disagree on the $N-d$ dimensions. The best way to resolve this problem is to carry out the modeling using the atlas method and develop local charts in the lower dimension d itself.

The method described in this chapter is useful to find the dimension of the manifold and the equations for the local charts based on a numerical data. Further, the method is useful in dealing with some issues involving the modeling and the stability analysis of systems based on a time series data. The issues related to over embedding– the delay embedding in a dimension much higher than required– will be discussed in detail in the next chapter. One effective solution to the problems generated by over embedding is to develop a model which is based on local charts and for this purpose a good estimate of the underlying dimension of an embedded data is required. Chapter 5 provides an example of such an application.

CHAPTER 4
Prediction based on Recurrence Neighborhoods

This chapter introduces a new method of prediction and modeling that exploits the property of recurrence in a dynamical system. According to this model the dynamics of the system can be approximated by a few overlapping recurrence neighborhoods with specific affine transformations for each of them. The model wants to exploit the inherent redundancy structure of the delay embedding for the purpose of reducing computational load which is inevitable in nonlinear analysis. Global analysis of the system is done in a higher dimensional space for this purpose. Local analysis of the system is done in a lower dimensional space using charts specific to the recurrence neighborhoods. It is observed that the recurrence timings– the time delay between two adjacent recurrences– reveal lot of information about the system. This property of recurrence when coupled with the redundancy of delay embedding opens up a new area of prediction for the special class of signals that share the property of recurrence. According to the model, if the recurrence timings and the affine transformations specific to the neighborhoods are known, the entire dynamics of the system can be predicted. The goal of this chapter is to demonstrate the model using data generated by a well studied dynamical system: the Duffing oscillator.

4.1 Introduction

Ideas discussed in this chapter focus on a particular class of dynamical systems that share a property known as Recurrence. Recurrence in general can be seen as a repetition of patterns in a time series generated by a system. A dynamical system is said to have the property of recurrence if a time series generated by the system keeps on visiting the neighborhood of a point. Mathematical theories related to recurrence were developed in 1987 by Eckmann *et al.* [30]. Analyzing recurrence patterns are important in the study of dynamical systems and finding them in a time series is a prominent method for data analysis [31, 32]. Owing to the rapid growth in the theory of nonlinear dynamics in the last three decades, the property of recurrence has

found applications in numerous fields as physiological data assessment, information theory, economics and biomedical time series analysis [33, 34, 35, 36].

Recall another important property of dynamical systems which was observed by Packard *et al.* when they could reconstruct the state–space of a system from a time series generated by the system [16]. Takens in 1981 proved a theorem according to which the state space of a system can be reconstructed in a high dimension using time–delayed copies of a measured variable, known as delay embedding [25]. There is a one–one similarity between the reconstructed space in the delay embedding and the original state space of the system.

The proposed model is based on a few ideas from Topology, Nonlinear Dynamics and Chaos theory– (i) recurrence patterns of a data series reveal lot of information about the system that generated the data (ii) delay embedding matrices have inherent symmetries and a redundancy structure. The model demonstrates that the best way to exploit the recurrence property of a signal for prediction is to couple it with the delay embedding procedure. The objective of this chapter is to introduce a novel method of modeling by coupling the properties (i) and (ii) to reduce the computational load that is inevitable in nonlinear signal analysis. The possibilities of the model are explored for a chaotic data generated by Duffing oscillator. The prediction results are excellent and it is demonstrated with duffing data. Chapters 5 and 7 demonstrates applications of the model for biomedical signal analysis.

In what follows, section 4.2 discusses the methodology– the proposed new model and the key ideas on which the model is built. Section 4.3 lists the details of the dynamical system under study. An algorithm for finding the recurrence timings of a time series and a neighborhood of a recurrent point in a delay embedding space is given in this section. Section 4.4 discusses the details of the transformation used for projection and the dynamics of the neighborhoods in both the high dimensional space R^N and the lower dimension R^d. It shows how to develop (i) conjugacy maps across R^N and R^d and (ii) evolution maps for various equivalence classes in R^d. Results of the data analysis are given in Section 4.5. Section 4.6 is the conclusion and it also has a discussion about the empirical values chosen for the projected dimension d.

4.2 Methodology

The proposed new model and the key ideas on which the model is built are discussed in this section. An important property of differential equations that model a given dynamics is that they generate a set of maps known as one–parameter group of diffeomorphisms. This property is reviewed in next section.

4.2.1 Affine Approximations in small Neighborhoods for a Dynamical System

A Differential equation generates maps at every time of its evolution. Consider a differential equation of the form $\dot{x} = f(x)$ that is defined on a manifold $M \subset R^n$. The flow of this system can be generated by the set of maps $\phi_t(x) = \phi(x,t)$ for every $x \in M$ and $t \in R$. The set of differentiable transformations $\phi_t : M \to M$ for every $t \in R$ such that $\phi_s(\phi_r) = \phi_{s+r} = \phi_{r+s}$ form a one–parameter group of diffeomorphisms [47]; where ϕ_0 is the identity map of M as $\phi_0(x) = x$ for every $x \in M$ and each map ϕ_t has a differentiable inverse ϕ_{-t} such that $\phi_t(\phi_{-t}) = \phi_{-t}(\phi_t) = \phi_0$. Hence given a set of initial conditions on the manifold M, its evolution after a specific time τ can be given by the specific map ϕ_τ.

Consider a point $x_1(0)$ which is recurrent with a recurrent time τ. Its evolution $x_1(\tau)$ belongs to a small neighborhood around $x_1(0)$. Let ϕ_τ be the transformation that represents the evolution. Even if ϕ_τ is nonlinear, it has an affine form for this small neighborhood. This follows from the properties of Taylor series expansion of analytic maps (refer Appendix C for the theorem and its proof). Hence one can simplify the dynamics in terms of affine transformations in small neighborhoods everywhere. This gives a possibility of using a set of overlapping neighborhoods with specific affine transformations for each of them to represent the local dynamics. Recalling the Hartman-Grobman theorem that the behaviour of a nonlinear system near an equilibrium point is qualitatively similar to the behaviour of a linear system [49], this idea can be seen as a generalization of this theorem. The dynamics can be simplified in terms of affine transformations in small neighborhoods everywhere and linear transformations in small neighborhoods at rest points. Since the affine map becomes a linear map at the rest point, this hypothesis reduces to the Hartman-

Grobman theorem at the rest point.

According to the proposed model, the affine maps specific to the neighborhoods can be empirically determined from a time series. The maps are developed in terms of the coordinates of the delay embedding to exploit the symmetries of the delay embedding for the purpose of modeling and prediction. The redundancy structure of the delay embedding matrix is discussed in next section.

4.2.2 Inherent Symmetries of the Delay Embedding Matrix

Takens established the delay embedding theorem which proved a one–to–one correspondence between the states of the system and the delay vectors constructed from the time series generated by the system [25]. According to this theorem, the reconstructed space in some higher dimension using time–delayed copies of a measured variable and the original state space of the system are diffeomorphic to each other as long as the embedding dimension is greater than twice the dimension of an underlying dynamics. If d is the dimension of the original dynamics and N is the embedding dimension for reconstruction, then the reconstructed system is said to be diffeomorphic to the original dynamical system for $N > 2d$.

There is a special structure to the delay embedding matrix as it contains lot of redundant information. Delay vectors that constitute the embedding matrix are interconnected to each other. Due to this property, intermediate vectors can be predicted from corresponding end vectors. If one knows where the initial and final vectors go, the evolution of the intermediate columns can be predicted. This property is demonstrated below for a 4×4 delay embedding matrix.

Consider a short segment of time series $x_0, x_1, ..x_n$ which has the property of recurrence is embedded in R^4 as shown below.

$$\begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_5 \\ x_3 & x_4 & x_5 & x_6 \end{bmatrix} \tag{4.1}$$

Let these x vectors evolve to y vectors after one recurrence cycle as,

$$\begin{bmatrix} y_0 & y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 & y_4 \\ y_2 & y_3 & y_4 & y_5 \\ y_3 & y_4 & y_5 & y_6 \end{bmatrix} \tag{4.2}$$

Assuming that evolution of the first and last vectors at the end of one cycle is known,

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \longrightarrow \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{4.3}$$

$$\begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \longrightarrow \begin{bmatrix} y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} \tag{4.4}$$

evolution of the middle vectors of the embedding matrix can be predicted from the evolution of end vectors. One can write the 3^d column as a sum of linearly transformed end vectors –1^{st} and 4^{th} columns as shown below.

$$\begin{bmatrix} y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} \tag{4.5}$$

This relation is true even though the delay vectors belong to a nonlinear time series. Since the column vectors of the embedding matrices are connected to each other, evolution of one of the vectors can be written as the sum of a few linear transformed evolution of other delay vectors as shown by Eq. 4.5. The aim of the model is to exploit this property of the delay embedding to reduce the computational load that is inevitable in nonlinear data analysis.

One goal of modeling a system based on a time series is to develop equations

from a given time series [18, 19] for prediction or a description of the dynamics [20, 21]. These type of modeling methods generally fall into two categories: Global methods that find equations valid for the entire statespace, and Atlas methods that develop local charts for small neighborhoods of the statespace [22, 23]. According to the proposed model, the delay embedding in a higher dimensional space is used as a scaffolding for the model to analyze the global structure of the dynamics. The delay embedding is followed by a projection to a lower dimension (which is closer to the dimension of the actual dynamics) for the local analysis of the dynamics. This projection will take care of two fundamental issues related to high dimensional models that describe a low dimensional dynamics: (i) the ambiguity that many models fit the data (ii) the fact that the models become less economical, requiring more amount of data as the embedding dimension goes up. Both of these issues are briefly reviewed in next section.

4.2.3 Issues of High Dimensional Models that describe a Low Dimensional Dynamics

Global embedding of nonlinear systems often require high embedding dimensions. When the data already belong to a projective space of the actual dynamics, a delay embedding of the data in a higher dimension opens up a possibility of numerous models fitting the data. If the data is generated by a non–linear dynamical system of dimension d, it would occupy a d dimensional sub manifold in the N–dimensional embedded space. One could have a large number of competing dynamical systems which will fit the data on d dimensions but disagree on the $(N-d)$ dimensions. (An example that demonstrate this ambiguity is discussed in Appendix D.) This will result in an ambiguity associated with high dimensional models and it can cause spurious inferences about the stability when the system is represented in a dimension higher than the required value [26].

Other of issues reported with higher dimensional models that represents a lower dimensional dynamics are: (i) increase in model dimension that expands the effect of noise (ii) use of higher order polynomials in the model equation that increases the number of coefficients to be estimated (iii) some orbits of the model statespace that

goes outside the region containing the time series, thus making the model no longer connected with the object under study [27]. In addition to taking care of the reported issues, there is also a practical reason why one prefers a lower dimensional model. A model in a lower dimensional space is economical in terms of data compared to its high dimensional counterpart.

The key idea proposed here for a time series that exhibit the property of recurrence is to represent the dynamics in a small neighborhood by an affine form. The existence of an affine form implies the possibility of a linear model $X \rightarrow AX$ with respect to a neighborhood. Now the model depends upon the property of the A matrix that is unique to that neighborhood. Empirical determination of A will require enormous amounts of data if the neighborhood belongs to a high dimensional space. If the neighborhood belongs to a delay embedding space R^N where N is much higher than the actual dynamics, then N^2 parameters are needed to be estimated for the A matrix which require at least N^2 neighborhood vectors. Hence, for an R^{800} embedding we need to estimate 160000 parameters for A matrix where as representation in R^3 requires only 9 parameters to be estimated; thus making the model in the lower dimensional space much more economical in terms of the available data required for modeling. (Also note that the transformation A in this case corresponds to a thin low dimensional manifold in R^{800} which will cause strong restrictions on A that are difficult to express in a high dimension. But one can effectively deal with A in if the dynamics is represented in a low dimension.)

4.2.4 A Solution to deal with the Issues of High Dimensional Embedding

According to the proposed model a delay embedding in a higher dimensional space is required for global analysis. To tackle the issues related to higher dimensional models, a practical solution based on Broomhead's theorem from topology is proposed here. According to this theorem, Finite Impulse Response (FIR) filters preserve all the information one wants to extract by embedding techniques (refer Appendix E for more details about the theorem) [48]. FIR filters that do not change the dimension of the data preserve embedding [48, 72, 5].

4.2.5 The Proposed New Model

Given a time series that exhibit the property of recurrence the proposed method of modeling is as follows:

- A delay embedding of the data is done in a higher dimensional space R^N. Global analysis of the system is done in R^N. The embedding dimension N and the Recurrent timings (the array that record the time of recurrence for the entire data) are empirically found from the data series.

- A Recurrent neighborhood is defined for a recurrent point in a high dimensional space R^N.

- The recurrent neighborhood in R^N is projected to a lower dimensional space R^d using an FIR filter and the local analysis of the system is done in R^d.

- Neighborhood vectors in the lower dimensional space are collected into a set of equivalence classes based on their recurrent times (the time delay between two adjacent recurrences). A set of affine maps are derived across these equivalence classes. Once the recurrent timings are known, an evolution of vectors in the lower dimensional space can be generated using the affine maps.

- Conjugacy maps are found between the recurrence neighborhoods reconstructed in R^N and R^d. Hence the evolution in R^N can be predicted from the evolution of the vectors in R^d.

Objective of this chapter is to explore these possibilities for a numerically generated data by an ideal dynamical system for a proper understanding and demonstration of the proposed method. The data generated by the Duffing oscillator was specifically selected for the following reasons: (i) it is a well studied system in nonlinear dynamics, and (ii) floquet exponents of the system for stability analysis can be calculated and verified experimentally as the magnitude of their sum is equal to the trace of the system [50].

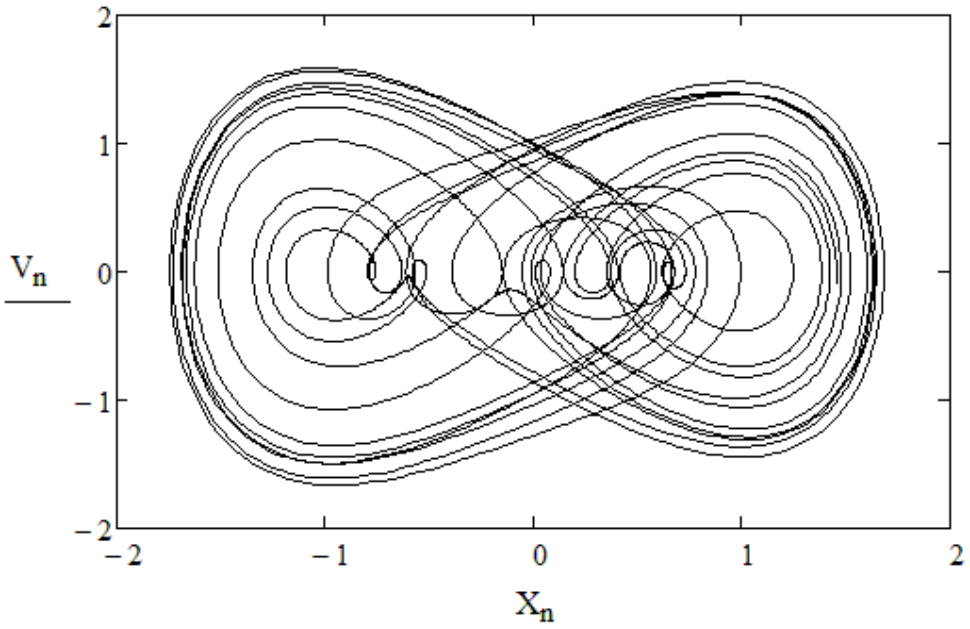

Figure 4.1: State space of the Duffing oscillator under Chaos

4.3 Analysis of data generated by Duffing Oscillator

This section contains the results of numerical analysis. The proposed method is used for the modeling and prediction of a data series generated by Duffing oscillator. Duffing Oscillator is a periodically forced oscillator with a nonlinear elasticity as represented by the second order differential equation of the following form.

$$\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx + \delta x^3 = Fcos(\omega t + \alpha) \tag{4.6}$$

whose state space representation is,

$$\frac{dX}{dt} = V \tag{4.7}$$

$$\frac{dV}{dt} = -cV - kX - \delta X^3 + Fcos(\omega t + \alpha) \tag{4.8}$$

A data was generated from this system under chaos for parameter values $c = 0.04496$, $k = 0$, $\delta = 1$,$\alpha = 0$, $F = 1.03$ using an initial condition $(1.226, 0.868)$. Fig. 4.1 represents a numerically generated state space of the oscillator.

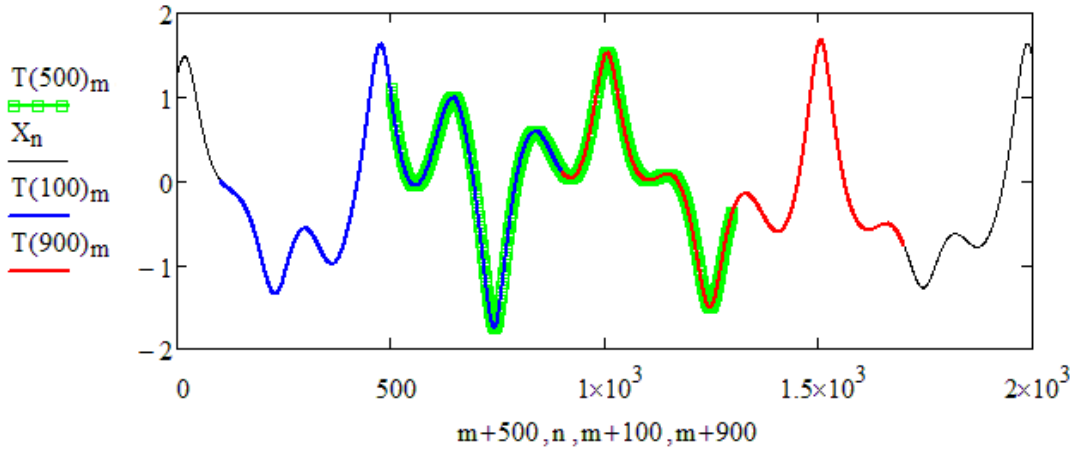

Figure 4.2: Data generated by the Duffing Oscillator under Chaos. Three overlapping delay vectors in R^{800} embedding of the data are highlighted: the end vectors (blue curve at 100 and red curve at 900) completely overlap the middle vector (green curve at 500).

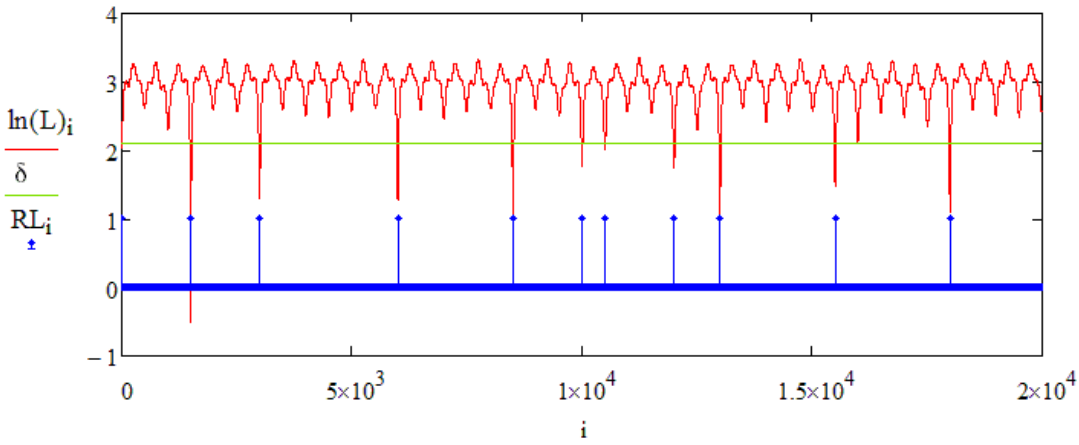

Figure 4.3: Recurrence plot of $ln(L)_i$ Vs i, the index of the delay vectors. Blue stems in the RT plot corresponds to the recurrent neighbors.

Duffing oscillator under chaos exhibits the property of recurrence. A time series generated by the system shows repetition of patterns and it keeps on visiting the neighborhood of that point. The time interval between two such consecutive recurrences is defined as the recurrent time τ. The variable τ is not a constant and it varies in general. Varying recurrent times is a common feature of time series generated by nonlinear systems. Study of recurrence behaviour of the time series reveals a lot of information about the system dynamics [30]. The recurrence timings and a neighborhood for the system are identified from a time series as explained in next section.

4.3.1 Identification of Recurrence Timings and Recurrent Neighborhood from a time series

The time series is embedded in a high dimensional delay embedding space R^N. A specific reference point is identified in R^N and the L_2 norm of the distances of all the vectors with respect to the reference point is recorded in an array L. A recurrence plot can be obtained by plotting $ln(L)$ Vs the index of the delay vectors. (The recurrence plot introduced by Eckmann *et al.* is a visualization of a 2D binary matrix with horizontal and vertical axes defining time, while each element of the matrix will represent either 1 or 0 depending on their measurements at those time indices are closer or not [30]. For the recurrence plot discussed here, there is a flexibility for defining the threshold as the distances are not measured in binary). Once the distances of all other vectors with respect to a reference vector is known, a neighborhood for that reference vector can be identified. Algorithm 4.1 shows the steps to empirically (i) identify a recurrence neighborhood in a high dimensional Takens' delay embedding space R^N and (ii) calculate the recurrence timings for the time series.

For numerical simulation, N was set higher than recurrence interval that was used for prediction. For the duffing oscillator simulation, since period–1 cycle consisted of 500 steps, N was set 800. The data analysis section also gives a prediction for the period–3 cycle which consisted of 1500 steps and in that case N was set 1800. Three overlapping delay vectors in R^{800} embedding of the data are shown in

Fig. 4.2. A recurrence plot $ln(L)$ Vs the index of the delay vectors R^{800} is shown in Fig. 5.4.

Algorithm 4.1 Identification of Recurrence Timings and Recurrence Neighborhood in a Delay Embedding Space R^N

1. Create a delay embedding of the data series in R^N. Each column vector of the delay embedding matrix can be seen as a point in R^N.
2. Choose one vector of the embedding matrix as the reference point X_0.
3. Define an array L to record the distance of all the vectors with respect to X_0.
4. Set a threshold δ for finding a the local minima of L in every recurrence cycle.
5. As the distance crosses δ find the local minima for that recurrence cycle.
6. Choose the point with minimum distance with respect to X_0 in every recurrence cycle as a member of recurrence neighborhood. The collection of all such vectors is mentioned as the recurrent neighborhood of the reference vector X_0.
7. Define an array RT to record the recurrence timings– time delay between consecutive recurrences.

4.4 Dynamics in the Neighborhoods of Manifold in R^d and R^N

According to the proposed model the global analysis of the system is done in the higher dimensional space R^N and the local analysis is done in a lower dimensional space R^d. A linear transformation A represents the FIR filter used for $R^N \rightarrow R^d$ projection. The A matrix has a dimension of $(d \times N)$ and it operates on an $(N \times 1)$ vector of R^N to generate a $(d \times 1)$ vector of R^d. The specific matrix A for the $R^N \rightarrow R^3$ projection, is given below.

$$\begin{aligned} A_{0,n} &= Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \\ A_{1,n} &= Imag\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \\ A_{2,n} &= Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_1}\right)\right) \end{aligned}$$

Parameters chosen for Projection: For the numerical simulation of the Duffing oscillator, number of points per period was set as 500. The constant h_0 was selected to match the number of points per period and h_1 was set $2h_0$. Hence the parameter values chosen for A matrix were: $h_0 = 500, h_1 = 1000, N = 800$ and $n = 1...N-1$. An additional value for the embedding dimension, $N = 1800$ was chosen for the numerical results explained in Section 4.5.

Regarding the similarity of dynamics in R^N and R^d, Let us review the concepts of similarity transformations and topological conjugacy briefly. Similarity transformation in Linear algebra is a conformal mapping that exists between two linear transformations as follows. Two $n \times n$ matrices A and B are called similar if $A = P^{-1}BP$, for some invertible $n \times n$ matrix P. If $x \to Ax$ and $y \to Bx$ then by the existence of the invertible relationship $y = Px$, the mapping $x \to Ax$ is equal to $x \to P^{-1}BPx$ [73]. There is a similar property which defines equivalence relations in topological spaces named topological conjugacy. Let X and Y be two topological spaces with two continuous evolution functions f and g such that $f : X \to X$ and $g : Y \to Y$. If there exists a homeomorphism $h : X \to Y$ such that $f = h^{-1} \circ g \circ h$, then the evolution functions f and g are said to be topologically conjugate to each other. This concept of topological equivalence is useful in the study of dynamical systems as it helps to analyze systems with similar dynamics [4].

According to the proposed model, the data is embedded in a higher dimensional space R^N and then projected to a lower dimension space R^d. A topological conjugacy between the neighborhoods reconstructed in R^N and R^d are empirically justified in this section by finding maps to go back and forth across the neighborhoods. The goal is to predict the evolution in R^N using the evolution of vectors in R^d. Projection of the recurrent neighborhood from R^N to R^d is simple using the linear transformation A. Following sub sections explains the process of going back and forth across the neighborhoods in the lower dimension R^d and the higher dimension R^N. In R^d, members of the neighborhood are classified into different equivalence classes based on their recurrent time τ. One can generate evolution maps in R^d for each of these equivalence classes as shown in section 4.4.2. The affine map shown in section 4.4.1 can be used for going back from R^d to R^N.

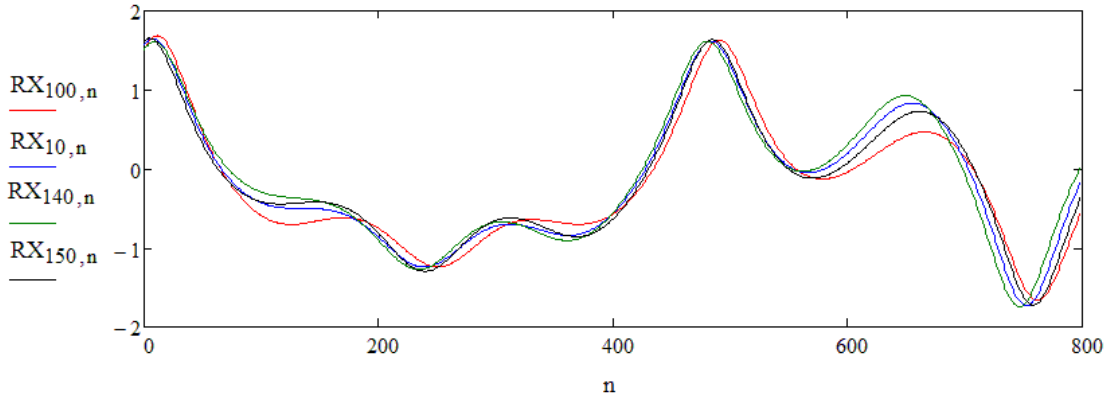

Figure 4.4: A few random vectors in the neighborhood of the Duffing data in R^{800}

4.4.1 Conjugacy maps defined across the neighborhoods in R^d and R^N

Projection from R^N to R^d is simple using the linear transformation A. But going back from the lower dimension R^d to the higher dimension R^N is not that straight forward especially if the dynamics is nonlinear. Hence a conjugacy map is developed in this section which is used to predict the vectors in the higher dimension R^N from the corresponding vectors in the lower dimensional space R^d. The method of singular value decomposition (SVD) and a nonlinear generalization of SVD [52] are used to find these maps.

Let the recurrent neighborhood matrices be denoted as RX in the higher embedding dimension R^N and RY the lower projected dimension R^d. For the numerical analysis the values were $N = 800$ and $d = 3$. Let the neighborhoods contain P vectors each. RX had dimension $(P \times 800)$ and RY had dimension $(P \times 3)$.

Total number of neighbors $P = 1000$ for the numerical simulations explained here. A few neighborhood vectors of the Duffing data in R^{800} are shown in Fig. 4.4. The set of affine transformations across these neighborhoods in R^d and R^N are empirically determined as follows.

Step 1: Find the centroid of the neighborhood and subtract it from the neighbors to translate the neighborhood to origin. Let rX and rY represent the recurrent neighborhood matrices translated to origin. Record the centroid of RX and RY as $\overline{RX}$ and $\overline{RY}$ respectively. A few neighborhood vectors for the Duffing

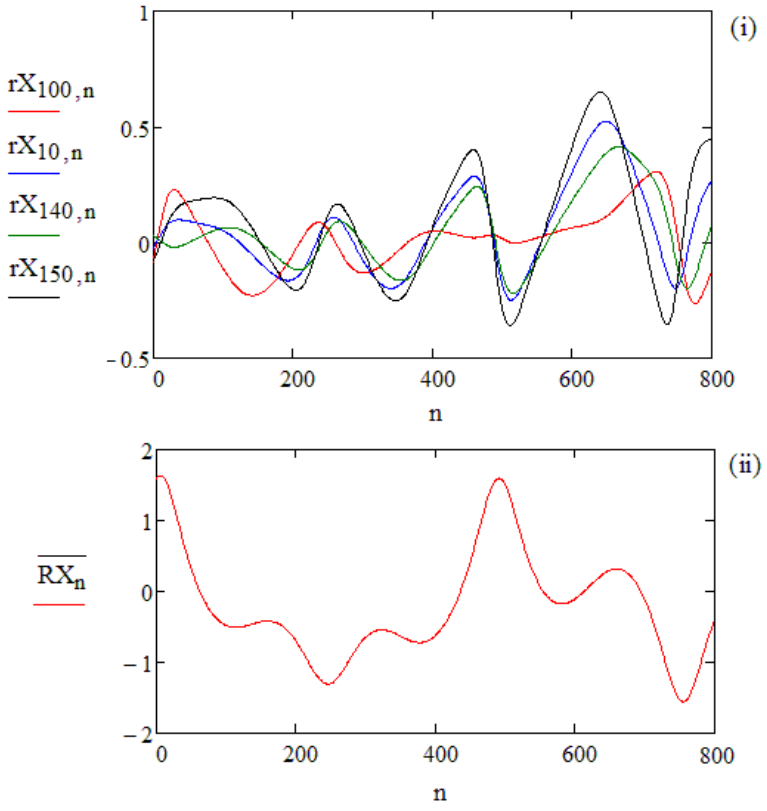

Figure 4.5: A few random vectors in the neighborhood of the Duffing data in R^{800} after translation to origin (i) and the centroid $\overline{RX}$ (ii)

data after translation to origin along with the centroid are shown in Fig. 4.5.

$$\overline{RX}_n = \frac{1}{P}\sum_{p=1}^{P} RX_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{4.9}$$

$$\overline{RY}_n = \frac{1}{P}\sum_{p=1}^{P} RY_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{d} \tag{4.10}$$

Neighborhoods translated to origin are:

$$rX_{p,n} = RX_{p,n} - \overline{RX}_n \qquad \text{for all p,} \quad \text{for n} = 1, 2 \ldots \text{N} \tag{4.11}$$

$$rY_{p,n} = RY_{p,n} - \overline{RY}_n \qquad \text{for all p,} \quad \text{for n} = 1, 2 \ldots \text{d} \tag{4.12}$$

Step 2: Define a conjugacy map T, across the neighborhood matrices rX

and rY such that,

$$(rY).T = rX \tag{4.13}$$

Step 3: rX of dimension $P \times N$, rY of dimension $P \times d$ are known matrices; T of dimension $d \times N$ can be calculated by pre–multiplying rX with the generalized inverse of rY.

$$T = (rY)^{-1}.rX \tag{4.14}$$

T represents a linear map across the neighborhoods rY, rX shifted to origin. This implies an affine map of the form ($\beta = T\alpha + \overline{RX}$) exists for every vector α in RY and β in RX. Hence the map T and the centroid $\overline{RX}$ are known, a vector in R^N can be generated from a vector in R^d. In this manner one can predict any of the recurrence vectors in R^N from their R^d counterparts. This affine transformation across the neighborhoods RX and RY represents the topological conjugacy between the manifolds in R^N and R^d.

4.4.2 Evolution maps across different equivalence classes in R^d.

Members who belong to the recurrence neighborhood in R^d can be classified into different equivalence classes based on the recurrent time τ. Fig. 5.5 shows the equivalence classes in R^3 which corresponds to various periods of duffing oscillator generated data. Numerically generated data of 600000 points (1200 cycles with 500 points per cycle) were used to plot the frequency depicted in this figure. X axis shows the equivalence classes corresponding to recurrent timings 500 (period 1), 1000(period 2), 1500(period 3) ... etc and Y axis shows the total number of members in each equivalent class for this finite data.

Once the equivalence classes in R^3 are identified, an evolution map for a particular equivalence class can be defined as explained below. Steps mentioned here are similar to the steps 1–3 in Section 4.4.1 except that the transformation is defined in R^d alone and the transformation matrix V has a dimension $d \times d$. Steps given below also shows the empirical values of the variables in R^3 for a particular equivalence class 1500 (period 3).

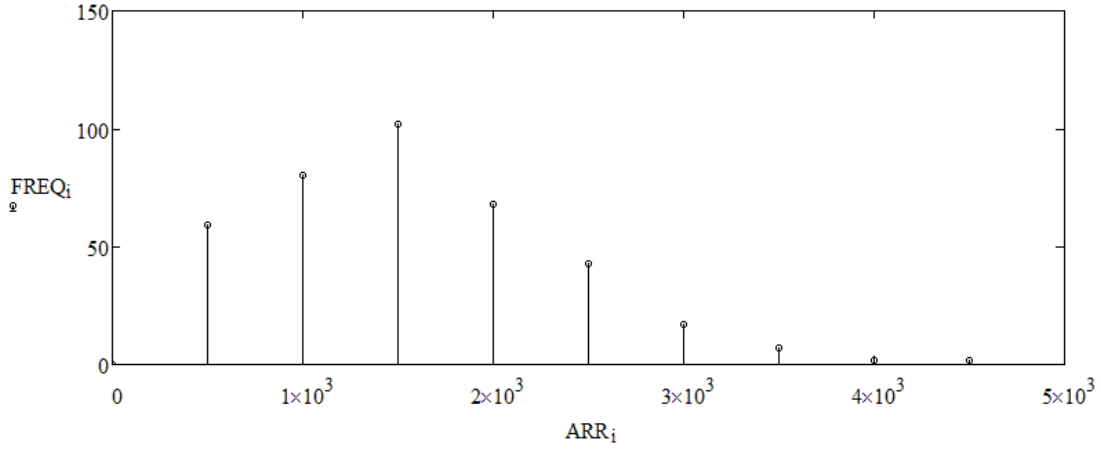

Figure 4.6: Equivalence classes in R^3 for the Duffing data. X axis shows the equivalence classes corresponding to recurrent timings 500 (period 1), 1000(period 2), 1500(period 3) ... etc. Y axis shows the total number of members in each equivalent class.

Let the Start and End matrices for the equivalence class 1500 in R^3 be denoted by SX, SY respectively. SX contains all the vectors that start with recurrent timing 1500 and SY contains all the vectors that end with recurrent timing 1500. Fig. 5.5 shows that the equivalence class 1500 contains 102 members. Hence SX and SY have dimensions of 3×102 each. Now we can develop affine transformations to reach SX from SY as follows. Once the parameters of the transformations are known, we can predict a SY vector given an SX vector.

Step 1: Find the centroid of the neighborhoods SX, SY and subtract it from the neighbors to translate the neighborhood to origin. Let sX and sY represents the neighborhoods translated to origin. Record the centroids of SX and SY as $\overline{SX}$ and $\overline{SY}$ respectively.

SX and SY have dimensions of 102×3 each. Numerical values of few vectors that constituted SX and SY matrices were,

$$SX^T = \begin{bmatrix} 0.4364 & 0.4634 & 0.3966 & 0.4258 & 0.4239 & \dots \\ -0.0658 & -0.0761 & -0.0511 & -0.0232 & -0.0975 & \dots \\ -0.0690 & -0.0771 & -.1400 & -0.0315 & -0.0303 & \dots \end{bmatrix}$$

$$SY^T = \begin{bmatrix} 0.4667 & 0.4393 & 0.4545 & 0.4690 & 0.4669 & \dots \\ -0.0860 & -0.0514 & 0.0129 & -0.0813 & -0.0936 & \dots \\ -0.0821 & 0.0274 & -0.0303 & -0.0803 & -0.0777 & \dots \end{bmatrix}$$

Numerical values of the centroid vectors $\overline{SX}$ and $\overline{SY}$ in R^3 were,

$$\overline{SX} = \begin{bmatrix} 0.4325 \\ -0.0583 \\ -0.0574 \end{bmatrix}$$

$$\overline{SY} = \begin{bmatrix} 0.4529 \\ -0.0438 \\ -0.0481 \end{bmatrix}$$

Step 2: Define a transformation V, across the matrices sX and sY such that;

$$sX.V = sY \tag{4.15}$$

Step 3: sX, sY both of dimension 102×3 are known matrices; V of dimension 3×3 can be calculated by pre–multiplying sY with the generalized inverse of sX. The generalized inverses of sX were found using both linear and nonlinear SVD procedure.

$$V = (sX)^{-1}.sY \tag{4.16}$$

Note that this transformation is a function of only the recurrent time τ. Empirically determined V matrix for the period–3 equivalence class ($\tau = 1500$) was,

$$V = \begin{bmatrix} 0.0506 & 0.0414 & 0.0035 \\ 0.0429 & 0.5758 & -0.1028 \\ 0.0130 & -0.0993 & -0.1304 \end{bmatrix}$$

Step 4: Once the V matrix is estimated, prediction for any random starting vector in 1500 equivalence class can be made. V multiplied by an R^3 vector of dimension 3×1 creates another vector in R^3. Given any random vector $\alpha \in SX$

subtract the centroid $\overline{SX}$ to translate it to origin. Let $r\alpha$ be the translated vector.

$$\alpha - \overline{SX} = r\alpha \tag{4.17}$$

Step 5: If V is known $r\alpha \in sX$ can be evolved to $r\beta \in sY$ according to the equation,

$$r\alpha.V = r\beta \tag{4.18}$$

Step 6: $r\beta$ belongs to the 1500– End equivalence class in R^3. The prediction for $\beta \in SY$ the original neighborhood can be made by adding centroid vector $\overline{SY}$ for that equivalence class as,

$$r\beta + \overline{SY} = \beta \tag{4.19}$$

Hence following steps 1–6 one can find the evolution of a vector that belongs to the 1500 equivalence class. In this manner evolution maps for all the equivalence classes can be found.

Once the recurrent timings for the entire time series is available, the algorithm 4.2 can be used to predict the evolution of vectors in R^N using their evolution in R^d.

Algorithm 4.2 Prediction of the evolution of vectors in R^N using their evolution in R^d

---

1. Start with an initial vector in the higher dimension space R^N.
2. Find the corresponding vector in the lower dimensional space R^d.
3. Locate the equivalence class in which the initial vector belongs to (using the given information about the recurrent timings)
4. Evolve the vector in R^d according to the recurrent timings using the transformations corresponding to the specific equivalence classes.
5. Use the conjugacy maps defined from R^d to R^N, to come back to the higher dimensional space at the end of every recurrent cycle.
6. Stitch the R^N vectors appropriately at their corresponding neighborhood locations (given by the recurrent times) for the prediction

---

4.5 Result of Data Analysis

This section contains the analysis results for the duffing data using the proposed new model. The data was embedded into R^N and then projected it to R^d by a linear transformation. Numerical simulation results are demonstrated here for two choices of the delay embedding dimension $N = 800$ and $N = 1800$. Value of d was set as 3. The conjugacy maps that take R^3 vectors to R^N vectors and evolution maps in R^3 were found as explained in previous sections. Both the standard and nonlinear SVD procedures were used to find these maps. Predicted signals were compared with the original signals using the scores explained below. Prediction results for few random vectors are demonstrated with figures in following sections using both the linear and nonlinear SVDs. Prediction using the standard SVD is discussed in Section 4.5.2 and nonlinear SVD is discussed in section 4.5.3. A comparison of the errors for each pair these predictions are given in Section 4.5.4. The scores for were calculated for the vectors that belong to the period 3 ($\tau = 1500$) equivalence class.

4.5.1 Scores for the Analysis of Results

Two types of scoring functions were used for comparing the prediction with the original signal. If x_i and y_i are the original signal and predicted signal respectively where $i = 1, 2, ...N$. Then the scoring functions were,

$$Q1(x, y) = 1 - \frac{MSE(x, y)}{VAR(x)} \tag{4.20}$$

$$Q2(x, y) = corr(x, y) \tag{4.21}$$

where,

$$MSE(x, y) = \frac{1}{N} \sum_{i=1}^{N} (y_i - x_i)^2$$

$$VAR(x) = \frac{1}{N} \sum_{i=1}^{N} {x_i}^2 - \left[\frac{1}{N} \sum_{i=1}^{N} x_i \right]^2$$

4.5.2 Prediction results using standard SVD

This Section explains the results where the transformation matrix in Eq. 4.16 was found by a generalized inverse based on the standard SVD.

1. Prediction from R^3 to R^{800}: Here we show the prediction for two random members: the 50^{th} member and the 120^{th} member in the R^{800} neighborhood. In Fig. 5.7, $P1$ is the affine prediction, RM_F is the centroid for the R^{800} neighborhood, and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.7 that the Affine predictions improves the centroid fit. Scores of affine prediction for the 50^{th} member were $(Q1 = 0.9955, Q2 = 0.9985)$ and that of centroid fit were $(Q1 = 0.9604, Q2 = 0.9818)$. Scores of affine prediction for the 120^{th} member were $(Q1 = 0.9975, Q2 = 0.9988)$ and that of centroid fit were $(Q1 = 0.9479, Q2 = 0.9780)$.

2. Prediction from R^3 to R^{1800}: Here we show the prediction for two random members again: the 50^{th} member and the 120^{th} member in the R^{1800} neighborhood. In Fig. 5.8, $P1$ is the affine prediction, RM_H is the centroid for the R^{1800} neighborhood,and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.7that the Affine predictions improves the centroid fit. Scores of affine prediction for the 50^{th} member were $(Q1 = 0.9896, Q2 = 0.9948)$ and that of centroid fit were $(Q1 = 0.9300, Q2 = 0.9753)$. Scores of affine prediction for the 120^{th} member were $(Q1 = 0.9785, Q2 = 0.9897)$ and that of centroid fit were $(Q1 = 0.9068, Q2 = 0.9697)$.

4.5.3 Prediction results using nonlinear SVD

This Section explains the results where the transformation matrix in Eq. 4.16 was found by a generalized inverse based on the nonlinear SVD.

1. Prediction from R^3 to R^{800}: Here we show the prediction for two random members: the 50^{th} member and the 120^{th} member in the R^{800} neighborhood. In Fig. 5.7, $P1$ is the affine prediction, RM_F is the centroid for the R^{800} neighborhood, and the error of the original signal with respect to the pre-

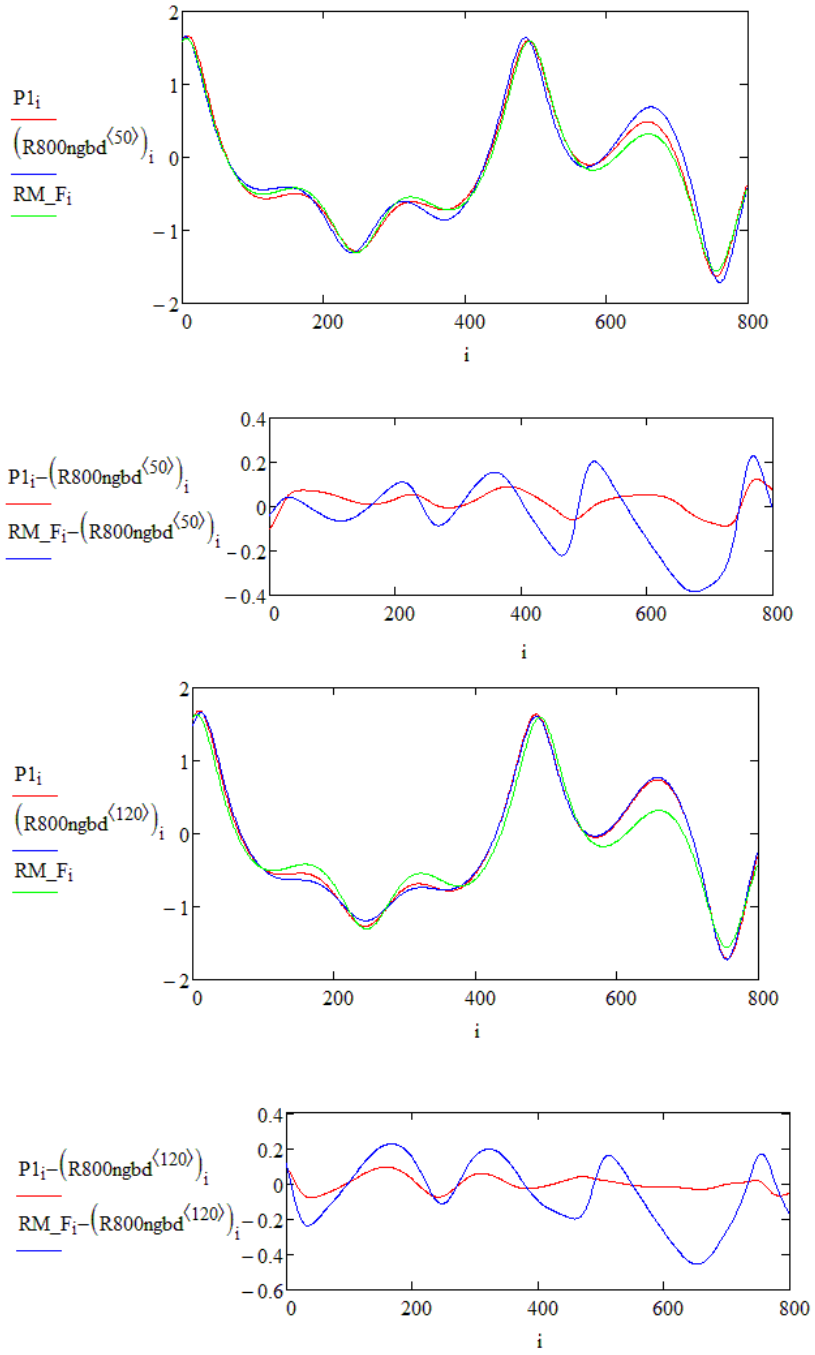

Figure 4.7: $P1$ **is the affine prediction, RM_F is the centroid for the** R^{800} **neighborhood, the error of the original signal with respect to the predicted signal and the centroid vector also shown. Scores of affine prediction for the** 50^{th} **member were** $(Q1 = 0.9955, Q2 = 0.9985)$ **and that of centroid fit were** $(Q1 = 0.9604, Q2 = 0.9818)$**. Scores of affine prediction for the** 120^{th} **member were** $(Q1 = 0.9975, Q2 = 0.9988)$ **and that of centroid fit were** $(Q1 = 0.9479, Q2 = 0.9780)$**.**

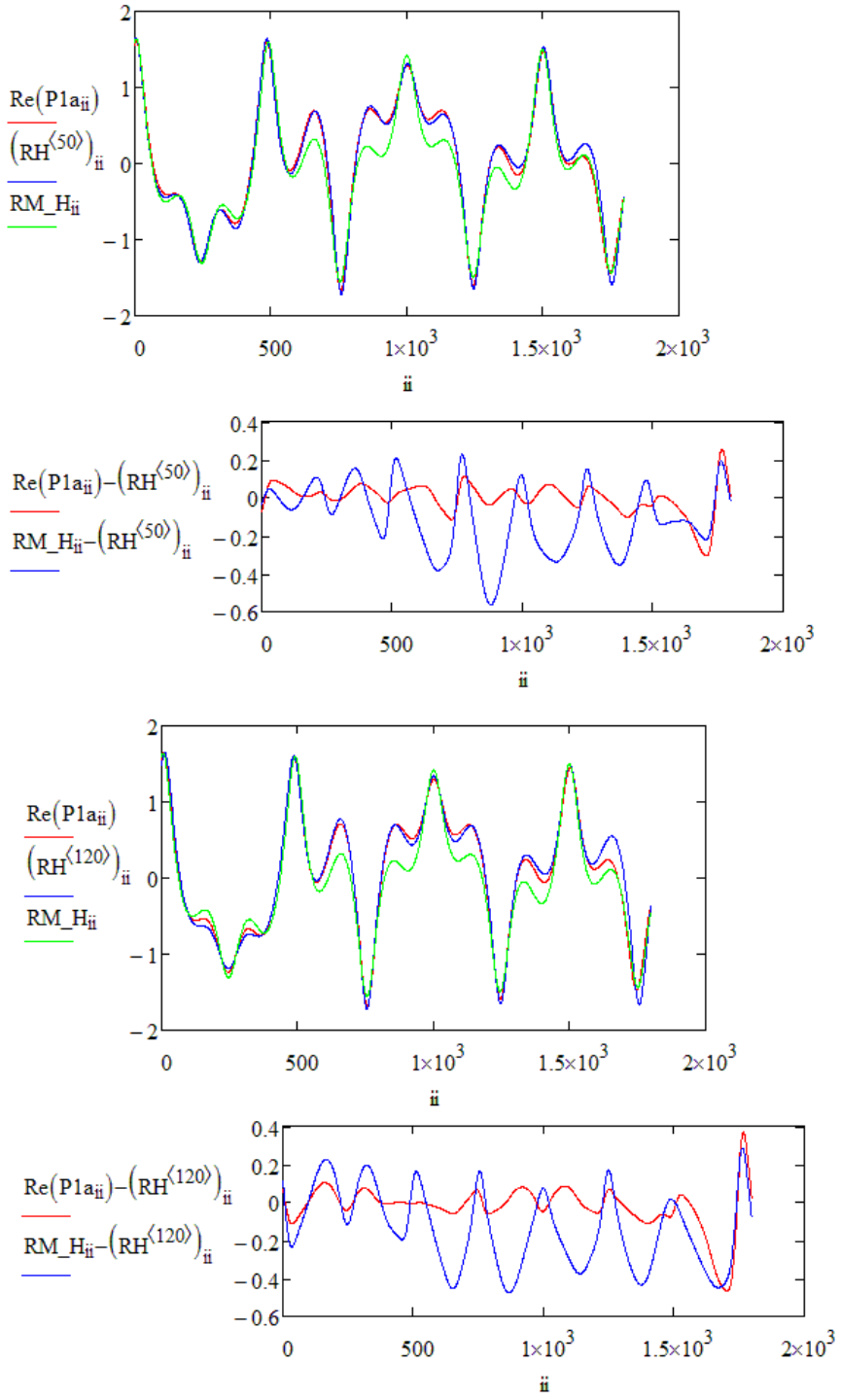

Figure 4.8: $P1$ **is the affine prediction, RM_H is the centroid for the** R^{1800} **neighborhood, the error of the original signal with respect to the predicted signal and the centroid vector also shown. Scores of affine prediction for the** 50^{th} **member were** $(Q1 = 0.9896, Q2 = 0.9948)$ **and that of centroid fit were** $(Q1 = 0.9300, Q2 = 0.9753)$**. Scores of affine prediction for the** 120^{th} **member were** $(Q1 = 0.9785, Q2 = 0.9897)$ **and that of centroid fit were** $(Q1 = 0.9068, Q2 = 0.9697)$**.**

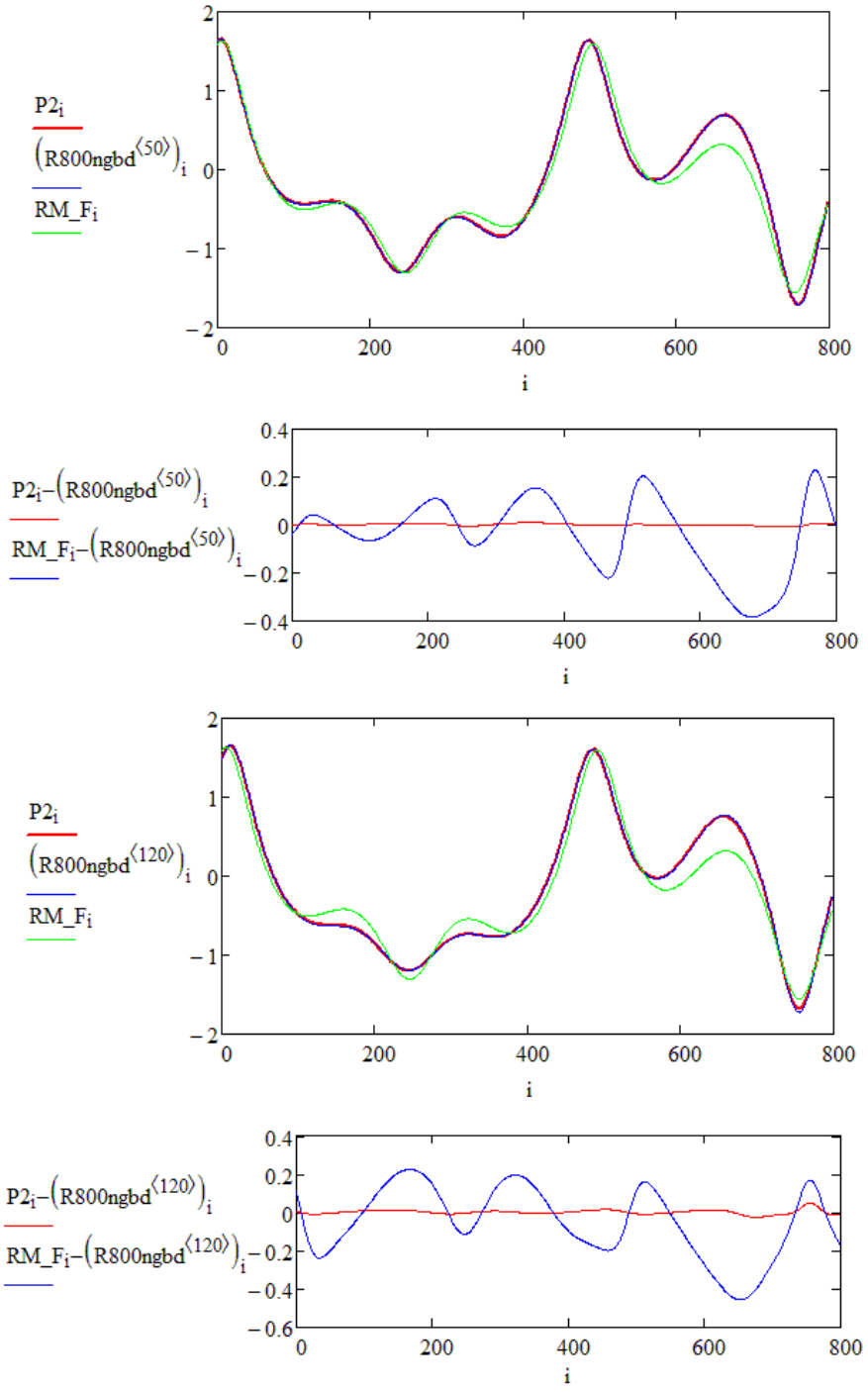

Figure 4.9: $P2$ **is the affine prediction based on nonlinear SVD, RM_F is the centroid for the** R^{800} **neighborhood, the error of the original signal with respect to the predicted signal and the centroid vector also shown. Scores of affine prediction for the** 50^{th} **member were** $(Q1 = 0.99997, Q2 = 0.99999)$ **and that of centroid fit were** $(Q1 = 0.9604, Q2 = 0.9818)$**. Scores of affine prediction for the** 120^{th} **member were** $(Q1 = 0.9998, Q2 = 0.9999)$ **and that of centroid fit were** $(Q1 = 0.9479, Q2 = 0.9780)$**.**

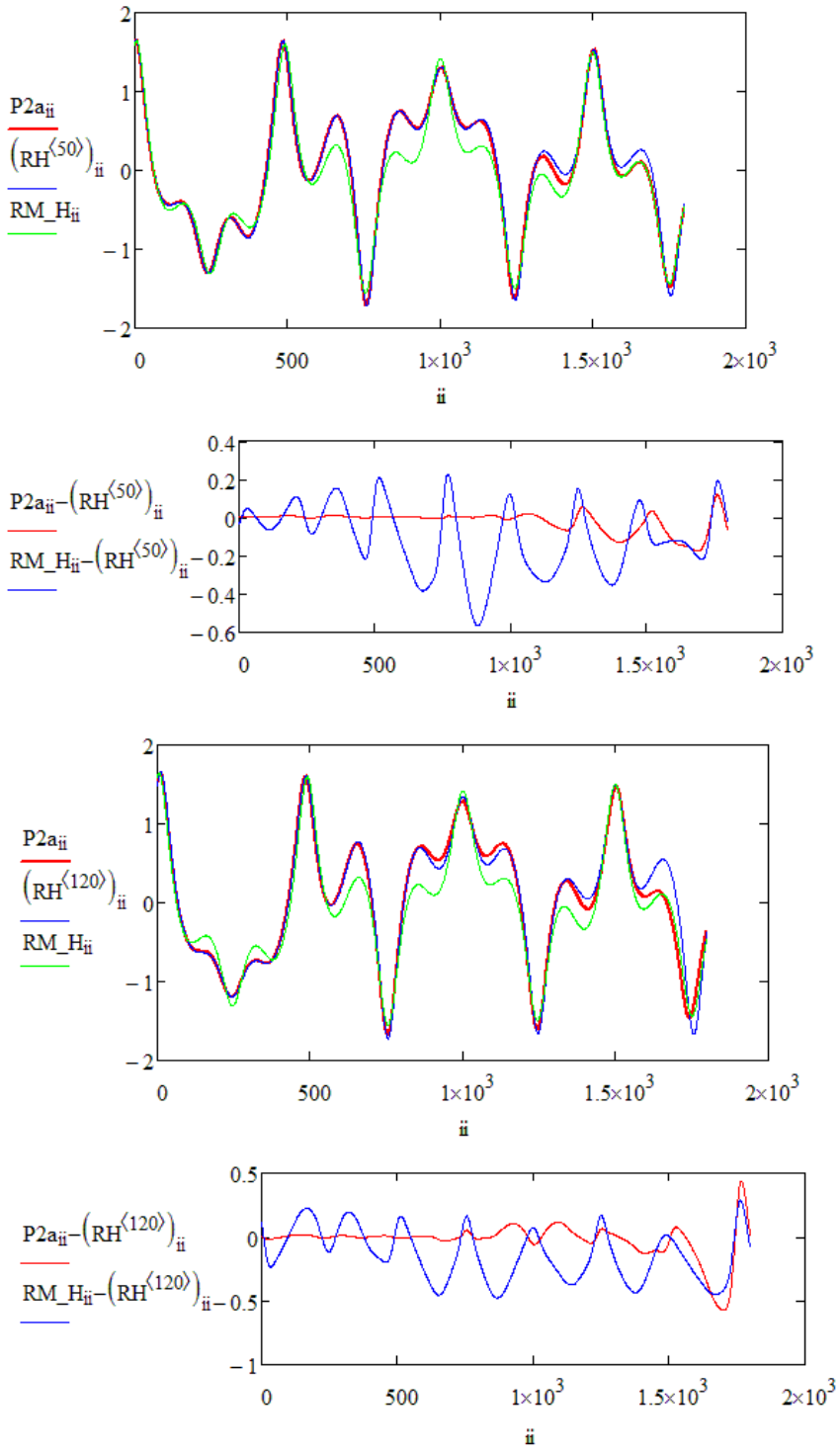

Figure 4.10: $P2a$ **is the affine prediction based on nonlinear SVD, RM_H is the centroid for the** R^{1800} **neighborhood, the error of the original signal with respect to the predicted signal and the centroid vector also shown. Scores of affine prediction for the** 50^{th} **member were** $(Q1 = 0.9949, Q2 = 0.9978)$ **and that of centroid fit were** $(Q1 = 0.9300, Q2 = 0.9753)$**. Scores of affine prediction for the** 120^{th} **member were** $(Q1 = 0.9691, Q2 = 0.9849)$ **and that of centroid fit were** $(Q1 = 0.9068, Q2 = 0.9697)$**.**

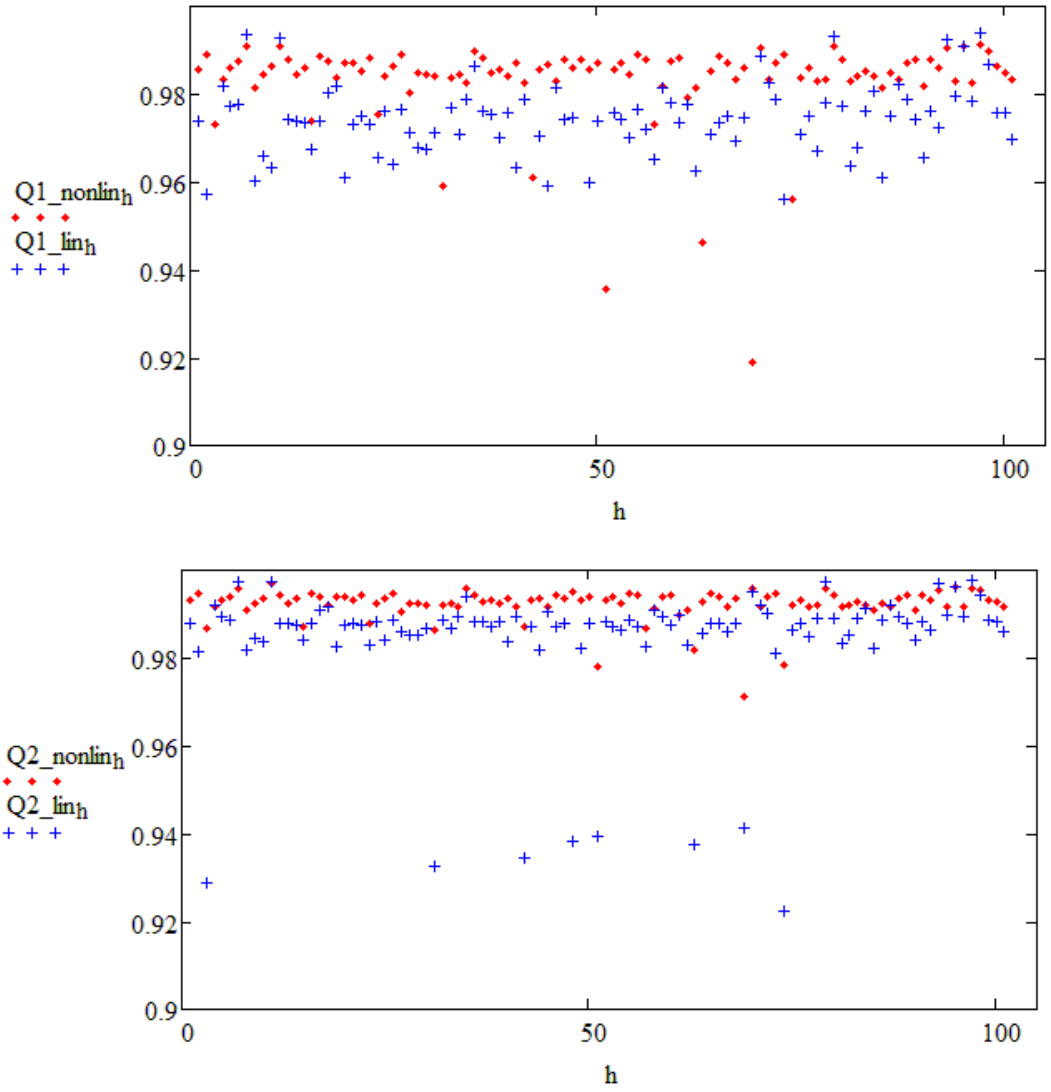

Figure 4.11: Scores Q1 and Q2 for both the linear and nonlinear SVD based prediction in R^{800} for the 102 members of 1500 equivalence class of vectors in R^3 evolution.

dicted signal and the centroid vector also shown. It is clear from the Fig. 4.9 that the Affine predictions improves the centroid fit. Scores of affine prediction for the 50^{th} member were $(Q1 = 0.99997, Q2 = 0.99999)$ and that of centroid fit were $(Q1 = 0.9604, Q2 = 0.9818)$. Scores of affine prediction for the 120^{th} member were $(Q1 = 0.9998, Q2 = 0.9999)$ and that of centroid fit were $(Q1 = 0.9479, Q2 = 0.9780)$.

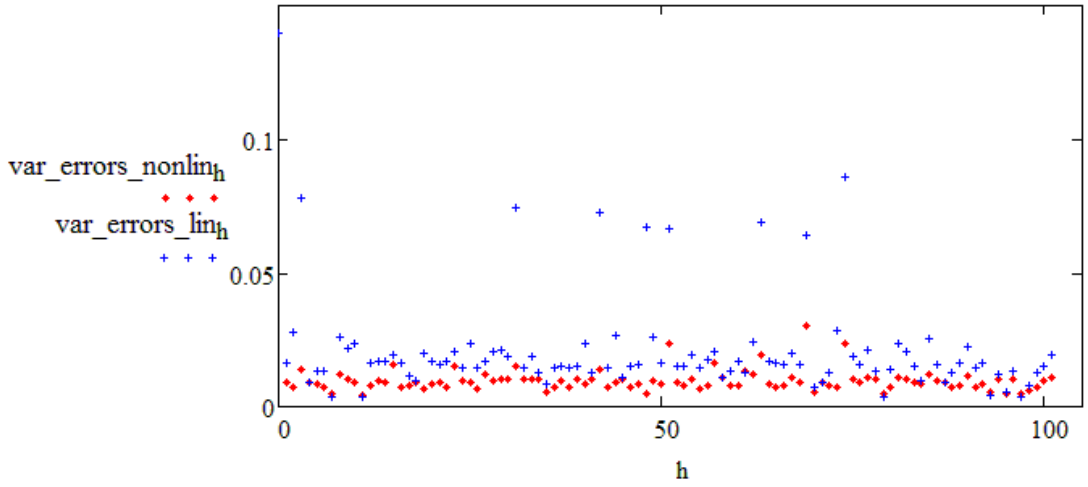

Figure 4.12: Variance of errors for both the linear and nonlinear SVD based prediction in R^{800} for the 102 members of 1500 equivalence class of vectors in R^3 evolution.

2. Prediction from R^3 to R^{1800}: Here we show the prediction for two random members again: the 50^{th} member and the 120^{th} member in the R^{1800} neighborhood. In Fig. 4.10, $P1$ is the affine prediction, RM_H is the centroid for the R^{1800} neighborhood,and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.7 that the Affine predictions improves the centroid fit. Scores of affine prediction for the 50^{th} member were $(Q1 = 0.9949, Q2 = 0.9978)$ and that of centroid fit were $(Q1 = 0.9300, Q2 = 0.9753)$. Scores of affine prediction for the 120^{th} member were $(Q1 = 0.9691, Q2 = 0.9849)$ and that of centroid fit were $(Q1 = 0.9068, Q2 = 0.9697)$.

The prediction scores for both the linear and nonlinear SVD based prediction results for entire members of a particular equivalence class was also found. Fig. 5.9 shows Q1 and Q2 scores of 102 members of 1500 equivalence class for R^3 to R^{800} prediction.

Table 4.1: Variance of errors between the original and predicted signal (for the three cases P1, P2, RM_F) for R^3 to R^{800} prediction of the 50^{th} and 120^{th} members

Neighborhood Vector	P1(std SVD)	P2(nonlinear SVD)	RM_F(centroid fit)
50	0.00234	0.00002	0.02415
120	0.00173	0.00014	0.03361

4.5.4 Variance of Errors for the Prediction based on linear and nonlinear SVD

Analysis results of the previous section shows that even the centroid based predictions give decent scores. Hence one can use them if no much information is available. Affine transformations based on the standard SVD gave better predictions. The nonlinear SVD based predictions are the best choice as they improve the scores significantly. Another score to compare the errors of these three cases of predictions was defined as follows.

If x_i and y_i are the original signal and predicted signal respectively, where $i = 1, 2, ...N$. Then the error is defined as $e_i = x_i - y_i$ and the variance of error is defined as,

$$VAR(e) = \frac{1}{N}\sum_{i=1}^{N} {e_i}^2 - \left[\frac{1}{N}\sum_{i=1}^{N} e_i\right]^2$$

Variance of errors for the three specific cases of prediction were calculated. For R^3 to R^{800} prediction of the 50^{th} and 120^{th} members, variance of errors (for the three cases P1, P2, RM_F depicted in figures) are listed in Table. 4.1. The variance of errors for both the linear and nonlinear SVD based prediction for the entire members of 1500 equivalence classes are shown in Fig. 5.10.

4.6 Conclusions

For the numerical simulations explained in this chapter the delay embedding dimension N were chosen as 800 and 1800. Ideally the dimension for the delay embedding should be greater than the recurrent interval one want to predict. For a 1500 recurrence cycle, choosing embedding dimension less than 1500, $N = 800$ gives

a partial prediction for the signal, where $N = 1800$ gives a full prediction (Section 7.6.1 of Chapter 7 discusses the issue of partial prediction this issue in detail).

Regarding the dimension of the projected space R^d, numerical analysis was done for projections to various lower dimensions i.e. for varying values of d, $d = 2, 3, 4$. Based on numerical simulations for data from $R^{800} \rightarrow R^2$ and $R^{800} \rightarrow R^4$, It was observed that the prediction based on the equivalence classes in R^3 and R^4 were equally good but the values in R^2 was slightly better out of these three choices. Hence it can be inferred that for the duffing case, even though the system dimension is greater than 2 as it is a $2D$ non–autonomous system, $d = 2$ would be enough for a local analysis of a neighborhood on the manifold. Stability analysis section of Chapter 6 demonstrates that selecting $d > 2$ can generate spurious floquet exponents for duffing data.

The goal of the chapter was to study the possibilities of the proposed new model for the data generated by duffing oscillator. It is demonstrated that the data can be modeled using the recurrence timings and transformations specific to the equivalence classes of the recurrence neighborhoods. The prediction results were excellent and the proposed methods are explored for the electrocardiogram (ECG) data in the next chapter. Simulation results for $R^{800} \rightarrow R^3$ projection demonstrated for duffing oscillator in this chapter can be used to make a comparison of similar projection of ECG data in the next chapter. Stability analysis for both the systems based on the data are done in Chapter 6. An application of some methods discussed in this chapter for the prediction of multichannel data using $R^N \rightarrow R^N$ transformations are discussed in Chapter 7.

CHAPTER 5
Topological structure of ECG signals

The goal of this chapter is to demonstrate the proposed new model, for analyzing an electrocardiogram (ECG) data of the human heart. The ECG data exhibits the property of recurrence and we will see in this chapter that the proposed model gives an excellent prediction for the data. According to the proposed model the dynamics of the system can be approximated by a few overlapping recurrence neighborhoods with specific affine transformations for each of them. For the special case of ECG data, the entire nonlinear structure of the data can be deduced from one small neighborhood as shown in this chapter.

5.1 Introduction

The dynamics of the heart is fascinating and it exhibits both nonlinear deterministic and stochastic aspects. A typical ECG signal has a repetitive structure in every cycle which can be connected to various actions of the heart during one pumping cycle [74]. Further, the signal exhibits nonstationarity properties as its statistical properties changes over time. The property of variation of recurrent intervals with time is a prominent feature of nonlinear signals. A variation in the beat–to–beat interval timings known as the Heart Rate Variability (HRV) is a typical characteristic of a healthy ECG measurement [65, 75]. Both stochastic and nonlinear models are used by researchers to capture important aspects about the heart dynamics [65, 76]. The analysis for the ECG signal presented in this chapter takes an alternative and fruitful point of view.

The motivation to work on this problem was the Physionet challenge 2010 whose goal was to find a short segment of missing data from one of the channels in a multichannel physiological data [37]. The aim was to develop a model that can predict various physiological signals based on the information available on other channels. While dealing with this issue we came across many questions. Let us focus on two of them for the purpose of signal analysis: (i) how to model an ECG signal

effectively? (ii) how is the heart rate linked to the heart dynamics? The methods demonstrated in this chapter attempt to give a solution to these questions.

The key objective of this chapter is to apply the method of prediction discussed in Chapter 4 for the analysis of ECG signals. The ECG signal exhibit the property of recurrence. A typical single channel ECG recording is shown by Fig. 5.1. Note that the beat–to–beat intervals or RR intervals of the ECG data is same as the recurrence intervals of the signal. If the reference vector is selected such that it matches a QRS peak location, then the address of all the recurrence neighbors will match with all the QRS peak locations of the signal. We have seen that there is lot of information that is available on the recurrence timings of the signal which can be used for the purpose of modeling and prediction. Hence, if the recurrence timings are known for the ECG data, the rest of the dynamics is quite simple. It is inferred that the proposed model works very effectively for ECG signals due to the excellent prediction it gives.

According to the model, If the affine maps specific to the equivalence classes and the conjugacy maps across the lower and higher dimensional spaces are known, a prediction can be done for the ECG signal using an additional information– the recurrent timings. The maps corresponding to the equivalence classes (each with a specific recurrent time) for the case of ECG could be seen as various possibilities for the heart dynamics and the HRV could be seen as a switching between these different possibilities. The objective of this chapter is to show how to model and predict an ECG signal using transformations specific to the recurrence neighborhood, given the recurrent timings. It is demonstrated that the entire ECG dynamics can be deduced by studying a recurrent neighborhoods.

The chapter is organized as follows. Sections 5.2 discusses the property of recurrence in ECG signals and lists an algorithm to identify the recurrent timings for the signal in a high dimensional delay embedding space. Section 5.3. discusses the details of the transformation that projects the data from the delay embedding space R^N to a low dimensional space R^d (Basic concepts of the proposed method are already are discussed in Chapter 4). Further it discusses how to develop conjugacy maps and evolution maps in R^N and R^d for the ECG signal for the purpose of

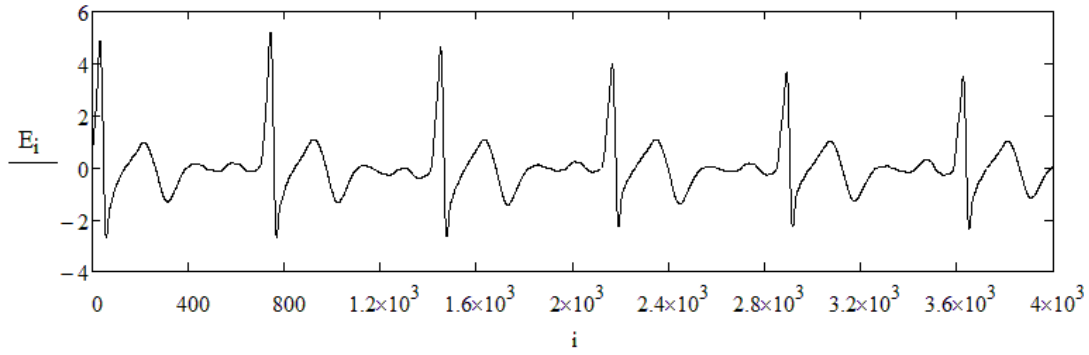

Figure 5.1: Data under study: an ECG recording of a healthy Male aged 36 years sampled at 1kHz

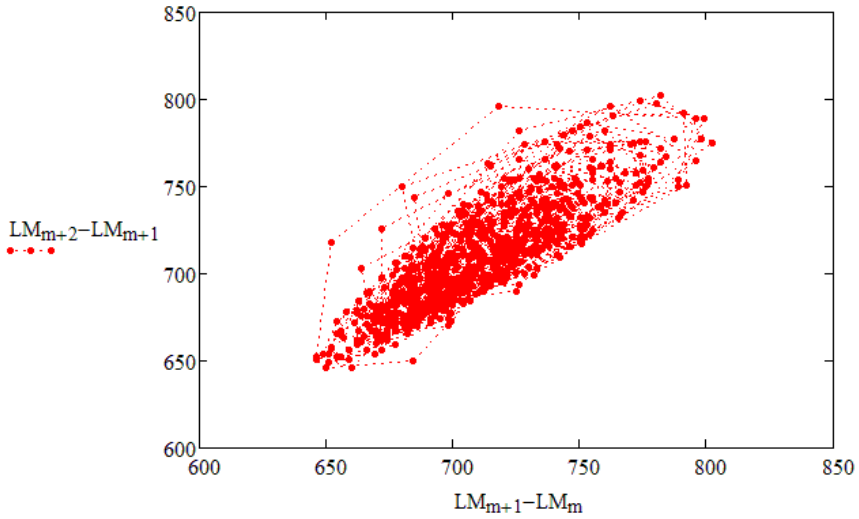

Figure 5.2: The heart rate variability plot for the ECG data depicted in Fig. 5.1

prediction. Results of the data analysis are given in section 5.4. Section 5.5 has a discussion about the empirical values chosen for the projected dimension d and section 5.6 contains the conclusions of the study.

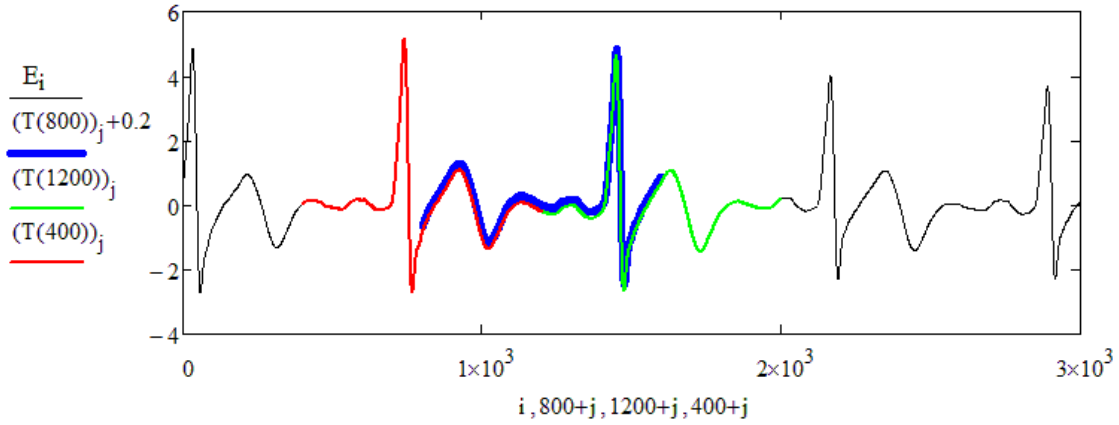

Figure 5.3: Three overlapping delay vectors in R^{800} embedding of the ECG signal. The blue curve (delay vector at 800) is overlapped by red curve (delay vector at 400) and green curve (delay vector at 1200).

5.2 Analysis of ECG data

The ECG data exhibits the property of recurrence as it can been seen by the repetition of wave patterns in every cycle. Fig. 5.1 shows the ECG recording of a healthy Male aged 36 years sampled at 1kHz (data courtesy [77]). The heart rate variability plot also known as RR variability plot (Beat–to–beat intervals plotted with a delay) for this data is depicted in Fig. 5.2.

5.2.1 Identification of Recurrence Timings and Recurrent Neighborhood for the ECG data

The recurrence timings and a neighborhood are identified from the ECG data as follows. The time series is embedded in a high dimensional delay embedding space R^N. The delay vectors can be seen as interconnected fibers and hence there is a possibility of predicting the intermediate fibers from corresponding end fibers. Fig. 5.3 shows three overlapping delay vectors in R^{800} delay embedding of the ECG data. For finding the recurrence timings and neighborhood of the data, start by selecting a reference vector in R^N and find the L_2 norm of the distances of all other vectors with respect to the reference point. Choose a threshold for the radius of the neighborhood and select vectors in every recurrence cycle whose distance fall within

the threshold. Select one vector which has the least distance to the reference vector in every cycle to be a member of the neighborhood. The time interval between the selected recurrent neighbors is recorded as an array of recurrent timings. An algorithm to find the neighborhood for the reference point and the recurrence timings are given by Algorithm. 5.1.

Algorithm 5.1 Identification of Recurrence Timings and Recurrence Neighborhood for the ECG data in the Delay embedding space R^N

1. Set an integer that is approximately equal to an average recurrence cycle of the time series as the initial value for N.
2. Create a delay embedding of the data series in R^N .
3. Choose a vector of the embedding matrix $X_{ref} \in R^N$ as the reference point. N should be selected such that this reference vector has only on QRS peak in it. (Ideally select a vector that has the QRS peak approximately towards the middle as X_{ref} for the reasons explained step 8)
4. Define an array L to record the L_2 norm of the distances of all the vectors with respect to X_{ref}.
5. Find the local minima of the distance curve L corresponding to each recurrence cycle. Local minima correspond to the location of the vectors that constitute the neighborhood (Set a threshold δ for finding the local minima. As the distance becomes less than the δ value, select the point with least distance from X_{ref} in that recurrence cycle as a member of the neighborhood.)
6. Once the locations of the recurrence vectors are known, we can find the recurrence variability– the distance between two adjacent recurrences. Define an array LM to record the recurrence variability.
7. Find the range of recurrence intervals– from LM_{min} to L_{max}
8. If X_{ref} is selected such that the QRS peak is towards the middle, then reset the value of N closer to the LM_{max}. If X_{ref} is selected such that the QRS peak is in the beginning then N value should be lower closer to LM_{min} to satisfy the condition that there should be one and only one peak in that vector.
9. Reset X_{ref}, N and L array repeating steps 2–8 for fine tuning.

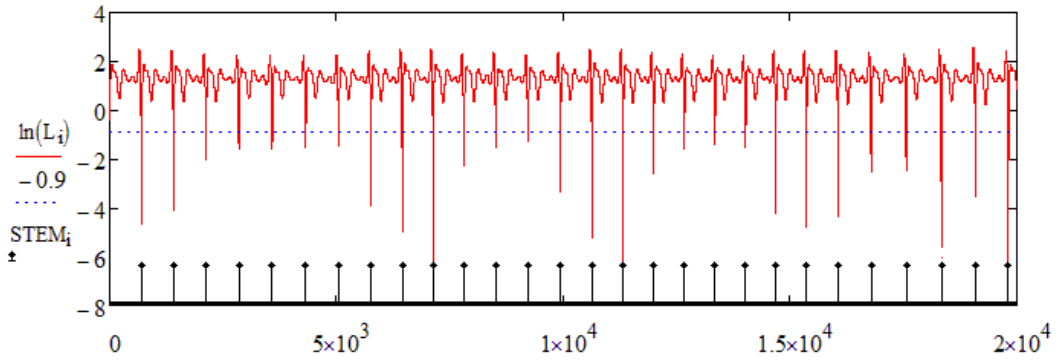

Figure 5.4: L_2 norm of the distance vectors and the location of the recurrent neighbors: L curve is the logarithm of the L_2 norm of the distance vectors. Black stems are the local minima of L that correspond to the recurrent neighbors.

Chosen Parameter Values

The ECG data under study (a short segment is depicted in Fig. 5.1) had a total of 604376 samples and it comprised of 840 recurrent cycles. Frequencies of each of these recurrence intervals are shown in Fig. 5.5. Recurrent intervals were in the range 646–802. Finding the range of recurrence intervals is important as it helps to fine–tune the required dimension N of delay embedding space by appropriately selecting the reference vector X_{ref}. Since the recurrence intervals do not vary too much in this special case of ECG data, there is a possibility of properly adjusting the embedding dimension N such that the recurrent neighborhood vectors in R^N do a proper overlap of the recurrence intervals in the time domain signal. This property is exploited for prediction demonstrated in the analysis section of this chapter.

For the numerical simulations $N = 800$ was set for the delay embedding. (Note: Though the occurrence of intervals greater than 800 is rare, choosing $N = 800$ will give a prediction short of few points in recurrence intervals $> N$. e.g. prediction short of 2 points in the 802 recurrence interval.) Fig. 5.4 depicts the distance of all vectors with respect to a reference vector $X_{ref} \in R^{800}$ and the locations of vectors that constitute the recurrence neighborhood. Though the recurrence locations change according to the selection of the reference vector X_{ref}, the recurrence vari-

ability LM will be invariant. LM is also invariant of N as long as X_{ref} contains one and only one QRS peak. Also note that when X_{ref} is selected as a vector starting with an exact QRS peak, the recurrence locations will be identical to the QRS peak locations of the ECG data. This procedure gives us a flexibility of finding the RR variability as the recurrence variability which is identical to it, is independent of the QRS peak locations. According to this method, one need not restrict to the exact QRS peak locations for finding a recurrence neighborhood. This could be used as an alternative for finding accurate QRS peak locations and heart rate which is an active area of research in biomedical signal processing community [78, 79].

We see in later sections how the entire structure of the state space can be deduced by looking at one small neighborhood. This property may not be generic to all dynamical systems, but for the special case of recurrence observed in ECG signal under study, we show how to derive the nonlinear structure from a single neighborhood by satisfying the conditions for embedding dimension N and the reference vector for the neighborhood X_{ref}.

5.3 Dynamics of the Neighborhoods in R^d and R^N

Once the neighborhood is constructed in the delay embedding space R^N, it is projected to a lower dimensional space R^d as per the model. The linear transformation A of dimension (3×800) is used for the $R^{800} \rightarrow R^3$ projection. For the numerical simulations demonstrated in this chapter parameter values were: $N = 800, d = 3, h_0 = N, h_1 = 2N, n = 0...N-1$.

$$A_{0,n} = Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \tag{5.1}$$

$$A_{1,n} = Imag\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \tag{5.2}$$

$$A_{2,n} = Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_1}\right)\right) \tag{5.3}$$

A topological conjugacy between the recurrence neighborhoods reconstructed in R^N and R^d are demonstrated below. Evolution of the vectors in R^N can be predicted by the corresponding evolution of vectors in R^d. Following sub sections

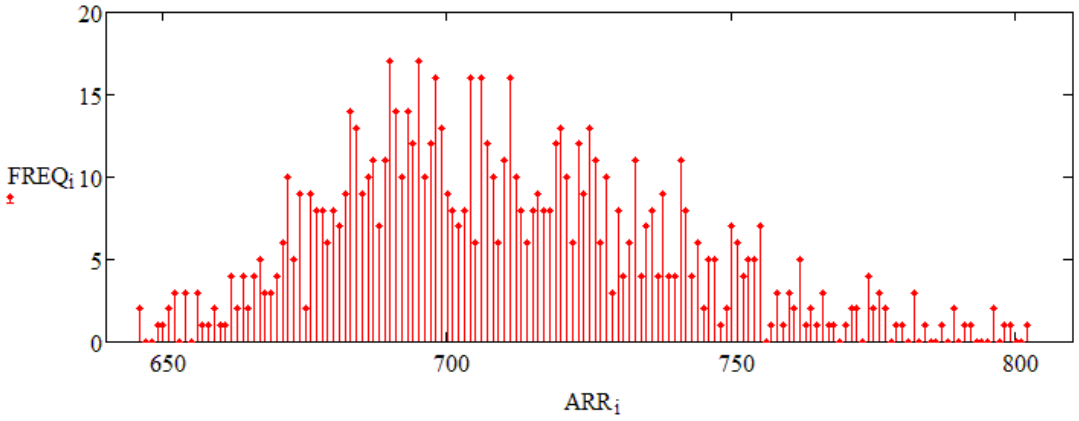

Figure 5.5: Equivalence classes in R^3 for the ECG data. X axis shows the equivalence classes corresponding to recurrence intervals in the range $646 - 802$ and Y axis shows the total number of members in each equivalent class.

explains the process of going back and forth between the lower dimension R^d and the higher dimension R^N. In R^d, members of the neighborhood are classified into different equivalence classes based on their recurrent time τ. One can generate evolution maps in R^d for each of these equivalence classes as shown in Section 5.3.2. The conjugacy map shown in Section 5.3.1 can be used for going back from R^d to R^N.

5.3.1 Conjugacy maps defined across the neighborhoods in R^d and R^N

Projection from R^N to R^d is a simple process using the linear transformation A. But going back from the lower dimension R^d to the higher dimension R^N is not that straight forward especially if the dynamics is nonlinear. This section explains the existence of affine transformations across the neighborhoods in R^d and R^N, and they are used to go from the lower dimensional space to the higher dimensional space.

Let the recurrent neighborhood matrices be denoted as RX in the delay embedding space R^N, and RY in the projected space R^d. For the numerical analysis the values set were $N = 800$ and $d = 3$. A total of 840 recurrence neighbors were available in each of these matrices. Hence RX has dimension 840×800 and RY has

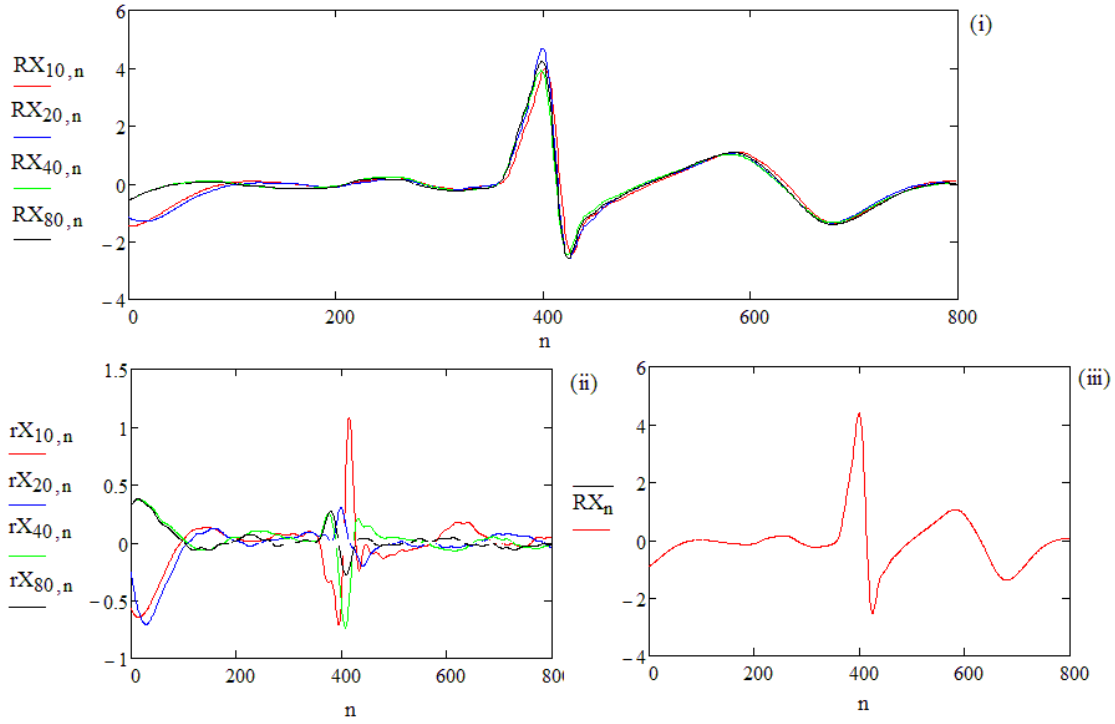

Figure 5.6: A few vectors in the neighborhood of the ECG data in R^{800} before translation to origin (i), after translation to origin (ii) and the centroid $\overline{RX}$ (iii)

dimension 840×3. If the parameters of the transformations are known, RX can be predicted from RY using the map developed in this section. Steps to generate affine transformation across RX and RY are listed below.

Step 1: Find the centroid of the neighborhood RX and RY. Let rX and rY represent the recurrent neighborhood matrices translated to origin. Record the centroid of RX and RY as $\overline{RX}$ and $\overline{RY}$ respectively. A few neighborhood vectors for the Duffing data after translation to origin along with the centroid are shown in Fig. 7.6.

Step 2: Define a conjugacy map T, across the matrices rX and rY such that;

$$(rY).T = rX \tag{5.4}$$

Step 3: rX of dimension 840×800, rY of dimension 840×3 are known matrices; T of dimension 3×800 can be calculated by pre–multiplying rX with the

generalized inverse of rY.

$$T = (rY)^{-1}.rX \tag{5.5}$$

T represents a linear map across the neighborhoods rY, rX shifted to origin. This implies an affine map of the form ($\beta = T\alpha + \overline{RX}$) exists for every vector α in RY and β in RX. Hence the map T and the centroid $\overline{RX}$ are known, a vector in R^{800} can be generated from a vector in R^3. This affine transformation across the neighborhoods RX and RY represents the topological conjugacy between the manifolds in R^{800} and R^3.

5.3.2 Evolution of the equivalence classes in R^3

Members of the recurrence neighborhood in R^3 can be classified into different equivalence classes based on their recurrence intervals. Fig. 5.5 shows the equivalence classes corresponding to various recurrence intervals of the ECG data. Once the equivalence classes in R^3 are ready, one can define transformation for a particular equivalence class. The steps mentioned here are similar to the steps 1-3 in section 5.3.1 except that the transformation V is defined in R^3 alone and the transformation matrix T has a dimension 3×3.

Analysis results for a particular equivalence class: 695 are given in this section. Fig. 5.5 shows that this equivalence class has 17 members. Let the Start and End matrices for this equivalence class in R^3 be denoted by SX, SY respectively. SX contains all the vectors that start with recurrent interval 695 and SY contains all the vectors that end with recurrent timing 695. Hence SX and SY have dimensions of 17×3 each. The explicit steps of finding the transformation across the neighborhoods in R^3 are discussed below.

Step 1: Find the centroid of the neighborhoods SX, SY and subtract it from the neighbors to translate the neighborhood to origin. Let sX and sY represents the neighborhoods translated to origin. Record the centroids of SX and SY as $\overline{SX}$ and $\overline{SY}$ respectively.

SX and SY have dimensions of 102×3 each. Numerical values of few vectors

that constituted SX and SY matrices were,

$$RX^T = \begin{bmatrix} -0.17084 & -0.22662 & -0.1791 & -0.17524 & -0.17983 & \dots \\ -0.00561 & -0.02428 & -0.02579 & 0.0003 & -0.01094 & \dots \\ 0.01052 & 0.007 & 0.04078 & 0.00469 & 0.00687 & \dots \end{bmatrix}.$$

$$RY^T = \begin{bmatrix} -0.18263 & -0.18227 & -0.20581 & -0.18806 & -0.18247 & \dots \\ -0.01661 & 0.00412 & -0.03492 & -0.01039 & -0.01833 & \dots \\ 0.01627 & -0.00361 & 0.02956 & 0.00557 & 0.01384 & \dots \end{bmatrix}.$$

$$S\bar{X} = \begin{bmatrix} -0.19318 \\ -0.02045 \\ 0.02182 \end{bmatrix}$$

$$S\bar{Y} = \begin{bmatrix} -0.19861 \\ -0.01841 \\ 0.01356 \end{bmatrix}$$

Step 2: Define a transformation V, across the matrices sX and sY such that;

$$sX.V = sY \tag{5.6}$$

Step 3: sX, sY both of dimension 17×3 are known matrices; V of dimension 3×3 can be calculated by pre–multiplying sY with the generalized inverse of sX. The generalized inverses of sX were found using both linear and nonlinear SVD procedure.

$$V = (sX)^{-1}.sY \tag{5.7}$$

Note that V is a function of the recurrent time τ. Empirically determined V matrix for the 695 equivalence class ($\tau = 695$) was,

$$V = \begin{bmatrix} 0.56279 & -0.86352 & -1.29201 \\ -0.51484 & 0.6968 & 0.08961 \\ 0.56033 & -0.92404 & -0.33528 \end{bmatrix}.$$

Step 4: Once the V matrix is estimated, prediction for any random starting vector in 695 equivalence class can be made. V multiplied by an R^3 vector of dimension 3×1 creates another vector in R^3. Given any random vector $\alpha \in SX$ subtract the centroid $\overline{SX}$ to translate it to origin. Let $r\alpha$ be the translated vector.

$$\alpha - \overline{SX} = r\alpha \tag{5.8}$$

$$\alpha = \begin{bmatrix} -0.17768 \\ -0.03366 \\ 0.0512 \end{bmatrix}$$

$$r\alpha = \begin{bmatrix} 0.0155 \\ -0.01321 \\ 0.02937 \end{bmatrix}$$

Step 5: If V is known $r\alpha \in sX$ can be evolved to $r\beta \in sY$ according to the equation,

$$r\alpha.V = r\beta \tag{5.9}$$

$$r\beta = \begin{bmatrix} -0.01782 \\ -0.01455 \\ 0.01104 \end{bmatrix}$$

Step 6: $r\beta$ belongs to the 695–End equivalence in R^3. The prediction for $\beta \in SY$ the original neighborhood can be made by adding centroid vector $\overline{SY}$ for that equivalence class as,

$$r\beta + \overline{SY} = \beta \tag{5.10}$$

$$\beta = \begin{bmatrix} -0.21643 \\ -0.03296 \\ 0.0246 \end{bmatrix}$$

Hence following steps 1–6 one can find the evolution of a vector that belongs to the 695 equivalence class. In this manner evolution maps for all the equivalence

classes can be found.

5.3.3 Evolution in R^d and going back to R^N

This section explains the process of going back and forth between the lower dimension R^d and the higher dimension R^N. Projection of the recurrent neighborhood from R^N to R^d is simple using the linear transformation A. In R^d, members of the neighborhood are classified into different equivalence classes based on their recurrent times. One can generate evolution maps in R^d for each of these equivalence classes as shown in section 5.3.2 and conjugacy maps developed in section 5.3.1 can be used for going back from R^d to R^N. Stitching the R^N vectors appropriately at their corresponding neighborhood locations according to the recurrent times gives a prediction for the time domain signal.

Algorithm 4.2 given in Chapter 4 can be used for predicting the evolution of vectors in R^N using their evolution in R^d given the recurrent timings and the maps.

5.4 Results of Data Analysis

This section contains the simulation results of the modeling and prediction algorithm for the ECG data. The data was embedded into R^N and then projected to R^d as per the model. The empirical values selected for the parameters for the numerical simulation results of this section are: $N = 800$ and $d = 3$. Conjugacy maps that take R^3 vectors to R^N vectors and evolution maps in R^3 were determined as explained in previous section. The techniques of standard and nonlinear SVD were used for generating the evolution maps in R^d. Predicted signals were compared with the original signals using two scores; $Q1$ based on mean squared error deviation and $Q2$ based on correlation (more details on scoring functions are given in Section 4.5.1 of Chapter 4). Prediction results for few random vectors are demonstrated with figures for both the cases of linear and nonlinear SVD based prediction. Prediction using the standard SVD is discussed in section 5.4.1 and nonlinear SVD is discussed in section 5.4.2. A comparison of the errors for each pair these predictions are given in section 5.4.3. Section 5.4.5 contains the scores for the entire vectors that belong to the equivalence class of recurrence interval 695.

5.4.1 Prediction results using standard SVD

This section explains the results where the transformation matrix in Eq. 5.14 was found by a generalized inverse based on standard linear SVD.

Prediction from R^3 to R^{800}: Here we show the prediction for two random members: the 60^{th} member and the 150^{th} member in the R^{800} neighborhood. In Fig. 5.7, $P1$ is the affine prediction, RM_F is the centroid for the R^{800} neighborhood, and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.7 that the Affine predictions improves the centroid fit. Scores of affine prediction for the 60^{th} member were $(Q1 = 0.9830, Q2 = 0.9957)$ and that of centroid fit were $(Q1 = 0.9589, Q2 = 0.9906)$. Scores of affine prediction for the 150^{th} member were $(Q1 = 0.9945, Q2 = 0.9981)$ and that of centroid fit were $(Q1 = 0.9668, Q2 = 0.9855)$.

5.4.2 Prediction results using nonlinear SVD

This section explains the results where the transformation matrix in Eq. 5.14 was found by a generalized inverse based on nonlinear linear SVD. **Prediction from R^3 to R^{800}:** Here we show the prediction for two random members: the 60^{th} member and the 150^{th} member in the R^{800} neighborhood using their R^3 counterparts. In Fig. 5.8, $P2$ is the affine prediction, RM_F is the centroid for the R^{800} neighborhood, and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.8that the Affine predictions improves the centroid fit. Scores of affine prediction for the 60^{th} member were $(Q1 = 0.9897, Q2 = 0.9958)$ and that of centroid fit were $(Q1 = 0.9589, Q2 = 0.9906)$. Scores of affine prediction for the 150^{th} member were $(Q1 = 0.9969, Q2 = 0.9987)$ and that of centroid fit were $(Q1 = 0.9668, Q2 = 0.9855)$.

Scores of R^3 to R^{800} Prediction for the entire equivalence class: The prediction scores for both the linear and nonlinear SVD based prediction for entire members of a particular equivalence class was found. Fig. 5.9 shows Q1 and Q2 scores of 17 members of 695 equivalence class for R^3 to R^{800} prediction.

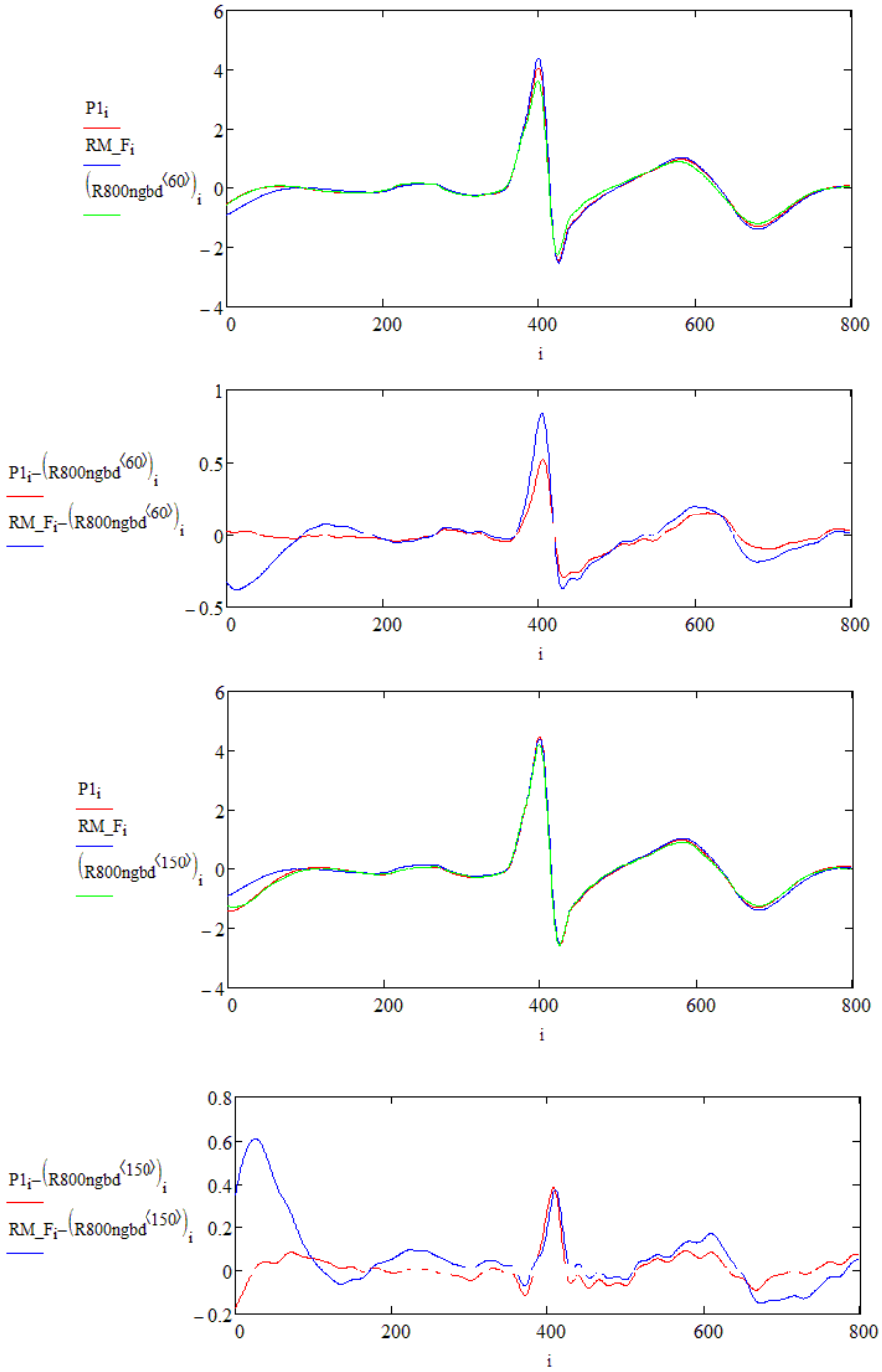

Figure 5.7: $P1$ **is the affine prediction, RM_F is the centroid for the** R^{800} **neighborhood, and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.7that the Affine predictions improves the centroid fit. Scores of affine prediction for the** 60^{th} **member were** $(Q1 = 0.9830, Q2 = 0.9957)$ **and that of centroid fit were** $(Q1 = 0.9589, Q2 = 0.9906)$**. Scores of affine prediction for the** 150^{th} **member were** $(Q1 = 0.9945, Q2 = 0.9981)$ **and that of centroid fit were** $(Q1 = 0.9668, Q2 = 0.9855)$**.**

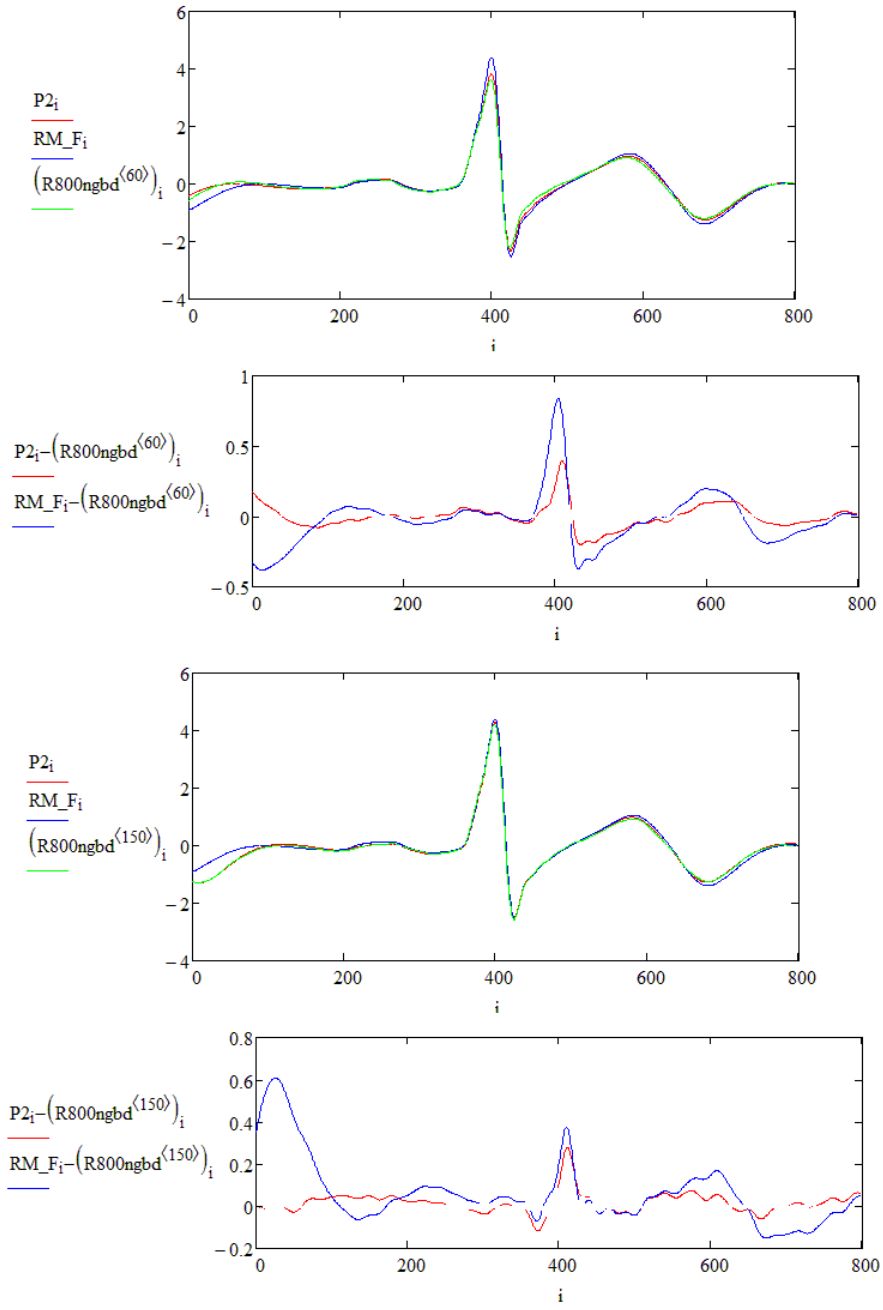

Figure 5.8: $P2$ **is the affine prediction, RM_F is the centroid for the** R^{800} **neighborhood, and the error of the original signal with respect to the predicted signal and the centroid vector also shown. It is clear from the Fig. 5.8that the Affine predictions improves the centroid fit. Scores of affine prediction for the** 60^{th} **member were** $(Q1 = 0.9897, Q2 = 0.9958)$ **and that of centroid fit were** $(Q1 = 0.9589, Q2 = 0.9906)$**. Scores of affine prediction for the** 150^{th} **member were** $(Q1 = 0.9969, Q2 = 0.9987)$ **and that of centroid fit were** $(Q1 = 0.9668, Q2 = 0.9855)$

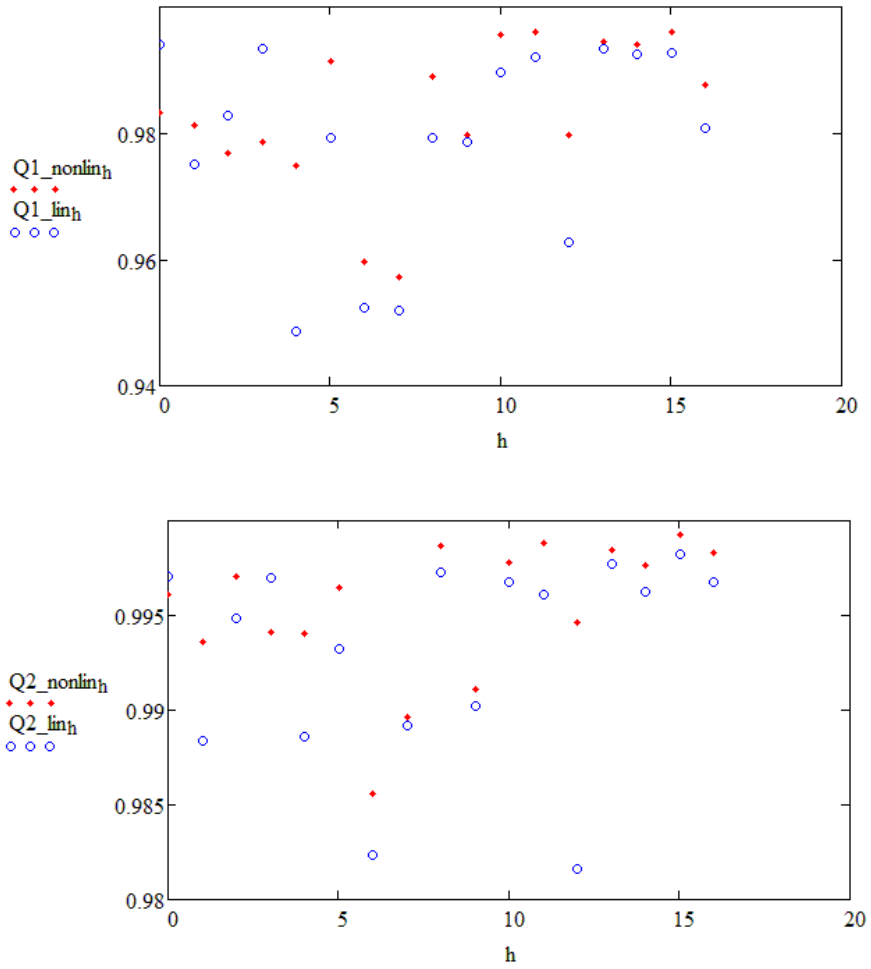

Figure 5.9: Scores $Q1$ and $Q2$ for both the linear and nonlinear SVD based prediction in R^{800} for the 17 members of 695 equivalence class of vectors in R^3 evolution.

5.4.3 Error Variances for the Prediction based on linear and nonlinear SVD

Given that the centroid based prediction has decent scores, one can use that if no much information is available. Affine transformations based on the standard SVD gave better prediction. Nonlinear SVD based affine transformations are the best choice as it improved the scores significantly. Error variance for these three cases of predictions were also found (more details about the error variance function is given in Section 4.5.4 of Chapter 4). Variance of errors for the three specific cases of prediction were calculated. For R^3 to R^{800} prediction of the 50^{th} and 150^{th}

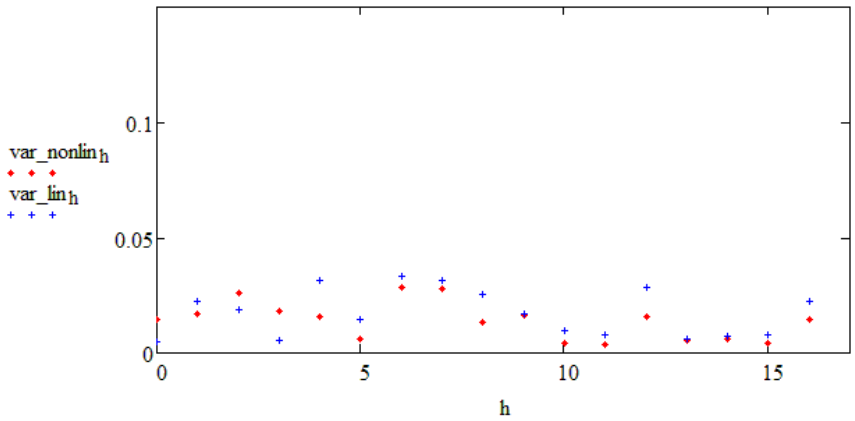

Figure 5.10: variance of errors for both the linear and nonlinear SVD based prediction in R^{800} for the 17 members of 695 equivalence class of vectors in R^3 evolution.

Table 5.1: Variance of errors between the original and predicted signal (for the three cases P1, P2, RM_F) for R^3 to R^{800} prediction of the 50^{th} and 150^{th} members

Neighborhood Vector	P1(std SVD)	P2(nonlinear SVD)	RM_F(centroid fit)
50	0.01273	0.00707	0.03515
150	0.0050	0.00239	0.02503

members, variance of errors (for the three cases P1, P2, RM_F depicted in figures) are listed in Table 5.1. Fig. 5.10 shows the variance of errors for both the linear and nonlinear SVD based prediction in R^{800} for the same 17 members of 695 equivalence class of vectors in R^3 evolution.

5.5 Topological structure of ECG signals

In the Analysis section we have shown the numerical results for ECG data projected from a higher dimension R^{800} to the lower dimension R^3. We have also analyzed the results for projections for various lower dimensions i.e. for $d = 2, 3, 4$. Based on numerical simulations for data from $R^{800} \rightarrow R^2$ and $R^{800} \rightarrow R^4$, we have observed that the prediction based on the equivalence classes in R^2, R^3 gave excellent scores with a slight decrease in R^4 prediction scores. Simulation results for $R^{800} \rightarrow R^3$ projection were explicitly demonstrated for the ECG data in this

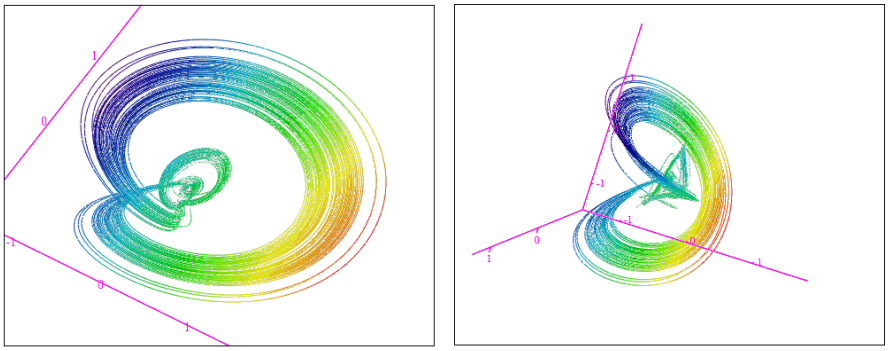

Figure 5.11: State space reconstructed in R^3

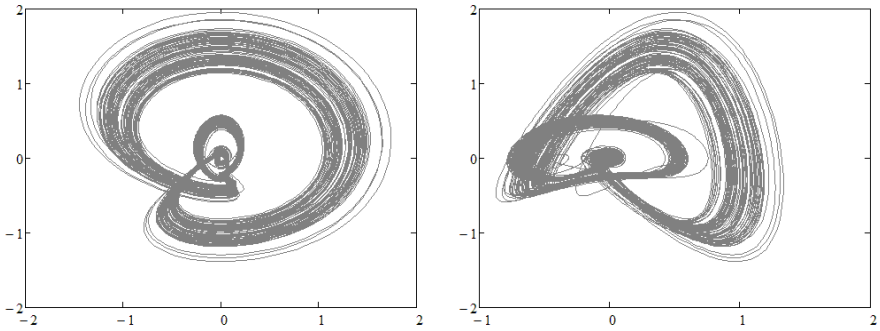

Figure 5.12: State space reconstructed in R^2

chapter, for the reader to make a comparison of results for the similar $R^{800} \rightarrow R^3$ projection of the duffing oscillator generated data in the previous chapter. We have noticed some properties of the transformations specific to the equivalence classes and that will be discussed in the stability analysis section of Chapter 6.

As we have discussed earlier ECG signal exhibits both stochastic and non-linear deterministic properties. In a nonlinear deterministic perspective, empirical estimates of correlation dimension of ECG data have been quite high– 15 or more [80]. In this chapter, we presented an analysis that takes an alternative perspective. We propose a model for a healthy heart, where the whole heart dynamics can be separated into two: (i) the dynamics that generate RR intervals (ii) the regulating effect of RR interval on pumping action of heart. It is not clear if the dynamics behind (i) is chaotic or stochastic as by analyzing a random looking signal alone we cannot specify if it is generated by deterministic chaos or a stochastic process as they

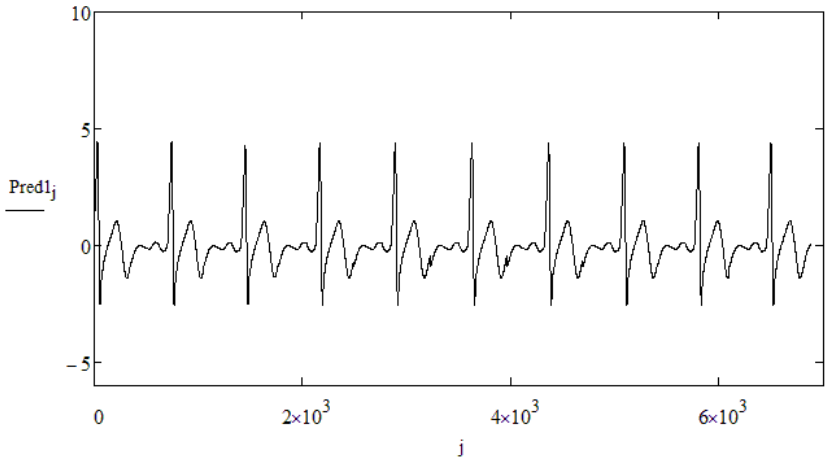

Figure 5.13: Reconstruction the time domain signal by stitching the embedding vectors appropriately at their neighborhood locations

are shown to be equivalent to each other [81]. But once the dynamics behind the impulses are kept apart, the rest of the heart dynamics is fairly simple with a lower dimensional structure. In Fig. 5.11 and 5.12 we show statespace of the ECG signal in the lower dimensional spaces R^3 and R^2. In both the $2D$ and $3D$ pictures we can see that the ECG data lies in a ribbon–like structure. Existence of this manifold like structure was utilized for the predictions demonstrated in this chapter. In a deterministic dynamics perspective we can see that the neighborhood points of the state space have similar future. Hence we assume that the equivalence classes as the collection of similar points who have similar future. Once we have a collection of these equivalence classes, the impulses determines the switching between them. For this ECG data, we found that a very good prediction can be obtained by using our model and some additional information about the exact occurrences of recurrence. Stitching the embedding vectors appropriately at their corresponding neighborhood locations (given by the recurrent times) we can reconstruct the time domain signal as shown in Fig. 5.13. Due to the specific property of these model maps, once they are known, one could start at any random initial guess and then converge to the actual data. Fig. 5.14 shows the convergence of 2 predictions from two random different initial conditions.

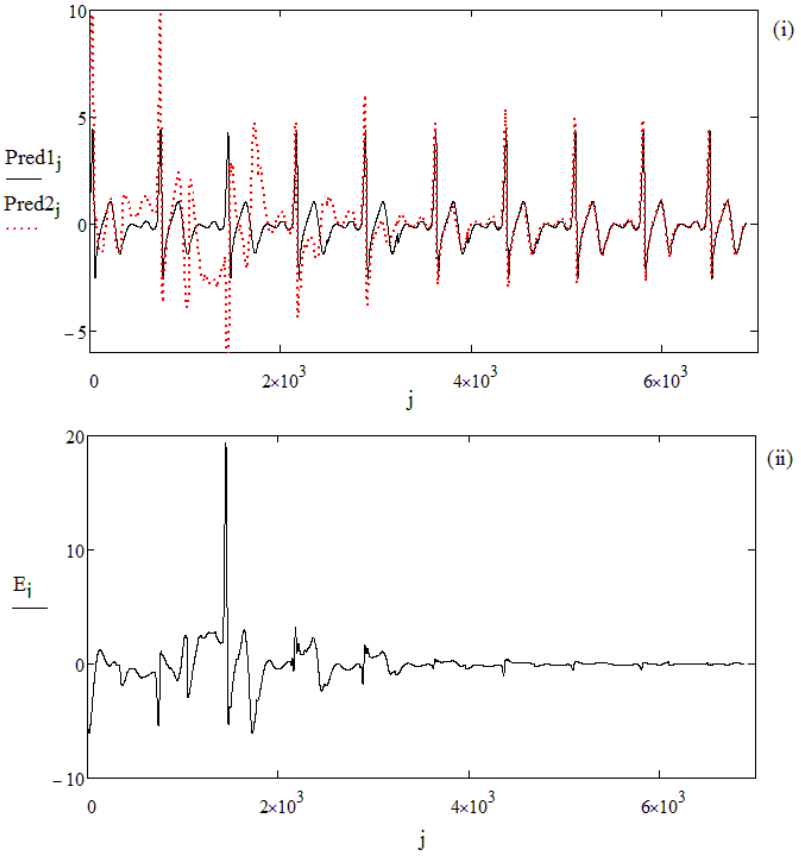

Figure 5.14: Convergence of 2 different predictions based on two different initial conditions (i) and the difference between the predictions $E = Pred2 - Pred1$ (ii).

5.6 Conclusion

We have demonstrated a method to predict ECG signal based on the (i) evolution maps specific to the equivalence classes in a lower dimension and (ii) the conjugacy maps that exists across the lower and higher dimensions. We have found that once the dynamics that is responsible for the variability in recurrences are separated, rest of the heart dynamics is very simple. We have shown that the data lies in a ribbon like manifold structure in a low-dimensional space which can be exploited for the purpose of prediction. In addition to that we have found some properties of the evolution maps in lower dimension due to which the predictions from any random initial guesses converges to the actual data. This property of the maps will be is explored in the stability analysis section of next chapter.

CHAPTER 6
Stability Analysis based on Maps in the Recurrence Neighborhood

An empirical method to develop affine maps specific to neighborhoods for data generated by recurrent dynamical systems was explained in the last 2 chapters using (i) a data series generated by the Duffing oscillator and (ii) an ECG measurement of the heart. This chapter analyzes the properties of these maps for the stability analysis based on data. A numerical method for calculating the floquet estimates from the data are explained and the fact that –representing a system in a dimension higher than the inherent dynamics of the system causes wrong inferences about the stability of the system– is verified for the Duffing data. In the case of ECG data, an estimate for the local dimension was made using the conjugacy property as the exact dimension of the dynamics is unknown. A theoretical justification for an experimental observation that the Heart Rate Variability (HRV) implies a healthy functioning of the heart is given in this chapter. It is demonstrated that HRV implies a stable dynamics where as a uniform heartbeat could result in instability based on the properties of the affine maps derived from the data.

6.1 Introduction

Previous chapters demonstrated that the set of the affine transformations (specific to the equivalence classes of recurrent neighborhood) and recurrence timings (intervals between consecutive recurrences) can be used for prediction and this was verified for the duffing and the ECG data. One important result that came out of the study was that the structure of the state space can be deduced by analyzing one or a few recurrence neighborhoods. The data generated by the duffing oscillator as specifically used to demonstrate the model as (i) it is a well studied system in non-linear dynamics and (ii) for the stability analysis, floquet exponents of the system can be calculated and verified experimentally as their sum is equal to trace of the coefficient matrix of the system (for duffing oscillator under study, it is equal to $-c$

where c is the damping coefficient of the system) [50]. One of the objectives of this chapter is to demonstrate an empirical method to find the floquet coefficient of the system based on a recurrent neighborhood. We further verify that if the estimated local dimension of the manifold is correct, the floquet exponents of the system can be accurately retrieved from the data.

Section 5.6 of Chapter 5 has briefly mentioned the convergence property of the affine maps specific to the ECG signal. For this case, once the maps and the recurrence timing are known, one could start at any random initial guess and then converge to the actual data. Hence if the maps are known, any information about the initial conditions are eventually unnecessary. It is demonstrated in this chapter that this convergence is a property of asymptotically stable maps. Another objective of this chapter is to explain a long standing puzzle about ECG signals is using this property of the affine maps. It has been found that a good HRV is an indication of a healthy functioning of the heart [75]. Using the maps derived from actual data we show that uniform recurrence intervals can indeed lead to instability, whereas non-uniform intervals might promote asymptotic stability. Here stability implies as situation where the individual heart cells in the ventricles are synchronized with each other thus inducing the rapid contraction of ventricles which is essential for discharging blood out of the heart. Instability could be seen as situation where the synchrony of the individual heart cells are lost thus resulting in a quivering of ventricles rather than a proper contraction.

In what follows, Sections 6.2 gives a brief review about the stability of equilibrium points and periodic orbits in a dynamical system. An empirical method to estimate floquet coefficients from data is given in Section 6.3. The method was verified by accurately estimating the floquet coefficients for the Duffing system. Section 6.4 demonstrates an extension of this method to calculate the floquet exponents for recurrent neighborhoods reconstructed in delay embedding space and various lower dimensional projected spaces. Stability analysis for the ECG data is explained in Section 6.5. Some theoretical details about the stability for Composition of maps are discussed in section 6.6. Section 6.7 discusses the stability of the composition of the affine maps specific to the ECG data under study and Section 6.8 is the conclusion.

6.2 Stability of equilibrium points and periodic orbits of a dynamical system

Nonlinear systems can have multiple equilibria and special solutions like periodic orbits. In general, a dynamical system can be said to be stable around its equilibrium points and periodic orbits if small perturbations around them do not significantly change the behaviour of the system. Consider a dynamical system represented by,

$$\dot{X} = F(X), \quad X \in R^d \tag{6.1}$$

A point $\bar{X}$ of this is system is said to be an equilibrium point if $F(\bar{X}) = 0$. Hence the system is at $\bar{X}$ at time t_0, it remains there for all time $t > t_0$.

The stability of equilibrium points in general can be defined based on the following notions of Lyapunov and Asymptotic stability. The system is said to be Lyapunov stable around $\bar{X}$ if for every $\epsilon > 0$, there exists a $\delta_1(\epsilon) > 0$ such that, if $\|x(0) - \bar{X}\| < \delta_1$, then $\|x(t) - \bar{X}\| < \epsilon$ for all $t \geq 0$. According to this condition, the solutions in the neighborhood of an equilibrium point $\bar{X}$ stay near $\bar{X}$ forever. Further the system is said to be Asymptotically stable around $\bar{X}$ if (i) it is lyapunov stable and (ii) also satisfies the condition of convergence that there exists $\delta_2 > 0$ such that if $\|x(0) - \bar{X}\| < \delta_2$, then $\|x(t) - \bar{X}\| \to 0$ as $t \to \infty$ [82].

Stability of a periodic orbit can be also defined in a similar manner as that of the equilibrium point. A poincare map P can be defined for the flow given by Eq. 6.1 as follows. Let S be a $d-1$ dimensional surface defined transverse to the flow of the system in R^d. P is a mapping defined on this surface S to itself such that,

$$x_{k+1} = P(x_k) \tag{6.2}$$

Where x_k is the k^{th} intersection of the flow on S. If $\bar{x}$ is a fixed point of the n^{th} return map P^n such that,

$$\bar{x} = P^n(\bar{x}) \tag{6.3}$$

Then $\bar{x}$ is said to be a period–n orbit of the system. Since periodic orbits can be seen as fixed points of the poincare map, it can be said to be stable if the

corresponding poincare map is stable at $\bar{x}$ [2].

Since a periodic orbit can be seen as a fixed point of the corresponding poincare map P, the periodic orbit is said to be stable if P is stable at the corresponding fixed point $\bar{x}$. To find the stability at $\bar{x}$ of P, consider a small perturbation δ_0 around it. Let the first iteration of δ_0 be δ_1 such that,

$$\bar{x} + \delta_1 = P(\bar{x} + \delta_0)$$

Taylor series expansion of the map P at $\bar{x}$ is,

$$P(\bar{x} + \delta_0) = P(\bar{x}) + JP(\bar{x})\delta_0 + O(\delta_0^2)$$

$JP(\bar{x})$ represents the Jacobian of P (linearisation of P) at $\bar{x}$. Neglecting the higher order terms $O(\delta_0^2)$, one can get a linear map for the evolution of perturbation δ as,

$$\delta_n = JP(\bar{x})\delta_{n-1} \tag{6.4}$$

Thus the evolution of perturbations $\{\delta_0, \delta_1, \ldots \delta_n \ldots\}$ depends upon the eigen values $\{\lambda_i\}$ of $JP(\bar{x})$. All perturbations will die out if all the eigen values are < 1 thus indicating the stability of the periodic orbit. These eigen values, that determine the stability of the orbit are known as characteristic multipliers or floquet multipliers of the corresponding periodic orbit [2]. Corresponding $\{\rho_i\}$ such that $\{\lambda_i = e^{\rho_i T}\}$ where T is the period of the orbit are called the floquet exponents of the system [83].

6.3 An empirical method to estimate floquet coefficients from data

Floquet multipliers that determine the stability of a periodic orbit of a dynamical system can be found in general only by numerical integration [2]. According to the floquet theory, if $\Phi(t)$ is the fundamental matrix of solutions of the dynamical system, then there exists a non–singular monodromy matrix C associated with the periodic solutions (of period T) such that $\Phi(t+T) = \Phi(t)C$. The eigen values $\{\lambda_i\}$ of C are known as the floquet multipliers of the system. Corresponding $\{\rho_i\}$ such

that $\{\lambda_i = e^{\rho_i T}\}$ are called the floquet exponents of the system [83]. The monodromy matrix C is analogous to the linearized Poincare map JP at the periodic orbit of the system mentioned by Eq. 6.4.

This section gives a numerical method to calculate the floquet exponents from the solutions of the system near a periodic orbit. For the simplicity of the equations mentioned in this section, without loss of generality assume that the given dynamical system belongs to R^3. Let the $2D$ affine vector space that represents the floquet modes in a small neighborhood around the periodic orbit of the system be:

$$\mathbf{x} = f + ae^{\rho_1 t}p + be^{\rho_2 t}q \tag{6.5}$$

Note that f, p, q are periodic functions of time, a, b are scalar constants, and ρ_1, ρ_2 are the floquet exponents. Corresponding linear vector space can be obtained by removing the constant function f which represents the trivial floquet mode (the floquet multiplier with value 1 as it corresponds to the perturbations along the periodic orbit). Thus the linear vector space that represents the nontrivial floquet modes is,

$$x = ae^{\rho_1 t}p + be^{\rho_2 t}q \tag{6.6}$$

Substituting $g = e^{\rho_1 t}p, h = e^{\rho_2 t}q$; where g, h are constant functions that do not change for this vector space, Eq. 6.6 can be written as,

$$x = ag + bh \tag{6.7}$$

The coefficients a, b represent the proportions how g, h vary for every members of this vector space. Let x_1, x_2 be two specific members of this vector space,

$$x_1 = a_1 g + b_1 h \tag{6.8}$$

$$x_2 = a_2 g + b_2 h \tag{6.9}$$

Define y as the evolution of x after a time period T such that $y(t) = x(t+T)$.

$$\begin{aligned} y(t) &= x(t+T) \\ &= ag(t+T) + bh(t+T) \\ &= ae^{\rho_1 T} g(t) + be^{\rho_2 T} h(t) \end{aligned}$$

Denote $\lambda_{\mathbf{1}} = e^{\rho_1 T}$ and $\lambda_{\mathbf{2}} = e^{\rho_1 T}$. $\lambda_{\mathbf{1}}, \lambda_{\mathbf{2}}$ represents the floquet multipliers around the specific periodic orbit with period T.

$$y = a\lambda_{\mathbf{1}} g + b\lambda_{\mathbf{2}} h \tag{6.10}$$

For two specific members y_1, y_2 of this space,

$$y_1 = a_1\lambda_{\mathbf{1}} g + b_1\lambda_{\mathbf{2}} h \tag{6.11}$$

$$y_2 = a_2\lambda_{\mathbf{1}} g + b_2\lambda_{\mathbf{2}} h \tag{6.12}$$

Eqs. 6.11 and 6.12 can be re-written in matrix form as,

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} = \begin{bmatrix} g & h \end{bmatrix} \begin{bmatrix} a_1\lambda_{\mathbf{1}} & a_2\lambda_{\mathbf{1}} \\ b_1\lambda_{\mathbf{2}} & b_2\lambda_{\mathbf{2}} \end{bmatrix} \tag{6.13}$$

Similarly the sets of equations Eq. 6.8, 6.9 can be written in matrix form as,

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} g & h \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \tag{6.14}$$

Define a transformation C across $\begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 & y_2 \end{bmatrix}$ as,

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} C = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \tag{6.15}$$

Substituting Eq. 14 and 15,

$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} C = \begin{bmatrix} a_1\lambda_{\mathbf{1}} & a_2\lambda_{\mathbf{1}} \\ b_1\lambda_{\mathbf{2}} & b_2\lambda_{\mathbf{2}} \end{bmatrix} \tag{6.16}$$

RHS of this equation can be also written as,

$$\begin{bmatrix} \lambda_{\mathbf{1}} & 0 \\ 0 & \lambda_{\mathbf{2}} \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} = \begin{bmatrix} a_1\lambda_{\mathbf{1}} & a_2\lambda_{\mathbf{1}} \\ b_1\lambda_{\mathbf{2}} & b_2\lambda_{\mathbf{2}} \end{bmatrix} \tag{6.17}$$

Hence,

$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} C = \begin{bmatrix} \lambda_{\mathbf{1}} & 0 \\ 0 & \lambda_{\mathbf{2}} \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \tag{6.18}$$

This is of the form $AC = LA$, then $C = A^{-1}LA$ which indicates that C, L are similarity transformations that share identical eigen values. Thus $\lambda_{\mathbf{1}}, \lambda_{\mathbf{2}}$ are the eigen values of C.

If x_1, x_2 and y_1, y_2 are numerical data of length k, then $\begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 & y_2 \end{bmatrix}$ are matrices of dimension $k \times 2$ each. Hence matrix C of dimension 2×2 can be determined using Eq. 6.15 as,

$$C = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{-1} \begin{bmatrix} y_1 & y_2 \end{bmatrix} \tag{6.19}$$

Since $\begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 & y_2 \end{bmatrix}$ represents the empirical values (of evolution of perturbations around the orbit), the monodromy matrix C and the floquet exponents can be empirically determined. Next section uses this method to numerically estimate the floquet coefficients for a Duffing oscillator data.

6.3.1 Verification of the method to find the Floquet Exponents for Duffing Oscillator from a data

Consider a data generated by the Duffing oscillator,

$$\frac{dX}{dt} = V$$

$$\frac{dV}{dt} = -cV - kX - \delta X^3 + Fcos(\omega t + \alpha)$$

for the parameter values $c = 0.04496$, $k = 0$, $\omega = 0.44964, \delta = 1$, $\alpha = 0$, $F = 0.7$.

For the numerical simulation results explained in this section, forcing period was set 13.9738 seconds and the number of points per period was 500. A period 3 trajectory of the system and a small neighborhood around that orbit were identified. Observe the evolution of n trajectories whose initial conditions are in the identified neighborhood. Numerical results using 2 neighboring trajectories (for $n = 2$) are given below. Identify a block of numerical data of length $k = 800$ each, for x_1, x_2 and y_1, y_2. Construct $\begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 & y_2 \end{bmatrix}$ matrices of dimension $k \times 2$ using this data.

The estimated C matrix of dimension 2×2 using Eq. 6.19 was,

$$\begin{bmatrix} 2.3863 & 5.2291 \\ -1.1993 & -2.5644 \end{bmatrix}$$

And the eigen values of the C matrix were, $\begin{bmatrix} -0.08905 - 0.37937i \\ -0.08905 + 0.37937i \end{bmatrix}$.

The eigen values of C matrix (denoted by $\lambda_{\mathbf{1}} = e^{\rho_1 T}$ and $\lambda_{\mathbf{2}} = e^{\rho_1 T}$) represents the floquet multipliers for the specific period (T) under consideration. The floquet exponents ρ_1, ρ_2 of duffing oscillator can be extracted from the eigen values as follows. If λ_1, λ_2 represents the eigen values of C for the period $3T$ (where

$T = 13.9738$) orbit then,

$$\rho_1 = \frac{ln(\lambda_1)}{3T} = -0.022484 - 0.04297i \tag{6.20}$$

$$\rho_2 = \frac{ln(\lambda_2)}{3T} = -0.022484 + 0.04297i \tag{6.21}$$

These empirically determined values can be verified to be correct as their sum is equal to the trace $= -c$ of the duffing oscillator (where c is the damping constant) under study. The sum was $\rho_1 + \rho_2 = -0.044964$ and the chosen parameter value for c was 0.04496.

If more number of neighbors ($n > 2$) are selected from the neighborhood, the dimension of C ($n \times n$) changes as it depends on the number of neighbors n. Minimum number of trajectories required for simulation is 2. The non-zero eigen values of C matrix remains invariant for $n \geq 2$. As the number of neighbors are increased $n > 2$, non–zero eigen values remains the same and the rest of $n-2$ values are exactly zero.

The results can be reproduced for different values of the forcing period F and periodicity T of the orbit. For a period 1 orbit where $F = 0.4$, the eigen values of the C matrix were,

$$\begin{bmatrix} -0.57926383 + 0.444915i \\ -0.57926383 - 0.444915i \end{bmatrix}.$$

Corresponding floquet exponents were,

$$\rho_1 = \frac{ln(\lambda_1)}{T} = -0.0224814 + 0.177949i \tag{6.22}$$

$$\rho_2 = \frac{ln(\lambda_2)}{T} = -0.0224814 - 0.177949i \tag{6.23}$$

And the sum of the exponents $\rho_1 + \rho_2 = -0.044963$ whose magnitude almost equals damping constant c as expected. Hence the numerical method successfully retrieved the floquet exponents of the Duffing oscillator system from a numerical data.

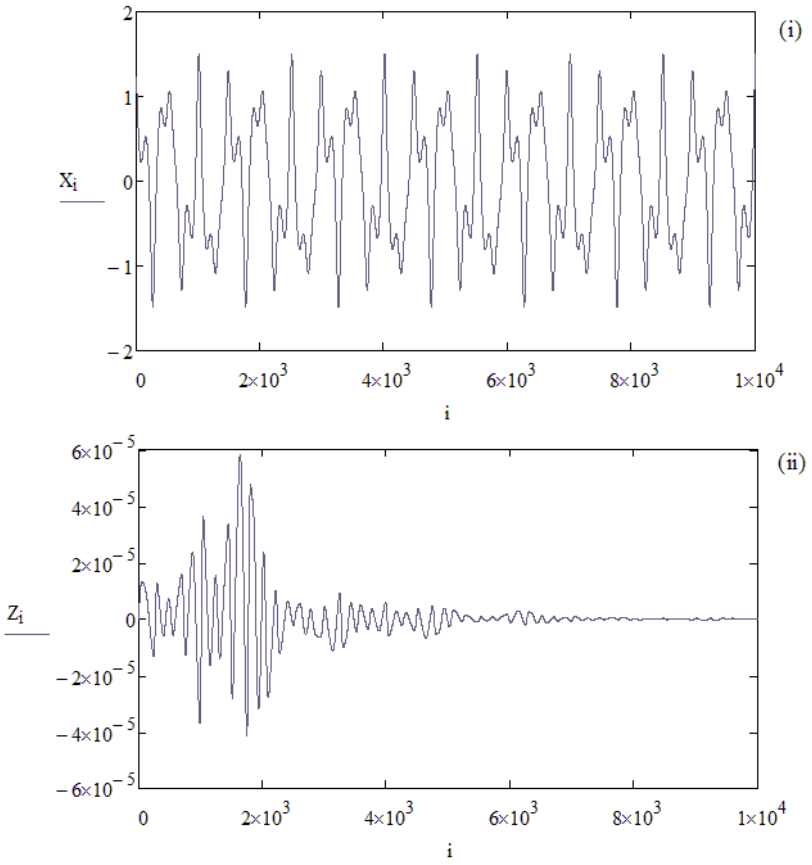

Figure 6.1: Evolution of a recurring orbit in the neighborhood of an attracting period 3 orbit of the Duffing oscillator (i). Evolution of the perturbations that finally settles in the periodic orbit (ii).

6.4 Floquet Exponents in a Recurrence Neighborhood

Floquet coefficients that determine the stability of a periodic orbit can be empirically determined by analyzing a neighborhood around the orbit of the system. Now assume that all we have is a measurement of the system but have no information about the dynamics that generated the data. The new proposed model (demonstrated in chapters 4 and 5) requires a delay embedding of the data in a higher dimension to reconstruct the global dynamics and further projection to a lower dimension for analyzing the local dynamics. This section focusses on the stability analysis of the recurrent orbits in recurrent neighborhoods reconstructed in both the higher and the lower dimensional spaces.

Consider a recurring orbit of the Duffing oscillator in the neighborhood of a periodic orbit. If the periodic orbit is attractive the trajectory eventually settles on the periodic orbit. Evolution of the recurring orbit X in the neighborhood of the period 3 orbit is shown in Fig. 6.1. The periodic orbit is attracting as the perturbations Z eventually vanish and the trajectory X finally settles on the orbit.

6.4.1 Recurrent Neighborhood in the higher dimensional space R^N

Construct a delay embedding of X in a higher dimensional space R^{800} and identify a neighborhood (as explained in section 4.6 of chapter 4). Create a matrix RX and RY of all the vectors that denote the start and end matrices of the equivalence class corresponding to period 3 recurrence (as explained in section 4.8 of chapter 4).

The data segment X had 60000 samples. For the numerical simulations explained in this section, 39 neighbors members were identified for the equivalence class in R^{800}. Thus the dimensions of RX and RY matrices were 800×39. Follow the steps explained below to find the transformation across RX and RY specific to the period 3 interval.

Step 1: Find the centroid of each of the neighborhoods and subtract it from the neighborhood vectors for translating the neighborhoods to origin. Let rX and rY matrices represents the neighborhoods (RX and RY matrices) translated to origin. Record the centroid of the neighborhoods RX and RY as $\overline{RX}$ and $\overline{RY}$.

Step 2: Define a transformation Γ, across rX and rY matrices such that;

$$rX.\Gamma = rY \tag{6.24}$$

rX, rY are of dimension 800×39 each. Γ of dimension 39×39 can be numerically estimated by pre-multiplying rY with the generalized inverse of rX.

$$\Gamma = (rX)^{-1}.rY \tag{6.25}$$

The eigen values of Γ matrix were,

$$\begin{bmatrix} -0.08900824983137 + 0.37937254166312i \\ -0.08900824983137 - 0.37937254166312i \\ -0.18447707434417 \\ 0.10496200761857 \\ -0.00002287163165 \\ 0.00000000000005 + 0.00000000001064i \\ 0.00000000000005 - 0.00000000001064i \\ 0.00000000000623 \\ \dots \\ \dots \end{bmatrix}$$

Since the floquet multiplier near the period–3 orbit is already estimated in section 6.3.1, it can be verified that the first 2 eigen values of the Γ matrix corresponds to the actual floquet coefficients of the system. But the rest of the eigen values of the Γ matrix are spurious values generated due to the over embedding of the system in a higher dimensional space. This verifies the fact [26] that the reconstructed manifold can produce spurious inferences about the stability of the dynamics when the embedding dimension is higher than required.

6.4.2 Recurrent Neighborhood in lower dimensional space R^d

This section verifies that the estimation of floquet coefficients can be done correctly only if the dimension of projected space is equal to the actual dimension of the system that generated the data. The numerical results for projections from R^{800} to various lower dimensional spaces R^2, R^3, R^4 are analyzed here. Estimation of floquet coefficients was done correctly when R^2 which matches the exact local dimension of the Duffing system.

The A matrix given below was used for projecting data from R^N to R^d. This section contains lists the estimated values of the floquet exponents in various lower

dimensions i.e. for d values $4, 3, 2$.

$$A_{0,n} = Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \tag{6.26}$$

$$A_{1,n} = Imag\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_0}\right)\right) \tag{6.27}$$

$$A_{2,n} = Real\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_1}\right)\right) \tag{6.28}$$

$$A_{3,n} = Imag\left(\frac{1}{N}\exp\left(\frac{-i2\pi n}{h_1}\right)\right) \tag{6.29}$$

For R^N to R^d projection, A matrix has a dimension $(d \times N)$. The constant h_0 was selected to match the embedding dimension and h_1 was set $2h_0$. Hence the parameter values chosen for projection were: $h_0 = 800, h_1 = 1600, N = 800$ and $n = 1...N-1$.

Let SX, SY represents the counterparts of RX, RY, the start and end matrices of the equivalence class corresponding to period 3 interval in R^d. Follow the steps explained below to find the transformation (specific to the period 3 interval in R^d) across SX and SY .

Step 1: Find the centroid of the neighborhoods SX, SY and subtract it from the neighbors to translate the neighborhood to origin. Let sX and sY represents the neighborhoods translated to origin. Record the centroids of SX and SY as $\overline{SX}$ and $\overline{SY}$ respectively. SX and SY have dimensions of $39 \times d$ each.

Step 2: Define a transformation Λ, across the matrices sX and sY such that;

$$sX.\Lambda = sY \tag{6.30}$$

Step 3: Estimate Λ of dimension $d \times d$ by pre–multiplying sY with the generalized inverse of sX.

$$\Lambda = (sX)^{-1}.sY \tag{6.31}$$

Transformation Λ specific to the period 3 equivalence class was found for R^4, R^3, R^2 projections. Floquet estimates for these 3 cases are discussed below.

Floquet exponents in R^4

Eigen values of the Λ matrix in R^4 was,

$$\begin{bmatrix} -0.089032 + 0.37939i \\ -0.089032 - 0.37939i \\ -0.0006535 \\ 0.0621768 \end{bmatrix}$$

And the corresponding estimates of the floquet exponents were,

$$\begin{bmatrix} -0.022479 + 0.04297i \\ -0.022479 - 0.042967i \\ -0.17492 + 0.07494i \\ -0.06626 \end{bmatrix}$$

Since we have already found the floquet multiplier near the period–3 orbit in section 6.3.1, we can see that the first 2 estimates correspond to the actual floquet coefficients of the system. But the last two are spurious values generated due to representation of the dynamics in a higher dimension.

Floquet exponents in R^3

Eigen values of the Λ matrix in R^3 was,

$$\begin{bmatrix} -0.088609 + 0.37927i \\ -0.088609 - 0.37927i \\ -0.01586 \end{bmatrix}$$

And the corresponding estimates of the floquet exponents were,

$$\begin{bmatrix} -0.022492 + 0.042945i \\ -0.022492 - 0.042945i \\ -0.01586 \end{bmatrix}$$

Again, we know that the first 2 of this array corresponds to the correct floquet exponents. But the last one is a spurious mode that gets generated.

Floquet exponents in R^2

When the projected dimension has a dimension identical to the actual dynamics, the eigen values of the map will correspond to the exact floquet modes. Eigen values of the Λ matrix in R^2 was,

$$\begin{bmatrix} -0.08902 + 0.37935i \\ -0.08902 - 0.37935i \end{bmatrix}$$

And the corresponding estimates of the floquet exponents were,

$$\begin{bmatrix} -0.022492 + 0.042945i \\ -0.022492 - 0.042945i \end{bmatrix}$$

In this case the sum of the floquet exponents adds up to the trace $(-c)$ of the system as expected. This simulation demonstrates that the stability estimates can be done empirically if the dimension that represents the dynamics is equal to the inherent dimension that generated the data.

6.5 Stability Analysis of ECG data

This section analyzes the stability of the affine maps specific to the ECG data (developed in chapter 5), which corresponds to the equivalence classes in the lower dimension space in R^3. These affine maps are empirically derived and are functions of only the recurrence timings. It was also shown that, given the set of maps and the recurrence timings, one could start at any random initial guess and then converge to the actual data. This property of convergence is a common case of asymptotically stable maps. Using this property of the affine maps a justification for an experimental observation about the heart rate variability is given in this section. It has been found that a reduced heart rate variability is an indication of unhealthy functioning of heart [75]. Using the maps derived from an empirical ECG data we show that uniform recurrence intervals can indeed lead to instability, whereas non-uniform intervals might promote asymptotic stability.

The maps corresponding to the equivalence classes as various possibilities for

heart dynamics. A variability in the heart rate implies a switching between these various possible dynamics. Hence one can propose a hypothesis that the switching process is determined by the impulses that reach the SA node and this switching process selects any one possibility of different possible combination of these affine maps. Here stability implies a situation where the individual heart cells in the ventricles are synchronized with each other thus inducing the rapid contraction of ventricles which is essential for discharging blood out of the heart. Instability could be seen as situation where the synchrony of the individual heart cells are lost thus resulting in a quivering of ventricles rather than a proper contraction.

6.6 Stability of composition of maps

For linear transformations, the gist of the stability result is quite simple. Consider 2 two linear transformations represented by matrices A and B with following properties: (i) both have their determinant < 1 indicating that they are area contraction maps and (ii) at least one of the eigenvalue has a magnitude is > 1 indicating that the maps are individually unstable. This will result in a possibility were a monotonous repetition of the transformations (corresponding to the sequence $AAA\ldots$ or $BBB\ldots$) is unstable but a switching of them ($ABAB\ldots$) might lead to stability.

Let A and B be,

$$A = \begin{bmatrix} 1.5 & 0 \\ 0 & 0.3 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.2 & 0 \\ 0 & 2 \end{bmatrix}$$

The determinants $\left|A\right| = 0.45$ and $\left|B\right|$=0.4 are < 1. But both A,B have one eigenvalue whose magnitude is > 1. Also note that both **A and B are commutative** as,

$$AB = BA = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.6 \end{bmatrix}$$

A monotonous repetition of the transformations $AAA\ldots$ or $BBB\ldots$ is unstable as on of the eigen values explodes. A finite composition of them are given below,

$$A^{20} = \begin{bmatrix} 3325.2567 & 0 \\ 0 & 10^{-10} \end{bmatrix}$$

$$B^{20} = \begin{bmatrix} 10^{-8} & 0 \\ 0 & 1048576 \end{bmatrix}$$

But a switching of them ($ABAB\ldots$) is stable as both the eigen values < 1. For a finite composition,

$$(AB)^{10} = \begin{bmatrix} 0.000006 & 0 \\ 0 & 0.006 \end{bmatrix}$$

Permutations of A, B in this combination of transformation does not matter as A and B both commutative.

Now lets consider two **non–commutative transformations C and D** such that $CD \neq DC$ with the same properties as above: Determinant < 1 and at least one eigenvalue whose magnitude is > 1.

$$C = \begin{bmatrix} 1.1 & 0.6 \\ 0.3 & 0.4 \end{bmatrix}$$

$$D = \begin{bmatrix} 0.5 & 0.3 \\ 0.1 & 1 \end{bmatrix}$$

The determinants $\left|C\right| = 0.26$ and $\left|B\right|$=0.47 are < 1.

The eigen values of $C = \begin{bmatrix} 1.3 \\ 0.2 \end{bmatrix}$ and the eigen values of $D = \begin{bmatrix} 0.446 \\ 1.05 \end{bmatrix}$

Since C and D are non–commutative,

$$\begin{bmatrix} 0.61 & 0.93 \\ 0.19 & 0.49 \end{bmatrix} = CD \neq DC = \begin{bmatrix} 0.64 & 0.42 \\ 0.41 & 0.46 \end{bmatrix}$$

A monotonous repetition of the transformations $CCC\ldots$ or $DDD\ldots$ is unstable. For a finite composition,

$$C^{20} = \begin{bmatrix} 155.5 & 103.66 \\ 51.83 & 34.55 \end{bmatrix}$$

C^{20} has eigen values $\begin{bmatrix} 190.05 \\ 10^{-9} \end{bmatrix}$

$$D^{20} = \begin{bmatrix} 0.26 & 1.42 \\ 0.47 & 2.61 \end{bmatrix}$$

and D^{20} has eigen values $\begin{bmatrix} 10^{-13} \\ 2.87 \end{bmatrix}$

Now we have considered switching of them ($CDCD\ldots$). For a finite composition,

$$(CD)^{10} = \begin{bmatrix} 0.44 & 0.85 \\ 0.17 & 0.33 \end{bmatrix}$$

This particular permutation $(CD)^{10}$ was stable as all the eigen values $\begin{bmatrix} 0.77 \\ 10^{-13} \end{bmatrix}$ had a magnitude less than < 1. Another permutation $(DC)^{10}$ of this combination too was stable as all the eigen values $\begin{bmatrix} 0.89 \\ 10^{-12} \end{bmatrix}$ had a magnitude less than < 1. But we found that some other permutations of the same combination of maps $[C^{10}D^{10}]$, $[D^{10}C^{10}]$, $[(DDCC)^5]$ were unstable as they had at least one eigen value > 1. The eigen values of these finite compositions were $\begin{bmatrix} 10^{-11} \\ 10.8 \end{bmatrix}$, $\begin{bmatrix} 10^{-10} \\ 10.8 \end{bmatrix}$, $\begin{bmatrix} 1.36 \\ 10^{-8} \end{bmatrix}$ respectively.

We have seen some numerical examples of matrices where combinations of transformations were stable even though the individual ones that made the combination were unstable. In the case of non-commutative maps we saw a possibility of some permutations of a combination of maps being stable while others not. Now we

will explore the theoretical details about the stability of nonlinear maps and their affine equivalents and will introduce few theorems on the stability of the composition of maps.

6.6.1 Asymptotic Stability of maps

A fixed point $\bar{x}$ of a map $x \mapsto F(x)$ is said to be Lyapunov stable if given any $\epsilon > 0$, there is a $\delta > 0$ such that $||F^n(x) - \bar{x}|| < \epsilon$ for all integers $n > 0$ for all $||x - \bar{x}|| < \delta$. And if for all x, $F^n(x) \to \bar{x}$ as $n \to \infty$ then $\bar{x}$ is said to be asymptotically stable.

Theorem 6.1. *If F is a nonlinear real analytic map, then F has an affine representation f.*

Proof : If F is analytic, it has a valid taylors series at any point a. For any small neighborhood of a,

$$F(a + x) = F(a) + J(a)x + o(x^2) \tag{6.32}$$

Where $o(x^2)$ represents all the higher order components of the Taylor series. $J(a)$ is the matrix of partial derivatives of F (i.e. Jacobian matrix) at point a. Then the affine approximation of F at a can be written as,

$$f_a(x) = F(a) + J(a)x \tag{6.33}$$

$f_a(x)$ can represent $F(x)$ in a neighborhood of a, if for every $\epsilon > 0$, there exits a $\delta > 0$ such that for all x, with $|a - x| < \delta$, $|F(a + x) - f_a(x)| < \epsilon$.

Comparing Eq. 6.33 with standard from of affine transformations $f(x) = Ax + c$, $A = J(a)$ is of dimension $(d \times d)$ and $c = F(a)$ is of dimension $(d \times 1)$ if map f is defined in R^d.

Theorem 6.2. *If $\bar{x}$ is the rest point of the affine map $f(x) = Ax + c$, then $f(x)$ can be also written as a linear map with respect to its rest point.*

Proof : $\bar{x}$ is the rest point of $f(x)$, hence $\bar{x} = A\bar{x} + c$ and $\bar{x} = (I - A)^{-1}c$. Define a new co-ordinate $y = x - \bar{x}$. Substituting for x we get, $y + \bar{x} = A(y + \bar{x}) + c$.

Thus $y = Ay$ which is a linear map.

Theorem 6.3. *Asymptotic stability of the map f depends on the eigen values of A. If all eigen values of $A < 1$; then the map f is asymptotically stable.*

Proof : An affine map of form $f(x) = Ax + c$ can be written as a linear map of form $y = Ay$ (Theorem 1). If the eigen values of the Jacobian matrix A has a magnitude < 1, then the perturbations eventually dies out. Let the eigen vectors $\{e_i\}$ of matrix A form a basis for the initial perturbation vector say δ_0 such that, $\delta_0 = \sum_{i=0}^{N-1} \alpha_i e_i$ for some scalars α_i. Its evolution can be represented as $\delta_1 = A\delta_0 = A \sum_{i=0}^{N-1} \alpha_i e_i = \sum_{i=0}^{N-1} \lambda_i \alpha_i e_i$. Iterating k times, $\delta_k = \sum_{i=0}^{N-1} {\lambda_i}^k \alpha_i e_i$. Hence if all $|\lambda_i| < 1$, then $||\delta_k|| \to 0$ causing the perturbations along all the eigen directions die out. Otherwise if $|\lambda_i| > 1$, then the perturbation along that e_i will grow [2].

Theorem 6.4. *Composition of affine maps are also affine.*

Proof : Consider two affine maps of the form $f_1(x) = A_1 x + c_1$ and $f_2(x) = A_2 x + c_2$. Then,

$$\begin{aligned} f_1(f_2(x)) &= A_1(A_2 x + c_2) + c_1 \\ &= A_1 A_2 x + A_1 c_2 + c_1 \\ &= Bx + e \end{aligned}$$

where $e = Ad + c$. Note that the Jacobian of the composition B is equal to the product of the individual Jacobian matrices A_1 and A_2.

Theorem 6.5. *An affine composition $f_1 f_2$ is asymptotically stable if the product of the individual Jacobian matrices $A_1 A_2$ has the magnitude all eigen values < 1.*

Proof : Same as proofs of Theorem 3 and 4. Further, the Jacobian matrix of the composition $f_1 f_2$ is the product of the individual Jacobian matrices A_1 and A_2. If the product has the magnitude of its eigenvalues < 1 then $f_1 f_2$ is asymptotically stable.

Theorem 6.6. *For the infinite composition of affine maps (of the form $fff\ldots$), where $f(x) = Ax + c$, the stability depends on the eigen values of the A. If the magnitude of all eigen values of A are < 1, then infinite composition is asymptotically stable.*

Proof : same as proof of Theorem 5.

Theorem 6.7. *For affine maps $f_1(x) = A_1x + c_1$ and $f_2(x) = A_2x + c_2$, If A_1 and A_2 commutes and if a particular composition $f_1f_2f_1f_2\ldots$ is unstable then all the permutations of that composition are unstable.*

Proof : If a particular composition $f_1f_2f_1f_2\ldots$ is unstable then the corresponding Jacobian $A_1A_2A_1A_2\ldots$ will have at least one eigen value with a magnitude > 1. Since A_1 and A_2 commutes, the order in which f_1 and f_2 they appear does not matter. Thus all permutations of this composition $f_1f_2f_1f_2\ldots$ are unstable.

Theorem 6.8. *If f_1, f_2 are stable; then the infinite composition $f_1f_2f_1f_2\ldots$ will eventually oscillate between exactly two points. The first point is the fixed point of $f_1f_2(x) = x$ the second is the fixed point of $f_2f_1(x) = x$.*

Proof : Assume that the compositions fg and gf have fixed points $\bar{x}$ and $\bar{y}$ respectively. Then $fg(\bar{x}) = \bar{x}, gf(\bar{y}) = \bar{y}$. Construct an infinite set S of finite compositions that constitute the above infinite sequence. Let the set be, $S = \{f, fg, fgf, fgfg, fgfgf, \ldots\}$. Note that each of the finite composition will settle to either $\bar{x}$ or $\bar{y}$. This does not guaranty convergence. But if f, g are stable infinite composition can be stable in the sense that it switches between the individual rest points $\bar{x}$ and $\bar{y}$. Hence $\bar{x}$ and $\bar{y}$ form a limit set for the infinite sequence of maps $fgfg\ldots$.

Theorem 6.9. *Any infinite composition of maps $f_1f_2f_3f_4\ldots$ are stable if (i) the magnitude of the eigen values of the Jacobian matrices corresponding to the finite compositions of the infinite composition of maps, form a Cauchy sequence and if (ii) limit point of the above Cauchy sequence is < 1. (i) and (ii) together is the sufficient condition for stability of the infinite composition.*

Proof : Construct an infinite set S of finite compositions of maps that constitute the infinite composition as, $S = \{f_1, f_1f_2, f_1f_2f_3, \ldots\}$. Define an infinite sequence of numbers that $d_1, d_2, d_3, \ldots$ that correspond to the magnitude of eigen values of the Jacobian matrices of individual composition of maps in the set S. If the limit point of the above sequence $d < 1$, then the infinite sequence is asymptotically stable.

6.7 Stability of Affine maps specific to the equivalence classes of the ECG data

In this section we will see the properties of the affine maps specific to the equivalence classes which were empirically determined from a healthy ECG data. Determinants and Eigenvalues of the Jacobians of the affine maps had following features.

1. Most of the Jacobians of the affine maps had a determinant < 1 (as shown in Fig. 6.2 whose X–axis shows the RR interval corresponding to the particular equivalence class whose evolution is specified by the affine map and Y–axis shows the determinant of the Jacobian of that affine map).

2. Most of them have at least one eigenvalue with magnitude > 1. (as shown in Fig. 6.3 whose X–axis shows the RR interval corresponding to the particular equivalence class whose evolution is specified by the affine map and Y–axis shows the first two eigen values of the Jacobian of that affine map. For most of them the first eigen value is > 1 and the second one is < 1).

3. HRV implies a composition of various affine maps. The corresponding Jacobian matrix of the composition map had a determinant < 1 and its eigenvalues had magnitudes < 1. Let us consider a composition of affine maps corresponding to a short ECG data segment that consist of 10 cycles with the actual RR pattern of $\{712, 702, 738, 706, 696, 787, 737, 692, 751, 736\}$. Most of these affine maps had at least one eigenvalue with a magnitude > 1. But the Jacobian

of the composition had eigen values $\begin{bmatrix} 0.83966 \\ -0.00027 \\ 10^{-11} \end{bmatrix}$. Hence the composition was stable even though the individual maps were unstable.

4. If there is no RR–variability, it implied a monotonous repetition of one of the affine maps. Consider a possibility of being stuck at a uniform heart rate of 674 repeatedly. Jacobian of the affine map specific to the equivalence class corresponding to an RR interval 674 had the following property: Determinant of the Jacobian was -0.55567. Eigenvalues were $\begin{bmatrix} -3.03673 \\ 0.91204 \\ 0.20063 \end{bmatrix}$. Now the monotonous repetition of the affine map corresponding to this equivalence class will result in exploding the eigen value whose magnitude is > 1. Eigen values of the Jacobian of the affine map corresponding to a repetition of 10 cycles of this map were $\begin{bmatrix} 66689.75999 \\ 0.39825 \\ 10^{-13} \end{bmatrix}$. Hence the monotonous repetition was unstable as the individual map itself was unstable.

Using Theorem 9 discussed in section 6.8, we can formulate a stability condition for the affine maps corresponding to various RR intervals of ECG data as follows. The Jacobian of most of these affine maps had a (i) determinant < 1 indicating that they are contraction maps and (ii) at least one eigenvalue whose magnitude > 1 indicating that the maps are individually unstable. The composition of these maps can be stable if they satisfy the conditions:

1. magnitude of the eigen values of the Jacobian matrices corresponding to the finite compositions of the affine maps form a Cauchy sequence

2. limit point of the above Cauchy sequence is < 1.

Condition (1) and (2) are sufficient for the stability of the composition.

Proof: same as the proof of the Theorem 6.9 discussed in section 6.6.

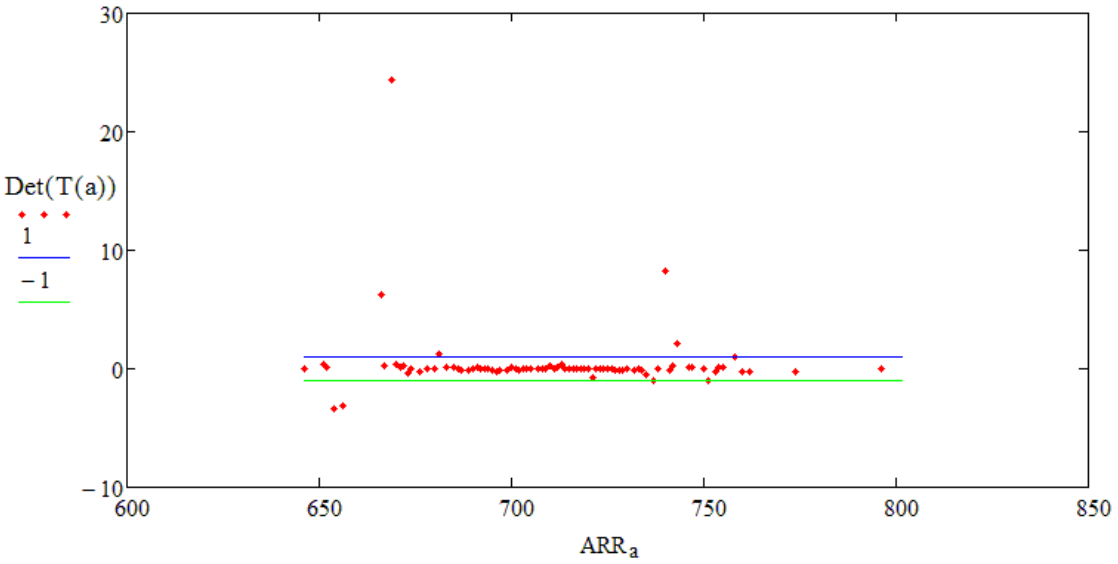

Figure 6.2: Determinants of the Jacobians of the Affine maps corresponding the equivalence classes. X–axis shows the RR interval corresponding to the particular equivalence class whose evolution is specified by the affine map and Y–axis shows the determinant of the Jacobian of that affine map.

6.7.1 Importance of being Non-Commutative

It does appear that the property of non-commutativity is essential for stability as it indicates a proper functioning of the heart. The numerical simulations confirmed an already known fact that a uniform heart rate results in an unstable condition. In retrospect this has been explained using the non-commutativity of the set of affine transformations which are empirically found. Thus we may conclude that the heart cells would not synchronize with uniform heart rate but could well do so in the case of a non–uniform heart rate. Fig. 6.4 shows the convergence of 2 farther initial conditions to an ECG data segment, for a combination of maps that corresponded to actual RR intervals (which showed a heart rate variability). Fig. 6.5 shows the explosion and divergence of two ECG trajectories starting on 2 nearby initial conditions, using one map (corresponding to the RR interval 674) which was used repeatedly (as in a uniform heart rate condition).

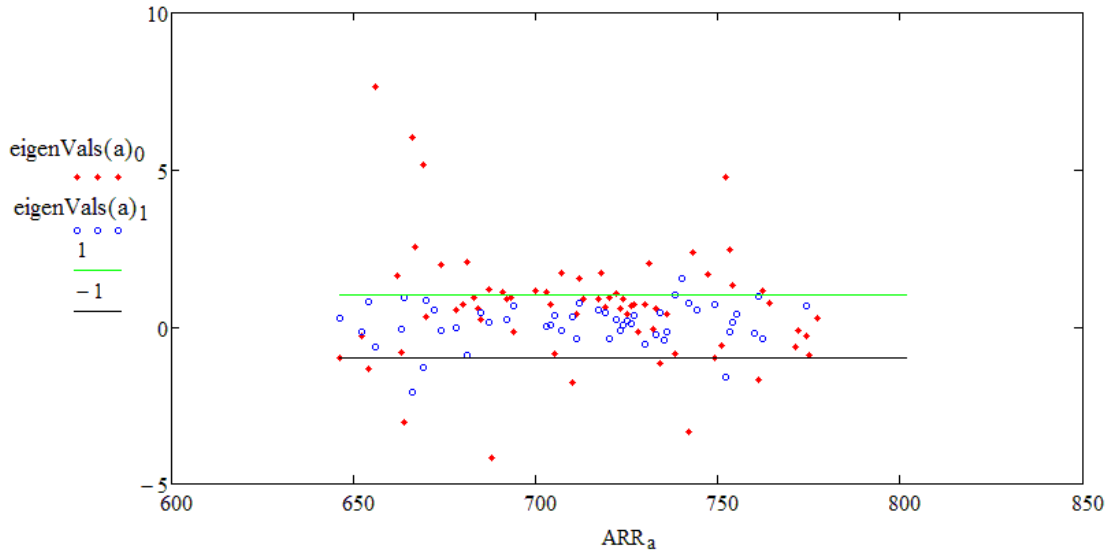

Figure 6.3: Eigen values of the Jacobians of the Affine maps corresponding the equivalence classes. X–axis shows the RR interval corresponding to the particular equivalence class whose evolution is specified by the affine map and Y–axis shows the first two eigen values of the Jacobian of that affine map. For most of them the first eigen value is > 1 and the second one is < 1.

6.7.2 Reduced Heart Rate Variability

A reduced heart rate variability which is very common in patients with heart conditions can be seen as a switching of the affine maps with a restricted symbolic sequence. A non–restricted symbolic sequence will imply that the system has the ability to select from a set of all possible permutations and combinations of the maps specific to various recurrence intervals. A restricted symbolic sequence denotes a condition where the topological entropy of the sequences are highly reduced, there by some permutations of certain combinations are not allowed. Switching of maps with a restricted symbolic sequence in this context may be similar to something as follows: If f_1, f_2 are the non–commutative maps, a restriction can be applied as an example: f_1 can be followed by f_2 at most say 3 times. There is also a possibility that some permutations of the composition of maps have unstable rest points (If one of them is unstable all of them are unstable. These maps may form their own

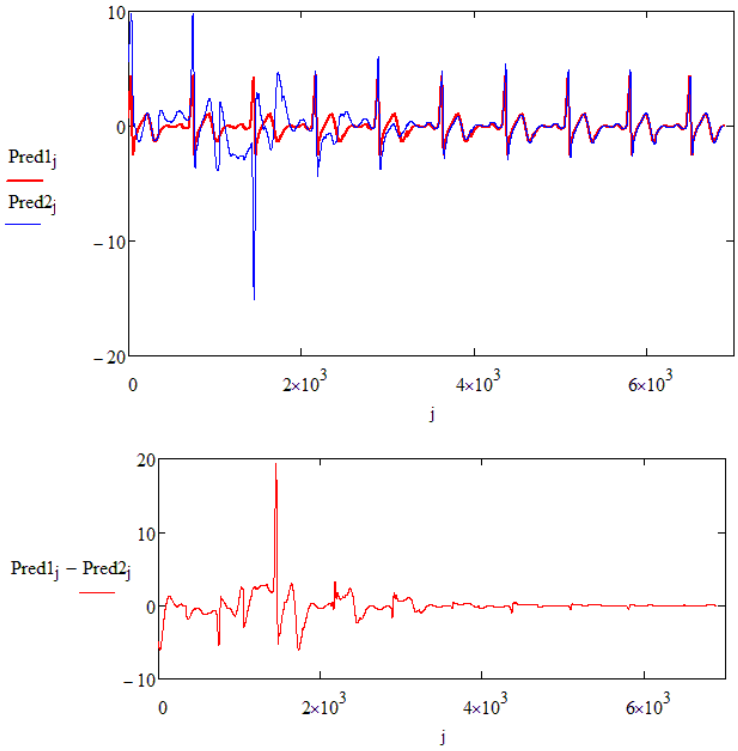

Figure 6.4: Convergence when starting on 2 farther initial conditions, using a combination of maps (non-uniform heart rate). Error of 2 trajectories that start on 2 nearby initial conditions, evolving using one map over and over again

omega limit sets of going from infinity to infinity). The switching sequences will create a dynamics like a cantor set. This sequence will have uncountable infinity of omega limit points which are in a compact attractor.

6.8 Conclusion

This chapter analyzed the properties of the maps specific to the equivalence classes for both the duffing oscillator generated data and the ECG data. The empirically derived affine maps across the equivalence classes were functions of only the recurrence timings. It is further verified that representing the data in a dimension higher than the original dimension of the system that generated the data can cause wrong stability conclusions. It is demonstrated for the duffing oscillator data that

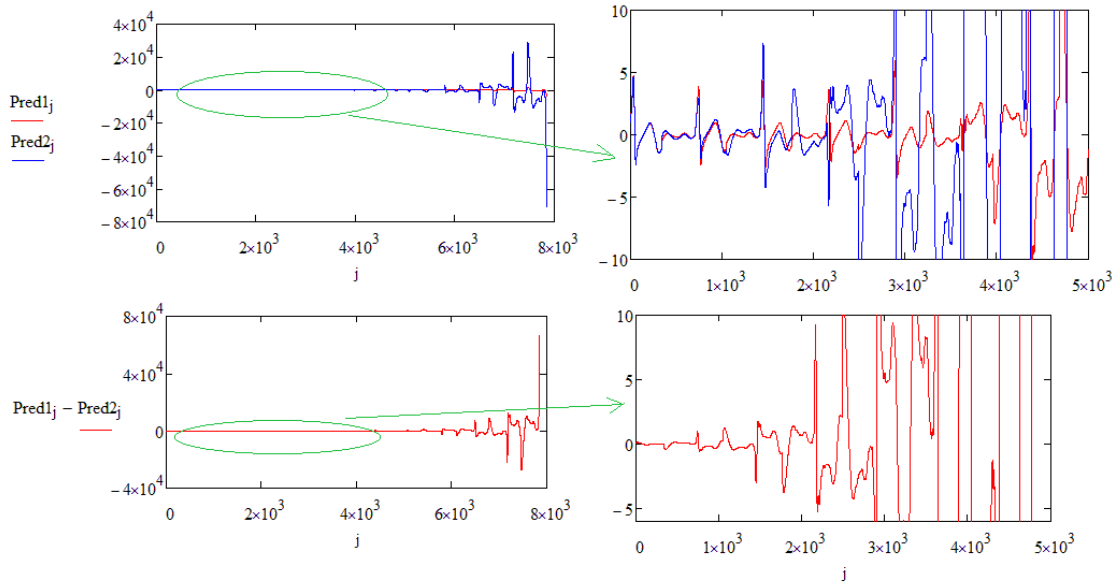

Figure 6.5: Divergence when starting on 2 nearby initial conditions, using one map over and over again (uniform heart rate). Error in 2 trajectories that convergence when starts on 2 farther initial conditions and evolving using a combination of maps.

for empirically determining the exact values of the floquet exponents, the neighborhood has to be represented in the proper local dimension of original system i.e. $d = 2$. A theoretical foundation of an experimental observation about the ECG data is discussed in this chapter. Heart rate variability is known to be a must for the healthy functioning of heart. Based on the properties of the affine maps, the chapter demonstrated that heart rate variability implies a stable dynamics where as a uniform heartbeat may lead to an instable dynamics.

CHAPTER 7
Predicting Multichannel data using Recurrent Neighborhoods

A contribution of the thesis to the existing body of knowledge in the field of biomedical signal processing is discussed in this chapter. This chapter extends of the methods of prediction explained in chapters 4 and 5 for processing multichannel data. An online data set of 300 patients published by Physionet as part of the challenge 2010 was used for data analysis. The challenge was to find a short stretch of missing data in a channel using the information available in other channels. The method of prediction based on recurrent neighborhoods gives a very accurate prediction for certain type of signals that have identical recurrent timings (as in the case of ECG and Blood pressure signals); but the method gave poor results when the recurrent timings of the signals were not identical (as in the case of ECG and Respiration signals. They had non–identical recurrence timings but interestingly they showed a phase locked behaviour, exploitation of which could be a future direction of research). Various local models based on the recurrent neighborhood were found for signals in the multichannel data for the purpose of prediction. The ideas were optimized for the data set A, for which missing data was separately provided and later for set B, for which online feedback was available. This chapter contains the results for all the data in set C for which the total score was ranked 6 among the participants of the challenge 2010 by Physionet.

7.1 Introduction

Real time monitoring of physiological signals are important in clinical research and diagnosis of diseases. It is reported that signal corruption and signal loss are very common in the hospital settings [37]. A trained personal may be able to deal with some of the signal distortions as loss of signals or random noises due to the cognitive abilities of human brain. It is a challenging task to develop an algorithm that can mimic some of these cognitive tasks to identify the signal corruption based on

available contextual information. Intelligent algorithms that can simulate cognitive skills to fill the gaps based on the context thus taking care of the signal corruption for the purpose of diagnosis and forecasting is a crucial in the area of biomedical research [37]. This chapter explores some algorithmic solutions that can be used for noise removal and signal reconstruction. The methods used here are extensions of the modeling method explained in chapters 4 and 5 to deal with the said issues.

An online data set of 300 ICU patients published by Physionet as part of the challenge 2010, named 'Mind the Gap' was used for the study. The data set had multichannel physiological ICU recordings of around 12–15 leads each that included various electro-cardiogram (ECG) leads, blood pressure (BP) leads, respiration(RESP) and fingertip plethysmogram (PLETH) outputs. The challenge was to find a short stretch of missing data in a channel using the information available in other channels. Though the recordings contains a measurement of various physiological functions, one can see that some of these signals, for example various ECG channels and BP channels could carry a lot of mutual information as they basically represent activities of the cardiovascular system. Similarly there a possibility of some relationship between different channels that are simultaneously recorded. A model that can exploit the redundancy and mutual relationships across channels to predict or reconstruct the lost or corrupt information is highly desirable under these circumstances.

Different methods for analysis and prediction were used by the groups that participated in the challenge to address these issues mentioned above [44]. Neural network based methods are known to be the best among successful methods that solve similar problems as many of them can give reasonable predictions with intense training sessions [37, 46]. Though it was once difficult to use neural networks for modeling complex signals such as physiological data, the development of novel training strategies gave hope for solving complex problems [84, 85, 86]. The winner of the challenge 2010, Rodrigues has used an efficient multi–layered neural network method that uses an adaptation of Geoffrey Hinton training strategy for the prediction between channels for multi–channel physiological data [45]. Sullivan et.al, another participant who scored high in challenge have reported that focused time

delay networks are more robust, computationally efficient, and accurate when compared to distributed time–delayed networks and nonlinear autoregressive networks to deal with the prediction problem [87]. Radial basis function (RBF) networks, another class of artificial neural networks too could give a set of reasonable prediction scores [88]. But the main drawbacks of neural network based methods are the following: (i) they are computationally very extensive (ii) most of them demand a long training time (iii) prediction is limited to the patterns or behaviors that are available during the training time otherwise that behavior is not learned or completely lost during prediction.

Apart from neural networks, some algorithms that were based on kalman filters, adaptive filters and their combinations also gave good quality predictions for the missing channels [38, 39]. Hartmann has used a robust method based on adaptive filters that identifies linear-static interconnections between channels as composite Infinite Impulse Response (IIR) transfer functions [38]. Though these interconnections could be time variant, they were assumed to be static over short time windows. A genetic algorithm that was used for the finding the filter coefficients and fitness function by minimizing the mean square error and maximizing the correlation guaranteed good prediction results. A collection of methods like linear regression, pattern matching, ECG derived Respiration techniques were also used for the prediction of different types of signals [40]. Some alternative approaches used algorithms based on average substitution, principal component analysis and wavelet decompositions for prediction [41, 42, 43].

The methods explored in this chapter exploits conjugacy that exists between different channels in a simultaneous recoding. Dynamic models were developed for various signals and these models were successfully used for the prediction across the channels. For some of the signals, when the recurrent timings were identical (as in the case of ECG and BP) the prediction results were excellent as shown in the data analysis section of this chapter.

The chapter is organized as follows. Section 7.2 lists the key objectives of the chapter. The methodology followed, the data set understudy and details of the proposed model are explained in section 7.3. Section 7.4 contains the results

of data Analysis. This section demonstrate the prediction across various channels with figures. The prediction scores for the entire data set C is given in section 7.5. Section 7.6 discusses the analysis results and the chapter concluds in section 7.7.

7.2 Objectives

- Extend the method of prediction based on recurrence neighborhoods for multichannel physiological data
- Demonstrate the application of the method of prediction and modeling for a multichannel physiological data set of 100 recordings and analysis the Prediction results

7.3 Methodology– Extension of the Proposed Model for analyzing Multichannel Data

The property of recurrence in signals was explored in detail in Chapter 5. An objective of this chapter to extend the method of modeling based on the recurrence property of signals to analyze a multichannel recording of physiological signals. These recordings contain simultaneous measurements of cardiovascular signals of a patient in hospital settings. The signals in these channels exhibit the property of recurrence as demonstrated in the following sections.

The model for prediction is based on an assumption that if two signals in the same multichannel recording exhibit the property of recurrence, there is a possibility of the existence of a topological conjugacy across the respective recurrent neighborhoods. This assumption is made as all the events that happen in these signals are related or the signals are synchronized in some generalized manner based on the facts : (i) all signals are simultaneous measurements of the same cardiovascular system and (ii) all share the property of recurrence– some of the signals (all ECG and BP channels) have identical recurrent timings with respect to each other; in the case of non–identical recurrent timings (as ECG, CVP, RESP), the signals are synchronized in a phase locked behaviour.

The concept of topological conjugacy property of dynamical systems was discussed in Chapter 4. If X and Y be two topological spaces with two continuous functions f and g such that $f : X \to X$ and $g : Y \to Y$; existence of a homeomorphism $h : X \to Y$ such that $f = h^{-1}gh$ implies that f and g are topologically conjugate to each other [4].

For the case of multichannel-channel physiological data, assume $x(t)$ and $y(t)$ represents two simultaneous channels that exhibit the property of recurrence. Let RX and RY be the recurrence neighborhoods reconstructed from the data $x(t)$ and $y(t)$ respectively. A topological conjugacy is said to be existing between the recurrence neighborhoods RX and RY, if there exists an invertible map that can go back and forth across the neighborhoods RX and RY. Section 7.3.4 discusses the existence of such a transformation across the neighborhoods. In general conjugacy implies that data in one channel can be transformed into data at another channel and vice versa. Note that the evolution equations that govern $x(t)$ and $y(t)$ could be stochastic, nonlinear, or any high–dimensional deterministic functions. The topological conjugacy that exists between the recurrent neighborhoods of $x(t)$ and $y(t)$ can be exploited for the purpose of prediction and modeling as demonstrated in this chapter.

7.3.1 Data set under Study

The online data resource of 300 sets of data provided by Physionet is used for the analysis and testing of our algorithms [51]. These are ICU patient monitor recordings that are simultaneously recorded multiple channel physiological data for each patient. This archive contains 3 data sets A, B, C each containing 100 recordings each that are 10 minute in length with a sampling frequency of 125 samples per second. The signals vary across the records and they include various ECG channels, continuous invasive blood pressure, intra-cranial pressure, central venous pressure, respiration and raw fingertip plethysmogram outputs. For a typical multichannel recording of length 10 minute, the last 30 seconds of one of the randomly chosen channel was replaced by a gap (a flat line signal) as shown in Fig. 7.2. The goal of the physionet challenge 2010 was to reconstruct the missing signal in each record us-

```
Source: record challenge/2010/set-c/c21
val has 8 rows (signals) and 75000 columns
(samples/signal)
Duration:    10:00
Sampling frequency: 125 Hz
Sampling interval: 0.008 sec

Row Signal   Gain       Base    Units
1   RESP     4093       2       pm
2   CVP      102.3667   307     mmHg
3   PLETH    1023       0       NU
4   II       4094       0       mV
5   III      2046       1024    mV
6   V        4094       0       mV
7   ABP      20.4733    307     mmHg
8   UAP      102.3667   307     mmHg

To convert from raw units to the physical units shown
above, subtract 'base' and divide by 'gain'.
```

Figure 7.1: info about c21m.dat.

ing the information available in other channels. These data set were made available online for the Physionet challenge 2010–Mind the Gap and is still available online to support for further study [51].

7.3.2 Identification of a Recurrent Neighborhood in a high dimensional Delay Embedding Space

Recurrent timings are important for determining the dimension of Delay embedding space. The conditions for the embedding dimension is same as that given in Section 5.2.1. of Chapter 5. These conditions guarantee that the columns of the recurrent neighborhood matrices do a proper overlap of the time domain signal. The study of a single neighborhood recurrence can yield all information about the dynamics for the right choice of embedding dimension N. Note that choosing a smaller embedding dimension than required will result in a non proper overlap of the time domain signal that give a partial prediction. In that case one needs to find multiple overlapping recurrent neighborhoods for the prediction of the entire signal (this will be discussed in detail in the following Section 7.6).

All figures used in this chapter to discusses the technical details of the modeling are generated using a representative data set 'c21.dat' from the set C [1]. The

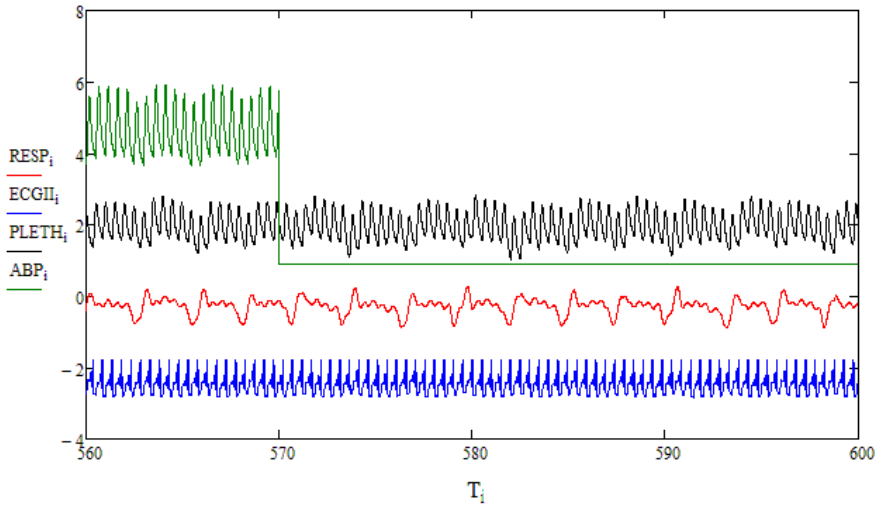

Figure 7.2: RESP, ECGII, PLETH, ABP signals of the 8–channel record *c21.dat*. Short data segments of the signals in the last 40 seconds are scaled and shifted to fit the figure. ABP signal missing in the last 30 seconds is denoted by a flat line.

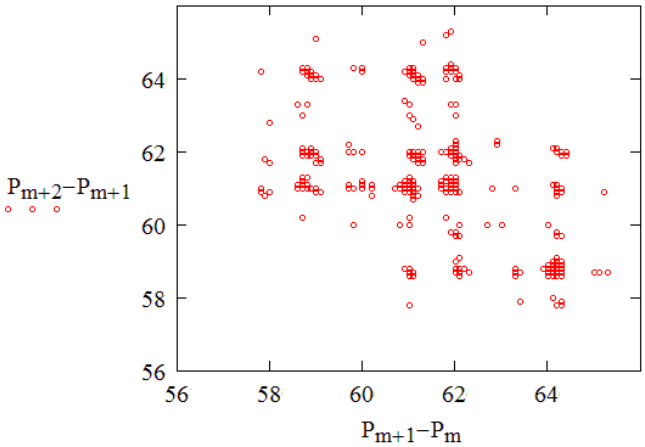

Figure 7.3: Heart Rate Variability (HRV) plot generated using ECGII signal of *c21.dat*

information file attached to this data record which contains the channel–signal details are shown in Fig. 7.1. All algorithms that were used for analysis and plotting figures for this chapters were developed in the technical software package MATHCAD version 14.0. Since the data set C understudy contains the vital signals of sick patients in ICU, they show a lesser heart rate variability (HRV) in general. One can make a comparison of the HRV of *c21.dat* shown in Fig. 7.3 to the HRV of a healthy person shown in Fig. 5.5 of Chapter 5.

The channel 4 – ECGII and channel 7– ABP of the *c21.dat* are used for the analysis to find the conjugacy across the recurrence neighborhoods of these data if it exists. Both the ECGII and the ABP signals have identical recurrent timings as shown by Fig. 7.4. An assumption is made that all the events that happen in signals ECGII and ABP are related or the signals are synchronized in some generalized manner based on the facts : (i) both are simultaneous measurements of the same cardiovascular system and (ii) both share identical recurrent timings. The data ECGII and ABP are embedded in a higher dimension R^N where $N = 111$ (values of N generally depend upon the range of recurrent timings and it varies across the data sets). Recurrence locations for the ECGII, ABP signals were identified using the Algorithm 5.1 mentioned in Section 5.2.1 of Chapter 5. Once the recurrent neighborhood matrices for both the signals ECGII and ABP are ready in R^N, the next step is to finding transformations across them which is explained in detail in the following section.

7.3.3 Modeling and Prediction across Recurrence Neigbhourhoods in R^N

Objective of the study is to predict a segment of the data that is missing in one of the channels in a simultaneous multichannel recording. Proposed method of prediction is as follows. Identify two separate channels in the multichannel data set: one of them is the channel where data is missing and the second one can be any other channel which has data available corresponding to time where the data is missing in the other channel. Modeling and prediction windows in the two selected channels of a multichannel recording are pictorially represented by Fig

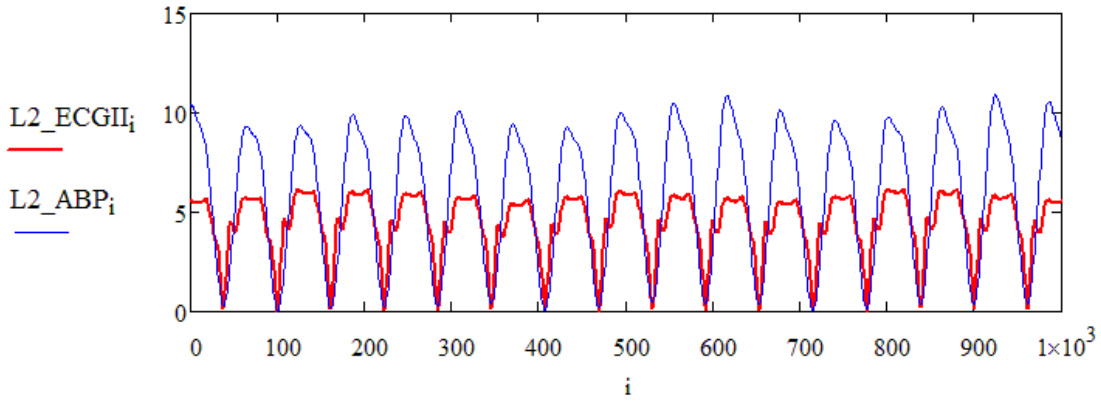

Figure 7.4: Identical recurrent timings for ECGII and ABP signals: L_2 norms of distances of the vectors with respect to a reference vector in R^{110} for ECGII and ABP data of ***c21.dat***

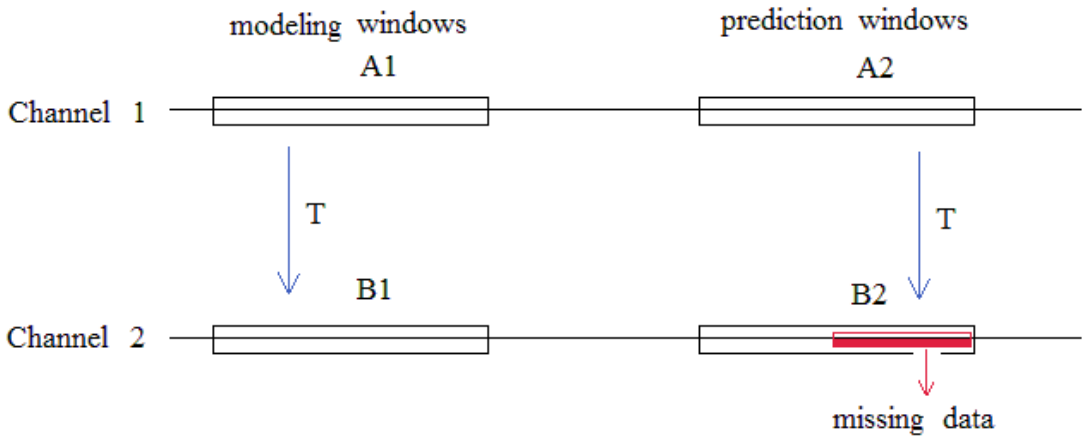

Figure 7.5: Modeling and prediction windows in 2 channels of a multi-channel recording. Channel 2 contains the missing data segment for which a prediction is made using data available in Channel 1. $A1, B1$ are the modeling windows across which a conjugacy map T is identified. T is used for prediction of data in the window $B1$ from window $A2$.

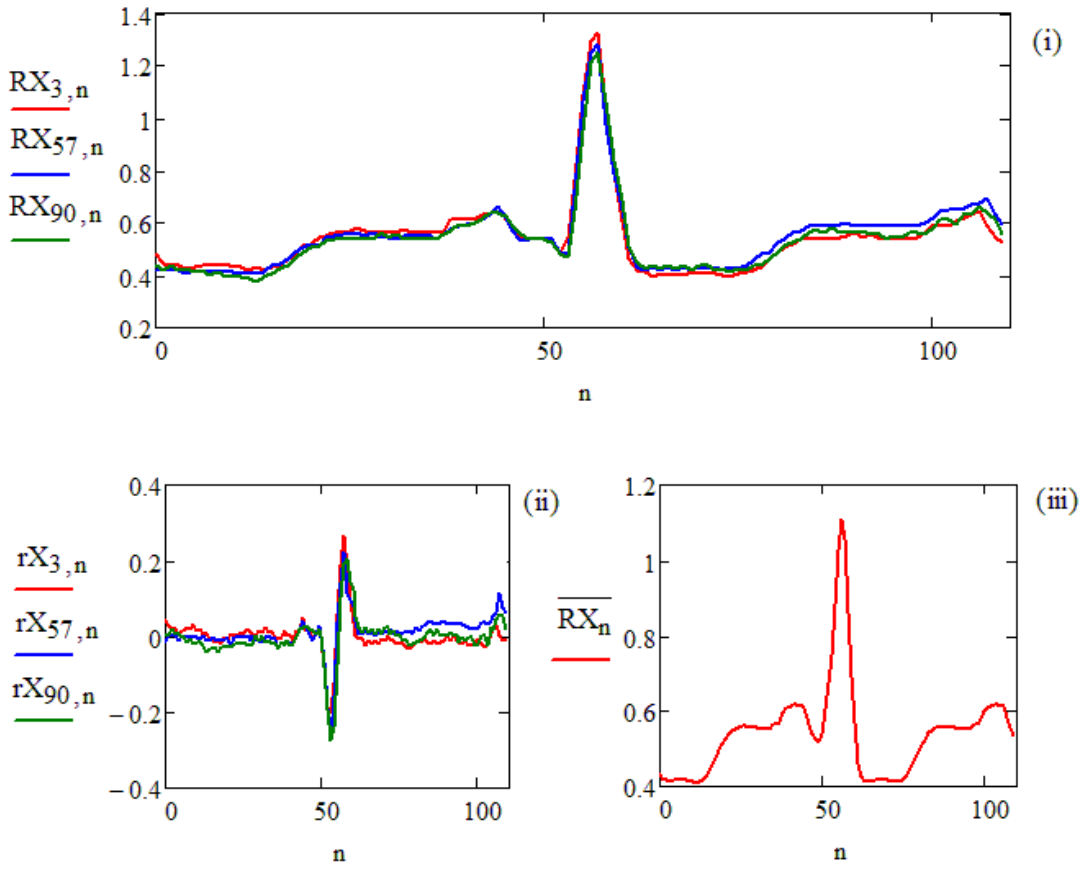

Figure 7.6: Few random neighborhood vectors in R^{110} before translation to origin (i) centroid of the neighborhood (iii) and the same vectors after translation to origin (ii) for the ECGIII data of *c21.dat*.

7.5. Channel 2 contains the missing data segment for which a prediction is made using data available in Channel 1. For finding a transformation that capture the relationship exists across the channels 1 and 2, recurrent neighborhoods are identified in the delay embedding space R^N for the data in the modeling windows $A1, B1$ in both the channels. A conjugacy map identified across the recurrent neighborhoods corresponding to the data in $A1, B1$ and this T is used for prediction of data in the window $B2$ from window $A2$. Modeling windows are chosen ahead but preferably closer to the prediction windows as an assumption is made that the parameters of the conjugacy map do not vary much across the windows.

7.3.4 Finding Transformations across the Recurrent Neighborhoods in R^N

Identify channels 1 and 2 for the prediction and modeling and record the recurrence timings for the signals in these channels using the Algorithm 5.1 given in Section 5.2.1. Once the modeling windows $A1, B1$ are identified; create recurrent neighborhood matrices denoted by RX, RY as explained in Section 5.2.1. This section explains the existence of an affine transformation across these matrices RX and RY. If the parameters of the transformation are known, RY can be predicted from RX.

Step 1: Find the centroid of each of the neighborhoods and subtract it from the neighborhood vectors for translating the neighborhoods to origin. Let rX and rY matrices represents the neighborhoods (RX and RY matrices) translated to origin. Record the centroid of the neighborhoods RX and RY as $\overline{RX}$ and $\overline{RY}$. If the neighborhood in R^N (represented by matrix RX, RY) contains P vectors each, the dimension of RX, RY are $(P \times N)$ and the centroids are:

$$\overline{RX}_n = \frac{1}{P}\sum_{p=1}^{P} RX_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{7.1}$$

$$\overline{RY}_n = \frac{1}{P}\sum_{p=1}^{P} RY_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{7.2}$$

Neighborhoods translated to origin are:

$$rX_{p,n} = RX_{p,n} - \overline{RX}_n \qquad \text{for all p,} \quad \text{for n} = 1, 2 \ldots \text{N} \tag{7.3}$$

$$rY_{p,n} = RY_{p,n} - \overline{RY}_n \qquad \text{for all p,} \quad \text{for n} = 1, 2 \ldots \text{N} \tag{7.4}$$

Fig. 7.6 depicts few neighborhood vectors before and after translation to origin along with the centroid for the case of ECGII data (channel 4 of the *c21.dat*).

Step 2: Define a transformation T of dimension $(N \times N)$, across the matrices rX and rY such that;

$$rX.T = rY \tag{7.5}$$

Step 3: Estimate T by pre-multiplying rY with the generalized inverse of

rX. (For the numerical simulation results explained in the following sections the generalized inverse of rX was calculated by using singular value decomposition by retaining approximately $N/2$ singular values).

$$T = (rX)^{-1}.rY \tag{7.6}$$

T represents the linear map across the neighborhoods rX, rY shifted to origin. This implies an affine map of the form ($\beta = T\alpha + \overline{RY}$) exists for every vector α in RX and β in RY. For the purpose of prediction an assumption is made that the parameters of the affine map does not change for the vectors in the neighborhoods in the prediction window $A2, B2$.

Let SX, SY represents the recurrent neighborhood matrices in the prediction windows $A2, B2$. SY is unknown as it represents the missing data in Channel 2. **Step 4:** Record the centroid of the neighborhoods SX as $\overline{SX}$. Since SY is unknown assign $\overline{SY} = \overline{RY}$. Let sX and sY matrices represents the neighborhoods (SX and SY matrices) translated to origin.

$$\overline{SX}_n = \frac{1}{P}\sum_{p=1}^{P} SX_{p,n} \qquad \text{for n} = 1, 2 \ldots \text{N} \tag{7.7}$$

$$\overline{SY} = \overline{RY} \tag{7.8}$$

Neighborhoods translated to origin are:

$$sX_{p,n} = SX_{p,n} - \overline{SX}_n \qquad \text{for all p,} \quad \text{for n} = 1, 2 \ldots \text{N} \tag{7.9}$$

Step 5: Define a linear map across the neighborhoods sX, sY. Find an estimate for the unknown variable sY using known variables T and sX.

$$sY = sX.T \tag{7.10}$$

Step 6: sY represents the neighborhood shifted to origin, to predict the neighborhood matrix SY corresponding to the window $B2$, use the affine transformation

of the form ($\beta = T\alpha + \overline{SY}$) for every vector α in SX and β in SY.

Step 7: SX matrix represent the vectors in a higher dimensional space R^N. Prediction for the window $B2$ can be reconstructed by stitching the embedding vectors appropriately at their corresponding recurrence neighborhood locations (recorded in the array of recurrence timings).

7.4 Data Analysis and Results

This section contains the simulation results of the modeling and prediction algorithm for a particular data set 'c21m.dat'. This data set has 8–channels and with various physiological signals (RESP, various ECG channels ECG II, III, V, PLETH, ABP, CVP). ABP is the missing channel which has a flat line for the last 30 seconds in the data set *c21.dat*. Various pairs of these channels are used to implement and demonstrate the algorithms in this section. For each pair of the signals, the signal available in the modeling window $B2$ (here onwards referred as the target signal) was masked and the steps mentioned in previous section 7.2.4 is used for prediction. The target signal and the resultant prediction were compared with a set of scores. Two scoring functions explained in Section ($Q1$ based on mean squared error and Q2 based on correlation)are used for comparing the prediction with the target signal. Prediction results for various pairs of signals in the data set *c21.dat* is demonstrated with figures (Fig. 7.7– Fig.7.18) in this section. The scores for missing signals in the entire data set (100 multichannel recordings in set C [1]) are given in tables (Table 7.1, 7.2 and 7.3). The scores were calculated for 3750 samples in the last 30 seconds of the 10 minute data set (sampling frequency is 125 Hz).

1. Prediction of ABP signal from ECGIII signal: ABP on 7–channel is predicted using the ECGIII on 5–channel. Fig 7.7 shows a short segment of data of these two channels. Both the ECG and ABP waves are repeated with similar time periods. (They share identical recurrent timings as shown in Fig. 7.4). Fig. 7.8 shows the predicted ABP signal and the original along with the error in prediction. Scores of Prediction are ($Q1 = 0.9757, Q2 = 0.9932$).

2. Prediction of ECGV signal from ECGIII signal: Prediction is done

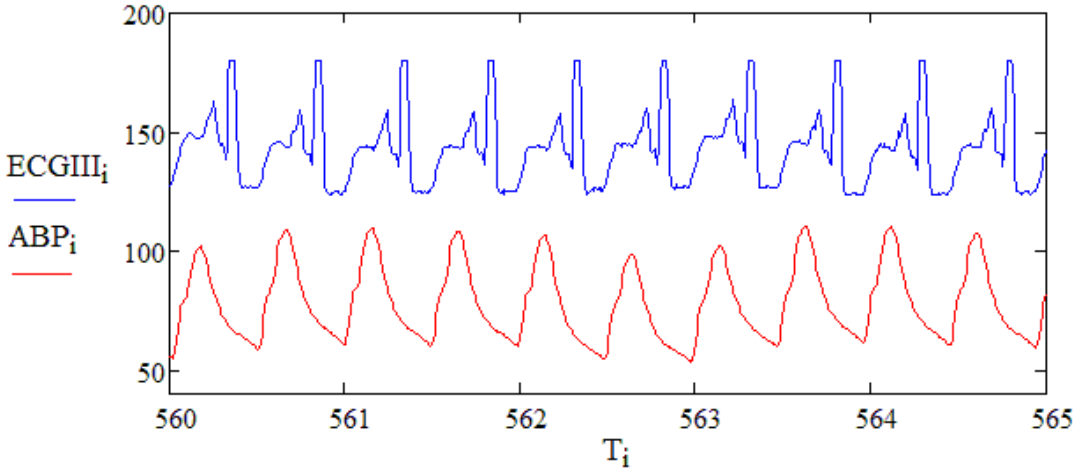

Figure 7.7: ECGIII and ABP signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

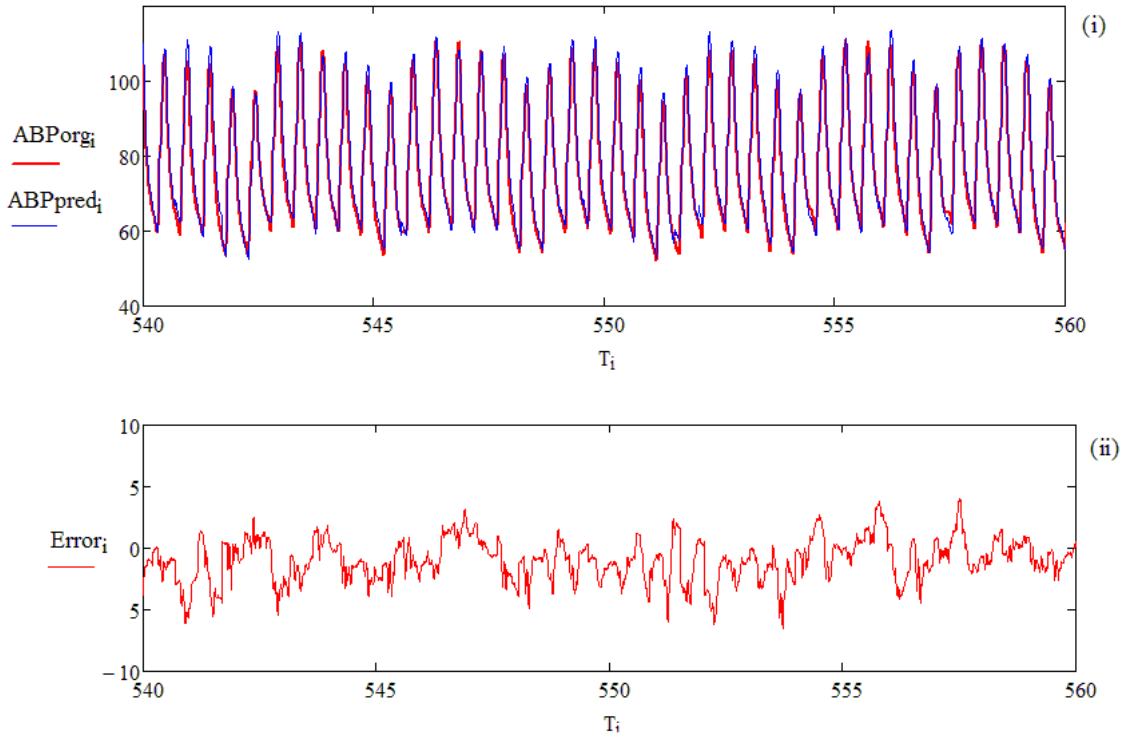

Figure 7.8: Predicted ABP signal and the original (i) and the error in Prediction, $Error = ABPorg - ABPpred$ (ii). Scores for Prediction ($Q1 = 0.9757, Q2 = 0.9932$)

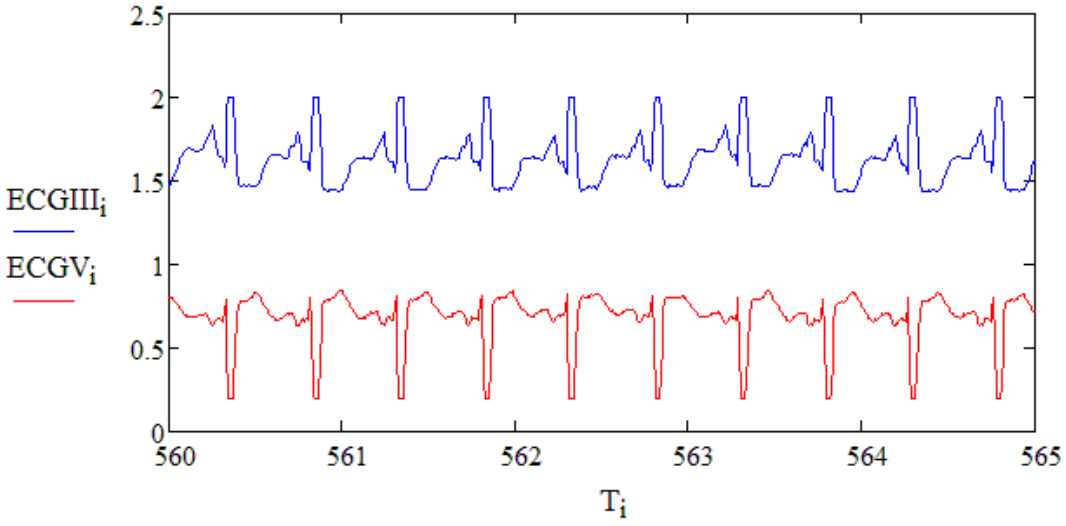

Figure 7.9: ECGIII and ECGV signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

for the 6–channel ECGV using the signal on 5–channel ECGIII. Fig. 7.9 shows a short segment of data of these two channels. Both the ECGIII and ECGV waves have identical time–periods. Fig. 7.10 shows the predicted ECGV signal and the original ECGV signal along with the corresponding error signal. Scores of Prediction are $(Q1 = 0.9933, Q2 = 0.9967)$.

3. Prediction of CVP signal from ECGII signal: Here the prediction was done for the 2–channel CVP using the signal on 4–channel ECGII. Fig. 7.11 shows a short segment of data of these two channels. Both the ECGIII and CVP waves have different time-periods of repetition. Fig. 7.12shows the predicted CVP signal and the original CVP signal along with the corresponding error signal. Scores of Prediction are $(Q1 = 0.6915, Q2 = 0.8508)$.

4. Prediction of CVP signal from RESP signal: Prediction is done for the 2–channel CVP using the signal RESP on 1–channel. Fig. 7.13 shows a short segment of data of these two channels. Both the CVP, RESP waves have identical time-periods of repetition. Fig. 7.14 shows the predicted CVP signal and the original CVP signal along with the error in prediction. Scores of Prediction are $(Q1 = 0.6149, Q2 = 0.8019)$. Note that the scores of RESP–CVP prediction are slightly lesser than the ECGII–CVP prediction scores.

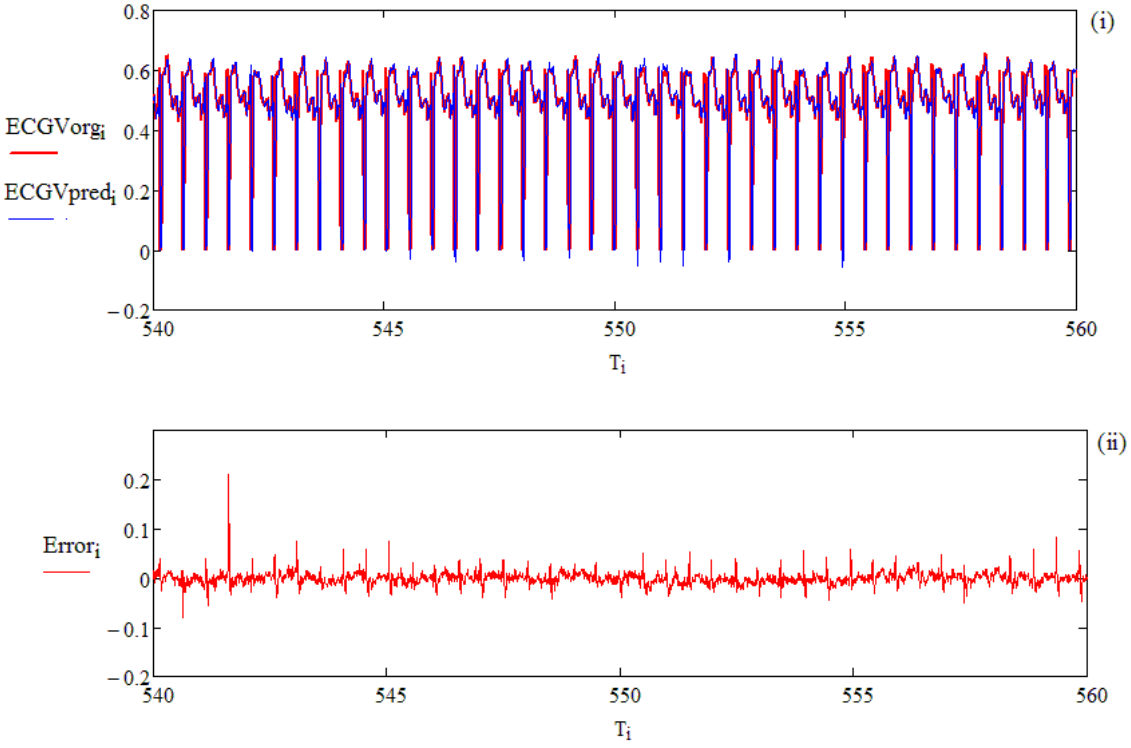

Figure 7.10: Predicted ECGV signal and the original (i) and the error in Prediction, $Error = ECGVorg - ECGVpred$ **(ii). Scores of Prediction** $(Q1 = 0.9933, Q2 = 0.9967)$**.**

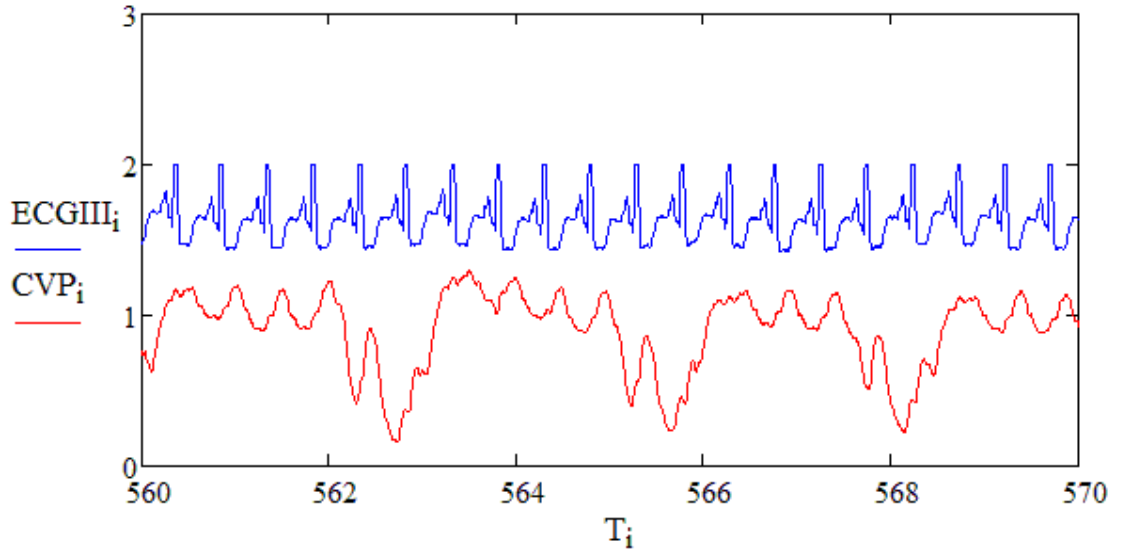

Figure 7.11: ECGIII and CVP signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

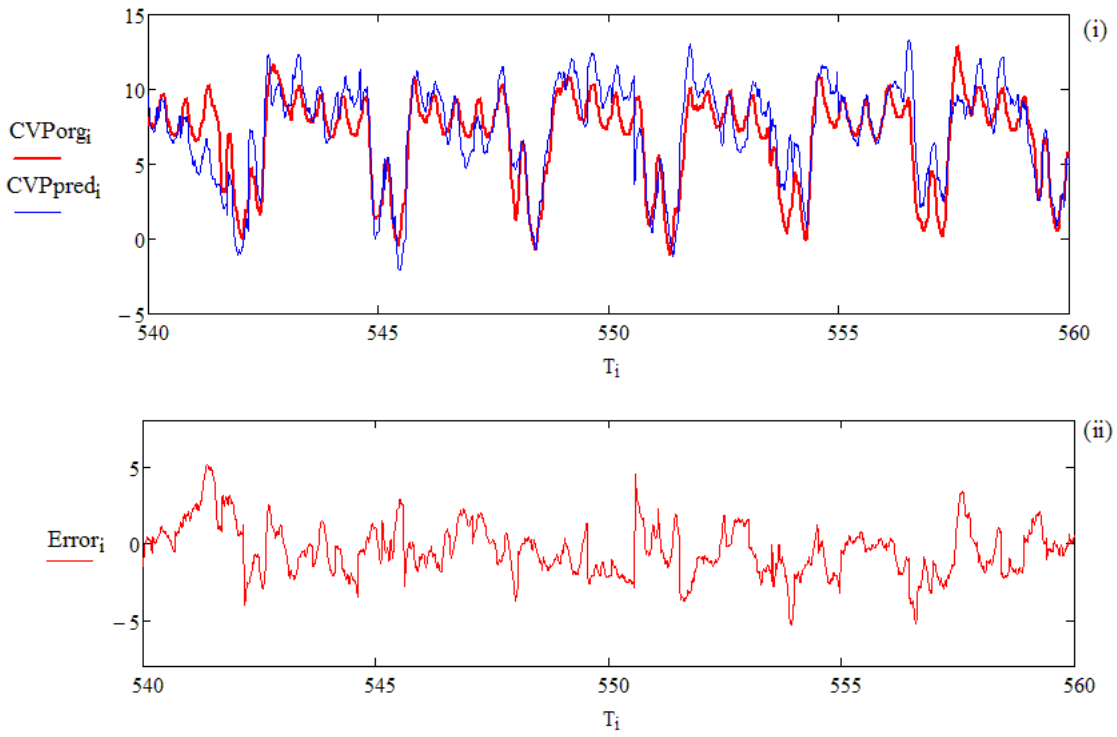

Figure 7.12: Predicted CVP signal and the original (i) and the error in Prediction, $Error = CVPorg - CVPpred$ (ii). Scores of Prediction are $(Q1 = 0.6915, Q2 = 0.8508)$.

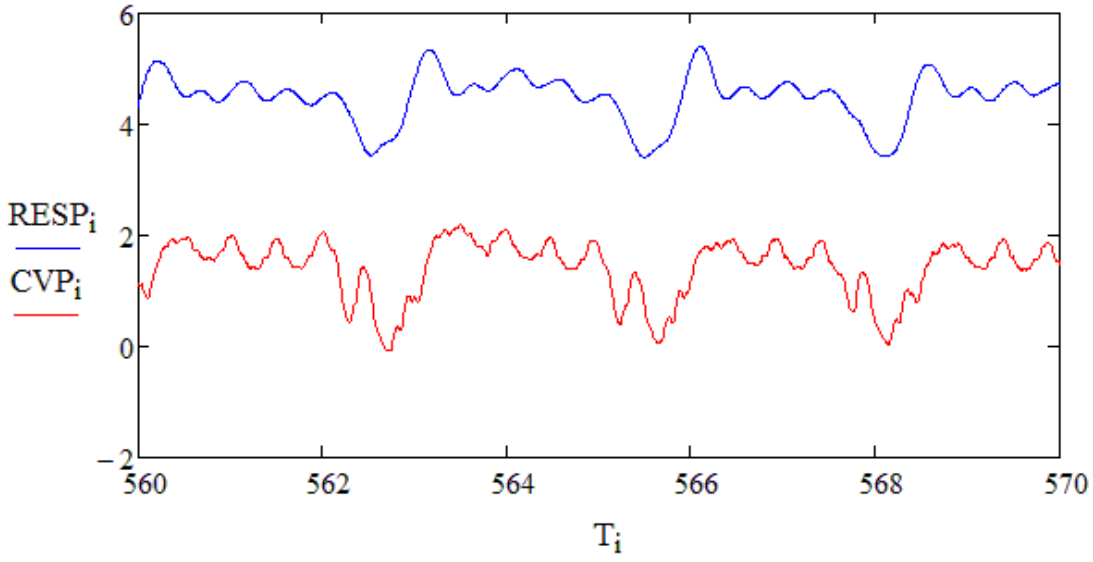

Figure 7.13: CVP and RESP signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

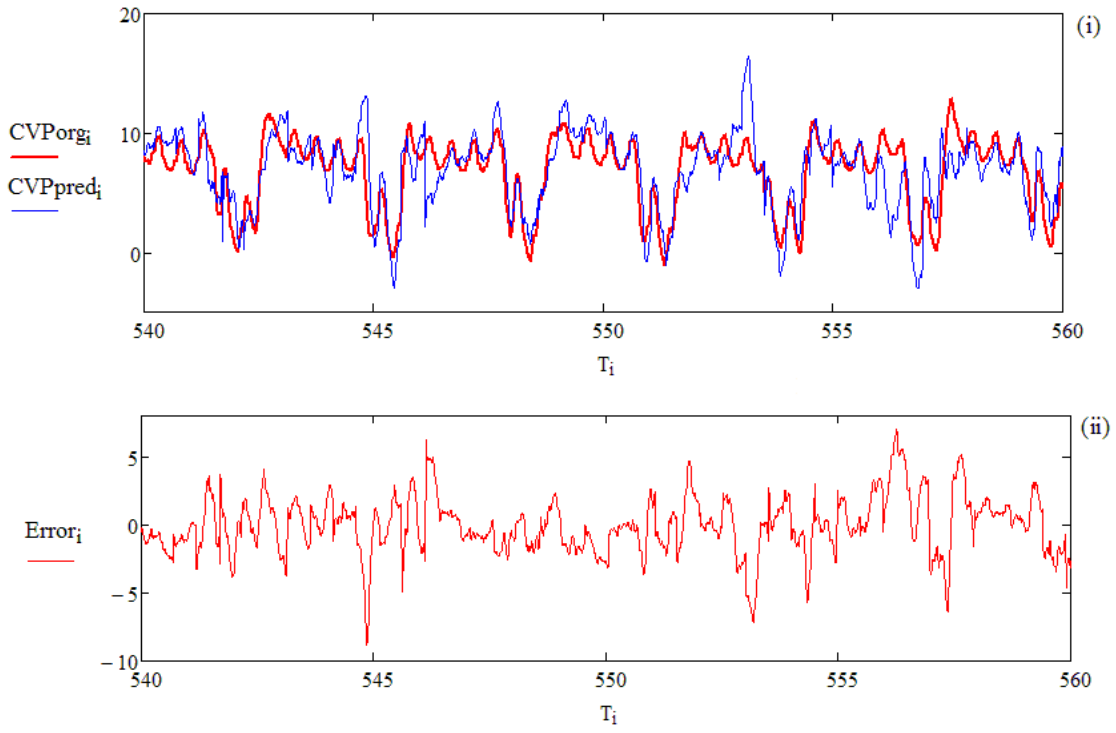

Figure 7.14: Predicted CVP signal and the original (i) and the error in Prediction, $Error = CVPorg - CVPpred$ **(ii). Scores of Prediction were** $(Q1 = 0.6949, Q2 = 0.8019)$**.**

5. Prediction of PLETH signal from ECGII signal: Here the prediction was done for the 3–channel PLETH using the signal on 4–channel ECGII. Fig. 7.15 shows a short segment of data of these two channels. Note from the figure that the high frequency component of the PLETH wave have identical repetition time with respect to ECGII signals. But there is low frequency component that repeats with a longer time period (almost a 5:1 ratio with the ECG wave). Fig. 7.16 shows the predicted PLETH signal and the original PLETH signal along with the corresponding error signal. Scores of Prediction were $(Q1 = 0.9162, Q2 = 0.9574)$.

6. Prediction of RESP signal from ECGV signal: Here the prediction was done for the 1–channel RESP using the signal on 6–channel ECGV. Fig. 7.17shows a short segment of data of these two channels. From the figure we can see that both the ECGV and RESP waves have different time-periods of repetition. Fig. 7.18 shows the predicted RESP signal and the original RESP signal along

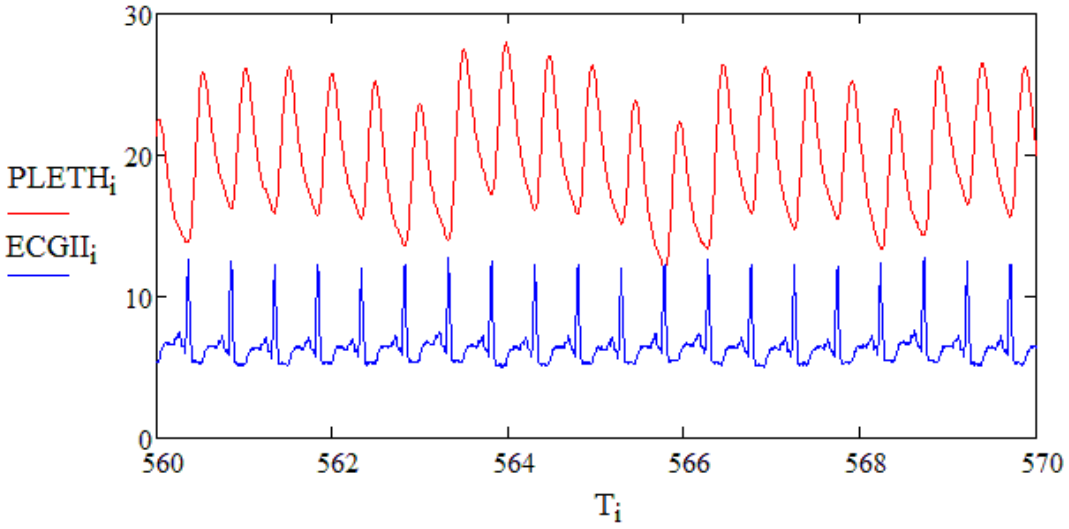

Figure 7.15: PLETH and ECGII signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

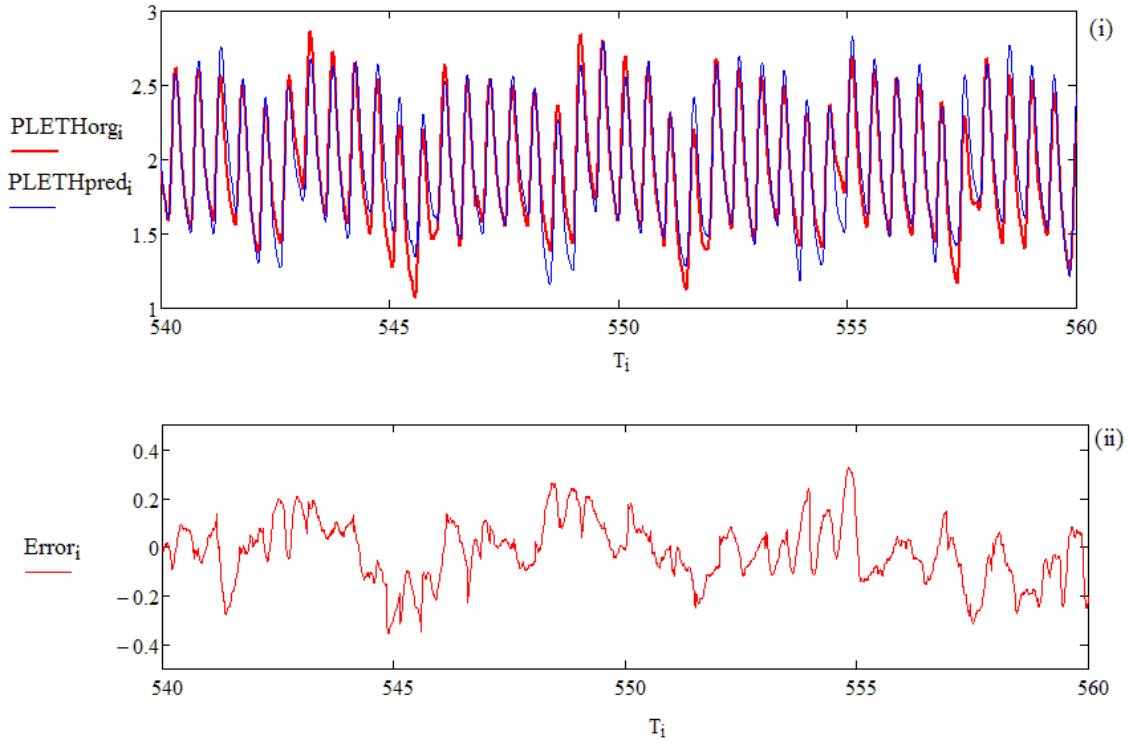

Figure 7.16: Predicted PLETH signal and the original (i) and the error in Prediction, $Error = PLETHorg - PLETHpred$ (ii). Scores of Prediction are $(Q1 = 0.9162, Q2 = 0.9574)$.

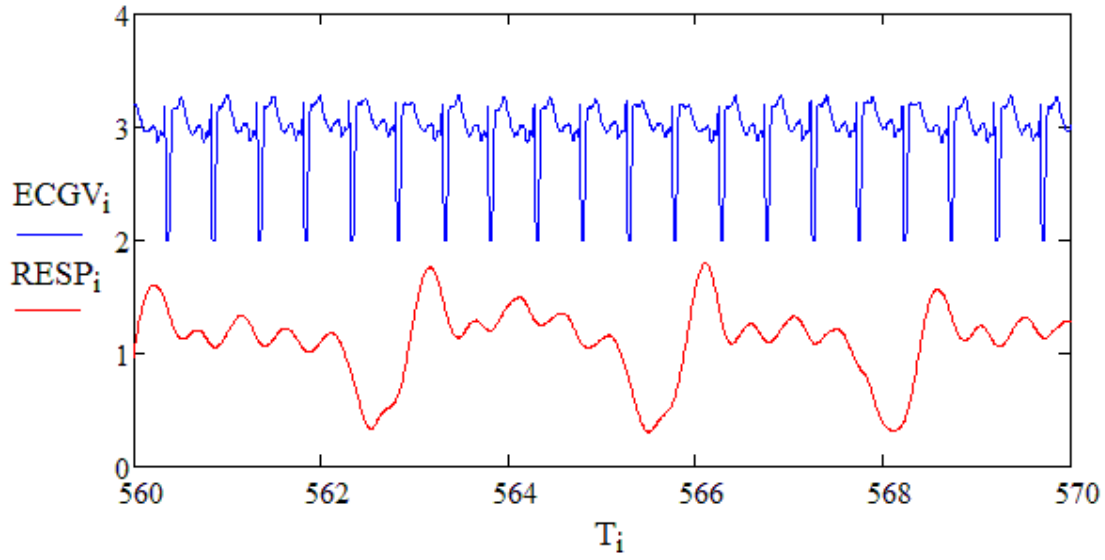

Figure 7.17: ECGV and RESP signals of *c21.dat* (a short data segment of both the signals scaled and shifted to fit the figure).

with the corresponding error signal. Scores of Prediction were $(Q1 = 0.6023, Q2 = 0.8248)$.

7.5 Prediction Scores for the entire Data set under study

Table. 7.1, Table. 7.2, Table. 7.3 contains the $(Q1, Q2)$ scores of ECGI, ECGII, ECGIII, ECGV, AVR prediction for the missing channel for the data records in Set C of Physionet Challenge 2010 archive [1]. Table. 7.4 and Table. 7.5 gives the scores for ABP, ICP, CVP signals. Table. 7.6 contains the PLETH prediction scores and Table. 7.7 has RESP prediction scores.

7.6 Discussion of Analysis Results

The introduced method of modeling was based on an affine transformation across the recurrent neighborhoods in a higher dimensional delay embedding space. This implied a possibility of developing transformations across the signals that are simultaneously recorded in a data set. For the prediction of $30s$ window data of length $30s$ was used for finding the conjugacy map. The parameters of the map

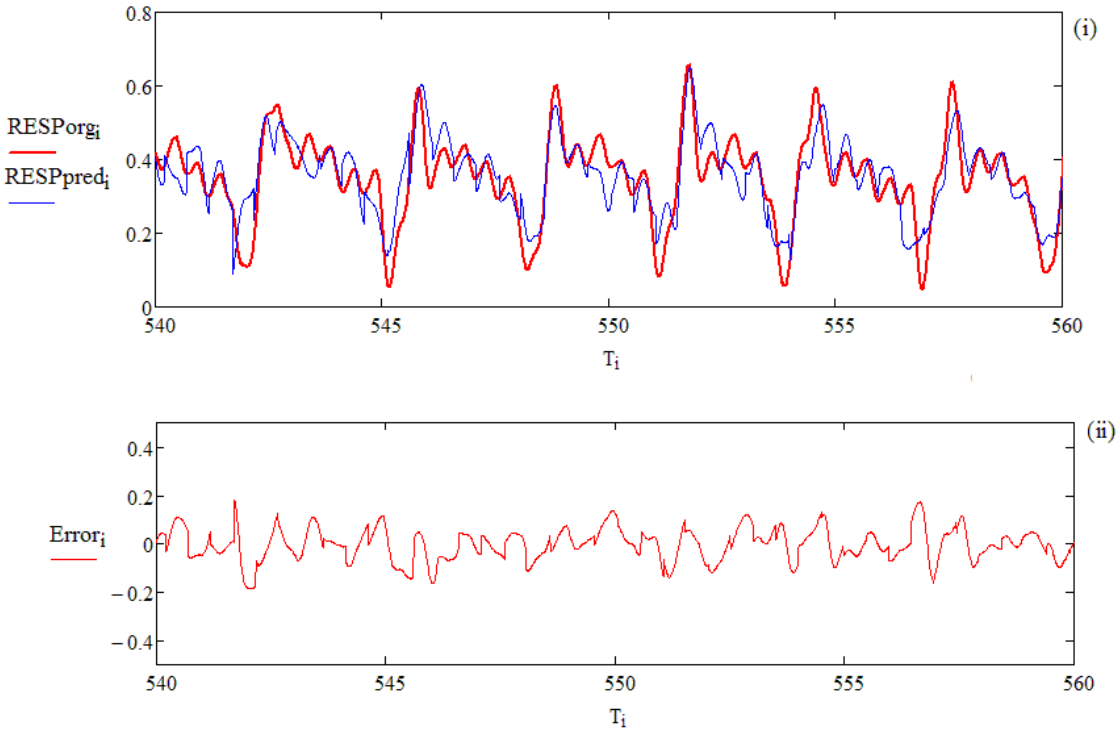

Figure 7.18: Predicted RESP signal and the original (i) and the error in Prediction, $Error = RESPorg - RESPpred$ (ii). Scores of Prediction are $(Q1 = 0.6023, Q2 = 0.8248)$.

Table 7.1: ECGII, ECGIII, ECGI Prediction Scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score Q1	Score Q2
c08	ECG II	0.9913	0.9965
c27	ECG II	0.9826	0.9914
c33	ECG II	0.7318	0.8623
c43	ECG II	0.9818	0.9909
c45	ECG II	0.9987	0.9994
c55	ECG II	0.9987	0.9994
c70	ECG II	0.9982	0.9991
c81	ECG II	0.9955	0.9978
c86	ECG II	0.9922	0.9961
c90	ECG II	0.9911	0.9958
c98	ECG II	0.9782	0.9898
c82	ECGIII	0.9923	0.9968
c84	ECGI	0.0121	0.6656

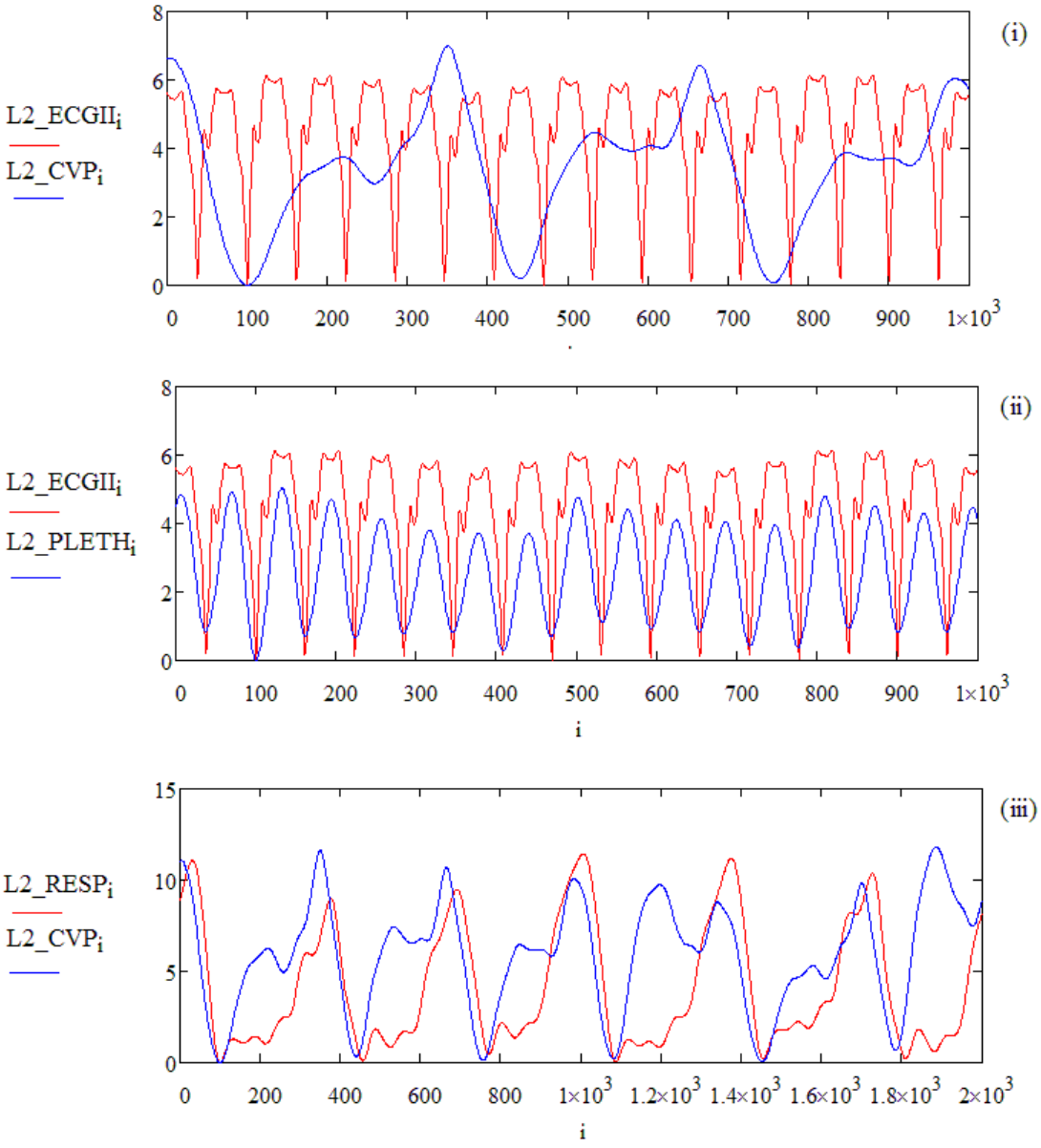

Figure 7.19: Recurrent timings for various signals in the data set *c21.dat*: (i) ECGII and CVP, (ii) ECGII and PLETH, (iii) RESP and CVP.

Table 7.2: ECGV Prediction Scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score Q1	Score Q2
c02	ECG V	0.8857	0.9429
c07	ECG V	0.9976	0.9988
c18	ECG V	0.9976	0.9988
c22	ECG V	0.8869	0.9430
c23	ECG V	0.9602	0.9804
c30	ECG V	0.6921	0.8516
c35	ECG V	0.9720	0.9859
c49	ECG V	0.9948	0.9974
c52	ECG V	0.4364	0.7286
c57	ECG V	0.9831	0.9922
c64	ECG V	0.9204	0.9611
c75	ECG V	0.9245	0.9626
c79	ECG V	0.9874	0.9937
c89	ECG V	0.9763	0.9882
c99	ECG V	0.8331	0.9272

Table 7.3: AVR Prediction Scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score Q1	Score Q2
c09	AVR	0.9850	0.9929
c10	AVR	0.9660	0.9834
c29	AVR	0.9973	0.9989
c38	AVR	0.2988	0.6334
c58	AVR	0.9754	0.9892
c71	AVR	0.8381	0.9169
c74	AVR	0.9971	0.9986
c78	AVR	0.9971	0.9985
c80	AVR	0.9051	0.9547
c88	AVR	0.9460	0.9745
c92	AVR	0.9736	0.9868

were developed based on a short data segment in the modeling window preceding the prediction window and they were assumed to be invariant across the windows for the purpose of prediction.

Analysis results on Table 7.1– Table 7.7 show that the predictions gives a good fit to the actual target for some of the signals of the data set under study. It is inferred that the model could potentially reflect the changes in the cardiovascular

Table 7.4: ABP prediction scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score Q1	Score Q2
c01	ABP	0.9229	0.9776
c04	ABP	0.9889	0.9951
c14	ABP	0.7590	0.8837
c19	ABP	0.9034	0.9723
c20	ART	0.9914	0.9958
c21	ABP	0.9724	0.9950
c31	ART	0.9691	0.9895
c34	ABP	0.9881	0.9956
c36	ABP	0.7650	0.8827
c37	ABP	0.9869	0.9945
c47	ABP	0.9553	0.9779
c56	ABP	0.9507	0.9952
c66	ABP	0.9320	0.9680
c68	ABP	0.9965	0.9984
c73	ABP	0.8399	0.9172
c76	ABP	0.9833	0.9920
c93	ABP	0.9751	0.9875
c95	ABP	0.9431	0.9803

Table 7.5: ICP, CVP prediction scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score Q1	Score Q2
c13	ICP	0.9336	0.9807
c17	ICP	0.5631	0.8900
c26	ICP	0.0000	0.9217
c53	ICP	0.9177	0.9882
c54	ICP	0.0000	0.8899
c11	CVP	0.0000	0.4347
c15	CVP	0.8323	0.9138
c41	CVP	0.0000	0.5117
c42	CVP	0.6050	0.8136
c46	CVP	0.0000	0.0000
c48	CVP	0.2570	0.6844
c59	CVP	0.8270	0.9147
c67	CVP	0.0000	0.3978
c72	CVP	0.7851	0.8928
c77	CVP	0.0000	0.1229

Table 7.6: PLETH prediction scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score 1	Score 2
c05	PLETH	0.0000	0.6145
c06	PLETH	0.0000	0.1192
c28	PLETH	0.8008	0.8975
c39	PLETH	0.9946	0.9973
c50	PLETH	0.2130	0.7856
c60	PLETH	0.8345	0.9180
c61	PLETH	0.0000	0.5775
c62	PLETH	0.0000	0.1641
c83	PLETH	0.9401	0.9697
c85	PLETH	0.0000	0.1961
c87	PLETH	0.1364	0.4020
c94	PLETH	0.0000	0.0000
c96	PLETH	0.9571	0.9829
c97	PLETH	0.7438	0.8647

Table 7.7: RESP prediction scores for data records in Set C [1]

Name of the data record	Predicted Channel	Score 1	Score 2
c00	RESP	0.5676	0.9055
c03	RESP	0.0000	0.4794
c12	RESP	0.2884	0.6764
c16	RESP	0.0000	0.0000
c24	RESP	0.4761	0.7165
c25	RESP	0.0000	0.1651
c32	RESP	0.0000	0.2691
c40	RESP	0.2395	0.6362
c44	RESP	0.8456	0.9198
c51	RESP	0.0000	0.2190
c63	RESP	0.5549	0.8054
c65	RESP	0.6169	0.7979
c69	RESP	0.0000	0.0337
c91	RESP	0.6995	0.8440

system efficiently for specific signals (ECGI, ECGII, ECGIII, ECGV, AVR, ABP). An excellent quality of reconstruction was obtained for all the ECG and BP signals. The PLETH, RESP, CVP scores were not as good as the ECG and BP scores. One reason for the bad prediction could be that either there is no connection between the channels (ECG Vs PLETH, RESP, CVP signals) or the connection was highly

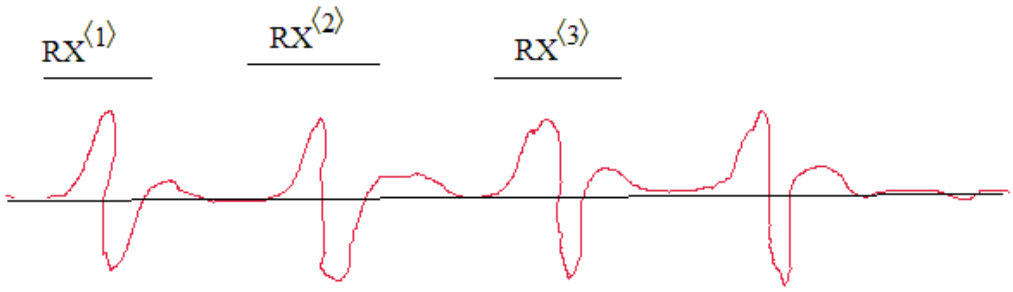

Figure 7.20: Pictorial Representation of insufficient embedding dimension resulting in recurrent vectors doing improper overlap of the time domain signal.

non-linear or the model developed was inefficient to extract the connection.

It is observed that RESP, CVP have identical recurrence patterns with respect to each other but a different recurrent timing with respect to the ECG signals as shown in 7.19. For the simulation results explained in the analysis section, recurrent neighborhoods and the recurrent timings were calculated with respect to the ECG /BP signals. There is an opportunity to improve the reconstruction of RESP, CVP signals by finding the recurrent neighborhoods and the recurrent timings with respect to them instead of ECG signal which we plan as a direction for future research.

7.6.1 When the Embedding dimension N Value is less than Required

An estimate of the embedding dimension is made such that the vectors of the recurrent neighborhood in R^N do a proper overlap of the time domain signal at the recurrent timings. When the embedding dimension is not adequate the embedding vectors of the neighborhood fail to overlap the time domain signal as shown in Fig. 7.20 which will result in an partial prediction of the signal. In this case, multiple recurrent neighborhoods can be used to predict the entire structure.

Fig. 7.21 show the partial prediction for ABP signals using ECGII signal of *c21.dat* when the embedding dimension was chosen as 30 instead of 110. In this case

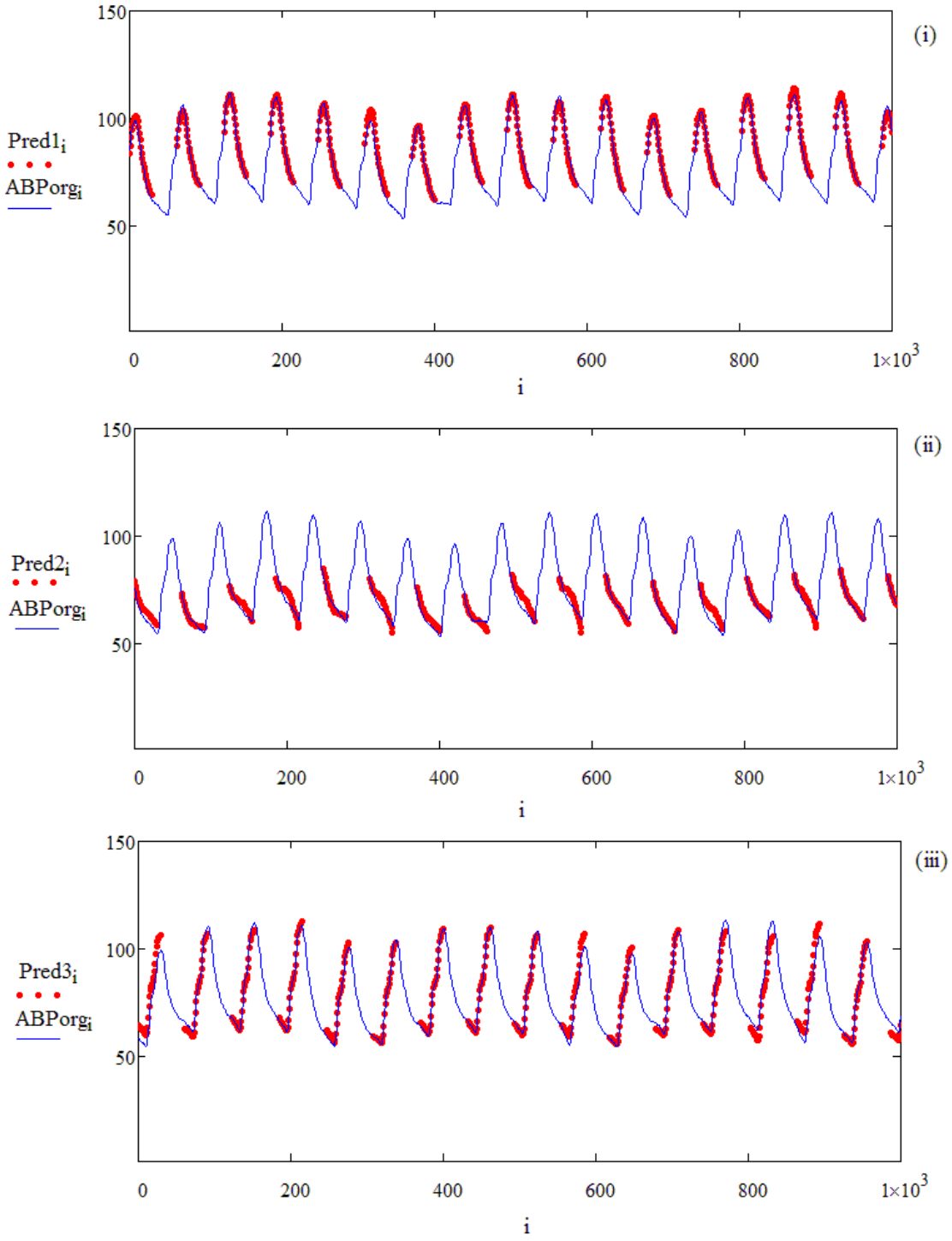

Figure 7.21: Pred1 (i), Pred2 (ii) and Pred3 (iii) are partial predictions for ABP signals using ECGII signal of *c21.dat* when the embedding dimension was chosen as 30 instead of 110.

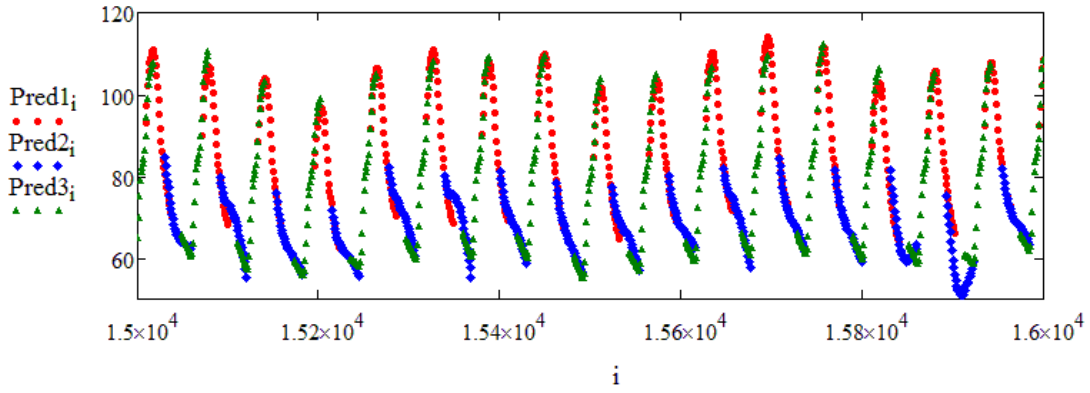

Figure 7.22: Patching up the partial predictions Pred1 (i), Pred2 (ii) and Pred3 (iii) of the from multiple recurrent neighborhoods in R^{30} depicted in Fig. 7.21 for a full prediction of the ABP signal.

multiple recurrent neighborhoods in R^{30} can be used for full prediction of signal. Fig. 7.22 shows the possibility of patching up the partial predictions for the full prediction of the entire signal.

7.7 Conclusions

This chapter demonstrated the efficacy of the model for the purpose of reconstruction of missing signals in a multichannel data set. It was observed during the data analysis that the scores of reconstruction is dependent on the quality of the signals that were available on the channels based on which the model was made. If the signals available in the modeling and prediction windows were noisy or clipped, prediction score were affected. The introduced method of modeling was based on an affine transformation across the recurrent neighborhoods in a higher dimensional delay embedding space. For the prediction of $30s$ window data of length $30s$ was used for finding the conjugacy map and the parameters of the map were assumed to be invariant across the windows for the purpose of prediction. Since the parameters of the map were developed based on a short data segment in the modeling window preceding the prediction window there was no need of extensive training strategies which drastically reduced the time of prediction. This is one advantage of the proposed method compared to the current best algorithms that deals with the same

issue of signal corruption and predictions. An immediate application of the method is the cure of loss/corruption of data in one of the channel, using the available data in other channels.

The results discussed in this chapter are motivated by Physionet Challenge 2010. The scores for 100 data records in set C given in Tables 7.1–7.7 had a total score of $(Q1 = 67.42, Q2 = 80.95)$ which ranked 6 among the participants of the challenge.

CHAPTER 8
Conclusions and Future Research Directions

8.1 Summary of Results

The foundation of this thesis was an exploratory research for finding new, effective methods for analyzing physiological signals. Modeling a system based on an observed time series is a difficult problem in general as the measurements contain only partial information about the system and further it could be noisy. The goal of the thesis is to develop a new methods of modeling a system based on a time series with a dynamical system perspective.

The method proposed in this thesis is specific to a special class of nonlinear signals that exhibits the property of recurrence. It is demonstrated here that the best way to make use of the recurrence property for prediction and modeling is to couple it with the topological technique of the delay embedding. The special structure of the delay embedding matrices are investigated as it contains lot of redundant information. One of the goals for the modeling was to exploit this redundant property to reduce the computational load that is inevitable in nonlinear data analysis.

The proposed model used a delay embedding at a higher dimensional space for the global representation of dynamics and a projection to a lower dimensional space for the local analysis of the system dynamics. The projection was done using an Finite Impulse Response (FIR) filters to tackle the issues of over embedding. According to Broomhead's theorem from Topology FIT filters preserve all the information one wants to extract by embedding techniques. Since a filtered time series of a system preserves the embedding, a projection from the high dimensional space to a lower dimensional space was used for the local analysis. This projection took care of two fundamental issues related to high dimensional models that describe a low dimensional dynamics: (i) the ambiguity that many models fit the data and (ii) the fact that the models become less economical, requiring more amount of data as the embedding dimension goes up.

As per the model, the recurrence neighborhoods were empirically determined

in the delay embedding space R^N and the projected low dimensional space R^d. The recurrent neighborhood in R^d is subdivided into various equivalence classes based on the recurrence times– the time delay between tow adjacent recurrences. A set of affine maps (which are functions of only the recurrent times) were derived across these equivalence classes. This procedure resulted in a possibility of simplifying the dynamics in terms of affine transformations in small neighborhoods. Thus, the set of overlapping neighborhoods with specific affine transformations for each of them was used to represent the local dynamics.

A topological conjugacy that exists between the recurrence neighborhoods reconstructed in both the lower and the higher dimensional spaces was demonstrated. Once the model is ready, given the recurrence timings (the interval between two consecutive recurrences) of the data, the entire data can be predicted. The proposed methods were demonstrated for the analysis of (i) data generated by the Duffing oscillator and (ii) an Electrocardiogram (ECG) record. It is shown that the entire nonlinear structure can be deduced from one or few overlapping neighborhoods for these data.

One of goals of the thesis was to introduce a new modeling method for nonlinear signals focusing on signal processing applications. A novel technique of nonlinear generalization of SVD was proposed for detecting and quantifying nonlinearity in a time series. This method is an extension of the standard SVD technique in linear algebra. Thesis demonstrated the technique of nonlinear SVD to retrieve nonlinearity from data generated by chaotic dynamical systems: Logistic map, Henon map, Van der Pol oscillator and Duffing oscillator. The recovery of parameters are shown in the following scenarios: (i) data generated by nonlinear maps and flows (ii) comparison of the method for both noisy and noise-free nonlinear data (iii) surrogate data analysis for both the noisy and noise-free cases, and also discussed two particular applications of the method: (i) Mathematical Modeling and (ii) Chaotic Cryptanalysis.

The method of nonlinear SVD is further extended to find the dimension of the system based on the time series. If the data on a manifold embedded in some R^N is available a method to compute the dimension of a manifold using nonlinear SVD is

discussed in the thesis. Two specific cases are discussed: For the simple case where the manifold is in the form of a lower dimensional affine subspace, the standard SVD is used to (i) calculate the dimension of the manifold and (ii) to get the equations which define the subspace. For the general case of manifolds, the nonlinear SVD is used (i) to search for an upper bound for the dimension of the manifold and (ii) to find the equations for the local charts of the manifold. This introduced method is highly useful in the context of the Takens' embedding- a technique that is used through out this thesis for data analysis. Finding a good estimate of the underlying dimension of an embedded data is a requirement while modeling a system based on local charts. Thesis provides examples of this type of modeling.

The method of modeling and prediction proposed in this thesis is expected to be of general utility for a large class of problems. One of the objectives of the thesis was to use these methods for the analysis of ECG data. First, the possibilities of the model were explored for a numerically generated data by an ideal dynamical system for a proper understanding and demonstration of the proposed method. The data generated by Duffing oscillator was specifically selected for this purpose as it is a well-studied system in nonlinear dynamics and it exhibited the property of recurrence. Then, the method was demonstrated for the analysis of the ECG record. For both the cases of the Duffing data and the ECG data the entire nonlinear structure was deduced from one or few overlapping neighborhoods.

A stability analysis for the data is done and it is verified using the Duffing data, that representing a system in a dimension higher than the inherent dynamics of the system leads to wrong inferences about the stability of the system. Since the exact dimension of the dynamics that governs ECG data is unknown, nonlinear SVD was used to find both the dimension of the manifold and the local dynamics. The local dimension of the manifold is equal to the local dimension of the system due to the conjugacy property. For the ECG data, an inference about the stability was made by studying the properties of affine maps specific to the neighborhoods. Heart rate variability, variability in beat-to-beat intervals of the heart, is known to be a necessity for healthy functioning of the heart. Based on the properties of the affine maps it is demonstrated that heart rate variability implies a stable dynamics

where as a uniform heartbeat could result in instability.

In addition to that dynamic models for various components of a multichannel physiological record– ECG (different channels), Arterial blood pressure, Central venous pressure, Intra Cranial Pressure, were developed and these models were used successfully for the prediction of one signal using others. The prediction results were excellent for the ECG and BP signals and the scores scores were low only when the signal was extremely noisy, lost or clipped. An immediate application of the proposed model is the cure of loss/corruption of data in one of the channel in a multichannel record, using available data in all the channels.

8.2 Contributions of the Thesis

While analyzing the signals understudy it was observed fact that there is lot of information available on RR intervals which can be used for the purpose of modeling. In addition to that the information available on RR intervals was used for the stability analysis in the thesis. The motivation for this study was the Physionet challenge 2010 which was about predicting a short segment of missing data from one of the channels in a multichannel physiological record. A few important questions which were under consideration while working on this issue were:

- Modeling various ECG and BP signals effectively by exploiting the recurrence property exhibited by the signals.
- Modeling a relationship that exists across various ECG channels and BP channels signals in a deeper physiological point of view as they are different measurements of the same cardiovascular system.
- Relationship between heart rate variability and heart dynamics.
- Theoretical explanation of a well known experimental observation that the lesser is a potential predictor for various heart ailments.

The research embodied in this thesis is attempt to answer these questions. Thesis demonstrates a an effective way to model various ECG and BP signals by exploiting the recurrence property. For some specific class of signals like various

ECG signals and Blood pressure, the cross predictions were excellent implying a generalized synchronization that exists between them. Thesis gives a theoretical explanation for a the observation that the good heart rate variability is linked to a healthy heart dynamics.

The proposed model in the thesis has a new perspective for analyzing biomedical signals. The model was developed using many ideas of Topology, Nonlinear Dynamics and Chaos theory. A contribution of the thesis to the existing body of knowledge in the field of physiological signal processing is demonstrated in the the thesis. Real time monitoring of vital physiological signals have an important relevance in clinical research and diagnosis. Corruption or loss of these signals are common as there is high chance for the signals to get distorted or lost or get immersed in noise in the hospital settings. A trained person may be able to deal with some of the noises and loss of signals in physiological data due to the cognitive abilities of human brain. But it is always challenging for an algorithm to mimic cognitive skills to recognize and ignore noises based on the contextual information. Taking care of the corruption and loss of signals for the purpose of diagnosis and forecasting is crucial in the area of biomedical research.

The proposed prediction methods were used to analyse an online data set of 300 multichannel records published by the Physionet. The introduced algorithms relied on the topological conjugacy between simultaneously recorded signals. For some specific class of signals like various ECG signals and BP, the cross predictions were excellent implying a generalized synchronization that exists between them. Analysis results showed that the predictions gives a good fit to the actual target for some of the signals of the data set under study. It is inferred that the model could potentially reflect the changes in the cardiovascular system efficiently for specific signals and a excellent quality of reconstruction was obtained for all the ECG and BP signals.

Thesis further discusses some new techniques and algorithms for nonlinear data analysis. A nonlinear extension of well known linear algebra technique SVD is proposed. Nonlinear SVD is proposed for finding the dimension of the manifold where the data resides. This method has potential for analyzing local dynamics

of the system as in general while analyzing the experimental data the system that generated the data is unknown.

Unique contribution of this thesis is that it has developed a new model for prediction by combining the recurrence properties of the signal and the topological ideas. The current best methods to deal with the missing data in multichannel records are mostly based on neural networks with different training strategies. But the main drawbacks of neural network based methods are the following: (i) they have a high computational load and long training time (iii) prediction is limited to the patterns that are available during the training sessions otherwise that behavior is not learned and lost during prediction. The proposed method gives an excellent prediction without any extensive training time. The possibilities of the model are explored using data generated by Duffing oscillator successfully and hence the model is expected to be potentially useful in many real-world applications.

8.3 Future Research Directions

The proposed method of prediction worked extremely well (in a range of scores 0.90 to 0.99 for most of cases) for cardiovascular signals (various ECG signals and some blood pressure (ABP,ICP) signals) that had identical recurrent timings. This can be attributed to the fact that they carried a lot of mutual information by basically representing activities of the same system. Also there was a possibility of some relationship between different channels that are simultaneously recorded from the same person. The model could efficiently exploit the redundancy and mutual relationships across some channels to predict or reconstruct the lost or corrupt as discussed in the last chapter of the thesis. But prediction were not that efficient for some other signals: plethmosgraphy (scores ranging around the value 0.5 with a highest score approaching 0.99) and CVP (score ranging around 0.5 with a highest score approaching 0.91) and Respiration (wide range of scores from 0 to 0.9 with 2 data sets scoring up to 0.92). Prediction scores for CVP and Plethmosgraphy were in general better than the prediction scores for the Respiration data. Plethmosgraphy signal had the same recurrence timings as the ECG and BP signals. CVP and Respiration signals had a different recurrence timings compared to ECG but they

were in phase synchronization with the ECG signals roughly in a $3:1$ ratio. There are some experimental observations about a feed back that exists between the cardiovascular system and the respiratory system. One can manipulate breathing rate to control heart rate. Meditation based techniques like deep breathing are known to reduce the blood pressure to a healthy lower rate there by improving the heart rate variability.

One of the key inference based on the study of the multichannel data set is that the prediction model was extremely accurate for signals that has identical recurrence timings given that it measured the activity of the same organ as in the case of ECG and ABP of the cardiovascular system. But the prediction was not excellent for signals that had (i) identical recurrent time but measured different activities (as in the case of ECG and Plethmosgraphy), (ii) different recurrence timings (as in CVP and RESP signals). It was also observed that even in the case of different recurrence timings ECG and RESP were synchronized in a phase locked behaviour though the recurrence timings were different. This opens up a new direction of research for improving the model (i) when signals have different recurrent timings (ii) when the signals have same recurrent timings but they represents vital measurements of different systems within the body. Another important extension of research would be developing an algorithm feasible for real time processing of multichannel data in medical labs.

On a dynamical system perspective, the proposed model relied on a hypothesis that the dynamics on the manifold can be simplified in terms of affine transformations in small neighborhoods everywhere. Thesis demonstrated the possibility of simplifying the dynamics in terms of affine transformations in small neighborhoods. The set of over lapping neighborhoods with specific affine transformations for each of them was used for the local analysis. The justification of this hypothesis was done numerically using on topological techniques of delay embedding and projection and the conjugacy transformations across the manifolds. Recalling the Hartman-Grobman theorem that the behaviour of a nonlinear system near an equilibrium point the system is qualitatively similar to the behaviour of a linear system [49], a generalization of this theorem could be worked out prove the hypothesis that the

dynamics of recurrent systems can be approximated by linear transformations in small neighborhoods at rest points, and affine transformation elsewhere. This could be possible of direction of future research for the local analysis of manifolds from experimental data. Additionally, thesis has shown that recurrence timings can reveal lot of the information about the system dynamics that can be used for prediction and modeling. The property of recurrence timings can be explored more in the context of symbolic sequences that are already proven to contain information about the initial condition of the system under study.

The thesis contains a peripheral result related to floquet theory, which did not fall into the main objectives of the thesis. Floquet theory deals with finding stability of the system around its periodic orbits. The floquet modes can be seen as affine spaces in the function domain. There are some standard methods that are available for finding various periodic orbits in the system for analyzing stability. The thesis demonstrated a method to find floquet modes of the recurrent orbits of duffing oscillator. The condition that orbit need to be periodic is not necessary as the recurrent orbits too can be seen as affine spaces. Floquet theory is important as the issue of stability is crucial in time series analysis. Ability to determine floquet coefficients from real time data could be a goal of future research.

APPENDIX A
Proof of Theorem 3.2

The Q linear algebraic homogeneous equations can always be written as,

$$GZ = 0 \tag{A.1}$$

where G is $Q \times N$ constant matrix of maximal rank Q, and Z is a column vector of dimension N representing the co-ordinates in R^N.

Now consider the transpose of the data matrix A. Each of the columns of A^T represents a vector which will also belong to the subspace governed by the Eq. 3.3 of Chapter 3. Recalling Eq. 3.6 of Chapter 3,

$$V^T A^T = W U^T$$

Now if we denote any of the columns of A^T by X and the corresponding column of the U^T by u,

$$V^T x = W u \tag{A.2}$$

$$x = VWu \tag{A.3}$$

But x lies on the subspace defined by $GZ = 0$.

Substituting Eq. A.3 for Z in Eq. A.1 we get,

$$GVWu = 0 \tag{A.4}$$

u is the equivalent co-ordinates of the data points in the U matrix. Since Eq. A.4 has to be true for all u,

$$GVW = 0 \tag{A.5}$$

Consider the matrix product $M = GV$. G has rank Q and V being an orthogonal matrix is of full rank N which is higher than Q. Now by using the Sylvester's

inequality for the product of two matrices [89], rank of the product GV is at least $Q + N - N = Q$. In fact, we know that it is Q because G has only Q rows.

Now consider the product of M with W and let the rank of W be R. Again by using the same inequality, the rank of MW is at least $Q + R - N$. But we know that it is zero, because of Eq. A.5. This implies that the rank of W is at the most $N - Q$. Since W is a diagonal matrix, it means that at least Q of the diagonal elements, the singular values must be zero. This proves the Theorem 3.2 for the linear subspace.

APPENDIX B

A verification of using the procedure of nonlinear SVD to find explicit nonlinear equations for charts on the manifold

Consider X data, 100 random numbers selected from a uniform noise distribution in the interval [0,1]. Generate Y data from X using a quadratic equation of the form $Y = aX + bX^2 + c$; where the coefficients were $a = 2, b = 3, c = 1.5$. Now the goal is to predict a, b, c from the data X, Y using the procedure of nonlinear SVD. The procedure to determine the coefficients of a quadratic equation are explained below.

step 1: Create a data matrix D with columns X, X^2, Y.

step 2: Remove the means $\bar{X}, \bar{X^2}, \bar{Y}$ from the columns X, X^2, Y to create a matrix A with columns $x1, x2, x3$, where $x1 = X - \bar{X}, x2 = X^2 - \bar{X^2}, x3 = Y - \bar{Y}$. Record the means: $\bar{X} = 0.50593, \bar{X^2} = 0.33601, \bar{Y} = 3.51989$

step 3: Do the SVD of A matrix. The singular values of the A matrix in this case were,

$$\begin{bmatrix} 14.62048 \\ 0.52504 \\ 1.21589 \times 10^{-15} \end{bmatrix}$$

And the third column of the V matrix were,

$$\begin{bmatrix} 0.53452 \\ 0.80178 \\ -0.26726 \end{bmatrix}$$

The 3^{rd} singular value can be considered to be zero because of the rounding error.

step 4: Using the Eq. 3.3 of Chapter 3, get a relationship between the third column of V matrix and the coordinates $x1, x2, x3$ as,

$$0.53452x1 + 0.80178x2 - 0.26726x3 = 0$$

$$x3 = \frac{0.53452}{0.26726}x1 + \frac{0.80178}{0.26726}x2$$

$$x3 = 2x1 + 3x2$$

Substitute for the actual data coordinates,

$$Y - \bar{Y} = 2(X - \bar{X}) + 3(X^2 - \bar{X^2})$$

$$Y - 3.51989 = 2(X - 0.50593) + 3(X^2 - 0.33601)$$

$$Y = 2X + 3X^2 + 1.5 \tag{B.1}$$

Hence the parameters of the quadratic equation are recovered.

APPENDIX C
Affine approximation of an Analytic Map in a small neighbourhood

Theorem C.1. *If F is a nonlinear real analytic map, then F has an affine representation f at any point a.*

Proof : If F is analytic, it has a valid taylors series at any point a. For any small neighborhood of a,

$$F(a+x) = F(a) + J(a)x + o(x^2) \tag{C.1}$$

Where $o(x^2)$ represents all the higher order components of the Taylor series. $J(a)$ is the matrix of partial derivatives of F (i.e. Jacobian matrix) at point a. Then the affine approximation of F at a can be written as,

$$f_a(x) = F(a) + J(a)x \tag{C.2}$$

for every $\epsilon > 0$, there exits a $\delta > 0$ such that for all x, with $|a - x| < \delta$, $|F(x + a) - f_a(x)| < \epsilon$. Eq. 2, is an affine transformation of the form,

$$f(x) = Ax + c \tag{C.3}$$

If f is defined in R^n, $A = J(a)$ is of dimension $(n \times n)$ and $c = F(a)$ is of dimension $(n \times 1)$. Note that Eq. C.3 can be linearized with respect to a rest point as shown by the following theorem.

Theorem C.2. *If $\bar{x}$ is the rest point of the affine map f, then f can be also written as a linear map with respect to its rest point.*

Proof : $\bar{x}$ is the rest point of f, hence $\bar{x} = A\bar{x} + c$ and $\bar{x} = (I - A)^{-1}c$. Define a new co-ordinate $y = x - \bar{x}$. Substituting for x, $y + \bar{x} = A(y + \bar{x}) + c$, hence $y = Ay$ which is a linear map.

APPENDIX D
Example: Ambiguities in Higher Dimensional Models that represents Lower Dimensional Dynamics

Consider a $2D$ dynamical system given by the following set of equations, generates data with an initial condition say (a, b);

$$\dot{x} = y \tag{D.1}$$
$$\dot{y} = f(x, y) \tag{D.2}$$

Assume that one finds a model for this system in $3D$ as,

$$\dot{x} = y \tag{D.3}$$
$$\dot{y} = z \tag{D.4}$$
$$\dot{z} = \frac{\partial f}{\partial x} y + \frac{\partial f}{\partial y} z \tag{D.5}$$

This gives us a freedom to choose an initial condition for the z variable. If one choose it as c, the corresponding $2D$ system would be,

$$\dot{x} = y \tag{D.6}$$
$$\dot{y} = f(x, y) + c - f(a, b) \tag{D.7}$$

And if $c = f(a, b)$ the model corresponds to the original $2D$ system. In general as c varies one gets different models of 2 dimensional manifolds in $3D$ [26].

APPENDIX E
Embedding Theorems

For a differentiable manifold M, a function $D : M \to R^N$ is an embedding if the image of the manifold $DM \subset R^N$ is also a manifold, and if the map between the manifolds $D : M \to DM$ is a diffeomorphism. Consider a dynamical system represented by (M, ψ) where M is a smooth manifold of dimension d on which the system evolves and ψ is the evolution function. Let the state of the system be denoted as, $x \in R^d$. Let h be an observation function of the system $h : R^d \to R$ that gives a time series $\{y\} = \{h(x)\}$ where $y \in R$.

Then N-dimensional delay vectors $\{z\}, z \in R^N$ can be constructed from the time series $\{y\}$ as $z_i = (y_i, y_{i+1}, \ldots y_{i+N-1})$. Each of these delay vectors $\{z\}$ are functions of the state variable x, since they are related to the evolution function ψ as, $z_i = (h(x_i), h(\psi(x_i)), \ldots h(\psi^{N-1}(x_i))$. Define the delay map $D_{\psi,h}(x)$ as,

$$D_{\psi,h}(x) = (h(x_i), h(\psi(x_i)), \ldots h(\psi^{N-1}(x_i)) \tag{E.1}$$

E.1 Takens' delay embedding theorem

Theorem E.1. *Let M be a compact manifold of dimension d. For pairs (ψ, h) with ψ a diffeomorphism of M and h a smooth real valued function on M, it is a generic property that $D_{\psi,h} : M \to R^N$ is an embedding if $N > 2d$. [25]*

This theorem proves a one–to–one correspondence between the statespace and the delay vectors generated from its time series y_i [25]. Hence for most of the evolution functions ψ and observation functions h, the delay map $D_{\psi,h} \in R^N$ is an embedding if N is large enough [90].

E.2 Broomhead's Theorem on Finite Impulse Response (FIR) filters that preserves embedding

The delay embedding theorem has found lots of practical applications as it opened a new method to retrieve information about the system from a time series. Takens theorem guarantied that the reconstruction would still be a copy of the original system and hence one could extract information about the original dynamics from the reconstructed system. But real world systems are far away from this ideal case, as in most of the practical cases we have only partial information about the system and the measurements itself could be noisy. So there was this important question whether the delay embedding is preserved even if the time series of a system is filtered. In 1992 Broomhead et.al. proved that finite order nonrecursive filters preserves all the information one want to extract by delay embedding techniques [48].

Consider the real valued measurements $\{y_j\}$ of the dynamical system (M, ψ). Delay embedding specified by Takens theorem assumes that the dynamical system evolves on an attractor which can be thought of as a subset of a finite dimensional differentiable manifold M. Apply a non-recursive filter to the measurements $\{y_j\}$ to get a new time series $\{u_j\}$ as follows,

$$\begin{aligned} u_j &= \sum_{k=1}^{n} b_k y_{j-k} \\ &= B * y_j \end{aligned}$$

where $B = b_1, b_2, \ldots b_n$ is a finite vector of filter coefficients. Then the theorem can be stated as follows,

Theorem E.2. *Let $\{y_j\}$ be a time series of measurements made on a dynamical system (M, ψ) which satisfies the hypotheses of Takens theorem. Then, for triples (B, ψ, y) it is a generic property that the method of delays, which constructs, from the time series $\{u_j\}$ vectors of the form $(u_j, u_{j-1}, \ldots u_{j-m+1})$ for any finite $m > 2d + 1$ and $u_j = B * y_j$ gives an embedding of M.*

LITERATURE CITED

[1] Mind the Gap – PhysioNet Computing in Cardiology Challenge 2010 Data archive Set C . [Online]. Available: http://physionet.org/physiobank/database/challenge/2010/set-c/

[2] S. H. Strogatz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry and Engineering (Studies in Nonlinearity)*. Reading, MA: Perseus Books, 1994.

[3] M. Lakshmanan and S. Rajasekar, *Nonlinear Dynamics, Integrability Chaos and Patterns*. Springer–Verlag Heidelberg, 2003.

[4] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems*. Springer, 1996.

[5] E. Ott, T. Sauer, and J. A. Yorke, "Coping with chaos. analysis of chaotic data and the exploitation of chaotic systems," *Wiley Series in Nonlinear Science, New York: John Wiley*, vol. 1, 1994.

[6] M. K. S. Yeung, J. Tegner, and J. J. Collins, "Reverse engineering gene networks using singular value decomposition and robust regression," *PNAS*, vol. 99, pp. 6163–6168, 2002.

[7] O. Edfors, M. Sandell, J. Van de Beek, S. Wilson, and P. Borjesson, "Ofdm channel estimation by singular value decomposition," *Communications, IEEE Transactions on*, vol. 46, no. 7, pp. 931–939, 1998.

[8] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *Automatic Control, IEEE Transactions on*, vol. 25, no. 2, pp. 164–176, 1980.

[9] B. Luo and E. Hancock, "Structural graph matching using the em algorithm and singular value decomposition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 10, pp. 1120–1136, 2001.

[10] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, "Numerical recipes in c: The art of scientific computing. 2," *Cambridge: CUP*, 1992.

[11] G. Strang, *Linear Algebra and Its Applications. Thomson–Brooks*. Cole, Belmont, CA, USA, 2003.

[12] D. S. Broomhead and G. P. King, "Extracting qualitative dynamics from experimental data," *Physica D: Nonlinear Phenomena*, vol. 20, pp. 217–236, 1986.

[13] A. Sahay and K. R. Sreenivasan, "Low-dimensional characterization of a local climate system," *Phil. Trans. R. Soc. Lond. A*, vol. 354, pp. 1715–1750, 1996.

[14] H. Abarbanel, R. Brown, J. Sidorowich, and L. Tsimring, "The analysis of observed chaotic data in physical systems," *Reviews of modern physics*, vol. 65, no. 4, p. 1331, 1993.

[15] B. A. Boukamp, "A nonlinear least squares fit procedure for analysis of immittance data of electrochemical systems," *Solid State Ionics*, vol. 20, no. 1, pp. 31–44, 1986.

[16] N. Packard, J. Crutchfield, J. Farmer, and R. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.

[17] M. Muldoon, R. MacKay, J. Huke, and D. Broomhead, "Topology from time series," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1, pp. 1–16, 1993.

[18] J. Cremers and A. Hübler, "Construction of differential equations from experimental data," *Zeitschrift für Naturforschung. A, A Journal of physical sciences*, vol. 42, no. 8, pp. 797–802, 1987.

[19] J. L. Breeden and A. Hübler, "Reconstructing equations of motion from experimental data with unobserved variables," *Physical Review A*, vol. 42, no. 10, p. 5817, 1990.

[20] K. Judd and A. Mees, "On selecting models for nonlinear time series," *Physica D: Nonlinear Phenomena*, vol. 82, no. 4, pp. 426–444, 1995.

[21] ——, "Embedding as a modeling problem," *Physica D: Nonlinear Phenomena*, vol. 120, no. 3, pp. 273–286, 1998.

[22] J. P. Crutchfield and B. S. McNamara, "Equations of motion from a data series," *Complex systems*, vol. 1, no. 3, pp. 417–452, 1987.

[23] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical review letters*, vol. 59, no. 8, p. 845, 1987.

[24] H. Whitney, "Differentiable manifolds," *Ann. of Math.*, vol. 37, no. 2, pp. 645–680, 1936.

[25] F. Takens, "Detecting strange attractors in turbulence," *Dynamical systems and turbulence, Lecture notes in mathematics*, vol. 898, no. 9, pp. 366–381, 1981.

[26] P. Vaidya and S. Majumder, "Embedding in higher dimensions causes ambiguity in determining equations from data," *The European Physical Journal-Special Topics*, vol. 165, no. 1, pp. 15–24, 2008.

[27] B. P. Bezruchko and D. A. Smirnov, "Constructing nonautonomous differential equations from experimental time series," *Physical Review series E*, vol. 63, no. 1; Part A, pp. 016 207–016 207, 2001.

[28] N. Marwan, "A historical review of recurrence plots," *The European Physical Journal Special Topics*, vol. 164, no. 1, pp. 3–12, 2008.

[29] E. Ott, *Chaos in dynamical systems*. Cambridge university press, 2002.

[30] J. Eckmann, S. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *EPL (Europhysics Letters)*, vol. 4, no. 9, pp. 973–977, 1987.

[31] J. P. Zbilut and C. L. Webber, "Embeddings and delays as derived from quantification of recurrence plots," *Physics letters A*, vol. 171, no. 3, pp. 199–203, 1992.

[32] C. L. Webber Jr and J. P. Zbilut, "Recurrence quantification analysis of nonlinear dynamical systems," *Tutorials in contemporary nonlinear methods for the behavioral sciences*, pp. 26–94, 2005.

[33] C. Webber and J. P. Zbilut, "Dynamical assessment of physiological systems and states using recurrence plot strategies," *Journal of Applied Physiology*, vol. 76, no. 2, pp. 965–973, 1994.

[34] A. D. Wyner, J. Ziv, and A. J. Wyner, "On the role of pattern matching in information theory," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2045–2056, 1998.

[35] A. Antoniou and C. Vorlow, "Recurrence plots and financial time series analysis," *Neural Network World*, vol. 10, no. 1/2, pp. 131–146, 2000.

[36] N. Thomasson, T. J. Hoeppner, C. L. Webber Jr, and J. P. Zbilut, "Recurrence quantification in epileptic eegs," *Physics Letters A*, vol. 279, no. 1, pp. 94–101, 2001.

[37] G. B. Moody, "The physionet computing in cardiology challenge 2010– mind the gap," *Computing in Cardiology*, vol. 37, pp. 305–308, 2010.

[38] A. Hartmann, "Reconstruction of missing cardiovascular signals using adaptive filtering," *Computing in Cardiology*, vol. 37, pp. 321–324, 2010.

[39] I. Silva, "Physionet 2010 challenge: A robust multi-channel adaptive filtering approach to the estimation of physiological recordings," *Computing in Cardiology*, vol. 37, pp. 313–316, 2010.

[40] Y. Li, Y. Sun, P. Sondhi, L. Sha, and C. Zhai, "Reconstructing missing signals in multi-parameter physiologic data by mining the aligned contextual information," *Computing in Cardiology*, vol. 37, pp. 449–452, 2010.

[41] P. Langley, S. King, D. Z. K. Wang, R. Giovannini, M. Bo-jarnejad, and A. Murray, "Estimation of missing data in multi- channel physiological time-series by reference timing channel and average substitution." *Computing in Cardiology*, vol. 37, pp. 309–312, 2010.

[42] R. Petrolis, R. Simoliuniene, and A. Krisciukaitis, "Principal component analysis-based method for reconstruction of fragments of corrupted or lost signal in multilead data reflecting electrical heart activity and hemodynamics," *Computing in Cardiology*, vol. 37, pp. 437–439, 2010.

[43] W. Wu, "Multi-parameter physiologic signal reconstruction by means of wavelet singularity detection and signal corre- lation," *Computing in Cardiology*, vol. 37, pp. 457–459, 2010.

[44] Mind the Gap – PhysioNet Computing in Cardiology Challenge 2010 archive of Papers. [Online]. Available: http://physionet.org/challenge/2010/papers/

[45] R. Rodrigues, "Filling in the gap: a general method using neural networks," *Computing in Cardiology*, vol. 37, pp. 453–456, 2010.

[46] S. Haykin, "A comprehensive foundation," *Neural Networks*, vol. 2, 2004.

[47] G. F. T. Del Castillo, *Differentiable manifolds: a theoretical physics approach*. Springer, 2011.

[48] D. Broomhead, J. Huke, and M. Muldoon, "Linear filters and non-linear systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 373–382, 1992.

[49] L. Perko, *Differential Equations and Dynamical Systems*. Springer–Verlag, New York, 1991.

[50] G. Floquet, "Sur les equations differentielles lineaires," *Annales Ecole Norm. Sup [2]*, pp. 47–88, 1883.

[51] Mind the Gap – PhysioNet Computing in Cardiology Challenge 2010 Data archive Sets A,B,C. [Online]. Available: http://physionet.org/physiobank/database/challenge/2010/

[52] P. G. Vaidya, P. S. Sajini Anand, and N. Nagaraj, "A nonlinear generalization of singular value decomposition and its applications to mathematical modeling and chaotic cryptanalysis," *Acta applicandae mathematicae*, vol. 112, no. 2, pp. 205–221, 2010.

[53] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "Svd-based methodologies for fetal electrocardiogram extraction," *IEEE International Conference Proceedings ICASSP*, vol. 6, pp. 3771 – 3774, 2000.

[54] J. Bhattacharya and P. Kanjilal, "On the detection of determinism in a time series," *Physica D: Nonlinear Phenomena*, vol. 132, no. 1, pp. 100–110, 1999.

[55] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1, pp. 189–208, 1983.

[56] A. Osborne and A. Provenzale, "Finite correlation dimension for stochastic systems with power-law spectra," *Physica D: Nonlinear Phenomena*, vol. 35, no. 3, pp. 357–381, 1989.

[57] D. Ruelle, "The claude bernard lecture, 1989. deterministic chaos: The science and the fiction," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pp. 241–248, 1990.

[58] G. Sugihara and R. May, "Nonlinear forecasting: An operational way to distinguish chaos from measurement error," *Nature*, vol. 344, pp. 734–741, 1990.

[59] L. Breiman and J. Friedman, "Estimating optimal transformations for multiple regression and correlation," *Journal of the American Statistical Association*, vol. 80(391), p. 580598, 1985.

[60] K. Chon, M. Korenberg, and N. Holstein-Rathlou, "Application of fast orthogonal search to linear and nonlinear stochastic systems," *Annals of biomedical engineering*, vol. 25, no. 5, pp. 793–801, 1997.

[61] S. Lu, K. Ju, and K. Chon, "A new algorithm for linear and nonlinear arma model parameter estimation using affine geometry [and application to blood flow/pressure data]," *Biomedical Engineering, IEEE Transactions on*, vol. 48, no. 10, pp. 1116–1124, 2001.

[62] Y. Nievergelt, "Total least squares: State-of-the-art regression in numerical analysis," *SIAM review*, vol. 36, no. 2, pp. 258–264, 1994.

[63] S. Lu and K. Chon, "Nonlinear autoregressive and nonlinear autoregressive moving average model parameter estimation by minimizing hypersurface distance," *Signal Processing, IEEE Transactions on*, vol. 51, no. 12, pp. 3020–3026, 2003.

[64] A. Porta, G. Baselli, S. Guzzetti, M. Pagani, A. Malliani, and S. Cerutti, "Prediction of short cardiovascular variability signals based on conditional distribution," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 12, pp. 1555–1564, 2000.

[65] A. Porta, S. Guzzetti, R. Furlan, T. Gnecchi-Ruscone, N. Montano, and A. Malliani, "Complexity and nonlinearity in short-term heart period

variability: comparison of methods based on local nonlinear prediction," *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 1, pp. 94–106, 2007.

[66] V. Marmarelis, "Identification of nonlinear biological systems using laguerre expansions of kernels," *Annals of biomedical engineering*, vol. 21, no. 6, pp. 573–589, 1993.

[67] P. G. Vaidya, "Monitoring and speeding up chaotic synchronization," *Chaos, Solitons & Fractals*, vol. 17, no. 2, pp. 433–439, 2003.

[68] R. Gonzalez, R. Woods, and S. Eddins, *Digital image processing using MATLAB*. Pearson Education India, 2004.

[69] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1, pp. 77–94, 1992.

[70] M. Kennel and S. Isabelle, "Method to distinguish possible chaos from colored noise and to determine embedding parameters," *Physical Review A*, vol. 46, no. 6, p. 3111, 1992.

[71] P. G. Vaidya and P. S. Sajini Anand, "Estimating the dimension of a manifold and finding local charts on it by using nonlinear singular value decomposition," *Topology Proceedings*, vol. 43, pp. 1–15, 2014.

[72] J. A. Suykens and J. Vandewalle, *Nonlinear Modeling: advanced black-box techniques*. Springer, 1998.

[73] H. Lauwerier, *Fractals: endlessly repeated geometrical figures*. Princeton Univ. Press, Princeton, NJ, 1991.

[74] [Online]. Available: http://en.wikipedia.org/wiki/Heart

[75] M. Malik and A. J. Camm, "Heart rate variability," *Clinical Cardiology*, vol. 13, no. 8, pp. 570–576, 1990.

[76] H. Kantz and T. Schreiber, *Nonlinear time series analysis*. Cambridge university press, 2004, vol. 7.

[77] D. Pradhan, *National Institute of Mental Health and Neurosciences (NIMHANS)*, Bangalore, India. PIN 560029.

[78] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software qrs detection," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 21, no. 1, pp. 42–57, 2002.

[79] Robust Detection of Heart Beats in Multimodal Data: the PhysioNet/Computing in Cardiology Challenge 2014. [Online]. Available: http://www.physionet.org/challenge/2014/

[80] R. Govindan, K. Narayanan, and M. Gopinathan, "On the evidence of deterministic chaos in ecg: Surrogate and predictability analysis," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 8, no. 2, pp. 495–502, 1998.

[81] P. G. Vaidya and N. Nagaraj, "Foundational issues of chaos and randomness: "God or Devil, do we have a choice?" in *Proc. of Foundations of Sciences (editor: B. V. Sreekantan), Published by Project of History of Indian Science, Philosophy and Culture (PHISPC), New Delhi, India*, 2006.

[82] M. Dahleh, M. A. Dahleh, and G. Verghese, "Lectures on dynamic systems and control," *A+ A*, vol. 4, no. 100, pp. 1–100, 2004.

[83] J. K. Hale, "Ordinary differential equations," *Pure and Applied Mathematics, John Wiley & Sons, New York*, 1969.

[84] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[85] G. E. Hinton, "Learning multiple layers of representation," *Trends in cognitive sciences*, vol. 11, no. 10, pp. 428–434, 2007.

[86] ——, "To recognize shapes, first learn to generate images," *Progress in brain research*, vol. 165, pp. 535–547, 2007.

[87] A. M. Sullivan, H. Xia, J. McBride, and X. Zhao, "Reconstruction of missing physiological signals using artificial neural networks," *Computing in Cardiology*, vol. 37, pp. 317–320, 2010.

[88] L. Silva, J. Duque, R. Tin's, and L. Murta-Jr, "Reconstruction of multivariate signals using $q-$gaussian radial basis function network," *Computing in Cardiology*, vol. 37, pp. 445–447, 2010.

[89] [Online]. Available: http://en.wikipedia.org/wiki/Rank_(linear_algebra)

[90] J. P. Huke, "The dynamical systems approach to nonlinear signal processing." [Online]. Available: http://personalpages.manchester.ac.uk/staff/jerry.huke/intro.pdf