Noise Reduction in Chaotic Systems using Unstable Periodic Orbits

by

Leon Beale

M.Sc. (Hons) University of Canterbury (1989)

This thesis is presented for the degree of

Doctor of Philosophy

Department of Mathematics

2018

Certification

This thesis entitled:

Noise Reduction in Chaotic Time Series using Unstable Periodic Orbits

written by: **Leon Beale**

has been approved by the Department of Mathematics at Macquarie University.

Paul Smith — Date: 20/11/2018

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Declaration

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than Macquarie University.

I also certify that this thesis is an original piece of research and has been written by me. Any help and assistance that I have received in my research work and preparation of the thesis itself has been appropriately acknowledged.

In addition, I certify that all information sources and literature are indicated in the thesis.

Date: 20/11/2018

Leon Beale

Acknowledgements

I am deeply grateful to my supervisor Professor Paul Smith. Without his continuous optimism concerning this work, enthusiasm, encouragement and support this study would hardly have been completed. I also thank Paul for his patience, sympathetic manner and understanding as I bridged 25 years of learning since my Masters degree, and balanced family and study responsibilities. Further, I wish to thank Deb Kane, my co-supervisor for her guidance and continuous encouragement.

I would also like to acknowledge my gratitude to the kind people from the Department of Mathematics at Macquarie University for their help and support. In particular I would like to thank Audrey Markowskei for her guidance with Matlab and Latex and continuous support.

Finally, I would like to thank my wife, Jessie and children Franklin, Daniel and Katelyn for their patience, understanding and positivity throughout the last few years as I took time out to pursue my passion. In particular, I thank my wife Jessie for her unwavering support of my research endeavours and belief in me.

Leon Beale

Sydney, Australia

2018

Abstract

In this thesis, a method is presented to approximate a noise-free low-dimensional continuous chaotic system (flow), using a single sample or multi-sampled scalar time series containing high levels of measurement noise or low levels of dynamical noise. The shadow-UPO noise reduction method (SUNR method) does not require the prior embedding of data and operates directly on the sampled time series, thus avoiding the limitations of Takens theorem and the estimation of embedding parameters when significant levels of noise are present. The method aims to overcome the well-documented severe limitations of directly filtering noise-infected chaotic time series, by focusing on nearly periodic orbit segments ('shadow-UPOs') shadowing the dense set of unstable periodic cycles (UPOs) that form the skeleton of a chaotic system, each of which is locally amenable to linear filtering techniques.

The innovation is two-fold and comes from firstly deconstructing the chaotic system into approximate cycles, where we are free to directly apply signal processing techniques, based on the specific type of noise. Secondly, shadow-UPOs are detectable in the presence of high noise using the observation that histograms constructed from recurrence matrices are highly robust to noise.

Shadow-UPOs are located, allocated to categorical bins, and filtered. We firstly utilise these to estimate the basis set of noise-free lower order UPOs, and estimate individual maximal Lyapunov exponents for each UPO. Secondly, we approximate the noise-free time series by replacing noise-infected near-cycles in the time series with their noise-filtered counterparts. The resultant time series are sufficiently noise-reduced that conventional algorithms can be used to subsequently estimate the Lyapunov exponents that would otherwise not be computable.

The method is illustrated in detail as a case study of the Rossler system, tested for various types of noise (uniform white, Gaussian white, high-frequency, coloured and dynamical) and also on several chaotic systems with a range of differing topologies (Chua, Rabinovich-Fabrikant, Lu-Chen, Lorenz). Goodness of fit metrics are defined, measured for each system and presented. We identified limitations of the recurrence method of detecting cycles when dealing with higher instability systems, and successfully modified the SUNR method for these.

List of Tables

Notation

m	The dimension of the delay vectors.
τ	The time delay used for the delay vectors.
x_i	An element of the time series.
f_i	An element of the approximation of the time series.
e_i	A residual or error element: $x_i - f_i$.
h	The time step used in numerical integration.
ϵ	The critical radius.
L_p	The L_p metric space.
d_p	The distance between two points using the L_p norm.
$d(x, y)$	The distance from x to y.
$RK4$	The 4^th order Runge Kutta numerical integration method.
GWN	Gaussian white noise.
UWN	Uniform white noise.
HFN	High frequency noise.
SNR	The signal to noise ratio.
FFT	The Fast Fourier Transform.
$SUNR$	The Shadow-UPO Noise Reduction method.
UPO	Unstable periodic orbit.
$R(i, j)$	The recurrence matrix.
Θ	The Heaviside function.
$H(i)$	The recurrence-histogram.
λ_{max}	The maximal Lyapunov exponent.
μ_x	The mean of the time series x.
σ^2	The variance of the time series x.
σ_{xy}	The cross-covariance between two time series x and y.
r_{xy}	The cross-correlation between two variables x and y.
MAE	The mean absolute error.
$RMSE$	The root-mean-square error.
MME	The mean maximum error.
R^2	The coefficient of determination.

Contents

Preface

Noise reduction of chaotic time series is inherently problematic in three main ways. Firstly, directly applying conventional noise-reduction filters to a noise-infected chaotic time series may destructively interfere with the system, effectively corrupting the underlying data. Chaotic signals are intimately bound with the noise in the system through the nonlinearities, and conventional noise reduction methods that are recursive in nature tend to corrupt the underlying signal in the process of reducing noise. Incorrect filtering will actually increase the correlation dimension calculated by standard algorithms and will add an apparent extra Lyapunov exponent to the system being studied. We are severely limited in the selection of conventional filters, to only finite impulse response filters (FIR), which fortunately includes the moving average type. The problem is exacerbated in part because the Fourier frequency power spectrum of a chaotic signal appears similar to broadband noise, so techniques based on frequency separation strategies also run into trouble.

Secondly, existing noise reduction techniques for chaotic systems usually rely on initially embedding the noise-infected time series in phase space. Embedding noise-infected data in phase space may result in a devalued system that is not topologically equivalent to the underlying noise-free system. Takens' theorem requires an infinitely long noise-free time series to guarantee topological equivalence between the true chaotic attractor and that derived from the phase space reconstruction. Algorithms to calculate the embedding parameters, the time delay, and embedding dimension, are also highly sensitive to noise. These restrictions reduce the robustness of the technique and limit the application to very low levels of added noise. Existing noise-reduction techniques are often highly effective with lower levels of measurement noise but generally do not perform well as the noise level becomes significant ($> 10\%$). There are few techniques to address dynamical noise in chaotic time series.

Thirdly, algorithms to directly calculate chaotic system invariants such as the maximal Lyapunov exponent from time series assume noise-free data and generally

perform poorly in the presence of noise. It is preferable to have a low residual noise time series approximant to use directly in these algorithms, rather than apply the algorithms to noisy data.

These three problems have significant consequences in the analysis of time series data collected from experimentation. In practice, measurement noise and dynamical noise are commonly present in experimental data. Sources of measurement noise include finite precision measurements, truncation errors, and missing data (both temporal and spatial). Dynamical noise arises from a feedback process wherein a system is perturbed by a small random amount at each time step (noise is added during the evolution of the system). An example of an experimental system with both types of noise present is that of a chaotic laser; currently available processing techniques of such measured time series data would greatly benefit from improved noise reduction techniques. It is important to reduce the noise as much as possible from the data in order to properly classify and quantify the underlying chaotic dynamical system.

In this thesis, a new method is presented to reduce noise from a chaotic time series. The shadow-unstable periodic orbit noise reduction method (SUNR method) approximates a noise-free low-dimensional continuous chaotic system (flow), using a single sample or multi-sampled scalar time series containing high levels of measurement noise or low levels of dynamical noise. The SUNR method does not require the prior embedding of data and operates directly on the sampled time series, thus avoiding the limitations of Takens theorem and the estimation of embedding parameters when significant levels of noise are present. The method aims to overcome the limitations described above, by focusing on nearly periodic orbit segments ('shadow-UPOs') shadowing the dense set of unstable periodic orbits that form the skeleton of a chaotic system, many of which are detectable using recurrence techniques. In essence, the noise-infected time series is deconstructed into a set of shadow-UPOs, each of which is locally amenable to linear filtering techniques. The shadow-UPOs are filtered and subsequently replaced back into the time series, resulting in an improved estimate of the noise-free time series.

The principal innovation presented in this thesis is that of recognising that a noise-infected time series may be approximated by a set of detected shadow-UPOs – even in the presence of high levels of noise. It is known that noise-free shadow-UPOs may be detected using recurrence methods. It is also known that recurrence histograms are somewhat robust in the presence of noise. We have utilised these two observations, to establish a shadow-UPO approximation of a time-series, specifically designed for data infected with significant levels of noise. We further improve the approximation by using detected partial (incomplete) shadow-UPOs that contain valuable dynamical information.The second innovation is to recognise that the individual detected noise-infected shadow-UPOs are amenable to targeted conventional linear filtering techiques, depending on the type of noise (if known). We can thus deconstruct a noisy data set into cycles, apply specialist noise filters, and subsequently reconstruct the vastly noise-reduced time series. The third innovation is to apply averaging to the bins of detected of shadow-UPOs, providing noise-reduced estimates of the individual lower order UPOs. We further are able to estimate individual maximal Lyapunov exponents for each basis UPO using the detected noise-infected shadow-UPOs. Fourthly, we demonstrate the successful application of the SUNR method to a variety of model chaotic systems. We identified limitations of the recurrence method of detecting cycles when dealing with higher instability systems, and successfully modified the SUNR method for these.

The method is illustrated in detail as a case study of the Rossler system, tested for various types of noise (uniform white, Gaussian white, high-frequency, coloured and dynamical) and also on several chaotic systems with a range of differing topologies (Chua, Rabinovich-Fabrikant, Lu-Chen, Lorenz). Goodness of fit metrics are defined, measured for each system and presented. We did not test the method on experimental data as this will be a comprehensive exercise, and it is planned for future work.

Chapter 1

Theoretical Foundations

1.1 Chapter Overview

In this chapter we start by providing a brief introduction to the theory of chaotic dynamical systems. The noise reduction methodology presented later in Chapters 4 and 5 requires concepts from many areas of chaos theory. Our intention is not to provide an exhaustive review but rather highlight and explain essential concepts that will be called upon later in this thesis. We firstly introduce dynamical system concepts including flows, nonlinearity and phase space. We present several definitions of chaos, emphasising the intimate inextricable linkage to unstable periodic orbits (UPOs). The Shadowing Lemma underpins the use of computed system models with added noise and our approximation using shadow-cycle sequences extracted from the time series. Finally we discuss UPOs, which are fundamental building blocks of chaotic attractors and the objects through which we facilitate noise reduction. We present some key results of the theoretical framework (periodic orbit theory), discuss briefly the detection of UPOs and review several papers that provided motivation for this research.

1.2 Chaotic Dynamical Systems

The invention of smaller, faster, personal computers in the second half of the twentieth century rapidly accelerated progress in understanding dynamical systems

defined by nonlinear equations of motion. These developments have enlightened the scientific community and profoundly altered our view and understanding of the physical universe. It has been said that alongside the development of general relativity and quantum mechanics, the discovery of chaos was the third great pillar in twentieth century physical science [59]. It has also made us revise some long held views, recognising that deterministic equations are not a guarantee of quantitative predictability and that in fact deterministic systems may behave as if they are stochastic.

Describing chaotic dynamical systems is relatively easy, whereas precisely defining chaotic dynamical systems is a much more complex matter. For this reason, we will loosely describe these systems here and provide more rigorous definitions later.

The term *chaos* was first used in the mathematical sense by Li and Yorke in 1975 [92] in their paper "period three implies chaos", defining chaos in interval maps. Mathematical chaos is the term used to describe the apparently complex behaviour exhibited by simple, orderly, deterministic systems. Chaotic behaviour appears erratic and random, somewhat like a system strongly perturbed by external random "noise" or the sophisticated dynamics of a system with many degrees of freedom. In reality, chaos can manifest itself in remarkably simple dynamical systems which may be noise free and have only a few degrees of freedom and these systems are in fact deterministic. In deterministic systems, provided we have exact knowledge of the initial state at a specific time, it is theoretically possible to exactly predict the future states of that system. Underlying the random-looking chaotic dynamics is a set of defining mathematical equations, which may be known or unknown. The objective of modern chaos theory is to reconcile this apparent dichotomy between randomness and determinism. These systems lie somewhere in the continuum between predictable regular periodic or quasi-periodic behaviour and unpredictable, stochastic behaviour. This was summarised nicely by Ian Stewart [136] as follows " Chaos is apparent stochastic behaviour occurring in a deterministic system".

Chaotic systems are defined by their remarkable *sensitive dependence on initial conditions* (SDIC), and the key factor of this feature is nonlinearity in the defining system equations. For such nonlinear systems, a small perturbation in a system parameter or initial conditions can lead to swift and radical changes in both the qualitative and quantitative evolution of the system. In practice, the sensitivity to even minute perturbations in initial conditions is so significant that prediction of long-term behaviour is impossible unless the system and its parameters are known with infinite precision. The presence of chaos in a dynamical system means that long-term predictions based on historic data are futile. This sensitivity to initial conditions was nicely encapsulated in the following description of chaos by Edward Lorenz:

> "Chaos: When the present determines the future, but the approximate present does not approximately determine the future".

Let us now explore some of the core concepts referred to above.

1.3 Dynamical Systems

A *dynamical system* comprises a *phase space* and an *evolution rule*. The phase space coordinates completely describe the state of the dynamical system at any point in time. The evolution rule specifies the future values of all *state variables*, given their current values. A "system" is a collective or aggregation of interacting parts. Several examples of dynamical systems with mathematical models include equations that describe the pattern of a heartbeat, a share market index, the trajectory of a comet, and the size of a population of rabbits in the wild. At any given time a dynamical system has a state given by a set of real numbers (a vector) that can be represented by a point in an appropriate phase space (a geometrical manifold). The evolution rule for a dynamical system is a function that outputs the future states of the system, given the current state. For deterministic dynamical systems, the evolution rule will yield a unique future state for any specified time period. Alternatively, there are stochastic systems where random events also affect the evolution of the state variables. Some evolution rules are defined in terms

of distance ("spatial chaos") rather than time ("temporal chaos"), however our discussion will relate to temporal chaos.

Dynamical system evolution is generally described using two types of transformation rules. Firstly, for continuous-time systems (*flows*) the system equations specify the time derivatives of the state variables in terms of their current (and possible past) values. The state variables are real numbers that vary continuously in time. The equation describing the motion of a simple pendulum is an example of a continuous-time dynamical system. In contrast, for discrete-time systems (*maps*) the system evolution is described recursively, with the future values of the state variables expressed as functions of the current (and possibly past) values. All physical systems are continuous-time systems at their most fundamental level. However in practice, it is often useful to describe the system configurations in discrete time and consider the system as it jumps from one configuration to the next. Experimental sampling of position, amplitude or intensity data are examples of systems modelled in discrete time.

Mathematically, dynamical systems are described by differential equations in continuous-time and by difference equations in discrete-time. Static systems are described by algebraic equations.

A deterministic dynamical system is defined by both a set of state variables (or a state vector $x(t)$) that describes the state of a system at point in time t and a dynamical law that specifies the temporal evolution of the corresponding state variables. Nonlinear dynamical systems are usually defined by systems of first order ordinary differential equations (ODEs). Many physical systems yield higher order systems of ODEs, but it is relatively simple to reformulate them into equivalent first order systems. We therefore do not lose any generality by confining our attention to the systems of first order ODEs. Also, numerical methods designed for higher order initial value problems usually require reformulation into first order systems as a first step.

An n^{th} order continuous-time system may be defined by the following differential equation:

$$x'(t) = f(x(t), p, t). \tag{1.1}$$

The system variables are components of the vector function x and $x(t)\,\epsilon\,\mathbb{R}^n$ is the state at time t. The vector x' refers to first time derivative of x. Also the vector $f : \mathbb{R}^r \to \mathbb{R}^n$, represents a smooth function of all of the system variables, at fixed values of the k parameters p. The *vector field* f is said to generate a flow:

$$\phi : \mathbb{R}^r \to \mathbb{R}^n, \tag{1.2}$$

where $r = n + k + 1$.

Given an initial condition, $x_0\,\epsilon\,\mathbb{R}^n$ and a time t_0, an *orbit* or *trajectory* (solution of Equation 1.1) passing through (or based at) x_0 at time t_0 is denoted as $\phi_t(x_0, t_0)$, where $\phi_t(x, t)$ is a smooth function satisfying:

$$\frac{d}{dt}\phi(x,t)|_{t=\tau} = f(\phi(x,t), p, t), \tag{1.3a}$$

$$\phi_{t_1+t_2} = \phi_{t_1} \circ \phi_{t_2}, \tag{1.3b}$$

$$\phi(x, t_0) = x_0. \tag{1.3c}$$

Note that f can depend explicitly on time, as would be the case for driven pendulum for example.

Because the time is explicit in Equation (1.1), f is said to be non-autonomous. An *autonomous differential equation* is one where the independent variable does not appear explicitly in the defining equations. If the independent variable is time, they are called *time-invariant* systems. In many physical laws, the independent variable is time. These laws are often formulated as autonomous systems because it is reasonably assumed that the laws of nature are constant in time. There is a well established canon of techniques to solve autonomous systems and non-autonomous

systems can be reformulated as an equivalent autonomous system by including a single additional ODE to the set of equations.

An n^{th} - order discrete-time system can be described by a difference equation of the form:

$$x(t+1) = F(x(t), p), \tag{1.4}$$

where the function $F : \mathbb{R}^{n+k} \to \mathbb{R}^n$ is recursive, outputting an updated x at the next time step, instead of a derivative as is the case with flows. The function F is often referred to as a map that takes the system from one time step to the next. A trajectory or orbit of a discrete system is a set of points $\{x(t+1)\}_{t=0}^{\infty}$ in $\mathbb{R}^n$.

The definitions for discrete systems are analogous to the ones described for continuous-time systems and therefore will be omitted. Note however, that the mapping function F might not necessarily be a bijection and thus the reverse mapping of x_{t+1} into x_t is not always uniquely defined. The implication is that state information is partially lost as we iterate forward in time, with future outcomes being unpredictable despite the system having a deterministic formulation.

Chaos may occur in nonlinear continuous-time systems with three or more degrees of freedom. Degrees of freedom in systems characterised by ODEs refers to the number of required first order autonomous ordinary differential equations. The Poincare-Bendixson Theorem mandates that for differential equations in two dimensions (i.e. the plane) only fixed points (time independent solutions) or limit cycles (periodic orbits) are possible, and thus three dimensions is the minimum required for chaos. In contrast, for discrete-time systems described by invertible maps, chaos only requires two or more degrees of freedom; noninvertible maps in one dimension can exhibit chaos. For discrete-time systems the degrees of freedom are the same as the number of components of the state vector $x(t)$. Linear systems of finite dimension can never exhibit chaos. To exhibit chaos, a dynamical system must be either nonlinear or infinite-dimensional.

A great deal of research involving the extraction of *unstable periodic orbits* (UPOs) from chaotic time series has been conducted for discrete time systems, as maps are generally much simpler to work with, and are more well behaved, than systems of differential equations representing flows. Also flows can be reduced to lower dimensional mappings using carefully oriented *Poincare sections*. Differential equations are used much more widely in science and engineering, and we shall thus focus on continuous time systems of equations and their associated flows, with the goal of designing a noise reduction method with real world applications.

1.4 Nonlinearity

The remarkable properties of chaos in a deterministic dynamical system are the result of at least one nonlinear term within the defining system equations.

Most nonlinear systems cannot be solved using analytical methods. In contrast, linear systems admit a superposition principle. They can be deconstructed into multiple components, each of which can be solved individually and then recombined into a solution. This principle often facilitates the simplification of highly complex problems and underpins many mathematical techniques like Fourier analysis. A linear system is simply the sum of its parts. In contrast, nonlinear systems cannot be deconstructed into manageable components and solved separately. They have to be analysed directly regardless of their complexity; and this is interesting and challenging.

The failure of the linear superposition principle for nonlinear ODEs, means that many of the "bread and butter" mathematical techniques (such as Laplace transforms and Fourier analysis) for solving linear ODEs have no utility for solving nonlinear ODEs.

Nonlinearity quite commonly arises in even the simplest systems with interactions between the components leading to emergent phenomena as chaos, solitons, fractals, and meta/multi-stability. The following quotation by Stanislaw Ulam nicely highlights this diversity:

> "Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals."

1.5 Phase Space

Dynamical systems theory (including chaos theory) seeks to explain the long-term qualitative behaviour of dynamical systems. The inclusion of nonlinearities renders systems of differential equations extremely difficult, if not impossible to solve analytically. Instead of aiming for exact solutions, the focus is generally shifted towards determining whether or not the system will settle into a steady state, finding the states that are possible and characterising how the system evolution varies with differing initial states. It also assists understanding to have a visual representation of the dynamical system.

The *phase space* (or state space) of a dynamical system is the set of all possible states that the system may take. Each possible state of the dynamical system is denoted by a unique point in the phase space. The minimum number of state variables to fully represent the system is its dimensionality. The state of a system at time t consists of all information needed to uniquely determine the future system states for times beyond t. A phase space plot is a powerful visual tool which may assist in holistic observation of the system and in developing mathematical strategies to further analyse the system. If we view the chaotic system through the lens of conventional linear techniques such as Fourier transforms, chaos may simply appear as broadband noise. In contrast, the structure of an attractor in phase space has remarkable fractal geometry and structure.

In a phase space, each system component is represented by an axis in a multidimensional space. A point is embedded in the phase space for each possible state of the system. The evolving system traces out a path through the high-dimensional space, referred to as an orbit. In aggregate, the phase space represents all states the system can assume. The geometrical object traced out by the evolving system may illuminate qualities and properties of the system that would otherwise go undetected. Although most commonly studied model chaotic systems are low-

dimensional (and very few experimental "hyperchaotic" behaviours have actually been detected), the phase space in general may contain a large number of dimensions. Figure 1.1 [85] shows the phase space for a well known chemical reaction (the Belousov-Zhabotinsky or B-Z reaction).

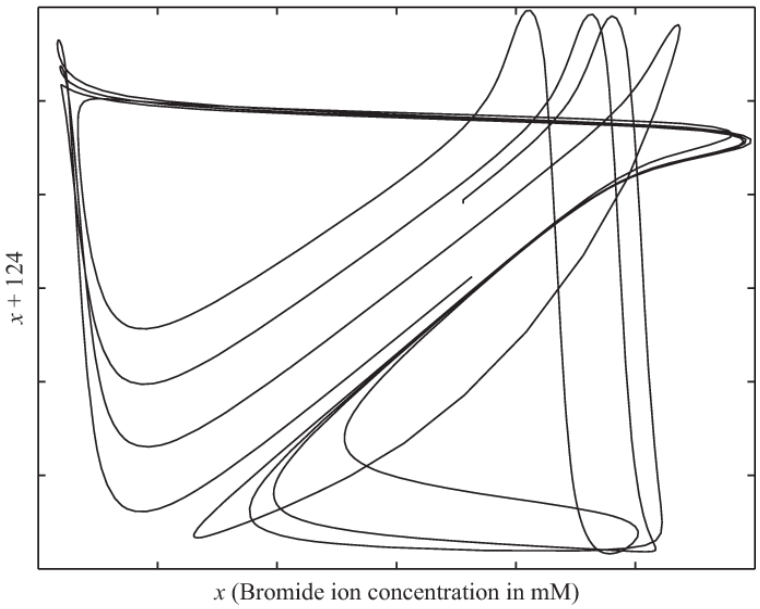


Figure 1.1: A "phase space" reconstruction in two dimensions of the chaotic dynamics of the bromide ion concentration in the classic Belousov-Zhabotinsky chemical reaction. The plot represents the evolution of the bromide ion concentration. The limit set is referred to as an *attractor*. The x axis is the measured bromide ion concentration, and the y axis is the same "x" variable lagged or delayed by 124 points.

1.6 The Definition of Chaos

Having described the key features of chaotic dynamical systems, let us now review a more rigorous definition of chaos. Whilst there is no universal mathematical definition of chaos, the most widely utilised definition of chaos is due to R. Devaney [38]. It nicely captures the essential features of chaotic dynamics and is applicable to both continuous systems and discrete iterated mappings. It provides an insightful perspective on the anatomy of chaos and in particular includes the important presence of a "skeleton" of unstable periodic orbits that lie within the chaotic attractor of the system.

Definition 1 *Chaos (Devaney)*
Let X be a compact metric space. A continuous function $f : X \to X$ is said to be chaotic on X if f has the following three properties:

1. *f is topologically transitive;*
2. *The periodic points of f are dense in X; and*
3. *f has a sensitive dependence on initial conditions.*

Let us briefly explain each condition in turn.

Definition 2 *Transitivity*
f is topologically transitive if for all non-empty open sets U and V of X, there exists a natural number k such that $f^k(U) \cap V$ is non-empty.

The implication of this condition is that in the fullness of time, orbits originating from points contained in U will densely fill the phase space X. *Topological transitivity*, also referred to as *topological mixing* means that eventually any given region in its phase space will intersect with any other given region. This mathematical concept of "mixing" aligns with the everyday meaning, like the mixing of coloured dyes for example. The outcome of this concept is that clusters of initially close points do not remain localised and are expelled into larger sets.

Definition 3 *Sensitive Dependence on Initial Conditions (SDIC)*
A continuous map f of a compact metric space X into X is said to possess sensitive dependence on initial conditions if there exists a positive ε such that for all $x \epsilon X$ and all $\delta > 0$ there is some y which is within a distance δ of x and for some k, $d(f^k(x), f^k(y)) > \varepsilon$. The number ε represents the maximum departure of orbits and is called the sensitivity constant of f.

In practice, it is extremely difficult to accurately specify the initial conditions of a dynamical system. This is a problem when coupled with sensitive dependence on initial conditions (SDIC) as the system will be unpredictable, and determinism will fail. In such systems we cannot approximate a given point with infinitesimally close neighbours as all nearby points have completely different future trajectories. The tiniest fluctuation in the current orbit may lead to wildly different future outcomes.

Sensitivity to initial conditions is commonly referred to as the "butterfly effect", an expression that emerged from the title of the paper by Edward Lorenz [94] entitled "Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?". This famous real world analogy has the flapping butterfly wings causing a small change in the initial conditions of the weather system. This subsequently initiates a chain of events ultimately resulting in a vastly different system state than if the butterfly had not flapped its wings. We can also interpret this phenomenon in terms of loss of information. A dynamical system initially contains a finite amount of information, but as a consequence of the SDIC eventually becomes unpredictable, meaning information has been lost. This is the case with weather modelling where forecasts are usually only possible for about 1 week ahead.

Definition 4 *Dense Periodic Points*
Let X be an interval in $\mathbb{R}$. Then a set A is said to be dense in X if for any $x \epsilon X$, any open interval containing x must intersect A. In other words, for each $\delta > 0$, the open interval $J = (x - \delta, x + \delta)$ contains a point of A. More generally, a set A is dense on a metric space X if for any $x \epsilon X$, any open set containing x must intersect A (contain a point of A).

The requirement that periodic points must be dense means that periodic orbits approach every point in the space arbitrarily closely. If we invert this concept, it means that every point on the chaotic attractor is arbitrarily close to some point on a periodic orbit. It is this central idea that we leverage in this thesis to develop approximations to the noise-free chaotic attractor using detected noise-infected periodic orbits.

It is a remarkable characteristic of chaotic systems that although most points are non-periodic (not on a cycle), we cannot distinguish these points visually from periodic points because there is always a periodic point arbitrarily close.

Let us briefly consider the strengths and weaknesses of the Devaney definition. The definition is precise and compact. However it does not incorporate the stretching and folding mechanism that is necessary for chaotic dynamics. As a result of this, it has been suggested that perhaps the Devaney definition is a consequence of chaos rather than a condition of chaos. Also, the definition emphasises periodic orbits rather than aperiodic, when it is the absence of periodicity that is the hallmark of chaos. However, as we shall see, the dense set of unstable periodic orbits lies at the very heart of chaos.

There is a surprising counter-intuitive consequence of the Devaney definition. The SDIC condition is mathematically redundant if the set X has an infinite number of elements. One would hardly imagine that the SDIC condition could be redundant [8].

Theorem 5 *(Banks et al. [8])*

Let X be an infinite metric space and f a continuous function on X. If f is topologically transitive and has dense periodic orbits, then f has sensitive dependence on initial conditions. (i.e. transitivity + density of periodic points $\Rightarrow$ sensitivity to initial conditions).

Transitivity and density are topological properties whereas sensitivity depends on a metric. This elegant result makes it clear that chaos is a property relying on the topological, and not the metric properties of a space. Also for

completeness, Assaf and Gadbois have proven that this is the only redundancy.

Theorem 6 *(Assaf and Gadbois [81])*

1. *Transitivity and sensitivity do not imply density of periodic points.*
2. *Density of periodic points and sensitivity do not imply transitivity.*

If we restrict our attention to intervals, M. Vellekoop and R. Berglund have shown that an even stronger result can be obtained [142]: on an interval, transitivity implies chaos.

P. Touhey [140] proposed a new and natural definition of chaos, equivalent to Devaney's. He combines the topological conditions of transitivity and density of periodic points into a single condition that yields a simple, concise, intuitive definition:

Definition 7 *(Touhey's Definition of Chaos)*
Let X be a metric space. A continuous function f from X to itself is said to be "Chaotic" on X if given any U and V, non-empty open sets in X, there exists a periodic point $p \epsilon U$ and a non-negative integer k such that $f^k(p) \epsilon V$.

With this definition any pair of non-empty open subsets of X shares a periodic point. This leads to the equivalencies of the following four statements; expressed in terms of periodic points of the function f:

1. f is "Chaotic",
2. f is (Devaney) Chaotic,
3. Any finite collection of non-empty open subsets of X shares a periodic point,
4. Any finite collection of non-empty open subsets of X shares infinitely many periodic orbits.

1.7 The Shadowing Lemma

The extreme sensitivity to initial conditions (SDIC) has serious implications for the construction of a chaotic dynamical system by a computer using an iterative numerical algorithm and an initial condition. A computed orbit will diverge exponentially from the true underlying orbit because of approximations and truncation errors inherent in the numerical and computing process. In a very short time the computer generated orbit and the true orbit will be de-correlated. This phenomenon brings into question the relationship between the actual physical behaviour of a chaotic system and any computed numerical model.

Fortunately the shadowing phenomenon (as described by Bowen's Shadowing Lemma) guarantees the existence of a true orbit that "shadows" the numerically generated orbit (called the *pseudo-orbit*) for relatively long times. By "shadowing" we mean it remains very close. Interestingly, this true *shadow-orbit* will be an orbit with a different initial condition than the numerical orbit. The end result is that the numerical model output does in fact "look like" a true orbit of the underlying chaotic system after all.

Definition 8 *Hyperbolic Set*
A subset Λ of a smooth manifold M is said to have a hyperbolic structure with respect to a smooth flow f if its tangent bundle may be split into two invariant sub-bundles, one of which is contracting and the other is expanding under f, with respect to some Riemannian metric on M. Such a subset is referred to as a hyperbolic set. In the case where the entire manifold M is hyperbolic, the flow f is called an Anosov diffeomorphism (more details appear in section 1.10).

Lemma 9 *Shadowing Lemma (Bowen [12])*
Given a map $f : X \to X$ of a metric space (X, d) to itself, define a $\varepsilon-$orbit as a sequence (x_n) of points such that x_{n+1} belongs to a $\varepsilon-$neighbourhood of $f(x_n)$, for each n. Then, near a hyperbolic invariant set, the following statement holds. Let Λ be a hyperbolic invariant set of a diffeomorphism f. There exists a neighbourhood U of Λ with the following property: for any $\delta > 0$ there exists $\varepsilon > 0$, such that any (finite or infinite) $\varepsilon-$orbit that stays in U also stays in a $\delta-$neighbourhood

of some true orbit. For all (x_n) *with* $x_n \epsilon U$*, with* $d(x_{n+1}, f(x_n)) < \varepsilon$*, there exists* (y_n) *with* $y_{n+1} = f(y_n)$*, such that for all n,* $x_n \epsilon U_\delta(y_n)$ *(where* $U_\delta(y_n)$ *is the open ball of radius* δ *centred at* y_n*).*

The theory states that every numerically computed orbit with rounding errors at every step, stays uniformly close to some true orbit (with a different initial point). The computer generated orbit is "shadowed" by a true one.

Let us now consider the further implications of adding noise to the chaotic mix. Noise perturbs the paths of individual orbits in a chaotic system. Low levels of added noise are unlikely to alter the qualitative properties of purely deterministic orbits greatly. We can interpret the noise-infected orbit through the same lens we viewed the orbit with computer truncation errors. If there is a true orbit shadowing the noise-infected orbit, then we can have some confidence that the noise-infected orbit will retain the same statistical properties for length scales exceeding those of the noise. Note that the existence of a shadowing orbit alone does not guarantee the statistical properties are the same.

Anasov and Bowen [4], [12] proved the result stated in the lemma for the limited case of "everywhere hyperbolic" (Anasov) dynamical systems. Anasov dynamical systems require that every point of the dynamics can be projected onto manifolds where the motion is either exponentially expanding or contracting. Anasov systems have no homoclinic tangencies, meaning the stable and unstable manifolds are never parallel.

The Anasov-Bowen construction relies the absence of homoclinic tangencies. It was previously thought that Anasov systems were the most common types of chaotic systems. However, Lai [88] has shown the contrary—that most dynamical systems do in fact have homoclinic tangencies. An interesting result by Farmer et al. [17] shows that dynamical systems commonly exhibit a sensitive dependence on parameters. Arbitrarily close to a parameter value that defines a chaotic attractor, there is another parameter value that generates a stable periodic attractor. A periodic orbit that is stable may be nudged by noise into becoming chaotic [27].

In this situation the shadowing lemma is not true in general and there is no shadow-orbit.

1.8 The Method of Delays

Takens [137] has shown that, if we measure any single state variable with sufficient accuracy for a long period of time, it is possible to reconstruct the underlying dynamic structure of the entire system from the single variable using delay coordinates and an embedding procedure. Takens provided the first example of a mathematical proof for reconstructing a diffeomorphic shadow manifold using delays of a single time series. The central idea, that was earlier demonstrated by Packard, Crutchfield, Farmer, and Shaw [109] is that under generic conditions, a shadow manifold M' can be created using time-delayed observations based on a single measurement function that is a smoothly invertible mapping with one-to-one mapping to M. Although different in details such as the dimension of the reconstructed space, the work by both contributors is similar in spirit.

Suppose that we have a measured time series of a single scalar variable x, consisting of N points, $x(N) = (x(t_1), x(t_2), ..., x(t_N))$. The *time delay vectors* constructed from the the single variable $x(t)$ are:

$$y(t_i) = [x(t_i), x(t_i + \tau), ..., x(t_i + (m-1)\tau)],\ i = 1, ..., N - (m-1), \qquad (1.5)$$

where τ is the *time delay* (number of time steps), m is the *embedding dimension* (dimension of the reconstructed space). As a result, the time series of N points will generate $M = N - (m-1)$ vectors in the phase space. Let us now define the concepts of an *immersion* and an *embedding*:

Definition 10 *Immersion*
An immersion is a differentiable function between differentiable manifolds whose derivative is everywhere injective. Formally, suppose $f : N \to M$ is a smooth map between manifolds. The map f is called an immersion if $D_p f : T_p M \to T_{f(p)} N$, is

an injective function at every point p of M, where T_pM denotes the tangent space of a manifold M at a point p in M.

Definition 11 *Embedding*
Suppose $f : N \to M$ is a smooth map between manifolds. The map f is called an embedding if f is an immersion which is a homeomorphism to its image.

A formal expression of Takens' theorem follows.

Theorem 12 *(Takens [137])*
Let M be a compact manifold of dimension m. For pairs (ϕ, y), where $\Phi : M \to M$ is a smooth diffeomorphism (an invertible function that maps one differentiable manifold to another such that both the function and its inverse are smooth) and $y : M \to \mathbb{R}$ a smooth function (at least C^2), it is a generic property that the $(2m + 1)$ delay observation map $\Phi_{(\phi,y)}(x) : M \to \mathbb{R}$ given by

$$\Phi_{(\phi,y)} = (y(x), y \circ \phi(x), ..., y \circ \phi^{2m}(x)), \tag{1.6}$$

is an embedding.

1.9 Chaotic System Invariants

Experimental time series created from deterministic dynamical systems appear stochastic when analysed with standard linear modeling techniques. As an example, a time series constructed using the logistic map has the same autocorrelation function as white noise. It is therefore critically important to be able to distinguish chaotic systems from stochastic ones. With noise-free linear systems we may identify and characterise the signal using Fourier analysis. The locations of the sharp peaks in the Fourier power spectrum characterise the physical system analysed. If we increase power input to the system or start at a new time, the signal phase will change but the location of Fourier spectrum peaks will be conserved; they will not change. The characteristic frequencies or harmonics of the linear system are dynamical invariants and may be used to qualitatively describe the underlying physics.

In a nonlinear dynamical system we have several key statistical metrics that characterise the deterministic chaotic evolution. These invariants we refer to are Lyapunov exponents, correlation dimension and entropy; each measuring important features of the chaotic dynamics and geometry. These metrics which are invariant under the dynamics and topological transformations are used as classifiers. We shall confine our detailed discussion to Lyapunov exponents as this will be a point of focus during the research discussion.

1.9.1 Stretching and Folding Mechanism

A chaotic attractor is a unique geometrical object created by the asymptotic states of a chaotic system. The dynamics on the chaotic attractor are characterised by the stretching and folding mechanism underlying chaos at the most fundamental level. This is illustrated in Figure (1.2). The stretching aspect causes nearby orbits to diverge. The folding aspect enforces boundedness by confining the dynamics to a finite region in a subspace that is the minimal space embedding of the chaotic attractor.

The folding creates a unique characteristic of chaotic attractors in that they are characterised by a non-integer dimension (there are a multitude of non-integer dimension measures). In non-chaotic systems, attractors such as fixed points and limit cycles are characterised with integer dimension. The rate at which nearby orbits diverge (as a result of SDIC) and fractal dimension are both properties of the dynamical system that are independent of any specific orbit. The stretching and folding process de-correlates nearby states on the attractor. The result is that the long-term state that evolves from an initial condition containing any uncertainty is entirely unpredictable.

The fractal dimensions characterise the scaling properties of the probability distribution of the data in phase space (the invariant measure); entropy measures do the same for transition probabilities from one part of the phase space to another. For regular dynamics, points in a small ball will be mapped into another small ball with all transition probabilities equal to zero except one. In contrast for a pure

noise process, the future of a small ball of points is disconnected from the initial state and completely unknown. Chaotic systems reside somewhere between these two scenarios and entropies measure this loss of information about the state of a system as a result of time evolution. For both fractal dimensions and entropy measures definitions must include infinitely small scales and infinite times.

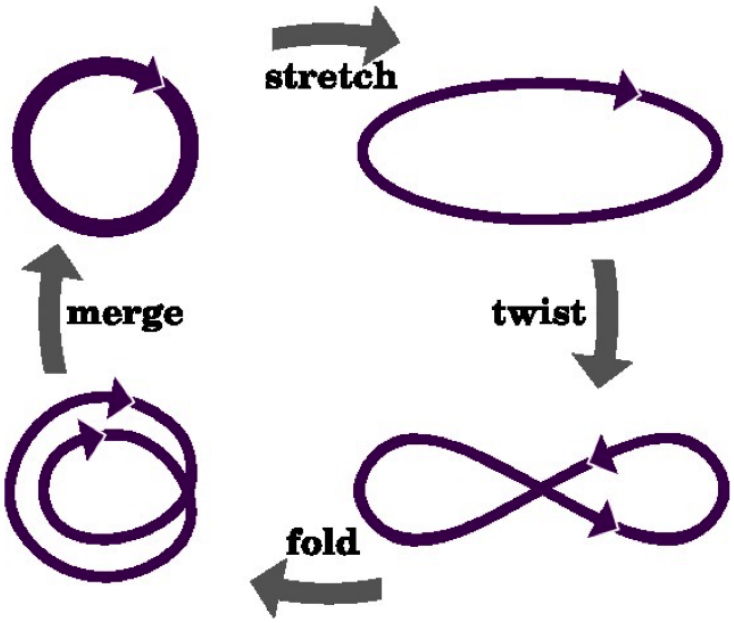


Figure 1.2: The stretching and folding mechanism that underpins chaos. The figure demonstrated on a unit cycle, producing a result that looks very much like the unstable periodic orbits we will later locate and extract from time series. The twist is not necessary for chaos.

1.9.2 Metric and Dynamical Invariants

Invariants can be grouped into two categories: those use for characterising attractors and those used for classifying attractors. Metric invariants (fractal dimension) and dynamical invariants (Lyapunov exponents) are restricted to characterising a given attractor.

Metric and dynamical invariants are real numbers that are invariant to coordinate transformations but not to changes in the system control parameter values. They are robust for a fixed experimental situation and are widely used. They are highly useful when one wants to identify quantities that remain unchanged when initial conditions on an orbit are changed or when perturbations are encountered during the orbit. There are a multitude of definitions for each of these measures of complexity as well as relationships between them, allowing us to compute one from the other.

1.9.3 Topological Invariants

Topological invariants refer to properties that are preserved under homeomorphisms. They are useful to classify different types of chaos. A *homeomorphism* is a bijective (one-to-one) mapping that is continuous and has a continuous inverse. A homeomorphic function preserves the topology of a set. The concept of entropy is an example of a topological invariant. The idea here is to uniquely classify the system dynamics (as represented by a finite time series sample) as a specific topological type. This is a little like the zoological classification of animals into species. These systems will remain topologically the same under system parameter changes. The development of this taxonomical model is still under development, and includes topological/geometric concepts such as linking numbers, relative rotation rates and template identification. There have been major contributions to this work by Mindlin and Gilmore [104] and Gilmore [57].

Interestingly the key topological concepts are formulated in terms of unstable periodic orbits (UPOs). Noise-reduction methods based around topological structure do not exist at this time (to the best of our knowledge) and are a potential area of future research. Locating, extracting and filtering UPOs from a noise-infected time series will provide useful inputs into this classification process.

UPOs are topological invariants that preserve the time ordering of data and exist in abundance in a strange attractor. The stretching and folding mechanisms that act to create a chaotic attractor also uniquely organise the UPOs embedded in

the chaotic attractor. The organisation of the UPOs within the chaotic attractor serves to identify the stretching and folding mechanisms that create the chaotic attractor. Periodic orbit theory is an area of study focused on determining system invariants using UPOs and was largely developed by Cvitanovic [28], [29].

1.9.4 Limitations of Estimated Invariants

There are several conceptual and practical challenges with all these invariant measures. There are however some limitations:

1. The underlying dynamical process is being observed through data collected using a measurement process, which itself may perturb the data. If the value of the measured quantity depends on the collection procedure then the information is devalued as a reliable quantifier.

2. Secondly, they work well when the time series is simulated from a model low-dimensional dynamical system, but most of them break down as soon as noise is added to the series.

3. Thirdly, they are mathematical concepts and are not defined in terms of time series of observations but rather in terms of infinitely finely sampled, infinite length data sets. The finite length and sampling frequency of the measured time series act to corrupt the invariant properties of the measured quantities.

One objective of this thesis is to determine the system invariants for noise-infected, chaotic systems. Many existing numerical algorithms to compute chaotic invariants require either prior knowledge of the defining system equations or a time series that is relatively noise free, as we shall see in Chapter 2. Both these requirements are highly restrictive in experimental situations. UPOs are topological invariants that can be extracted from the noise-infected time series, noise-filtered, and provide an excellent basis to calculate invariants.

1.9.5 Lyapunov Exponents

Lyapunov exponents are a measure of the average rate of divergence of neighbouring orbits (perturbed initial conditions) on a chaotic attractor, normally represented by the symbol λ. These exponents are frequently used to diagnose the presence of chaos and also may be used in the calculation of other invariant quantities such as the fractal dimension of a chaotic attractor. Two orbits on a chaotic attractor with nearby initial conditions will diverge exponentially at a rate specified by the largest Lyapunov exponent [93].

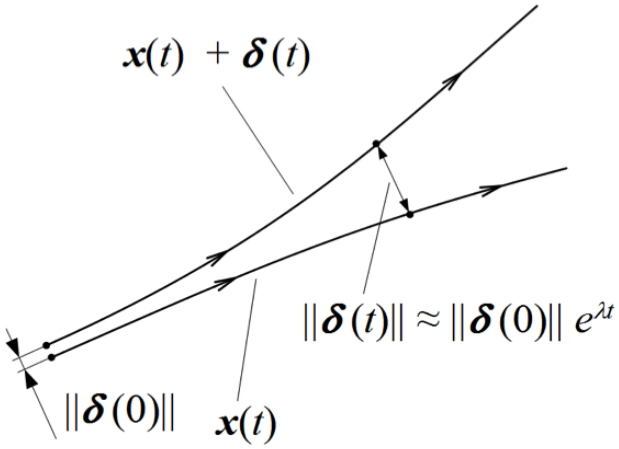


Figure 1.3: A Lyapunov exponent is derived by considering the divergence of two orbits that originate from nearby points.

Consider any two nearby orbits on a chaotic attractor:

$$x(t) = f^t(x_0), \tag{1.7}$$

$$x(t) + \delta x(t) = f^t(x_0 + \delta x_0), \tag{1.8}$$

that start out in close proximity and subsequently diverge exponentially with time. In some finite time they will be far apart, separated by the extent of the

accessible phase space. For small t we have

$$\|\delta x(t\| \approx e^{\lambda t} \|\delta x_0\|\,, \tag{1.9}$$

where λ, the mean rate of separation of orbits is the largest Lyapunov exponent (hereafter referred to as the *maximal Lyapunov exponent* or $\lambda_{\max}$).

The maximal Lyapunov exponent, calculated over a long time period, is an average global measure of the rate at which nearby orbits diverge. The signs of the individual Lyapunov exponents, particularly the positive ones, provide insight into the chaotic system's dynamics. The existence of at least one positive Lyapunov exponent indicates the presence of chaos, but is not definitive. Perron's counterexample [90], [87] shows that a negative largest Lyapunov exponent does not guarantee stability, and similarly that a positive largest Lyapunov exponent does not in general guarantee chaos. Thus Lyapunov exponents in isolation are indicative of chaos, but are insufficient to confirm chaos. The magnitude of the most positive Lyapunov exponent is also a measure of the time scale over which chaotic behaviour may be predicted. Similarly the magnitude of the most negative Lyapunov exponent indicates the time scale for transients to decay [49]. A dynamical system with more than one positive Lyapunov exponent is referred to as hyperchaotic [119].

1.9.5.1 Lyapunov Exponents for Flows

Let us now derive an expression for the Lyapunov exponents of a flow. Consider a ball of points in n-dimensional phase space and follow the orbit of each point in the ball as time evolves. On a short time scale, this initial group of points will either collapse to a single point, remain unchanged as a ball, or deform into an ellipsoid [58] approximately as shown in Figure 1.4. The rate of deformation of this small ball is the Lyapunov exponent. For the existence of a bounded attractor, we require the overall dynamics to be dissipative (globally stable). The contraction rate must exceed the expansion rate and the sum of all the Lyapunov exponents must be negative.

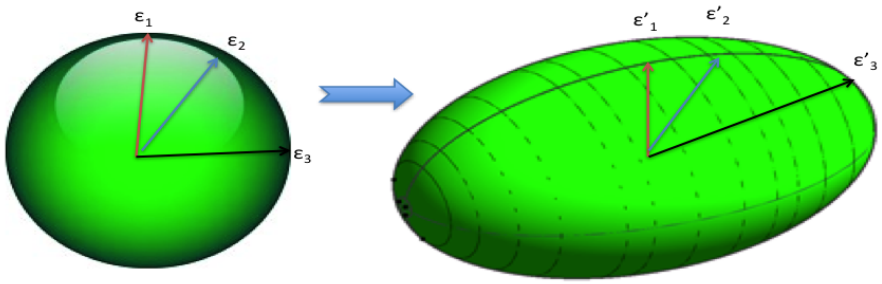


Figure 1.4: The time evolution of an initially small sphere of principal axis $\varepsilon = \varepsilon_1 = \varepsilon_2 = \varepsilon_3$. With increasing time, the initial rotationally symmetric region (ball) is gradually deformed into a ellipsoid.

Let us take a d-dimensional autonomous system of ordinary differential equations:

$$\frac{dx(t)}{dt} = F(x; p), \tag{1.10}$$

where x is a vector of d state variables and p is a vector of parameters. A solution ϕ_t for this system starting from some initial condition x_0 would yield an orbit $x(t)$. If a small initial perturbation $\epsilon(0) = \epsilon_0$ is applied to that orbit, the perturbed orbit becomes $\phi_t(x_0 + \epsilon_0)$. A Taylor series expansion gives

$$\phi_t(x_0 + \epsilon_0) = \phi_t(x_0) + J(t).\epsilon_0 + O(\epsilon_0^2), \tag{1.11}$$

where O denotes the order of the correction factor and $J(t)$ is the Jacobian matrix given by the linearisation of ϕ_t around the point x_0:

$$J(t) = (\frac{\partial \phi_t(x_0)}{\partial x_0}) = \frac{\partial x(t)}{\partial x(0)}, \tag{1.12}$$

$$J_{(i,j)} = \frac{\partial x_i(t)}{\partial x_j(t)}, \tag{1.13}$$

where $x_i(t)$ is the i^{th} component of the state vector x at time t.

Thus, after linearising Equation (1.10) about the perturbation, the following equation for the perturbation is obtained:

$$\frac{d\epsilon}{dt} = J.\epsilon(t). \tag{1.14}$$

An initial perturbation, $\epsilon(0) = \epsilon_0$, will evolve according to:

$$\epsilon(t) = \Phi(t)\epsilon(0), \tag{1.15}$$

where $\Phi(t)$ is the fundamental (transition) matrix solution of Equation (1.10) associated with the orbit $x(t)$.

For an appropriately chosen ϵ_0, the exponential rate of expansion or contraction in the direction of ϵ_0 on the orbit passing through x_0 defines the Lyapunov exponent along that direction:

$$\lambda_i = \lim_{t \to 0} \frac{1}{t} \ln \frac{\|\epsilon(t)\|}{\|\epsilon_0\|}, \tag{1.16}$$

where $\|.\|$ denotes the vector norm on $\mathbb{R}^d$.

The determinant of the Jacobian matrix describes the overall contraction of phase-space volume (i.e. the dissipation in the system) and the eigenvalues describe the divergence of nearby orbits. The Lyapunov exponents quantify the rate of expansion of these eigenvalues:

$$\lambda_n = \lim_{t \to 0} \frac{1}{t} \ln \left\| (n^{th}\, eigenvalue\, of\, J(t)) \right\|. \tag{1.17}$$

A pair of close initial points will diverge at a rate dominated by the largest Lyapunov exponent. This is unless the vector joining the respective starting points is precisely orthogonal to the eigenvector associated with the largest eigenvalue of the Jacobian.

The *maximal Lyapunov exponent* is most commonly calculated and discussed as it is easiest to numerically calculate from a time series. It also generally yields the greatest insight into the dynamics of the system. For experimental data, Lyapunov exponents other than the maximal exponent have not found a lot of utility so far.

We can interpret the d Lyapunov exponents that comprise the Lyapunov spectrum as rates of expansion or contraction in the direction of the principal axes of the infinitesimal ball shown in Figure 1.4. Wolf et al. [145] provide the following geometrical interpretation. Since the phase has d dimensions, a set of d linearly independent orthonormal basis vectors, $y_1, y_2, ..., y_d$ can be defined for this space. Applying a small initial perturbation, $\epsilon_1, \epsilon_2, ..., \epsilon_d$, along each of these directions, d Lyapunov exponents may be also defined, λ_i, by Equation (1.16) . By arranging the d principal axes of the ellipsoid in order from the most rapidly expanding to most rapidly contracting, we may arrange the corresponding Lyapunov exponents as follows:

$$\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_d. \tag{1.18}$$

This is the *Lyapunov exponent spectrum*. For small times t, the length of the first principal axis is proportional to $e^{\lambda_1 t}$. Similarly, the area defined by the

first two principal axes is proportional to $e^{(\lambda_1+\lambda_2)t}$. More generally, the volume determined by the first k principal axes is proportional to $e^{(\lambda_1+\lambda_2+\lambda_3+...+\lambda_k)t}$. The Lyapunov spectrum can thus be defined so that the exponential growth of a k-volume element is given by the sum of the k largest Lyapunov exponents. We can also interpret the volume change created by the expanding principal axes as information created by the system. The *Kolmogorov entropy* (K), representing the average rate of information gain is equal to the total of the positive Lyapunov exponents.

For all orbits $x(t)$, with the exception of a fixed point, one of the λ_i will be zero, meaning that perturbations occurring along the orbit neither diverge nor converge. For a stable limit cycle, all other λ_i will be negative, indicating that perturbations made orthogonal to that orbit will decay back onto it. If $\sum \lambda_i < 0$ then the system is *dissipative*, meaning that volumes in the phase space will contract overall. If a system is dissipative and also has at least one positive Lyapunov exponent, then that system is chaotic.

1.9.5.2 Lyapunov Exponents for Mappings

As already indicated, many physical systems can be modelled using finite-difference equations or "mappings" of the form:

$$x_{k+1} = f(x_k, \mu), \tag{1.19}$$

where x is a (single or multi-dimensional) vector of state variables, μ is a vector of parameters, and k is an index indicating the iteration of the map. For example mathematical biologists, often model population dynamics using discrete maps where k counts successive generations of species.

Many mathematical techniques for continuous chaotic systems (flows) involve the application of a Poincare section of the flow, resulting in mapping. Experimental data that are digitally sampled or model chaotic systems obtained from numerical integration using some finite time interval, Δt, can also be interpreted as a map. This type of finite-time mapping is employed to numerically estimate

Lyapunov exponents from time series data. The difference between data deriving from a map and data deriving from a flow is somewhat arbitrary. However data deriving from a continuous flow will always exhibit at least one zero Lyapunov exponent, for orbits perturbed along the flow neither diverge nor converge.

To define the Lyapunov exponent for discrete-time systems, consider how a one-dimensional map:

$$x_{k+1} = f(x_k, \mu), \tag{1.20}$$

evolves when started from two initial states, x_0 and $x_0 + \epsilon_0$, that are initially very near to each other (i.e., ϵ_0 very small). The separation between these two points after n iterations of the map is defined as ϵ_n. If this separation evolves approximately as:

$$|\epsilon_n| \approx |\epsilon_0| \, e^{\lambda n}, \tag{1.21}$$

then λ is the Lyapunov exponent.

By taking logarithms of both sides we find a more precise expression for the divergence of the two resulting orbits:

$$\lambda \approx \frac{1}{n} ln \left| \frac{\epsilon_n}{\epsilon_0} \right| = \frac{1}{n} \ln \frac{|f^n(x_0 + \epsilon_0) - f^n(x_0)|}{\epsilon_0}. \tag{1.22}$$

We now take the limit of Equation (1.22) as $n \to \infty$. First, the remaining term inside the logarithm is expanded using the chain rule:

$$\frac{f^n(x_0 + \epsilon_0) - f^n(x_0)}{\epsilon_0} \approx (f^n)'(x_0) = \prod_{i=0}^{n-1} f'(x_i). \tag{1.23}$$

Substituting this expression back into Equation (1.22), one obtains:

$$\lambda \approx \frac{1}{n} \ln \left| \prod_{i=0}^{n-1} f'(x_i) \right| = \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)| \, . \tag{1.24}$$

Finally, the limit of this expression is taken as $n \to \infty$ to define the Lyapunov exponent for the orbit starting at x_0:

$$\lambda = \lim_{n\to\infty} \{\frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|\}. \tag{1.25}$$

For multi-dimensional mappings, this definition is extended to yield a spectrum of Lyapunov exponents, $(\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_d)$ as before. These Lyapunov exponents defined for discrete maps have the same interpretations and implications as those defined for continuous-time systems.

1.9.5.3 Global and Local Lyapunov Exponents

The *global Lyapunov exponents* measure the average expansion or contraction of infinitesimal perturbations to an orbit over a long time. If we have a very large time series of noise-free data, where the time evolving orbit traverses the entire volume of the attractor, we can calculate an estimate of this average. The global Lyapunov exponent is useful in diagnosing the presence of chaos in a system and possibly as an input to calculate other invariants. If we want to know more about orbital divergence, we need to define local measures. In reality the orbit evolves along a series of UPOs, sometimes completing entire circuits and other times transitioning to the vicinity of another UPO due to the instability in the system.

Lyapunov exponents will vary along the orbital path and it is this context that we suggest that *short time Lyapunov exponents* are a more meaningful measure of orbital divergence than averages. The definitions for local and short time Lyapunov exponents are as follows:

$$Global : \lambda_\infty = \lim_{n\to\infty} \frac{1}{n} \ln \frac{\|\delta x(t)\|}{\|\delta x(0)\|}, \tag{1.26}$$

$$Short\,Time : \lambda_T(x(t), \delta x(t)) = \frac{1}{T} \ln \frac{\|\delta x(t+T)\|}{\|\delta x(t)\|}, \tag{1.27}$$

$$Local : \lambda_{local} = \lim_{t\to 0} \frac{1}{T} \ln \frac{\|\delta x(t+T)\|}{\|\delta x(t)\|}. \tag{1.28}$$

These are simply Lyapunov exponents defined over a shorter time periods. The short time Lyapunov exponent is defined on a finite time interval, which is particularly useful in the context of UPOs. The *local Lyapunov exponent* is the limiting version of the short time Lyapunov exponent when the time interval tends to zero. It measures how fast infinitesimal perturbations to a orbit at a given point expand or contract after k time steps subsequent to the perturbation.

The short time Lyapunov exponents will usually vary materially with location on the attractor. This variation is exacerbated when k is small, meaning that predictability will vary greatly over the attractor. The local and short time local measures are both dependent on initial points. The short time Lyapunov exponent is also dependent on the length of the time interval. See [1] and [40], respectively, for further mathematical details.

These localised measures are a more practical way of presenting Lyapunov exponents as they articulate more information than a single number. We propose that a more meaningful representation of the divergence concept is in terms of Lyapunov exponents of the set of lower order UPOs. UPOs are dense on the attractor and any finite sampled (time series) orbit can be represented by an approximation consisting of lower order UPOs. UPOs can be viewed as a basis or skeleton for the chaotic attractor and each has a unique Lyapunov exponent. Representing the attractor as a set of lower order UPOs and the Lyapunov exponent each is a meaningful characterisation of the attractor. The concept has more merit than arbitrarily calculating divergence for an orbit segment. We know that the orbit will spend a significant amount of time near a small number of different types of

UPO, and for each of these we know the orbital divergence. We represent the Lyapunov signature as follows:

$$L_k = (\lambda_1, \lambda_2, ...\lambda_k), \tag{1.29}$$

where we have selected the first k UPOs, ordered by period length from least to greatest, to represent the system.

1.9.5.4 Units of Measurement for Lyapunov Exponents

The Lyapunov exponent is defined using the natural logarithm (base e) and it is measured in units of 1/time. When considering the information - theoretic context, the Lyapunov exponent may be expressed in base 2, where it equates to bits per second. It may then be interpreted as the rate at which information about the state of the system is being created or destroyed [108]. We shall use base e when referring to the Lyapunov exponents in this thesis.

1.10 Periodic Orbits in Chaotic Systems

We have seen that periodicity plays a crucial role in chaos theory. Periodic points are at the heart of definitions of chaos and the infinite set of UPOs comprise the skeleton of a chaotic attractor. To better understand the connections between cycles and chaos we need explore the theory of *Smooth dynamics*, which is the study of differentiable maps and flows.

Hyperbolic dynamical systems are a class of smooth dynamical systems where the differential yields strong local, and sometimes global information about the underlying dynamics. The hyperbolic behaviour is characterised by the expanding and contracting sub-manifolds at a point and is caused by the stretching and folding mechanism underpinning chaos. The tangent space at each point partitions into contracting (or stable) and expanding (or unstable) subspaces. These expanding and contracting sub-manifolds provide valuable insight and quantifiable information about individual orbits and the topology of the system. Importantly, the theory of hyperbolic dynamical systems provides a mathematical formalism for the study of chaos.

Definition 13 *Stable and Unstable Manifolds*

The stable and unstable manifolds $W^s(x_0)$ and $W^u(x_0)$ of a saddle equilibrium point x_0 are

$$W^s(x_0) := \{x \epsilon \mathbb{R}^n \mid \lim_{t\to\infty} \phi^t(x) = x_0, \tag{1.30}$$

$$W^u(x_0) := \{x \epsilon \mathbb{R}^n \mid \lim_{t\to\infty} \phi^{-t}(x) = x_0. \tag{1.31}$$

where ϕ^t is the flow defined in Equation 1.2.

Definition 14 *Hyperbolic Attractor*
A hyperbolic attractor requires the following two conditions to hold:

1. *There exist stable and unstable manifolds at each point of the attractor whose dimensions, n_s and, n_u, are the same for each point on the attractor, with $n_s + n_u = d$, where d is the dimension of the phase space.*

2. *There exists a constant $K > 1$ such that for all points, x, on the attractor, if a vector u is chosen tangent to the unstable manifold, then*

$$\|Df(x)u\| \geq K \|u\|, \tag{1.32}$$

and if u is chosen tangent to the stable manifold

$$\|Df(x)u\| \leq \|u\| / K, \tag{1.33}$$

where $Df(x)$ denotes the Jacobian matrix of the map f evaluated at the point x.

Whilst not all systems of interest are hyperbolic, hyperbolic systems remain important as they are more amenable mathematically. Much of the rigorous theory of dynamical systems is prefaced with the assumption that the system is hyperbolic. It is believed that much of the theory of hyperbolic systems translates also to non-hyperbolic systems, but there are few rigorous results.

1.10.1 Unstable Periodic Orbits

Unstable periodic orbits (UPOs) comprise the skeleton of, and are fundamental in the definition of, a chaotic attractor. Within the chaotic attractor there are densely embedded, an infinite number of UPOs of all possible periods.

We can view the time evolution of a chaotic orbit as traveling close to a UPO and eventually jumping into the vicinity of another UPO. Time spent in the vicinity of a UPO depends on the relative instability, noting that all these periodic orbits are unstable. Figure 1.5 shows various UPOs detected from the Lorenz system by Yamagita et al. in [147].

The set of UPOs is a dynamical invariant; their number, distribution, and properties are conserved under a change of coordinates. Various chaotic invariants such as Lyapunov exponents, fractal dimension measures and topological entropy may be calculated using UPOs. Note that these calculations usually require a

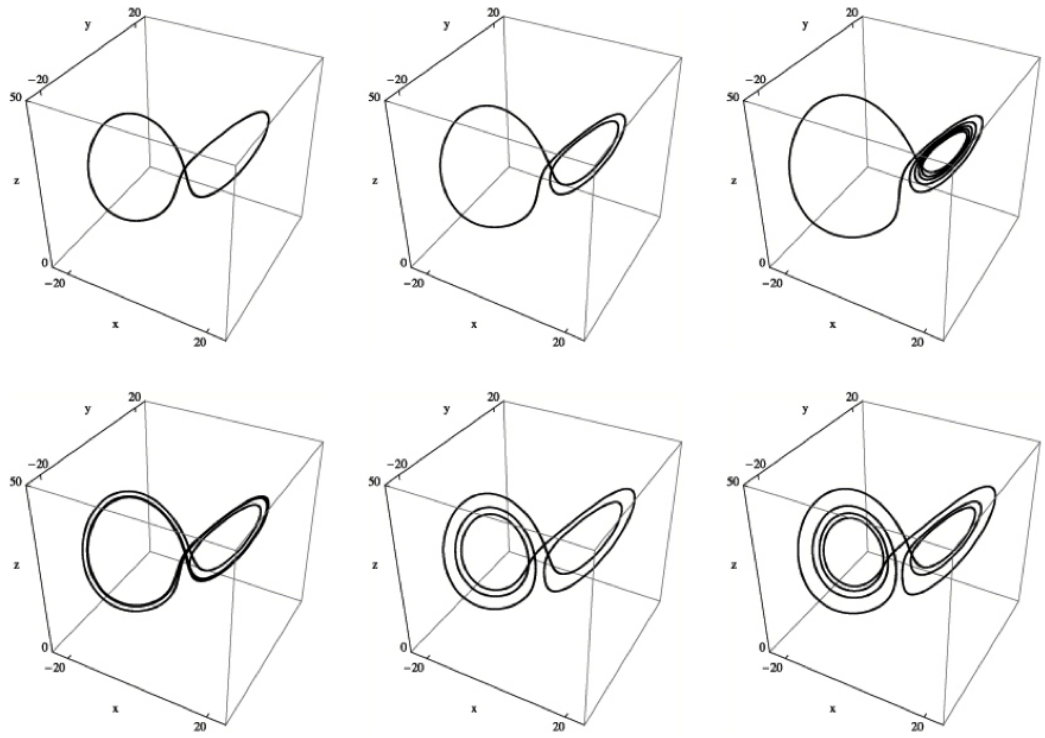


Figure 1.5: UPOs in Lorenz system

complete set of lower order UPOs and any calculation is subject to the errors created in the determination of each UPO.

The closure of the set of UPOs defines a chaotic attractor, and generic chaotic attractors can be approximated hierarchically using UPOs up to a given length. UPO periods can vary according to the topology of the system. For example the Rossler attractor has been shown to only admit UPOs that are multiples of a fundamental period

$$T_0 \cong 2\pi. \tag{1.34}$$

The Lorenz system does not have this integral subharmonic period structure.

UPOs provide a means to understand the rich dynamics and topology of a chaotic system. They also provide an excellent basis for developing approximations to the noise-free system, when our starting point is a noise-infected time series.

1.10.2 Periodic Orbit Theory

Periodic orbit theory provides a formal framework for the unstable periodic orbits of a system. This knowledge enables us to utilise the properties of individual solutions, such as their periods, location and stabilities, to estimate system quantifiers like Lyapunov exponents, fractal dimension and entropies. Periodic orbit theory expresses all long time averages of the chaotic dynamics in terms of cycle expansions [28], [29]. Sums over periodic orbits (cycles) are ordered hierarchically according to the orbit length.

In 1988, Cvitanovic [28] argued that for low dimension deterministic dynamical systems, UPOs provide a detailed invariant characterisation with the following virtues:

1. UPO symbol sequences are topological invariants – they provide the spatial layout of the chaotic attractor;

2. UPO eigenvalues are metric invariants – they provide the scaling of each section of the attractor;

3. UPOs are ordered hierarchically – lower order UPOs provide a good approximation to the attractor and errors due to neglecting longer period UPOs are bounded;

4. UPOs are robust – eigenvalues of lower order UPOs vary slowly with smooth parameter changes;

5. Lower order cycles can be accurately extracted from experimental data.

Cvitanovic predicted that future investigations of experimental strange attractors would use deterministic noise smoothing techniques. This never happened to any major degree, but is the focus of this thesis. He further demonstrated the UPOs and their associated eigenvalues provide a theoretically optimal measurement of the invariant properties of a dynamical system.

We now summarise some of the key results of periodic orbit theory. The following results require the assumptions that the attractor of f is both hyperbolic and mixing. By mixing, we mean that for any two subsets A_1, A_2 in the phase space, we have

$$\lim_{i\to\infty}\mu[A_1 \cap f^i(A_2)] = \mu(A_1)\mu(A_2), \tag{1.35}$$

where μ is the natural measure of the attractor (see Definition 1.2 of Transitivity). In other words the system will evolve over time so that any given open set in phase space will eventually overlap any other given region.

Denote the magnitudes of the eigenvalues of the Jacobian matrix for the p times iterated map f^p evaluated at the j^{th} fixed point by $\lambda_{1j}, \lambda_{2j}, ...\lambda_{nj}$. Suppose that the number of unstable eigenvalues ($\lambda_{ij} > 1$), is given by n_u, and further, that we order them as follows:

$$\lambda_{1j} \geq ...\lambda_{n_u j} \geq 1 \geq \lambda_{(n_u+1)j} \geq ....\lambda_{nj}. \tag{1.36}$$

Let L_j denote the product of unstable eigenvalues at the j^{th} fixed point of f^p,

$$L_j = \lambda_{1j}\lambda_{2j}.....\lambda_{n_{uj}}. \tag{1.37}$$

Then the principal result of the periodic orbit theory is the following. Given a subset A of phase space, we may define its natural measure to be

$$\mu(A) = \lim_{p\to\infty}\mu_p(A), \tag{1.38}$$

where

$$\mu_p(A) = \sum_j L_j^{-1}. \tag{1.39}$$

The detailed mathematics underpinning these relationships are provided in Grebogi and Ott [61], [62]. This result leads to several important consequences, for

example, it can be shown that the Lyapunov exponents of f are given by:

$$ln\lambda_p = \lim_{p\to\infty} \frac{1}{p} \sum_j L_{j,}^{-1} \tag{1.40}$$

where for each p, L_j is defined by Equation (1.35). An analogous result exists for the topological entropy:

$$h_T = \lim_{p\to\infty} \frac{1}{p} ln(N_p), \tag{1.41}$$

where N_p denotes the number of fixed points of the map f^p.

1.11 Detecting UPOs

There are now numerous examples of UPOs being detected in real life high dimensional chaotic systems. UPOs were detected and analysed in a barotropic ocean model by Kazantsev [79]. Rempel and Chian [113] discussed intermittency in space plasma dynamics using UPOs. Kato and Yamada [76] and Kawahara and Kida [78] extracted UPOs from fluid dynamics models and used these to characterise turbulence properties. McKenzie [101] developed a recurrence plot based test to search for chaotic behaviour in stock market indices and concluded from the results that non-chaotic nonlinear behaviour is present. Gilmore [56] conducted a similar exercise for currency exchange rates, finding similar results. These are just a few of many examples.

As interest in UPOs has grown, so has the number of techniques to identify and capture the lower order cycles. We note, importantly, that the instability of UPOs makes them difficult to detect. Increasingly complex algorithms have developed over time. Broadly these techniques can be divided into two groups; those utilising recurrence matrices and those using more sophisticated Newton-Raphson type methods. The Newton-Raphson methods reduce the continuous dynamical system to a discrete one using Poincare sections. The method of So et al. [133], [132] transforms the time series data using the local linear dynamics along an orbit,

concentrating the transformed data on the periodic orbits. UPOs are subsequently located and extracted by searching for peaks in a finite grid approximation of the distribution function of the transformed data. Another approach, is to "stabilise" UPOs using matrix transformations to convert unstable orbits into stable ones [33], [34], [121], [122]. There are also methods using the Newton-Raphson-Mees method [102], [22], with a damping coefficient, which eliminates the troublesome Poincare section [116]. The Newton-Raphson methods are excellent for noise-free data but do not accommodate noise in the data set and as one would expect, the more sophisticated techniques are highly sensitive to even low levels of noise. We will therefore not discuss these further.

Eckmann et al. [41] introduced the concept of recurrence plots as a tool to study dynamical systems in 1987. Pioneering work in the area of using recurrence plots to detect UPOs was done by Auerbach et al. [6], and is presented in the highly cited paper, "Exploring chaotic motion through periodic orbits". A simple close returns technique was applied to the Henon map to extract lower order UPOs and calculate invariants. The two key findings presented were:

1. For the Henon map one can extract all the UPOs of order n, for n not too large, directly from the chaotic orbit, and calculate their stabilities (Lyapunov exponents), and

2. this information can be used to describe important properties of general chaotic sets.

This research has arguably been a major catalyst for the subsequent development of UPO detection using recurrence techniques. Using a very simple form of close returns and a noise-free time series of 200,000 points, all Henon UPOs up to period 10 were detected. The eigenvalues of all the cycles were calculated using a numerical technique and were with 1%—2% of true values, with errors in the worst case being a factor of 2. Further, the topological entropy was estimated using knowledge of the number of periodic orbits of period n and Equation (1.39).

Whilst this case study demonstrated the considerable strengths of employing this approach, there are some considerations:

1. Close returns were applied to the Henon map in order to find UPOs. Whilst we favor the method of close returns as the best UPO detection technique, the method was applied to a simple two-dimensional map, which is relatively easy to work with in contrast with higher-dimensional sampled flows or experimental data. UPOs for simple maps may contain 1—10 points, whereas for a flow we may find 60 points in the period 1 cycle alone.

2. UPOs of all low-order periods may be detectable, but for highly unstable systems this may require extremely long time series. The time series used contained 200,000 points, which is long. In practice we use single samples or repeated samples of 10,000 points. Processing large numbers of data points can be problematic unless considerable processing power is employed.

3. Calculation of the Lyapunov exponent required considerable matrix multiplication, which can be a source of accumulated errors. Again, with simple maps with a few points in each cycle, the matrix multiplication holds. However, with say 50 points in a cycle, multiplying fitted Jacobians at each point can result in aggregated errors.

4. Finally, and most importantly, the method as presented does not allow for noise. The addition of noise will ensure complete UPOs of some periods are not detectable and the invariant calculations require all UPOs for the calculation. Otherwise the method simply does not work.

Lathrop and Kostelich [89] extended this work to an experimental flow in 1989 in a further landmark paper. The system studied was the Belousov-Zhabotinski (BZ) chemical reaction[149] and a time series of 65,000 points was used. They observed that 95% of points were clustered closely to UPOs up to period 8. Interestingly, the BZ system exhibits cycle lengths of integral multiples of a fundamental period. The BZ attractor and period 1, 2 and 3 cycles are shown below in Figure 1.6 from the paper cited above.

Importantly, the authors presented a method to calculate Lyapunov exponents using the eigenvalues of each detected UPO type that did not rely on matrix multiplication of Jacobians at each point of a UPO. Their method accommodated flows. They subsequently estimated the information dimension using the Kaplan - Yorke conjecture relationship and achieved reasonable results. Noise-infected data was not considered as part of this research.

These two papers ([6] and [89]) laid the foundations for techniques that detect UPOs using the method of recurrence matrices, and efforts to calculate estimates of invariants using the UPO set.

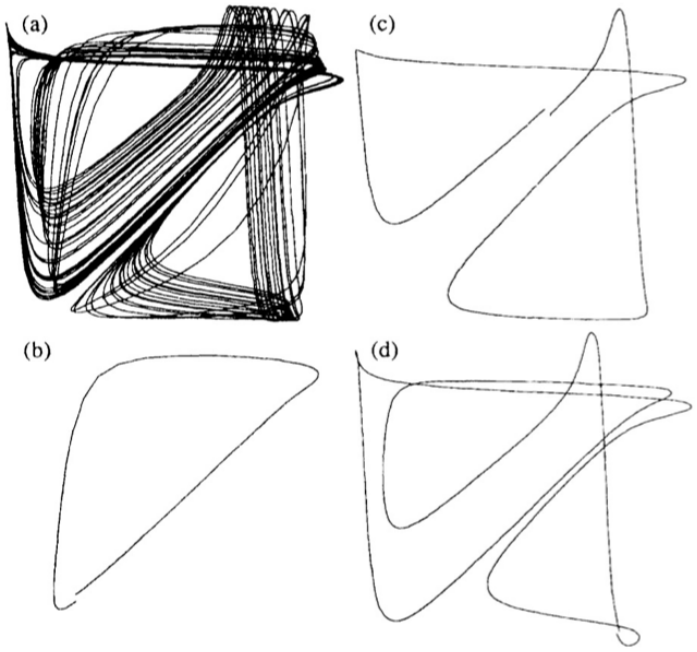


Figure 1.6: (a) BZ Attractor, (b)–(d) Trajectories near UPOs of period 1,2 and 3.

1.12 Summary and Discussion

In this chapter we have introduced the building blocks and conceptual tools required for the development of a noise-reduction technique for chaotic time series.

We firstly have introduced the fundamental concepts of maps and flows, nonlinearities and phase space. The Devaney definition of chaos is presented as it provides insight on the anatomy of chaos; revealing the significance of unstable periodic orbits (UPOs) at the most fundamental level in chaos. We also discuss Touhey's definition of chaos, which is expressed entirely in terms of periodic points. We have described and formulated Lyapunov exponents for flows and maps, and argued that the short time Lyapunov exponent set derived from the basis set of lower order UPOs is a meaningful description of orbital divergence during the evolution of a chaotic system in phase space.

The concepts introduced so far allow us to next introduce unstable periodic orbits and periodic orbit theory. The periodic orbit theory of Cvitanovic provides a real framework for the application of UPOs in numerical methods. UPOs comprise the skeleton of the attractor, are ordered hierarchically and are topological invariants. After transients are eliminated, all points within a time series lie within a small distance of a UPO and thus the set of lower order UPOs provides an excellent basis through which we can approximate the time series (and thus the attractor).

Detection of UPOs from a time series is the final tool we harness. There is a large volume of research on locating and extracting UPOs from a time series. Much of it relies on Newton-Raphson type methods, is computationally complex and very effective for noise-free data. However, we are occupied with highly noise-infected data and these techniques fail immediately in this case. Instead we will be using recurrence methods or close returns plots to locate UPOs. These methods are robust in the presence of noise and will be discussed in Chapter 3.

Finally we discuss and review two key papers that form the motivation and platform for the research in this thesis. The pioneering work of Auerbach et al.

[6] in finding the UPOs of a simple noise-free Henon map by using close returns, and subsequently estimating invariants has been highly influential. Lathrop et al. [89] extended this to flows by considering the experimental BZ system. Neither of these authors specified addressed noise. Extending this type of approach to accommodate noise is the focus of this thesis.

Chapter 2

Noise Reduction Techniques

2.1 Chapter Overview

In this chapter we start by briefly defining, comparing and discussing measurement and dynamical noise as well as defining the signal to noise ratio. We summarise how noise may actually induce chaos in a dynamical system and that linear noise reduction tools cannot simply be applied to chaotic systems without serious consequences. Non-noise related features of chaotic systems that are problematic in the context of numerical mathematics are reviewed as these have implications for numerical noise-reduction techniques. We further discuss the numerical algorithms used to directly calculate chaotic invariants (Lyapunov exponents and correlation dimension) and how noise affects them. Finally we categorise and critically review the multitude of noise reduction methods that have been proposed by researchers, with a particular emphasis on the few techniques that utilise the properties of UPOs.

2.2 Types of Noise

The simplest description of noise is that it is the "unwanted" part of the signal [74]. In reality, deterministic systems are a nice theoretical construct but are unlikely to exist in nature as all systems interact with their surrounding environment and are to some degree contaminated. Noise is problematic. It has been shown to

interfere with or even invalidate numerical algorithms used for system identification [148], [77] and parameter estimation [130]. Nonlinear deterministic dynamical systems can evolve in highly erratic patterns autonomously, without the influence of added noise. Although the generalised theory of chaotic dynamical systems is now well established, many practical challenges arise when when translating the theoretical concepts into methods for analysis of experimental or real world data time series. The deterministic reference frame is best considered as a subset of a broader structure involving fluctuations of the environment and of the chaotic system itself.

In signal processing, noise is a general term for unwanted variations that perturb a signal during detection, storage, transmission, processing, or conversion. Noise is classified at a highly detailed level by its statistical properties (often referred to as the "colour" of the noise) and by how it modifies the underlying or intended signal. There are a multitude of very specific definitions. For measurement noise (which gets added to the signal) categories include white noise, pink noise, black noise, Gaussian noise, uniform noise, flicker noise, brown noise, Cauchy noise, just to name a few. We will not require this level of detail just yet and will focus on the broadest groupings of measurement (or observation) noise and dynamical (or system) noise.

For an excellent classification of mathematical schemes by which noise is imposed on a dynamical system, the reader is referred to the paper by Argyris, Andreadis, Pavlos and Athanasiou [5]. We shall follow their nomenclature, categorisation and formulation. When noise interferes with the evolution of a dynamical system, it is called dynamical noise and may be formulated as an additive or multiplicative expression. In contrast measurement noise does not influence the evolution of the system. This is the key point of distinction and both types of noise may take the form of an additive or multiplicative expression.

We consider a flow (continuous in time) dynamical system in the Euclidean space $\mathbb{R}^n$ defined by the equation:

$$x' = f(x, \mu, t), \tag{2.1}$$

where: $x \epsilon \mathbb{R}^r$, $\mu \epsilon \mathbb{R}^k$ are vectors and $t \epsilon \mathbb{R}$.

2.2.1 Measurement Noise

Measurement noise, also known as additive or observation noise does not enter directly into the evolution of the dynamical system and is by far the easier to deal with. In the presence of measurement noise we must study a new evolution $X(t)$ by directly applying the noise at a point in time to $x(t)$ (the solution of Equation (2.1)). In its most general form $X(t)$ is defined by:

$$X(t) = K(x(t), t, w), \tag{2.2}$$

where $x = \{x_1, x_{2,}...x_r\}$, $w = \{w_1, w_{2,}....w_r\}$, K is a function and w is a noise.

Measurement noise is further specified by whether it is an additive contribution or multiplicative contribution to the new evolution equations:

$$Additive: \qquad X(t) = x(t) + w, \tag{2.3}$$

$$Multiplicative: \qquad X(t) = h(w \cdot x(t)), \tag{2.4}$$

where h is a map.

The measured orbit is corrupted by measurement noise. Sources include finite precision measurements, truncation errors, and missing data (both temporal and spatial). If the additive error term is modeled by a random variable for example, then an identically and normally distributed (IND) process may be written as:

$$w \sim N(0, \sigma^2_{noise}), \tag{2.5}$$

where σ^2 is the variance of the noise.

2.2.2 Dynamical Noise

Dynamical noise, in contrast, appears as a disturbance that influences the evolution of the dynamical system and is formulated as follows:

$$x' = f(x, \mu, t, w), \tag{2.6}$$

where $x \,\epsilon\, \mathbb{R}^n$, $\mu \,\epsilon\, \mathbb{R}^k$ are vectors, $t \,\epsilon\, \mathbb{R}$ and w is a noise.

Dynamical noise arises from a feedback process wherein a system is perturbed by a small random amount at each time step (noise is added during the evolution of the system). Dynamical uncertainty refers to external fluctuations interacting with and changing internal variables in the underlying system. Whilst measurement uncertainty obscures the state vectors, dynamical uncertainty changes the actual dynamics.

When the noise is added to the right hand side of Equation (2.6) we call it additive dynamical noise. If noise is included as a perturbation within the function f in Equation (2.6) we refer to it as multiplicative dynamical noise. The general forms are as follows:

$$Additive: \qquad x' = f(x, \mu, t) + w, \tag{2.7}$$

$$Multiplicative: x' = f(x, g(\mu, w), t), \tag{2.8}$$

where $x \,\epsilon\, \mathbb{R}^n$, $\mu \,\epsilon\, \mathbb{R}^k$ are vectors, $t \,\epsilon\, \mathbb{R}$ and w is a noise and g a vector function.

Dynamical noise and measurement noise are two notions of the error that may not be distinguishable a posteriori, based on the data only. Both descriptions can be consistent to some extent with the same signal [25]. For strongly chaotic systems that are everywhere hyperbolic (Axiom A systems), measurement noise and dynamical noise can be mapped onto each other [134].

For a noise reduction scheme to be practically useful, one should require that it works at the very least in reducing measurement noise, and that it does not produce nonsense in the case of dynamical noise.

Generally, dynamical noise induces much greater problems in data processing than measurement noise, since in the latter case a nearby noise-free orbit of the underlying deterministic system exists and can indeed be found to some degree of accuracy by a proper noise reduction procedure. Also, what we interpret as dynamical noise sometimes may be a part of higher dimensional deterministic dynamics evolving with a small amplitude. Even if this is not so, dynamical noise may be essential for the observed dynamics. For an interesting example using the logistic equation (with the parameter a=1.9408), where Gaussian noise interacts with the dynamics to nudges the orbit off the periodic orbit into a repeller, see [135]. This example nicely illustrates that dynamical noise may have more severe effects than simply smearing out some small-scale deterministic structures.

We will assume that any researcher employing the noise-reduction technique presented in this thesis, does have knowledge of the type of noise present in the measured time series, although does not necessarily need to know the magnitude of the noise. Most results presented in this thesis will address high levels of measurement noise and low-level additive dynamical noise. Similar expressions are defined for maps and the categorisation is presented collectively in Figure 2.1 below, which follows [5]. We shall provide specific formulations for measurement and dynamic noise employed for our research in Chapter 5.

Finally, we wish to point out that not all noise is introduced to a system by external factors. Many real physical systems produce noise-infected output as a result of intrinsic noise (e.g. thermal). Whether this noise acts as dynamical or measurement noise depends on the type of process. In [146], a chemical reaction is investigated by Wu and Kapral. The process dynamics, when being well stirred, are completely described by three coupled ordinary differential equations. If the stirring is inadequate (or the diffusion process too slow), the dynamics in different sections of the container become slightly out of phase. A global observable variable,

like the concentration of one reactant thus becomes infected with noise. In this example the intrinsic noise has all the characteristics of measurement noise.

Additive/Measurement Noise

System	Additive	Multiplicative
Continuous in time $X'=f(x,\mu,w)$	$X(t)=x(t)+w$	$X(t)=h(w,x(t))$
Discrete in time $x_{n+1}=F(x_n,\lambda,n,w_n)$	$X_n=x_n+w_n$	$X_n=H(w_n,x_n)$

Dynamical/System Noise

System	Additive	Multiplicative
Continuous in time $x'=f(x,\mu,w)$	$x'=f(x,\mu,t)+w$	$x'=f(x,g(\mu,w),t)$
Discrete in time $x_{n+1}=F(x_n,\lambda,n,w_n)$	$x_{n+1}=F(x_n,\lambda,n)+w_n$	$x_{n+1}=F(x_n,\gamma(\lambda,w_n),n)$

Figure 2.1: Mathematical definitions of measurement and dynamical noise. We will be using additive measurement noise and additive dynamical noise in our modelling.

2.3 Signal-to-Noise Ratio

The signal-to-noise ratio is defined as the ratio of the power of the signal (wanted information) to the power of background noise (unwanted information). The noise-infected signal may have a non-zero mean that can be subtracted off. Then if the variance of the noise-free signal is σ^2_{signal} and of the noise signal is σ^2_{noise}, then signal to noise ratio (SNR) is defined as:

$$SNR = \Big(\frac{\sigma^2_{signal}}{\sigma^2_{noise}}\Big), \tag{2.9}$$

An alternative approach, which we use throughout this thesis, is defined in terms of decibels:

$$SNR_{dB} = 10.\log_{10}\Big(\frac{\sigma^2_{signal}}{\sigma^2_{noise}}\Big) = 20.\log_{10}\Big(\frac{\sigma_{signal}}{\sigma_{noise}}\Big). \tag{2.10}$$

To calculate the SNR for a noise-infected time series, when we know the noise-free time series and noise vector (noisy time series - noise-free time series), we use:

$$SNR_{TS} = 20.\log_{10}\Big(\frac{rms(TS_{noise-free})}{rms(TS_{noise-vector})}\Big) = 20.\log_{10}\Big(\frac{\sigma_{TS_{noise-free}}}{\sigma_{noise-vector}}\Big). \tag{2.11}$$

2.4 Noise Induced Chaos

The presence of noise itself can actually instigate chaos in a dynamical system. Crutchfield et al. [26], [27] observed noise truncating a period-doubling cascade and inducing chaos. Gao et al. [50], [52] have further confirmed and extended Crutchfield's work, finding that there are three necessary conditions required to instigate chaos:

1. Firstly, the noise level has to be within a specific narrow band. Noise less than the minimum is insufficient to trigger chaos, and noise above the band maximum will dampen and destroy the induced chaos.

2. Secondly, when the noise level of the signal is within the narrow band, the adjacent chaotic states should still behave chaotically on finite scales.

3. The third and most critical condition requires that the periodic state, when subjected to weak noise, should undergo a process that is considerably more diffusive than Brownian motion. In the case of a specific period-doubling cascade, the entire period-doubling sequence is terminated. The cascade is masked by noise otherwise. Hwang et al. [69] have observed these conditions in a semiconductor laser system.

It is worth briefly touching on the statistical framework of noise. The underlying dynamical systems we work with in this thesis are *Markov processes* in discrete time (as they are sampled). In a Markov process, the existing state completely determines the probability distribution of the future states. Real dynamical systems exhibiting dynamical noise are also examples of Markov processes. In experimental situations, the actual state of the Markov process itself is unknown (we cannot measure it directly), and we rely on experimental measurements, usually a scalar time series, that provide incomplete phase space information. These Markov process as just described are referred to as *hidden Markov processes* (HMP).

Siefert et al. [129] have developed a quantitative method to differentiate between dynamical and measurement noise. Their method uses the theory of diffusion processes (the theory of Kramers-Moyal coefficients) to estimate the magnitude of both dynamical and measurement noise. The method does not require knowledge of the system equations. For the case of a chaotic dynamics concurrently infected by dynamical and measurement noise, they show how to numerically estimate the magnitude of both types of noise. As a consequence, they present a criterion to verify the correct embedding for chaotic dynamics infected with dynamical noise. Similarly, Bottcher et al. [11] developed an approach using a broad group of Langevin processes that describe a variety of complex dynamical systems. Interestingly, these models rely on the underlying Markov properties holding, and Kleinhans et al. [80] have shown that the presence of measurement noise spoils these Markov properties.

2.5 Finite Impulse Response Filters

We need to distinguish between recursive and non-recursive linear filters. The most general linear filter takes a sequence x_k of input points and produces a sequence y_k of output points by the formula:

$$y_n = \sum_{k=0}^{M} c_k x_{n-k} + \sum_{j=0}^{N-1} d_j y_{n-j-1.} \tag{2.12}$$

The $M + 1$ coefficients c_k and the N coefficients d_j are fixed and define the filter response. This filter produces each new output value using the current M previous input values, and from its own N previous output values:

1. If $N = 0$, there is no second sum in Equation (2.12) and the filter is called a *finite impulse response (FIR) filter* or non-recursive filter.

2. If $N \neq 0$, then it is called an *infinite impulse response (IIR) filter* or recursive filter.

2.6 Limitations in Noise-Filtering Non-Linear Data

Noise reduction for a time series may be considered as the filtering of a noisy signal to extract a relatively clean signal. By definition, the noise is the unwanted part of the data. An ideal filter would clean any noise from the time series, leaving a pure signal, without interfering with the underlying dynamics in any way. We could then reconstruct the chaotic attractor and measure the chaotic invariants using established algorithms to fully characterise the attractor. When it comes to nonlinear dynamical systems, the non-linearity of the system creates a multitude of problems and it is difficult to apply a filter and not interfere with underlying signal and produce flawed results.

It is difficult to find many simple and effective data filters in the body of literature relating to nonlinear dynamical systems. Specialist non-linear filters do

exist but are few and generally recursive in nature. We will see later that recursive filters interfere with the underlying chaotic signal and cannot be employed for the purposes of this thesis.

Most noise filters are in essence linear, and are inadequate when applied to non-linear time series data. This deficiency is compounded further by the estimation error implicit in each of the input assumptions that are usually required. Linear filters that rely specifically on frequency considerations, to distinguish signal from noise, are problematic when applied directly to chaotic data.

Fourier analysis is the conventional linear model adopted by practitioners to approach the signal-noise problem. Data are transformed from the time domain to the frequency domain, where we assume high-frequency components are mostly noise. Components at frequencies greater than a fixed cut-off are attenuated. The modified frequency domain signal is finally inverted to provide a noise-reduced signal. Unfortunately the frequency domain power spectrum of a chaotic signal has a broadband structure, making it indistinguishable from noise. As the signal inhabits all frequencies, a frequency cut-off approach is inadequate and will in fact alter the dynamics of the noise-reduced signal. It is as if the noise is deeply intertwined with the signal as a result of the nonlinearities in the system. For a useful comparison between Fourier spectra of regular and chaotic systems the reader is referred to [39].

There is a plethora of other linear filters available [21], all specifically developed to target noise with a particular characteristic. Each is very useful in the correct context, but unfortunately none of these will be directly applicable to the entire chaotic time series. However we do note that if large enough segments of the time series can be viewed as approximately linear or periodic, then these noise-reduction tools will be directly applicable and most useful.

It is surprising that the simple process of filtering a chaotic time series may result in a time series with different underlying dynamics. Thus any dynamical invariants calculated using such a noise-reduced chaotic time series will not be

representative of the original system. Research by Badii et al. [7] reveals that a low-pass filter effectively generates an additional Lyapunov exponent, depending on the cut-off frequency used. For a sufficiently low cut-off frequency, the low-pass filter may increase the fractal dimension of the reconstructed chaotic attractor. Mitschke et al. [105] have presented an experimental example of this phenomenon using data measured from an electronic circuit.

Fortunately, Broomhead et al. [13] proved that finite impulse response (FIR) filters (finite-order and non-recursive filters), such as moving average filters, do not have this effect. Therefore the universe of applicable filters for chaotic time series is restricted to FIR filters and is applicable to non-stationary time series because chaotic time series are generally strongly non-stationary.

2.7 Embedding of Noise-Infected Data

According to Takens' theorem [137], when the embedding or reconstruction in phase space is carried out correctly, the reconstructed dynamics are topologically equivalent to the dynamics on an attractor in the original phase space. This means that topological invariants of the system are conserved. The invariant properties of the system that are conserved under the topological equivalence property of the diffeomorphism (we actually only need a homomorphism) include singularities, closed orbits, attractor geometry and flows. For infinite noise free data sets the choice of reconstruction parameters such as time delay is almost arbitrary (providing the embedding dimension m is sufficiently high).

In the more realistic case of noise-infected data, the embedding process, even optimised, is severely limited in the presence of noise. Takens' theorem simply does not allow for noise and reconstructing an attractor with noise-infected data provides no guarantee that invariants are preserved. In practice, the embedding parameters (m and τ) must be determined numerically and numerical algorithms to determine these perform poorly, particularly for m when noise is present. Many noise-reduction techniques for chaotic time series start with embedding the data in phase space. This immediately restricts the procedure to very low levels of noise

and arguably invalidates it for more significant levels of noise. Similarly, numerical algorithms to estimate invariants also pre-embed the data and suffer the same flaw.

At this time, the effects on the Takens' embedding of noise, both measurement and dynamical, are not thoroughly understood. There is some empirical evidence that the fidelity of the embedding may be somewhat resilient to measurement noise, shown in [64]. Further, a more complex theory developed in [135] also suggests such a robustness.

As a result of the concerns raised above, the method presented in Chapters 4 and 5 uses the scalar time series directly and avoids embedding the noise-infected data.

2.8 Challenges in Approximating the True Dynamics

Having ruled out traditional non-linear techniques and IIR linear techniques, let us consider the constraints in developing a new approach. Any algorithm to reduce noise and uncover the true dynamics must address several key challenges that result from the special nature of chaotic dynamics. Before proceeding to more detailed discussions on noise reduction, we should mention that noise, although problematic, is not the only factor affecting the integrity of analysis to find the underlying signal. The following summary is distilled from an excellent summary of non-noise factors presented in [83].

2.8.1 Measurement Function

The evolution of a low dimensional dynamical system may be expressed either as a set of ordinary differential equations (ODEs) or as a discrete mapping. Although most real systems evolve continuously in time, data are always sampled discretely. The noise-free dynamics of the underlying system can be obtained by starting with some initial condition x_0 and iterating a function f via:

$$x_{n+1} = f(x_n). \tag{2.13}$$

The state of the system at any time t_n is represented by a vector y_n in some m-dimensional phase space. A very long orbit will eventually define an attracting set for practical purposes (theoretically we require the infinite time series) as all the possibilities in phase space are eventually traversed. This attracting set, and the invariant measures defined on it are the natural features to study from a nonlinear dynamics perspective.

If the equations defining the system dynamics are not explicitly known, this phase space is not directly accessible by a researcher. However, in typical situations, points on the dynamical attractor in the system phase space have a one-to-one correspondence with measurements of a limited number of variables. This is a powerful fact as by definition a point in phase space carries complete information about the current system state. A one-to-one correspondence means the phase space can be identified by measurements. This result is formalised in Takens' Theorem (see Section 1.8).

Assume we can measure m variables simultaneously:

$$y(t) = (y_1(t), y_2(t), ..., y_m(t)). \tag{2.14}$$

This m-dimensional vector may be considered as a function of the system state $x(t)$ in the full system phase space.

$$y = F(x) = (f_1(x), f_2(x), ..., f_m(x)). \tag{2.15}$$

F is referred to as the *measurement function* and the m-dimensional vector space occupied by the y vectors is known as the the reconstructed space (or embedding space). The measurements are represented as a vector-valued function F of x, as each measurement is a well-defined function of x (a consequence of uniqueness of the states in the system). The measurement function F defines a

one-to-one mapping between the original attractor states in phase space and the set of m observation vectors, provided the embedding dimension m is sufficiently large.

Experimental results will vary from theoretical formulations for a number of reasons:

1. We usually cannot obtain y_n directly and instead rely on the scalar values generated by the measurement function F,

2. Perturbations of y_n, caused by a stochastic process or fluctuations in system parameters may drive the system into being stochastic rather than deterministic states.

3. The measured time series often represents a short, finite, segment of system evolution. Depending on the ergodicity of the system the orbit segment may explore most of the attractor, or malinger in a small neighbourhood. Estimation error is always present from random fluctuations or due to the discretisation involved in sampling.

Whilst Takens embedding theorem is underpinned by topological considerations, we can understand the concept intuitively as follows. Suppose we can measure a single system variable and its derivatives up to some order m. If the dimension of the system is less than m, we then have sufficient information to construct a set of m differential (or difference) equations. These m equations are sufficient to completely define the dynamical system. Measuring m derivatives and measuring the system at m different time intervals are equivalent, provided we use a sufficiently high sampling rate, m.

Remarkably, provided certain conditions are met, the reconstructed attractor contains the same information as the original system. The assumptions required are that m is sufficiently large, the sampling rate is high enough, the time series is sufficiently long and the measurement function F is twice differentiable [137].

These conditions are unlikely to be met when working with real world data. Firstly, the concept of obtaining "proxy" derivatives of the system state by recording successive observations is greatly diminished when the data is infected with noise. Higher order derivatives are extremely difficult to calculate numerically. Further, the differentiability of F is violated by digitisation (both sampling and quantisation) of the data. Regardless, experimenters generally assume that provided the conditions are loosely met the topology of the embedded attractor reasonably resembles the topology of the underlying attractor M; and the sequence of embedded points evolves under a deterministic process that is approximately equivalent to the evolution operator ϕ.

In order to focus the discussion in this thesis on the influence of noise, we will assume throughout that the measured data are otherwise well behaved. By this we mean that the signal would be to some extent predictable by exploiting an underlying stationary deterministic rule if it were not for the noise. This is the case for data sets that can be embedded in a low dimensional phase space, which are stationary and which are not too short. Violation of each one of these requirements leads to further complications that will not be addressed here.

2.8.2 Drifting Problem

Let us assume that m dimensions are sufficient to produce an embedding. We can then write:

$$y_{m+1} = f(y_1, y_2, ..., y_m) + \eta_{m+1}, \tag{2.16}$$

where η_{m+1} represents an measurement noise term, and y_i are the time series elements.

The subscripts denote sequential elements in the time series and we refer to this a forward-in-time representation of the dynamics. If we seek to find an approximation f^* to f, so we may estimate y_{m+1} as:

$$y^*_{m+1} = f^*(y_1, y_2, ..., y_m). \tag{2.17}$$

We could start with the first m time series entries $y_1, y_2, ..., y_m$ and output a prediction of y_{m+1} as y^*_{m+1}. Then use $y_2, y_3, ..., y_{m+1}$ to predict y_{m+2} as y^*_{m+2} and so forth. This approach will adjust all the time series entries after the m^{th}.

This process will fail to reduce noise if iterated by running each updated time series back through the algorithm. For a chaotic time series, this approach will produce output time series that drift further and further away from the original due to errors in each y_i being amplified as a result of the sensitivity to initial conditions. In summary, we cannot extrapolate using previous points.

2.8.3 Curvature Problem

For some statistics derived from a chaotic time series, the curvature effect arising from the curved shape of the attractor becomes highly material. From a technical perspective, the curvature effect is closely linked to the limit assumption used in calculating invariants (e.g. correlation dimension).

This effect is highly material with numerical techniques applied to time series to estimate invariants such as correlation dimension and Lyapunov exponents and arguably a severely limiting factor.

The sparsity patterns in relatively small scales also causes measurement bias and is related to the lacunarity of the attractor. At this time, the relationship between noise level and the bias in estimates caused by macroscopic geometric effects has not been explained. Coban et al. [23] discuss a model for the curvature effect based on the distribution of estimation errors.

2.8.4 Errors-in-Variables Problem

Many existing noise reduction methods rely heavily on the use of linear approximations, in seeking to find a superior replacement orbit either on the attractor directly or on a linearised subspace. The Euclidean or L_2 metric, used for determining optimality is favoured in nonlinear systems analysis and thus the method of least squares is employed. An example is the algorithm of Sano and Sawata

[117] used to determine Lyapunov exponents, which relies heavily on least squares calculations along the orbit.

The classical least squares problem assumes that:

$$y_i = ax_i + b + \epsilon_i, \tag{2.18}$$

where each observation y_i is a linear function of the independent variable x_i, and the random variables ϵ_i are normally distributed with mean 0 and variance σ^2. We assume that the only error occurs in the measurement of y_i, as the values of x_i are known exactly.

This assumption is invalidated when one deals with noisy input data. In fact, all the observations are measured with some error. The classical least-squares problem must be replaced by:

$$y_i = a(x_i + \delta_i) + b + \epsilon_i. \tag{2.19}$$

This is referred to as the errors-in-variables model. Here δ_i and ϵ_i are independent, normally distributed random variables. The classical least squares estimate is biased, i.e. the slope a is underestimated by an amount that depends on the variance of the δ_i and is independent of the number of observations. Each iteration of a noise reduction method, using ordinary least squares is inherently inaccurate unless corrections are made to reduce the bias.

2.8.5 Ill-Conditioned Least Squares Models

Most low dimensional attractors have a fractal structure with points tending to form a Cantor set of layers. The layers may be indistinguishable from a curve because of the limited resolution and size of typical data sets. Sometimes the points captured in a small ball surrounding a typical orbit may be coplanar, leading to ill-conditioned least squares problems. The covariance matrix of measurements is nearly singular. The numerically computed solution of an ill-conditioned least squares problem has a large relative error, and the Jacobian matrix cannot be

calculated accurately. We may employ Singular Value Decomposition (SVD) to detect these situations, and numerically computed singular values are seldom zero due to round-off error. Large values of the condition number can be used instead to detect ill-conditioning. The problem is exacerbated when data are only known to finite accuracy.

2.8.6 Outliers and influential points

Influential points are relatively common and a serious issue in noise reduction methods. These are small groups of points residing relatively far away from the rest, and as a consequence of the stretching and folding mechanism, lead to difficulties in accurate estimation of the Jacobian matrix. They may be outliers, resulting from anomalies in the data, but in usually arise from the striated structure of the attractor. Influential points are a heuristic notion; there is no formal definition. Unlike ill conditioning, influential points do not necessarily affect the accuracy of least squares solutions. Judgement is required in considering whether to discard influential points.

2.8.7 Loss of Information

Usually the embedded orbit is constructed from a measured scalar time-series through the use of delay vectors. The trajectory-based methods produce a replacement orbit that is usually not realisable as a delay vector reconstruction from a scalar time series. This implies there is some loss of information about underlying dynamics.

2.8.8 Summary of Limitations

There are many challenges in developing a noise-reduction technique for a chaotic time series as we have seen in the previous section. For a more detailed review of problems arising with estimating dynamics from data sets, see Kostelich [82]. The noise is tightly intertwined with the nonlinear signal and it is near impossible to separate the two without corrupting the underlying signal itself. Recursive filters (IIR) damage the underlying signal, and this rules out most traditional filters, both

linear and non-linear. In particular, direct application of Fourier filters runs into problems as the chaotic signal looks like broadband noise and separation criteria are difficult to define.

In the context of developing new, tailored techniques, there are further challenges. In particular, we cannot extrapolate existing data (recursive concept again) and there are issues with the application of least squares to determine optimality as a result of noise in the data and the unusual geometry of chaotic attractors. Further, the curvature of the attractor provides limitations for any stepping trajectory-based approaches.

Methods relying on firstly embedding the noise-infected data set are immediately compromised as a result of data quantisation and finite sampling time. Further, Takens theorem does not guarantee topological equality of an embedding if noise is present. In such a case, one is unable to quantify whether the approximation is valid or not.

The issues discussed in Section 2.8 are not just theoretical. While conducting research for this thesis, nearly all of these problems were encountered in the data analysis. In summing up, the most important considerations for an effective noise-reduction filter for chaotic time series are that it must not be recursive (IIR) and ideally not rely on an embedding as a starting point. All techniques will need to also deal with curvature and the unusual geometry of chaotic systems. Techniques that approximate the time series and attractor with a set of cycles (UPOs) or approximate cycles will immediately remove many of the challenging barriers. We summarise the attributes of an effective/ideal noise-reduction technique for chaotic data below in Figure 2.2.

ATTRIBUTES OF EFFECTIVE NOISE-REDUCTION TECHNIQUE FOR A CHAOTIC TIME SERIES
Separate noise from signal (directly or indirectly)
Must not corrupt underlying signal (FIR method, locally linear approximation holds)
Handle a range of noise contamination from small to large. (> 10%)
Maximize use of information (e.g. time-ordering, geometry)
Prefer technique that operates directly on scalar time series and adds embedding techniques where appropriate
Must be simple and easily programmable
Require low-level of user input/judgment
Work on experimental data (not just model systems)
Accommodate low levels of dynamical noise, as well as measurement noise
Immune to scaling considerations and allow for curvature
Must not have unintended consequences or artefacts. (e.g. spurious dimensions added).
Filtered time series or attractor must be superior to the original in terms of minimization of a suitably defined error metric

Figure 2.2: Summary of attributes of an ideal noise-reduction technique for experimentally determined chaotic time series

2.9 Estimating Invariants from Noise-Infected Data

Noise complicates the entire field of nonlinear time series analysis immensely, in that many algorithms are rendered useless by a few per cent of noise corruption to the signal. Numerical algorithms to estimate the Lyapunov exponents (usually the most positive) and correlation dimension are of greatest importance.

Kostelich and Yorke [85] demonstrated with simple numerical experiments that a noise level of 1% of the time-series extent makes it impossible to measure the correlation dimension of the attractor using distances less than 3% of the attractor extent.

Abarbanel [2] found that the negative Lyapunov exponent cannot be determined when the noise level is as small as 10^{-4}, and none of the exponents can be determined with satisfactory precision in the presence of 1% noise. This is true even if the minimum required embedding dimension is known. Higher embedding dimensions lead to the additional problem of spurious negative exponents.

Noise has a material impact for the numerical algorithms used to calculate the global structural invariants of chaotic dynamics. Noise reduces the effective scaling ranges for computations, since most of them have been derived under noise-free assumptions. The range of scales accessible from the data set is a constraint, limited above by the overall extent of the attractor and below by the amplitude of the noise. With experimental data, noise makes all reconstructions non-invertible and the quality of the reconstruction is a question of degree. Nearest neighbours are commonly measured and utilised in determining noise reduction algorithms and noise does corrupt this process also.

We also need to appreciate that higher dimensional dynamics with small amplitude may appear nearly identical to dynamical noise. From another viewpoint, rather than separating noise from the underlying deterministic system, the challenge is that of separating the low-dimensional dynamics from a higher-dimensional complex system.

2.9.1 Determining Lyapunov Exponents from Noisy Data

Numerical methods to determine Lyapunov exponents tend to estimate the largest or maximal Lyapunov exponent ($\lambda_{\max}$). Other methods focus on determining just the positive exponents and some estimate the full Lyapunov exponent spectrum.

Techniques to estimate the spectrum of Lyapunov exponents for dynamical systems with known mathematical models are well established. These include algorithms published by Wolf et al. [145], Benettin et al. [9] and Shimada and Nagashima [127].

Determination of these exponents from experimental time series is significantly more difficult. Existing algorithms may be broadly categorised into two groups:

1. Trajectory-based, real space or direct methods (Wolf et al. [144], Rosenstein et al. [114], Kantz [73]), and

2. Perturbation, tangent space or Jacobian methods (Sano & Sawada [117], Eckmann et al. [42], Brown et al. [15], and Kruel et al. [86]).

A useful comparison of the relative efficiency and accuracy of methods is provided by Geist et al. [55].

All experiments are contaminated by noise in some form. Dynamical noise may originate from within the system itself and certainly some level of measurement noise is unavoidable in almost any experiment. Unfortunately, the presence of such noise will lead any experiment to exhibit some form of SDIC, exponential or otherwise. Using the descriptive language of Deissler and Farmer [37], a chaotic system ($\lambda_{\max} > 0$) is a noise amplifier. After a perturbation occurs, the effects of the perturbation are magnified over time by the systems intrinsic dynamics. Thus, both the system itself and external perturbations contribute to the systems unpredictability. A non-chaotic system ($\lambda_{\max} < 0$) is a noise muffler: the effects of external perturbations decay asymptotically to zero over time. Any unpredictability is solely due to perturbations.

Most of the commonly used methods for determining Lyapunov exponents implicitly make the apriori assumption that the data come from a noise-free, strictly deterministic system. Consequently, these methods are prone to false positives in which non-chaotic data are misidentified as being chaotic [138], [11].

Franca and Savi [48] conducted a comparison of the key numerical algorithms for determining Lyapunov exponents using an analytical model of a damped driven pendulum to provide examples of periodic and chaotic motions. They then superimposed uniformly distributed random noise onto the resulting signals. The simulated data was used to compare the relative sensitivity of the algorithms of Sano and Sawada [117], Wolf et al. [145], Rosenstein [114], and Kantz [73], [75] to measurement measurement noise. They found that the algorithms of Sano and Sawada and Wolf et al. algorithms were especially sensitive to measurement noise, whereas the algorithms from Rosenstein and Kantz were less so.

Many examples exist in the published literature where the estimation of Lyapunov exponents fails to distinguish chaos from noise in time series data (e.g., [19], [138], [69], [126]), which can lead to erroneous conclusions of "chaos" even for periodic (e.g., [52]) or completely random (e.g., [126]) datasets that are not chaotic at all.

Strictly speaking, the Lyapunov exponents "are not rigorously defined in the presence of external noise" [144]. Several authors have questioned the validity of even defining Lyapunov exponents in the presence of noise [110], [144], [44]. Without being able to determine with certainty that the cause of the SDIC found in a given dataset was because of a deterministic process, an accurate diagnosis of chaos is generally not possible. Thus, findings of positive Lyapunov exponents, should never be taken in isolation as conclusive proof of chaos without further evidence. Taking this a step further, Ellner and Turchin [44] have suggested that "strict separation between chaotic and stochastic dynamics" may be "unnecessary and misleading" for certain biological systems.

Lyapunov exponent estimation algorithms based on orbital divergence are not robust to dynamical noise. In a chaotic system, some (or possibly all) of the long-term unpredictability is internally generated by nonlinear causal relationships among the state variables. Paired orbits will diverge even if $\lambda < 0$, simply because they did not experience the same sequence of perturbations, so the estimate of λ is positively biased, according to Sayers et al. [120]. Consider, for example, a system

influenced by dynamical noise. No matter what the phase space structure of the system is, two orbits starting nearby will always diverge if they are influenced by different sequences of random perturbations. The same is true if the perturbations are deterministic. Ellner [45] points out that the key to defining chaos in systems with exogenous elements is to observe whether nearby orbits diverge when subjected to precisely the same exogenous perturbation and present a method based on mappings. A model for the effects of dynamical noise within chaotic dynamics is still an open problem.

Gao et al. [51] have devised a scale-dependent Lyapunov exponent measure (SDLE), along with an efficient algorithm to separate noise from chaos. The authors define a framework, grouping different types of motion together and identifying different scale ranges where the various types of motions are manifested. The authors assert the algorithm can be applied with some accuracy to short noisy time series and readily classify the full spectrum of motions.

Finally, note that when driven by external noise, local instabilities in otherwise stable dynamical motion can cause behaviour that looks quite similar to deterministic chaos (see [37]). Microscopic fluctuations are amplified to create irregular macroscopic variations in both the amplitude and phase of a signal. In contrast to deterministic chaos, however, chaotic-looking behaviour generated by local instabilities vanishes when external noise is eliminated. This behaviour is easily confused with deterministic chaos. In an experiment where local instabilities are suspected the best method to make the distinction from deterministic chaos is to add external noise to the dominant source, and test for linear scaling of the amplitude of noisy behaviour with the amplitude of the external noise.

In summary, existing numerical methods to estimate Lyapunov exponents are not robust in the presence of any noise that is not extremely small.

2.9.2 Determining Fractal Dimensions from Noisy Data

Research on the effect of noise on dimension measurement has generally focused on correlation dimension. There are other methods of measuring dimension but

the correlation dimension is widely used. It has the advantage of being straightforwardly and quickly calculated, of being less noisy when only a small number of points is available, and is often in agreement with other calculations of dimension.

In the case of dimension estimation, however, the effect of low-amplitude noise is often not as significant as other effects. One might anticipate that the fractal scaling within a given volume will break down at length scales equal to the noise amplitude. However unless the system amplifies noise excessively, one does not expect the scaling to be affected at length scales much larger than the noise amplitude. Although noise is amplified along the expansion directions of a chaotic attractor, this effect does not have much influence on the dimension estimation because the noise is amplified back onto the attractor. In other words, the noise is drawn to the attractor and consequently has little effect on the scaling. Thus at relatively low SNRs there is still a good range over which a fractal may be scaled.

Takens et al. [123] propose a simple method to estimate the correlation dimension of a noise-infected attractor. The method uses the fact that the noise induces a bias in the observed differences of orbits, which tend to appear further apart than they are. A correlation integral function is determined that allows for this effect of noise, which is also assumed to be strictly bounded in magnitude. This function is based on a rescaling of the interpoint distances on the attractor. The assumption is that noisy orbits on the attractor stay within the proximity of the uncorrupted orbits of the underlying (chaotic) dynamical system. The rescaled correlation integral may be used to approximate the correlation dimension of the underlying chaotic system. The noise is concurrently estimated, and in general this will be underestimated.

Argyris et al. [5] investigated the effects of noise on the correlation dimension D_2 of chaotic attractors using a broad range of well known discrete and continuous time dynamical systems. They initially chose small values for all parameters and increased then gradually. Maximum values for parameters were chosen so the attractor retained its geometric structure and a finite correlation dimension. Numerical results revealed the presence of noise inflates the correlation dimension.

For dynamical noise, the geometric structure of an attractor is unstable in the sense that as the values of parameters are increased, the geometrical structure of the attractor vanishes.

2.10 Principal Component Analysis for Signal Noise Separation

Later we will approximate the chaotic dynamics by deconstructing it into a set of shadow-UPOs, each of which is approximately a cycle and is amenable to linear noise reduction techniques. Broadly speaking, there are two main techniques for separating signal from noise for a linear system, PCA and Fourier -based filters.

Principal Component Analysis (PCA) is a widely used technique to reduce dimensionality of a data set and arises across multiple disciplines like engineering, signal analysis, mathematics, physics and statistics, and in each case usually has a different name. The various statistical methods like Principal Component Analysis (PCA), Independent Components Analysis (ICA) and Kernel Principal Components Analysis (KPCA) all refer to techniques or algorithms that utilise singular value decomposition (SVD) as part of their method. The distinction here is that in many cases SVD is a single step in a more complicated process. In engineering these SVD method is referred to as Principal Orthogonal Decomposition (POD). The method is also known as Singular Spectrum Analysis.

It is also important to note the distinction between PCA as an attractor reconstruction method and as a noise reduction tool. All subsequent discussion concerning PCA will be in the context a method to reduce noise from a signal. Also, to avoid confusion, we will refer to PCA as the noise reduction method, noting that it uses SVD as part of the process.

Principal Component Analysis (PCA) reduces a large set of variables to a far smaller set, whilst retaining as much information as possible. The assumption is that whilst all of the original variables might be present in the expressions for prin-

cipal components, the number of principal components required to represent the data may be less than the number of original variables. PCA transforms a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables referred to as principal components. The first principal component accounts for as much of the variability as possible (i.e., has the greatest possible variance). Each subsequent principal component has the highest variance possible with the requirement that it is orthogonal to the preceding components. For a set of n observations in p variables, the number of unique principal components is $\min(n-1, p)$.

The principal directions are defined by the eigenvectors of the covariance matrix and form an uncorrelated orthogonal basis set. They are ranked using their eigenvalues with a higher eigenvalue denoting more explained variance. The method is sensitive to the relative scaling of the original variables.

The SVD (Singular Value Decomposition) deconstructs a matrix Z into the form

$$Z = S\Sigma C^T, \tag{2.20}$$

where S is a orthogonal $N \times n$ matrix of eigenvectors of the covariance matrix XX^T, with $N \gg n$ (in this context), C is a orthogonal $n \times n$ matrix of eigenvectors of $Z^T Z$ and, Σ is the $n \times n$ diagonal matrix consisting of singular values i.e. $diag(\sigma_1, \sigma_1, ..., \sigma_n)$, where entries $\sigma_1 \geq \sigma_1 \geq ... \geq \sigma_n \geq 0$.

For noise-reduction of the discrete measured time series $\{v_i\}$ with $i = 1, 2, ...N_T$, where N_T is the total number of data points, we generate a series of vectors

$$\{z_i \, \epsilon \, \mathbb{R}^n \; : i = 1, 2, ..., N\}, \tag{2.21}$$

and use these to construct a trajectory matrix X as follows:

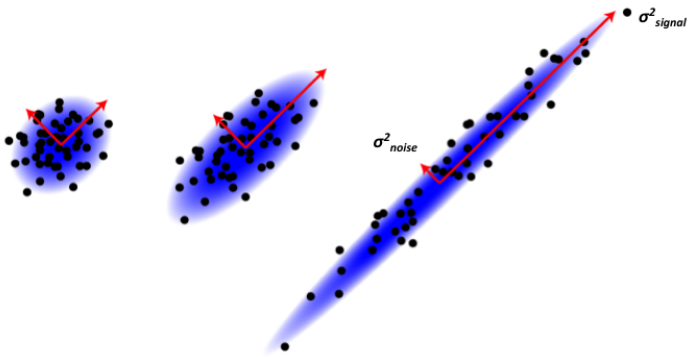


Figure 2.3: PCA: A dimension reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information from the larger set. We can use this method to separate noise from the signal.

$$Z = \begin{bmatrix} z_1^T \\ z_2^T \\ \vdots \\ z_N^T \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \\ v_2 & v_3 & \dots & v_{n+1} \\ \vdots & \vdots & \dots & \vdots \\ v_N & v_{N+1} & \dots & v_{N+n-1} \end{bmatrix}, \tag{2.22}$$

where $N = N_T - (n-1)$ is the embedding dimension. The trajectory matrix is just the Takens embedding in phase space with $\tau = 1$. If we write $ZC = S\Sigma$, we can interpret the orbit as exploring an n-dimensional ellipsoid, where $\{c_i\}$ represent directions and the $\{\sigma_i\}$ represent lengths of the principal axes of the ellipsoid.

Noise causes all the singular values of the trajectory matrix Z to be non-zero. If the noise is white, all the singular values $\overline{\sigma}_i^2$ of the noise-free signal will be uniformly shifted:

$$\sigma_i^2 = \overline{\sigma}_i^2 + \sigma_{noise}^2 \qquad , i = 1, 2, ..., k \tag{2.23}$$

$$\sigma_{k+1}^2 = ... = \sigma_n^2 = \sigma_{noise}^2, \tag{2.24}$$

where σ_{noise}^2 are the singular values of the noise floor. We can in this case write the trajectory matrix Z as:

$$Z = \overline{Z} + NZ = [S_1 \quad S_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} C_1^T \\ C_2^T \end{bmatrix}, \tag{2.25}$$

where $\overline{Z}$ is the deterministic part of Z, NZ is the noise-dominated part, $S_1 \epsilon \mathbb{R}^{N \times k}$, $S_2 \epsilon \mathbb{R}^{k \times k}$ and $C_1 \epsilon \mathbb{R}^{n \times k}$.

To separate the trajectory matrix Z from the noise-dominated part, one can estimate the deterministic part $\overline{Z}$ by using least squares or the minimum variance estimate. The least squares estimate of $\overline{Z}$ is given by

$$\overline{Z_e} = S_1 \Sigma_1 C_1^T, \tag{2.26}$$

and the minimum variance estimate is given as

$$\overline{Z_e} = \frac{S_1}{\Sigma_1}(\Sigma_1^2 - \sigma_{noise}^2 I_k) C_1^T, \tag{2.27}$$

where I_k is the $k \times k$ identity matrix.

2.10.1 Relation to Fourier Analysis

PCA process has similarities with Fourier analysis. As with PCA, Fourier analysis involves expansion of the original data in an orthogonal basis:

$$x_{pq} = \sum c_{pk} e^{i.2\pi qk/m}. \tag{2.28}$$

If we normalise the vector $\{e^{i.2\pi qk/m}\}\pi$ and name it v'_k, then:

$$x_{pq} = \sum b_{pk} v'_{qk} = \sum u_{pk} \prime s'_k v'_{qk}, \tag{2.29}$$

which generates the matrix equation $X = U'S'V'^T$, which is similar to Equation (2.20) above. The key difference is that whilst the set $\{v'_k\}$ forms an orthonormal basis, the set $\{u'_k\}$ does not. The PCA process, of filtering by concentrating on the largest singular values, is similar to the low-pass cut-off approach of Fourier analysis.

We will be using the Fourier filter to reduce noise from detected cycles in Chapters 3 to 5. This Fourier technique could be replaced by PCA or supplemented with PCA to reduce low levels of remaining residual noise. We prefer the Fourier approach as it allows flexibility to target specific frequency ranges and also to utilise signal processing strategies designed for specific types of noise.

2.11 Review of Noise-Reduction Methods

We now look at directly separating noise from a time series. This includes techniques specifically modelling and extracting the noise, exploiting the dynamical aspects by extracting orbits, projective maps that separate signal and noise and the fitting global functions to parts or all of the dynamics. Also, and importantly, we discuss techniques that utilise the UPO approximation of a chaotic time series.

The following diagram shows a useful taxonomy of methods.

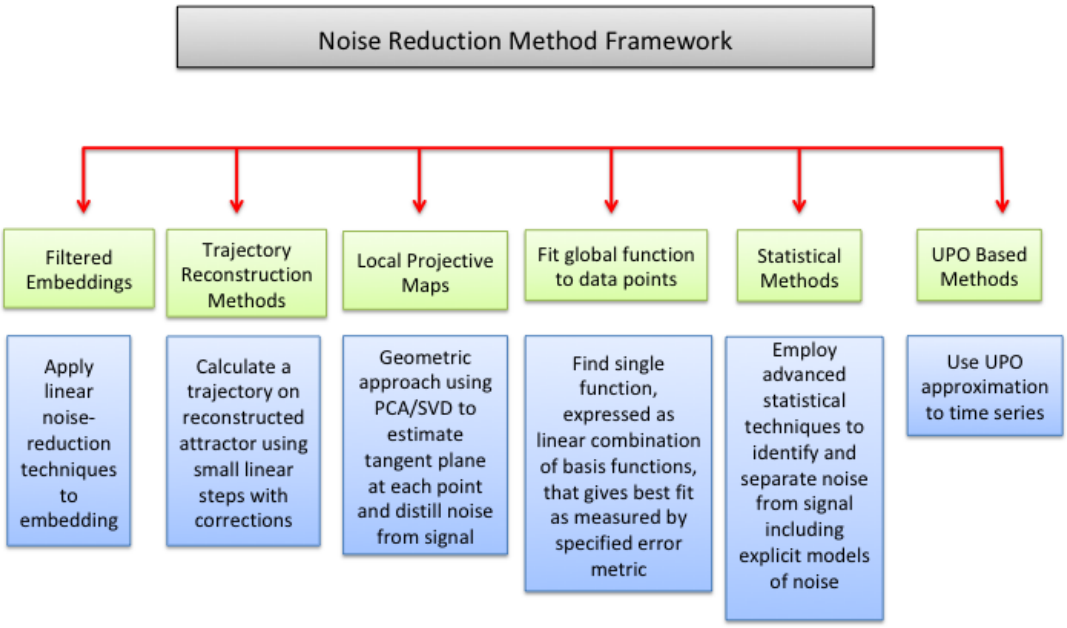


Figure 2.4: Taxonomical chart, broadly grouping noise-reduction methods that follow the same approach.

All these approaches have their advantages and limitations, and it is fair to say most work well with well-understood classic model chaotic systems with added

Gaussian noise. The real challenge is to address real life systems with higher noise levels and different types of noise. Thus there is a need for robust techniques that counter the many difficulties highlighted previously and yield a good attractor characterisation.

The key reference in this area, summarising noise-reduction methods was presented by Kostelich and Schreiber [83] in 1993 and this excellent material is heavily referenced for the following summary.

The emphasis of this literature review is pre-embedded techniques, trajectory-based methods and local projective maps as they form the core foundation of research to date. The common thread running through all these methods is that they seek either to employ an optimal linearisation to sections of the data where a linearisation might accurately represent the signal, or to deconstruct the signal into linear subspaces so the signal will separate from the noise as best as is possible. These are common techniques, used frequently in applied mathematics.

We shall not consider advanced statistical techniques as they more appropriately fall into the scope of statisticians. The fitting of global approximants is touched on briefly as there is limited research in this area. Finally we look at the very limited research into noise-reduction in chaos using UPOs, and recent developments in the field.

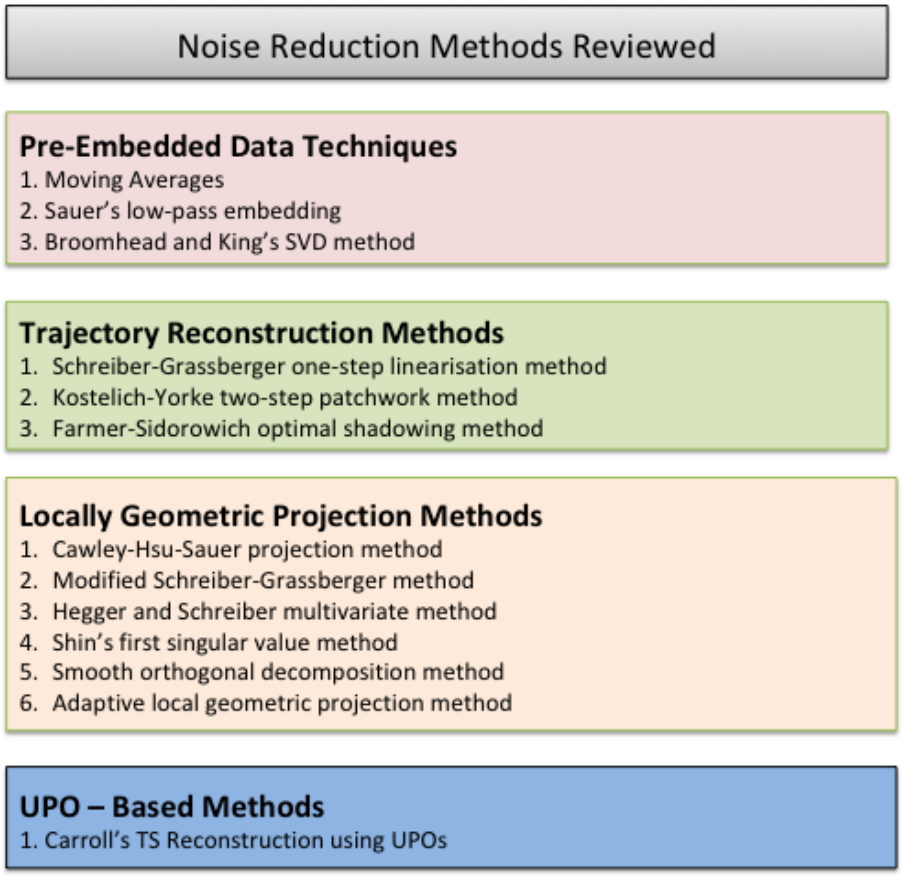


Figure 2.5: Summary of the noise-reduction techniques reviewed.

2.11.1 Pre-Embedded Data Methods

2.11.1.1 Moving Averages

The moving average filter is a simple low pass finite impulse response filter (FIR) and thus does not destroy the underlying chaotic dynamics. Whilst being optimal in the time domain in the sense of smoothing out white noise, it does have limitations and is not an end solution in itself. The moving average is an exceptionally good smoothing filter (output in the time domain), but an exceptionally bad low-pass filter (output in the frequency domain). It is the worst filter for signals with

a broadband profile in the frequency domain as there is no means to distinguish one band of frequencies from another [131]. The moving average is highly effective if our goal is to smooth out the fluctuations uniformly on the signal, to facilitate the easier application of a numerical method requiring smoothness. We are mostly limited to short-length centered moving averages. However, as a noise reduction method it is left wanting unless noise has a mean of zero. The smoothing inherent in this transformation assists greatly as a pre-processing tool, particularly in locating and identifying UPOs. It is an important part of the noise-reduction method presented later in this thesis and will be discussed in Chapter 3. The following two methods are focused on filtering the embedded data.

2.11.1.2 Sauer's (Low-Pass Embedding) Method

Sauer [118] designed a four-step numerical iterative algorithm to reduce noise from a discretely sampled input signal. It assumes that noise is additive and that the user has an estimate of the noise. The first step of the method is to smooth data locally into meaningful neighbourhoods, and this is achieved using a Fourier transform directly with a selected window size of w. The highest $\frac{1}{2}n$ frequency contributions are attenuated (set to zero) for a selected even number n. The remaining smoothed data set is subjected to the inverse FFT. This provides a smoothed version of the windowed section of the signal. The idea here being to reduce noise sufficiently to justify the subsequent embedding. The smoothed data is subsequently embedded in phase space and organised into neighbourhoods of size r, where r is a rough estimate of the magnitude of the noise.

SVD is then employed to calculate the principal directions of the data set. The principal directions connect a fixed anchor point, say the center of mass of the neighbourhood, to the embedded points. The remaining directions are eliminated, resulting in a noise-reduced signal.

Finally, to minimise the introduction of new correlations in the noise from the algorithm, the corrections are adjusted to ensure they add to zero. If the noise is uncorrelated with the signal, the random noise in each embedding coordinate will have mean zero. A multiple of the calculated correction is added to the raw data.

The multiple is a number between 0 and 1, is small for the first iteration through the data set, and is incremented by a fixed value for each pass as the data becomes more consistent with a deterministic process. The Sauer method is summarised as follows.

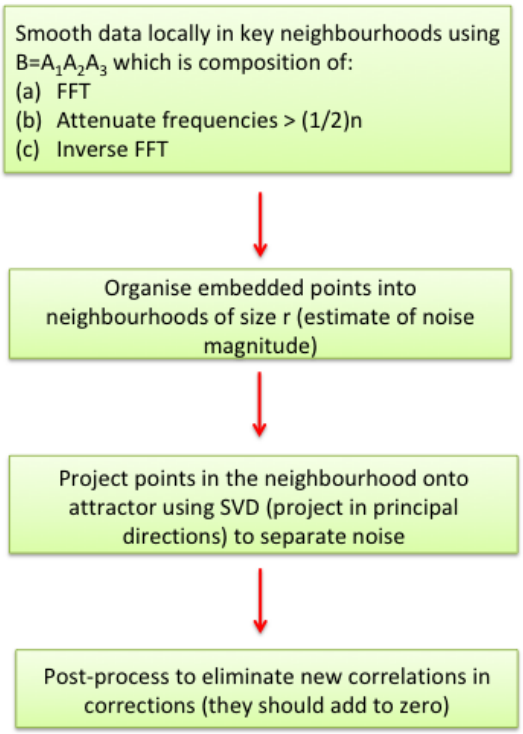


Figure 2.6: Summary of steps in Sauer low-pass embedding method.

Sauer thus utilised both a Fourier transformation and an SVD transformation; two of the most powerful tools in signal analysis. The method worked well as measured by the SNR with Rossler and Lorenz systems. The are several challenges with the application of the model. Firstly, finding the appropriate windows within the time series to apply low-pass Fourier filtering. Our SUNR method presented in this thesis, will take a similar approach, but we define our windows using detected complete and partial shadow-UPO sequences. Secondly, the initial low-pass filter will not specifically address the "type" of noise that is present. Also the residual

noise may be sufficiently high to invalidate the subsequent embedding. Finally, the process of grouping points into noise-based neighbourhoods and applying SVD may not eliminate any noise already removed using the Fourier low-pass filter.

The highlight of this work by Sauer is a very useful identity showing how the application of multiple linear operators to the time delay vectors will still produce an embedding. Define x_i as a point on attractor that describes the state of the system at time i. For measurement function $h(x)$, we have $y_i = h(x_i)$ is the i^{th} point of the time series. If we define the embedding map, $E : A \to \mathbb{R}^n$:

$$E(x_i) = [h(x_i), h(x_{i-1}), ...h(x_{i+n-1})] = (y_i, y_{i+1}, ...y_{i+n-1}). \tag{2.30}$$

Sauer et al. [118] have shown we are not restricted to E and can define a more general embedding using linear combinations of the input time series:

$$F(x_i) = P \begin{pmatrix} x_i \\ x_{i+1} \\ \vdots \\ x_{i+w+1} \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1w} \\ \vdots & \vdots & \vdots \\ p_{m1} & \cdots & p_{mw} \end{pmatrix} \begin{pmatrix} x_i \\ x_{i+1} \\ \vdots \\ x_{i+w+1} \end{pmatrix}, \tag{2.31}$$

P is a $m \times w$ matrix, which transforms the set of time-delay vectors. P produces an embedding provided the rank of P is sufficiently large and P does not collapse periodic points of the attractor of integral period less than equal to n. This insight allows linear transformations of delay vectors, but does not preclude more complex nonlinear transformations and in particular it does admit many of the linear methods used commonly today.

Assuming we are conducting the attractor reconstruction using the method of delays (noting most other methods can be shown to be mathematically equivalent), we may apply a linear filter P, which we know will preserve the embedding. This underpins the application of a range of signal processing techniques to the data, in particular the Fourier transform of the data.

2.11.1.3 Broomhead and King's Application of PCA

Broomhead and King [14] were arguably the first researchers to take the Principal Components Analysis (PCA) concept from signal analysis and apply it to non-linear dynamical systems. They structured the delay windows (or results taking n multiple samples of the time series) into a trajectory matrix. PCA was then used to reconstruct the attractor, with the added advantage of eliminating data redundancy and having a built-in noise filter. We explained the core concepts of finding the embedding dimension and separating noise from signal using principal axes in Section 2.10.

Interestingly, Grassberger et al. [60] have found that filtering based on basic SVD is about as efficient as Fourier-based Wiener filtering, and usually better than simple low-pass filtering. Broomhead et al. also employed the concept of taking multiple readings (multisampling), so one may exploit the redundancy in the data and employ Principal Components Analysis (PCA). Suppose we take p time series of measurements and the time series at time i is denoted by

$$x = \{y_i^1, y_i^2, ...y_i^p\}. \tag{2.32}$$

Each of the p time series contains the same dynamical information and similar amounts of noise. The series with the largest variance has the smallest relative noise level (highest SNR). We can employ the previous Sauer identity stating that linear combinations of the delay vectors will preserve an embedding. Almost every nonsingular linear combination Q_i contains the same information about the dynamics, where

$$Q_i = \sum_{k=1}^{p} b_k y_i^k. \tag{2.33}$$

One can select the Q_i with the highest SNR ratio. This can be achieved by maximising the variance of Q_i subject to the constraint $\sum_{k=1}^{p} b_k y_i^k = 1$. The required vector is the eigenvector corresponding to the largest eigenvalue of the

$p \times p$ covariance matrix C whose $(k,l)^{th}$ component is

$$C_{kl} = \left\langle y^k y^l \right\rangle - \left\langle y^k \right\rangle \left\langle y^l \right\rangle, \tag{2.34}$$

where the angular brackets denote the average value over all time steps. The method [14] was illustrated on the Lorenz system without specific testing against a range of noise types or levels.

2.11.2 Trajectory-Based Reconstruction Methods

The methods described in the previous section are in general linear methods applied to sections of the time series that are restricted in either the time or frequency domain. They do not exploit the underlying dynamical behaviour to identify noise perturbations and extract them from the underlying deterministic orbits. The sequence of elements in the time series provides valuable time-ordering information that is desirable to utilise if at all possible. Trajectory-based techniques follow the evolution of the orbit in time, averaging forwards and backwards in time to smooth small segments and reduce the noise.

2.11.2.1 Schreiber - Grassberger One-Step Linearisation Method

Schreiber and Grassberger [125] designed a method where past and future observations are used to update or correct one or more observations in the middle, with the goal of avoiding the drifting phenomenon. They define the dynamical relationship between past and future values (including noise) through the implicit relationship

$$F(y_1, y_2, ...y_{m+1}) + \eta_{m+1} = 0. \tag{2.35}$$

They then take a linear approximation to F to find a least-squares estimate of the value in the centre of the sequence

$$y^*_{(m/2)} = \sum_{k=1,\, k\neq m/2}^{m+1} b_k y_k + c, \tag{2.36}$$

The coefficient $b_{(m/2)}$ is excluded to prevent a trivial fit.

To execute this method, given a time series $y_1, y_2, ..., y_{m+1}$, one locates several closely matching sequences of $m + 1$ observations. Write the middle value in each sequence as a linear combination of the other entries in the time series as per Equation (2.36). The linear coefficients (except $b_{m/2}$) are determined using a least-squares fit. In a similar way, use $y_2, ..., y_{m+1}, y_{m+2}$, to determine a new set of coefficients to estimate $y^*_{(m/2+1)}$, and so on to the end of the time series. This process adjusts all but the first and last $m/2$ values. The output is a less noisy time series if the linearisation is an accurate representation of the dynamics.

Schreiber [124] has presented a simpler version of this method, which yields good results for short and noisy data sets. Essentially, the linear approximation in Equation (2.36) is replaced with a constant term, which can be determined with much less data. Only the central coordinate in the delay window is corrected as only this coordinate is optimally controlled from past and future (its value is fixed along both stable and unstable manifolds). The key optimiser in this approach is the size of the neighbourhoods, ε. For the examples given, a value of ε of about three times the amplitude σ of the noise works best. This method has the present coordinate x_i replaced by its mean value in N_i^ε:

$$x_i^{corr} = \frac{1}{N_i^\varepsilon} \sum_i^\varepsilon x_j, \tag{2.37}$$

where N_i^ε is the number of neighbours of x_i for which

$$sup\{|x_{j-k} - x_{i-k}|, ... |x_{j-l} - x_{i+l}|\} \equiv \|x_j - x_l\|_{\text{sup}} < \varepsilon. \tag{2.38}$$

That is, all segments of the orbits, which are close during a time lasting from k iterations in the past to l iterations in the future, are employed in Equation (2.37).

The experimental studies were presented using the Henon Map and the Mackay-Glass equation, both with added Gaussian noise. Results showed that the method works best with moderate amounts of data and noise levels above 1%. This method has formed the basis or motivation for many of the more advanced local projection methods we shall discuss.

2.11.2.2 Kostelich-Yorke Two-Step Method

Kostelich and Yorke [84] leverage the delay-vector reconstruction method to reduce noise. The evolving path on the chaotic attractor in phase space contains non-local signal information. Their two-step iterative algorithm detects and corrects errors in orbits resulting from noise. The method integrates a dynamical learning technique with a least-squares orbit adjustment technique to create a modified orbit.

The process firstly involves constructing a small ball around each point in the time series. The subsequent evolution of the set of points encased in each ball is then followed. A linear approximation to these dynamics is computed using the *Eckmann-Ruelle linearisation* technique. This technique is shown in Figure 2.7, referenced from [85]. Since we are taking discrete samples of the original signal, we can consider points on the reconstructed attractor as iterates of an unknown nonlinear map f. Assume the dynamics are given by

$$x_{n+1} = f(x_n), \tag{2.39}$$

where f is an unknown function, which is assumed to be at least piecewise differentiable. Eckmann and Ruelle [43] propose the local linear approximation of f

$$f(x_{ref}) = Ax_{ref} + b, \tag{2.40}$$

that is computed using least squares. Here A is a $m \times m$ matrix approximating the Jacobian matrix of partial derivatives of f evaluated at x_{ref} and b is an m-vector. Suppose x_{ref} has k neighbours within a suitably small neighbourhood

U. Linear regression is used to find the matrix A and vector b that minimize the sum of squares

$$\sum_j \|x_{j+1} - (Ax_j + b)\|^2 , \tag{2.41}$$

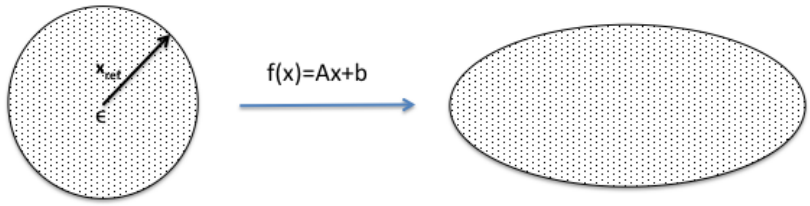
where the sum runs over all indices j such that $x_j \epsilon U$. A different A and b are computed for each neighbourhood on the reconstructed attractor.

To avoid the drift problem explained above, the linearisation must be followed by a second step to adjust the orbit to be more consistent with the dynamics. Trajectories are now adjusted to ensure distances are small between each point and its pre-image, each fitted point and the original observation, and the fitted point and its image. The desired orbit is thus $\{x^*_{i+k}\}_{k=0}^{p}$ that minimizes the sum of squares

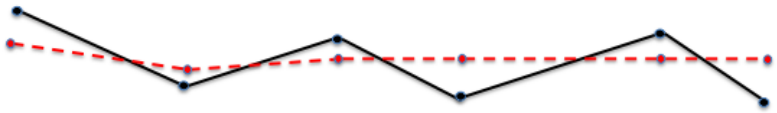
$$S = \sum_{k=0}^{p} \left\|x^*_{i+k} - f(x^*_{i+k-1})\right\|^2 + w \left\|x^*_{i+k} - x_{i+k}\right\|^2 + \left\|x^*_{i+k+1} - f(x^*_{i+k})\right\|^2 , \tag{2.42}$$

where f denotes the estimated dynamics at each point.

This outputs a sequence of scalar values. The distances between each point and its image are weighted twice as heavily as the distance to the original observation when $w = 1$. Larger values of w can be utilised when the input contains large isolated errors so the orbit is not moved excessively.



First stage of Kostelich-Yorke method. A collection of points on an ϵ -ball about the reference point x_{ref} is used to find a the Eckmann-Ruelle linear approximation there. Points on the attractor are used post-reconstruction, rather than the time-series.



Second stage of Kostelich-Yorke method. The (black) trajectory defined by the sequence $\{x_i\}$ is adjusted to a new trajectory (red) which is more consistent with the dynamics.

Figure 2.7: Summary of steps in two-step Kostelich Yorke Method.

As with all numerical methods this has some limitations. In relation to the Eckmann-Ruelle linearisation, there are several difficulties which have been considered in detail by Kostelich in [82]. Two important considerations are as follows. The accuracy of the approximation depends on how well a linear map describes the dynamics and some regions of the attractor may contain few observations. A larger ball size increases the number of available points but makes the nonlinearities more prominent. This situation becomes more significant as the dimension of the attractor increases. Also the linear least squares process produces biased estimates of the matrix A and vector b because errors in measurement are present in all the observations; the "errors in variables" problem. In relation to the correction step, the minimisation exercise can be complicated. In particular, the Jacobian

of f must be available. This step is more complex if f is not a piecewise linear function. Strategies for this case are presented by Davies [36].

Kostelich and Yorke tested their method on experimental data captured from the flow of noise-infected vortices in a Couette-Taylor experiment, where the attractor is a limit cycle. The frequency domain profile is comprised of a fundamental period and its various subharmonics above the noise floor. The two-step noise reduction method improved resolution of the limit cycle and the noise floor was reduced significantly. Power was conserved in all harmonics, and some harmonics that were were previously hidden by noise were revealed. The produced results superior to those obtained using low pass (Butterworth) filter. When applied to a Henon map with 1% added noise, it was found that 79% of that noise was removed.

2.11.2.3 Farmer-Sidorowich Optimal Shadowing Method

Hammel extended the original work of Anosov and Bowen relating to shadowing in Axiom-A dynamical systems, to non-hyperbolic systems. Hammel [65] then utilised the techniques employed in the proof of the shadowing lemma for a noise-reduction method. The constraint defining a shadowing orbit is replaced by a linear approximation.

For a specified noise-infected orbit y, we require an orbit x that minimises the L_2 distance between x and y, with the constraint $x_{t+1} = f(x_t)$ for $t = 1, 2, ..., N - 1$. This problem can be solved with the method of Lagrange multipliers and is equivalent to minimising:

$$K = \sum_{t=1}^{N} \|y_t - x_t\|^2 + 2 \sum_{t=1}^{N-1} [f(x_t) - x_{t+1}]^T \lambda_t, \tag{2.43}$$

where $\{\lambda_i\}$ are the Lagrange multipliers.

These equations are differentiated with respect to unknowns, and we search the resulting system of equations for an extremum. f is usually nonlinear so there is no closed form solution and a numerical method is required. Farmer and Sidorowich [47] use Newton's method, expanding f about a trial solution $\{x_t^*\}$:

$$f(x_t) \approx f(x_t^*) + f'(x_t^*)(x_t - x_t^*), \tag{2.44}$$

$$f'(x_t) \approx f'(x_t^*). \tag{2.45}$$

One can now formulate the problem more efficiently by defining the following expressions:

$\Delta_t = x_t^* - x_t =$ the deviation from the true orbit,

$\gamma_t = y_t - x_t^* =$ the deviation from the noise-infected orbit,

$\epsilon_t = x_{t+1}^* - f(x_t^*) =$ the deviation from determinism, and

$J_t = f'(x_t^*) =$ the Jacobian at x_t^*.

We seek a minimum of $\sum_{t=1}^{N} \|\gamma_t\|^2$, subject to the constraint that $\epsilon_t = 0$ for $t = 1, 2, ...N$.

Farmer and Sidorowich define an optimal solution, in the least-mean-squares sense, to the shadowing problem and applied it as a noise-reduction method. They employ a learning technique, and exploit the expanding and contracting behavior to achieve a noise reduction. The linearised equations are now:

$$\gamma_t = J_t^T \lambda_t - \lambda_{t-1} - \Delta_t, \tag{2.46}$$

$$\Delta_{t+1} = J_t \Delta_t + \epsilon_t. \tag{2.47}$$

This can be written in matrix form as follows:

$$\begin{pmatrix} -1 & J_1^T & \cdots & \cdots & \cdots & \cdots \\ J_1 & 0 & -1 & \cdots & \cdots & \cdots \\ \cdots & -1 & -1 & J_2^T & \cdots & \cdots \\ \cdots & \cdots & J_2 & 0 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & 0 & -1 \\ \cdots & \cdots & \cdots & \cdots & -1 & -1 \end{pmatrix} \begin{pmatrix} \Delta_1 \\ \lambda_1 \\ \Delta_2 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ -\epsilon_1 \\ \gamma_2 \\ -\epsilon_2 \\ \vdots \\ -\epsilon_N \end{pmatrix}. \tag{2.48}$$

If M denotes the $(2N-1)\times(2N-1)$ matrix in Equation (2.48), we can write this concisely as:

$$Mv = w. \tag{2.49}$$

When f is a d-dimensional dynamical system the elements of M are each $d \times d$ matrices.

The next step is to invert M, and solve for v. This yields an updated approximation of the true shadowing orbit $\{x_t\}$ that is superior to the test solution $\{x_t^*\}$, after iterating by Newton's method (replacing x_t^* by x_t after each iteration). For optimal noise reduction, N must be as large as possible. Unfortunately the structure of M is problematic as it becomes ill-conditioned for chaotic dynamical systems. Approximate homoclinic orbits further cause M to be nearly rank deficient.

The net result is that these factors make it difficult to invert M when N is large. The ill-conditioning problem is addressed by using the method of manifold decomposition. Manifold decomposition works by projecting the solution onto the stable and unstable manifolds. It requires that orbits are free of homoclinic tangencies. When the angular separation between stable and unstable manifolds is too small, this method fails. The problem with homoclinic orbits can be tackled using SVD - this algorithm is commonly used for inverting nearly rank deficient matrices. It works well for short orbits with good noise reduction, even with relatively high levels of initial noise. The difficulty here is that using SVD to invert a $n \times n$ matrix is computationally expensive and for large n this process is slow.

Farmer and Sidorowich blend these two techniques in a hybrid method. One determines the stable and unstable manifolds at each point. A search is conducted looking for points that have small angles between the stable and unstable manifolds. Once these points are located, a SVD is carried out, inverting M for short orbits (usually < 50 points) centered on these points. Manifold decomposition is then carried out for the full time series. The resulting orbit is tested for

determinism. If not, then again perform the SVD on the places where it is not.

The method is tested on the Henon map with added noise and performs well. If the function f is known, this method can in theory reduce noise to within the margin of computer error over long orbits. If the function f is unknown and is to be learned from the data (i.e. from embedding and global function fits), the ability to reduce noise is constrained to the margin of error of the learning algorithm and the length of the time series. Initial noise levels need to be low for the method to converge. Results were favourable with the initial SNR as low as 50 dB.

Several features of this work are interesting and useful. Firstly, these authors identify the real issues that exist as a result of homoclinic tangencies; a common phenomenon now known to be present in most chaotic systems. Further, they utilise both SVD and manifold decomposition to address the problems. Clearly this method is computationally expensive, complex, requires user judgment and best suited for situations where one has knowledge of the underlying dynamical function f. Shadowing techniques in general, to the best of our knowledge, have only been tested on two-dimensional maps.

2.11.3 Locally Geometric Projection Methods

The more promising methods to date for chaotic systems with moderate levels of noise employ orthogonal projections on linear hyper-surfaces, followed by the application of linear mathematical methods. There has been a significant amount of research published in this area, most of it focusing on stepwise improvements over the generic PCA approach. The methods/algorithms discussed in this section are provided to illustrate key progress points in development of these methods.

The methods presented are designed for noise reduction with unknown dynamics and take a non-parametric approach. The only assumption on the underlying dynamics is smoothness. The algorithms are always based on the local analysis of the data, which are scalar time series. These methods can all be viewed within a common structure – the projection of neighbourhoods on suitable linear manifolds. The key points of differentiation between methods are explained in terms of the metric considered, neighbourhoods selected, and the goal pursued.

All methods start with the attractor reconstructed in phase space. It is a subset of a smooth manifold in an m-dimensional phase space, so one may estimate the local tangent plane at each point using singular value decomposition techniques. The central concept is that noise leaks out from any point on the chaotic attractor into dimensions higher than those of the attractor's tangent space at that point. By projecting the embedded data down from a higher dimensional space onto the lower-dimensional tangent subspace we attenuate some of the noise from the signal.

The creation of a coordinate system by diagonalising a covariance matrix is widely known concept in signal processing. A basis derived in this manner produces an optimum compression of information. To isolate points in a set of interest (i.e. a orbit in $\mathbb{R}^n$) to a given accuracy the PCA basis minimises the number of components to be specified, thus eliminating redundancy. For a given embedding dimension, the average error caused by projecting onto the first v basis vectors is minimised if only the first v singular vectors are retained. The standard basis when we implement the method of delays is $\mathbb{R}^n$. In contrast to PCA, this yields

the worst possible information compression since the projections of the orbit onto the basis vectors are all equal.

2.11.3.1 Single Framework for Projective Methods

Mera and Moran [103] have created a unified framework for projective noise reduction algorithms, relying on a theorem which shows how to define the linear subspace which best fits data in $\mathbb{R}^n$ with respect to a given metric.

Theorem 15 *(Mera and Moran [103])*
Let A be a $n \times n$ symmetric positive definite matrix and let $\mathcal{L}_p$ be the set of p-dimensional linear subspaces of $\mathbb{R}^n$. The orthogonal projection of $v \epsilon \mathbb{R}^n$ onto $T \epsilon \mathcal{L}_p$ with respect to metric $\delta_{A^{-1}}$ is given by:

$$P_T v \equiv \min_{w \epsilon T} v A^{-1} w^T. \tag{2.50}$$

We say that the linear subspace $T \epsilon \mathcal{L}_p$ is the best linear subspace in $\mathcal{L}_p$ for the points $\{Z_k, k = 1, 2, ..., N_v\} \epsilon \mathbb{R}$, with respect to the metric $\delta_{A^{-1}}$, if it minimises over all $T \epsilon \mathcal{L}_p$:

$$\Lambda(T) = \frac{1}{N_v} \sum_{i=1}^{NV} (Z_k - P_T Z_k)^t A^{-1} (Z_k - P_T Z_k) \tag{2.51}$$

This best linear subspace can be obtained through the eigenvectors of the $n \times n$ matrix:

$$M_{ZZ} = \frac{1}{N_m} \sum_{k=1}^{NV} Z_k Z_k^t \tag{2.52}$$

Definition 16 *A system of vectors $\{w_1, w_2, ...w_n\} \epsilon \mathbb{R}^n$ is called an orthonormal system of eigenvectors of M_{ZZ} in the metric δ_A if they satisfy:*

1. *There exist real numbers λ_i (i.e. eigenvalues) such that $M_{ZZ} w_i = \lambda_i A w_{i,} 1 \leq i \leq n$, and*

2. *$w_i^t A w_j = \delta_{ij}$, for all i,j.*

Theorem 17 *(Mera and Moran [103])*
Let $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n$ *be the eigenvalues of* M_{ZZ} *in the metric* δ_A*, let* $\{w_i, i = 1, ..., n\}$ *be the corresponding orthonormal eigenvectors and, for* $1 \leq p \leq n$*, let* B *be the* $n \times (n-p)$ *matrix whose columns are the last* $n-p$ *eigenvectors* $\{w_{p+1}, w_{p+2}, ...w_n\}$*. Then the best linear subspace in* $\mathcal{L}_p$ *with respect to the metric* $\delta_{A^{-1}}$ *is given by:*

$$T_p = span\{Aw_1, Aw_{2,}.....Aw_p\}, \tag{2.53}$$

$$P_{T_p}Z = (1 - ABB^t)Z,\ Z\,\epsilon\mathbb{R}^n. \tag{2.54}$$

2.11.3.2 The Cawley-Hsu-Sauer Method

Cawley and Hsu [18] and Sauer [118] suggested that noise in the observations may be reduced by projecting the observations onto the subspace spanned by a suitable collection of singular vectors at each point on the attractor. Both methods are very similar although the specific embedding methods used differ slightly. The method involves linear pre-filtering, local linearisation and the use of fitted maps. The embedding space is $\mathbb{R}^m$ for a suitable m and the L_2 metric is used. They take the m-dimensional identity matrix as the matrix A in Theorem 12.

If we let U_i be a neighbourhood of X_i^m and let $\langle X_i^m \rangle_{U_i}$ be the centre of mass of the points within U_i. The algorithm reduces the noise by projecting the data $Z_j = X_j^m - \langle X_i^m \rangle_{U_i}, X_j^m\,\epsilon\,U_i$ onto the best d-dimensional subspace T_d. The computed estimate of x_j is:

$$x_j' = \langle X_i^m \rangle_{U_i} + \alpha Z_j + (1-\alpha)P_{T_d}Z_j,\ \ 0 \leq \alpha < 1 \tag{2.55}$$

This procedure connects the noise-reduced time series to the original data through the term αZ_j, in line with the goal of a pointwise reduction. The authors justify this step as a means to soften the effects of rare statistical outliers. Since each point of the scalar time series appears as a component of m consecutive delay

vectors, the algorithm provides multiple estimations for almost all of the points of the scalar time series. This problem is addressed by averaging such estimations.

Grassberger et al. [60] calculated a mathematical relationship between the Cawley-Hsu-Sauer local linear projection method and the one-step Schreiber-Grassberger trajectory method after providing a formalism encompassing both methods. They concluded that the Cawley-Hsu-Sauer method is very efficient for oversampled time series, but does have a significant theoretical drawback. The first and last components of delay vectors are corrected; this is problematic due to instabilities along the stable and unstable manifolds. Davies [35] showed that the Cawley-Hsu-Sauer method is equivalent to noise reduction by the method of gradient descent.

2.11.3.3 Modified Schreiber-Grassberger Method

This is an extension of the Schreiber and Grassberger one-step trajectory-method discussed earlier, developed by Grassberger et al. [125] after comparing the Cawley-Hsu-Sauer approach with the Schreiber-Grassberger one-step method, where they showed the methods are closely related. This "optimised" method was constructed to combine the strengths of these two methods and avoid their disadvantages where possible. The key lesson from the analysis of these methods is that one should not make large corrections to the outer components of delay vectors and also not project by changing only a single coordinate.

Recall that after a previous embedding of the data in $\mathbb{R}^m$ for a suitable (odd) value of $m > 2d + 1$, the one-step algorithm proceeds to assign as the estimate of the central coordinate of each m-vector X_i^m, a linear function of its $m-1$ remaining coordinates. This linear function is estimated using a neighbourhood U_i of X_i^m and it outputs a single estimate of each data point of the scalar time series. The one-step method can be proven to be equivalent to the projection of X_i^m on the best subspace of dimension $m-1$ with respect to the metric $\delta_{A^{-1}}$, corresponding to a diagonal matrix A with all its entries equal to zero, except the central coordinate which is set equal to one. Since such a matrix is non-invertible, the authors take a diagonal matrix whose diagonal entries are all very small (they use 0.001) and

a single 1 in the central entry of the diagonal. The central coordinate plays a special role. If the coordinate computed as a linear function were to be the last coordinate, it would be ill-defined along the unstable manifold and a correction of the first coordinate would be ill-defined along the stable manifold. Only the central coordinate correctly takes into account information about both the past and the future.

The Grassberger et al. method [60] further modifies this algorithm, so instead of correcting just the central coordinate for each point X_i^m it corrects several more coordinates. How many of these will be corrected and the dimension of the projection subspaces are parameters of the algorithm since each data point of the scalar time series appears as a central component of several delay vectors, the average of such estimations gives the final estimate of the point. In the Mera-Moran framework, the metric $\delta_{A^{-1}}$ is given by a diagonal matrix A with very small entries, with the exception of the central coordinates which are equal to ones.

Testing was carried out the Lorenz, Rossler and Henon systems. For highly-sampled flows the performance of the method is shown to be as good as the Cawley-Hsu-Sauer limit. The optimal algorithm and the best parameters depend on the noise level, the sampling rate, and the length of the time series. The method is slower than conventional linear filters but gives superior results, particularly for low noise levels.

2.11.3.4 Hegger and Schreiber's Multivariate Method

Hegger and Schreiber [66] extended the Grassberger-Schreiber one-step method to multivariate time series. To capture the past information and the future estimate of x_i the algorithm operates in $\mathbb{R}^{3d}$ taking the time series:

$$Z_i = (X_{i-1}, X_i, X_{i+1})\ \epsilon\, \mathbb{R}^{3d},\ i = 2, ..., N-1. \tag{2.56}$$

The hypothesis is that the clean time series satisfies the following linear relationship for points z_j near to z_i:

$$x_j \sim B_i x_{j-1} + C_i x_{j+1} + d_i, \tag{2.57}$$

where B_i and C_i are $d \times d$ matrices and d_i is an d-dimensional vector, calculated by solving the optimisation problem:

$$\min_{B,C,D} \sum_{j:X_j \epsilon U} \|X_j - (B_i X_{j-1} + C_i X_{j+1} + d_i)\|^2, \tag{2.58}$$

where U_i is a neighbourhood of Z_i. Let $X_i^{corr} = B_i X_{j-1} + C_i X_{j+1} + d_i$ be the estimate given by the linear model described above. The algorithm takes as the estimate:

$$x_i' = (1-\alpha)X_i + \alpha X_i^{corr},\ 0 < \alpha \leq 1. \tag{2.59}$$

It can be proved that X_i^{corr} are the central coordinates of the orthogonal projection of Z_i on the best linear d-dimensional subspace with respect to the metric $\delta_{A^{-1}}$ where A is a diagonal matrix having all the entries almost null except for the d central coordinates, which are ones. The method was tested successfully on the Ikeda map and the Lorenz system. If multivariate data is available the method is superior to scalar methods applied to the single coordinates.

2.11.3.5 Shin's First Singular Value Method

Shin et al. [128] provide a simple iterative SVD method, which aims to improve the quality of an embedding reconstruction by addressing the noise issue. They point out that for the purposes of noise reduction (not reconstruction), one can maximise the signal to noise ratio by finding a way to use only the first singular value. This is important as attractors reconstructed using SVD methods may be degraded by the contribution of singular values that are only slightly greater than σ_{noise}. To achieve this, the first singular value must contain most of the energy of the deterministic signal, which will generally be the case when dealing with low-dimensional systems.

The method is simple in that only two parameters are required, the sampling rate and the embedding dimension. A sinusoidal signal is used as a case study, since the first singular value carries most of the energy of the signal. The singular values are calculated explicitly, and an expression for the energy carried by the first singular value is calculated, showing its dependence on sampling time and embedding dimension. Given an estimated embedding dimension, the optimal sampling frequency is deduced. For low dimensional systems, a sampling rate of approximately ten times the cut-off frequency is shown to be sufficient. Once it is ensured that the energy of the signal is compressed towards the first singular value, then that singular value alone can be used to estimate x'_i using Equations (2.26) or (2.27). This will maximize the SNR of the recovered signal. We then get

$$x'_{e_1} = \sigma_1 S_1 C_1^T, \tag{2.60}$$

$$x'_{e_1} = \Big(\frac{\sigma_1^2 - \sigma_{noise}^2}{\sigma_1}\Big) S_1 C_1^T, \tag{2.61}$$

where S_1 and C_1 are the first columns of the corresponding singular vectors in equation Equations (2.26) or (2.27).

This procedure can be considered as an optimal filtering as the singular vector associated with the largest singular value is also the FIR filter which maximises the output power. Thus Equations (2.60) or (2.61) may be interpreted as the optimal FIR filter. Since x'_{e_1} is only an estimate of the true deterministic part of x', the recovered signal is not noise free and several iterations are required. The First Singular Value Method is summarised in Figure 2.8.

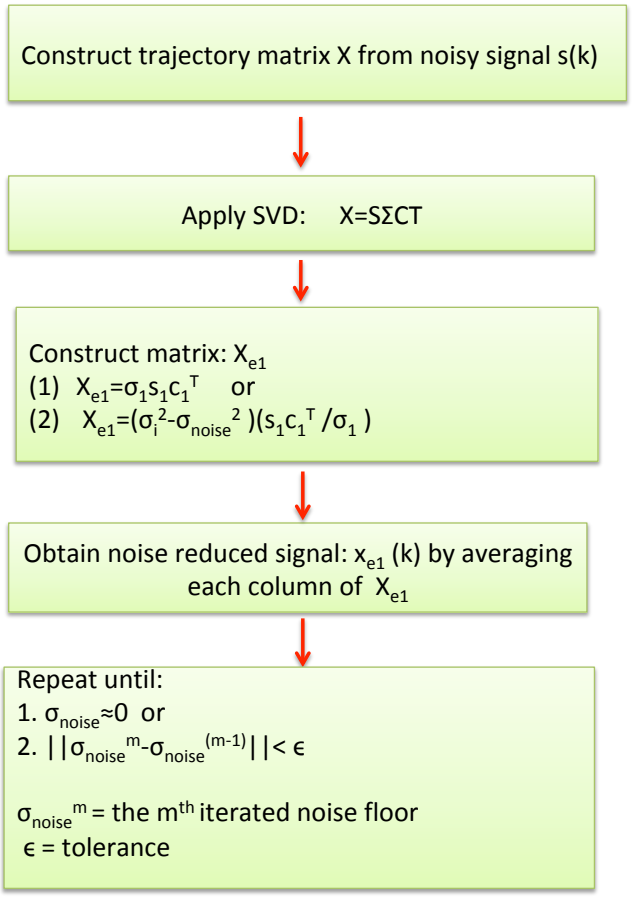


Figure 2.8: Summary of steps in Shin's First Singular Value Method.

2.11.3.6 Adaptive Local Geometric Projection

Leontitsis et al. [91] proposed an adaptive method to increase effectiveness of noise reduction using local geometric projection. This study focused on one aspect of local geometric projection, the selection of the signal subspace eigendirections, showing its significance and highlighting the fact that subsequent algorithms should build in this flexibility. The signal subspace is selected using the most significant eigendirections of a neighbourhood and the remaining ones define the noise subspace. A limitation of the other methods is that the choice of the number of largest, most significant eigenvectors (defining the signal-space) is fixed before the algorithm begins and remains constant for all local neighbourhoods, without accomodating their specific structure. The authors show that the number q of principal eigenvectors does vary by neighbourhood. The generic fixed value of q is based on a trial-and-error approach and a best guess. This adaptive approach allows q to vary from neighbourhood to neighbourhood.

A simple criterion is presented to distinguish the significant eigendirections using the maximum logarithmic difference between neighbourhood eigendirection lengths. It is assumed that at least one eigendirection corresponds to the noise subspace. This adaptive approach improves the algorithm.

The method is demonstrated using the Henon map, Lorenz flow and the Ikeda map with added noise at 10%—30% as well as the NASDAQ Composite index. Note the estimated noise in the NASDAQ was around 53%. Results were markedly superior to those using a fixed parameter in defining principal eigendirections.

2.11.4 Global Function Interpolation Models

Global model approaches seek to find a single function f^* which gives the best fit to f, where the sum of squares is taken over all the data. The most common approach is to express f^*as a linear combination of a set of k basis functions:

$$f^* = \sum_{j=1}^{k} \alpha_i \psi_j. \tag{2.62}$$

The coefficients α_j are selected to minimise the L_2 approximation error, which is a linear optimisation problem that can be solved by standard techniques. A popular choice of basis functions ψ_j are radial basis functions [67], [72]

$$\psi_j = \phi(\|c_j - x\|). \tag{2.63}$$

All basis functions have the same functional form and are distinguished only by the different center points c_j. A variety of choices are possible for $\phi(r)$ with examples including Gaussian functions, exponential functions, low-order polynomials, and rational functions. The coefficients are usually obtained by minimising

$$e^2 = \frac{1}{N} \sum [S_{i+m+1} - f^*(s_i, s_{i+1,.....s_{i+m}})]^2, \tag{2.64}$$

where the sum includes all the elements in the time series. The numerical method of choice for this least-squares problem is singular value decomposition (SVD).

The success of the method depends on the choice of the form of the basis functions. In addition, one needs a strategy for choosing the centers c_j and a criterion for deciding how many basis functions are required. More basis functions leads to closer approximations of the input time series. Of course the original map f is not known, so as more centers are chosen we model more details of the noise-infected data. When the number of basis functions equals the number of input points, we have an interpolation between all the observations: the interpolation is exact on the observations, the noise is interpreted as part of the dynamics, and the function f^* cannot be used for trajectory adjustment.

Although global function fits have some appeal, the choice of basis functions introduces some bias. For instance, the accuracy of the dynamical approximations at each point on the attractor depends in a nontrivial way on both the shape of the basis functions, the distribution of the centers, and the curvature of the orbits. The main advantage of global models is that they can provide stable fits even for small amounts of data. An excellent reference for a summary of global nonlinear approximants is by Aguirre and Letellier [3].

2.11.5 Noise Reduction using Unstable Periodic Orbits

There are few, if any, formalised noise-reduction techniques in the literature that utilise the UPO approximation of a time series. The only publication we can identify of significance in this area is that of Carroll [16]. Carroll's work here has served as motivation for our proposed SUNR (Shadow-UPO noise reduction) method.

Carroll used the set of detected UPOs to build an approximation to the chaotic time series by constructing all possible sequences of detected cycles and then selecting the best fit. This technique was presented in the context of chaos being used as a communication signal. Communication signals are subject to large amounts of measurement noise, and the purpose of the method was to produce an approximation that recovered high-level properties, for example the type of attractor, rather than detecting an exact copy of the noise-free signal. Carroll constructed an approximate skeleton of the attractor by stringing together the detected UPOs. This process is summarised in Figure 2.9.

Carroll's work is set in the context of communication signals, and in particular recognising an encrypted signal under the noise. In this case the central idea is that an algorithm that detects UPOs under the noise layer and reconstructs a time series according to a specified set of rules will act as a solid encryption methodology.

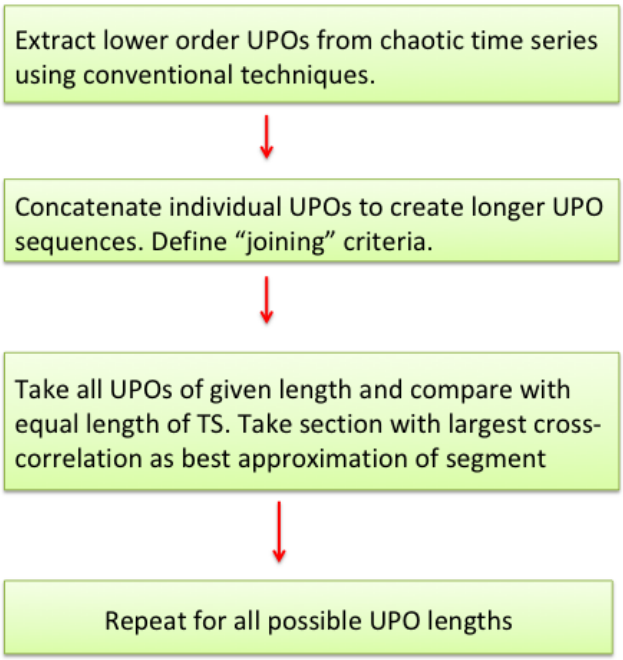


Figure 2.9: Summary of steps in Carroll's UPO-based method.

If we are to use this research in a slightly different context, as a basic platform to build upon to more accurately approximate a noise-free time series, then several improvements can be made:

1. Although Carroll's paper is framed in the context of noise-reduction it is mostly concerned with being able to "identify" a signal under the noise. No explicit reduction of noise is carried out. Shadow-UPOs detected from the time series are amenable to linear noise-reduction techniques. Further, specialist signal processing methods can be used that target different types of noise.

2. Carroll uses the method of close returns to detect shadow-UPOs, clearly aware of the resilience to noise of this method. The method can be modified to improve cycle detection rates for significantly higher levels of noise.

3. In the paper UPOs are concatenated in longer and longer sequences and tested for correlation with segments of the noise-infected time series to determine the best location of the approximating segment. This is not necessary as the actual location of individual UPOs in the time series can be recorded and stored. After shadow-UPOs are extracted and filtered, they can be returned to their original location in the time series.

4. Incomplete or partial cycles were not discussed given the original encryption context. However these appear in great numbers, contain valuable information and can be used.

5. Non-UPO segments of the time series are discarded. These contain valuable information about the dynamical system.

6. There are joining discontinuities that are problematic if the time series is to be subsequently used in numerical algorithms requiring smoothness, to calculate invariants. This problem is addressed by replacing cycles back their original location after filtering.

In Chapters 4 and 5 we present a new noise-reduction method (the SUNR method), which was motivated by this work. Features of the SUNR method include:

1. We detect shadow-UPOs using a modified recurrence matrix method. This enhances detection rates for all noise levels and detecting cycles at higher noise levels that otherwise would not be detected.

2. We reduce the noise on each cycle using a targeted strategy, depending on the type of noise.

3. We utilise complete and partial cycles (defined as $> 50\%$—80% of UPO period), which cover a high percentage of the time series.

4. We replace noise-reduced cycles back into their original location in the time series, reducing joining discontinuities.

5. We utilise the entire time series. Sections not filtered using complete or partial cycles are addressed with a simple centered moving average. The moving average is a FIR filter and does not corrupt the underlying signal and does not introduce a phase lag. It also only applies to a small section of the time series.

6. We use multi-sampling where possible to further enhance results.

2.12 Recent Developments in Noise Reduction

The rapid development in noise-reduction techniques for chaotic dynamical systems occurred primarily in the two decades between between 1980 and 2000. There has been a steady flow of research published since then, most of it as improvements and extensions of the core techniques developed during that period. We now briefly summarise several of the key developments since 2000.

2.12.1 Local Approximation using Polynomials

A novel variation on the local projective geometric method approach was published by Jafari et al. in 2012. The data is initially embedded in phase space, and the authors assert the method is insensitive to the choice of embedding parameters. The conventional local projection model involves projecting measured orbits onto the surface of a low dimensional attractor using PCA. The lower dimensional attractor contains the signal and not the noise and is approximated using a local tangent space. Typically a local tangent plane is fitted using a point and its near neighbours. This can be thought of as a first-order linear approximation.
Jafari et al. choose to fit an $n-$degree polynomial in the local tangent space rather than a plane. An adaptive process is prescribed that selects the value of n by taking into account the local stretching and folding of the attractor. The method requires that curvature is continuously calculated along the curve, which is problematic when noise is present. Multiple iterations are recommended to reduce this problem. Segments of the attractor with low curvature are modelled by polynomials of degree 1—3 and areas of high curvature are modelled with polynomials of degree 2—4. The curves are fitted using least squares. This enhances the effectiveness and

reduces the goodness of fit errors. Figure 2.10, from [71], illustrates the method, showing a section of noisy orbit approximated with a fitted polynomial.

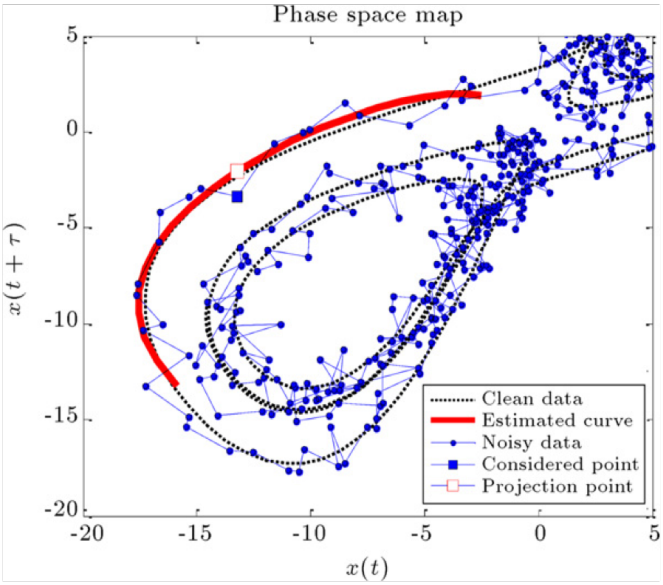


Figure 2.10: The red curve is a n-polynomial fitted using least squares with the degree n estimated based on curvature.

The method is tested on the Lorenz system and also on real sunspot data. For added noise levels of up to 20% the method results in a 50% improvement in the SNR. This method is an excellent variation of the local projective geometry method. In particular this method incorporates the curvature of the attractor. The limitations are around measuring the curvature accurately throughout the system evolution, fitting a polynomial that does reflect the true geometry and potential end-point discontinuities between piecewise fitted polynomials.

2.12.2 Wavelet Transforms for Noise Reduction

There has been developing interest in reducing noise from chaotic signals using wavelet transforms. At their most basic, wavelets are 'mini waves', existing for a finite time then smoothly dying away. A wavelet transform differs from the Fourier

transform in that the wavelet transform includes time as well as frequency information, unlike the Fourier transform which utilises only frequency information. This is very useful for non-stationary dynamical systems, like chaotic systems.

The Fourier Transform has a problem with resolution that may be likened to the Heisenberg uncertainty principle in physics. One can either accurately locate the frequency or the time of a signal, but not both. This dichotomy is illustrated in Figure 2.11, as shown in [30]. Decomposing a signal into wavelets instead of just frequencies can give superior resolution in the transformed domain. When a Wavelet Transform is applied, the signal is transformed into the wavelet domain, rather than the frequency domain.

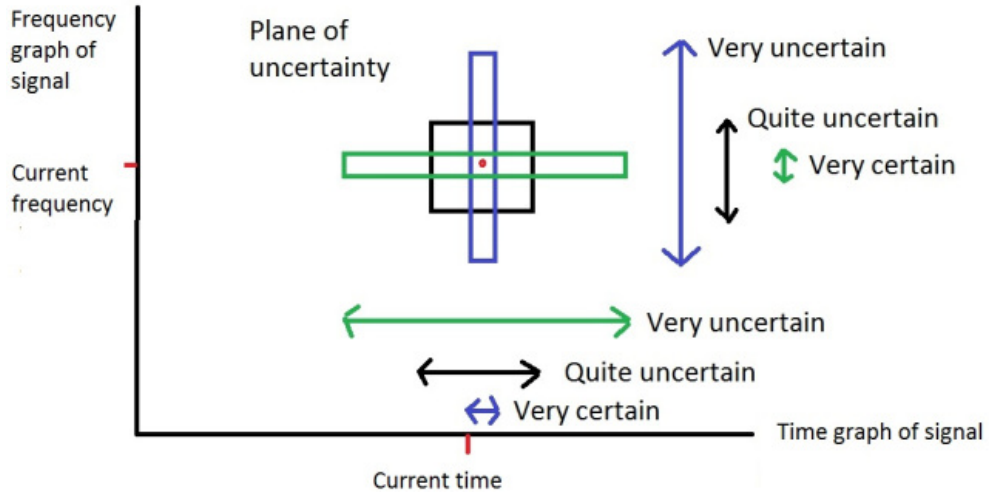


Figure 2.11: For non-stationary signals we may know the frequency or the time but not both.

The chaos noise-reduction problem using wavelets is formulated as follows. Let us model the chaotic attractor x_t using the relation:

$$x_t = z(x_{t-1}), \tag{2.65}$$

where z is a real function with a bounded support.

Assume the chaotic system is perturbed by noise and that as a result the observable data is no longer x_t but u_t and the observed attractor is no longer z but z^ϵ defined by:

$$u_t = z^\varepsilon(u_{t-1}). \tag{2.66}$$

We now seek a transformation on z^ε so as to get an estimate of z. The input signal is z^ϵ which we assume to be a squared-integrable function: $z^\varepsilon \, \varepsilon \, \mathcal{L}^2(\mathbb{R})$. We want to deconstruct z^ϵ using a discrete dyadic real wavelet basis. This wavelet basis is a countable subset of the function space $\mathcal{L}^2(\mathbb{R})$. The wavelet functions are defined as:

$$\psi_{j,k} : t \, \epsilon \, \mathbb{R} \longmapsto 2^{\frac{j}{2}} \psi(2^j t - k), \tag{2.67}$$

where $\psi \, \epsilon \, \mathcal{L}^2(\mathbb{R})$ is the real *mother wavelet*, $j \, \epsilon \, \mathbb{Z}$ is the *resolution level* and $k \, \epsilon \, \mathbb{Z}$ is the *translation parameter*. Popular mother wavelet types are the *Haar wavelet* and the *Daubechies wavelet*.

There is a well defined process for calculating wavelet coefficients $z^\varepsilon_{j,k}$ which are the projections of z^ε on the subspace generated by the vector $\psi_{j,k}$. The inverse wavelet transform consists of reconstructing the observed signal, z^ε, using each translation and resolution level. The reconstructed signal for a given resolution level j is called the *detail signal*:

$$D_j : t \, \epsilon \, \mathbb{R} \longmapsto \sum_{k \, \epsilon \, \mathbb{Z}} z^\varepsilon_{j,k} \psi_{j,k} \tag{2.68}$$

The entire reconstructed signal is the sum of the detail signals:

$$z^\varepsilon = \sum_{j \, \epsilon \, \mathbb{Z}} D_j \tag{2.69}$$

This is the wavelet noise-reduction framework and Garcin et al. [53], [54], equate the noise reduction problem to that of filtering the wavelet coefficients using a *thresholding function* ϕ. The aim of the thresholding function is to select the most significant vectors $\psi_{j,k}$ from the wavelet basis and eliminate the noisy

terms of the wavelet deconstruction, analogous to the PCA process. An expression for the probability density of the wavelet coefficients is provided by the authors. Interestingly, for dynamical noise, the expression is exact. A method is presented to filter the wavelet coefficients of a discrete dynamical system infected with weak noise, and henceforth construct estimates of the pure signal. They present the example of a logistic and a Lorenz chaotic dynamical system and test on financial data (oil prices and NOK/USD exchange rate), with good results.

Applying wavelets as a means of reducing noise from chaotic systems is an active research area. Results to date indicate the method is effective with lower levels of added noise (e.g. 10%) and has significant potential as a method to address dynamical noise.

2.12.3 Dynamical Noise: Kalman-Takens Filter

If a parametric, mathematical model is available for a time series, filtering of low-level dynamical noise may be possible using a Kalman-style filter. Although the original Kalman filter applies to linear systems, in recent years the model has been modified to accommodate non-linear model equations with close to optimal forecasting capability. These include in particular, the extended Kalman filter and the unscented Kalman filter, both which work on nonlinear systems.

The recursive algorithm runs in real time and is a two-step process, comprising a prediction phase and an update phase. During the first predictive stage the current values of state variables and their uncertainties are estimated using the Kalman filter. The subsequent (noise-infected) values of the state variables are then observed and used to revise the original estimates. The revision uses a weighted average with the weights structured to favour estimates with higher certainty. It is not necessary to assume the errors have a Gaussian distribution when applying the Kalman filter. However, should the errors be normally-distributed the filter will output the exact conditional probability estimate.

The difficulty is that in many cases a parametric model will not be known for an experimental time series. Recently, a method has been introduced that

combines Takens' non-parametric attractor reconstruction with Kalman filtering [63]. The dynamics are reconstructed non-parametrically using delay vectors and substituted for the parametric model. The Kalman-Takens' algorithm has been shown to filter additive noise-infected data with results comparable in performance to parametric filtering techniques that utilise a mathematical model. Hamilton, Berry and Sauer [64] studied the effect of the Kalman-Takens filter in the presence of dynamical noise and found the performance of the method to be almost as good as that achieved using the full parametric model.

Although one might expect the presence of dynamical noise to hamper the attractor reconstruction using delay coordinates, Hamilton et al. did not find it prohibitive. There is still work to be done on optimisation of the algorithm. This technique is designed to be applied to the entire time series, but is unlikely to be applicable to individual detected low-period shadow-UPOs unless high frequency sampling is used. The method is also dependent on the estimation of the noise covariance matrices required for the Kalman Filter: the noise covariance matrix Q and the measurement noise covariance matrix R. These quantify the amount of noise in the system. Process noise refers to the dynamical noise in the process. Q tells how much variance and covariance there is. The diagonal of Q contains the variance of each state variable, and off diagonal entries contain the covariance between the different state variables. R contains the variances of the measurements. Q is in state space and R is in measurement space and the Kalman filter matrix H converts these to a common space, and in nonlinear systems that must be linearised that in some manner. Finally, delay-coordinate embedding of a noise-infected time series is not guaranteed to be topologically equivalent to the noise-free time series and the low-noise assumption of equivalence is likely to weaken as the amount of noise increases.

2.12.4 Smooth Orthogonal Decomposition Method

Local projective noise reduction using SVD does not utilise the temporal characteristics of the time series. It only uses geometrical or topological properties of the data. Chelidze [20] has developed an extension of the local projective noise reduction method which accommodates both topological and temporal characteristics of the time series. Rather than using SVD to identify the tangent subspace of a chaotic attractor, Chelidze identifies a smooth subspace that locally embeds the attractor. This is achieved by using smooth orthogonal decomposition (SOD) of a bundle of nearest neighbour orbit strands. Temporal smoothness is imposed on the noise-reduced time series by confining orbits to these subspaces.
Arrange the data representing n simultaneous measurements of d state variables into a $n \times d$ matrix Y, with each column of Y having zero mean. The key idea is to find a linear coordinate transformation of this matrix Y:

$$Q = Y\Psi, \tag{2.70}$$

where the columns of $Q \epsilon \mathbb{R}^{n \times d}$ are new *smooth orthogonal coordinates* (SOCs), and *smooth projective modes* (SPMs) are given by the columns of $\Psi \epsilon \mathbb{R}^{d \times d}$.

The smooth orthogonal decomposition is obtained by formulating Equation (2.70) as an eigenvalue problem.

Chelidze tested the SOD method on model data using the Lorenz model and a double-well Duffing oscillator with added noise. A broad range of metrics were used, including power spectral densities, SNR's, correlation sum, short-term orbital divergence and visual inspection of reconstructions. The SOD method was compared with POD under a number of noise scenarios. POD performs well with low noise levels, but fails at higher noise levels and loses any trace of determinism at 80% noise.

2.12.5 Higher Order and Multi-scale Refinements

Recent research by Moore et al. [106] has focused on efficiency improvements in local projective methods, motivated by the second order refinement by Kantz and

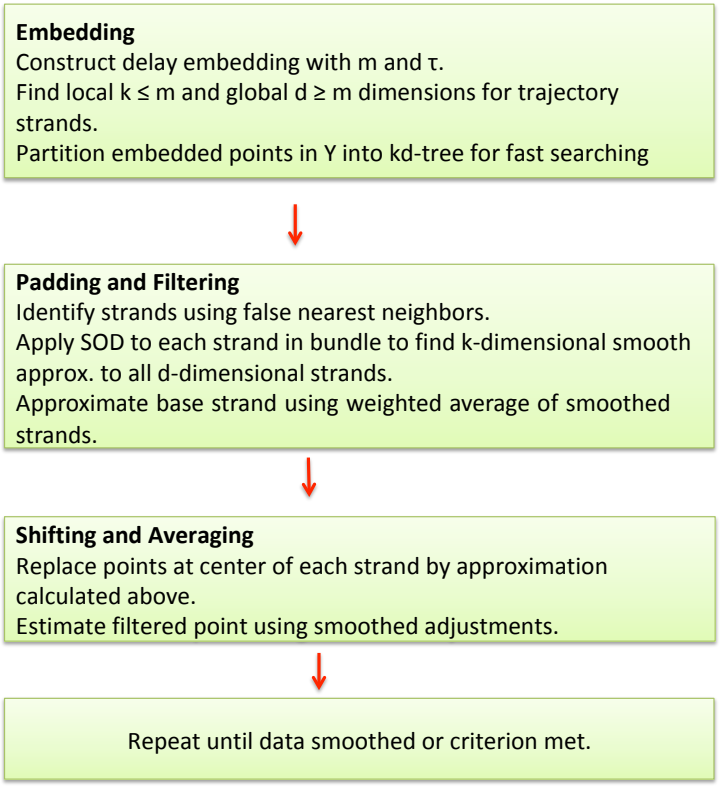


Figure 2.12: Summary of steps in the Smooth Orthogonal Decomposition Method. Note that the process of using orbit strands, and also that of delay coordinate embedding results in truncations in the matrix Y. These truncations cause unwanted edge effects and are remedied in the padding and filtering step of the algorithm.

Schreiber in 2004 [74]. Kantz and Schreiber observed (from the Sauer method) that the centres of mass of local neighbourhoods are not in general situated on the attractor. To more correctly incorporate attractor geometry, they introduced a second order correction, where first and second order centres of mass of local neighbourhoods in phase space are combined into a weighted sum. This is then used as the origin of the projective subspace.

Moore et al. extend the refinement approach of Kantz and Schreiber, developing a mathematical model using higher order and multi-scale noise filters. The authors assert that for low and moderate levels of added noise, a dual scale filter tends to outperform existing local projective noise filters. They measure the increase in SNR from 8 iterations of the dual scale filter and compare results with the methods of Sauer, Cawley-Hsu, and Kantz-Schreiber using the Henon and Ikeda maps as well as the Lorenz system. The method appears to be intended for long orbits with a relatively low sampling rates. Their initial finding was that extending these geometric considerations alone to higher order filters did not produce noise-reduction that was superior to that produced by existing methods. However, when the method was supplemented with statistical analysis to minimise attenuating error, it resulted in filters that perform better than existing methods. The findings are summarised as follows:

1. As noise increased from a low level (1%), higher order filters became relatively more effective.

2. As the noise level achieved moderate levels (30%), geometrically-based filters with high error attenuation begin to dominate in performance.

3. At the highest noise level considered (100%), conventional generic local projective filters were superior.

The results are interesting and show excellent noise reductions when the optimal order-scale scenario is selected.

2.13 Summary and Discussion

This chapter highlights the difficulties in removing noise from chaotic signals and summarises the central concepts of existing techniques. It is not a simple process and has resulted in a large canon of research. For practical purposes we broadly define noise as measurement noise or dynamical noise, the key difference being the recursive or feedback aspect of dynamical noise. The deeply nonlinear nature of chaos implies signal and noise are closely intertwined and difficult to separate by conventional linear or nonlinear techniques. The addition of noise to an evolving dynamical system can in itself lead to a transition to chaos. Noise and SDIC both radically alter the paths of individual orbits in a chaotic system. Fortunately we are provided with some protection via the Shadowing Lemma which guarantees the existence of a pseudo-orbit arbitrarily close to the true orbit, albeit the true path originating from a different starting point. Numerically detected orbits are in fact representative of the true dynamics of the system.

Filtering noise from chaotic systems is extremely tricky. Many filter types that involve a recursive or feedback element (IIR filters) actually interact with the signal itself, corrupting the results. We are severely limited in the selection of conventional filters to just finite impulse filters (FIR), which fortunately include moving averages. A further challenge is that the frequency domain power spectra of chaotic data looks a lot like broadband noise and ad hoc application of Fourier techniques that separate noise based of frequency considerations immediately run into trouble. Incorrect filtering will actually increase the correlation dimension and add an extra Lyapunov exponent to the system being studied. It has been demonstrated that the addition of small amounts of noise will render the common numerical algorithms to determine global structural invariants useless.

Whilst the interaction of filter with noise-infected signal is the most significant challenge, there is a raft of other issues to be overcome with chaotic systems. Included are curvature, drifting, the application of least squares with errors-in variables, ill conditioned least squares and outliers. There is also a range of errors arising from the use of measurement function in a practical situation. Finally,

there is a difficulty with embedding noise-infected data that is on par with the "Heisenberg uncertainty" type of noise-filter interaction discussed above. Takens theorem does not apply to digitised or noise-infected data and an embedding of such data cannot be guaranteed to be topologically equivalent to the real system.

We draw attention to the very useful Sauer identity stating that linear combinations of the delay vectors will preserve an embedding. Research by Grassberger et al. [60] comparing basic linear techniques for noise reduction reveals that filtering based on basic SVD is about as efficient as Fourier-based filtering, and usually better than simple low-pass filtering, another very useful result. We will need to apply linear transformations to orbits (detected shadow-UPO segments). We prefer Fourier techniques over PCA for noise-filtering due to the greater transparency involved (ability to view results in detail in the frequency domain) and the superiority of the FTT in allowing targeted solutions depending on the frequency domain power distribution of the noise.

These considerable complexities have resulted in the development of numerical techniques, specifically designed to reduce noise from chaotic data. We have provided a broad classification system of known noise-reduction algorithms based on the specific approach taken. Techniques applied to pre-embedded data vary from the simple application of moving averages to more sophisticated approaches like the method of Sauer, which applies a simple Fourier method and SVD to specific neighbourhoods. Most published noise-reduction techniques to date either follow step-wise along an orbit as time evolves or are local geometric projections on an appropriate subspace. Both techniques are seeking opportunities to locally apply linearisation or to locally separate signal and noise using a PCA technique. We have reviewed examples of each main type of key current noise-reduction technique. Many methods are highly complex, unavailable in Matlab form and for these reasons unlikely to be useful for non-mathematician experimenters.

The difficulty with trajectory models and local geometric projections is they start with an embedding. So do the numerical algorithms to calculate metric and dynamical invariants. We can reasonably assume that the Takens' conditions hold

approximately with low levels of noise and sufficiently sampled data. However, once measurement noise increases (say $> 10\%$), the validity of these methods becomes immediately compromised and questionable. Further, methods that step along an orbit run into scaling problems as noise becomes significant. We fully appreciate that in practical situations, with high noise levels, that any result may be preferable to no result. However, perhaps we can refine our approach to not rely on the embedding.

We propose that there is scope to develop improved noise-reduction techniques using UPOs. There has not been the vigorous development of techniques using UPOs, as was seen for trajectory and local geometric projection approaches. The key reference is that of Carroll [16], where a replacement time series is reconstructed using UPOs stitched together in a manner to optimise cross-correlation. This work is set in the context of communications and is a motivator of the method presented in this thesis.

An attractor can be approximated by the sum of lower order UPOs. In later chapters we will show how shadow-UPOs can be accurately detected from a highly noise-infected time series, and as these are "approximately periodic", the universe of filtering tools expands vastly to include most conventional signal processing techniques, including Fourier analysis and PCA. We can use these noise-reduced shadow-UPOs to estimate the true noise-free UPOs and approximate the noise-free time series. Also, the detection of shadow-UPOs does not necessarily require the use of an embedding, eliminating a major concern about the validity of the result. The time series can thus be represented by the filtered shadow-UPO sequences provided the coverage of the time series by UPOs is sufficiently large and the chaotic system is not too unstable. The approximated time series can be used in conventional numerical algorithms to determine invariants if residual noise is sufficiently low. If not, the approximation can be refined further using the existing techniques applicable to low noise levels. The detection of multiple copies of UPOs (to within a tolerance) allows averaging and provides as an output the set of noise-reduced lower order UPOs. The collection of detected noise-reduced shadow-UPOs lower order UPOs is valuable as the Lyapunov exponents of each UPO can be

directly estimated from these. Multiple sampling of the noisy time series, coupled with averaging can enhance these results significantly.

In the literature reviewed, sometimes a noise-reduction method has been tested on a very simple two-dimensional map such as the Henon map and in other cases a method has been tested on a more difficult flow such as the Lorenz system. Methods are seldomly tested on multiple chaotic systems. Often methods are tested for selected noise levels and the maximum noise levels that can be removed are not stated. As a result it is difficult in some cases to understand the noise-levels for which the method is effective and whether it performs better on some systems than others. These factors, and the complexity of many models, make it difficult to compare methods and appreciate which are most effective in practice.

Chapter 3

Detection of Cycles in Noise-Infected Chaotic Time Series

3.1 Chapter Overview

Before we can reduce the noise from the cycles comprising a chaotic dynamical system, we must first locate the cycles within the noise-infected time series. Chapter 4 and 5 describe the development of the Shadow-UPO Noise Reduction (SUNR) Method. One of the aims of this new method that we present is to detect, capture and filter all complete and partial cycles that can be detected within the constraints of the sampling regime. Naturally, we must assume the chaotic time series to be analysed is of sufficiently low instability to allow the evolution of orbits that remain in the vicinity of a UPO for at least one complete cycle or more; this is necessary to facilitate the analysis.

In this chapter we describe the construction of the recurrence plot and the factors affecting the information derived from it. There are two main methods in the literature for detecting cycles in a noise-free time series. Firstly, the Newton-Raphson type methods are highly effective at detecting cycles, but often require seeded starting points and are highly sensitive to small levels of noise. They also

usually require the careful use of Poincare sections which are sensitive to placement direction. These models are thus not applicable in the context of medium to high noise levels. The other cycle detection technique is uses *close returns plots* or *recurrence plots* and histograms derived from the time series data. Strictly speaking recurrence plots are close returns plots using embedded data. The original concept of recurrence plots was devised and introduced by Eckmann et al. [41] with the goal of detecting stationarity. However, both expressions are arguably now used interchangeably. For consistency we will subsequently use the terms "recurrence plots" or "recurrence matrices", even if we are using scalar data. This is our preferred approach as it searches for cyclic sequences of all periods at once and may accomodate small amounts of noise.

We illustrate this approach for the noise-free Rossler system, looking closely at the construction of the recurrence plot and the detection of cycles using a histogram constructed from the recurrence matrix. There are a number of control factors that determine the effectiveness of this technique. We examine each and consider limitations of this method. We next consider the effect of noise on recurrence plots and histograms and highlight that, although the histograms are somewhat resilient to noise, the underlying cycles are "broken" and the number of detected complete cycles falls off rapidly with increasing noise.

We modify the recurrence histogram method to accomodate high levels of measurement noise with the goal of maximising use of available information. We examine whether pre-embedding of the time series offers any advantages over use of the scalar time series, and pre-process the data to enhance the detection rates significantly. We also collect suitably long partial or incomplete cycles as in unstable systems these are numerous and contain valuable information about the system dynamics; these may be subsequently noise-filtered.

3.2 Nomenclature of Cycles

Attractors in chaotic regimes contain a dense set of UPOs, and unstable cycles of all periods are present. There will be a minimum, fundamental period that we will term T_1 .

Subsequent cycles will named in order of increasing period $T_2, T_3, ..., T_n$ where n cycles are detected. In some cases, for example the Rossler system, there will be a fundamental period T_1 and a series of sub-harmonics, with $T_k = k.T_1$ where k is an integer.

In other cases there will be a series of harmonics, revealed as peaks on the recurrence histogram, not necessarily defined as integral multiples of the fundamental period. In the Lorenz system, inspection of the symbolic dynamics reveals there are multiple geometric possibilities for periods of each cycle. Although there are multiple geometric possibilities, not all cycle types actually appear in simulations.

3.3 The Rossler Time Series

We now introduce the Rossler system [115] which we shall use as a heuristic example throughout this thesis to demonstrate all aspects of our noise reduction methodology (SUNR method). This system of three non-linear ordinary differential equations was originally created and studied by Otto Rossler. His intention was to construct a mathematical model that behaved similarly to the Lorenz attractor, but was easier to study. The Rossler system contains a single nonlinearity and is lightly unstable (maximum Lyapunov exponent $\lambda_{\max} = 0.079$ s^{-1}). The Rossler equations are presented in Equations 3.1a–3.1c. We build our model Rossler time series through numerical integration of these system equations using the 4^{th} order Runge-Kutta method. For our modelling purposes, we use system parameters of $(a, b, c) = (0.2, 0.2, 5.7)$, a time step of 0.1 and a starting point of $(x_0, y_0, z_0) = (-8.3029, -3.7887, 3.3094)$. We will be using 10,000 model points in the time series throughout this chapter, unless stated otherwise. The numerically

integrated noise-free Rossler attractor is shown in Figure 3.1. In Figure 3.2 we show the FFT spectrum of the Rossler time series, where we can see a well-defined peak at the fundamental period T, but also observe amplitudes in many other frequencies.

$$\frac{dx}{dt} = -y - z, \tag{3.1a}$$

$$\frac{dy}{dt} = x + ay, \tag{3.1b}$$

$$\frac{dz}{dt} = b + (x - c), \tag{3.1c}$$

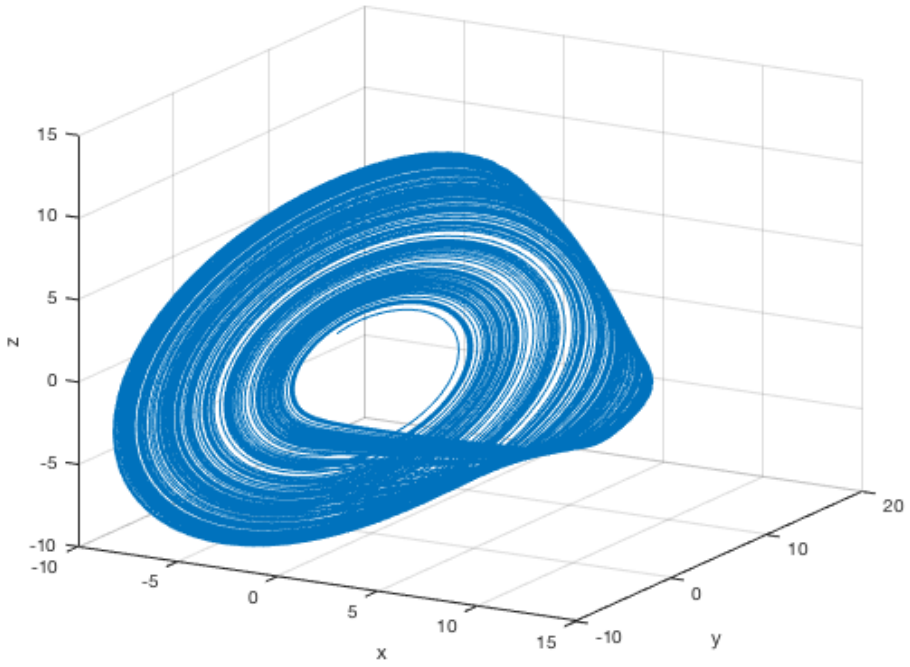


Figure 3.1: Noise-free Rossler attractor obtained by numerically integrating Equations (3.1a)–(3.1c) using the 4^{th} order Runge-Kutta method. It is constructed from 10,000 points, sampled at a time step of 0.1.

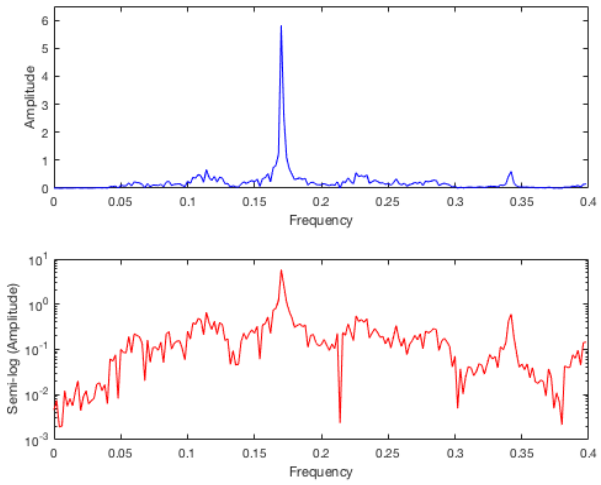


Figure 3.2: FFT amplitude spectrum of noise-free Rossler time series. The bottom figure is in semi-log format. Whilst there is a clear peak, we note the broadband nature of the signal.

3.4 Detecting Noise-Free Shadow-UPOs

3.4.1 Recurrence Matrices

The phase space orbit represented by a scalar time series evolves near a given UPO for a time, orbiting (or "shadowing") around it before being repelled into another region of the attractor. These shadowing orbits are referred to as *pseudo-UPO* orbits or *shadow-UPOs* and we shall use the latter term. The evolving orbit can be considered as "jumping" from a path near to one UPO to the next. The orbit may complete multiple circuits or simply repel away before completing a single circuit, depending on the level of instability of the UPO. All UPOs are unstable, but some have greater instability than others. That is, orbits leave the vicinity of different UPOs at slower or faster rates. UPOs that are more unstable will be detected more frequently and others that are highly unstable may be difficult to detect at all. Nearly all UPOs will be detected at least as a incomplete partial cycle.

Recurrence plots are specifically designed to detect this cyclic behavior. They have been demonstrated to be highly useful in detecting cycles in time series data, and a substantial body of research has now been established in this area.

Henri Poincare [111] first introduced the formal concept of recurrences. In the modern context [41] introduced recurrence plots as an analytical tool for experimental time-series and as a medium to visualise the periodic behavior of an orbit through phase space. Subsequent research and publications by [143] and [99] have advanced the study of recurrence plots to an independent field of study. The paper by Marwan et al., "Recurrence plots for the analysis of complex systems" [99] is a definitive reference in this area and considers the effectiveness of the method as inputs vary.

Higher-dimensional phase spaces can only be visualised by projection into the sub-spaces of dimension two or three. Building a recurrence plot enables us to investigate various aspects of the m-dimensional phase space orbit of time sequential data through a concise two-dimensional representation. They are a valuable tool for systems with varying underlying dynamics as they do not require the data to be stationary. Non-stationarity is commonly the case with chaotic systems. Also they may reveal hidden patterns and correlations in complex data series. Most importantly, they also provide a simple robust means to detect periodic cycles in the presence of noise.

The set of UPOs of an attractor is a dynamical invariant; their number, distribution and properties "unfold" the structure of chaotic orbits. They can be used indirectly to calculate other invariants, such as Lyapunov exponents, fractal dimension and topological entropy. Close returns resulting from near-periodic orbit paths on the chaotic attractor manifest themselves as diagonal line segments on the recurrence plot.

Let us consider the orbit T, represented by a finite discrete time series of length N in the phase space of a dynamical system

$$T = \{\overrightarrow{x_i}\}_{i=1}^{N}. \tag{3.2}$$

The *recurrence matrix* of this system is defined as:

$$R(i,j) = \begin{cases} 1 & \text{if } \vec{x_i} \approx \overrightarrow{x_j} \\ 0 & \text{if } \vec{x_i} \not\approx \vec{x_j} \end{cases} . \tag{3.3}$$

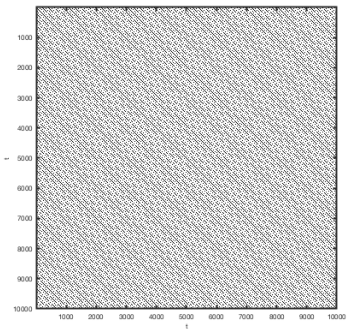
The expression $\overrightarrow{x_i} \approx \overrightarrow{x_j}$ means that the two points are considered equal to within a small error distance ε, which is commonly referred to as the *threshold corridor* or *critical radius* when referring to the L_2 distance metric. A *recurrence* occurs when the orbit returns "near" to a location it has visited previously, within a distance ε. This is further explained in Section 3.4.2. A compact version of this equation is:

$$R(i,j) = \Theta(\varepsilon - \|\overrightarrow{x_i} - \overrightarrow{x_j}\|), \qquad 1 \leq i,j \leq N, \tag{3.4}$$

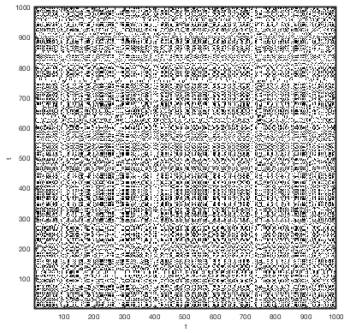
where Θ is the Heaviside function, defined by:

$$\Theta(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases} . \tag{3.5}$$

The recurrence matrix contains information on every possible $\varepsilon-$recurrence in binary notation with an $\varepsilon-$recurrence stored as a 1 and a non-recurrence by a 0. All states of the system are compared with all other states and thus there are N^2 elements in the recurrence matrix. Recurrence points are marked on the recurrence plot with a black dot and non-recurrence points are marked with a white dot (i.e. are unmarked). The recurrence plot is a thus a square plot comprised of black dots denoting recurrence between indexed points to within a distance of ε. The horizontal and vertical axes are labelled "t" and "t +i" respectively. There is an upward diagonal in all recurrence plots arising from the fact that self-compared vectors result in zero distances. We refer to this as the *identity line*.



(a) Periodic function.



(b) Gaussian white noise.

Figure 3.3: Recurrence plots for: (a) A periodic function y=sin(5πt/100). Note the clear structure within the plot, and (b) Gaussian white noise with mean 0 and standard deviation 4.0. Note the snowy structure within the plot.

Deterministic time series will produce a recurrence plot with many short line segments parallel to the identity line. In contrast, the recurrence plot of white noise, will reveal no structure and a "snowy" appearance. These are illustrated in Figure 3.3.

Building the recurrence matrix from the time series is an important first step in locating cycles within a time series. From the recurrence matrix we will determine the fundamental period of the chaotic system and other harmonics. The determination of cycle periods using recurrence data is somewhat resilient to noise. Our testing indicates periods remain determinable at added noise levels up to approximately 10%. Knowledge of the fundamental period and various harmonics will allow us to locate the noise-infected cycles in the time series. For a detailed theoretical analysis of the effect of noise on recurrence plots, the reader is referred to [139].

The diagonal form of the recurrence plot is a useful visual tool, with cycles appearing as diagonal lines. However for locating and extracting cycles it is preferable to see cycles appearing as horizontal lines, which is far more intuitive. This can be achieved by instead determining the matrix with (i, j) entry,

$$R(i+j-1(mod\,N),j) \qquad 1 \leq j \leq N. \tag{3.6}$$

The *horizontal recurrence matrix* and plot can either be calculated directly from the formula or by permuting the columns of the recurrence matrix $R(i,j)$. A simple example is as follows. Let D be the difference matrix for a time series of 4 points: x_1, x_2, x_3 and x_4.

$$D_{diag} = \begin{pmatrix} \|x_1 - x_1\| & \|x_2 - x_1\| & \|x_3 - x_1\| & \|x_4 - x_1\| \\ \|x_1 - x_2\| & \|x_2 - x_2\| & \|x_3 - x_2\| & \|x_4 - x_2\| \\ \|x_1 - x_3\| & \|x_2 - x_3\| & \|x_3 - x_3\| & \|x_4 - x_3\| \\ \|x_1 - x_4\| & \|x_2 - x_4\| & \|x_3 - x_4\| & \|x_4 - x_4\| \end{pmatrix}. \tag{3.7}$$

The elements under the diagonal are duplicates of entries above the diagonal. For horizontal plots we want the data in the following format:

$$D_{horz} = \begin{pmatrix} \|x_1 - x_1\| & \|x_2 - x_2\| & \|x_3 - x_3\| & \|x_4 - x_4\| \\ \|x_1 - x_2\| & \|x_2 - x_3\| & \|x_3 - x_4\| & \|x_4 - x_1\| \\ \|x_1 - x_3\| & \|x_2 - x_4\| & \|x_3 - x_1\| & \|x_4 - x_2\| \\ \|x_1 - x_4\| & \|x_2 - x_1\| & \|x_3 - x_2\| & \|x_4 - x_3\| \end{pmatrix}. \tag{3.8}$$

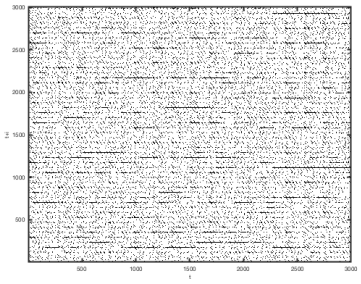
We can see that the j^{th} column of D_{horz} is obtained by permuting the j^{th} column of D_{diag} upwards by $(j-1)$ steps. This relationship may be expressed as:

$$D_{horz}(i,j) = D_{diag}(i+j-1(mod\,4),j). \tag{3.9}$$

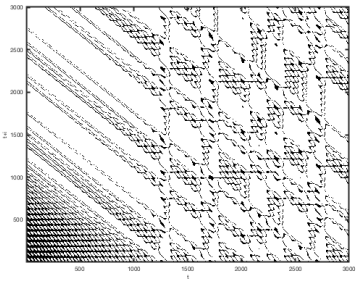
In Figure 3.4 we show horizontal recurrence plots of 1,000 points for specific model chaotic systems that we will be using later. These plots have been derived from the time series for the x-component of each of the systems, which are defined in Chapter 5. One can see the remarkable range of geometric patterns reflecting the diverse topology of the different systems.

In the following subsections we discuss factors affecting the construction of R including the critical radius, sampling frequency, length of the time series, distance

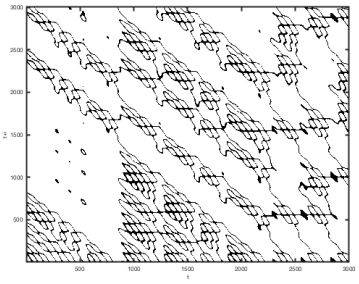
metric, maximum cycle length and initial point of the time series.



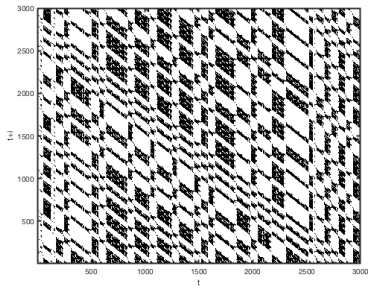
(a) Rossler system.



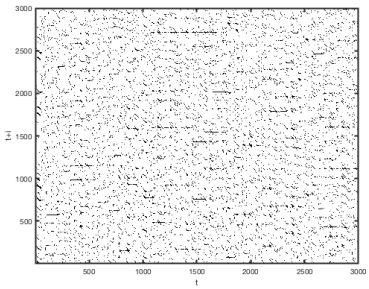
(b) Lorenz system.



(c) Chua system.



(d) Rabinovich-Fabrikant system.



(e) Lu-Chen.

Figure 3.4: Recurrence plots of 1,000 points (a) Rossler system, with a=b=0.2, c=5.7 and the critical radius 1.5; (b) Lorenz system with ρ=28, b=8/3, σ=10 and the critical radius 1.8; (c) Chua system with α=9, β=100/7, b=-5/7 and the critical radius 0.4; (d) Rabinovich-Fabrikant system with a=0.98, b=0.1 and the critical radius 0.2, and (e) Lu-Chen system with a=36, b=3, c=30, u=0 and critical radius 0.8. The system equations and time steps are defined in Chapter 5.

3.4.2 Critical Radius

For analysis of chaotic systems the selection of the critical radius parameter ε is paramount as the orbit will revisit a past state in a "near periodic" manner, but never exactly. The measure of "nearness" is provided by ε. The results are sensitive to the choice of ε. If we select ε too small, we will find few recurrences (false negatives) and little information about the dynamical system. We will not be able to follow the orbit along a complete cycle, and will only find disconnected partial cycle sequences. If we select ε too large, then we will capture artefacts (false positives) and not be able to discern real cycles. Thus there is somewhat of a juggling act in selecting a "best" value of ε. Further, the addition of noise will distort the signal and require a larger ε value. This is balance between collecting too many points and not enough points is illustrated in Figure 3.5.

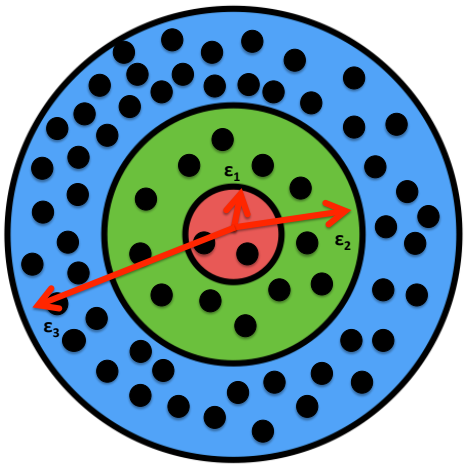


Figure 3.5: The critical radius is paramount in constructing the recurrence matrix from the data points (dots). If we select a value that is too small (ε_1) then we will not capture enough periodic points. Alternatively, if we select a value that is too large (ε_3) we will capture points that are not periodic but simply in the neighbourhood.

There are many different types of recurrence plots, each focusing on a specific aspect of the presentation, in particular varying the definition of the neighbour-

hood ε_i of the i^{th} point. The most commonly used neighbourhood is that with a fixed radius $\varepsilon_i = \varepsilon$, for all i. A fixed radius means that $R(i,j) = R(j,i)$ and thus the recurrence matrix is symmetric. This symmetry results in duplication of information necessary (and thus processing) to detect and extract cycles and it will be eliminated by modifying the recurrence matrix.

As already indicated, the choice of ε is more of an art than a science. Theil et al. [139] quantified the errors arising from measurement noise, the probabilities of false positives and false negatives, and simultaneously minimised these to provide a their criterion for selection of an "optimal" critical radius ε. Their numerical simulations indicated that:

$$\varepsilon \geq 5\sigma, \tag{3.10}$$

where σ is the standard deviation of the measurement noise.

If ε is smaller, the measurement noise will become a dominant effect when detecting recurrence points. Alternatively, if ε is close to the standard deviation of the underlying process, the density of recurrent points will be too great to discern any detailed structure of the underlying dynamics. This rule of thumb in selecting ε is applicable to a broad class of dynamical processes including higher dimensional systems. We estimate this value $\varepsilon_{5\sigma}$ for each of the model systems examined in the application of our noise-reduction algorithm and summarise the results in Table 3.1. We express the critical radius, calculated using $\varepsilon = 5\sigma$ as a percentage of the maximum extent of the attractor, defined as the maximum L_2 distance between any two points on the attractor. This provides a means of comparison with other rules for the critical radius which are expressed as such a percentage.

This minimum value of ϵ sets a reasonable critical radius for noise levels less than 5%. However, it is apparent from our computations that the standard deviation of the noise increases at a far greater rate than the width of the attractor, resulting in impracticably large critical radii for higher noise levels. At this time there is no maximum bound for the critical radius; and this is a topic for future research.

By itself the critical radius is arbitrary so we scale it to the size of the chaotic attractor. We express the critical radius as a percentage of the *maximum attractor extent*, i.e. the maximum L_2 distance between any two points on the attractor. This inflates the critical radius when noise inflates the breadth of the attractor. Many researchers restrict ε to between 2% and 5% of the time series or maximum attractor extent.

Noise level		**Rossler**	**Lorenz**	**Chua**	**R-F**	**Lu Chen**
1%	5σ	0.26	0.40	0.07	0.05	0.40
	% Max Extent	1%	1%	1%	1%	1%
5%	5σ	1.28	1.97	0.34	0.24	1.96
	% Max Extent	5%	3%	4%	6%	3%
10%	5σ	2.57	3.93	0.69	0.47	3.97
	% Max Extent	9%	7%	9%	10%	7%
25%	5σ	6.35	9.82	1.71	1.19	9.86
	% Max Extent	19%	16%	18%	22%	16%
50%	5σ	13.06	19.54	3.41	2.41	19.61
	% Max Extent	32%	27%	32%	33%	27%
100%	5σ	25.86	39.32	6.87	4.73	39.74
	% Max Extent	44%	37%	40%	43%	41%

Table 3.1: Summary of calculated values of ϵ=5σ for a group of model chaotic systems with a range of additive (Gaussian white) noise. This is expressed as a percentage of the maximum attractor extent, defined as the maximum L_2 distance between any two points on the reconstructed attractor.

Noise level	**Rossler**	**Lorenz**	**Chua**	**R-F**	**Lu Chen**
1%	1.4	2.8	0.4	0.2	2.9
5%	1.4	2.8	0.4	0.2	2.9
10%	1.5	2.9	0.4	0.2	2.9
25%	1.7	3.1	0.5	0.3	3.2
50%	2.1	3.7	0.5	0.4	3.7
100%	2.9	5.3	0.9	0.5	4.9

Table 3.2: Critical radius determined as 5% of the maximum L_2 distance between any two points on the reconstructed attractor.

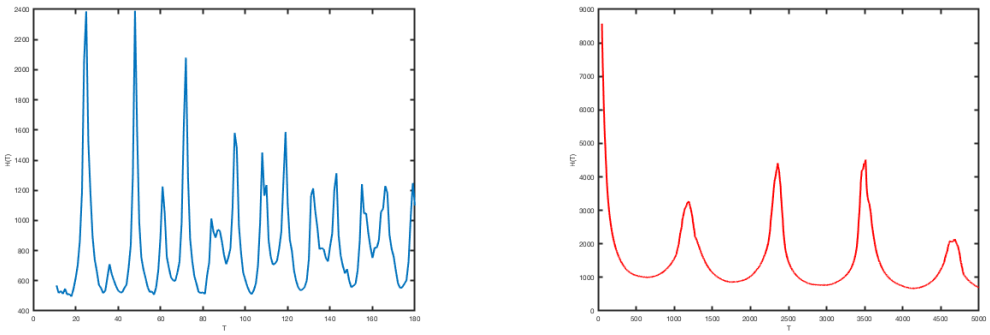
We calculate 5% of the time series width and attractor width for each of our model systems and for a range of added noise levels; with results collected in Table

3.2. We have found that as a rule of thumb, these 5% levels are a good guide to setting the critical radius for noisy data in our modelling. Sometimes we will use a higher percentage and sometimes less, depending on the geometry of the attractor.

3.4.3 Sampling Frequency

The *recurrence histograms* (which we will define and describe in more detail in Section 3.4.8) constructed from recurrence matrices have been shown to be robust in the presence of noise [104] and with respect to starting point when assembling the time series.

As with all signal analysis, the result is highly sensitive to sampling frequency and some trial and error is necessarily involved to find a good representation of the system; neither under-sampled or over-sampled. Recurrence plots are sensitive to sampling frequency. If we are restricted to sampling and recording a finite number of points then the sampling frequency needs to be calibrated to ensure the data captures the salient features of the underlying chaotic system. If we sample too infrequently (*under-sample*) we are missing sensitive information encoded in the data about the underlying dynamics. Similarly if we *over-sample*, we will have difficulty in isolating the salient features for analysis. Under-sampling results in aliasing in Fourier analysis. An estimate of the fundamental period can be made from an initial sample and subsequent sampling times are calibrated to ensure sufficient points are detected for this period. It is important to ensure the shortest period cycle is well populated with data points (at least 50) for the application of noise-reduction techniques later. In Figure 3.6 we provide examples of recurrence histograms for under-sampled and over-sampled data.



(a) Over-sampled.

(b) Under-sampled.

Figure 3.6: (a) Recurrence-histogram from over-sampled Rossler system with time step of 0.01. Although we have a large number of data points, they correspond to a short time period and at best will allow us to identify the first few shadow-UPOs only and over emphasise small fluctuations, (b) Recurrence-histogram from under-sampled Rossler system with time step of 0.25. The data is under-sampled and we do not have adequate information to determine clear peaks. It is likely we are missing information in between peaks.

3.4.4 Length of Time Series

We use model time-series of 10,000 points, noting that in practical situations it may be difficult to process extremely long time series (in excess of 20,000 points) given the dimensions of the recurrence matrix is proportional to N^2. If multi-sampling is available, then the effectiveness of the method is enhanced significantly. There is no obvious advantage to having a single long time series as the sensitivity to initial conditions ensures repeated samples with different initial points will likely provide sufficient information about the attractor. In Chapter 5, we will show how multi-sampling can be effectively exploited to obtain much improved estimates of UPOs, maximal Lyapunov exponents and reconstructed time series. Time series length is a very significant factor in the computing resource requirements (of time and memory).

3.4.5 Distance Metric

A number of distance metrics may be used for determining nearness of points. We use the L_2 or Euclidean distance measure for determining periodicity as it is the most intuitive.

3.4.6 Maximum Cycle Length

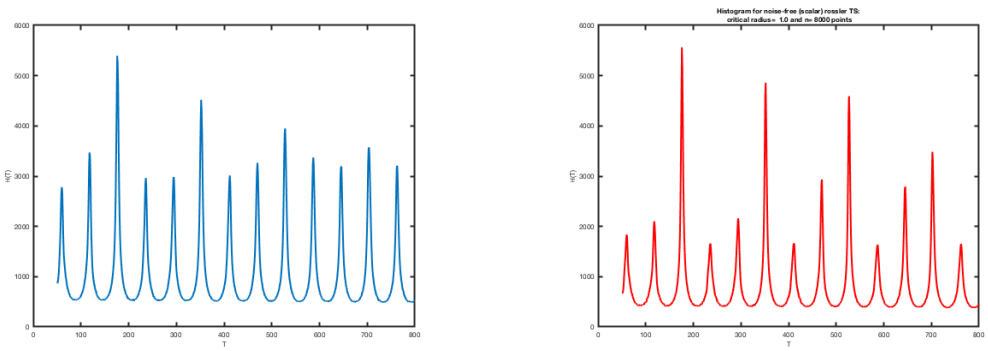
Longer cycles pose a particular problem as increased length results in greater opportunities for unavoidable transverse perturbations to appear. This may result for example from experimental noise or floating point arithmetic on a computer. Obviously there are limitations in maximum cycle size that arise from the recurrence matrix detection method and length of time series. A single cycle cannot be longer than $N/2$, where N is the time series length when using this methodology. The larger issue is the reduced trust in accuracy of longer and longer detected cycles.

Fortunately a consequence of periodic orbit theory is that a chaotic system can be well approximated by the sum of shorter length shadow-UPOs [28], [29]. We can invest greater trust in the reliability of these and they are also present with greater density within the system. Shorter cycles generally occur more frequently. Multiple detected copies of short length shadow-UPOs are what is needed for our method of noise-reduction via averaging.

Translating this into practical terms means specifying the maximum period to be detected. As a general rule, we restricted the maximum detected shadow-UPO length to $10T$, where T is the fundamental period.

3.4.7 Initial Point

The response of recurrence histograms to a change in initial point depends on the ergodicity of the chaotic system; whether it travels equally to different locations in time or tends to linger in specific neighbourhoods of the attractor. For the Rossler system we note the recurrence histograms are robust in relation to initial point provided the time series is sufficiently long to capture the dynamics. Experimentation reveals different starting points generally result in the same identified peaks but with different relative heights. This assumes the orbit travels relatively uniformly around the attractor and does remain in small areas for long periods. One can easily use the recurrence histogram to find an upper bound for the highest detectable period for the given data set. In Figure 3.7, we show the effect on the recurrence histogram for the Rossler system of having different initial points. The peaks are still present at the same frequencies, but with different relative heights. Multi-sampling the chaotic system from different starting points improves results greatly.



(a) Initial point A.

(b) Initial point B.

Figure 3.7: (a) Recurrence-histogram for Rossler system with initial points A (-8.3029, -3.7887, 3.3094), (b) Recurrence-histogram for Rossler system with initial point B (0.5010, 2.8851, 4.1555). The same harmonics are detected as above, albeit with different relative peak height.

3.4.8 Recurrence Histogram

Having selected an appropriate critical radius ε and constructed the reduced recurrence matrix, we now determine the fundamental period T of the system and the periods of other cycles. This is achieved by constructing a histogram that summarises the information in the recurrence plot. The histogram summarises the number of periodic points for each value of i:

$$H(i) = \sum_{j=1}^{N} \Theta(\varepsilon - \|x_i - x_j\|), \tag{3.11}$$

where Θ is the Heaviside theta function as defined in Equation (3.5).

If the time series is chaotic, the histogram will contain a series of peaks, sometimes but not always evenly spaced. In Figure 3.8 we show the recurrence histogram constructed using the scalar values in time series. In Figure 3.9 we construct the recurrence histogram using the embedded time series.

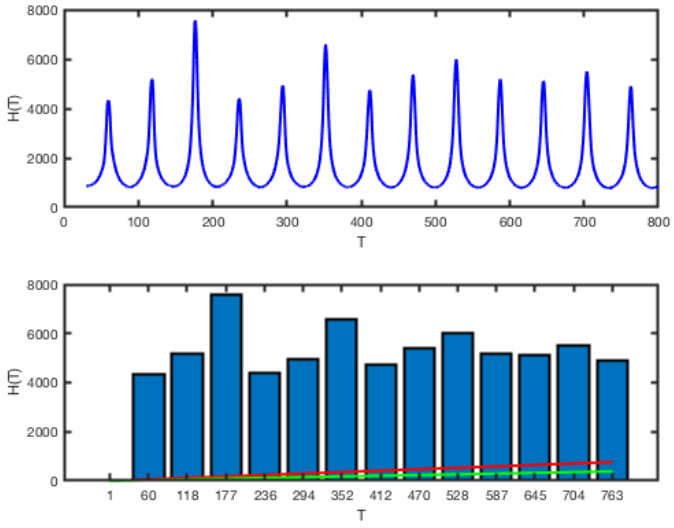


Figure 3.8: Top: Histogram derived from scalar time series, showing number of points on Rossler attractor associated with each recurrence time. Bottom: Block histogram showing cycle periods of identified peaks. The red and green lines are explained in the text below.

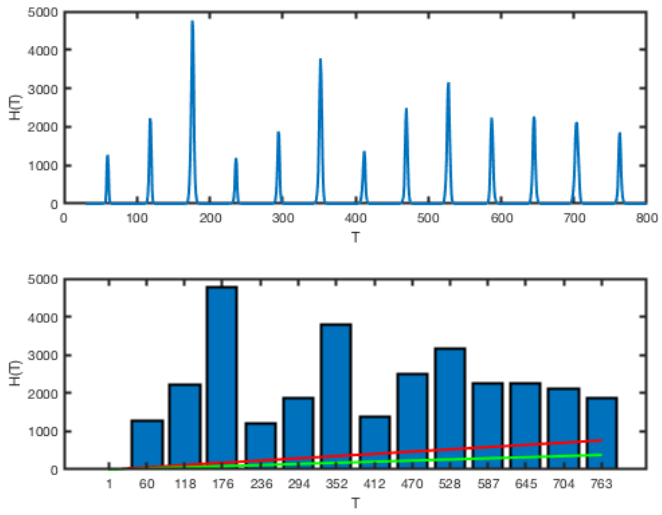


Figure 3.9: Top: Histogram derived from embedded time series, showing number of points on Rossler attractor associated with each recurrence time. Bottom: Block histogram showing cycle periods of identified peaks.

The peaks are identified as representing the fundamental period of the system and subsequent harmonics. The peaks are obvious from the upper histogram in Figure 3.8 and are identified and captured using a peak detection algorithm in Matlab. The block histogram in the lower plot shows only the cycle periods associated with the peaks and the frequency of points detected for each. The diagonal red line in Figures (3.8) and (3.9) represents the period of each cycle and thus represents the minimum number of detected points that must be in a bin for a complete cycle to exist. In the examples provided the height of the histogram is considerably greater than the red line showing that multiple complete cycles are possible for each period. The green line represents the minimum number of points required for a partial cycle of a given period.

The histogram counts the frequency of recurrence by summing along the rows in the recurrence matrix, counting up all the 1's. Note that a high peak of a given period does not guarantee than many shadow-UPOs of this period have been detected. It simply means a lot of periodic points have been detected. In

the case of a relatively stable shadow-UPO many complete cycles will be collected. In the case of a relatively unstable UPO we may see many incomplete partial cycles appearing but few if any complete cycles. Also, stand alone uncorrelated points which appear as fuzz on the scalar close returns plot, will be counted and contribute to the height of the histogram peak (although not materially).

The non-zero baselines in the scalar data histogram correspond with diagonal segments on the recurrence plot. As mentioned above, these are due to the fact that in stationary data sets, upward trending segments of data are always followed by downward trending segments, which create these spurious close returns. These spurious points settle relatively uniformly in the base. This is "cleaned up" with embedding as is evidenced in Figure 3.9.

In Figure 3.10 we show the effect on the recurrence histogram of varying the critical radius. We can see that the effect of increasing the critical radius is to sharpen the peaks, increasing the height and reducing the width. Although not researched here it is suggested that perhaps the definition of the optimal critical radius could be defined in terms of optimally resolving the peaks.

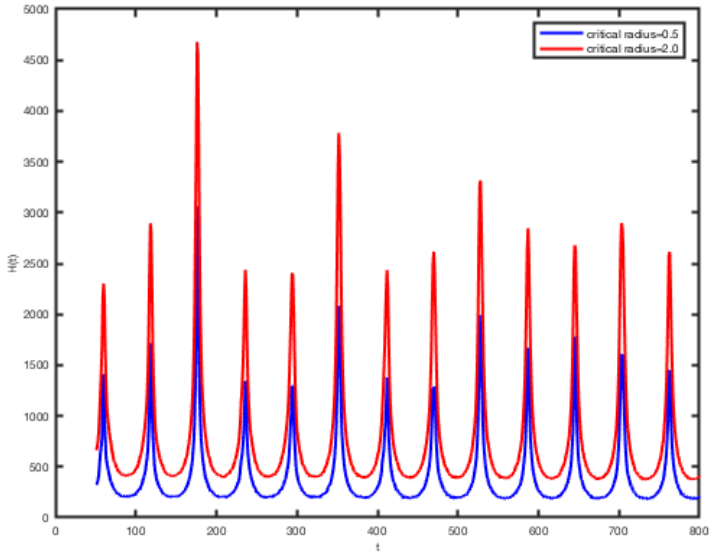


Figure 3.10: Effect on the recurrence histogram of increasing the critical radius from $\varepsilon = 0.5$ to $\varepsilon = 2.0$.

3.4.9 Is Embedding Necessary?

It is worth considering whether or not there is any advantage from firstly embedding the time series in phase space and using these points to build the recurrence matrix, or simply constructing the recurrence matrix using using the scalar time series directly.

Increasing the embedding dimension "cleans up" the recurrence plot, removing the "fuzz" [98]. This corresponds to removing single recurrence points (from the uncorrelated states) and emphasises the horizontal structures (from the correlated states). It is worth noting however, that spurious single points do not interfere with the process of finding sequences of recurrent consecutive points (UPOs). Thus although the appearance of uncorrelated points as "fuzz" on the recurrence plot in unattractive, it does not cause any problems.

Techniques have been developed for quantifying the structure in recurrence plots developed from experimental data. Collectively these techniques are referred to as Recurrence Quantification Analysis (RQA) and the key set of quantifying techniques was devised by [141]. Each of these measures was explored by [70] and they found that the same results could be obtained with or without embedding.

Regardless of measurement noise, it so happens that the errors associated with the estimations involved in finding a suitable embedding dimension m and time delay τ, can severely affect the result. Embedding can result in a significant amount of spurious correlations in the system which then manifest in the recurrence plot. Marwan et al. [99] show this effect can even produce distinct diagonally orientated structures in a recurrence plot of a time series of uncorrelated values if the embedding is high enough, although diagonal structures should be extremely rare for such uncorrelated data.

This is best understood by considering uncorrelated Gaussian noise by itself and analytically calculating the correlations that are induced from a non-optimal or unsuitable embedding. Since the process is uncorrelated, the detected correlations must be arising from the embedding process. The auto-covariance function of D_{ij}^2 is constructed where $D_{ij} = \|\vec{x_i} - \overrightarrow{x_j}\|$ is the distance function measured in the L_2 metric. Interestingly, the resulting expression reveals that there will be peaks in the auto-covariance function if the step size h, $h+(j-1)$, or $(j-i)-h$ are equal to one of the first (m-1) multiples of the time delay τ. These "false" peaks are not present when no embedding is used (m=1). These spurious correlations induced by the embedding lead to unwanted small-scale structures in the recurrence plot. Further, the embedded data may result in false peaks in the histogram subsequently used to determine the fundamental harmonic and subharmonics. The net result is that if one wishes to use embedding, assuming noise is very low, the embedding parameters must be chosen carefully and accurately.

Whilst acknowledging the sound theoretical arguments above, we still see considerable utility in preparing recurrence matrices and plots using both scalar time series data and embedded data, if possible. The primary source of information

is the scalar data, however the embedded data can be most useful in clarifying sections of the scalar–based recurrence histogram without clear peaks. We note that the false peaks on the embedded data recurrence histogram will at best be very small and peaks aligned on both the scalar-data recurrence histogram and embedded-data recurrence histogram can be compared and valuable information deduced. There are cases where the embedded data is clearly superior. It is important to keep in mind we are usually seeking only the first ten or so cycles.

With experimental data from an unknown system or with material levels of noise, we have fewer options. We earlier highlighted than Takens' theorem does not necessarily hold in the presence of experimental noise, and so unfolding the data into phase space as a starting point appears somewhat problematic from the outset. With a noise-infected embedding one simply cannot guarantee that the reconstructed noise-infected is topologically equivalent to the underlying system and that the dynamical invariants are preserved. In the case of an experimental time series with significant noise, we thus have little choice but to work with the scalar time series directly as estimation of embedding parameters is near impossible with material noise.

If we are restricted to using a scalar time series, for purposes of identifying UPOs, the recurrence histogram can be significantly improved by taking multiple time series samples, each from a different initial point (we use the midpoint of the current time series as the initial point for the new series) and combining results. In Chapter 5 we will apply the SUNR method to multiple noise-infected time series samples and use the noise-reduced approximant for each to directly estimate the maximal Lyapunov exponent. We conducted this experiment on several different chaotic systems of varying topology. For each system tested we found that by averaging the Lyapunov exponents from multiple time series we obtain a considerably more accurate result than that obtained from a single time series.

3.5 Detecting Noise-Infected Shadow-UPOs

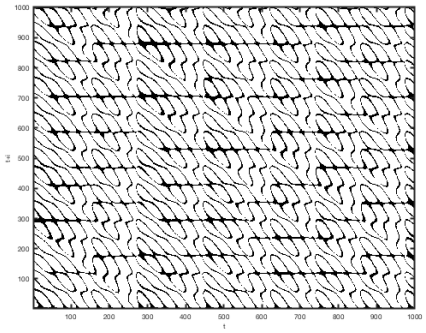
We now consider the effects of noise on the recurrence plot and the recurrence histogram. The methods we are about to describe assume low dimensionality of the chaotic system and thus that the system most likely has a single positive Lyapunov exponent.

3.5.1 Recurrence Plot

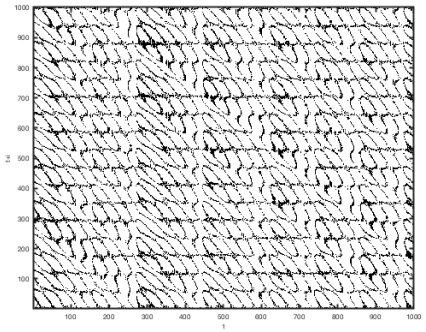
The inclusion of measurement noise results in fewer detected complete cycles and more incomplete partial cycles, in comparison with a noise-free scenario. As one would expect this is because random noise will occasionally cause detected points to fall outside the critical radius that otherwise would have been detected within the critical radius. The illustrations in Figures 3.11a–3.11d reveal how, with increasing noise, the horizontal lines in the recurrence plot corresponding to cycles become fractured.

3.5.2 Recurrence Histogram

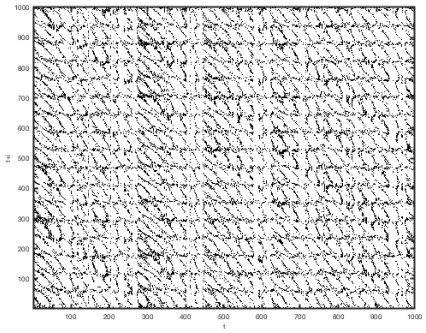
The recurrence histogram is remarkably resilient to noise in the presence of medium to high levels of added noise [104]. It is this attribute of the recurrence methodology that we exploit to locate and extract shadow-UPOs even in the presence of high noise levels. Figure 3.12 illustrates the effect of increasing the Gaussian white noise to 100%. Despite some degradation in the clarity of peaks, they are still easily detectable. One interpretation of this result is that the intrinsic periodicity appears to remain a dominant feature for high levels of added noise. This result similarly applies for dynamical noise, where although the cycles are deformed by the noise, for moderate noise levels the periodicity appears to remain intact. Although we can determine the system harmonics in the presence of significant noise, this does not guarantee we can detect full unbroken cycles.



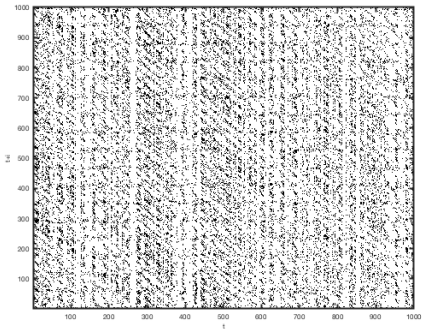
(a) GWN=0%.



(b) GWN=25%.



(c) GWN=50%.



(d) GWN=100%.

Figure 3.11: Recurrence plots of the Rossler system with Gaussian white noise increased from 0% - 100%. Note the disintegration of the horizontal lines with increasing additive noise.

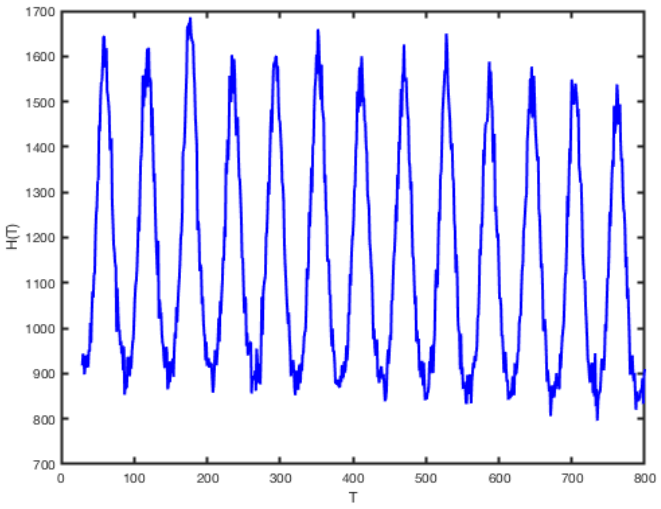


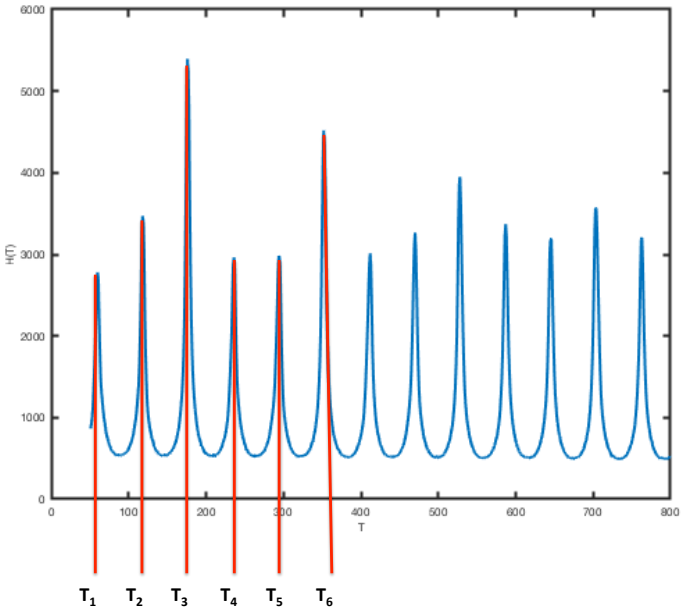
Figure 3.12: Effect of increasing noise levels on the Rossler recurrence histogram. The robustness of the histogram is clear with peaks representing harmonics still clearly detectable at 100% noise.

3.5.3 Location of Cycles in the Time Series

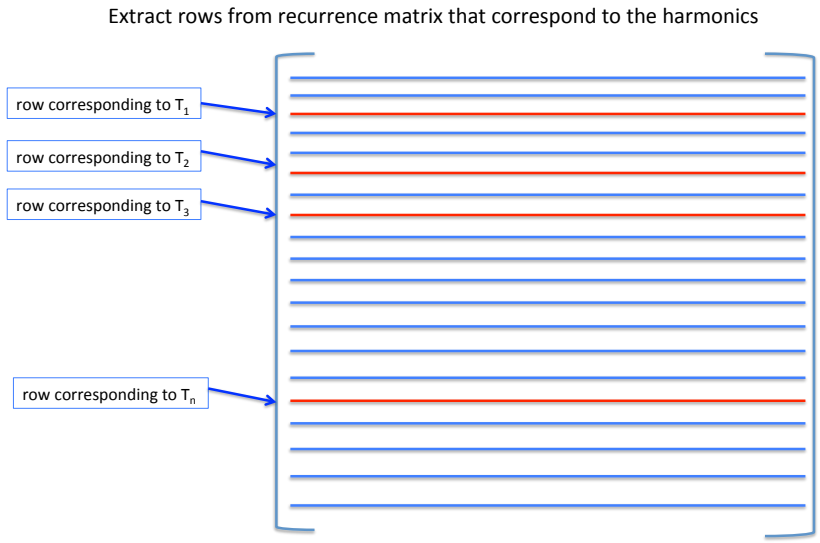
The recurrence histogram allows us to detect the periods of cycles. Peaks on the histogram correspond to the period of the most frequently detected periodic points. This information allows us to confine our search of the recurrence matrix to only rows relating to these specific periods. Importantly, the presence of well-defined peaks does not guarantee that we will find large volumes of complete cycles of these periods. The number of detected complete cycles depends on the instability of the chaotic system and level of noise present. The periods of the composite cycles may adhere to a simple fundamental harmonic and associated sub-harmonics or a more complex harmonic relationship. We detect the peaks using a peak detection algorithm and extract the rows from the recurrence matrix that correspond to the histogram peaks for more detailed analysis. This results in a *reduced recurrence matrix*. This process is illustrated in Figure 3.13.

Each row of the reduced recurrence matrix corresponds to cycles of a specific period and holds important information on them. The elements of a specified row

are binary; either a '0' or a '1'. For a row corresponding to period T, the presence of a '1' at a point indicates that another equivalent point has been detected T points later in that row. The orbit has passed this point x_i and subsequently passed within a small ball centered at x_{i+T} constructed using the critical radius ε. A consecutive unbroken sequence of '1's represents points on a shadow-UPO of period T. The detected sequence may be of length $L < T$ and in this case we have an incomplete or *partial cycle*, where the orbit has moved away from the unstable cycle. In other cases we may find a long sequence, encompassing multiple circuits of a cycle, before the orbit finally breaks free. How long the orbit remains on a cycle depends on the level of instability of that cycle. Shadow-UPOs of some periods may be relatively less unstable than others, and appear multiple times along the row of the reduced recurrence matrix as complete cycles or even multi-circuit sequences. Other highly unstable cycles may only appear as incomplete partial sequences of length less than T.



(a) The periods identified from the peak identify harmonics of the system and identify the relevant rows in the recurrence matrix to search for consecutive periodic points. .



(b) We extract rows in the recurrence matrix relating to the system harmonics.

Figure 3.13: Locating and extracting rows from recurrence matrix that relate to the system harmonics (UPO periods).

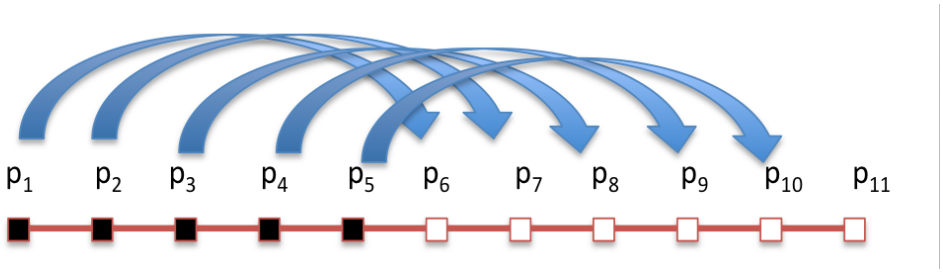


Figure 3.14: For $p_1 - p_5$ to be deemed a cycle of period 5, the entire sequence must repeat itself. This means $p_6 = p_1, p_7 = p_2, .., p_{10} = p_5$ and so for every cycle detected there is a second cycle and thus the cycles appear in conjugate pairs.

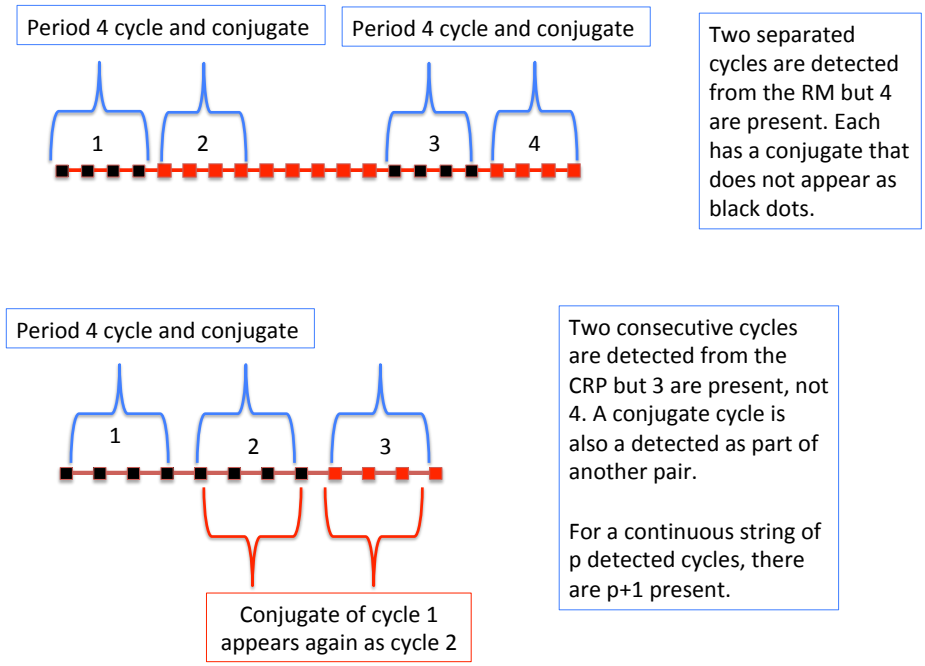


Figure 3.15: Two scenarios are presented where two cycles are detected as sequences of black dots on the recurrence plot. Top: Conjugate pairs are separated in time and 4 cycles are detected. Bottom: Shadow-UPOs are consecutive and the conjugate also appears as a detected cycle as part of another shadow-UPO conjugate pair.

3.5.3.1 Complete Cycles

The detection in a row of a sequence of '1's longer the period T, means that we have detected a cycle of period T but it also provides more information just this. This method of detection requires cycles to appear in "conjugate" pairs. "Conjugate" in this context refers to occurring in pairs. This is illustrated in Figure 3.14 for a simple cycle. For a given critical radius, confirmed detection of a cycle requires an unbroken sequence of consecutive points all of which must occur again exactly T time steps later. There are several possibilities as to how shadow-UPOs and their conjugates may appear, and it important to account for them correctly. Several possibilities are illustrated in Figure 3.15. The recurrence methodology has a very "strong" criteria for a cycle to be detected. Each cycle must have a successive *conjugate cycle*.

3.5.3.2 Partial Cycles

These are sequences of periodic points that are not long enough to constitute a complete circuit. Figure 3.16 shows a typical section of a row of period T in the reduced recurrence matrix. The partial cycle also has a conjugate. Partial cycles occur much more frequently than complete cycles and are valuable sources of information. Clearly a very short partial sequence is less reliable than a near complete cycle. Figures 3.17a–3.17b provide illustrations of various levels of cycle completion.

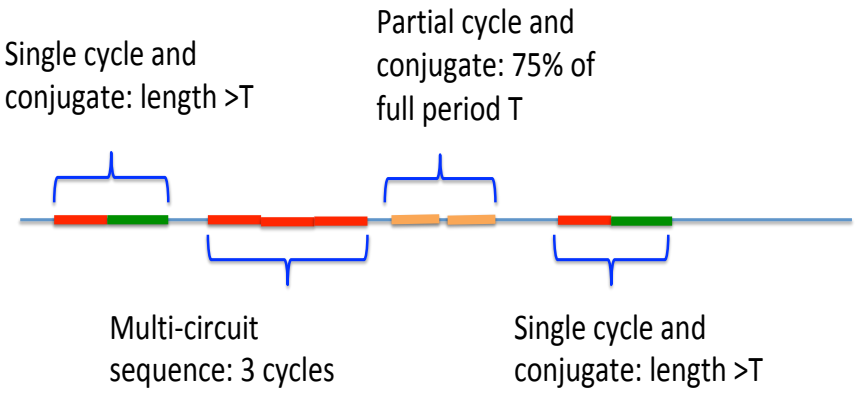
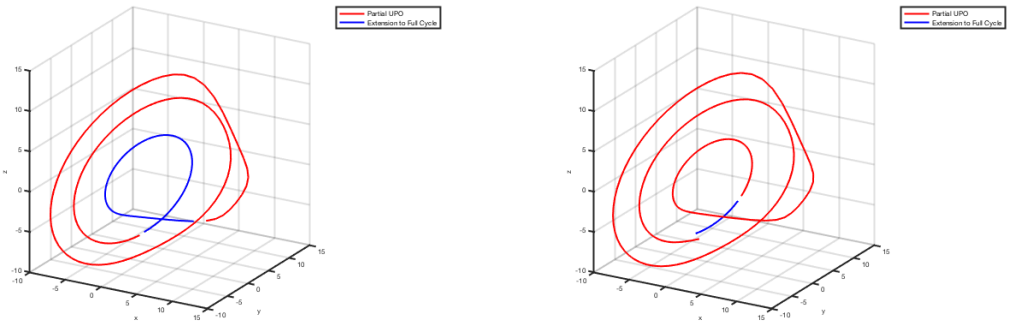


Figure 3.16: Example of sequencing pattern of periodic points occurring along row of reduced recurrence matrix corresponding to period T.

To utilise the data embodied in the partial cycles, we need to decide the minimum percentage of the period that is acceptable. This quantity is user-defined and judgment is involved, with a key consideration being the degree of instability of the extension path. If the orbits are well behaved after the cut off at the minimum partial percentage, then perhaps a relatively high threshold is acceptable. If orbits immediately diverge rapidly after the cut-off then perhaps a lower threshold is required. For the relatively well-behaved Rossler system, we set the minimum partial percentage in the range 50%–75% of the period T.

3.5.3.3 Artefact Cycles

There is an unintended consequence of the recurrence matrix method that results in repeated multiple circuits of a lower order cycle being also captured as a higher order cycle. These are detected cycles that do not have a fundamental period equal



(a) Low completion % .

(b) Higher completion %.

Figure 3.17: (a) Partial Rossler shadow-UPO (T=3) with completion percentages 64.2% of a full period, shown in red. The orbits are extended for illustration purposes in blue, to the end of the period, (b) Partial Rossler shadow-UPO (T=3) with completion percentages 94.3%. Clearly we will attribute greater trust to the 94.3% partial cycle here than the 64.2% cycle above. Sections in blue appear to be a continuation of the shadow-UPO but are (often just) outside the critical radius.

to the measured period, and we refer to them as "artefact" cycles as they are a consequence of the recurrence matrix detection methodology.

This is best illustrated by a simple example. Consider a cycle of period T that traverses 4 laps of the cycle. This will appear in the recurrence plot as an unbroken sequence of length $3T$ and will result in a detected sequence of length $4T$ once we account for the conjugate cycles. This will necessarily also appear as a cycle of period $2T$, with its conjugate being of length $2T$. One can see these are not cycles of minimum period $2T$ but are of a lesser period; these are accounted for within the lower order detected cycles and are an artefact of the method. This is illustrated with an example in Figure 3.18.

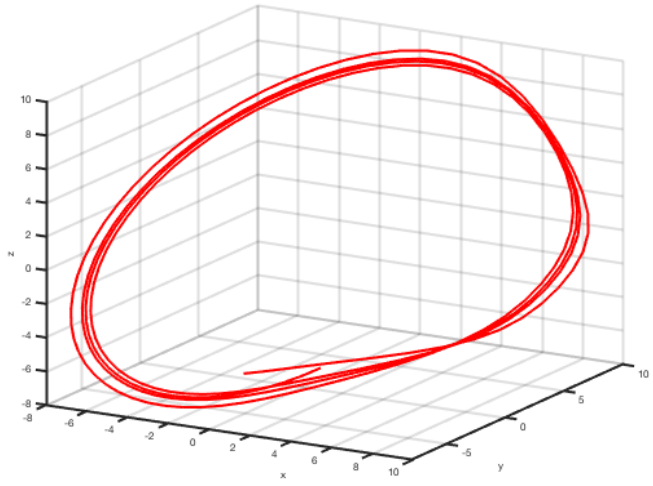


Figure 3.18: Example of artefact cycles. We see a sequence consisting of 4 consecutive circuits of a period 1 shadow-UPO. This will also be detected as two consecutive period 2 shadow-UPOs. The latter are redundant and must be deleted from the UPO location matrix.

These artefact cycles should be deleted from the cycle location matrix, as they are misleading when it comes to averaging binned cycles later. In the context of approximating the noise-free time series by replacing the relevant sections of the noise-infected time series with noise-reduced cycles, the artefact cycles are inconsequential and will most likely just be filtered twice. However, it is incorrect to treat these artefacts as distinct higher order cycles when seeking a complete set of all detectable and noise-reduced UPOs for other purposes.

There are a number of ways to eliminate these artefact cycles. Firstly, they can be visually observed in the summary data and deleted. We highlight 3 examples in Table 3.3 for the Rossler time series. This is an easy task with a small number of detected UPO sequences, but more difficult with multi-sampling or longer data sets.

A simple algorithm can be constructed to identify artefact cycles. Starting with lower order cycles and progressing to higher order cycles we can search for

the following attributes:

1. At least 4 complete consecutive circuits are required for an artefact cycle to occur. These are relatively uncommon and multi-circuits with more 4 loops can be identified and isolated.

2. For each of these multi-circuits sequences of period T, search for cycles of period $2T$, with starting points nearby. Due to the use of a critical radius and noise, the start points may differ by a few points (we require start points to be within 1–3 points).

3. Note that we are generally dealing with cycles of maximum period $10T$. In the case of eliminating artefact cycles from higher order period multiples, one needs to test the factorisation of the longer cycle for multiples of the lower order cycle being tested. For example a detected sequence of length $20T$ (i.e. a period $10T$ and conjugate) could be two $5T$ cycles-conjugate pairs.

As a general rule, the application of (a) and (b) together will suffice to remove any artefact cycles.

Start	End	Period T	Period Multiple	Number Full Cycles	Extra Points	Recurr. Matrix Row Number
1738	1878	60	1	2	21	1
2794	3106	60	1	5	13	2
3908	4050	60	1	2	23	3
4552	4745	60	1	3	14	4
6248	6503	60	1	4	16	5
7360	7672	60	1	5	13	6
565	992	118	2	3	74	7
1855	2266	118	2	3	58	8
2796	3113	118	2	2	82	9
6250	6565	118	2	2	80	10
7362	7679	118	2	2	82	11
860	1264	176	3	2	53	12
2269	2671	176	3	2	51	13
3322	3784	176	3	2	111	14
4026	4428	176	3	2	51	15
4786	5742	176	3	5	77	16
5663	6205	176	3	3	15	17
6485	6854	176	3	2	18	18
6775	7237	176	3	2	111	19
1564	2049	236	4	2	14	20
4899	5660	352	6	2	58	21
5311	6064	352	6	2	50	22

Table 3.3: Location matrix for detected full shadow-UPOs of noise-free Rossler attractor with critical radius =1.0 and 8,000 points. These locate the first 22 UPO sequences in the time series. A similar matrix is compiled for partial cycles. Artefact period 2 cycles highlighted in blue. These must be deleted as they are duplicates of the period 1 cycles shown in green.

3.6 Augmented Detection Rate using Moving Average Filter

In Figure 3.11 we showed the disintegration of the horizontal line sequences in the recurrence plot that occur as a result of measurement noise. We also demonstrated the high resilience of the recurrence histogram to noise and that even with high levels of added noise we can still determine the periods of the underlying UPO set. The difficulty is that we are now only able to detect a few, if any complete (and not fractured) UPOs, even though we know the UPO periods.

The effect of noise is to fracture the horizontal lines representing UPO sequences in the recurrence plot, greatly reducing the power of the recurrence method to capture UPOs. We have explored ways to "restore" the broken sequences in order to extract complete cycles (start and end points) in a manner that preserves the information in the sequences. Interestingly, we have found that the simple act of pre-processing the data with a narrow-window centered moving average achieves this goal and significantly enhances the UPO detection rate. We typically use an 11-point window, that is, 5 points either side of the central point. The temporal autocorrelation associated with moving averages is negated by using a centred moving window, preserving the phase of the data with no phase lag introduced. The moving average smears the noise sufficiently to bridge the fractures, restoring continuity. It brings individual perturbations that are "breaking" the continuity of cycles back to within the detection hyper-cylinder of critical radius ε. Results for 11 point and 21 point moving average windows are shown in Table 3.4. The 11 point moving average appears well-calibrated, but the 21 point window has detected too many cycles (false positives). We subsequently have found that this idea of recovering the signal from a recurrence plot using a moving average (low-pass) filter was originally discovered by Mindlin and Gilmore in [104].

Increasing the filter window length increases the number of cycles detected. For small window size the moving average pre-filter enhances the quality of the UPO capture process significantly. However there are obviously limitations to the filter

Noise Level	**No Pre-Filter**	**11 Point MA Pre-Filter**	**21 Point MA Pre-Filter**
0%	82	82	116
1%	82	84	118
5%	74	86	122
10%	12	88	108
20%	0	94	106
50%	0	10	70
100%	0	0	2

Table 3.4: Effect of centred moving average pre-filter on number of shadow-UPOs detected.

window length, after which the signal is smeared. We prefer to keep the window length as short as possible as the goal is not to utilize it as a low-pass filter but rather to restore the continuity of complete cycle sequences. Before using the method of moving averages, one should first be able to decide the length of the smallest cycle associated with the data. As a general rule, we like to maintain at least 50 points in the fundamental or period 1 UPO, and in this case 11 points (5 either side of the central value) lies well within cycle length. In Table 3.4 we summarise the effect for various noise levels, of the length of the moving average pre-filter on the number of complete cycles detected in the Rossler example. We can see that no complete shadow-UPOs are detectable from the data with added noise in excess of 10%. The 11 point filter results in more cycles being detected from noisy data than with no pre-filter, particularly for the higher noise levels (up to 50% noise). The 21 point filter results in even more cycles being detected, but may be detecting spurious cycles that are not shadow-UPOs. In Figure 3.19 we show the recurrence plot for a pre-filtered Rossler time series with 100% added noise. This can be compared with Figure 3.11d, where most structure has disintegrated due to the noise. Figure 3.19 shows clear structure, similar to the noise-free scenario shown in Figure 3.11a, and highlights the restorative effect of the moving average.

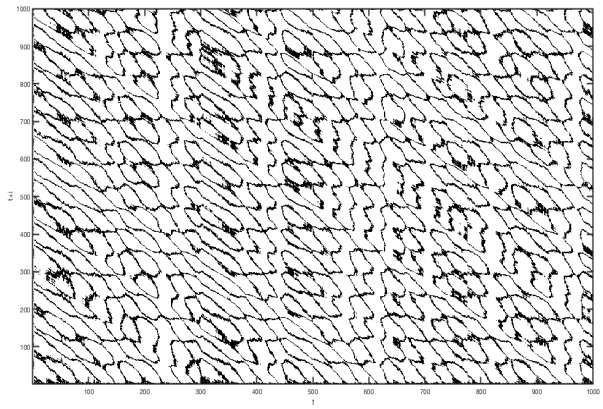


Figure 3.19: Recurrence plot of Rossler system with 100% added Gaussian white noise, illustrating the restorative effect (compare with Figure 3.11d) the moving average pre-filter has on the horizontal shadow-UPO lines.

This approach not only restores UPO continuity for detection purposes, but also preserves phase synchronicity and does not corrupt the underlying chaotic signal. This is the step required to enable the use of the powerful cycle detecting capability of the recurrence matrix method in high noise scenarios.

Finally, we reiterate that it is important not to draw any conclusions from the relative height of the recurrence histogram peaks. The purpose of the histogram is to locate the fundamental period and other harmonics of the chaotic system and the peaks simply identify these. The peaks represent numbers of detected periodic points and this alone does not guarantee the existence of complete cycles within the recurrence matrix. A tall peak may simply represent a multitude of short broken sequences of consecutive periodic points from a relatively unstable cycle. A shorter peak may nicely represent several complete cycles from a relatively less unstable cycle. Also note that artefact cycles inflate the height of their relevant histogram peaks. If we construct a revised histogram, after the elimination of artefact cycles, we will see shorter peaks as period multiple increases.

3.7 Summary and Discussion

In this chapter we describe a modified recurrence matrix method to locate shadow-UPOs in a noise-infected chaotic time series.

We have selected recurrence methods to locate cycles for several reasons. Firstly, the more sophisticated, Newton-Raphson type methods are not appropriate for an environment with medium to high noise. They are complex for the non-mathematician, rely on initial conditions for convergence, usually require an embedding and may require well-placed Poincare sections. Recurrence matrix methods are more effective in the presence of significant noise. Detection of cyclic orbit-sequences shadowing UPOs need not rely on an embedding so the issues with finding correct time delays, embedding dimensions and Takens theorem in the presence of noise are removed. Secondly, the recurrence histograms constructed using detected periodic points are resilient to high levels of noise, almost conserving periodicity; and one can thus locate cycles accurately under difficult conditions.

Next we describe how to construct recurrence matrices and recurrence histograms. The system harmonics are distinguished as peaks on the recurrence histogram and this information identifies the rows of the recurrence matrix to search for periodic sequences. Without modification, the recurrence histogram is quite resilient to noise, clearly revealing peaks at noise levels up to 10%, sometimes more. We apply a search algorithm to the rows of the reduced recurrence matrix corresponding to each detected period, looking for unbroken sequences of consecutive points (denoted as 1's) and recording the location in the time series of detected complete cycles and sufficiently long partial cycles. These will later be extracted for noise-reduction.

The critical input for recurrence matrix methods is the critical radius ε. We tested the criteria that states an optimal minimum value of $\varepsilon \geq 5\sigma$, using 5 different model chaotic systems. We found that this rule appears to be reasonable for noise levels less than 5%, but quickly inflates to impractical levels in the presence of higher levels of noise. Instead we prefer to define a maximum, recommending

5%—10% of the attractor extent, depending on attractor geometry and suggest testing carefully. We found this approach works better with higher noise levels.

We discuss the use of embedded data versus scalar data to locate system harmonics. For all systems modelled, we found that recurrence histograms constructed from scalar data were adequate and in many cases superior to those using the embedded data. This is further supported by research and has a theoretical basis [70]. Whilst acknowledging that use of embedding noise-infected data in the close returns matrix is logically weak, we still regard the embedded data as highly useful adjunct in the practical context. In some cases the embedded histogram will greatly assist in identifying histogram peaks that are not resolved adequately in the scalar histogram. We have found it beneficial to run both scenarios and use the scalar histogram as the base case. We also considered the effects on the recurrence histogram of varying the initial point and sampling frequency. We emphasise the advantages of using multiple time series samples whenever possible.

It is important when extracting cycles, to understand how cycles are manifested in the recurrence matrix and the processing considerations of the methodology. The cycles necessarily appear in conjugate pairs, there may be a phase difference between detected cycles of the same period and "artefact" cycles naturally arise as a consequence of the method. Artefact cycles are consecutive multiple circuits of certain lower order cycles appearing as higher order cycles and we describe how to remove them. It is best to store cycle information as indices linking cycle location back to time series.

We next examined the effect of noise on recurrence plots and recurrence histograms. As noise increases we observe that the horizontal lines on the recurrence plot become fractured. This is the effect of noise perturbing underlying cycle points to outside the critical radius tube. This phenomenon eventually manifests in the peaks of the recurrence histogram becoming less and less distinguishable. Even if peaks remain distinguishable, the subsequent search of the recurrence matrix results in mostly incomplete (broken) cycles. We modify the recurrence method to improve detection rates in the high noise situation and to maximise use of the

data.

To accomodate higher noise levels, we pre-process the noise-infected time series with a narrow-window (11 point) centered moving average that smears the larger noise fluctuations and effectively repairs cycle continuity, so that they may be detected. Importantly, this is not being used as a noise filter but as a technique to enhance the effectiveness of locating cycles. Cycle information, including phase is preserved provided we ensure sampling frequency allows the moving average window length to form a small portion of a single cycle. Pre-processing the time series in the manner allows us to resolve cycles with noise levels up to 50% and in some cases 100%, detecting nearly as many cycles as in the noise-free scenario.

Partial or incomplete cycles, defined as sequences comprising more than a certain percentage of a period (say 50%) are detected and stored also. The rigour of the recurrence method requires the partial cycle sequence is replicated T time steps later in the time series. These contain valuable information and are very frequent due to the instability in chaotic systems. In some cases they are all we can detect. Short partial cycles are of little use as they are highly unstable and are not representative of the underlying UPO, and some judgment is required in defining the minimum accepted length of a partial cycle. However, nearly complete partial cycles are sufficiently stable to detect, may be filtered and will be useful in approximating the noise-free time series in Chapter 5.

Chapter 4

Shadow UPO Noise Reduction Method

4.1 Chapter Overview

Chaotic motion can be viewed as an aperiodic orbit wandering among and shadowing an infinite set of UPOs. The attractor is the closure of, and thus can be approximated by, the set of UPOs and the unstable cycles can be imagined as a sort of skeleton supporting the dynamics of phase space. In particular, they provide an invariant topological characterisation of the dynamics, so estimating a basis set of lower order UPOs from noisy data is a worthwhile objective. If we can estimate the set of noise-free UPOs, then we have the ingredients to qualitatively define the chaotic system and also to calculate qualitative measures such as the maximal Lyapunov exponent.

This chapter describes the application of our new shadow-UPO noise reduction (SUNR) method to chaotic time series with lower instability (i.e. at the lower end of the range $0 < \lambda_{\max} \leq 1$). In these systems the instability is sufficiently low that we are likely to be able to detect sufficient complete shadow-UPOs of all periods to estimate a full noise free basis set of individual lower order UPOs. We cannot specify exactly how low the maximal exponent must be. The criteria of suitability here is that we can find at least 50% of the time series using detected

complete cycles. Otherwise the system is too chaotic for this approach and we defer further consideration to the time series approximation model presented in Chapter 5. Examples of chaotic systems where this technique is appropriate include the Rossler and Lorenz systems.

We use the Rossler system to demonstrate the reduction of measurement noise using the SUNR method. We apply the methods of Chapter 3 to detect and capture complete and partial cycles from the model (measurement) noise-infected Rossler time series. We use the detected and binned shadow-UPOs to estimate a set of noise-free lower-order basis UPOs and demonstrate the process using multi-sampling. The binned cycle data are used to numerically calculate the short time Lyapunov exponent for each lower order UPO. We have included a number of figures to illustrate the applied mathematical processes required and to provide a visual aid in understanding the dynamics.

Before demonstrating the SUNR technique, we begin by developing the framework for constructing model time series for experimentation and testing. We define and discuss the various types of measurement noise (uniform noise, Gaussian noise, high-frequency noise and pink noise) and dynamical noise that we will be using to test our approximation of the chaotic noise-free time series in Chapter 5.

4.2 Modelling Measurement and Dynamical Noise

The model Rossler time series is constructed by numerically integrating the Rossler system equations using the 4th order Runge-Kutta method, as discussed in Section 3.1. This time series will be corrupted by various types of noise and the various stages of the SUNR process will be examined.

4.2.1 Uniform White Noise (UWN)

The term "white noise" does not refer to a specific type of noise signal, but to a statistical model describing a random signal that has equal power distributed across all frequencies. Thus the FFT power spectrum in the frequency domain

of white noise is flat. White noise is comprised of all frequencies (with constant power at all these frequencies) and thus is analogous to white light emitting all the frequencies in the same proportion. When we use discrete time, as in sampling, white noise is a discrete uncorrelated signal with zero mean and a finite variance. Being uncorrelated in time does not restrict the values a signal can take. Any distribution of values is possible (although it must have zero DC component). The values could be uniformly distributed or assume a Gaussian or Poisson distribution for example. White noise is the generalised mean-square derivative of the Wiener process or Brownian motion.

If we let $R_1(i)\,\epsilon\,[-1/2, 1/2]$ be a random number, n be the percentage noise factor, x_{nf} be the noise-free time series, and σ_{nf} be the standard deviation of the noise-free time series, then the noise-infected time series x_{uwn} is:

$$x_{UWN}(i) = x_{nf}(i) + nR_1(i)\sigma_{nf}. \tag{4.1}$$

This serves as a simple model of a noise-floor for experimental apparatus (baseline instrument noise measurement). For example, thermal noise produced in active electronic components tends to be both uniform and white.

In Figure 4.1 we show the time series for the Rossler system, both noise-free and with 25% added uniform white noise. In Figure 4.2 we show a time series of just the uniform noise and its frequency domain profile after a FFT, where we see all frequencies present.

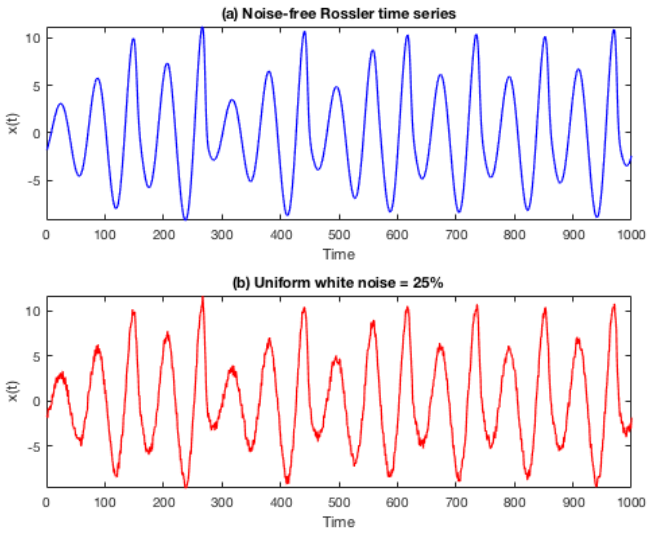


Figure 4.1: 1,000 points of the x-coordinate time series for the Rossler attractor with (a) the noise-free scenario and (b) 25% added uniform white noise.

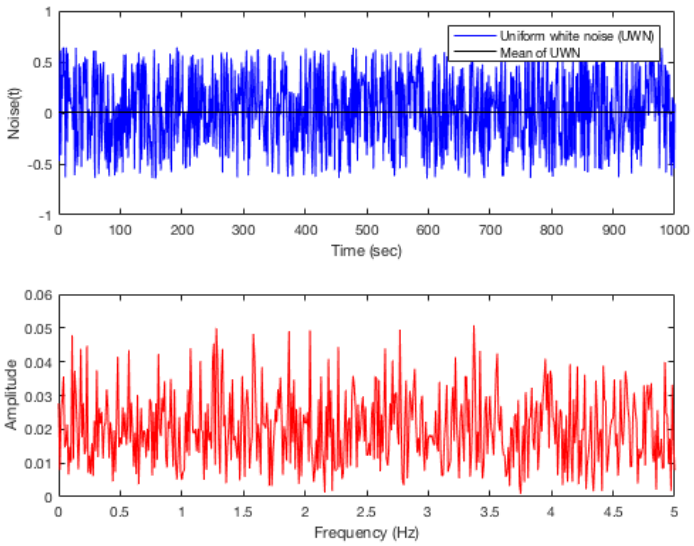


Figure 4.2: (a) Uniform white noise (25% of Rossler signal), and (b) the FFT of the uniform white noise in (a).

In the case above we have added uniform noise with mean 0. If the added noise has a non-zero mean, we will see a spike in the FFT plot. This is commonly referred to as "DC bias". If for example, we add uniform white noise, using $R_1(i)\,\epsilon\,[0,1]$, rather than $R_1(i)\,\epsilon\,[-1/2,1/2]$ then the mean is equal to $1/2$. This tall spike dominates the frequency spectrum and masks whether or not energy in the signal is uniformly distributed in the frequency domain. Subtracting the mean from the noise signal results again in the spectrum shown in Figure 4.2. It is for this reason we construct our model noise-free time series with zero mean. When dealing with an experimental time series, we need to subtract the mean from the data before proceeding with the analysis. The Rossler attractor with 25% added uniform noise is shown in Figure 4.3.

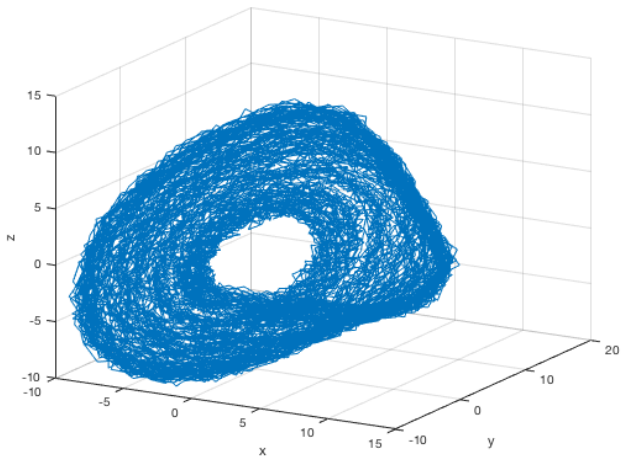


Figure 4.3: Rossler attractor with 25% uniform white noise added using Equation 4.1.

4.2.2 Gaussian White Noise (GWN)

If each sample has a normal distribution with zero mean, the signal is said to be Gaussian white noise. Since the mean of this noise is zero, we do not need to adjust for a DC bias.

If we let $R_2(i)\,\epsilon\,[-\infty, +\infty]$ be a random number that is normally distributed, n be the noise factor and σ_{nf} be the standard deviation of the noise-free time series, then the noise-infected time series x_{GWN} is:

$$x_{GWN}(i) = x_{nf}(i) + nR_2(i)\sigma_{nf.} \tag{4.2}$$

Additive white Gaussian noise (GWN) is a basic noise model used in information theory to model the effect of many random processes that occur in nature. It has a normal distribution in the time domain with mean 0. The central limit theorem of probability theory states that the summation of many random processes converges to a Gaussian or normal distribution. Examples in nature are the thermal vibrations of atoms in conductors (referred to as thermal noise or Johnson-Nyquist noise) and black body radiation from warm celestial objects like the Earth and Sun.

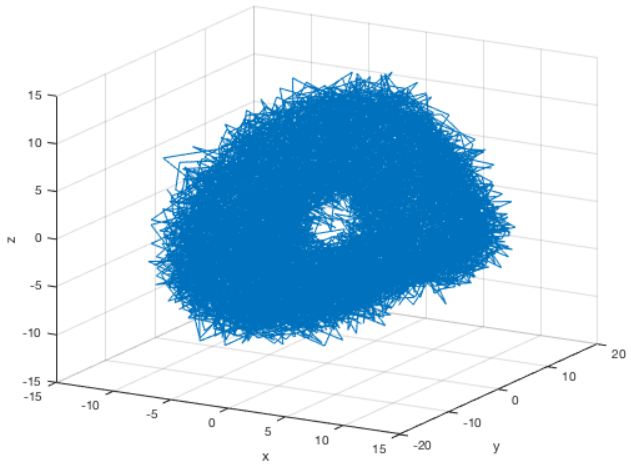


Figure 4.4: Rossler attractor with 25% Gaussian white noise added using Equation 4.2.

-

4.2.3 High Frequency Noise (HFN)

High Frequency Noise (HFN) is more difficult to construct and involves directly building the noise vector in the frequency domain. If we define $R_2(i)$, σ_{nf} and n as before and define k as the minimum frequency after which the high frequency noise acts on the signal. Let $y(i) = nR_2(i)\sigma_{nf}$ for $i = 1, ..., N$, and let $z = FFT(y)$. Define w as follows:

$$w(j) = \begin{cases} 0 \quad \text{for } j < k, \\ z(j) \text{ for } j \geq k. \end{cases} . \tag{4.3}$$

Then the high frequency noise vector is

$$HFN(i) = IFFT(w)(i) \quad i = 1, ..., N, \tag{4.4}$$

where IFFT denotes the inverse FFT. The high frequency noise time series is finally added to the noise-free signal as follows:

$$x_{HFN}(i) = x_{nf}(i) + HFN(i) \;\; i = 1, ..., N, \tag{4.5}$$

In Figure 4.5 we show the high-frequency noise signal in the time domain and frequency domain. In Figure 4.6 the HFN is added to the Rossler signal. The difficulty of meaningfully reducing HFN is exemplified in Figure 4.7.

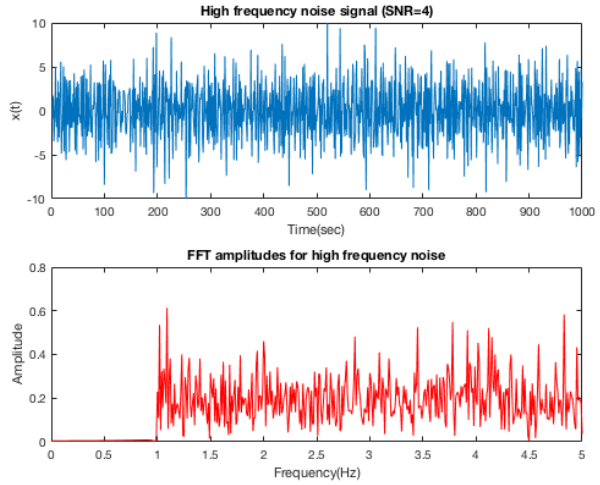


Figure 4.5: Top: High frequency noise time series in time domain. Bottom: Frequency domain, where high frequency noise is added for all frequencies above f=1.0.

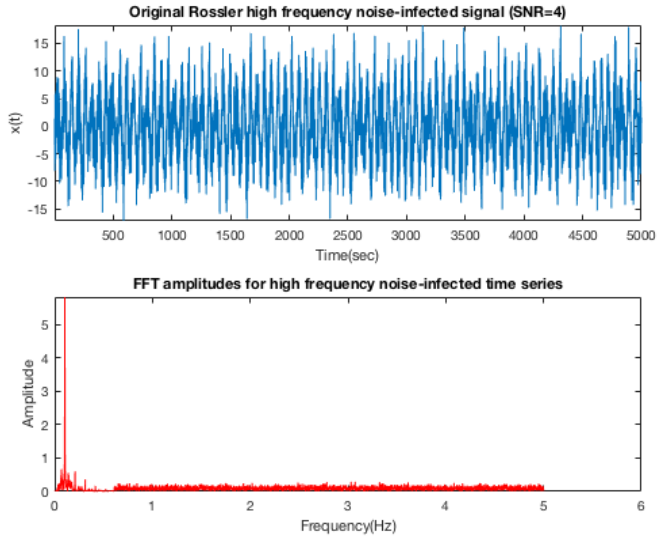


Figure 4.6: Example of adding high frequency noise (HFN) to noise-free Rossler signal. In this case HFN is added to frequencies > 1.0.

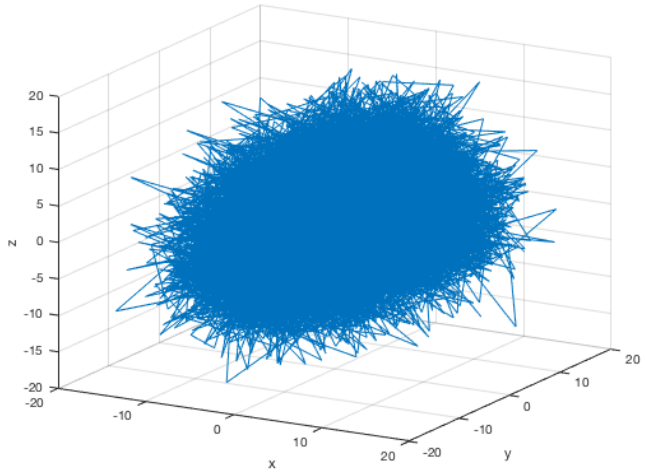


Figure 4.7: Rossler attractor with added high frequency noise. The SNR is 4.0 dB for this example.

4.2.4 Coloured Noise ($1/f^{\mu}$ noise)

Coloured noise (or $1/f^{\mu}$ noise) refers to a signal or process with a power density spectrum (in the frequency domain) that obeys an inverse power law of the signal frequency. The value of the power μ determines the "colour" of the noise. We will illustrate using pink noise or $1/f$ noise, so named from the pink appearance of visible light with this power spectrum. For pink noise, each halving or doubling of the frequency carries an equal amount of noise energy.

Pink noise is pervasive and occurs in a multitude of physical, biological and economic systems. Examples of pink noise in physical systems include waves lapping on the beach, fluctuations in tide and river heights, leaves rustling on trees, the flow of traffic and the pulsation of a Quasar. It is omnipresent in nearly all electronic devices (referred to as flicker noise) as resistance in the solid-state electronics. Biological examples include reflex times, heart beat rhythms and neural activity. Pitch and loudness fluctuations in speech and music are pink noises.

If we define $R_2(i)$, σ_{nf} and n as before and let $y(i) = nR_2(i)\sigma_{nf}$ for $i = 1, ..., N$, and let $z = FFT(y)$. Define w as follows:

$$w(j) = F(f_j)z(y), \;\; j = 1, ..., N, \tag{4.6}$$

where $F(f_j)$ represents the weighting at frequency f_j. Then the pink noise vector is:

$$PN(i) = IFFT(w)(i) \;\; i = 1, ..., N, \tag{4.7}$$

where again IFFT denotes the inverse FFT. The pink noise time series is finally added to the noise-free signal as follows:

$$x_{PN}(i) = x_{nf}(i) + PN(i). \tag{4.8}$$

In Figure 4.8 we show the noise-free Rossler time series and also the same time series with pink noise added and SNR of 4.0 dB. In Figure 4.9 we have isolated the pink noise time series and also show its profile in the frequency domain. The skewing of energy distribution towards the lower frequencies is apparent and we can see that a low-pass filter will be of little use here. We show the pink noise infected Rossler attractor in Figure 4.10.

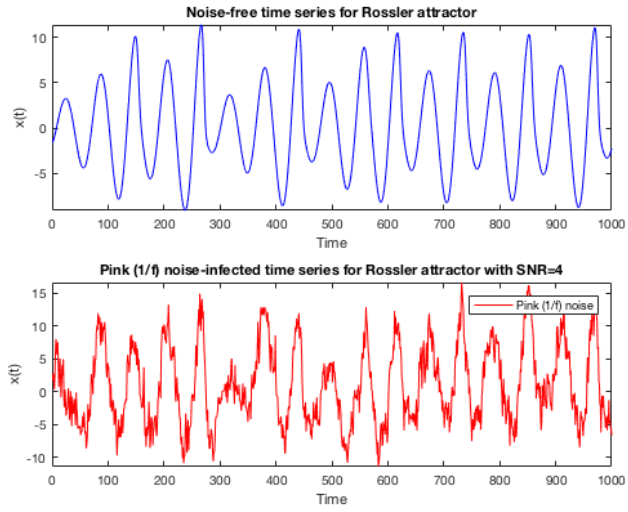


Figure 4.8: Top: 1,000 points of the x-coordinate time series for the noise-free Rossler attractor. Bottom: contains added pink $(1/f)$ noise with SNR=4.0 dB.

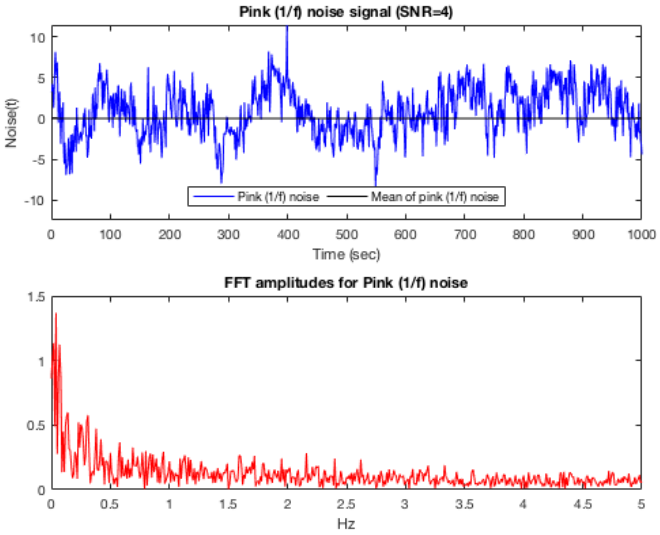


Figure 4.9: Pink or $1/f$ noise, extracted from the signal (top) and its FFT (bottom). The SNR is 4.0 dB for this example.

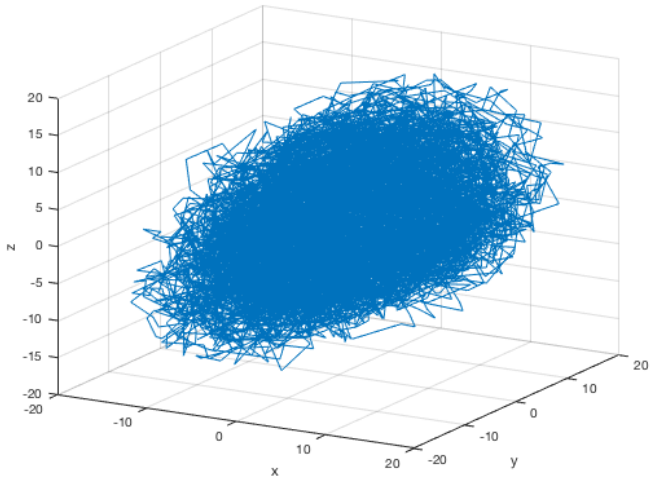


Figure 4.10: Rossler attractor with added pink $(1/f)$ noise and the SNR is 4.0 dB.

In Chapter 5 we will explore scenarios involving the four types of measurement noise we have introduced above. We further consider scenarios with low levels of added dynamical noise.

4.2.5 Dynamical (or System) Noise

As discussed in Chapter 2, measurement noise does not influence the evolution of the system. In contrast, dynamical noise appears as a disturbance that influences the evolution of the dynamical system. Dynamical noise is added to each iterate of the variable in the numerical integrator in the following manner. If the numerical integrator (e.g. Runge-Kutta method) is denoted by F, with step-size h, and the time-series is of length N, then the numerical integration process for noise-free data can be described as:

$$x_{i+1} = F(x_i, h) \qquad i = 1, ..., N-1. \tag{4.9}$$

$$\tag{4.10}$$

We add the dynamical noise as follows:

$$x_{i+1} = F(x_i + N_i, h), \tag{4.11a}$$

$$N_i = nx_iR_2(i),\ i = 1, ..., N-1, \tag{4.11b}$$

where n is the noise factor used to roughly calibrate to the desired SNR ratio. We cannot pre-define the dynamical noise level as it varies randomly per iterate as a function of $R_2(i)$. We calculate the SNR for the time series at the end of the numerical integration process. The best we can do to control the overall SNR is to vary n. Often, the iterates exponentially diverge rapidly.

Thus at the end of each iteration we add an increment to the iterate x_{i+1}; the increment is found by multiplying x_i by the product of a noise factor and a random number with Gaussian distribution (we could alternately use a uniformly distributed random number). In practice only small increments can be added, as the numerical integration process will diverge rapidly otherwise.

In Figure 4.11 we show the noise-free Rossler time series (blue) and the Rossler time series infected with dynamical noise. We observe the difference from measurement noise, which appears as fluctuations around the (intact) underlying signal. The dynamical noise actually disturbs and perturbs the geometry of the underlying signal itself, as well as introducing a phase lag. In Figure 4.12 we have overlaid the two time series shown in Figure 4.11 for comparison purposes and observe that the dynamical noise has altered the geometry and phase of the signal itself. In Figure 4.13 we have isolated the dynamical noise time series and calculated its FFT. We provide a closer view of the FFT in Figure 4.14. The dynamical noise manifests at all frequencies, and in particular at the key harmonics associated with the noise-free time series. The broadband nature of this noise makes it unlikely that we can remove noise with a simple frequency selection strategy. In Figure 4.15 we show the dynamical noise-infected Rossler attractor, which appears very smooth in contrast to the measurement noise scenario, which contains many jagged variations around the underlying path. The noise is "baked" into the signal making it extremely difficult to remove.

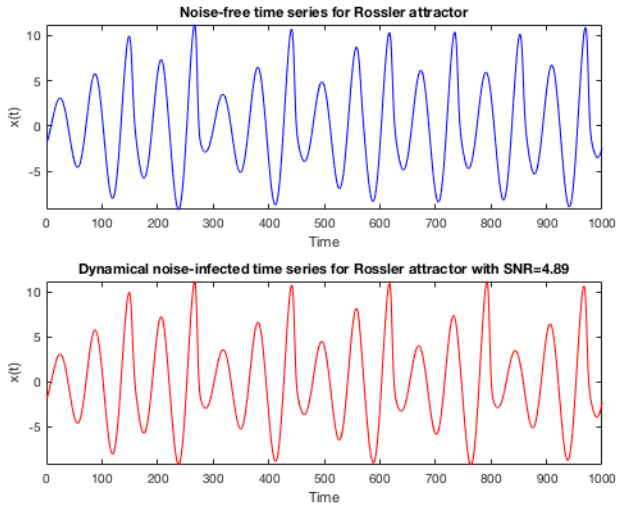


Figure 4.11: Rossler attractor with added dynamical noise and SNR = 4.89 dB.

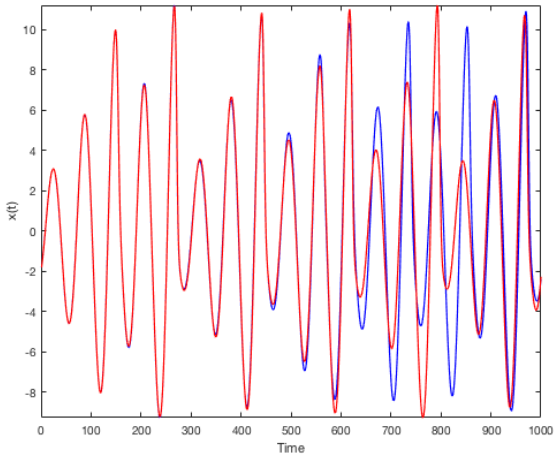


Figure 4.12: We have overlaid the dynamical noise-infected signal (red) with the noise-free signal.

.

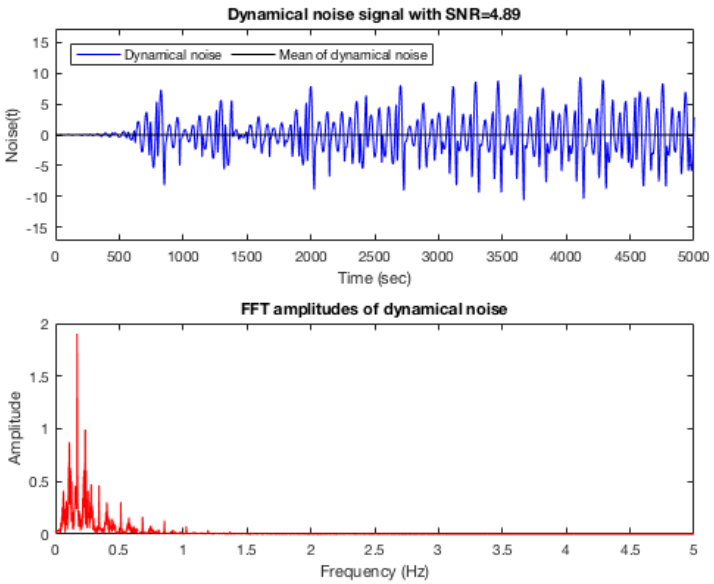


Figure 4.13: Gaussian dynamical noise, extracted from the signal.

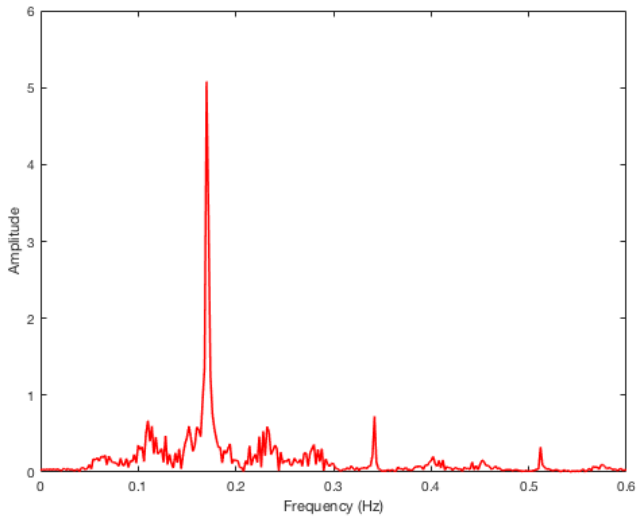


Figure 4.14: Closer view of FFT of Gaussian dynamical noise, extracted from the signal.

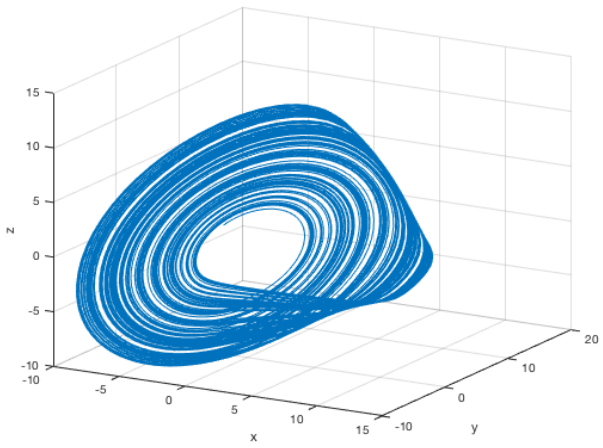


Figure 4.15: Rossler attractor with Gaussian dynamical noise and SNR= 4.89 dB. The attractor appears very smooth in contrast with the measurement noise scenario.

4.3 Heuristic Example: The Rossler System

We now demonstrate the application of the process for estimating lower order UPOs using a numerically integrated model time series for the Rossler system with step-size $h = 0.1$ and 25% added white Gaussian noise. As in Chapter 3 all time series will contain 10,000 points unless stated otherwise. We described the detection and extraction process for shadow-UPOs using recurrence matrices and histograms in Chapter 3. In this chapter we will focus on processing the detected shadow-UPOs to estimate the set of lower order UPOs.

4.3.1 Filter and Critical Radius Choices

Detection and extraction of shadow-UPOs requires the construction of an $N \times N$ recurrence matrix, where N is the length of the time series. This is the most computationally demanding aspect of the approach. The time required to construct a recurrence matrix grows as $O(N^2)$. This is still considerably faster than the $O(N^3)$ required by the Newton-method style algorithms in the literature for noise-

free UPO detection. To minimise processing time, our model noise-infected time series are limited to containing 8,000—10,000 points. For this practical reason, the analysis of multiple samples of the time series is easier to work with than a single long time series and will capture the same information, provided starting points are sufficiently varied. In the example in this section, we used Rossler time series with parameter values as defined in Section 3.3 and time series of 8,000–10,000 points.

A short pre-filter window is used with 5 points either side of the central value (11 point window). For higher levels of noise this may be extended to a wider window size. The application of this pre-filter allows us to detect and identify cycles with noise levels of 25%—50% or even greater, in contrast with 5%—10% for the unadjusted noise-infected time series. It also ensures that small sections of the noise-infected time series that cannot be extracted and noise-filtered directly at least have the noise reduced somewhat.

For the model Rossler system, we use a critical radius $\varepsilon = 1.6$, which corresponds to approximately 5% of the attractor extent.

4.3.2 Results and Coverage Ratio

Having identified the harmonics in the attractor and extracted the complete and partial shadow-UPO sequences, we next examine how successful this set is at covering the time series from which it is extracted. If we are to use the set of extracted shadow-UPOs to approximate the time series, it needs to cover a significant percentage of that time series. For the time series defined above, Figure 4.16 shows the coverage attained using complete cycles only. Figure 4.17 shows the coverage by partial cycles only and Figure 4.18 shows the result of using both complete and partial cycles. We can see the overall coverage rate is excellent and that partial cycles can make a significant contribution.

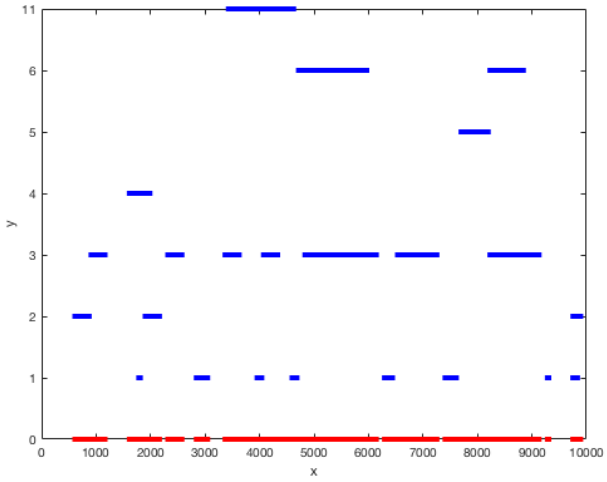


Figure 4.16: Coverage of time series by detected complete cycles is shown in the above plot. The blue lines represent coverage for each period (vertical axis) and the lower red line represents the overall coverage of 78.6%.

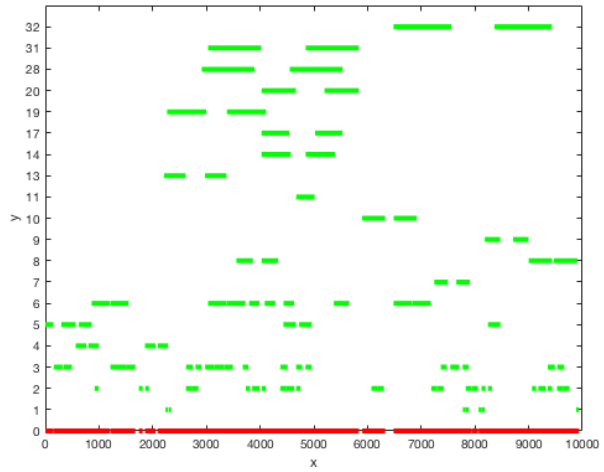


Figure 4.17: Coverage of time series by detected partial cycles ($> 50\%$ period) is shown in the above plot. The green lines represent coverage for each period and the lower red line represents the overall coverage of 89.6%. The vertical axis represents period expressed as a multiple of the fundamental period.

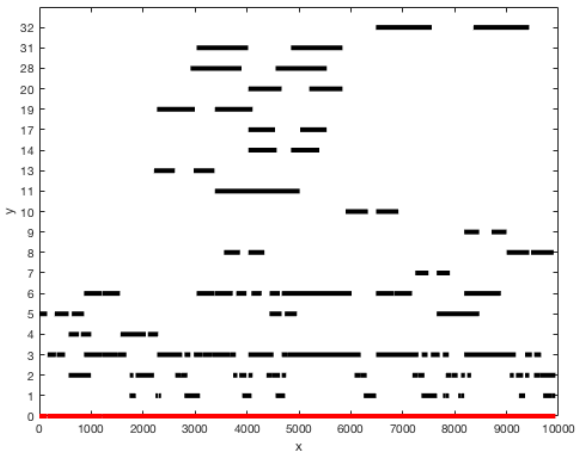


Figure 4.18: Coverage of time series by all detected cycles (complete and partial) is shown in the above plot. The black lines represent coverage for each period and the lower red line represents the overall coverage of 98.6%. The vertical axis represents period expressed as a multiple of the fundamental period.

Reviewing the coverage plots above, complete cycles comprise nearly 80% of time series, partial cycles comprise 96% and the combination of both covers 99% of the time series. We also see that all periods up to $10T$, where T is the fundamental period, are present in this single sample; either as complete cycles or partial cycles. We summarise variation in cover rates as the critical radius varies for the Rossler system in the Table 4.1 below. We observe that the rate of improvement in the cover ratio for complete cycles increases most rapidly until the critical radius equals 1.6, then falls off. Partial cycles provide a high level of coverage throughout and thus are a valuable source of information.

Critical Radius	Cover Ratio Full Cycles	Cover Ratio Partial Cycles	Cover Ratio All Cycles
0.5	36.8%	40.4%	67.1%
0.6	36.8%	47.7%	74.2%
0.7	36.8%	58.0%	85.6%
0.8	36.8%	49.9%	74.4%
0.9	46.5%	62.4%	81.5%
1.0	48.8%	57.7%	81.2%
1.1	55.5%	67.8%	89.1%
1.2	52.5%	67.3%	87.0%
1.3	55.6%	74.8%	93.7%
1.4	58.6%	81.0%	94.9%
1.5	70.5%	83.9%	97.8%
1.6	78.6%	89.6%	98.6%
1.7	83.2%	89.0%	96.9%
1.8	83.2%	91.0%	98.1%
1.9	83.8%	93.0%	98.4%
2.0	87.5%	95.6%	98.5%
2.1	89.8%	96.1%	99.5%
2.2	91.3%	97.5%	99.5%
2.3	91.4%	97.7%	100.0%
2.4	91.4%	98.3%	100.0%

Table 4.1: Variation in coverage levels of the noise-free Rossler time series with critical radius.

4.4 Reducing Noise from Detected Cycles

The technique adopted to reduce noise from individual cycles is dependent on the type of noise. If for example we are dealing with measurement high frequency noise, the approach adopted will be a low-pass filter or cut-off in the frequency domain after a FFT transformation. For noise types that inhabit a broad range of frequencies like Gaussian and uniform white noise, the cut-off has limited success and simply averaging a bin full of equivalent cycles is likely to get a better result. We demonstrate the averaging approach using the Rossler system. In Chapter 5 we will illustrate using more targeted signal processing approaches in the frequency domain.

4.4.1 Highly Populated Bins

The simplest way to establish a set of noise-reduced UPOs is to take the average within each bin containing multiple copies of continuous shadowing orbits that are very close to the underlying UPO at every point. Bins containing the most copies will provide the highest quality estimate of the underlying noise-free UPO. Usually we see a reduction in the detection frequency as the periodic order of cycles increases, however this is highly dependent on the stability of the cycle and the visitation time by the sampling orbit, and is no hard and fast rule.

The collection of shadow-UPOs extracted from the time series are to be sorted into bins of equal period. In the case of evenly spaced peaks on the histogram, corresponding to a cycle structure of a fundamental period and sub-harmonics, there should be little confusion about which cycle type each is. However in the case where peaks are not evenly spaced and some are relatively close to others, it is important they are allocated correctly. It is important to note that detected complete shadow-UPOs of the same period are almost certain to have different initial and final points and will need to be cyclically permuted to align before averaging. Figures 4.19 and 4.20 below illustrate this point. To synchronise the phase of all cycles in a bin, we select the first cycle X as the anchor. We then compare all other cycles Y with X and determine the phase shift to synchronise them. This is achieved by finding the phase lag corresponding to the maximum of the cross correlation r_{xy}. In Figure 4.21 we show a shadow-UPO after it has been cyclically permuted to line up with the "anchor" cycle and we observe the small "bend" that arises. The averaging of cycles smoothes out these bends. Also, in these figures we see the small discontinuity between start and end points caused by the instability inherent in the dynamics. The size of the gap is limited by the critical radius, ensuring we do not capture cycle sequences that are too divergent from the true UPO.

An unintended consequence of the recurrence method is that in certain cases repeated multiple circuits of a cycle are also captured as higher order cycles. These detected cycles have lower periods, this is a consequence of the recurrence histogram cycle detection methodology. These "artefact" cycles must be eliminated

when searching for cycles of unique period, and the methodology to do this was discussed in Chapter 3.

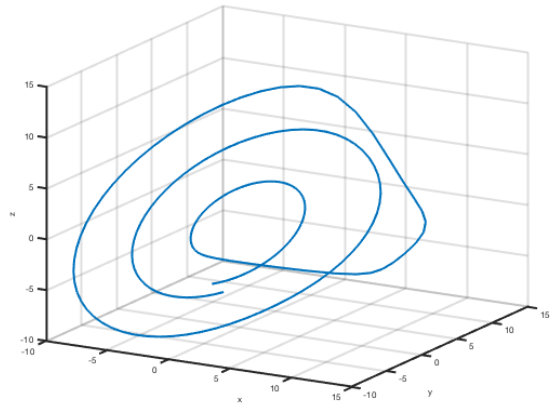


Figure 4.19: Detected shadow-UPOs will display a discontinuity between start and end points, as the orbit is unstable and is shadowing the underlying UPO. This is a period 3T noise-free cycle.

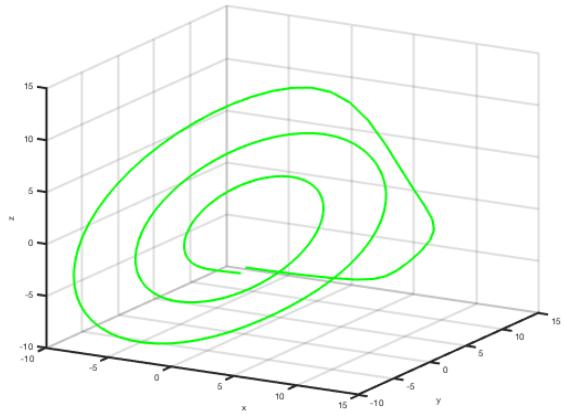


Figure 4.20: Another detected period 3T shadow-UPO from noise-free Rossler system. Each has a different start-end point and these are out of step when represented as a time series or delay vector. One shadow-UPO must be cyclically permuted to align with the other before they may be averaged.

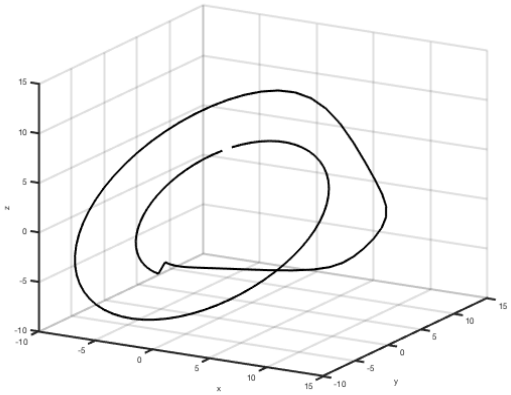


Figure 4.21: A period 2T noise-free cycle that is cyclically permuted by 94 points to align with the anchor. Note the "bend" is formed from the divergent start-end points of the un-permuted cycle.

We now illustrate the averaging of binned shadow-UPOs using the simple case of Rossler system with 5% added GWN. Twenty shadow-UPOs of period 3T were identified. The shadow-UPOs in each bin are first cyclically aligned by maximising the cross correlation between each pair of cycles. Once optimally aligned they are averaged point-wise. The average is shown in Figure 4.22 in red and the individual cycles in blue. We can see nice clustering around the average in this case.

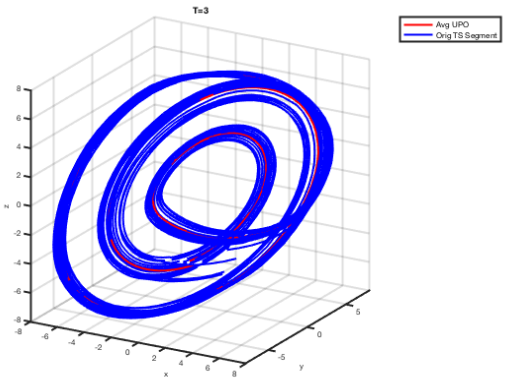


Figure 4.22: Plots of 20 detected period 3T shadow-UPOs and their average.

4.4.2 Lightly Populated Bins

If repeat sampling is possible, we suggest taking many samples and heavily populating the bins, so a very accurate average is produced for common cycles, and we find at least 10 occurrences for "rare" highly unstable periods.

In the event of having few detected pairs, then we are limited in the quality of averaging. If a cycle is detected, then a minimum of 2 copies will be detected as a result of the recurrence matrix methodology. The rare exception is where the conjugate cycle, detected using the scalar recurrence matrix, ends after the last point of the delay vector (but still lies within the scalar time series). We exclude these conjugates as there is insufficient data to find the delay vector. We can average all the detected cycles in a bin to estimate the underlying noise-free UPO, regardless of whether we have 100 cycles or 2 cycles. If there are only 2 copies of a cycle then this is the best we can do for the fixed length single time series sample.

We prefer to take a different approach to UPOs where we only have detected a small number of complete cycles. We note that all detected cycles are approximations to the underlying UPO and are orbits passing nearby and eventually veering off. Thus there is a "gap" or discontinuity between start and end points of the detected cycles, as highlighted in Figure 4.19 above.

We apply a Fourier filter to the "least unstable" of the few detected cycles. Before applying the Fourier filter we transform the cycle to join the end points. This is a simple linear end-point smoothing transformation, preserving the essential features of the cycle and an example for a period $5T$ shadow-UPO is illustrated in Figure 4.23. The "least unstable" example of the underlying UPO is the detected cycle with shortest discontinuity between the start and end point using the L_2 metric. The smallest gap corresponds to "least unstable" of the detected cycles as it divergences the least during the measured time period and is the best approximation to the closed underlying UPO. The end-point transformation is reversed after the Fourier filter is applied. In Figure 4.24 and Figure 4.25 we illustrate this process for the simple case of noise-infected Rossler shadow-UPO of period $11T$

using a time series of 8,000 points and a critical radius of 1.6. We were only able to detect a single pair of these *11T* cycles and thus cannot apply averaging.

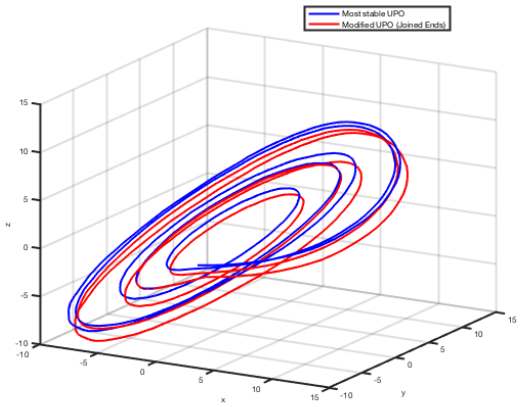


Figure 4.23: Original period 5T Rossler shadow-UPO is shown in blue. The red line shows the transformed UPO with end point smoothing.

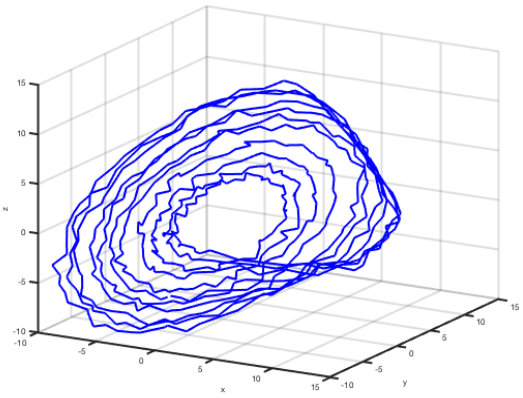


Figure 4.24: Plot of the period 11T "most stable" detected shadow-UPO containing 25% GWN from a time series of 8,000 points (smallest end-gap).

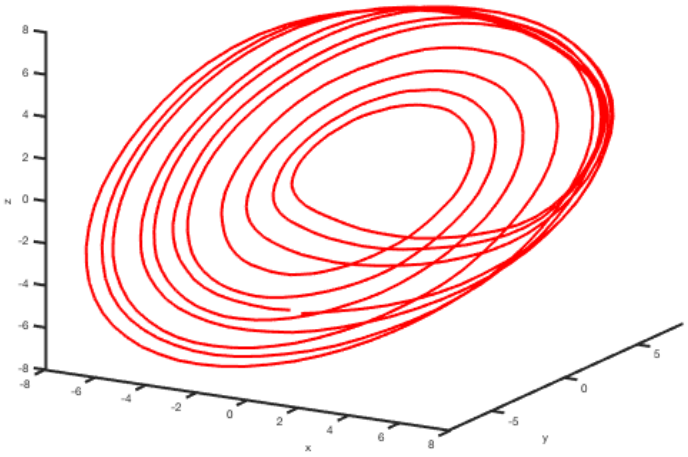


Figure 4.25: Plots of the period 11T "most stable" detected higher frequency shadow-UPO after the application of a Fourier filter, removing unwanted harmonics directly.

4.5 UPO Estimation using Multi-sampling

In this section we illustrate how to use multi-sampling for estimates of the lower order noise-free Rossler UPOs. We take multiple samples of noise-free data and then with 25% added Gaussian white noise. Multiple samples of time series were constructed of the Rossler system, using the mid point of the current sample as the initial point of the subsequent sample. The results of 50 samples of 8,000 points are shown below. The use of multiple samples is preferred as bins containing more cycles of a given period will yield a better average and a single noise-free Rossler time series is unlikely to yield all lower order UPOs. Some will just be too unstable to locate using a single sample. For the single Rossler time series used in Section 4.4, we were unable to detect the period 7 cycle. However, with 50 time series samples, we were able to detect 4 full period 7 cycles.

4.5.1 Lower Order UPOs from Noise-Free Time Series

In Figure 4.26 we show a recurrence histogram, aggregating the results of using 50 noise-free time series. Each time series was generated using the methods described previously and searched for shadow-UPOs. Complete shadow-UPOs are in blue and the partial shadow-UPOs are in red. Shadow-UPOs of periods T, $2T$, $3T$ and $6T$ are very dominant with large numbers detected. We will be able to derive good approximations to UPOs of these periods. In contrast, the period $7T$ shadow-UPO is very difficult to detect and we need to make the best of the examples detected. We can also see the significance of partial cycles for periods $4T$, $5T$, $7T$ and $8T$ where the numbers of complete cycles are relatively small. Tables 4.2 and 4.3 summarise the number of cycles detected and the percentage of the totals contributed by each period. We detected shadow-UPOs of up to period $12T$ and see that those of periods T-$3T$ comprise the 63% of total detected cycles.

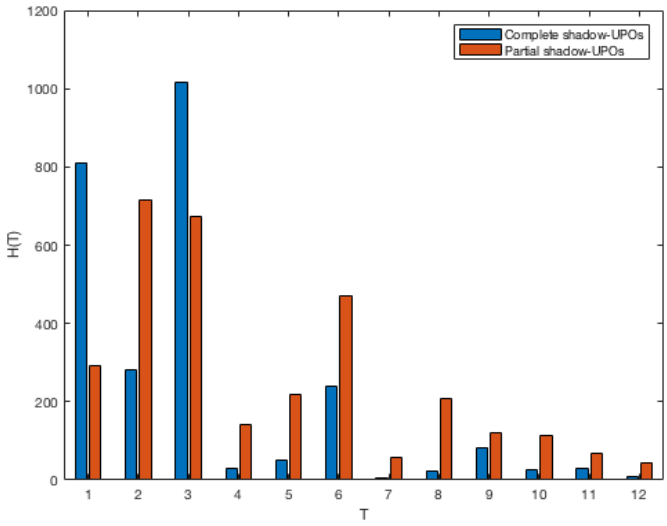


Figure 4.26: Histogram representation of complete and partial cycles detected from 50 noise-free Rossler system time series samples, each of length 8,000 points.

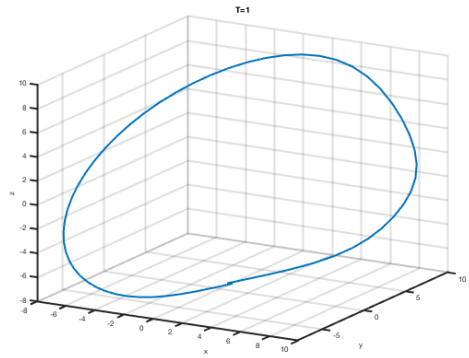
Period	Number of Full Cycles	Number of Partial Cycles
1	811	293
2	282	714
3	1015	675
4	29	140
5	49	219
6	240	472
7	4	56
8	24	207
9	83	122
10	26	113
11	31	67
12	10	45

Table 4.2: Summary table of number of detected cycles by period from 50 noise-free time series samples.

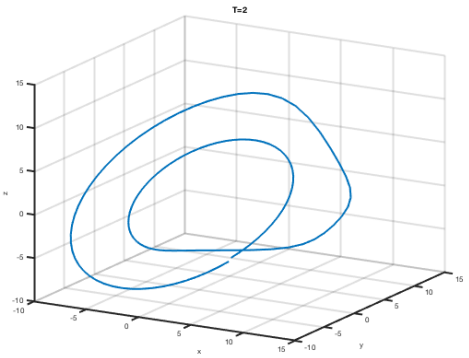
Integer Period	% Full Cycles	% Partial Cycles	% Total
1	31%	9%	19%
2	11%	23%	17%
3	39%	22%	30%
4	1%	5%	3%
5	2%	7%	5%
6	9%	15%	12%
7	0%	2%	1%
8	1%	7%	4%
9	3%	4%	4%
10	1%	4%	2%
11	1%	2%	2%
12	0%	1%	1%

Table 4.3: Summary table of number of detected cycles by period from 50 noise-free time series samples. Full cycles comprise 45% of the total. Approximately 63% of detected cycles are of periods T=1 to T=3.

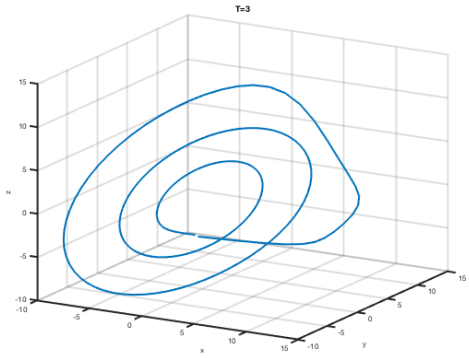
In Figures 4.27 and 4.28 we present estimates of all the lower order UPOs up to period $12T$. Each was calculated by averaging the appropriate bin of synchronised shadow-UPOs.



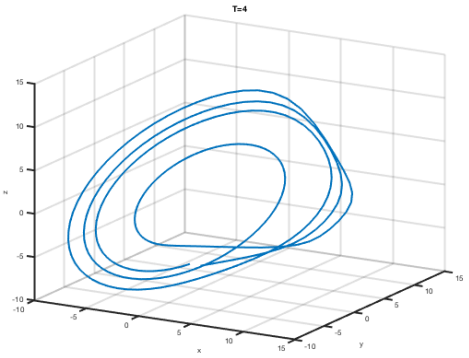
(a) Period 1 cycle.



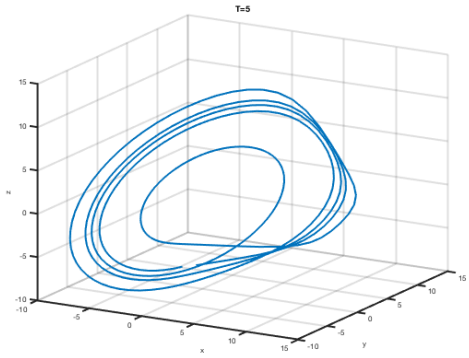
(b) Period 2 cycle.



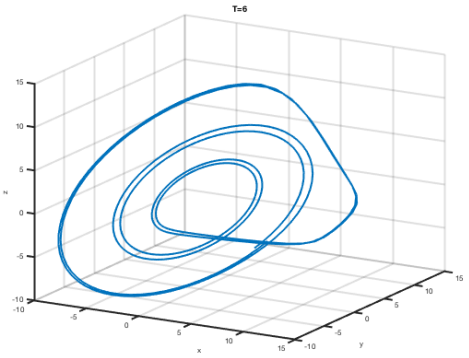
(c) Period 3 cycle.



(d) Period 4 cycle.

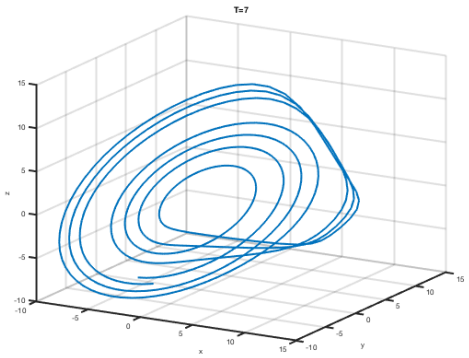


(e) Period 5 cycle.

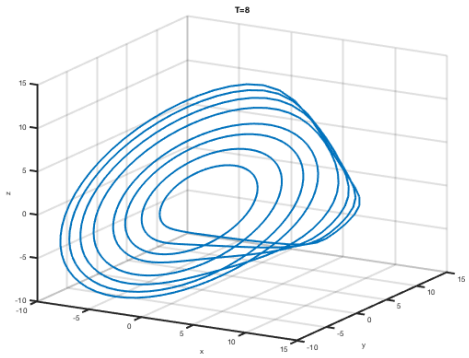


(f) Period 6 cycle.

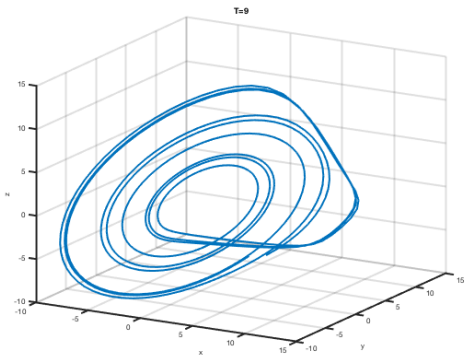
Figure 4.27: Estimates of complete set of lower order noise-free Rossler UPOs (periods 1-6), calculating by averaging detected cycles from 50 samples of the noise-free time series.



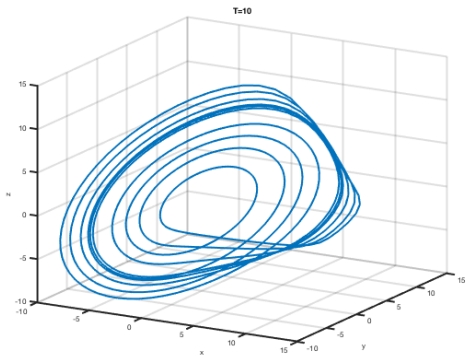
(a) Period 7 cycle.



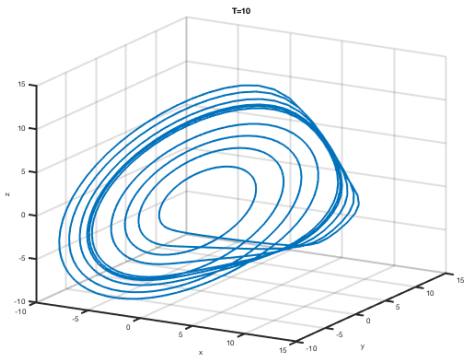
(b) Period 8 cycle.



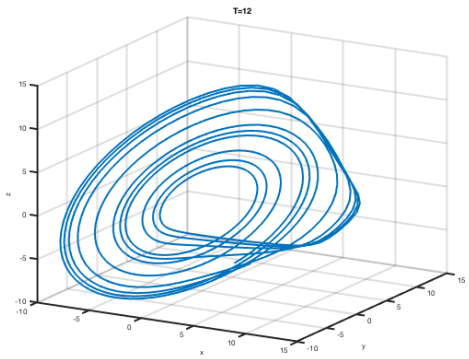
(c) Period 9 cycle.



(d) Period 10 cycle.



(e) Period 11 cycle.



(f) Period 12 cycle.

Figure 4.28: Estimates of complete set of lower order noise-free Rossler UPOs (periods 7-12), calculating by averaging detected cycles from 50 samples of the noise-free time series.

4.5.2 Lower Order UPOs from Time Series with 25% Added Noise

Figure 4.29 shows the recurrence histogram aggregating the numbers of shadow-UPOs detected from 50 noise-infected Rossler time series. The pattern is similar to that in Figure 5.26, with the lower order periods T–$3T$ dominating the result. In Tables 4.4 and 4.5 we summarise the numbers of noise-infected shadow-UPOs detected, noting that now nearer to 80% of the cycles detected are periods T–$3T$. Also partial shadow-UPOs have increased as a proportion of total detected cycles, and complete cycles reduced as expected due to noise.

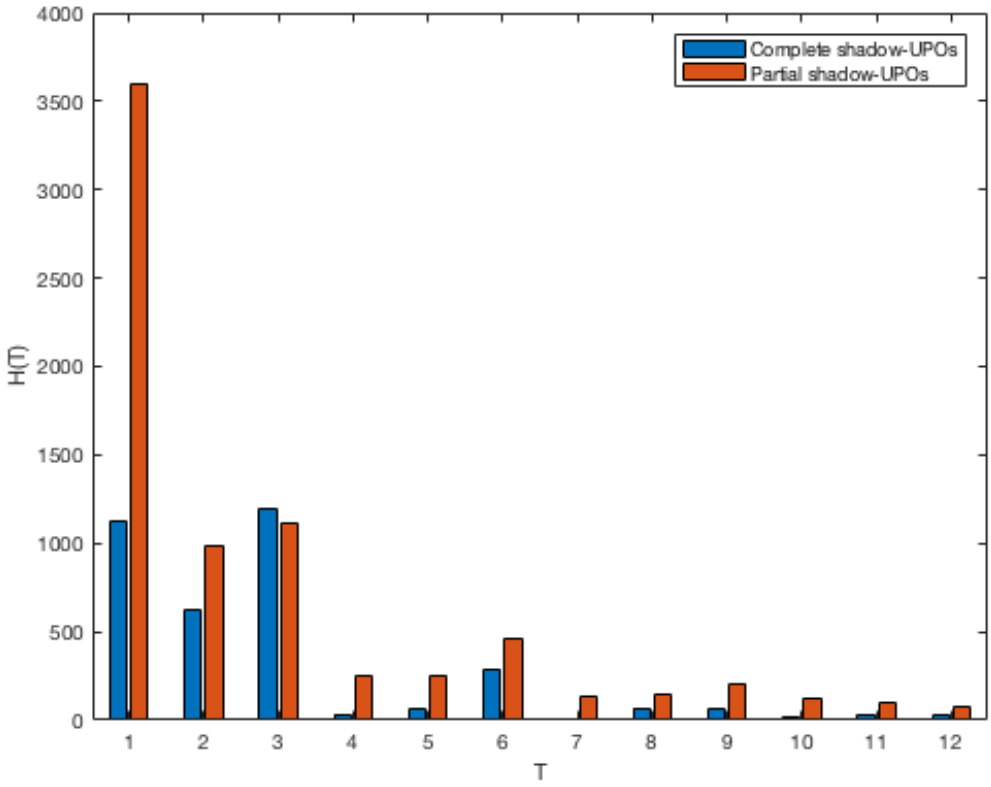


Figure 4.29: Histogram representation of complete and partial cycles detected from 50 pre-filtered noise-infected (25%) Rossler system time series samples, each of length 8,000 points.

Period	Number of Full Cycles	Number of Partial Cycles	Total Cycles
1	1124	3600	4724
2	623	990	1613
3	1201	1118	2319
4	34	246	280
5	59	253	312
6	288	463	751
7	2	134	136
8	69	143	212
9	59	209	268
10	22	127	149
11	28	97	125
12	**27**	**79**	**106**

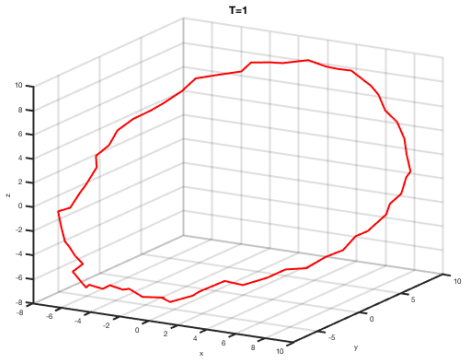
Table 4.4: Full and partial cycles detected from 50 noise-infected time series samples of 8,000 points from the pre-filtered Rossler system with 25% added GWN noise, each of length 8,000 points. There is a significant increase in the number of T=1 partial cycles as expected.

Period	% Full Cycles	% Partial Cycles	% Total
1	32%	48%	43%
2	18%	13%	15%
3	34%	15%	21%
4	1%	3%	3%
5	2%	3%	3%
6	8%	6%	7%
7	0%	2%	1%
8	2%	2%	2%
9	2%	3%	2%
10	1%	2%	1%
11	1%	1%	1%
12	1%	1%	1%

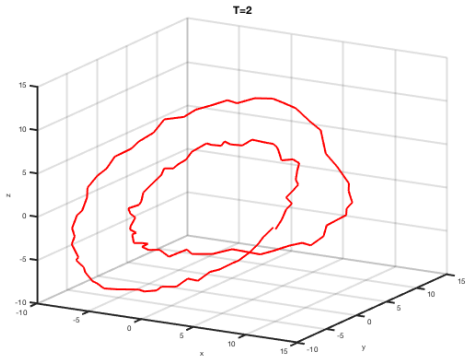
Table 4.5: Summary table of number of detected cycles by period from 50 noise-infected time series samples. Full cycles comprise 32% of the total number of cycles detected.

In Figure 4.30 we provide several examples of individual noise-infected shadow-UPOs prior to noise-filtering. In Figure 4.31 we show the estimates of noise-free

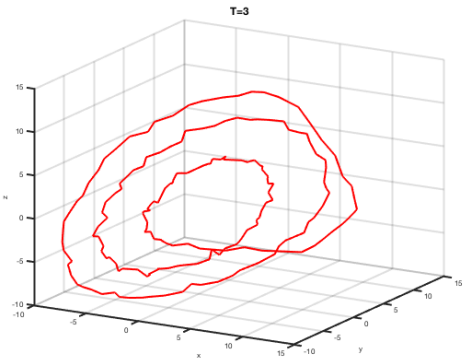
UPOs after averaging the binned data. The results in this section show that we can still get good estimates of the lower order UPOs using noise-infected data, however we rely on a smaller number of detected complete cycles and the approximation is not as accurate using bins that are lightly populated.



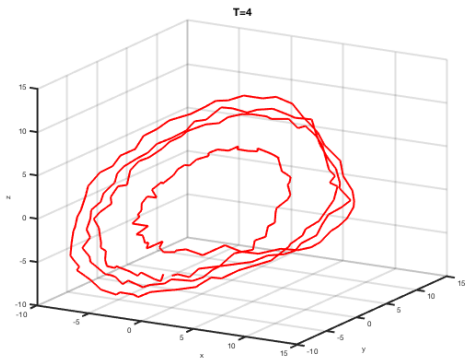
(a) Period 1 cycle.



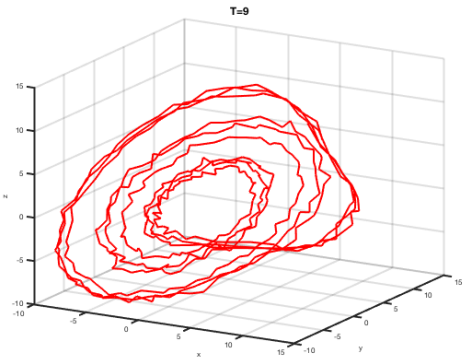
(b) Period 2 cycle.



(c) Period 3 cycle.

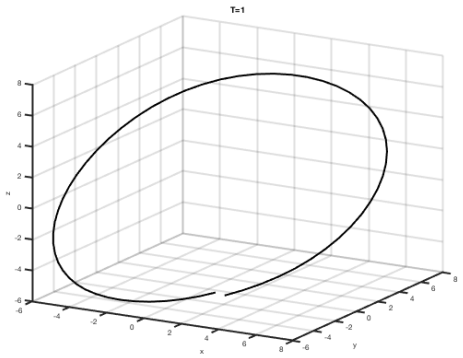


(d) Period 6 cycle.

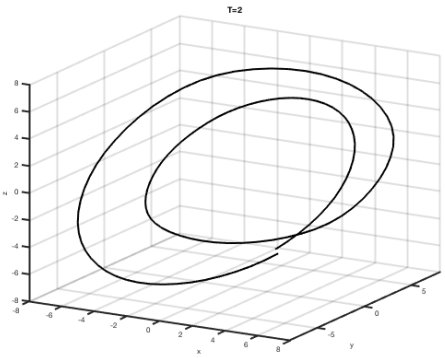


(e) Period 9 cycle.

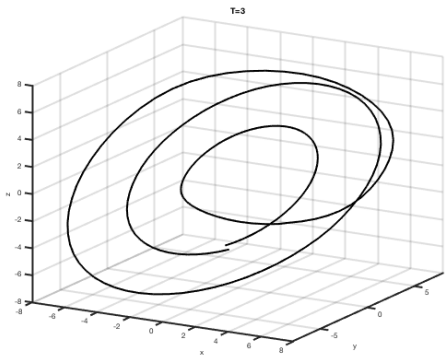
Figure 4.30: Examples of original 25% noise-infected Rossler shadow-UPOs (periods T, 2T, 3T, 6T and 9T).



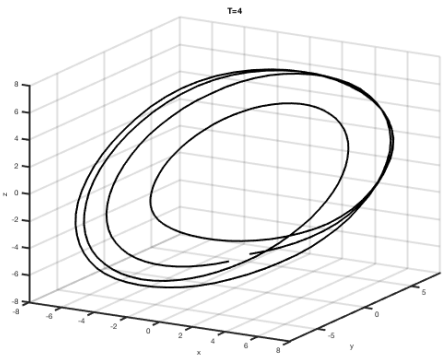
(a) Period 1 cycle.



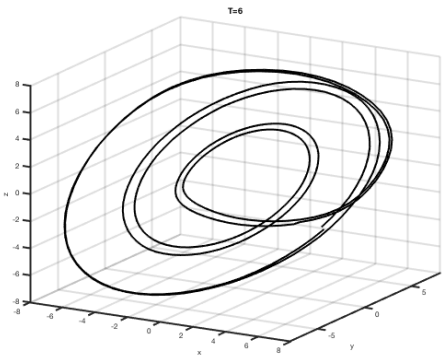
(b) Period 2 cycle.



(c) Period 3 cycle.



(d) Period 6 cycle.



(e) Period 9 cycle.

Figure 4.31: Estimates of Rossler UPOs (periods T, 2T, 3T, 6T and 9T) after averaging the detected and synchronised noise-infected shadow-UPOs.

4.6 Lyapunov Exponent of Individual UPOs

Using the SUNR method described in previous sections, we obtain collections of shadow-UPOs in bins. We have shown that if we can accumulate sufficiently large numbers of noise-infected cycles in each bin, then they may be individually filtered or averaged to reduce noise and estimate the underlying lower order UPOs directly. Large pools of cycles containing measurement noise appear to respond well to averaging and provide solid estimates of the UPOs. In particular, averaging can reduce broadband noise that is difficult to address using Fourier techniques involving frequency selection.

The lower order UPOs form a useful basis for describing the attractor. The attractor may be approximated by UPOs and retain its inherent characteristics. The average maximal Lyapunov exponent is a useful indicator of the presence of chaos and the degree of orbital divergence within a chaotic system. However, as discussed in Chapter 1, within a local neighbourhood, orbits shear apart or converge in accordance with the short time Lyapunov exponents representing the specific dynamics of that neighbourhood. Short time Lyapunov exponents are defined in Equation (1.27). The length of each lower order UPO provides an excellent measure of "short time". We propose that a highly meaningful definition of orbital divergence is defined in terms of the short time Lyapunov exponents of the lower order UPOs.

During the time evolution of an orbit, it will journey along a path "near" a lower order UPO. It could be traveling on a much higher order cycle near the lower order cycle, but for quantitative purposes it does not matter. If we can calculate a short time Lyapunov exponent for each lower order UPO, then this will approximately describe the deformation of path from the start to the end of the cycle. Furthermore to calculate the short time Lyapunov exponent for the cycle, we only need start and end point data. We therefore wish to calculate short time Lyapunov exponents for UPOs of each lower order period and refer to these as *UPO Lyapunov exponents*.

The data stored in the bins when estimating the noise-free basis set of lower order UPOs can be used to directly calculate a short time Lyapunov exponent for each cycle period. The UPO Lyapunov exponents may then be combined in a weighted average to estimate the global or average maximal Lyapunov exponent.

The methodology we adopt to estimating local periodic Lyapunov exponents is based on a technique used by Sano and Sawada [117] to compute the Lyapunov exponents of a noise-free time series, but we shall modify it for cycles. The Sano-Sawada technique identifies a set of points within a small ball of radius ε of a selected point. A matrix of difference vectors is constructed between each point in the ε-ball and the selected point x_j. Each point in the ball is evolved forward by n time steps and again the matrix of difference vectors is constructed. A linearisation is constructed between the start and end difference matrices:

$$\xi(t) = A^t.\xi(0), \tag{4.12}$$

where $\xi(t)$ is the resultant difference matrix at time t and A^t is a linear operator which maps tangent vector $\xi(0)$ to $\xi(t)$ and is in fact the Jacobian $J(t)$.

The Lyapunov exponents are determined as the eigenvalues of $J(t)$. Another point is then selected with sufficient neighbours within an ε-ball and the procedure is repeated, and so forth. The final set of Lyapunov exponents is calculated by averaging all those determined along the orbit. We shall modify this process for cycles, where we know that once suitably synchronised, the start points all fall within an ε-ball for an appropriately selected value of ε.

Define a small ball of radius ϵ, centered at the j^{th} entry in the time series; x_j . Let us find the set of points $\{x_{r_i}\}$ with $i = 1, 2, ..., N$ included within this ball:

$$\{y^i\} = \{(x_{r_i} - x_j) : \|(x_{r_i} - x_j)\| \leq \epsilon\}, \tag{4.13}$$

where y^i is the displacement vector between x_{r_i} and x_j. We use the L_2 (Euclidean) norm for the distance calculation.

After the passing of a time interval $\tau = p\Delta t$, the cycle point x_j will now progressed to x_{j+p} and the neighbouring points $\{x_{r_i}\}$ within the ball will have progressed to $\{x_{r_i+p}\}$. Now if we consider displacement vectors, our initial displacement vector $y^i = (x_{r_i} - x_j)$ is mapped to the set

$$\{z^i\} = \{(x_{r_{i+p}} - x_{j+p}) : \|(x_{r_i} - x_j)\| \leq \epsilon\}. \tag{4.14}$$

We assume the radius ϵ is sufficiently small for the displacement vectors $\{y^i\}$ and $\{z^i\}$ to be reasonable approximations of the tangent vectors in the tangent space. We can now represent the evolution of y^i to z^i as the matrix equation:

$$z^i = B_j y^i. \tag{4.15}$$

The $N \times N$ matrix B_j is an approximation at x_j of the flow map J^t.

The next step is to optimize the estimation of the flow map B_j using the data stored in $\{y^i\}$ and $\{z^i\}$ and for this the least-squares error algorithm is applied. This minimises the average of the L_2 distance norm between z^i and $B_j y^i$ with respect to all components of B_j as follows:

$$\min_{B_j} S = \min_{B_j} \sum_{i=1}^{N} \left\| z^i - B_j y^i \right\|^2. \tag{4.16}$$

The eigenvalues of B_j are the estimates of the local cycle Lyapunov exponents at x_j. This is illustrated by the diagram in Figure 4.32 below.

In order to apply this technique to binned shadow-UPOs we need to synchronise the cycles as they will not necessarily all be in phase. We can synchronise cycles using the time lag associated with maximum cross-correlation values as before. The method requires we construct a matrix Z of start points and a matrix Y of end points for each bin. Applying noise-reduction methods to each shadow-UPO in a given bin is unlikely to yield much improvement given we are using single points and will deal with the noise through the averaging process inherent in using a set of cycles.

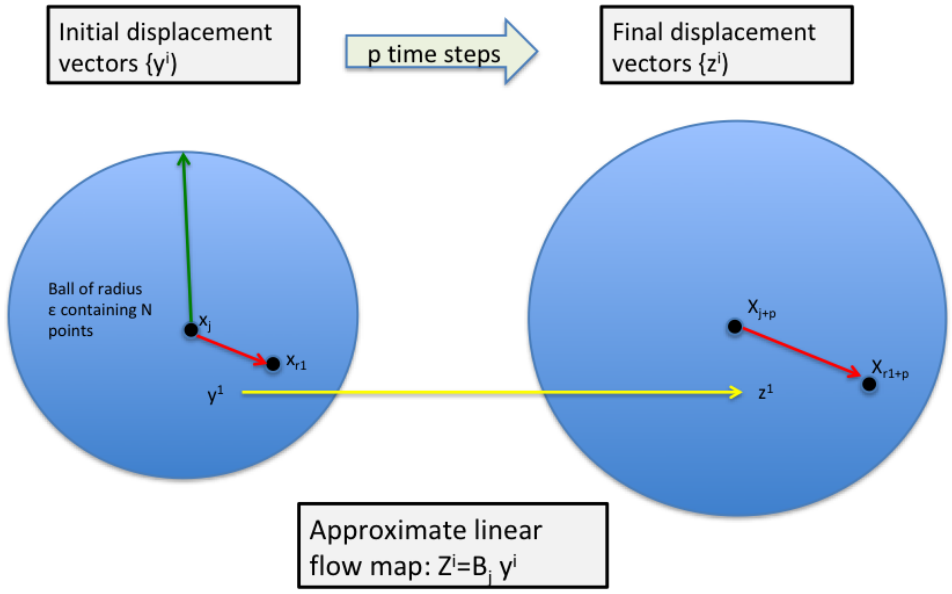


Figure 4.32: We are fitting an approximate linear mapping to the displacement vectors in the ball of radius ε centered on x_1 and following their evolution for p time steps (1 period), where the central point will be x_{1+p}. The Lyapunov exponents for a cycle bin will be determined as the eigenvalues of the $N \times N$ matrix B_j .

Figures 4.33 shows the importance of synchronising shadow-UPOs as initial points may be well out of step with each other. In Figure 4.33 we show as red dots the starting points of 383 individual cycles, collected from 12 sample time series of 10,000 points each. They are widely distributed. We need to synchronise cycles, aligning the start points as closely as possible to calculate difference vectors between points included in a ε-ball centred at an "anchor" point. In Figure 4.34 we show the same 383 cycles after they have been aligned using the cross-correlation of each with an "anchor" cycle. The start points are now sufficiently close to calculate a matrix of difference vectors. The "anchor" point is identified as that

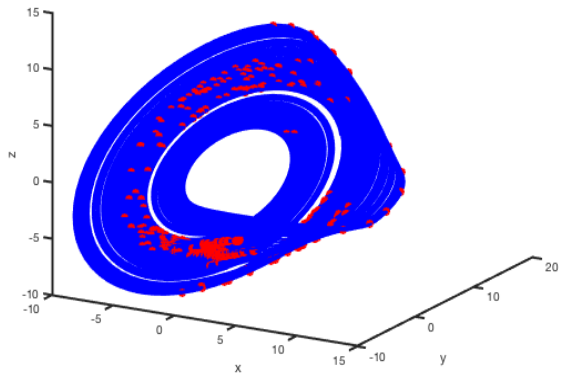


Figure 4.33: The set of detected complete period 3T cycles from 12 samples of 10,000 points. In total there are 383 cycles shown in blue. The start points are illustrated in red and are highly dispersed.

with minimum total L_2 distance to all the other points. We also eliminate any isolated points that remain too far from our "anchor" point, in this case a distance of more than 1.6.

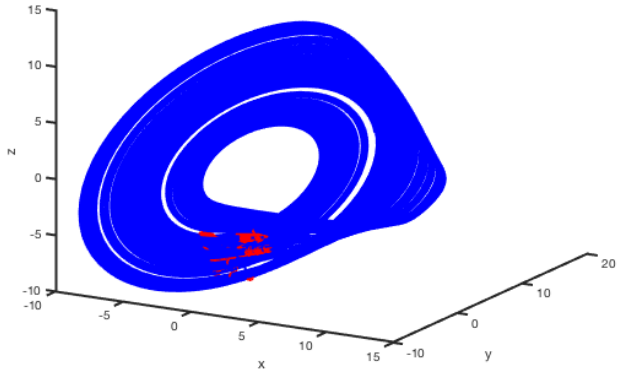


Figure 4.34: The set of detected 383 period 3T cycles, with start points synchronised.

Period	Weight	Noise=0% UPO LE	Noise=25% UPO LE
1	15.2%	0.081	0.087
2	10.9%	0.059	0.063
3	35.5%	0.084	0.090
4	2.4%	0.060	0.065
5	3.0%	0.081	0.086
6	15.9%	0.081	0.087
7	2.8%	0.063	0.068
8	3.2%	0.084	0.090
9	3.9%	0.087	0.094
10	7.1%	0.087	0.093

Table 4.6: Maximal Lyapunov exponents calculated for individual Rossler UPOs (Periods 1-10), using the Sano-Sawada methodology described above, adapted to bundles of shadow-UPOs. The weighted average estimates of the average global Lyapunov exponent are 0.08 (1/s) and 0.09(1/s) using noise-free and noise-filtered cycle bundles, respectively.

The estimated noise-free values and noise-reduced values of the UPO Lyapunov exponents are collected in Table 4.6 above. We found that 10 samples of the time series were sufficient obtain reasonable estimates of the short time Lyapunov exponents. The weights are calculated for both noise-free and 25% noise-infected shadow-UPOs from the pre-filtered data using the proportion of periodic points of a given period to the total. The total points of a given period are calculated by summing across the row of the reduced recurrence matrix. The noise-free and noise-reduced UPO Lyapunov exponents are reasonably close. By constructing a difference matrix using many periodic points collected within a ball and calculating the start-end point vectorial divergence after one period, we are also effectively averaging out noise without having to individually filter entire cycles. The weighted average estimates of the average global Lyapunov exponent are 0.08 s^{-1} and 0.09 s^{-1} using noise-free and noise-filtered cycle bundles, respectively.

4.7 Summary and Discussion

In this chapter we described the process of creating model chaotic time series with added noise for the purpose of testing the shadow-UPO based noise reduction (SUNR) method, a methodology for detection, extraction, processing and noise-filtering of the detected shadow-UPOs from a noisy time series. The time series is simply created through numerical integration of the system equations. We show how to generate time series infected by any of uniform white, Gaussian white, high frequency and pink measurement noise. We also specify the process for adding dynamical noise in Section 4.2.5. At this stage we focus on a single chaotic system, the Rossler system, to illustrate the SUNR method. The process will be repeated in Chapter 5 for several other systems with differing levels of added noise.

Our focus in this chapter is addressing chaotic time series with lower instability (i.e. at the lower end of the range $0 < \lambda_{\max} \leq 1$). For these systems the recurrence histogram model for detecting harmonics and shadow-UPO locations will detect a sizeable number of complete cycles. This may not be the case for more strongly chaotic systems and the necessary modifications are addressed in Chapter 5. We provide a heuristic example using the Rossler system, and show how to produce estimates of the basis set of noise-free lower order UPOs, using binning and simple averaging. We treat bins with many detected cycles differently than those with only a few. Highly populated bins are amenable to averaging after synchronising detected orbits, as many will be out of step. If the noise has specific characteristics amenable to a frequency attenuation technique, then we suggest individually filtering each cycle in the frequency domain using a FFT. For less populated bins we select the "least unstable" cycle for Fourier filtering, where the selected cycle has the smallest gap between start and end point. The averaging approach is particularly useful when noise is broadband in nature (e.g. Gaussian) and it is difficult to eliminate using frequency selection approaches.

Using a single noise-free Rossler time series we can detect most different lower order UPOs but are unlikely to find all. Some will just be too unstable to locate using a single sample. For the Rossler system the period 7 cycle is difficult to

detect for example. We thus apply multi-sampling and confirm the benefits of this approach. We usually get many examples of each cycle. Initially, using a noise-free Rossler time series we detect all UPOs up to period 12 and using averaging we obtain excellent approximations of the underlying UPOs. We then apply the method to a Rossler time series with 25% Gaussian white noise added, and again find all of the lower order cycles. The noise-reduced UPO approximations appear to closely match the estimates from noise-free data, and the goodness of fit considerations are addressed in Chapter 5.

Finally we propose that the set of lower order UPOs coupled with their associated short time Lyapunov signature form a meaningful description of the chaotic system. The problem of orbital divergence is then discretised down to a set of UPO Lyapunov exponents describing the divergence depending upon which UPO the orbit is traveling on or near to. A short time Lyapunov exponent can be calculated for each UPO period using the data bins we have already populated to determine the noise-free UPO basis set. The binned data is not noise-filtered as the method employed to determine the UPO Lyapunov exponents only uses start-end points and multiple cycles, implicitly noise-filtering the result. To calculate an average short time Lyapunov exponent using synchronised clustered shadow-UPOs, we modify a technique used by Sano and Sawada [117] that was applied to general segments of orbit. This numerical technique was also originally used by Lathrop and Kostelich [84] in the context of UPOs for the B-Z system, but not as a noise reduction tool. The average maximal Lyapunov exponent for the system may be estimated using a weighted average, using the relative proportions of periodic points. We illustrate for the Rossler system with 25% added GWN, obtaining a set of maximal UPO Lyapunov exponents and estimates of the global average Lyapunov exponent that are approximately equal to the true result.

Chapter 5

Adaptions for Higher Instability Systems

5.1 Chapter Overview

In this chapter we consider the noise-reduction problem for chaotic dynamical systems with greater instability than those discussed in Chapter 4. In these systems the instability is such that we may not be able to detect sufficient complete cycles of all periods to estimate a full noise free basis set of lower order UPOs. The instability may result in only specific periods being detectable, fewer complete cycles and more partial cycles; or both. For these time series we can still apply the binning and filtering version of the SUNR method that was described and illustrated in Chapter 4. However we may obtain an incomplete set of noise-free basis UPO estimates, each with less statistical credibility than those derived from bins with large numbers of detected cycles.

For these systems with higher instability we may not be able to obtain noise-free estimates of the lower order UPOs, but we can still estimate the noise-free time series. Provided the residual noise is sufficiently low in the approximated noise-free time series, we may then characterise the chaotic dynamical system by using this approximation in the conventional numerical algorithms to determine invariants. This alternative approach involves individually noise-filtering all complete and

partial shadow-UPOs and using these to construct an approximation of the noise-free time series. In this way we maximise use of the available information.

For such chaotic systems the SUNR method is again, for consistency, illustrated using a noise-infected time series derived from the chaotic Rossler system and we explore the effect of different types of measurement noise and different levels of noise. We detect, extract and process complete and partial cycles from the noise-infected data using our modified recurrence plot and recurrence histogram. For comparative purposes, noise is reduced from detected cyclic sequences using a simple Fourier low-pass filter. We emphasise that filtering individual shadow-UPOs allows greater noise reduction if we know the type of noise and can apply specific signal processing techniques. We illustrate with an example using pink noise and a targeted approach to separating the signal from the noise with excellent results. The noise-reduced approximation of the underlying time series is then used to directly calculate the maximal Lyapunov exponent using the Wolf algorithm and results are discussed. We also test the SUNR method on time series infected with dynamical noise, analyse the results and identify the limitations. Based on the observation that dynamic noise infected time series remain smooth but deformed, we present a simple alternative approach for dynamical noise using multi-sampling.

Finally, we test the SUNR method on several other chaotic systems of varying topology including the Chua, Lorenz, Rabinovich-Fabrikant and Lu-Chen systems. As in Chapter 3 and 4, all time series considered will contain 10,000 points unless stated otherwise. We measure goodness of fit results and estimate the maximal Lyapunov exponent in each case. The SUNR method works well for the relatively "low-instability" Rossler, Lorenz and Chua systems and this is confirmed by our goodness of fit results. The Rabinovich-Fabrikant system is introduced as an example of where the recurrence matrix detection method runs into difficulty. The Lu-Chen system is included as an example of a "higher-instability" system and we vary our approach to detecting cycles for these systems to get more favourable results. We apply the SUNR method to multi-sampled data taken from the Rossler, Lorenz and Chua systems, averaging the estimated noise-free Lyapunov exponent results and see that this approach significantly improves the computed results.

Before we summarise the results described above, we first define a set of performance metrics to measure the goodness of fit between the approximation and the noise-free time series. In all cases the SNR is reduced; however, this single measure is inadequate to measure the overall effectiveness of the approximation and better metrics are required for this purpose.

5.2 Goodness of Fit Measures for Time Series

The SUNR method produces as output an approximation of the noise-free time series, and we need to assess the goodness of fit of that approximation. There has been much discussion in the literature around the best method to test for equality of time series. The goal is to find a distance or statistical measure that will indicate the level of similarity or dissimilarity between time series. An excellent summary of methods is provided in the user guide for the TSdist software [107]. Following the categorisation introduced in [46], the time series distance measures can be divided into four categories and we describe each briefly.

Shape-based measures: This category of distance measures is based on directly comparing the data values and the shape of the series in different ways. Primary focus is on L_p distances derived from the norms. These are defined distance metrics that can only compare series of the same length. Also included in this class are distance measures designed to accommodate variations in the time series under comparison. The Dissim distance for example is designed for when sampling rates are different. In order to overcome the restrictions of rigid distances such as Euclidean distance, other similarity measures have been specifically developed. A popular approach is Dynamic Time Warping (DTW) [10]. This distance measure is specifically designed able to deal with transformations such as local warping and shifting, and allows the comparison between series of different lengths.

Edit-based Measures: Edit distance calculates the similarity between two sequences of strings based on the idea of counting the minimum number of edit operations (delete, insert and replace) that are necessary to transform one sequence into the other.

Features-based Measures: This category of distance measures focuses on extracting a set of features from the time series and calculating the similarity between these features. Examples are Pearson's correlation and Fourier coefficients based distance.

Structure-based measures: Structure-based distances are designed to identify higher-level structures in long series. Structure-based distances include (i) model-based approaches, where a model is fit to each series and the comparison is made between models, and (ii) complexity-based models, where the similarity between two series is measured based on the quantity of shared information.

As well as direct goodness of fit measures like those described above, there are also measures derived from the time series like dynamical invariants; in particular the maximal Lyapunov exponent. We will also be calculating these and comparing with the noise-free scenario.

5.3 Goodness of Fit Measures Applied

Our view is that there is no single perfect method to measure the quality of the approximation; it is best to apply a number of different metrics that measure the most important qualities of the specific approximation and view the results collectively. Given we have two time series of the same length and sampling frequency, we apply the L_p distance measures, also calculate a features-based (statistical) metric and a measure of phase synchronicity. We will measure the goodness of fit between the approximation and original noise-infected data for both the scalar time series and the associated delay vectors. The distance measures can diverge exponentially rapidly if the two time series are even slightly out of phase so ensuring phase synchronicity is imperative. Our goodness of fit (GOF) measures are summarised as follows.

5.3.1 Distance Measures

Let $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ be two time series. The entries in these time series may be a single scalar value or a m-dimensional point in phase space (a delay vector). The L_1, L_2 and L_∞ distances are $d_1 = \sum_{i=1}^{n} |x_i - y_i|, d_2 = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2}$ and $d_\infty = \max |x_i - y_i|$ respectively.

The goodness of fit measure using the L_1 norm is the *mean absolute error*, $MAE = d_1/n$. The L_1 norm accommodates outliers without giving them a higher weighting and produces a lower error in such cases when compared to the L_2 norm.

The goodness of fit measure using the L_2 norm is the *root mean square error*, $RMSE = d_2/\sqrt{n}$. The *RMSE* represents the standard deviation of the residuals (prediction errors) and is a measure of the spread. Since the errors are squared before they are averaged, the *RMSE* gives a relatively high weight to large errors. The *RMSE* is a measure of the "quality" of an estimator, it is always non-negative, rotationally invariant, and values closer to zero are better. Like the standard deviation, the *RMSE* has the same units of measurement as the quantity being estimated.

The goodness of fit measure using the L_∞ norm is the *mean maximum error*, $MME = d_\infty/n$. This is where the distance between two vectors is the greatest along any coordinate dimension. It highlights the "worst" dimensional fit at each point and we would like this as small as possible.

5.3.2 Coefficient of Determination or R^2

In statistics, the *coefficient of determination* denoted by R^2 is a measure of predictability between a dependent variable and an independent variable. R^2 represents the proportion of the variance in the dependent variable that results from using the independent variable to predict the dependent variable. It measures the effectiveness of the model at replicating observed outcomes. Let our time series x have n values denoted by $x_1, x_2, ...x_n$. Each value of x has an associated

predicted (or modelled) value $f_1, f_2, ...f_n$, where f is the prediction vector.

Define the set of residuals (or errors) as $e_i = x_i - f_i$ and the mean of the observed data $\overline{x}$ as usual by:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{5.1}$$

The variability of the time series can be expressed using three "sums of squares" quantities; the total sum of squares SS_{tot}, the explained sum of squares SS_{reg} and the residual sum of squares SS_{res}. These are defined as $SS_{tot} = \sum_{i=1}^{n}(x_i - \overline{x})^2$, $SS_{reg} = \sum_{i=1}^{n}(f_i - \overline{x})^2$, and $SS_{res} = \sum_{i=1}^{n}(x_i - f_i)^2$ respectively.

The coefficient of determination, R^2 is defined as:

$$R^2 \equiv 1 - \frac{SS_{res}}{SS_{tot}}. \tag{5.2}$$

R^2 is related to the *proportion of unexplained variance* (PUV), with the ratio term comparing the variance of the errors in the model (unexplained variance) with the data's total variance as:

$$R^2 \equiv 1 - PUV. \tag{5.3}$$

For example if $R^2 = 0.64$, then 64% of the variability between the two variables is explained using the dependent variable. The outstanding 36% of the variability is still unaccounted for. The coefficient of determination, R^2, will take values between 0 and 1.

There are a number of interpretations of the coefficient of determination in use. Firstly, it is commonly used in the context of least squares regression. In that case R^2 represents the square of the Pearson correlation coefficient relating the regressor and the response variable. The other context is broader, where predicted values are generated from a more general model, not necessarily linear

least squares regression. This is the context with which we utilise R^2. In this case, R^2 is calculated as:

$$R^2 = r_{xf}^2, \tag{5.4}$$

where r_{xf} is the correlation coefficient between the x data values and the modelled values. The R^2 value is thus a measure of the goodness of fit of the predictor that is constructed from the modelled values.

5.3.3 Phase Lag using Cross Correlation

The purpose of this metric is to ascertain the amount of any phase shift introduced as a result of the approximation process. For the comparison between the approximant and noise-free time series we determine the overall maximum of the cross-correlation between them and the lag associated with the maximum. If we are given two time series, x_t and y_t , we can delay x_t by T samples and then calculate the *cross-covariance* between the two signals as follows:

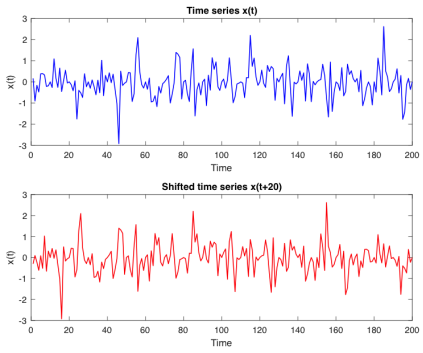
$$\sigma_{xy}(T) = \frac{1}{(N-1)} \sum_{t=1}^{N} (x_{t-T} - \mu_x)(y_t - \mu_y). \tag{5.5}$$

where μ_x and μ_y are the respective means of each time series with N samples. The cross-correlation function is the normalised version:

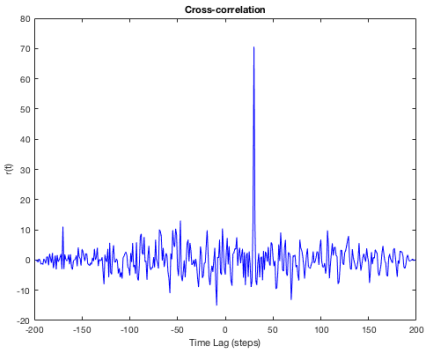
$$r_{xy}(T) = \frac{\sigma_{xy}(T)}{\sqrt{\sigma_{xx}(0)\sigma_{yy}(0)}}. \tag{5.6}$$

Note that $\sigma_{xx}(0) = \sigma_x^2$ and $\sigma_{yy}(0) = \sigma_y^2$ are the variances of the respective signals so that

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}. \tag{5.7}$$



(a) Phase-shifted time series fig:a



(b) Cross-correlation

Figure 5.1: (a) Two identical signals that are out of phase by 20 time steps, (b) Plot of cross-correlation between the two time series that are out of phase by 20 steps shown in Fig (5.1). Note the clear peak at time lag=20.

We express the phase lag (if any) as number of time steps required to ensure maximum of the cross-correlation between the time series.

5.4 Heuristic Example: The Rossler System

We demonstrate the SUNR process using a numerically integrated model time series for the Rossler system with step-size $h = 0.1$ and 25% added white Gaussian noise. The noise-infected attractor is shown in Figure 4.8.

5.4.1 Detecting and Extracting Complete and Partial Cycles

As shown in Chapter 3, we construct the recurrence histogram which is highly resilient to measurement noise. We apply a critical radius of 1.6 (approximately 5% of the attractor extent) to detect periodic points, using the scalar time series without embedding. We observe the horizontal lines on the recurrence plot corresponding to sequences of consecutive points with the same period. We summarise the detected shadow-UPOs in Table 5.1, where each period is a multiple of the fundamental period T (approximately 60 time steps).

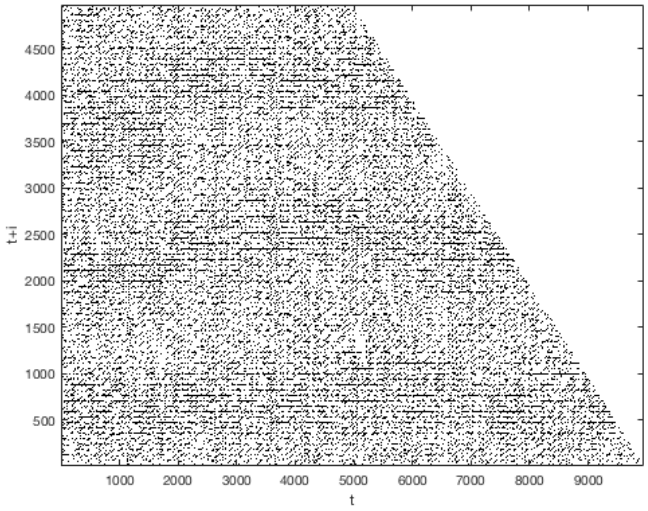


Figure 5.2: Horizontal recurrence plot using pre-filtered scalar Rossler time series, containing 25% added white Gaussian noise. Horizontal segments represent shadow-UPOs.

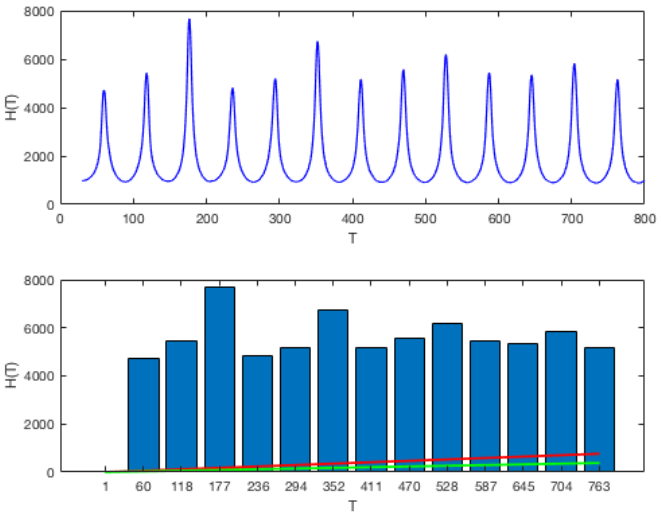


Figure 5.3: Histogram derived from the recurrence matrix. Peaks represent periodic points occurring with the greatest frequency. The associated frequencies are shown more clearly on the bar chart.

Period Number	**Period**	**Number of Full Cycles**	**Number of Partial Cycles**
1	60	27	18
2	118	13	20
3	177	12	33
4	236	0	6
5	294	3	6
6	352	4	9
7	470	0	2
8	528	0	2
9	587	0	2
10	645	0	2
Total		59	100

Table 5.1: Detected Rossler shadow-UPOs of each period. Partial cycles are sequences of length $\geq 50\%$ of a period.

The recurrence plot and recurrence histogram are shown in Figures 5.2 and 5.3 respectively. Histogram peaks are identified using the peak detection algorithm in Matlab. The "findpeak" function returns a vector with the local maxima (peaks) of the input signal vector. A local peak is a data sample that is either larger than its two neighbouring samples; and they represent periodic points occurring with greatest frequency. In the case of a chaotic system with relatively low instability (small maximal Lyapunov exponent) we will likely detect many complete cycles at periods corresponding to the peaks. For chaotic systems that are highly unstable (large maximal Lyapunov exponent) we may only detect partial or incomplete cycle sequences. The location of peaks, in the presence of high levels of measurement noise, is enhanced considerably by pre-filtering the noise-infected data with a centered moving average. In Figure 5.3 the associated periods are shown more clearly on the lower bar chart. The red line represents the height of a complete cycle period and the green line represents the minimum number of points for a partial cycle to be included. The bar heights must be above these lines for at least one shadow-UPO or partial shadow-UPO to be detected. In this example there are many periodic points detected in excess of the minimum.

Using the information on peak periods, we extract the relevant rows relating to these periods from the recurrence matrix and search for sequences of consecutive periodic points. The recurrence matrix method implicitly requires two consecutive complete cycles to be present to acknowledge that cycle. Thus complete cycles appear in consecutive conjugate pairs and this is accommodated in our methodology.

5.4.2 Filtering Complete Cycles to Reduce Measurement Noise

After linearly transforming the detected cycles using an end-point smoothing transformation to ensure they are closed (end points join), we next apply an appropriate noise filter to each complete cycle. This transformation is inverted after noise filtering. In this case, for illustration purposes we use a Fourier filter in the frequency domain, removing all amplitudes attached to frequencies > 0.8. See Figures 5.4–5.7 below for an example of a detected cycle of period $5T$ that is noise-filtered using a cut-off (low-pass) filter.

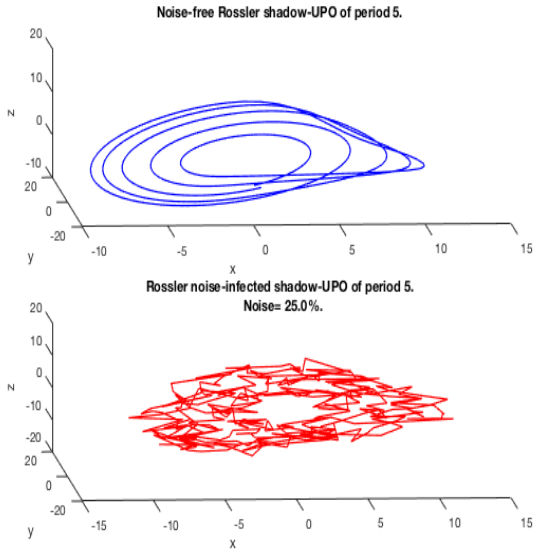


Figure 5.4: Complete period $5T$ cycle. Top: Noise-free. Bottom: 25% added Gaussian white noise.

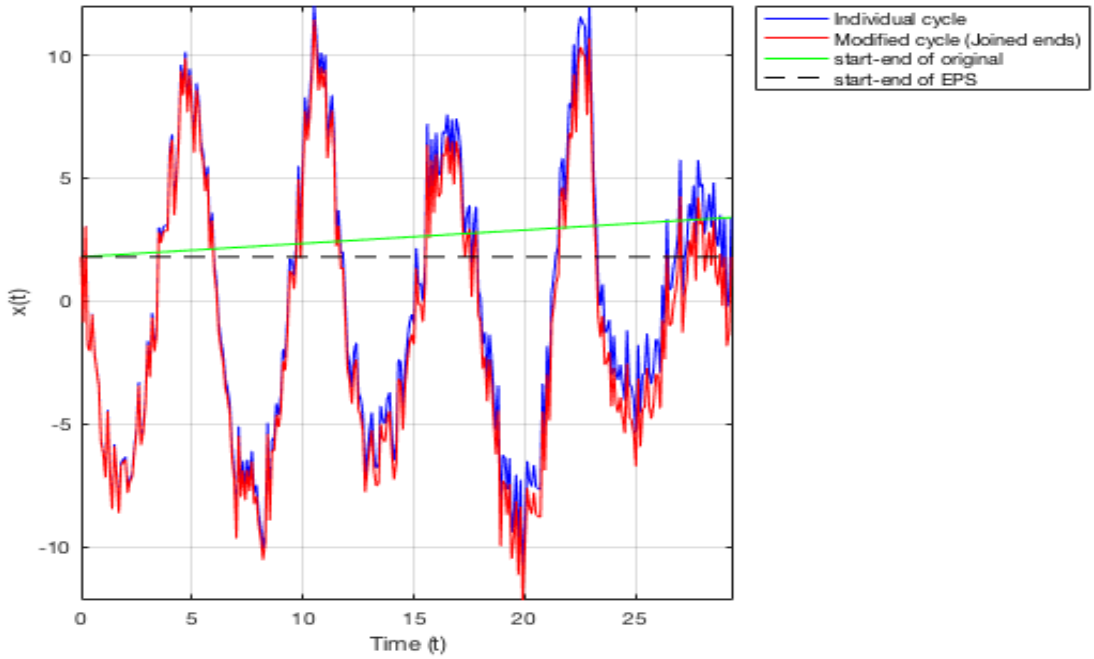


Figure 5.5: Trace of x-coordinate for noisy period $5T$ shadow-UPO, before and after end point smoothing.

In this illustrative example, the Fourier filter acts as a low pass filter and is similar to the application of a moving average. In practice, the type of noise will vary and more advanced signal processing techniques can be applied to the closed cycle. This can be expected to yield significantly better results than with a simple low-pass filter.

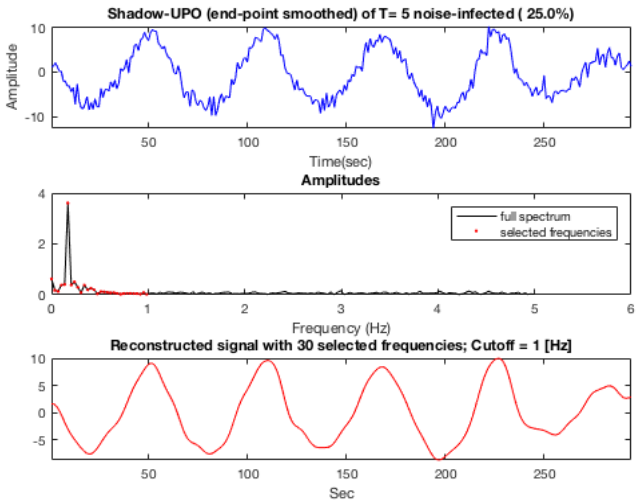


Figure 5.6: Period $5T$ shadow-UPO before (Top) and after (Bottom) the application of the low-pass filter. There is a marked improvement in quality of the cycle.

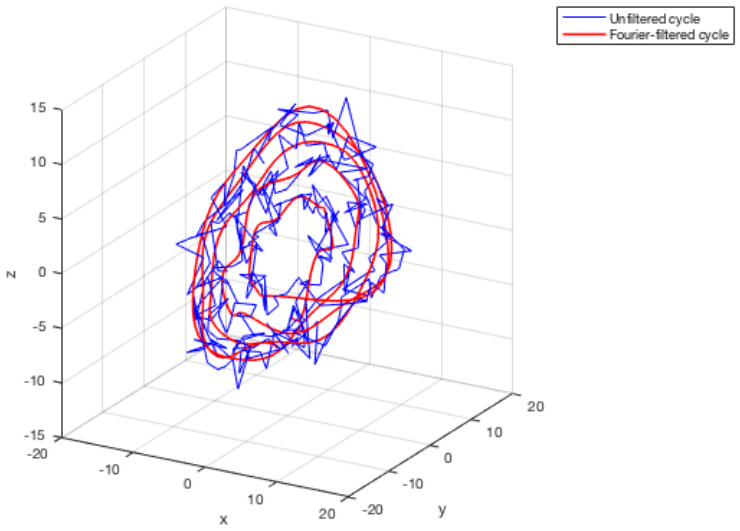
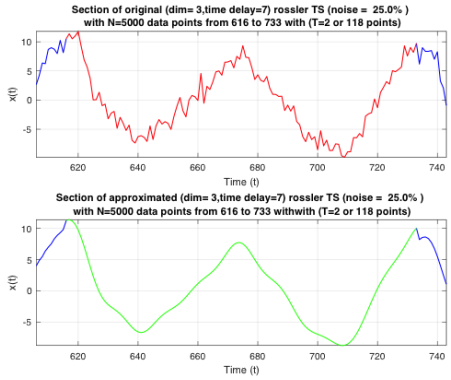
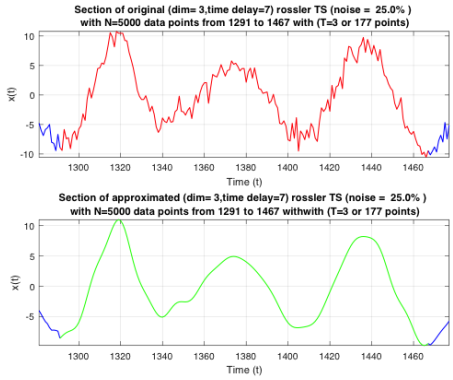


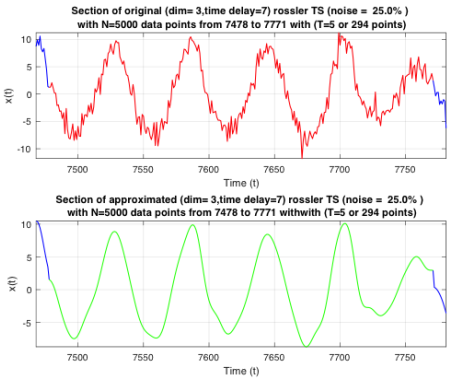
Figure 5.7: Period $5T$ shadow-UPO. Direct comparison in phase space of cycle before (blue) and after the application of the low-pass filter (red).



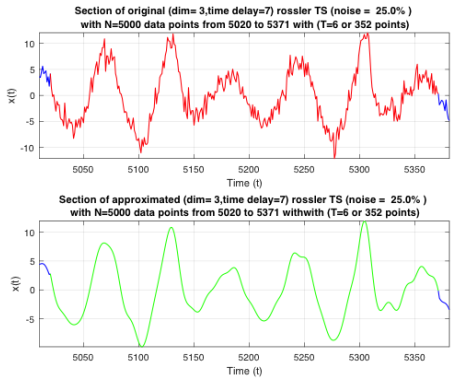
(a) Period 2 cycle.



(b) Period 3 cycle.



(c) Period 5 cycle.



(d) Period 6 cycle.

Figure 5.8: One can see the noise-filtered cycles (period 2, 3, 5 and 6) merge reasonably smoothly with the underlying time series.

In this case, and for the other systems tested later, we have found a fixed frequency cut-off is applicable for cycles of all periods, however individual cut-offs can be varied by period if required.

The plots in Figure 5.8a–5.8d show the cycles of several periods after noise-reduction using the Fourier filter. One can see a clear improvement, especially given the high levels of added white noise.

5.4.3 Filtering Partial Cycles to Reduce Measurement Noise

We now utilise the valuable information contained in partial cycles which in this case are required to be of length greater than 50% of the period. The detection criteria demanded by the recurrence matrix method means that each detected partial cycle sequence of period T has been repeated T points later in the time series. There are a multitude of these partial cycles in any chaotic dynamical system of higher instability, where most orbits, more often than not, do not complete two full trajectories in the vicinity of a UPO. In this case there are 104 partial cycles detected and they contain valuable information. Directly applying Fourier techniques to these incomplete cycles will result in artefacts appearing due to the Gibbs phenomenon induced by discontinuities. These will be problematic is providing a filtered partial cycle.

To make these partial cycles more amenable to Fourier methods, the simplest way is to "mirror" the partial signal x of length L. If it is of length L, create a temporary new closed cycle x_{mirror} as follows:

$$x_{mirror}(i) = \begin{cases} x(i) \text{ for } i = 1, 2, .., L, \\ x(2L - i) \text{ for } i = L + 1, ..., 2L. \end{cases} \tag{5.8}$$

The new 'hybrid' partial cycle is of length $2L$ and consists of the concatenation of the partial signal x and its "flipped" mirror image in the line $x = L$. The temporary new cycle is not supposed to represent the true cycle as we are missing information on the completion of the cycle. It is merely a useful construct for the purposes of noise filtering. The new hybrid partial cycle now has starting and end points matching, it is symmetrical and the FFT amplitudes are doubled at each frequency. We are assuming that the partial cycle is sufficiently long enough to retain informational integrity with respect to the true cycle. After using the same cut-off employed for complete cycles, we take the first half of noise-filtered x_{mirror} as our signal. We found this technique to be very effective for partial cycles. We illustrate an example of a hybrid partial cycle of period $7T$ below in Figure 5.9.

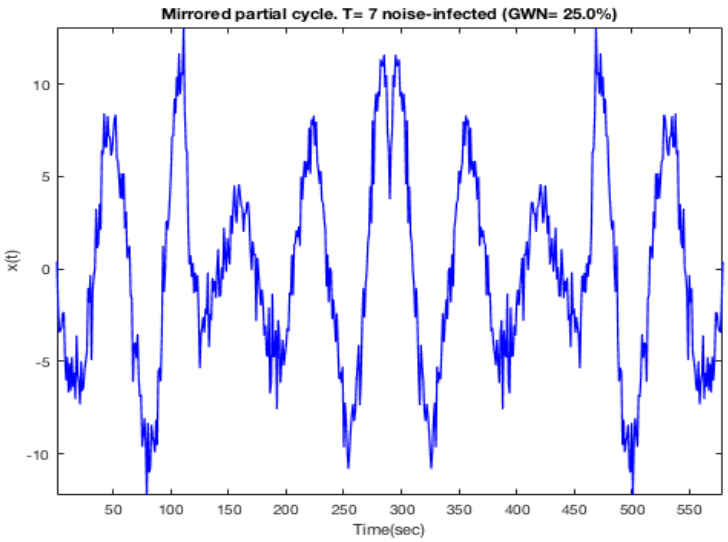


Figure 5.9: Hybrid partial cycle of period $7T$ constructed by "mirroring" the original partial cycle. This is an artificial construct in order to apply Fourier filtering to the partial cycle without start-end point artefacts becoming a problem.

In Figure 5.10 we show a period $7T$ partial shadow-UPO before and after applying a Fourier filter. We have colour coded the partial cycle and see that it joins smoothly either side in the time series. The same period $7T$ cycle is shown in three-dimensional phase space in Figure 5.11. The fully noise-reduced Rossler attractor is shown in Figure 5.12 and one can see visually that it is a considerable improvement of the noisy attractor. In Figure 5.13 we show the noise-filtered partial cycles of periods $1T$, $2T$, $3T$, $4T$, $6T$, $8T$, $9T$ and $10T$.

These detected and noise-reduced complete and partial shadow-UPOs will be placed back into the same location in the approximating time series. The location of each cycle (the start and end point) is stored in a matrix when the cycles are detected using the recurrence matrix method. This location indexing makes it easy to extract the cycles for noise-reduction and to return them subsequently. By returning filtered cycles to their original location in the time series, we reduce any "joining" discontinuities significantly. This is illustrated by the smooth joins in Figure 5.10 and Figure 5.13.

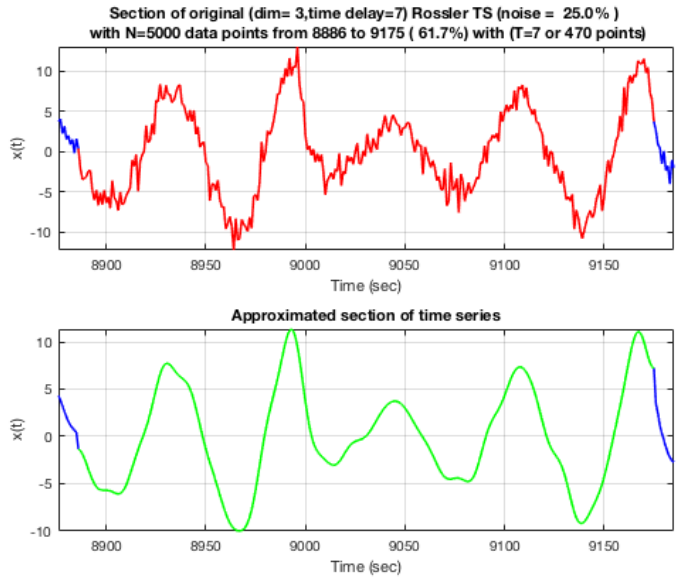


Figure 5.10: The joins between the underlying time series are relatively smooth. Top: Pre-filtered with a moving average. Bottom: Noise-filtered partial cycles (Bottom)

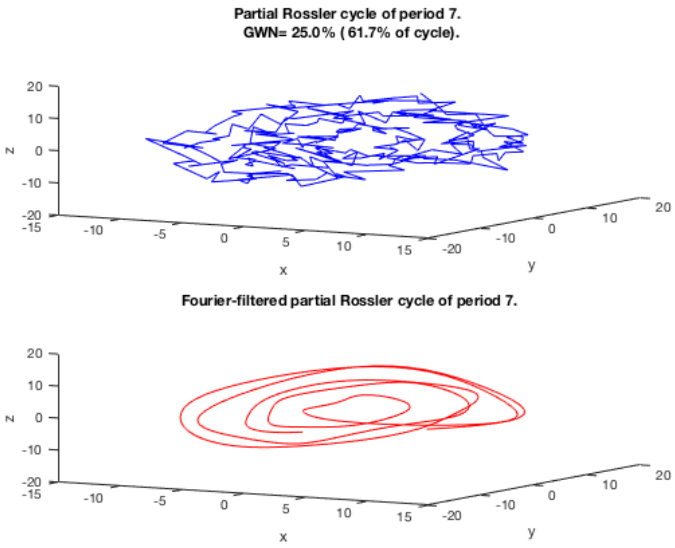


Figure 5.11: Partial cycle of period $7T$, shown in phase space before and after noise-reduction.

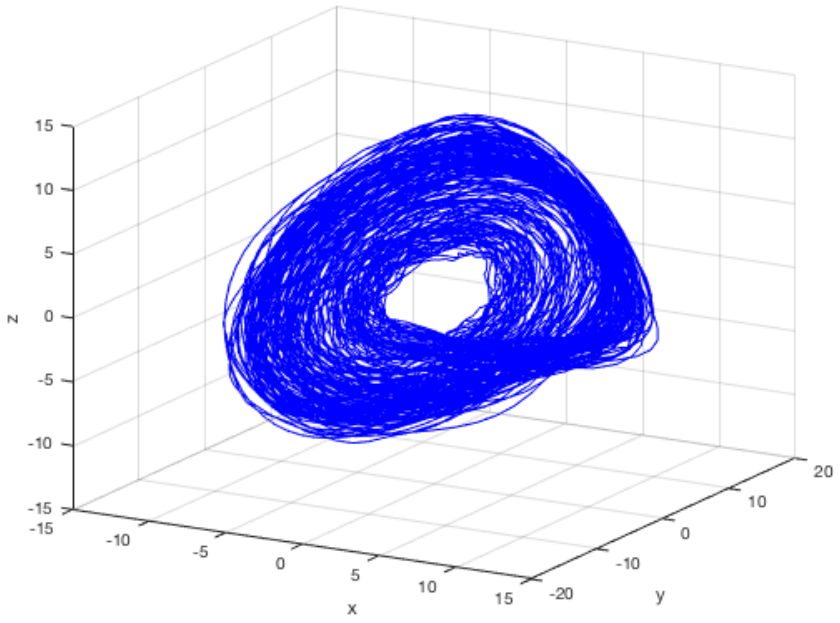


Figure 5.12: The original noise-infected attractor and approximation using noise filtered complete and partial detected shadow-UPOs. The coverage rate is 97.93%.

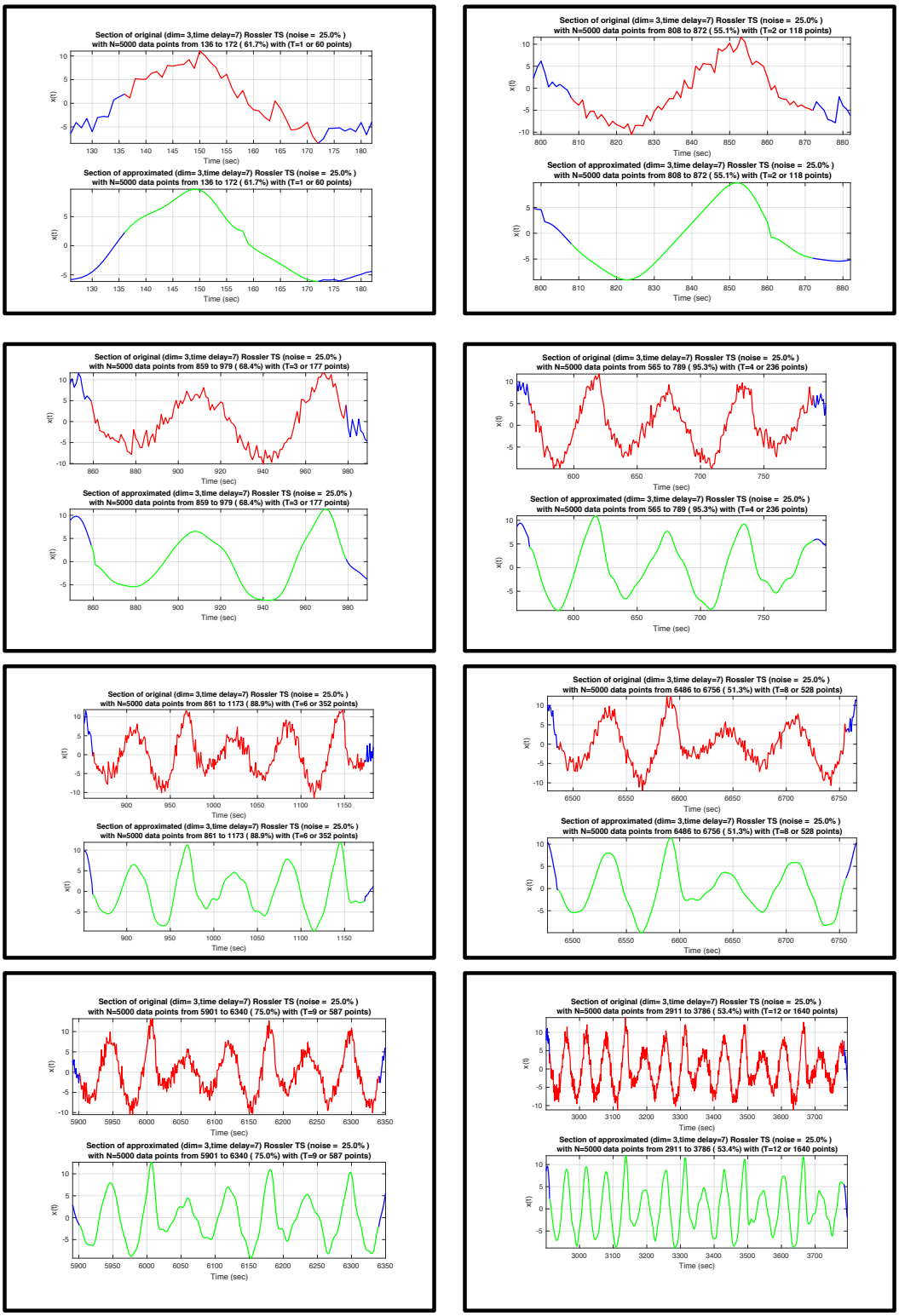


Figure 5.13: Noise-filtered partial cycles of period *1T*, *2T*, *3T*, *4T*, *6T*, *8T*, *9T* and *10T*.

5.4.4 Goodness of Fit of Approximant

In this section, we replace the noise-infected sequences of the moving average pre-filtered time series with the corresponding noise-reduced sequences and test the goodness of fit using the measures defined previously. We compare the goodness of fit of the final approximant to the original (noise-free) time series. We also compare the noise-infected time series and the moving average pre-filtered time series to the noise-free time series.

These comparisons are summarised in Table 5.2 below. A brief explanation of each comparison is as follows. The noise-infected time-series is compared to the noise-free (clean) time series. This is our baseline position against which we may compare an approximant. The moving average pre-filtered time-series is compared to the noise-free (clean) time series. This is our "worst case" noise filter as the moving average (low-pass) filter is the worst filter for signals with a broadband profile in the frequency domain as there is no means to distinguish one band of frequencies from another. The SUNR time series approximation is compared to the noise-free time series. The goodness of fit metrics quantify the effectiveness of the approximation. For our noise-reduction model to be successful, we require a significant improvement in the metrics when compared with the base scenario (noise-infected compared with noise-free). The least we will expect from our noise-reduction method is to produce results equivalent to the moving average comparison. For illustrative purposes we are using a simple low-pass filter and results are expected to be similar to those resulting from the moving average filter and indeed this is the case. The advantage of the SUNR method is that we can specifically tailor the noise-reduction technique for each cycle, based on the type of noise, to maximise the goodness of fit.

Comparisons are made using delay vectors constructed from the various series (with embedding dimension 3 and time delay 7).

The data in Table 5.1 shows that in this case the approximant is superior to the noise-infected time series and delay vector as expected on all metrics.

Delay Vectors	Noise vs. Clean	MA vs. Clean	Approx. vs. Clean
MAE	3.06	1.31	1.27
RMSE	4.90	1.01	0.89
MME	1.08	0.43	0.45
R^2	93.85%	98.73%	98.87%

Table 5.2: Goodness of fit of the SUNR approximation to the Rossler time series with 25% added GWN.

5.4.5 Variation of Approximation with Noise Type

Most research in the area of noise reduction is restricted to additive Gaussian white noise. Also the few research papers highlighting the apparent robustness of recurrence histograms to noise consider only Gaussian white noise and, to the best of our knowledge, this conclusion has not been confirmed for other types of noise. We now explore the consequences for cycle detection, and for the approximation more broadly, as we vary the type and level of the added noise.

Firstly we compare results for each type of measurement noise with a SNR fixed at 4.0 dB (25% added noise) and summarise these in Table 5.3.

Noise Type	**Number of Full Cycles**	**Number of Partial Cycles**	**Cover Ratio Full Cycles**	**Cover Ratio All Cycles**
High Frequency	62	124	64%	98%
Gaussian white	66	106	65%	96%
Uniform white	104	120	92%	99%
1/f	69	153	41%	80%

Table 5.3: Detection and coverage rates for Rossler system, by type of additive noise.

We observe that cycles are more difficult to detect when noise is confined to a section of the frequency domain (high frequency and coloured noise) rather than broadly across all modes. Typically the required critical radius is higher in these cases corresponding to 5%–10% of the (wider) maximum attractor extent. We observe that the robustness of the recurrence plot detection method remains intact in the presence of higher levels of added noise and when the noise type is varied.

Noise Type	MAE	RMSE	MME	**Phase Lag**	$\mathbf{R^2}$
Noise vs. Clean (HFN)	7.80	31.70	2.71	0	62.4%
Approx. (HFN)	0.51	0.28	0.19	0	99.6%
Noise vs Clean (GWN)	3.06	4.90	1.08	0	93.8%
Approx. (GWN: cut-off=1.0)	1.44	1.09	0.54	0	98.6%
Approx. (GWN: cut-off=2.5)	2.18	2.51	0.80	0	96.9%
Noise vs Clean (UWN)	0.96	0.41	0.32	0	99.5%
Approx. (UWN: cut-off=1.0)	0.46	0.11	0.17	0	99.9%
Approx. (UWN: cut-off=2.5)	0.65	0.21	0.24	0	99.7%
Noise vs. Clean (1/f)	7.77	31.75	1.91	0	60.2%
1/f (cut-off)	6.88	24.90	1.43	0	68.9%
1/f (targeted filter)	1.74	6.71	0.17	0	91.3%

Table 5.4: Comparison of goodness of fit metrics for SUNR method approximant. We compare each with the base case of the noise-free delay vector versus the 25% noise case (listed as Noise vs. Clean). All approximations are in phase with the noise-free time series.

As expected the method performs best for high frequency added noise where the cut-off is definitive and the noise is neatly removed. For this case the improvement over the noise-infected time series and fit is excellent.

In Table 5.4 we summarise the goodness of fit results for the SUNR approximations for various types of measurement noise. We provide two cut-off scenarios for both the Gaussian white noise and Uniform white noise, representing aggressive and relaxed cut-off scenarios. The case where the cut-off value is 1 corresponds to the cut-off used for the high-frequency noise and retains just the key peaks. The case where the cut-off value is 2.5 is less aggressive and retains many higher frequency amplitudes. These give us a range of results. These results for Gaussian white noise are vastly superior to the base case but slightly inferior to the high frequency noise case as noise is spread across all frequencies and some noise is retained at the lower frequencies and not cut off. The results for uniform white noise are noticeably better than for Gaussian white noise. We may conclude that a simple targeted cut-off approach will provide excellent results for high frequency noise, Gaussian white noise and uniform white noise without requiring a deeper knowledge of signal analysis.

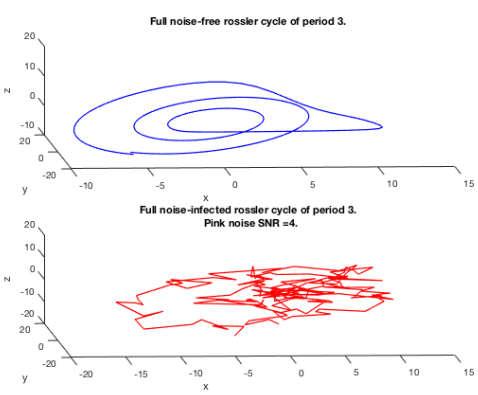
In the next section we consider pink noise in more depth.

5.4.6 Targeted Noise-Filtering Example: Pink Noise

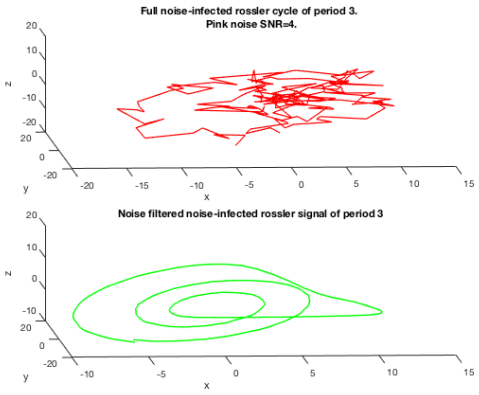
As expected, the results for pink noise using a straight cut-off are a little better than the base case but not by much. This is because much of the noise is retained in the lower frequencies and low-pass truncation is simply inadequate. We also had to apply a longer moving average pre-filter (window of 21 points) and a higher critical radius of 2.8 ($> 5\%$ of attractor extent) to detect a useful number of the cycles. This is illustrated in Figure 5.14(a) where we can see the shadow-UPO, after the low-pass filter, is still very noisy.

If however we apply a more targeted signal processing technique, superior results can be achieved. We can firstly approximate the SNR using a noise-tracking algorithm. There are a number of algorithms available to estimate the amplitude of the measurement noise without apriori knowledge of the underlying signal. An example is the method for non-stationary nonlinear data by Hu et al. [68]. There are also simpler updating methods involving sliding windows. The noise can be modelled by combining the estimated noise amplitude with the frequency domain geometry ($1/f$). With this information in hand we reduced each frequency bin using a calculated gain constant to remove the pink noise. In Figure 5.14(b) we see the improvement of using a targeted signal processing method. The period $3T$ cycle is now significantly noise-reduced. The goodness of fit results of the approximated time series are shown in Table 5.4 and are a considerable improvement over the low-pass filter.

The key point illustrated here is the SUNR method converts the intractable problem of filtering a high-noise chaotic time series into that of individually noise-filtering a series of shadow-UPOs where the results will be as good as the signal processing technique adopted. Fortunately there is a rich canon of signal processing techniques, tailored to each type of noise, that can be exploited to optimise results.



(a) Pink noise with FFT filter.



(b) Pink noise with targeted filter.

Figure 5.14: (a) Cycle of period 3, with added pink noise having a SNR=4.0 dB (top) after being filtered with a simple cut-off. Experimenting with the cut-off does not improve the result much. This type of noise requires a different approach, (b) The same cycle of period 3, with noise-removed using a more sophisticated technique targeted at pink noise; a far better result.

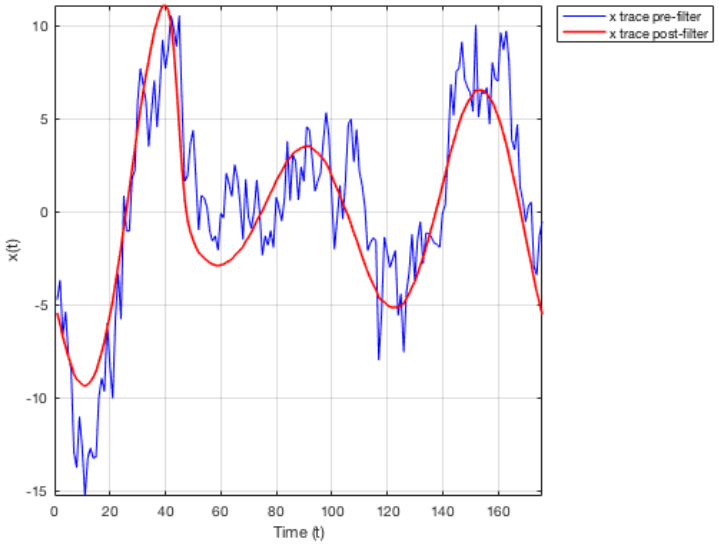


Figure 5.15: Trace of x - coordinate versus time of period $3T$ shadow-UPO (see Figure 5.14). We show the pink noise-infected (SNR=4.0 dB) shadow-UPO in blue. The noise-filtered version in red was obtained by modeling and subtracting the noise.

5.4.7 Variation of Approximation with Noise Level

Now let us explore how the approximation varies as added noise increases. We use Gaussian white noise in this case to illustrate the results. We note that coverage rates reduce with increasing noise, as one would expect. We used a critical radius of 1.8 for noise levels up to and including 25%. For the 50% and 100% added noise scenarios we used critical radii of 5% of the noise-infected attractor maximum extent (distance between two points) values of 2.4 and 2.8 respectively. We used a moving average window of 11 points (5 points either side of central value) for all noise up to 50%. For the 100% we needed to double the moving average window width to 21 points to obtain satisfactory results.

Table 5.5 summarises the goodness of fit results and we observe that the fitted approximations are always superior to that of the base case of noise-infected time series versus noise-free time series in all cases. In particular the important average L_2 distance for the approximant is around only 20% of that for the base case. The R^2 statistical measure shows considerable superiority for higher noise levels (50% and 100%).

We also use the approximation of the noise-free time series to calculate the maximal or positive Lyapunov exponent using the Wolf numerical algorithm [145]. We have established a reference set of maximal Lyapunov exponents for each of the chaotic systems studied in this thesis for comparison purposes. These Lyapunov exponents were calculated directly using noise-free data and the Matlab LET toolbox, as well as checking published results. They are as follows: Rossler ($\lambda_{\max} = 0.07$ s^{-1}), Lorenz ($\lambda_{\max} = 0.91$ s^{-1}), Chua ($\lambda_{\max} = 0.28$ s^{-1}), Rabinovich-Fabrikant ($\lambda_{\max} = 0.07$ s^{-1}) and Lu-Chen ($\lambda_{\max} = 1.75$ s^{-1}).

The Wolf algorithm typically fails for signals with noise in excess of 5%–10%. The results show that the SUNR filtered signal produces reasonable results up to 25% noise. Thereafter, residual noise is presumably affecting the results somewhat. Note the trend is clear with increasing noise. However, even at 100% noise we are still seeing a positive maximal Lyapunov exponent, signalling the system

is chaotic. This is a considerable improvement in determining the maximum Lyapunov exponent with higher noise levels.

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	99.33%	Base	0.61	0.20	0.22	99.75%	0.079
		Fitted	0.34	0.07	0.13	99.91%	0.066
10%	99.33%	Base	1.22	0.78	0.43	99.02%	
		Fitted	0.60	0.19	0.22	99.75%	0.064
15%	99.34%	Base	1.85	1.79	0.67	97.76%	
		Fitted	0.85	0.38	0.32	99.52%	0.070
20%	99.35%	Base	2.47	3.18	0.89	96.02%	
		Fitted	1.12	0.66	0.43	99.17%	0.071
25%	98.36%	Base	3.07	4.88	1.08	93.88%	
		Fitted	1.39	1.03	0.51	98.70%	0.084
50%	79.17%	Base	6.11	19.61	2.17	75.38%	
		Fitted	2.70	3.85	0.98	95.17%	0.111
100%	85.53	Base	12.27	78.04	4.32	2.10%	
		Fitted	5.27	14.82	1.90	81.50%	0.155

Table 5.5: Variation in goodness of fit for Rossler system (time step=0.1), with added Gaussian white noise level. The maximal (positive) Lyapunov exponent is estimated using Wolf's algorithm. The value derived for the noise-free time series is LE=0.079 (base e) (1/sec).

5.5 Filtering Dynamical Noise from Time Series

5.5.1 Locating and Extracting Cycles

In section 4.2 we discussed the effect of dynamical noise on a signal. Measurement noise is a random variation around a true underlying signal. Regardless of the amount of added noise, the signal is still on the "inside". Dynamical noise is very different in that it actually deforms the signal itself into a new signal. Thus one cannot strictly speak of "separating" the signal from dynamical noise. They are not combined and separable. At best we can try to undo the effect of the dynamical noise somewhat.

We showed earlier how the influence of dynamical noise perturbs and deforms the underlying signal, changing its shape. Provided the dynamic noise does not deform the signal too much, we can still locate the cycles using recurrence techniques. In Table 5.6 we summarise the numbers of complete and partial cycles detected from the dynamical noise-infected Rossler time series.

Period Number	**Period**	**Number of Full Cycles**	**Number of Partial Cycles**
1	60	14	12
2	118	15	14
3	177	23	15
4	236	0	8
5	294	0	6
6	352	4	8
7	470	4	0
8	528	0	10
9	587	0	2
10	645	0	0
Total		60	75

Table 5.6: Detected Rossler shadow-UPOs from dynamical noise-infected time series (time step = 0.1). Partial cycles are sequences of length $\geq 50\%$ of a period.

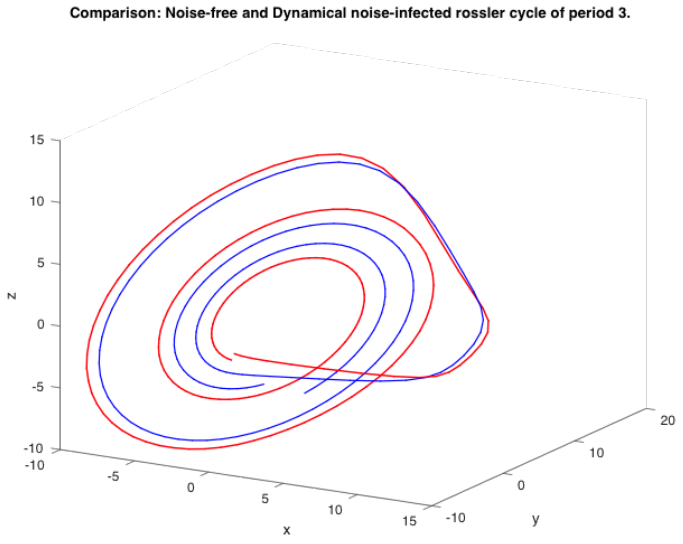


Figure 5.16: Period 3T shadow-UPO without noise shown in red. The addition of dynamical noise deforms the entire orbit (blue) and may shift the phase.

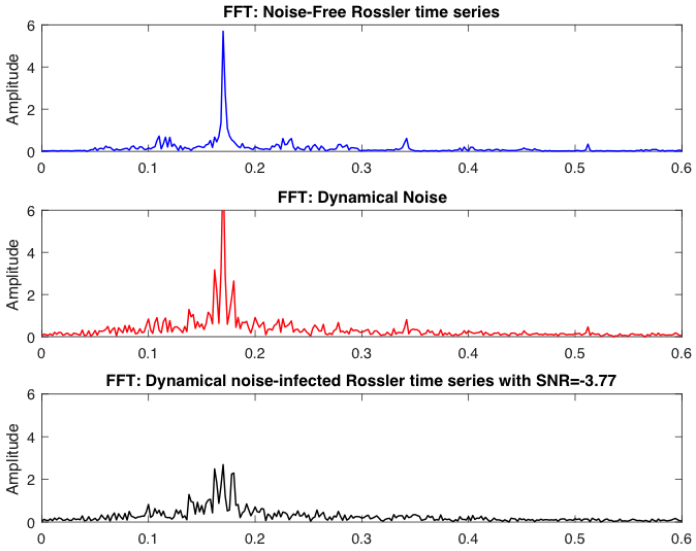


Figure 5.17: FFT of noise-free signal (blue), the dynamical noise (red) and the noise-infected signal (black). There is no easy way to recover the unperturbed signal.

Figure 5.16 shows a noise-free period $3T$ shadow-UPO (red) and the effect of adding dynamical noise (blue). The dynamical noise can be seen to deform the cycle and cause a phase shift. From these illustrations we can conclude that Fourier techniques that rely on frequency separation between signal and noise cannot succeed here. Figure 5.17 shows that the dynamical noise inhabits the same frequencies as the signal.

The removal of dynamical noise from time series is a complex area of research, without a lot of published material and to the best of our knowledge, no off-the-shelf software solutions or commonly used approach.

5.5.2 SUNR Approach with Multi-sampling

We attempted several simple approaches to reducing dynamical noise from detected shadow-UPOs. Application of a band-pass filter around the central peak still results in a very poor fit, significantly worse that a moving average (which also yields a poor result). Transforming the signal using logarithms, with the sign of each value preserved also does not improve the results (this method may improve results for multiplicative measurement noise).

The clear finding here is that any frequency domain separation method will not suffice as the dynamical noise is intermeshed with the signal at the same frequencies and a different technique is required. Cycles are irrevocably changed by dynamical noise and the best we can hope for is that the lower order cycles of each period, whilst being deformed, remain within the critical radius (as required by the Shadowing Lemma). Interestingly, we have found that the perturbed cycles are still detectable for a relatively low signal-to-noise ratios and this suggests an approach to reducing the effect of the dynamical noise. As the dynamical noise increases the reduction in number of cycles counted in the lower-period bins in a measure of when the model comes under stress. As long as we detect the same number of cycles in each bin as for low-levels of dynamical noise, the deformations caused by increasing dynamical noise further are insufficient to break the model.

We demonstrated earlier in Chapter 4 how to obtain a complete set of additive noise-reduced lower order UPOs for the model Rossler system using multiple time series samples. Shadow-UPOs were detected, binned and averaged, with the objective of measurement noise averaging out in highly populated bins. We can adopt a similar approach for a dynamical noise-infected system provided multiple samples could be collected. The approximation is usually dominated by lower order cycles (periods *1T–3T*), and we typically capture a great number of these. The dynamical noise added will be typically Gaussian or uniform, both with a mean of zero. The central idea is that the cycle deformations for lower order cycles will even out, with "overs" cancelling "unders"; at least enough to provide a better estimate of the noise-free UPO. For some iterations, if the dynamical noise is sufficiently low the subsequent iteration will remain on the same path. In other cases the orbit will be nudged onto another . The detected deformed cycles are essentially a blend of nearby near-UPOs with the orbit jumping around from one to the other occasionally. Averaging a pool of these will arguably provide an estimate of the underlying UPO, albeit not perfect.

Period Number	**Period**	**Number of Cycles Averaged**
1	60	194
2	118	101
3	177	219
4	236	6
5	294	11
6	352	50
7	470	6
8	528	12
9	587	2
10	645	4

Table 5.7: Shadow-UPOs detected from 20 samples of a dynamical noise-infected time series, each of length 10,000 points.

A summary of numbers of detected shadow-UPOs is presented in Table 5.9. The corresponding recurrence histogram for 10 samples of 10,000 points is shown in Figure 5.18, where we see a large number of period *T–3T*. We estimated the

noise-free UPOs by averaging the binned cycles of each period as before. Figure 5.19 shows the 219 period *3T*, dynamical noise-infected shadow-UPOs that were captured from 10 Rossler time series samples. The average is shown in red. Results were similar for the lower order cycles with period *T–6T*.

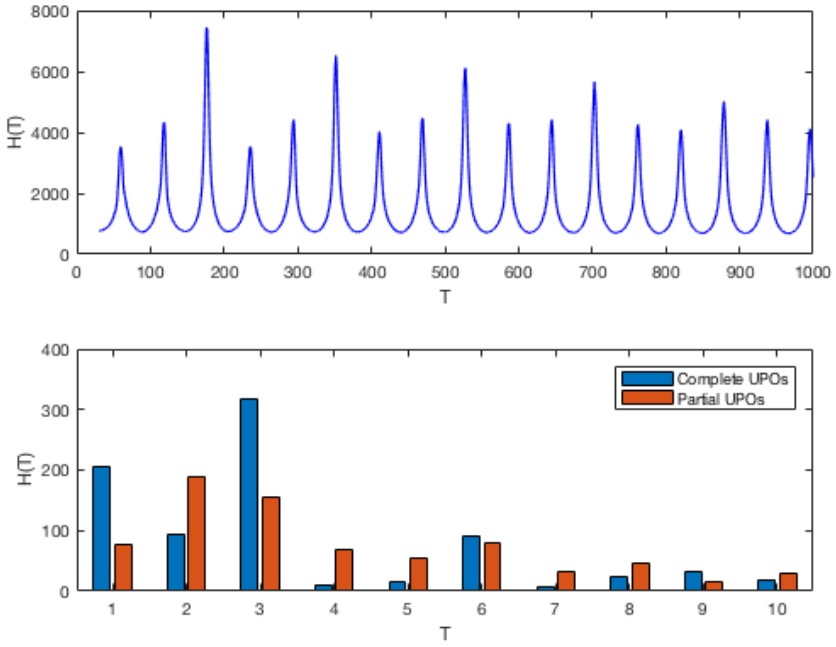


Figure 5.18: Recurrence histogram of aggregated detected Rossler shadow-UPOs from 10 time series of 10,000 points each. We clearly have a great number of periods *T–3T* cycles and partial cycles. The average SNR for these 10 samples is 4.64 dB.

We subsequently compared the UPO estimates derived from the dynamical noise-infected time series with those derived from 10 samples of the noise free time series. Results are shown below in Figure 5.20.

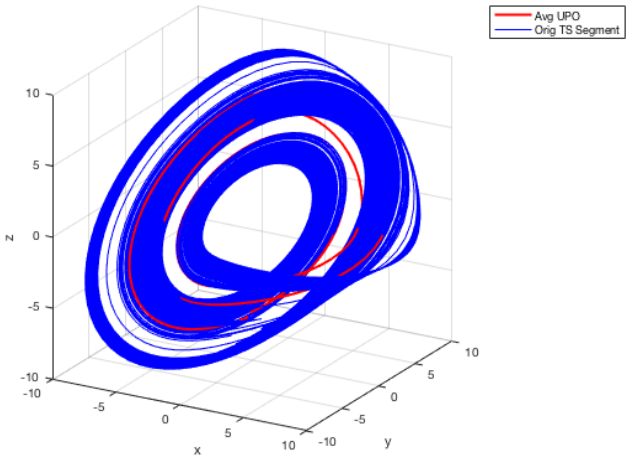
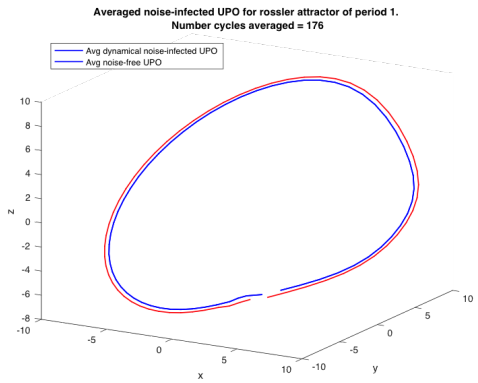


Figure 5.19: We averaged the 219 period *3T* dynamical noise-infected Rossler shadow-UPOs that were captured from 10 time series samples. Results were similar for the lower order cycles with period *T*—*6T*.

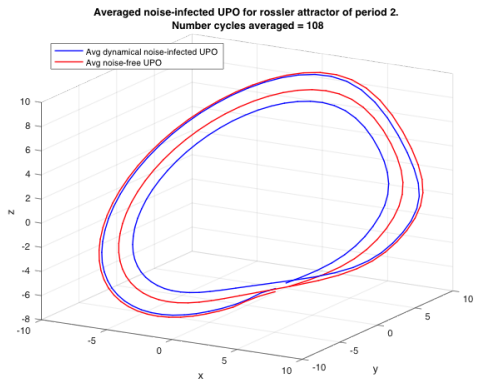
We replaced only complete cycles in a dynamical noise-infected model Rossler time series of 10,000 points with estimated noise-free cycles of each period less than *10T* derived from multi-sample averaging. We did not utilise partial cycles and deleted cycles requiring permutation of more than 3 time steps to synchronise with the most common cycle orientation. This reduced complexity and still retained large numbers of cycles for averaging.

Comparison	**SNR (dB)**	MAE	RMSE	MME	$\mathbf{R^2}$
5,000 points					
Noise vs. Clean		7.07	33.55	1.81	58.2%
Approx vs. Clean	3.78	6.69	23.11	1.69	71.2%
10,000 points					
Noise vs. Clean		9.53	46.91	2.44	41.2%
Approx vs. Clean	5.14	6.33	23.55	1.69	70.5%

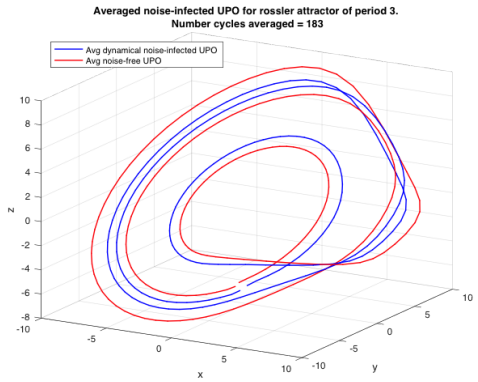
Table 5.8: Goodness of fit of the approximant derived by averaging equivalent cycles from multiple samples. We used 15 samples of 5,000 and 10,000 points respectively, with dynamical noise added to determine average cycles.



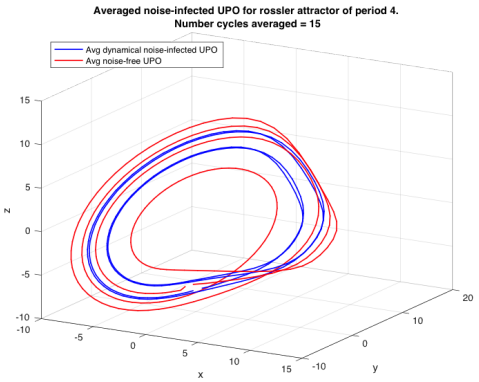
(a) Period 1 cycle.



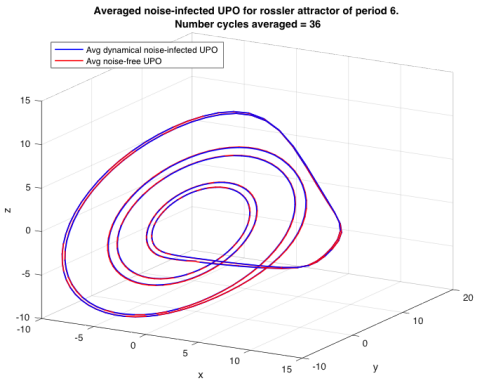
(b) Period 2 cycle.



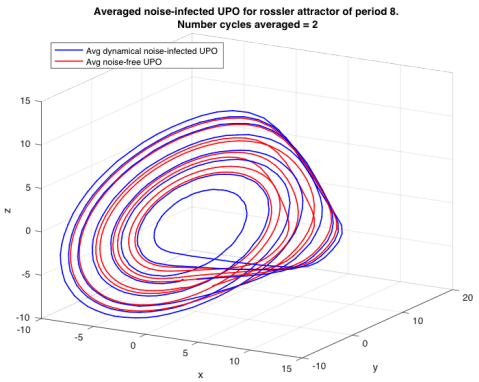
(c) Period 3 cycle.



(d) Period 4 cycle.



(e) Period 6 cycle.



(f) Period 8 cycle.

Figure 5.20: Comparison between noise-free UPOs and estimates using multiple-samples of a dynamical noise infected time series. One can see that these are reasonable approximations.

We tested this approach on many individual time series, each with a SNR around this level and the example shown in the table is representative of these cases. The average SNRs of the samples used to estimate UPOs were 3.78 dB (around 25% noise) and 5.14 dB (around 20% noise) respectively. The approximant shows improvement on all the metrics when compared to the original noise-infected case but the results at this point are not highly compelling. We made the following observations:

1. This approach will likely provide a reasonable approximation for low levels of dynamical noise. However, any perturbation added in this manner to a chaotic system quickly inflates and the resultant SNR is only broadly controllable.

2. Changing the length of the time series used in multi-sampling appears to make little difference. The goodness of fit results obtained time series samples of 5,000 points are similar to those for sample time series of 10,000 points.

3. Limiting the approximation to using only lower order averaged cycles, for example periods T–$3T$ only also appears to make little difference. One might expect that shorter cycles are less "deformed", but this does not appear to be the case. These cycles recur frequently along the evolution of the orbit and are deformed.

We conclude that this approach does show promise in a challenging situation where few, if any, techniques work. Further research is required in this area to improve the precision of the approximation.

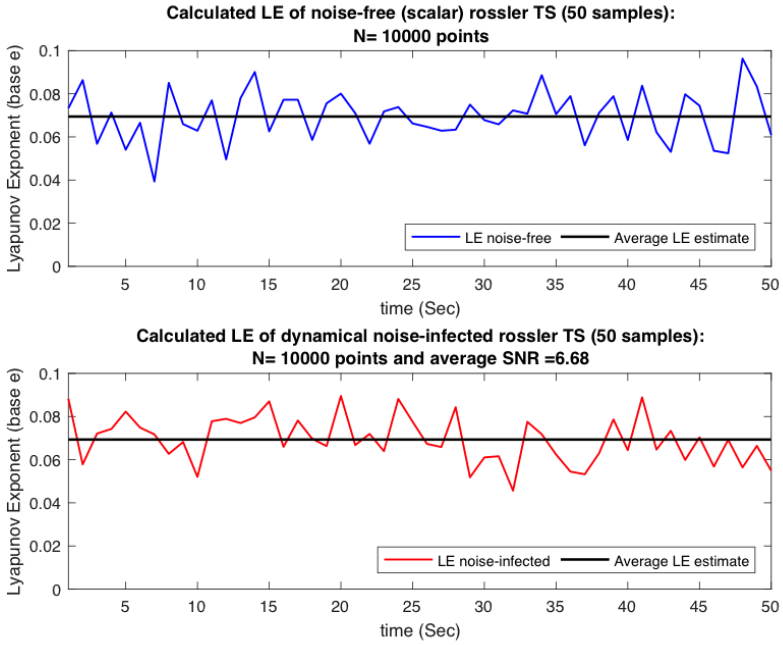
5.5.3 Simple Approach to Dynamical Noise

The end goal of approximating the noise-free time series is to have "usable" data that may be deployed within conventional numerical algorithms (that generally require noise-free data) to calculate invariants, in particular the maximal Lyapunov exponent. One would expect that any noise would result in diffusion of the attractor and an over-estimate of positive Lyapunov exponents. However, a key observation so far has been that dynamical noise-infected data, although deformed,

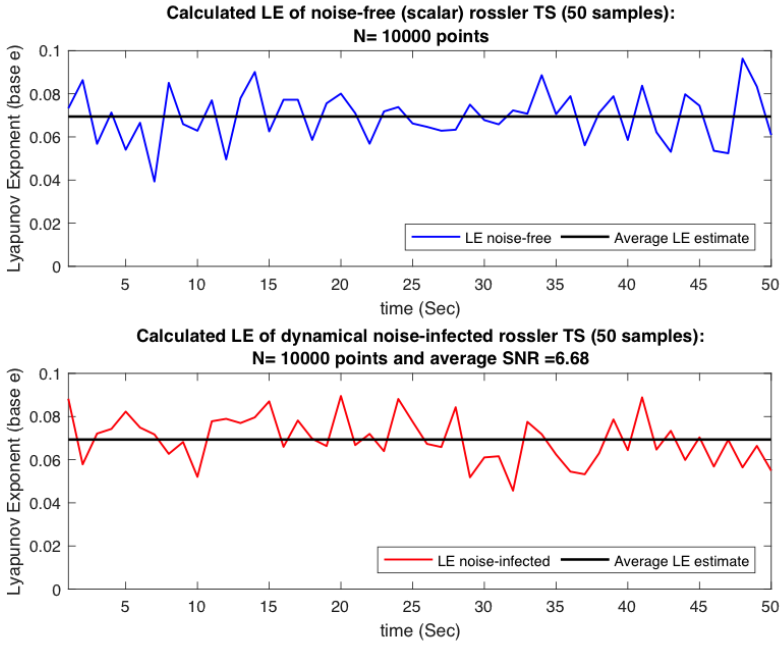
remains smooth and appears to show some resilience in retaining the integrity of structure. This is evidenced by recurrence histograms retaining integrity with clear peaks and the average of binned deformed cycles of a given period closely resembling the cycles in the noise-free data.

With this in mind, we tested the strength of the Wolf numerical algorithm for calculating maximal Lyapunov exponents when the time series data contains dynamical noise. Note that we found this method, which tracks the evolving orbit in small steps, failed with relatively low measurement noise. Such a trajectory-based method requires smooth steps forward in time and a relatively low point density (replacement points outside the noise scale). Surprisingly, we found this method to be highly robust in the presence of dynamical noise. We constructed multiple dynamical noise-infected Rossler time-series, running each directly through the Wolf algorithm and estimating the maximal Lyapunov exponent. Subsequent time series were started from the mid-point of the previous time series and up to 50 time-series were tested. Similarly, we constructed 50 noise-free Rossler time series using the same methodology and compared results.

These results are illustrated in Figure 5.21. The averages determined from the noise-free data and dynamical noise-infected data are approximately equal. For the SNR=5.35 dB case the averages are 0.0703 s^{-1} and 0.0696 s^{-1} respectively and for the SNR=6.68 dB case the averages are 0.0694 s^{-1} and 0.0695 s^{-1} respectively. This exercise was repeated multiple times and in all cases the results very close to each other. This would seem a good result considering the level of noise involved.



(a) SNR=5.35 dB



(b) SNR=6.68 dB

Figure 5.21: (a) Maximal Lyapunov exponents calculated for 50 Rossler time series, both noise-free and with dynamical noise using the Wolf algorithm. Two scenarios are presented (a): SNR=5.35 dB and (b) SNR=6.68 dB.

This suggests that a pragmatic approach for a chaotic system infected with dynamical noise is to simply estimate the maximal Lyapunov exponent using multiple time-series sampling and the Wolf algorithm directly; averaging the results. We also tested this approach on the Lorenz system with results shown in Figure 5.22.

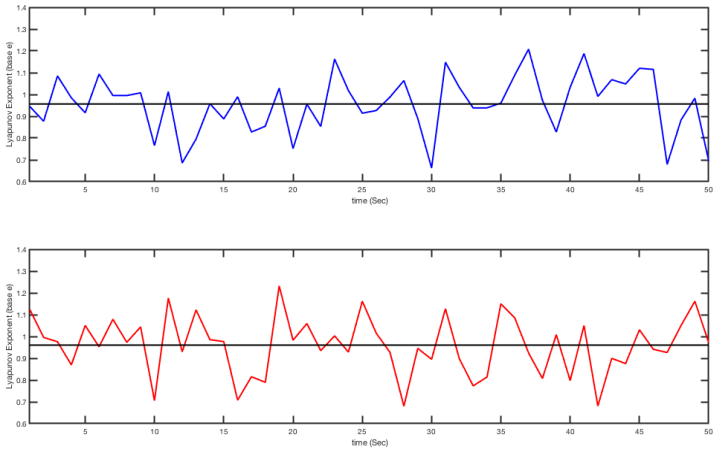


Figure 5.22: Maximal Lyapunov exponents calculated for 50 Lorenz time series, both noise-free and with dynamical noise using the Wolf algorithm. The average noise-free LE equals 0.97 s^{-1} (Top plot) and the average noise-infected LE equals 0.96 s^{-1} (Bottom plot). The SNR equals -1.4 dB indicating the signal is heavily infected by noise.

To test whether this approach was only applicable to systems with relatively low instability, we tested it on a highly chaotic system; the Lu-Chen system. The maximal Lyapunov exponent for this system is in the range 1.5–1.8 s^{-1} and is thus highly unstable. The results are shown above in Figure 5.23.

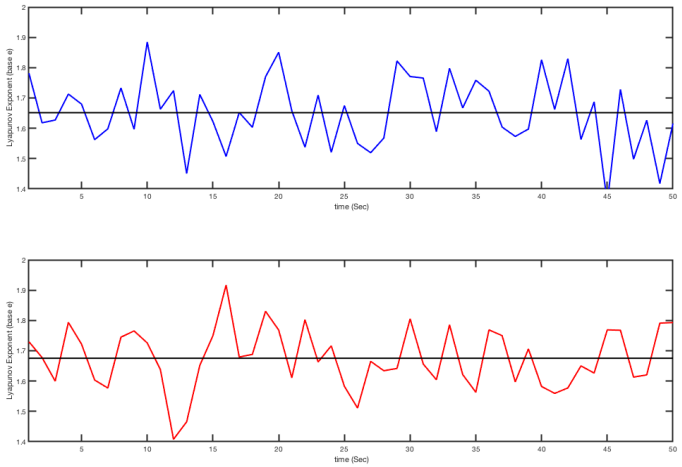


Figure 5.23: Maximal Lyapunov exponents calculated for 50 Lu-Chen time series, both noise-free and with dynamical noise using the Wolf algorithm. The average noise-free LE equals 1.65 s^{-1} (Top plot) and the average noise-infected LE equals 1.68 s^{-1} (Bottom plot). The SNR equals -2.9 dB indicating the signal is swamped by noise.

It is interesting to see that for all 3 systems, the average LE calculated using the dynamical noise-infected time series is indeed very close to that determined from the noise-free time series, albeit with more volatility between readings for the the more chaotic systems.

5.6 Results for Other Chaotic Systems

We have demonstrated in the case study above that the approximation to the noise-free Rossler time series using the SUNR method will show marked improvements in goodness of fit metrics, when compared to the time series infected with measurement noise. Also the resultant noise-reduced time series will generally have sufficiently low residual noise that the conventional Wolf algorithm can be deployed to determine the maximal Lyapunov exponent. We now tackle several other lower dimensional chaotic systems with differing topologies as follows.

5.6.1 The Chua System

The Chua system equations [100] are as follows:

$$\frac{dx}{dt} = \alpha(y - x - h(x)), \tag{5.9a}$$

$$\frac{dy}{dt} = x - y + z, \tag{5.9b}$$

$$\frac{dz}{dt} = -\beta y, \tag{5.9c}$$

where $h(x) = bx + \frac{1}{2}(a-b)(|x+1| - |x-1|)$. The parameters chosen are $(\alpha, \beta, a, b) = (0.9, 100/7, -8/7, -5/7)$, time step $h = 0.05$ and initial point is $(x_0, y_0, z_0) = (0.1, 0.1, 0.1)$. The numerically integrated Chua attractor is shown in three dimensional phase space in Figure 5.24.

The Chua circuit is the simplest electronic circuit exhibiting chaos, and its double-scroll attractor has been the subject of much study. The Chua equations, like the Rossler equations, only contain only one nonlinearity, in this case added through the piecewise-linear function $h(x)$ defined by the parameters a and b. These define the slope of the inner and outer segments of $h(x)$. (By contrast, the Lorenz equations contain 3 nonlinear terms, each consisting of the product of two variables). The geometrical structure of the attractor is remarkable, containing an

infinite number of oppositely directed, concentric cycles. Locally, the geometry of each cross-section appears to be a fractal across all scales.

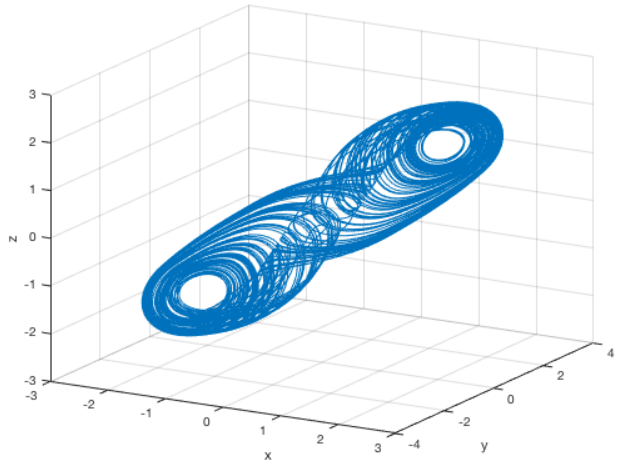


Figure 5.24: Attractor for the noise-free Chua system, obtained by numerical integration of Equations 5.17a–5.17c using 10,000 points and a time step of 0.05.

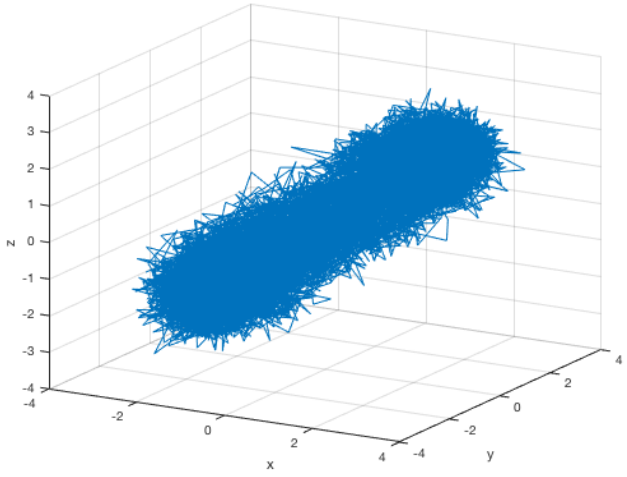


Figure 5.25: The Chua attractor with 25% added Gaussian white noise.

The recurrence histogram for the noise-infected Chua time series is shown in Figure 5.26. Peaks identifying the fundamental period and other harmonics are identifiable and we were able to extract many complete and partial shadow-UPOs. In Figure 5.27 we provide an example of a complete period $2T$ shadow-UPO, that was extracted and noise-filtered using a simple Fourier low pass filter. The same cycle is shown in phase space in Figure 5.28, where we can better see the improvement from filtering.

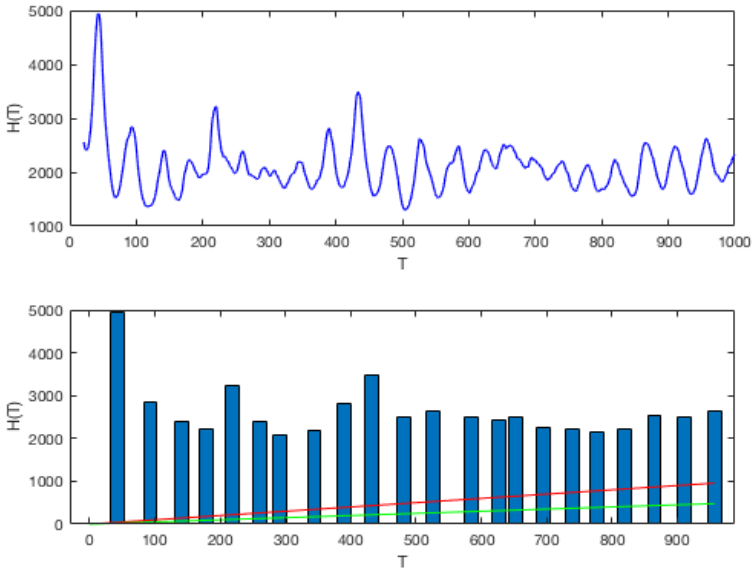


Figure 5.26: Recurrence histogram for Chua system with critical radius 0.5.

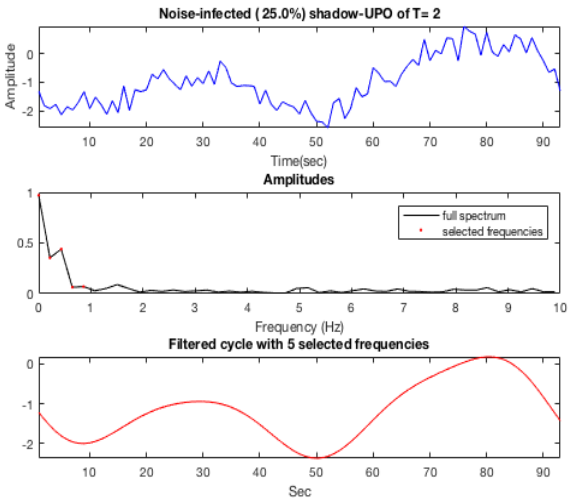


Figure 5.27: The top plot shows a Chua shadow-UPO of period $2T$, infected with 25% added Gaussian white noise. The FFT is shown in the middle plot, and the noise-reduced cycle is shown in the bottom plot.

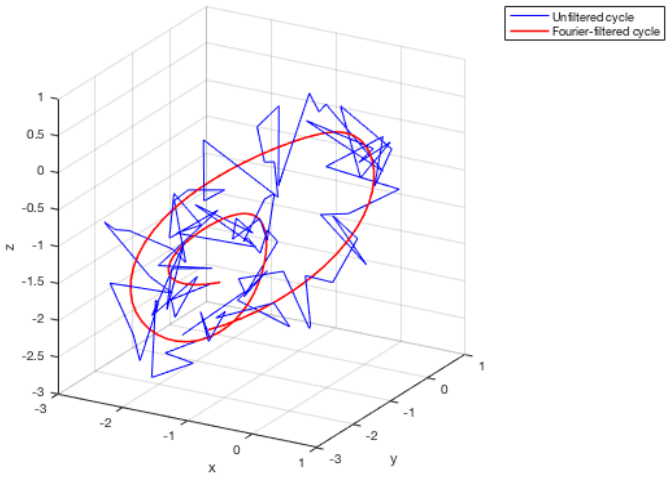


Figure 5.28: Chua shadow-UPO of period $2T$, both infected with 25% added Gaussian white noise (blue) and noise-filtered using the FFT (red) .

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	93.41%	Base	0.16	0.01	0.06	99.76%	0.28
		Fitted	0.08	0.00	0.02	99.91%	0.27
10%	93.03%	Base	0.33	0.06	0.12	99.00%	
		Fitted	0.13	0.01	0.04	99.83%	0.25
15%	92.09%	Base	0.49	0.12	0.17	97.80%	
		Fitted	0.18	0.02	0.05	99.67%	0.26
20%	92.09%	Base	0.65	0.22	0.23	96.06%	
		Fitted	0.23	0.03	0.06	99.48%	0.29
25%	85.67%	Base	0.83	0.36	0.30	93.58%	
		Fitted	0.28	0.04	0.08	99.22%	0.27
50%	53.90%	Base	1.65	1.44	0.49	74.32%	
		Fitted	0.56	0.17	0.16	97.00%	0.27
100%	28.23%	Base	3.28	5.64	1.18	52.14%	
		Fitted	1.03	0.56	0.29	90.23%	0.26

Table 5.9: Variation in goodness of fit of the SUNR approximant for Chua system with added Gaussian white noise. The maximal (positive) Lyapunov exponent is estimated using Wolf's algorithm. The value derived for the noise-free time series is LE=0.28 (base e)(1/sec). The time step is 0.05.

In Table 5.9 we summarise the goodness of fit results for the SUNR approximant for the Chua system for added GWN up to 100%. The distance metrics show excellent reductions throughout the entire noise range and in particular at the higher noise levels. At 50% added noise the distance metrics are similar to the values for an unfiltered time series with 15% added noise. At 100% the coverage rate is low (28.23%) and the result is dominated by the effects of moving average pre-filter. The R^2 values show a good fit, retaining values greater than 99% up to 25% noise and 97% at 50% noise. The estimates of the maximal Lyapunov exponent fluctuate a little but are close estimates of the true value ($\lambda_{\max} = 0.28$ s^{-1}).

5.6.2 The Lorenz System

The Lorenz system [95] is defined by the following nonlinear equations:

$$\frac{dx}{dt} = \sigma(y - x), \tag{5.10a}$$

$$\frac{dy}{dt} = rx - y - xz, \tag{5.10b}$$

$$\frac{dz}{dt} = xy - bz, \tag{5.10c}$$

where the parameters are $(\sigma, r, b) = (28, 8/3, 10)$, the time step $h = 0.01$ and initial point is $(x_0, y_0, z_0) = (1, 1, 1)$. The noise-free Lorenz attractor is shown below in Figure 5.29.

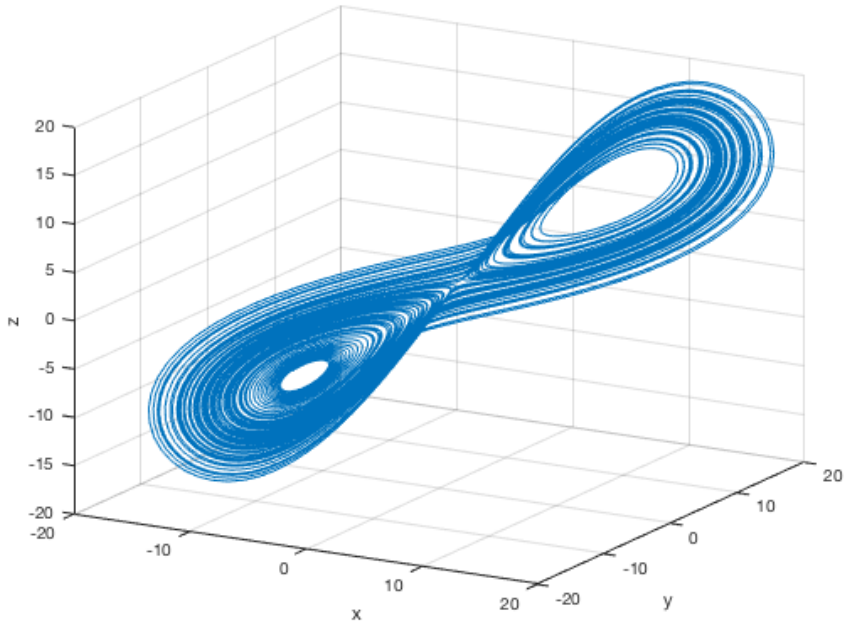


Figure 5.29: Lorenz chaotic attractor (noise-free) obtained by numerically integrating Equations 5.19a–5.19c using the 4th order Runge-Kutta method. There are 10,000 points with time step of 0.01.

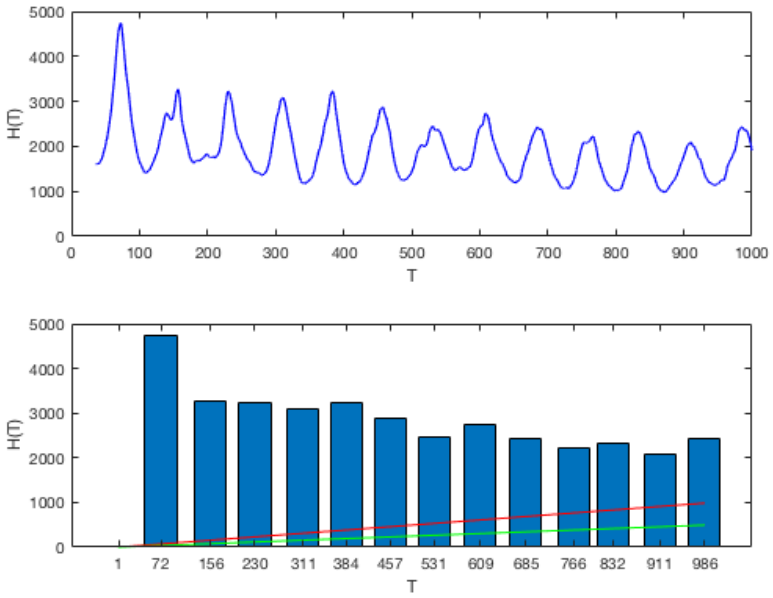


Figure 5.30: Histogram derived from recurrence matrix of Lorenz system.

The recurrence histogram for the noise-infected Lorenz time series is shown in Figure 5.30. Peaks identifying the fundamental period and other harmonics are clearly identifiable and we were able to extract many complete and partial lower-order shadow-UPOs. Most of the shadow-UPOs detected were of the low periods T–$3T$, but they provide sufficient coverage (>75% of time series). In Figure 5.31 we show examples of the lower order shadow-UPOs detected. In Figure 5.32 we provide an example of a period $3T$ shadow-UPO before and after noise-filtering.

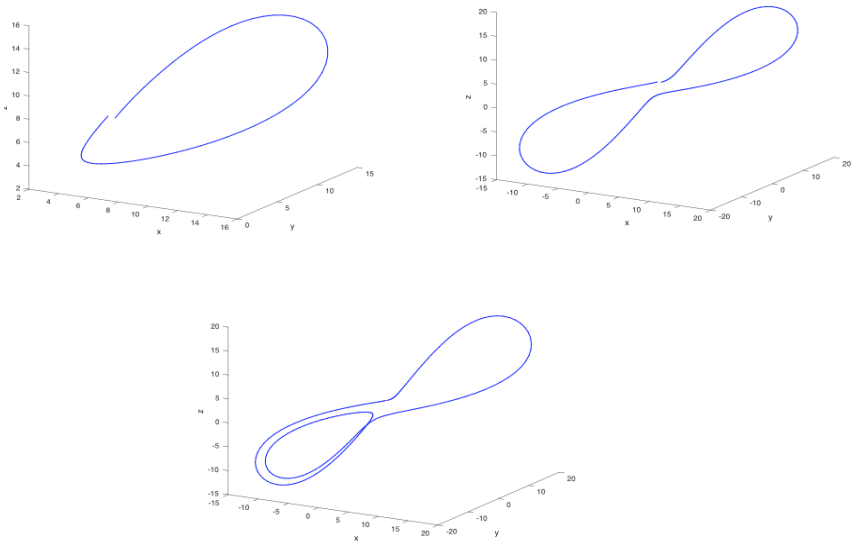


Figure 5.31: Examples of detected lower order shadow-UPOs.

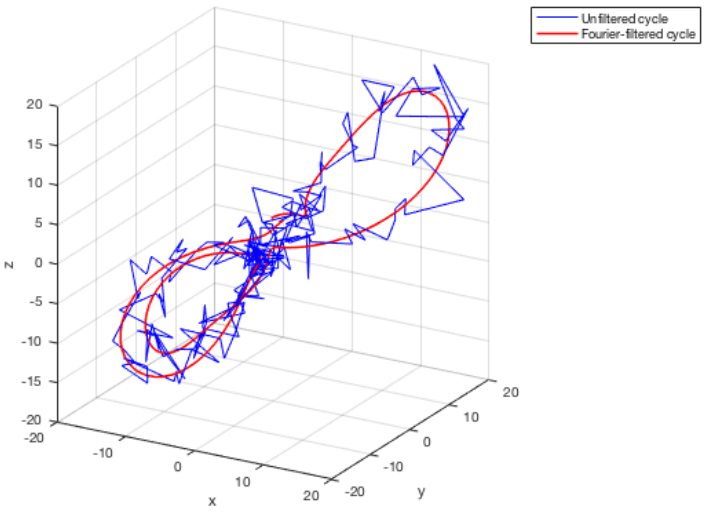


Figure 5.32: Noise-reduced cycle of period 3T.

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	85.34%	Base	0.94	0.47	0.34	99.75%	0.93
		Fitted	0.46	0.15	0.12	99.92%	0.96
10%	82.56%	Base	1.91	1.92	0.69	98.97%	
		Fitted	0.72	0.30	0.20	99.84%	0.94
15%	81.78%	Base	2.79	4.06	0.99	97.82%	
		Fitted	0.97	0.51	0.25	99.73%	0.93
20%	80.26%	Base	3.79	7.45	1.34	96.00%	
		Fitted	1.26	0.84	0.35	99.55%	0.94
25%	78..49%	Base	4.68	11.34	1.69	93.90%	
		Fitted	1.55	1.29	0.46	99.32%	0.96
50%	50.24%	Base	9.34	45.85	3.32	75.36%	
		Fitted	3.03	5.07	0.82	97.21%	0.99
100%	26.66%	Base	18.72	181.74	6.60	52.40%	
		Fitted	6.01	19.61	1.63	89.60%	0.95

Table 5.10: Variation in goodness of fit of the SUNR approximant for Lorenz system with added Gaussian white noise. The maximal (positive) Lyapunov exponent is estimated using Wolf's algorithm. The value derived for the noise-free time series is LE=0.93 (base e). The time step is 0.01.

In Table 5.10 we summarise the goodness of fit results for the SUNR approximant for the Lorenz system for added GWN up to 100%. The distance metrics show excellent reductions throughout the entire noise range and particular at the higher noise levels. At 50% added noise the distance metrics are similar to the values for an unfiltered time series with 15% added noise. At 100% the coverage rate is low (28.23%) and the result is dominated by the effects of moving average pre-filter. The R^2 values show a good fit, retaining values greater than 99% up to 25% noise and 97% at 50% noise. The estimates of the maximal Lyapunov exponent fluctuate a little and are a little higher the true value ($\lambda_{\max} = 0.91$ s^{-1}). Estimates of the maximal Lyapunov exponent using numerical algorithms can vary slightly depending on the calibration of the algorithm (the input parameters) and our noise-free estimate is $\lambda_{\max} = 0.93$ s^{-1}, so our starting position is slightly higher than the true value. The highest estimate is ($\lambda_{\max} = 0.99$ s^{-1} for noise at 50%) is only 6.5% from our noise-free estimate.

5.6.3 The Rabinovich-Fabrikant System

The Rabinovich–Fabrikant equations [112] are a set of three coupled ordinary differential equations exhibiting chaotic behavior for certain values of the parameters. It is not a highly unstable system with the maximal Lyapunov exponent typically less than 0.5 s^{-1} for the examples studied in the literature. We include this system as an example where the recurrence matrix method of cycle detection runs into problems. This system was originally presented as a mathematical model of the stochasticity arising from the modulation instability in a non-equilibrium dissipative medium. It is a simplification of a complex nonlinear parabolic equation modelling various physical systems, such as wind waves on water, and concentration waves during chemical reactions in a medium where diffusion may occur. The equations model a real physical system (not an artificial model) and a rigorous mathematical analysis cannot be carried out due to the strong nonlinearity.

The Rabinovich–Fabrikant (RF) system is an example of a chaotic system that is difficult to analyse [32], largely due to the quadratic and cubic terms in the system equations. Danca et al. have found that different attractors can be obtained for the same parameters by using different methods of numerical integration or different step sizes in the integration. It also thus poses real challenges to numerical methods for ODEs [31]. The system behavior depends mostly on the parameter a, and to a lesser extent on b. The system has several different chaotic attractors with different shapes (see [31]). Also, having five equilibrium points, the RF system is topologically non-equivalent to many classical systems, such as the Lorenz and Rossler systems. The Rabinovich–Fabrikant system is defined by the following nonlinear equations:

$$\frac{dx}{dt} = y(z - 1 + x^2) + bx, \tag{5.11a}$$

$$\frac{dy}{dt} = x(3z + 1 - x^2) + by, \tag{5.11b}$$

$$\frac{dz}{dt} = -2z(a + xy), \tag{5.11c}$$

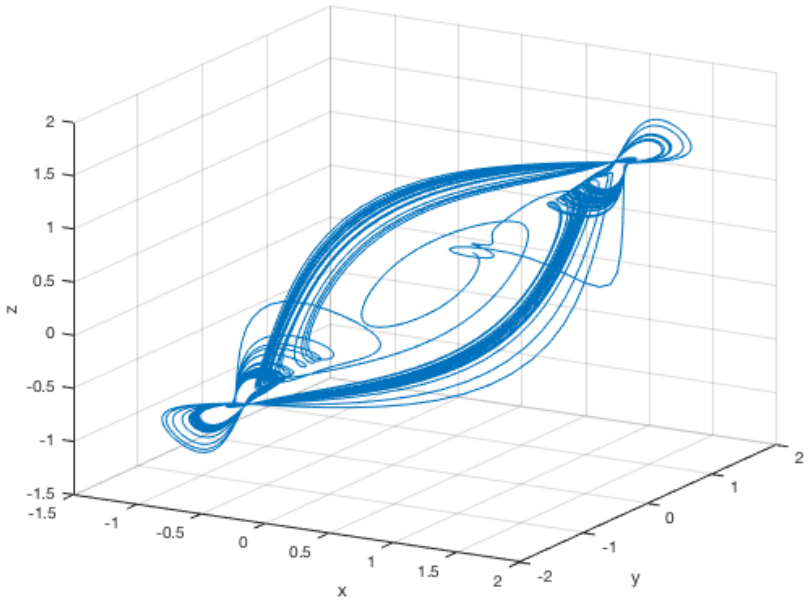
where the parameters are $(a, b) = (0.98, 1.0)$, time step $h = 0.2$ and initial point

is $(x_0, y_0, z_0) = (0.1, 0.1, 0.1)$.

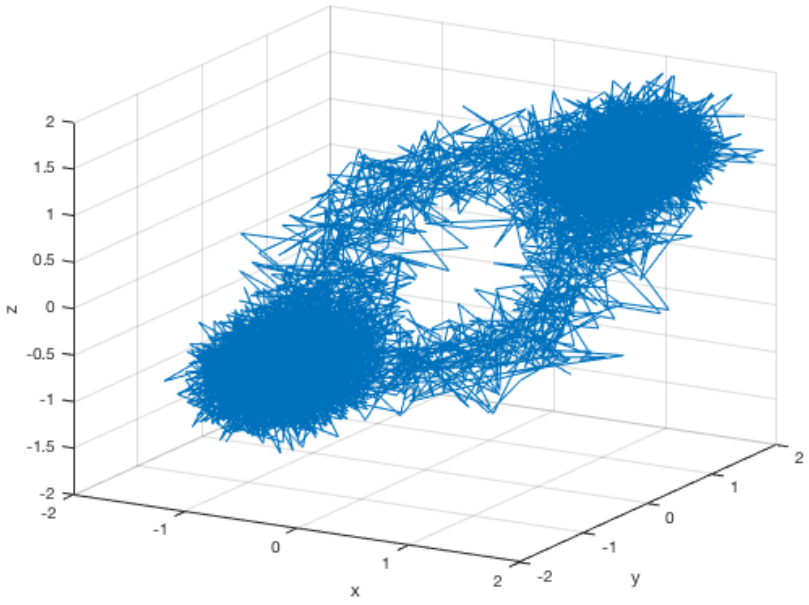
The numerically integrated Rabinovich–Fabrikant attractor is shown in three dimensional phase space in Figure 5.32a. Every orbit in the system is invariant under the transformation:

$$T(x, y, z) \to T(-x, -y, z). \tag{5.12}$$

The Rabinovich–Fabrikant system is remarkable in that a detailed study reveals so many chaotic features, for example, the coexistence between several types of attractors, cycling chaos, hidden attractors and transient chaos. One of the most interesting features is that different attractors can be obtained for the same parameters by using different step sizes in the integration. For some values of the parameter a, the results depend drastically on the step-size, the initial conditions, and the numerical methods used. The standard Runge-Kutta method RK4, utilised in this thesis, generally gives more accurate results, although in some cases these are strongly dependent on the step-size h.



(a) Noise-free



(b) GWN=25%

Figure 5.33: (a) Rabinovich-Fabrikant chaotic attractor (noise-free) with parameters (a, b) = (0.98, 1.0), time step=0.2 and initial point =(0.1,0.1,0.1). This is our illustrative case. (b) The same Rabinovich-Fabrikant chaotic attractor with 25% added Gaussian white noise.

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	52.73%	Base	0.11	0.01	0.04	99.75%	0.079
		Fitted	0.09	0.01	0.03	99.76%	0.073
10%	49.94%	Base	0.23	0.03	0.08	99.00%	
		Fitted	0.11	0.01	0.04	99.69%	0.069
15%	49.96%	Base	0.34	0.06	0.12	0.98%	
		Fitted	0.14	0.01	0.05	1.00%	0.074
20%	51.36%	Base	0.46	0.11	0.17	95.88%	
		Fitted	0.18	0.02	0.06	99.34%	0.061
25%	47.71%	Base	0.57	0.17	0.20	93.74%	
		Fitted	0.21	0.02	0.07	99.15%	0.073
50%	47.12%	Base	1.13	0.67	0.40	75.36%	
		Fitted	0.38	0.04	0.09	99.76%	0.076

Table 5.11: Goodness of fit of the SUNR approximant. Clearly the approximation is superior to the noisy time series on all goodness of fit metrics and we are able to directly calculate an estimate of the maximal Lyapunov exponent. The time step is 0.2.

We apply the SUNR technique illustrated previously to locate and filter the time series. Analysis of the goodness of fit data in Table 5.11 shows a very similar results to those obtained from the Chua and Rossler systems. The distance metrics (particularly the RMSE) show excellent reductions throughout the entire noise range including at the higher noise levels. As with the other systems, for 50% added noise the distance metrics are similar to the values for an unfiltered time series with 15% added noise. The coverage rate remains close to 50% for all noise levels and is almost completely comprised of partial shadow-UPOs. If we restricted ourselves to complete cycles only, we would be unable to produce any approximant. The recurrence histogram is shown in Figure 5.34 and we can see clear peaks. Thus for this very difficult system we can construct an approximation using only partial cycles. The R^2 values show a good fit, retaining values greater than 99% up to 50% noise. The estimates of the maximal Lyapunov exponent fluctuate a little but are close estimates of the true value ($\lambda_{\max} = 0.079$ s^{-1}). The recurrence matrix method requires the orbit to complete two loops near a UPO before being "detected". This is a strong condition and in this case we capture few complete cycles (although they clearly exist). As stated, we do capture a great number of partial cycles. This

outcome is a result of recurrence matrix methodology rather than because of high instability in the system.

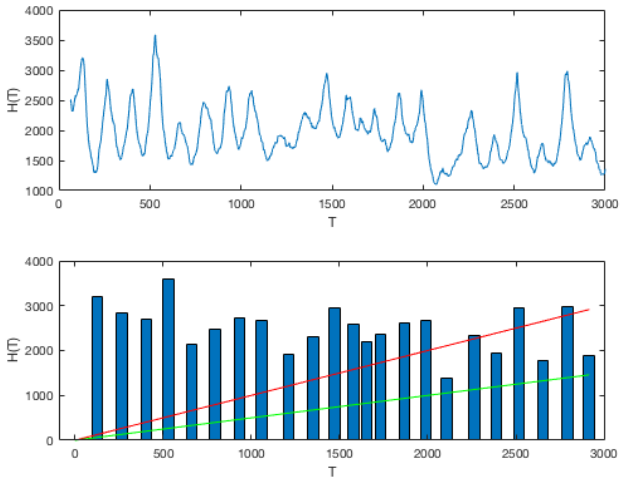


Figure 5.34: The recurrence histogram for the Rabinovich-Fabrikant chaotic attractor with 25% added Gaussian white noise. Note although we identify periods with peaks, we find few complete cycles but many partial cycles.

There are three cases where the recurrence histogram approach may not yield useful results. Firstly, when orbits that are highly chaotic (very unstable) and unlikely to remain on a specific cycle for any appreciable time and we may not detect complete cycles. The Lu-Chen system, which we will examine shortly, is such an example. This was clear from the recurrence plot in Figure 3.2e where few horizontal lines of any length were visible. In this case we may detect partial cycles only. A second situation is where cycle periods are not distributed as discrete subharmonics of a fundamental period, but rather as a near continuous spectrum. An example is a system that spirals outwards, taking a few more steps in each subsequent cycle. A third situation is where the orbit evolution is complicated and where cycles do not complete double loops near a UPO. They may complete a (complex) single cycle pattern, before moving on and perhaps retracing a similar single cycle later. The Rabinovich–Fabrikant system appears to exhibit this type of evolution based on observation of the time series and slowly evolving the system

forward in time. The cycles of the RF system are "bunched", remaining in a tube-like formation and are stable enough to detect partial cycles with discrete periods. For these situations the recurrence matrix methodology may be found wanting and other approachs to the cycle detection phase of the SUNR need to be considered.

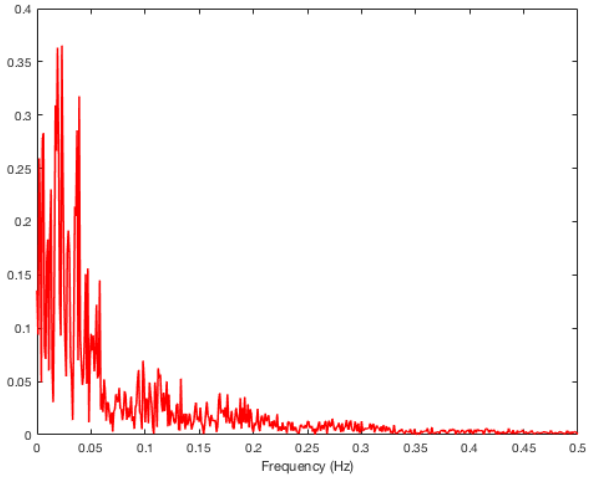


Figure 5.35: The FFT of the noise-free time series reveals a complex mixture of frequencies within the Rabinovich-Fabrikant system.

One might expect a single fundamental frequency peak with wide sides representing the orbits, to be visible in the Fourier spectral response (Figure 5.35); however the orbit does not evolve by traversing single laps of the attractor loop. A closer look at nearby detected cycles reveals a multitude of paths where for example the trajectory may orbit the small end nodes a single time or several times before proceeding. This behavior makes it difficult to categorise individual UPOs. An example is provided in Figure 5.36. In systems with sub-harmonics and simple distinct cycle patterns, the recurrence method works extremely well and provides strength and rigor in specifically identifying individual cycles. However our objective is primarily to identify "near-cycles" for noise-reduction purposes and all we require is that start and end points are near each other with relatively smooth and uncomplicated dynamics in between. In these situations where the attractor does not travel widely and fill out the available space, we will tend to see long

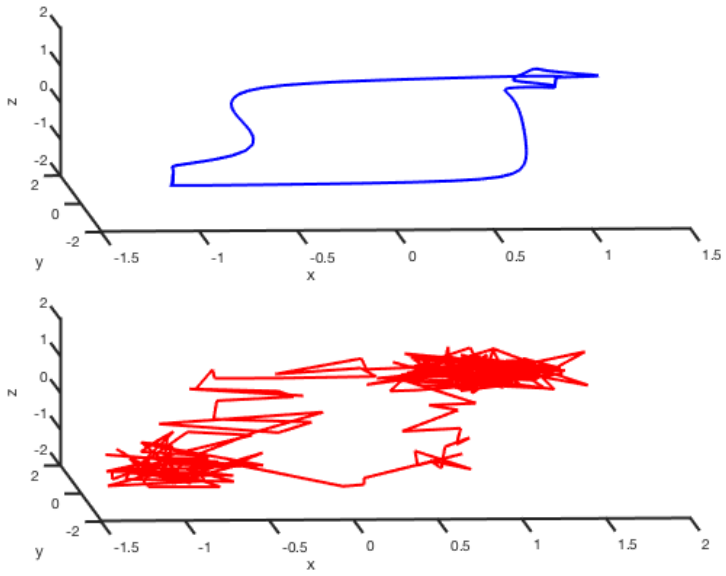
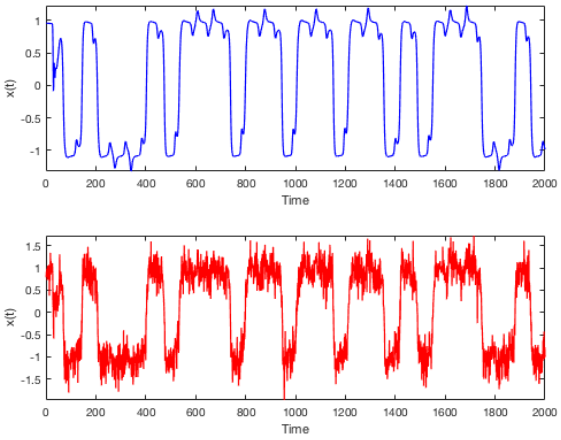


Figure 5.36: A cycle of the Rabinovich-Fabrikant chaotic attractor, both noise-free (Top) and with 25% added Gaussian white noise (Bottom). The path followed by the orbit is highly complex and not easily deconstructed into simple UPOs. However, there are a multitude of cycles present, which can be individually filtered provided they can be extracted.

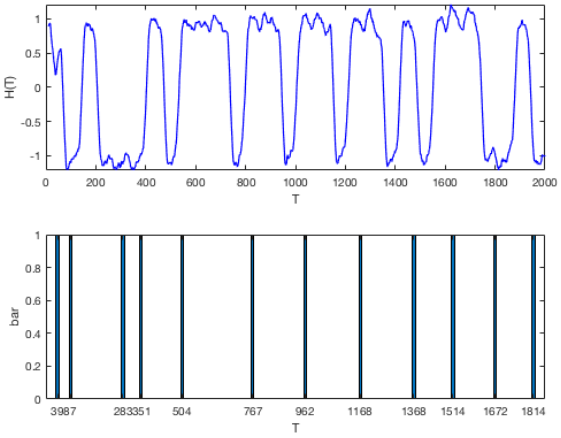
repetitive, similar looking cycles. These may not be shadow-UPOs but we can still easily extract and filter these, provided our goal is simply to reduce noise from the time series.

This limitation of the recurrence method matrix suggests we modify the cycle detection methodology employed within the SUNR method. The cyclical structure of the Rabinovich–Fabrikant system is particularly obvious from the attractor and also from the time series as shown in Figure 5.37 below. We see in Figure 5.37a that the cycles are well formed and clear, although without any apparent order of appearance. In order to more easily locate and filter cycles, we relax the strong close returns condition of requiring at least two consecutive circuits to be deemed a cycle. Instead we identify the local minima as start and end points of cycles. We first pre-process the data with a moving average filter to increase resolution of the local minima and detect the minima by modifying the peak detection algorithm

in Matlab. In Figure (5.37b) we have segmented the noisy time series into approximate cycles using local minima and requiring that start-end points are within a given distance tolerance of each other (in this case 0.75). Individual cycles are filtered as before and replaced back at their original location in the time series. We illustrate several of the detected cycles in Figure 5.38 below.

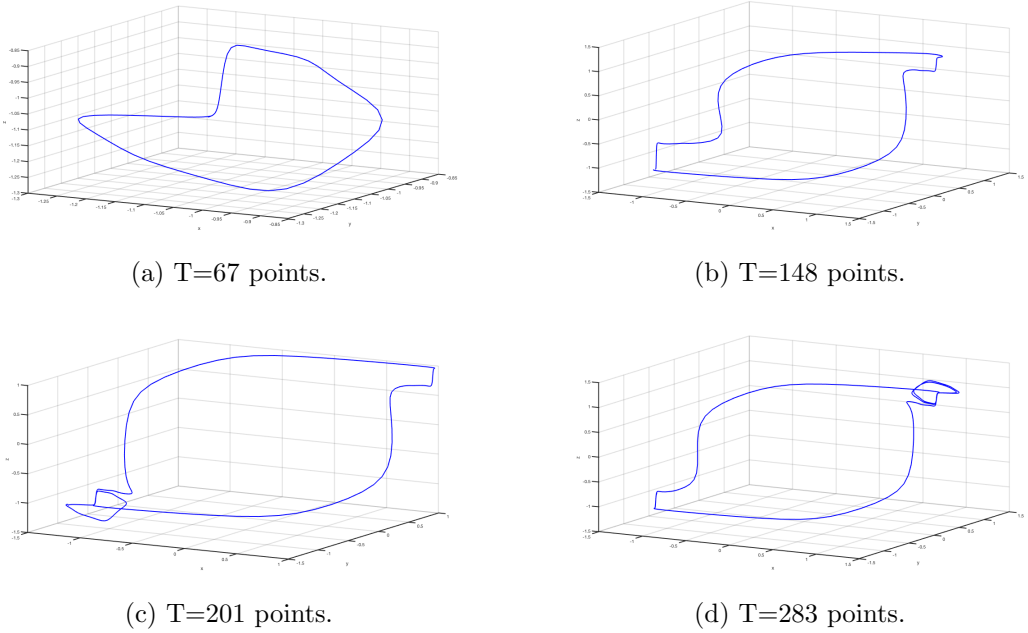


(a) Rabinovich-Fabrikant time series.



(b) Local minima.

Figure 5.37: (a) Time series for the Rabinovich-Fabrikant chaotic attractor, both noise-free (Top) and with 25% added Gaussian white noise (Bottom). (b) Minima (below) of noise-infected Rabinovich-Fabrikant system.



(a) T=67 points.

(b) T=148 points.

(c) T=201 points.

(d) T=283 points.

Figure 5.38: Cycles from the noise-free Rabinovich-Fabrikant system.

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	96.51%	Base	0.11	0.00	0.04	99.76%	0.079
		Fitted	0.10	0.00	0.03	99.74%	0.073
10%	96.45%	Base	0.23	0.03	0.08	99.00%	
		Fitted	0.11	0.01	0.04	99.68%	0.087
15%	96.43%	Base	0.34	0.06	0.12	97.74%	
		Fitted	0.14	0.01	0.05	99.57%	0.079
20%	96.56%	Base	0.46	0.11	0.17	95.59%	
		Fitted	0.17	0.02	0.05	99.38%	0.079
25%	99.11%	Base	0.57	0.17	0.20	93.84%	
		Fitted	0.20	0.02	0.06	99.24%	0.072
50%	98.91%	Base	1.15	0.69	0.41	74.93%	
		Fitted	0.33	0.06	0.10	97.88%	0.076

Table 5.12: Goodness of fit of approximant to Rabinovich-Fabrikant system based on segmentation into cycles using local minima. Results are similar to those obtained from the SUNR method. The time step is 0.2.

The goodness of fit of the SUNR approximant determined using local minima are summarised in Table 5.12. The results are very similar to those using the recurrence matrix method of detecting cycles, albeit with higher coverage rates. The similarity of results is supportive of the SUNR approach, but more importantly is illustrative of the fact that chaotic attractors, although very periodic in structure are often very different in how they evolve along cycles in time. Finding cycles may require some creativity in approach and require examination of the system. Note that in all cases there was no phase lag between the noise-free time series and the approximation. Finally, in Figure 5.38 we show the approximation to the noise-free Rabinovich-Fabrikant attractor, that was constructed using the local minima approach to cycle detection.

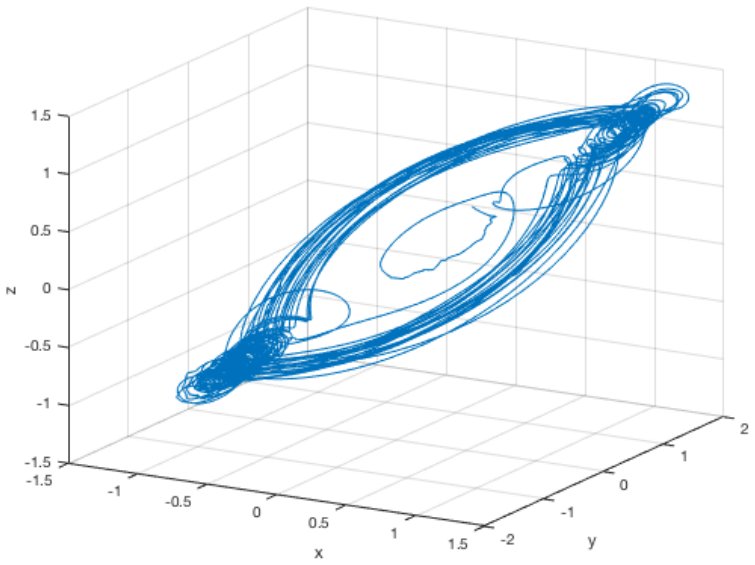


Figure 5.39: Attractor for noise-reduced Rabinovich-Fabrikant system using time series local minima.

5.6.4 The Lu-Chen System

Multiscroll attractors (attractors with multiple scrolls in a single attractor)) include the Lu-Chen attractor, the Rabinovich-Fabrikant attractor, the modified Chen chaotic attractor, the Duffing attractor, and the modified Chua chaotic attractor. The Lu-Chen chaotic system [97], [96] is an example of a system that lies topologically "in between" the Lorenz/Rossler and Lu-Chen/Chua systems. The Lu-Chen system is is defined by the following nonlinear equations:

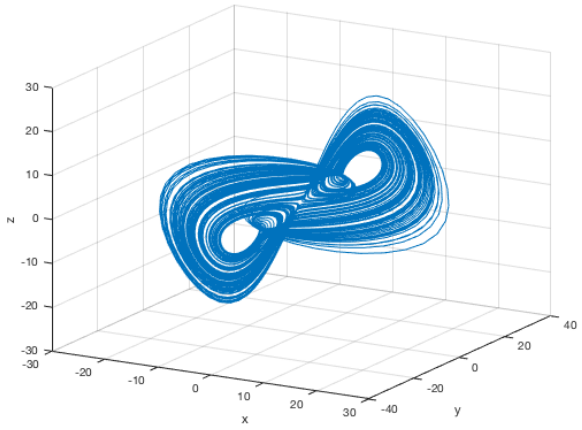
$$\frac{dx}{dt} = a(y - x), \tag{5.13a}$$
$$\frac{dy}{dt} = x - xz + cy + u, \tag{5.13b}$$
$$\frac{dz}{dt} = xy - bz, \tag{5.13c}$$

where the parameters are $(a, b, c, u) = (36, 3, 20, 0)$, time step $h = 0.01$ and starting point is $(x_0, y_0, z_0) = (0.1, 0.3, -0.6)$. The noise-free Lu-Chen attractor is shown in Figure 5.40a.

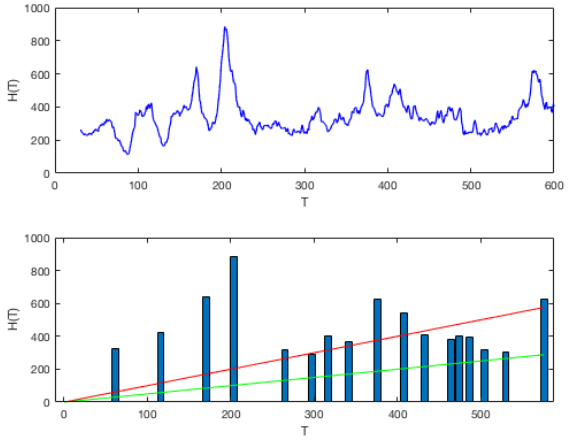
The Lu-Chen system is highly chaotic ($\lambda_{\max} = 1.75$ s^{-1}) and the recurrence matrix approach to detecting shadow-UPOs does not identify cycles. This is an example of a system where the instability is such that the orbit will not evolve near to any specific UPO long enough for it to be detected. This instability was observed visually in Figure 3.2e where horizontal lines of any significant length were not visible on the recurrence plot. This recurrence histogram is shown in Figure 5.40b and we see that peaks are not clearly discernable. Shadow-UPOs of many periods are present but we cannot detect any material amount of complete or partial cycles. There are clearly complete cycles present in this system but they are not detectable using the recurrence matrix method.

Nonetheless, the recurrence matrix approach is still an excellent first option approach as the rigor involved does ensure cycles are near real UPOs and can be discretely binned and analysed, with the goal of calculating a basis set of lower order noise-free UPOs. The recurrence plot histogram also provides valuable

information about cyclic point distribution, and in fact, we can often still detect the system harmonics, even if we cannot recover the actual cycles.



(a) Noise-free Lu-Chen attractor.



(b) Recurrence-histogram.

Figure 5.40: (a) The double-scroll Lu Chen attractor using a time step of 0.01, (b) Histogram derived from recurrence matrix of Lu Chen system. Several peaks are identified but no complete cycles are identified within those peak harmonics. There are periodic points of all frequencies present, but this does not guarantee an unbroken repeated sequence will occur.

Unlike the Rabinovich-Fabrikant system, the Lu-Chen does not appear to evolve in highly complex cycle patterns. However, it is unlikely to complete two complete consecutive loops near a UPO and is thus also not detectable using recurrence methods. For some attractors, particularly those of a more unstable nature, the orbit will simply not evolve in this manner. We show the frequency domain profile of the Lu-Chen signal after transformation using the FFT which shows a very broad distribution of frequencies and emphasises the instability in this system. The Lu-Chen system may be well be an example of a system with adjacent periods that are just a few points apart, with the orbit spiralling in and out. This hypothesis requires further investigation.

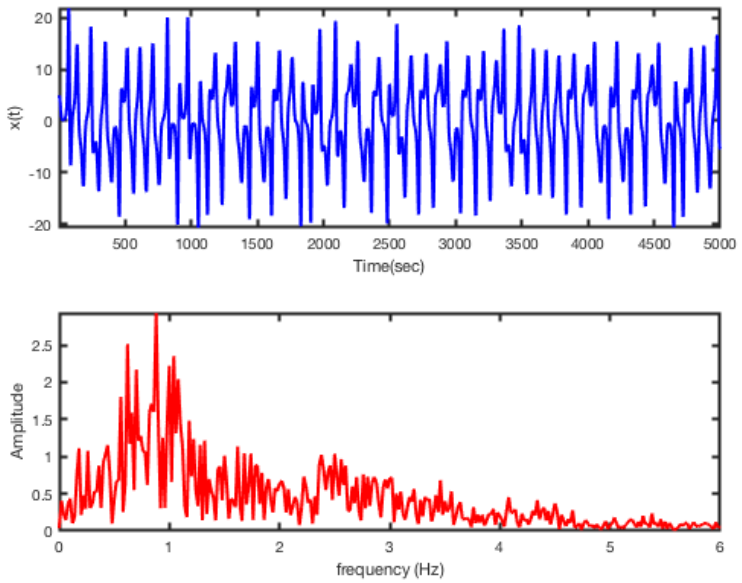
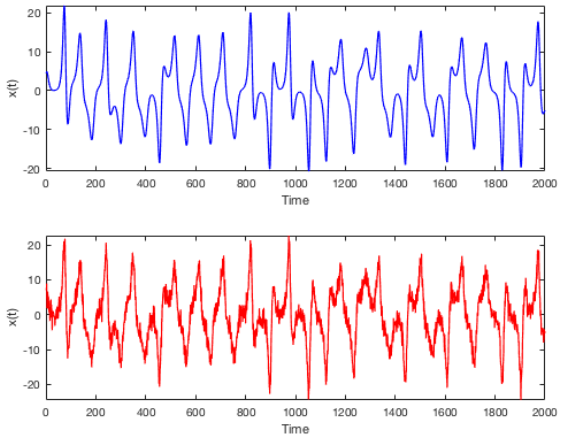


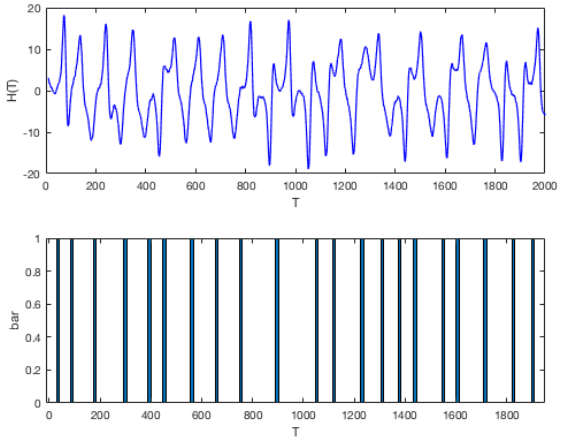
Figure 5.41: The FFT of the noise-free Lu-Chen system. One can see the almost continuous spectrum of cycle periods that define the attractor.

As an alternative to the recurrence matrix method, we again identify the local minima as start and end points of cycles. We follow the previous methodology, deconstructing the time series into approximate cycles, by first pre-processing the data with a moving average filter and detecting the minima. In Figure 5.42b we have segmented the noisy time series into approximate cycles using local minima

and requiring that start-end points are within a given distance tolerance of each other (in this case 0.5). Individual cycles are filtered as before and replaced at their original location in the time series.

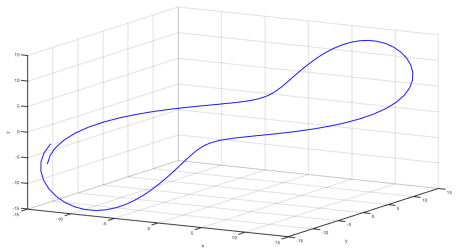


(a) Lu-Chen time series.

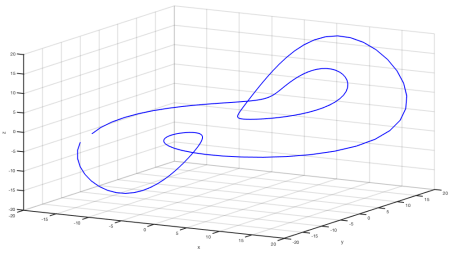


(b) Local minima.

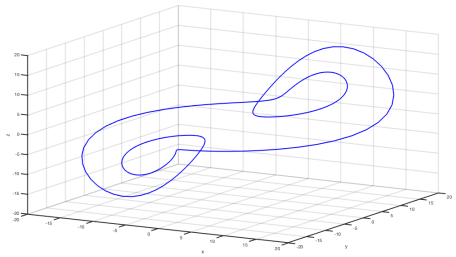
Figure 5.42: (a) Time series of Lu Chen system, both noise-free (top) and noise-infected (below). The periodicity, although irregular, is obvious, (b) Local minima (below) of noise-infected Lu Chen system.



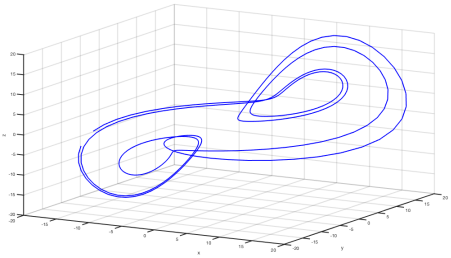
(a) T=100 points.



(b) T=170 points.



(c) T=206 points.



(d) T=372 points.

Figure 5.43: Cycles from the noise-free Lu-Chen system.

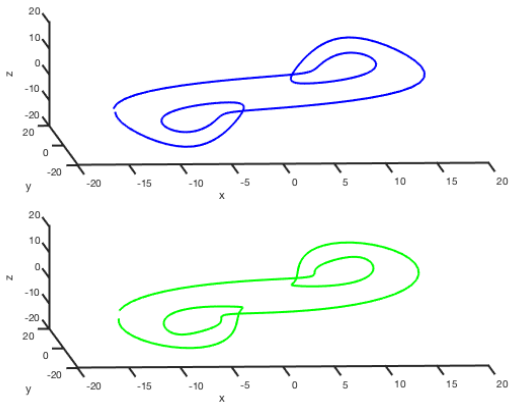


Figure 5.44: Gaussian white noise-infected Lu-Chen cycle of 202 points, before (Top) and after FFT filter (Bottom) to remove 25% added GWN.

The highly cyclic structure of the system is visible in the time series shown in Figure 5.42a and in Figure 5.43 we show 4 of the detected cycles using the local minima approach. In Figure 5.44 we show a noise-free cycle of period 202 time steps and the noise-reduced cycle after the application of the Fourier filter. We now consider the goodness of fit of the approximant.

Noise Level	**Coverage %**	**Comparison**	MAE	RMSE	MME	$\mathbf{R^2}$	**LE**
5%	97.87%	Base	0.93	0.45	0.33	99.76%	1.74
		Fitted	1.31	1.35	0.40	99.28%	1.82
10%	97.10%	Base	1.88	1.85	0.66	99.01%	
		Fitted	1.58	1.94	0.54	98.96%	1.93
15%	96.71%	Base	2.84	4.22	0.99	97.74%	
		Fitted	1.74	2.14	0.55	98.85%	1.82
20%	90.53%	Base	3.82	7.65	1.33	95.89%	
		Fitted	1.93	2.45	0.58	98.69%	1.95
25%	94.49%	Base	4.73	11.65	1.67	93.75%	
		Fitted	2.07	2.77	0.64	98.51%	2.02
50%	96.68%	Base	9.60	48.27	3.40	74.09%	
		Fitted	3.08	5.35	0.86	97.12%	2.31
100%	96.75	Base	18.78	185.54	6.75	4.54%	
		Fitted	5.32	14.68	1.39	92.13%	2.37

Table 5.13: Goodness of fit of approximant to Lu Chen system based on segmentation into cycles using local minima. The maximal (positive) Lyapunov exponent is estimated using Wolf's algorithm. The value derived for the noise-free time series is LE=1.75 (base e)(1/sec). The time step is 0.01.

These goodness of fit results are considerable improvements over the base case that compares the noise-free and noise-infected signals. The improvements are noticeably not as great as for the chaotic systems of lower-instability presented as examples so far. The estimate of the maximal Lyapunov exponent inflates with the noise and is reasonable, in the context of the high instability, up to 20%–25% added noise. For highly unstable chaotic systems it is unlikely we will recover the maximal Lyapunov exponent precisely and will have to accept some error in the result. Finally in Figure 5.45 we show the SUNR approximation to the Lu-Chen system with 25% added GWN.

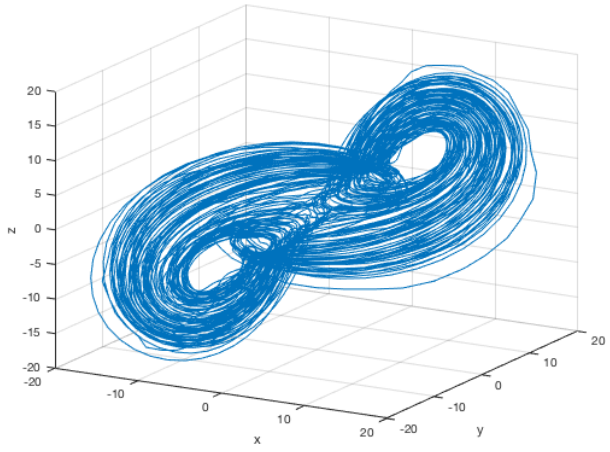


Figure 5.45: Approximation of Lu-Chen attractor, after reducing 25% added Gaussian white noise. The time step is 0.01.

5.7 Improving Results by Multi-Sampling

In most experimental situations, the experimenter will be able to capture multiple samples of a time series, rather than just a single set of points. We have modelled this situation for each of our 3 main model chaotic systems: Rossler, Lorenz and Chua. We constructed a looping model, using randomised starting points, that builds a model noise-infected (GWN) time series and steps through each process of the SUNR method, finally outputting an approximation of each noise-free time series. We then use the SUNR approximant to estimate the maximal Lyapunov exponent using Wolf's algorithm. The results are shown in Figures 5.46–5.48 below. In each case we constructed 50 time series samples of 10,000 points each and added 25% GWN.

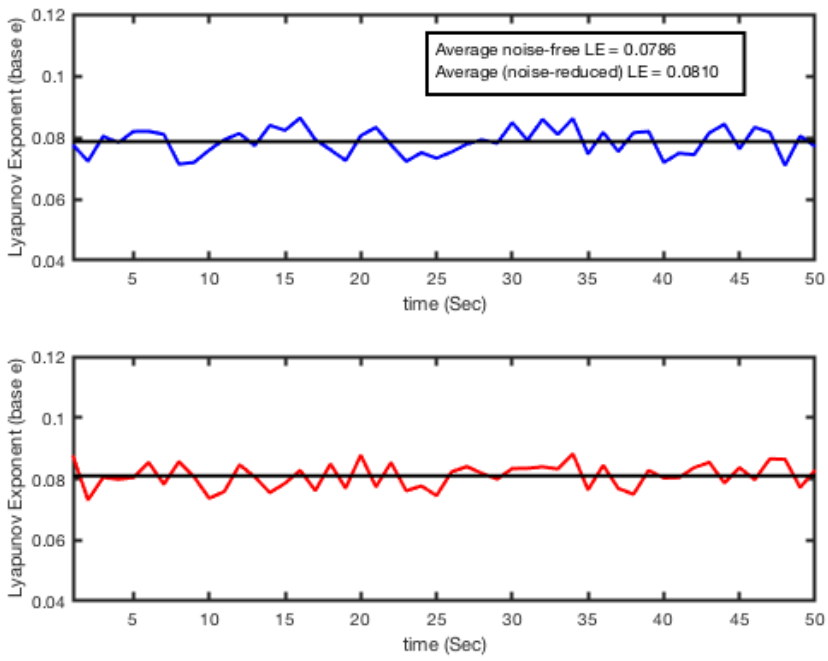


Figure 5.46: Estimate of maximal Lyapunov exponent using 50 Rossler time series samples. Estimate using noise-free samples is 0.079 s^{-1} (shown in blue) and noise-infected samples is 0.081 s^{-1} (shown in red).

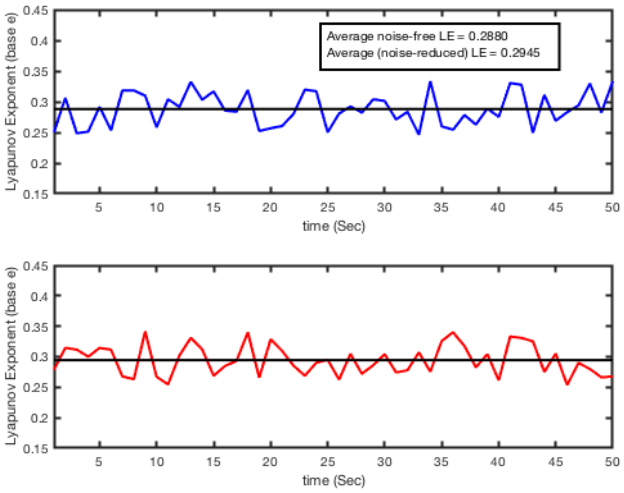


Figure 5.47: Estimate of maximal Lyapunov exponent using 50 Chua time series samples. Estimate using noise-free samples is 0.29 s^{-1} (shown in blue) and noise-infected samples is 0.29 s^{-1} (shown in red).

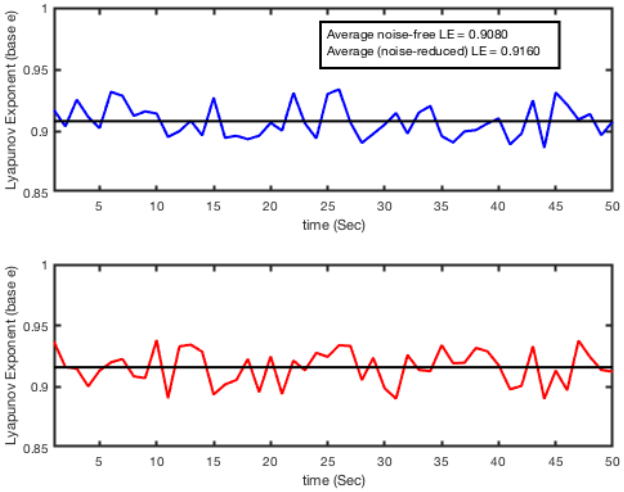


Figure 5.48: Estimate of maximal Lyapunov exponent using 50 Lorenz time series samples. Estimate using noise-free samples is 0.91 s^{-1} (shown in blue) and noise-infected samples is 0.92 s^{-1} (shown in red).

The multi-sampling averages out variations in the SUNR estimates of the maximal Lyapunov exponent and provides a more accurate result. In Table 5.5 we summarise the estimate of the average noise-free maximal Lyapunov exponent using both the noise-free data and the noise-infected data.

System	LE (Noise-free)	LE (25% GWN)	True Value
Rossler	0.079	0.081	0.079
Chua	0.29	0.29	0.28
Lorenz	0.91	0.92	0.91

Table 5.14: Summary of average estimates of maximal Lyapunov exponents using 50 time series samples and the SUNR method.

5.8 Summary and Discussion

Our focus in this chapter is adapting the SUNR method to chaotic time series with higher instability (i.e. at the higher end of the range $0 < \lambda_{\max} \leq 1$). In these systems the instability is such that we may not be able to detect sufficient complete cycles of all periods to estimate a full noise-free basis set of individual lower order UPOs. The instability may result in not all periods being detectable, fewer complete cycles and more partial cycles; or both. For these systems with higher instability we can still estimate the noise-free time series. This will also utilise the valuable information contained in the partial shadow-UPOs. Provided the residual noise is sufficiently low in the approximated noise-free time series, we may then determine invariants by using this approximation in the conventional numerical algorithms. This alternative approach involves individually noise-filtering all detected complete and partial shadow-UPOs and using these to construct an approximation of the noise-free time series. In this way we maximise use of the available information.

The SUNR method for these higher instability chaotic systems will produce an approximation to the noise-free time series. We thus start by defining a set of goodness of fit metrics for comparing two time series for equality. The set of GOF metrics we use includes the three point-wise distance metrics (MAE, RMSE and MME), the coefficient of determination (R^2) and phase lag.

We provide a heuristic example using the Rossler system with 25% added GWN, detecting and extracting complete and partial cycles using the noise-resilient recurrence histogram. We describe methods to reduce the noise from both complete and partial shadow-UPOs. Complete shadow-UPOs have a small start-end point discontinuity, limited in size by the critical radius. We apply an end-point smoothing (EPS) transformation to these complete cycles, to eliminate artefacts of the Fourier transform before applying a FFT and filtering noise in the frequency domain. The linear EPS transformation is inverted after filtering. Partial cycles are transformed in hybrid complete cycles by concatenating the partial cycle and its mirror image ("mirroring"). Again, this reduces the problems arising from incom-

plete cycles when the FFT is applied. These are similarly filtered in the frequency domain, inverted and the original partial shadow-UPO is recovered. We applied this methodology to a single Rossler time series and the goodness of fit metrics were significantly improved relative to the noise-infected state. As expected, the results of using a simple cut-off low-pass filter produced results similar to that of a moving average filter.

A strength of this method is that if we have knowledge of the type of measurement noise we can use a targeted filter in the frequency domain rather than a simple cut-off, which is a worst-case approach for any broadband type of noise. We illustrate using the low-pass filter with the knowledge that any targeted noise-reduction filter will produce superior results and testing the SUNR method this way produces the most conservative results. We apply the simple low-pass filter to time series with different types of added noise (Uniform white, Gaussian white, high-frequency and pink coloured) and observe that detection of unique cycles is better when noise is broadband and not localised in the frequency spectrum. It is also more difficult to remove noise from a signal in these cases and there is a detection-noise reduction trade off. We provide a simple example using pink noise to demonstrate how targeted noise-reduction strategies based on the type of noise will produce excellent results. By deconstructing the time series into approximate cycles, we may apply the best noise-reduction technique locally to each cycle.

We consider the variation in results for the Rossler example as measurement noise is increased from 0% to 100%. The SUNR approximation to the noise-free time series has sufficiently low residual noise to produce a good estimate of the maximal Lyapunov exponent for noise levels up to 25%. Between 50% and 100% the Lyapunov exponent estimate increases by a factor of two, most likely due to residual noise being a little high. To address noise at 50% or 100% level, the approximation can be re-filtered by a more sophisticated technique to reduce residual noise further.

The SUNR method was applied to the Rossler example with dynamical noise added but estimates of the maximal Lyapunov exponent were not very accurate

(possibly because we cannot use partial cycles). We see how dynamical noise deforms the shadow-UPOs but retains smoothness and occupies the same frequencies as the signal in the frequency domain, making it highly problematic. We subsequently tried multi-sampling and collected shadow-UPOs in bins. If we assume the dynamical noise to have a Gaussian distribution with mean zero, then the deformations should average out over sufficiently many samples (the chaotic attractor is bounded in phase space). The detected shadow-UPOs were averaged as in Chapter 4 and we compared the estimated noise-free UPOs with those obtained using the noise-free time series. We observed that the estimated noise-free UPOs were reasonably good approximations to the true UPOs, although not highly accurate. The dynamical noise based estimates are sufficiently close to the true UPOs for qualitative use, for example classifying the chaotic attractor.

Noting that dynamical noise appears to conserve cycles (up to a point) and does not destroy smoothness in the cycles, we found that directly using a dynamical noise-infected Rossler time series in the numerical algorithm by Wolf to estimate maximal Lyapunov exponents did not cause the algorithm to fail. Results did vary from one time series to another, so we used multiple samples (50) and averaged the result. Surprisingly we found the average determined by multi-sampling was a very good estimate of the noise-free result. The exercise was repeated multiple times and tested on the Lorenz and Chua systems with excellent results also. Further research is required to understand the mechanism here, but this is a useful result.

Finally we presented results after application of the SUNR method to several other systems of varying topology, including the Chua, Lorenz, Rabinovich-Fabrikant and Lu-Chen systems. The SUNR method works well for the relatively low-instability Rossler, Lorenz and Chua systems, reducing high levels of measurement noise and producing sensible estimates of the maximal Lyapunov exponent.

The Rabinovich-Fabrikant system is presented as example of a low-instability system with topology that is problematic for the recurrence matrix method of detecting cycles. It has a relatively small maximal Lyapunov exponent (0.07) and yet we did not detect any complete shadow-UPOs but many partial shadow-

UPOs. This system evolves, tracing out complicated single cycles that do not repeating immediately. The recurrence method of detecting cycles will not detect the complete cycles. We did however detect large numbers of partial cycles and obtained stable goodness of fit results using the recurrence method. This is a case demonstrating the value of partial shadow-UPOs as without these we would not get a result at all. Realising the limitations of the recurrence matrix method for this system, we modified the cycle detection process employed by the SUNR method, using local minima in time series to partition the time series into "cycle-like" segments (which are not necessarily shadow-UPOs). We then subsequently applied the FFT noise filter to each cycle and obtained good estimates of the maximal Lyapunov exponent. Goodness of fits were similar for both detection methods, further validating the results.

For highly unstable systems, the SUNR method simply fails to work as we cannot detect sufficient complete or partial shadow-UPOs. This is evidenced using the Lu-Chen system which is highly chaotic with maximal Lyapunov exponent of 1.75 s^{-1}. Again, realising the limitations of the recurrence matrix method for highly unstable systems, we used the local minima cycle detection approach. We obtained estimates of the maximal Lyapunov exponent (between 1.75 s^{-1} and 2.0 s^{-1}). This modified approach yielded a result when the recurrence method did not, highlighting that we need to adapt the cycle detection methodology to the topology of the chaotic system.

Finally we generated multiple samples for each of the Rossler, Lorenz and Chua systems, obtaining a SUNR approximation to each sample. We then applied the Wolf algorithm to each approximated time series, averaging the results. This averaging process provides accurate estimates of the maximal Lyapunov exponent for the systems tested.

Chapter 6

Summary and Outlook

6.1 Thesis Summary

In this thesis we have presented, tested and assessed a noise-reduction method (the shadow-UPO noise reduction method) that is applicable to scalar time series sampled from low-dimensional chaotic dynamical systems with medium to high levels of measurement noise or low levels of dynamical noise. The method was developed to tackle the difficulties associated with reducing noise from sampled chaotic time series.

6.1.1 Challenges in Noise Reduction and Invariant Estimation from Chaotic Time Series

The direct application of many conventional noise-reduction filters to a chaotic time series may destructively interfere with the system, effectively corrupting the underlying data. Chaotic signals are intimately bound with the noise in the system through the nonlinearities, and conventional noise reduction methods that are recursive in nature tend to corrupt the underlying signal in the process of reducing noise. Incorrect filtering will actually increase the correlation dimension calculated using standard algorithms and will add an apparent extra Lyapunov exponent to the system being studied. We are severely limited in the selection of conventional filters to just finite impulse response filters (FIR), which fortunately includes the

moving average type. The problem is exacerbated in part because the Fourier frequency power spectrum of a chaotic signal appears similar to broadband noise, so techniques based on frequency separation strategies also run into trouble.

Existing noise reduction techniques for chaotic systems usually rely on initially embedding the noise-infected time series in phase space. Embedding noise-infected data in phase space may result in a devalued system that is not topologically equivalent to the underlying noise-free system. Takens' theorem requires an infinitely long noise-free time series to guarantee topological equivalence between the true chaotic attractor and that derived from the phase space reconstruction. Algorithms to calculate the embedding parameters, the time delay, and embedding dimension, are also highly sensitive to noise. These restrictions reduce the robustness of the technique and limit the application to very low levels of added noise. Existing noise-reduction techniques are often highly effective with lower levels of measurement noise but generally do not perform well as the noise level becomes significant ($> 10\%$). There are few techniques to address dynamical noise in chaotic time series.

Algorithms to directly calculate chaotic system invariants such as the maximal Lyapunov exponent from time series assume noise-free data and generally perform poorly in the presence of noise. It is preferable to have a low residual noise time series approximant to use directly in these algorithms, rather than apply the algorithms to noisy data.

6.1.2 Conceptual Framework

The SUNR method is constructed based on three important observations. First is the idea that the highly cyclic composition of a chaotic attractor provides an excellent means to partition the attractor for noise-reduction purposes. Replacing an intractable noise-infected chaotic time series with a collection of near-cycles, all of which are amenable to targeted linear noise-filtering techniques, is the basis of an approach. Further, the cycles (UPOs) are well established within a formal framework. The formalism of the periodic orbit theory of Cvitanovic provides a

useful framework for the application of numerical methods to UPOs. The set of lower order UPOs provides a natural basis through which we can approximate the time series (and thus the attractor). UPOs are dense on the chaotic attractor meaning that all points within a time series lie within a very small distance of a UPO, usually a lower-order UPO. Many orbits "shadow" the UPOs (and thus we refer to them as shadow-UPOs), with the orbit segments being very nearly periodic but may have a small gap between the start point and end point due to the instability in the system.

UPOs are natural features of a chaotic system. They are the most fundamental element in the definition of chaos, a consequence of the stretching and folding mechanism that underpins chaos. They comprise the "skeleton" of the attractor, are ordered hierarchically and are topological invariants. Noise and sensitive dependence on initial conditions both radically alter the paths of individual orbits in a chaotic system. However we are provided with some protection through the Shadowing Lemma that assures us that there is a shadow-UPO arbitrarily close to the true path, albeit the truth path with a different starting point. Numerically detected shadow-UPOs are representative of the true dynamics of the system. We also propose that the set of lower order UPOs coupled with their associated short time Lyapunov signature form a meaningful description of the chaotic system. The problem of orbital divergence is then discretised down to a set of UPO Lyapunov exponents describing the divergence depending upon which UPO the orbit is traveling on or near to.

The second observation is that cycle detection using recurrence matrices/plots and recurrence histograms is quite robust in the presence of higher noise levels (up to 10%), with peaks representing harmonics of the system remaining detectable when noise is present. This observation appears to have been first made in the early 1990's by Mindlin and Gilpin [104] in the context of the topology of chaos, where they considered close returns plots. They noted resilience to noise of the recurrence-histogram and commented on this. This means that the periods of the UPOs and location of shadow-UPOs may be determined in the presence of measurement noise.

The third consideration is that if we deconstruct the time series into a set of shadow-UPOs, then specific signal-processing techniques based on the type of noise may be bought to the fore to remove noise more effectively. Most noise-reduction techniques in chaos research assume the noise is GWN and do not discriminate between types of noise. Not all noise is the same. With this approach we have the flexibility to bring more sophisticated noise-reduction solutions tailored to the character of the noise.

These are the fundamental ingredients to construct an approximation of a noise-infected time series: we know the system may be approximated by cycles, we can locate the cycles in the presence of high levels of noise and we can utilise specialised noise-reduction techniques.

6.1.3 The Shadow-UPO Noise Reduction (SUNR) Method

We developed this conceptual framework into a noise-reduction method, with the objective of maximising use of cycle information (shadow-UPOs) derived from the noisy time series. This required a number of innovations.

Firstly we found that recurrence-histograms constructed from just the scalar time series are more than adequate to locate system harmonics and often superior to those constructed from pre-embedded data. Thus we avoid the problematic pre-embedding of highly noise-infected data.

We also modified the recurrence-histogram cycle detection methodology to detect and locate shadow-UPOs in the presence of higher noise levels. The recurrence-histogram will facilitate identification of the UPO periods (system harmonics) with very high noise levels (up to 200% in some cases), but the shadow-UPOs will not be detectable from the recurrence plot for noise levels greater than 10%. Increasing the noise level fractures the horizontal lines on the recurrence plot that represent complete and partial shadow-UPOs. We found that pre-processing the data with a short-window centered moving average filter (low-pass filter) will enable detection of cycles at very high levels of added noise (up to 100%). This phenomenon was

discovered by Eckmann et al. in 1987 [41]. This pre-processing of data is not being used as a noise-filter but rather as a technique to enhance the effectiveness of locating cycles and importantly, it does not introduce a phase-shift in the time series. Smoothing of the noise recovers cycle detection rates almost to noise-free levels, with the smoothed noise bridging and weakly re-connecting the cycle sequences that were fractured by the noise.

Critical radius ε and sampling frequency are the two key factors in locating and extracting shadow-UPOs. For the critical radius, we tested the rule of thumb that recommends an optimal minimum value of $\varepsilon \geq 5\sigma$, where σ is the standard deviation of the noise [139] and found it to be reasonable for noise levels less than 5%, but not suitable for higher levels of noise. Also we require the critical radius as large as possible without comprising integrity. We find that using 5%–10% of the maximum attractor extent as a practical maximum critical radius works well for higher noise levels. Also, the result is highly sensitive to sampling frequency and some trial and error is necessary involved ensure a good representation of the system is collected, neither under-sampled or over-sampled. It is important to ensure the shortest cycle is well populated with at least 50 data points.

There are a number of technical issues to understand when locating, extracting, processing and storing cycles and we have described them within the thesis. In particular one needs to aware the detected shadow-UPOs occur in conjugate pairs, that detected shadow-UPOs may be out of phase and require synchronisation for comparison or subsequent processing; artefact cycles will necessarily be detected and must be removed from the accounting (and we provided an algorithm to do so).

The subsequent treatment of the detected and binned shadow-UPOs depends on the instability of the underlying chaotic system. We have determined that these techniques work well on systems with maximal Lyapunov exponent less than 1. This includes most well known low dimensional chaotic systems. The SUNR method takes a different approach depending on where the Lyapunov exponent lies within the instability range $(0 < LE_{\max} < 1)$ s^{-1}. The key differentiator

here is whether we detect large numbers of complete shadow-UPOs (lower instability) or few complete cycles (higher instability). There is no quantitative rule here and this can be simply determined by observation of the recurrence plot.

For systems with lower instability many complete cycles will be detected and can be used to estimate all the complete basis set of lower order UPOs. One should detect, extract and bin the lower order complete cycles. For bins containing significant numbers of shadow-UPOs, first synchronise these cycles using cross-correlation and them average them pointwise. The averaging process will reduce the noise and is particularly useful for broadband noise, where Fourier-based frequency separation approaches will not perform well. If the noise has specific characteristics amenable to a frequency attenuation technique, then we suggest individually filtering each cycle in the frequency domain using a FFT. For poorly populated bins, select a single cycle, the "least unstable" shadow-UPO. This is the one with smallest L_2 distance between start and end points and noise is then reduced using a Fourier filter. The binned data may also be used to estimate short time Lyapunov exponents for each UPO using a modified version of the Sano-Sawada algorithm. The modified algorithm is applied to the cluster of synchronised cycles of a given period and does not require the shadow-UPOs to first be noise filtered as the method implicitly averages the data. The average maximal Lyapunov exponent for the system may be estimated using a weighted average, using the relative proportions of periodic points.

For systems of higher instability, there will be insufficient detected complete cycles to estimate the individual noise-free UPOs. In this case approximate the noise-free time series instead, by individually Fourier filtering each detected complete and partial shadow-UPO. Although we cannot estimate individual UPOs, we can now utilise partial shadow-UPOs which are present in large numbers and contain valuable information. All noise filtered shadow-UPOs are reinserted back at their original location in the pre-filtered time series. This means that the remaining sequences of points that have not been detected as shadow-UPOs are at least noise-reduced by the moving average. In our examples, these typically comprise a small percentage of the time series (less than 10%–20%). Before applying

the targeted noise-reduction strategy in the frequency domain, we transform both the complete and partial shadow-UPOs to ensure they are complete cycles (closed ends). This avoids artefacts of the Fourier transform that arise from incomplete cycles. For complete shadow-UPOs apply a linear end-point smoothing transform is applied to close the small start-end gap. For partial shadow-UPOs construct a hybrid cycle by concatenating the cycle and its "mirrored" version to ensure the Fourier filter works as described. After all transformed smooth cycles are noise-filtered, the transformations are reversed, leaving the noise-reduced complete and partial shadow-UPOs.

The time series approximation can and should be made for lower-instability time series also as it utilises the valuable partial shadow-UPO information and it may provide a more robust estimate of the maximal Lyapunov exponent than that calculated using the estimated individual UPOs.

Finally we emphasise the importance of using multiple samples of the time series. The SUNR method will produce results for systems of lower or higher instability with single sufficiently long time series. If multi-sampling is available, then the effectiveness of the method is multiplied and this is demonstrated in our results. There is no specific advantage to having a single long time series, as generally the ergodicity will ensure that repeated samples with different initial points provide sufficient information about the attractor.

6.1.4 Results from the SUNR Method

We used a model Rossler time series with 25% added GWN noise as a heuristic example to illustrate the SUNR method in detail. The Rossler example was used as a test case case for the methods designed for both low-instability and high-instability chaotic systems. Although the Rossler system has low-instability, we also analyse it using methods for high-instability systems.

For the lower-instability case of a single Rossler time series we can estimate most lower order UPOs but not all periods up to $12T$; in particular period $7T$ is difficult

to detect. However, using multi-sampling we detect many complete cycles of each period and calculate excellent noise-free estimates of all UPOs with periods from $1T$ to $12T$, where T is the fundamental period. These compare well with those derived from noise-free multi-samples. For the lower-instability chaotic system we also calculate the short time Lyapunov exponents using the methodology laid out in Chapter 4. The noise-reduced versions are similar to the noise-free estimates and the weighted average estimates of the global maximal Lyapunov are similar to each other and the true result.

We defined a set of goodness of fit metrics applicable to time series to measure the effectiveness of the noise-reduced time series approximation. These include three distance measures (AME, RMSE, MME), a statistical measure (the coefficient of determination, R^2) and a phase lag measure. We then approximated the single 25% noise-infected Rossler time series using the SUNR method, noise-filtering all detected transformed cycles in the frequency domain after a FFT. The final approximation is obviously a significant improvement on the noisy time series and demonstrates results similar to a moving average; which is what we expect in this example. The method yields excellent results for higher levels of noise (up to 50%–100%) and all noise types tested (Uniform white noise, Gaussian white noise, high-frequency noise and pink noise).

The SUNR method was applied to the model Rossler system with dynamical noise added. For a single time series (10,000 points) the results were not strikingly good. Our research illustrates how dynamical noise deforms the shadow-UPOs but retains smoothness and occupies the same frequencies as the signal in the frequency domain, making it highly problematic. We subsequently tried multi-sampling and collected complete shadow-UPOs in bins. If we assume the dynamical noise to have a Gaussian distribution with mean zero, then the deformations should largely average out over sufficiently many samples (because the chaotic attractor is bounded in phase space). The detected shadow-UPOs were averaged and compared with the estimated noise-free UPOs obtained using the noise-free time series. We observed that the estimated noise-free UPOs were reasonably good approximations to the true UPOs, although not highly accurate. The dynamical noise based estimates

are sufficiently close to the true UPOs for qualitative use, for example classifying the chaotic attractor. Observing that dynamical noise appears to conserve cycles (up to a point) and does not destroy smoothness in the cycles, we experimented with directly using a dynamical noise-infected Rossler time series in the numerical algorithm by Wolf to estimate maximal Lyapunov exponents. The Wolf algorithm relies on smoothness of data and this is one reason it fails with relatively low levels of measurement noise. However this is not the case with smoother dynamical noise. Results did vary from one time series to another, so we used multiple samples (50) and averaged the result. Surprisingly we found the average determined by multi-sampling was a very good estimate of the noise-free result, with maximal Lyapunov estimates of $\lambda_{\max} = 0.7$ s^{-1} for the Rossler system, $\lambda_{\max} = 0.96$ s^{-1} for the Lorenz system and $\lambda_{\max} = 1.68$ s^{-1} for the highly chaotic Lu-Chen system. The exercise was repeated multiple times and tested on the Lorenz and Chua systems with excellent results also. Further research is required to understand the mechanism here, but this is a useful result.

We next tested the SUNR time series approximation on a single noise-infected time series of other chaotic systems of varying topology including the Chua, Lorenz, Rabinovich-Fabrikant and Lu-Chen systems. The SUNR time series approximation works well for the chaotic Rossler ($\lambda_{\max} = 0.07$ s^{-1}), Chua ($\lambda_{\max} = 0.28$ s^{-1}) and Lorenz systems ($\lambda_{\max} = 0.91$ s^{-1}), all of which have maximal Lyapunov exponents less than 1 s^{-1}. The SUNR method produced a noise-free time series approximation with significantly reduced levels of measurement noise on all the goodness of fit metrics. Estimation of maximal Lyapunov exponent using the noise-reduced time series approximation in the Wolf algorithm produced accurate estimates.

For highly unstable chaotic systems the recurrence-histogram cycle detection method simply did not detect enough cycles to construct a good approximation to the noise-free time series. No orbit remains on a single shadow-UPO long enough to provide useful information. This is evidenced by the Lu-Chen system which is highly chaotic, with maximal Lyapunov exponent at least 1.75 s^{-1}. The SUNR method simply failed to work for this system as we could not detect sufficient shadow-UPOs. Realising the limitations of the recurrence matrix method for

highly unstable systems, we tried a different approach, using local minima to partition the time series into "cycle-like" segments. For the Lu-Chen system, where we could not get a result at all using recurrence matrices, we did obtain a result but note the approach could not accomodate very high levels of noise (>25%).

The Rabinovich-Fabrikant system is an example of a system with a small maximal Lyapunov exponent (< 0.1 s^{-1}) which behaves as an extremely unstable system. We obtained stable goodness of fit results using the recurrence method, where most detected cycles were partial cycles and obtained consistent estimates of the maximal Lyapunov exponent for various levels of noise. We also tested the method of using local minima to detect cycles (not shadow-UPOs) and we obtained similar results to those obtained using recurrence matrices. Using local minima we find many complete cycles that do not immediately repeat and are thus not detected using the recurrence methodology.

Our testing demonstrated that the recurrence-histogram method for detecting shadow-UPOs works well for most chaotic systems, but not all chaotic systems and its usefulness depends on the topology of the attractor. The recurrence method of detecting shadow-UPOs has been utilised on a small number of chaotic systems in the literature (e.g. Henon map and Rossler system) where is works well and one could draw the conclusion it translates equally to other chaotic systems. It appears to work well for systems with low to moderate instability where the orbit travels several times in the vicinity of a UPO before shearing off onto another cycle. The exceptions occur when the system is so unstable that the orbit never travels near a specific cycle for very long, or when the orbit has a complicated evolution where cycles are traced out but not consecutively. The advantages of the recurrence-histogram method greatly outweigh any disadvantages and it should be a first option. We highlight that in these exceptional cases other techniques should be applied for cycle detection and noise-filtering purposes and that this is an area with more research is required.

Finally we conducted multi-sampling testing for each of the Rossler, Lorenz and Chua systems, obtaining a SUNR time series approximation of the noise-free time

series from each sample. We then applied the Wolf numerical Lyapunov exponent algorithm to each approximated time series, averaging the results. This averaging process provides a much more accurate estimate of the maximal Lyapunov exponent of the noise-free chaotic systems tested.

The SUNR method should not be viewed as a cure for all noise-infected time series. Rather, the method is targeted at reducing as much noise as possible from a medium to high noise-infected signal, to produce a result that is far more useful than the starting position for subsequent use. The resultant signal will still contain some noise, but at greatly reduced levels. This presents a superb opportunity for the application of the other, more sophisticated trajectory-based or locally projective techniques that can retain integrity and cleave further low levels of residual noise.

Whilst we have researched multi-sampling for a number of model systems and tested the efficacy of the SUNR method, it would be natural to now extend the method to experimental systems such as lasers. The value would be two-fold; we could assess the efficacy of the SUNR method on a physical system and secondly explore the various types of noise for such a system.

6.2 Limitations of SUNR Method

6.2.1 Chaotic System Factors

As with all numerical methods applied to experimental data, there are a number of limitations. The limitations of the SUNR method depend primarily on two factors, the degree of instability of the chaotic dynamical system and the effectiveness of the technique used to detect and locate shadow-UPOs under a layer of noise.

The SUNR Method is applicable to experimentally determined time series of data from low dimensional chaotic dynamical systems. Numerous chaotic systems in nature appear to be well represented by a low dimensional attractor so the method has practical applications. In an experimental context there are numerous

systems that lend themselves to this technique. An example is laser research for chaotic dynamical regimes, where multiple time series samples may be easily captured for analysis. Similarly, in the medical research context, multiple time series may be collected of heartbeat data.

The SUNR method requires that the chaotic system is not too highly unstable. Our experimentation indicates that the maximal Lyapunov exponent should be not much larger than 1. Fortunately this includes most common chaotic systems like the Rossler, Lorenz, Ueda, Duffing and Chua systems. For systems with extreme instability, the evolving orbit will not remain on any shadow-UPO for any significant period of time and we are thus unable to detect any meaningful cycle information for noise-reduction. Such systems, loosely speaking, are at the far end of the spectrum between chaotic and turbulent and special techniques will need devised to accommodate their dynamics for noise reduction.

Chaotic systems within the acceptable range of instability, may be partitioned in two groups—those with detectable complete lower order shadow-UPOs and those at the other end of the instability range where we detect some complete but mostly partial shadow-UPOs. Deciding which group describes the chaotic system under examination belongs to is a simple exercise. For the lower instability group, we may approximate a basis set of lower order noise-free UPOs and estimate their short time Lyapunov exponents and also approximate the noise-free time series. For the chaotic systems at the other end of the instability range, we are limited to an approximation of the noise-free time series, which may subsequently be used in numerical algorithms to estimate the invariants.

6.2.2 Shadow-UPO Detection Factors

The SUNR method is generally applicable to a broad range of measurement noise types (e.g. Gaussian, coloured). For best results it requires a knowledge of the type of noise and the quality of the noise reduction depends on the availability of specialist linear signal processing techniques applicable to the noise type. The type of noise can usually be characterised from the frequency domain power spectrum

and a knowledge of the physical and experimental system. Fortunately there are a great number of tailored noise-reduction techniques available, targeting specific types of noise.

The testing results indicate that the SUNR method can address measurement noise levels up to 100% of signal. The recurrence-histogram method can identify the chaotic system harmonics for very high levels of noise (we could still resolve histogram peaks at 200% measurement noise). The research indicates that the limiting factor is finding the start and end points of the complete and partial shadow-UPOs when noise is greater than 10%. The use of the moving average pre-filter increased the detection threshold from 10% noise to 50%–100% noise, but there is a limit. It is also applicable for dynamical noise at low levels where cycle sequences retain start to end point stability without significant phase shift.

For systems with sufficiently low instability to admit complete and partial cycles, the subsequent detection of shadow-UPOs using the recurrence-histogram depends on the combination of sampling frequency and critical radius. There is currently no optimal relationship between these two factors and some experimentation and trial and error is required. Visualisation of the recurrence plot and recurrence-histogram assists this process greatly. Further, the recurrence-histogram approach necessarily requires two consecutive circuits of a given cycle for it to be detected as a cycle. This excludes single cycles that repeat, but not immediately. For these cases, other techniques need to be developed that exploit the cyclic nature of chaotic systems.

This method does require user judgment and oversight and at present cannot be fully automated. It is a multi-step process that requires user input and judgment at each step. The payoff is a method that will provide a reasonable result where many other methods fail, and it is relatively simple and practical.

6.3 Future Research

Surprisingly, there has been little research into chaotic noise-reduction techniques using topological invariants such as UPOs. Given the significance of these in the architecture of chaotic attractors at a fundamental level, the opportunity exists to greatly expand research in this area. The geometry of chaotic attractors is often very complex and somewhat confronting. Fortunately, for most common chaotic systems this complexity can be reduced to a set of shadow-UPOs that are individually relatively simple to work with. UPOs are dense on the attractor and this alone necessitates that every point is very near a UPO, and commonly a lower order UPO in particular. This immediately opens doors and provides pathways to improved approximation techniques.

There is some scope for further research in optimising the recurrence-histogram cycle-detection approach. Essential to the process of separating signal from noise is locating and extracting lower order shadow-UPOs. It provides a very powerful probe to identify signals masked by significant measurement noise. The recurrence-histogram approach is a first priority as it strictly identifies all repeated sequences of points for all periods in the data and the periods corresponding to histogram peaks represent the most frequently occurring periodic points. The methodology captures cycles of the same period in conjugate pairs, both contained within a tube having radius equal to the critical radius; this is a strong condition. This information is useful in searching for cycle sequences, however a recurrence-histogram peak does not guarantee the existence of complete cycles of that period, but just a greater likelihood of detecting them. There is scope to modify this approach further for improved efficiency.

It would be interesting to determine the relationship between sampling frequency, critical radius and noise level analytically, searching for possible relationships, bounds and an optimisation rule. Several levers drive the effectiveness of the recurrence matrix detection method. Foremost is the critical radius, which defines a thin tube surrounding a given shadow-UPO, within which we consider any other cycle equivalent provided the entire cycle is encompassed within the tube. We

found that the minimum value of $\varepsilon \geq 5\sigma$ for the critical radius is only suitable for noise levels <10%. Our scenario modelling suggests setting the critical radius less than 5%–10% of the attractor extent. This is somewhat arbitrary; it is unlikely to be truly applicable for all systems.

The recurrence-matrix method is excellent for detecting system harmonics and locating shadow-UPOs, provided the data is pre-processed with a moving average as demonstrated in this thesis. This approach has been shown to be effective for most commonly studied chaotic systems (with maximal Lyapunov exponent <1 s^{-1}) with high levels of any many types of measurement noise. This approach has its limitations when systems are highly chaotic and almost turbulent. However, there are systems lying in between these instability regimes. The recurrence-histogram of some systems will not detect a usable number of shadow-UPOs, but many cycles are clearly present. There are chaotic systems that are sufficiently unstable so as to not traverse at least two circuits of a cycle, which is the minimum requirement for detection using the recurrence method. In all likelihood the orbit may just jump from cycle to cycle and at best complete single circuits or at worst complete only partial cycles. The recurrence histogram may fail to show clear well-defined peaks and in such a case, we need to seek other approaches to detect cycles. For the Rabinovich-Fabrikant system we observed this dynamic and utilised the single cycles in the time series data directly. We observed the presence of many complete single cycles, often recurring later, but never completing two circuits consecutively. Just as recurrence-histograms are robust in the presence of noise, so too are local maxima and minima in the time series, especially if a centered moving average pre-filter is applied first. We identified the cycles using local minima and the criterion that a cycle is completed at a subsequent local minima, provided the L_2 distance between the two minima is very small. We were careful to view the individual cycles and ensure that we were filtering proper cycles. This "work-around" provided good results; however there is an opportunity to more fully develop this search model or others that will accomodate this type of orbital instability.

Our objective is to locate and extract all cycles that are detectable. The frequency with which complete cycles appear depends both on the ergodicity and the

level of instability of the system. If the orbit does not traverse near a UPO very often, then our time series sample may not include it. If the system has high instability, we may not travel near UPOs long enough to detect them. In our research, we assume the instability of the chaotic system is fixed and that the exercise is one of detection efficiency. However, there are merits to exploring the use of invariant stability transforms to change the stability of the UPOs in the underlying system. These stabilising techniques have been used in Newton-Raphson type UPO detection algorithms for noise-free data [121][24]. The topological invariance of UPOs may allow such transformations to work in the presence of high levels of noise. A stabilised highly noise-infected time series will still be amenable to the SUNR method and such a process would shift the system down the instability spectrum, resulting in a far better approximation.

Partial shadow-UPO data may be more fully utilised. There are usually large numbers of partial cycles, where perhaps 50%–75% of the complete cycle of a given period is detected. For shadow-UPOs with periods having instability at the higher end of the acceptable range, often only partial cycles will be detected. These contain valuable information about the system and often cover a high percentage of the time series. Captured partial cycles that are sufficiently long, of periods stable enough for us to have already captured multiple samples, could be used in improving averages and constructing estimates of complete cycles. A collection of partial cycles, where no complete cycles have been detected is also useful as each partial cycle fulfils our stability criteria on its path. Methods could be developed to reconstruct complete cycles by blending the relevant sequences from different partial cycles and averaging. It would also be useful to develop criteria for the minimum percentage of a complete cycle for a partial cycle to be included, rather than the 50%–75% we use based on observing the future evolution of examples.

Dynamical noise is inherently problematic and few techniques exist to identify the underlying chaotic signal with even low levels of dynamic noise present. For the chaotic systems studied we observe that dynamical noise smoothly deforms cycles, but the end points remain remarkably well anchored and with very small phase shift. This is almost as if cycles are conserved (but deformed) under dy-

namical noise provided instability is less than a certain level. These observations provide a good starting point for deeper research in systems with instability that is not too great. The smoothness of the noise is a great asset when applying numerical algorithms that incrementally step forward in time in small steps. It would be useful to test dynamically noise-infected time-series in a wider range of conventional numerical algorithms to calculate invariants, using multi-sampling and averaging (as we did with the Rossler system) and explore the stability of the results. Also existing algorithms for invariants may be able to be modified slightly to accomodate the characteristics of dynamical noise.

A major research project that would be highly valued is a rigorous like-for-like numerical performance comparison of the available noise reduction techniques for chaotic systems. This would require testing a broad set of chaotic systems under a set of scenarios involving different types of noise, including dynamical noise, and using a wide range of noise levels. Many published research papers involve a single "standard" chaotic system and noise scenarios are confined to simple GWN. Computer code would need to be constructed directly from the research papers. This would be a significant undertaking, as Matlab code is generally not available for existing methods (many are from 20 years ago) and such a comparison would be highly useful to assess the relative merits of various techniques. There are no such comparisons in the majority of existing papers for these reasons. For these reasons we also did not compare the SUNR method with other techniques in a like-for-like comparison. It would be a useful exercise for a researcher to construct the code for all major methods (make it available), compare the methods and publish the results.

We have tested a range of model systems using several known types of measurement noise (UWN, GWN, HFN, Coloured noise). The ability to reduce known types of noise using existing signal processing noise reduction tools is a strength of the method. We appreciate that other types of noise exist and these need to be addressed in real world data. This is an area of focus for future research, where additional techniques may need to be developed.

Noise is closely intertwined with the signal in chaotic systems and over the years researchers have constructed many remarkable algorithms to separate the two. All the methods have their merits and pitfalls, as described in Chapter 2 and the field in still open to further developments. The literature shows there was an intense focus on local geometric projection noise reduction methods during the mid 1990s. Around this time UPO detection research shifted to more theoretical aspects and practical applications focused on controlling chaos. It appears the streams never converged and the field is still open for further developments. There is considerable room to advance the field by working in terms of the cycle approximation where the vast body of mathematics relating to cycles and that from signal processing is available and can be deployed.

Bibliography

[1] Abarbanel, H., and Abarbanel, H. D. I. The analysis of observed chaotic data in physical systems. *Reviews of Modern Physics 65*, 4 (1993), 1331–1392.

[2] Abarbanel, H. D., Brown, R., and Kennel, M. B. Local Lyapunov exponents computed from observed data. *Journal of Nonlinear Science 2*, 3 (1992), 343–365.

[3] Aguirre, L. A., and Letellier, C. Modeling nonlinear dynamics and chaos: a review. *Mathematical Problems in Engineering 2009* (2009).

[4] Anosov, D. V. *Geodesic flows on closed Riemann manifolds with negative curvature*. American Mathematical Society Providence, 1969.

[5] Argyris, J., Andreadis, I., Pavlos, G., and Athanasiou, M. The influence of noise on the correlation dimension of chaotic attractors. *Chaos, Solitons Fractals 9*, 3 (1998), 343–361.

[6] Auerbach, D., Cvitanović, P., Eckmann, J.-P., Gunaratne, G., and Procaccia, I. Exploring chaotic motion through periodic orbits. *Physical Review Letters 58*, 23 (1987), 2387–2389.

[7] Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., and Rubio, M. Dimension increase in filtered chaotic signals. *Physical Review Letters 60*, 11 (1988), 979.

[8] Banks, J., Brooks, J., Cairns, G., Davis, G., and Stacey, P. On Devaney's definition of chaos. *The American Mathematical Monthly 99*, 4 (1992), 332–334.

[9] Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. part 1: Theory. *Meccanica 15*, 1 (1980), 9–20.

[10] Berndt, D. J., and Clifford, J. Using dynamic time warping to find patterns in time series. In *KDD Workshop* (1994), vol. 10, Seattle, WA, pp. 359–370.

[11] Böttcher, F., Peinke, J., Kleinhans, D., Friedrich, R., Lind, P. G., and Haase, M. Reconstruction of complex dynamical systems affected by strong measurement noise. *Physical Review Letters 97*, 9 (2006), 090603.

[12] Bowen, R. ω-limit sets for axiom a diffeomorphisms. *Journal of Differential Equations 18*, 2 (1975), 333–339.

[13] Broomhead, D., Huke, J., and Muldoon, M. Linear filters and non-linear systems. *Journal of the Royal Statistical Society. Series B (Methodological)* (1992), 373–382.

[14] Broomhead, D., and King, G. P. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena 20*, 2 (1986), 217–236.

[15] Brown, R., Bryant, P., and Abarbanel, H. D. Computing the Lyapunov spectrum of a dynamical system from an observed time series. *Physical Review A 43*, 6 (1991), 2787.

[16] Carroll, T. Approximating chaotic time series through unstable periodic orbits. *Physical Review E 59*, 2 (1999), 1615.

[17] Casdagli, M. Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena 35*, 3 (1989), 335–356.

[18] Cawley, R., and Cawley, R. Local-geometric-projection method for noise reduction in chaotic maps and flows. *Physical review. A, Atomic, Molecular, and Optical Physics 46*, 6 (1992), 3057–3082.

[19] Cencini, M., Falcioni, M., Olbrich, E., Kantz, H., and Vulpiani, A. Chaos or noise: Difficulties of a distinction. *Physical Review E 62*, 1 (2000), 427.

[20] Chelidze, D. Smooth local subspace projection for nonlinear noise reduction. *Chaos: An Interdisciplinary Journal of Nonlinear Science 24*, 1 (2014), 013121.

[21] Chen, C.-T. *Linear system theory and design*. Oxford University Press, Inc., 1998.

[22] Chua, L. O., and Parker, T. S. Practical numerical algorithms for chaotic systems, 1989.

[23] Coban, G., Büyüklü, A. H., and Das, A. Linear least squares estimate of noise level in chaotic time series via l norm correlation sum.

[24] Crofts, J. J., and Davidchack, R. L. Efficient detection of periodic orbits in chaotic systems by stabilizing transformations. *SIAM Journal on Scientific Computing 28*, 4 (2006), 1275–1288.

[25] Crutchfield, J. Prediction and stability in classical mechanics. *Senior Thesis, University of California, Santa Cruz* (1979).

[26] Crutchfield, J., and Huberman, B. Fluctuations and the onset of chaos. *Physics Letters A 77*, 6 (1980), 407–410.

[27] Crutchfield, J. P., Farmer, J., and Huberman, B. A. Fluctuations and simple chaotic dynamics. *Physics Reports 92*, 2 (1982), 45–82.

[28] Cvitanović, P. Invariant measurement of strange sets in terms of cycles. *Physical Review Letters 61*, 24 (1988), 2729.

[29] Cvitanović, P. Periodic orbits as the skeleton of classical and quantum chaos. *Physica D: Nonlinear Phenomena 51*, 1 (1991), 138–151.

[30] Dallas, G. Wavelets for dummies: Signal processing, Fourier transforms and Heisenberg. 2014. *Disponıvel em:"lt; https://georgemdallas.*

wordpress. com/2014/05/14/wavelets-4-dummies-signal-processing-fourier-transforms-and-heisenberg (2017).

[31] Danca, M.-F., Feckan, M., Kuznetsov, N., and Chen, G. Looking more closely at the Rabinovich-Fabrikant system. *International Journal of Bifurcation and Chaos 26*, 02 (2016), 1650038.

[32] Danca, M.-F., Kuznetsov, N., and Chen, G. Unusual dynamics and hidden attractors of the Rabinovich-Fabrikant system. *Nonlinear Dynamics 88*, 1 (2017), 791–805.

[33] Davidchack, R. L., and Lai, Y.-C. Efficient algorithm for detecting unstable periodic orbits in chaotic systems. *Physical Review E 60*, 5 (1999), 6172.

[34] Davidchack, R. L., Lai, Y.-C., Klebanoff, A., and Bollt, E. M. Towards complete detection of unstable periodic orbits in chaotic systems. *Physics Letters A 287*, 1 (2001), 99–104.

[35] Davies, M. Noise reduction by gradient descent. *International Journal of Bifurcation and Chaos 3*, 01 (1993), 113–118.

[36] Davies, M. E., and Davies, M. E. Linear recursive filters and nonlinear dynamics. *Nonlinearity 9*, 2 (1996), 487–499.

[37] Deissler, R. J., and Farmer, J. D. Deterministic noise amplifiers. *Physica D: Nonlinear Phenomena 55*, 1 (1992), 155–165.

[38] Devaney, R. *An introduction to chaotic dynamical systems*. Westview press, 2008.

[39] Dumont, R. S., and Brumer, P. Characteristics of power spectra for regular and chaotic systems. *The Journal of Chemical Physics 88*, 3 (1988), 1481–1496.

[40] Eckhardt, B., and Yao, D. Local Lyapunov exponents in chaotic systems. *Physica D: Nonlinear Phenomena 65*, 1 (1993), 100–108.

[41] Eckmann, J.-P., Kamphorst, S. O., and Ruelle, D. Recurrence plots of dynamical systems. *EPL (Europhysics Letters) 4*, 9 (1987), 973.

[42] Eckmann, J.-P., Kamphorst, S. O., Ruelle, D., and Ciliberto, S. Liapunov exponents from time series. *Physical Review A 34*, 6 (1986), 4971–4979.

[43] Eckmann, J.-P., and Ruelle, D. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics 57*, 3 (1985), 617.

[44] Ellner, S., Gallant, A. R., McCaffrey, D., and Nychka, D. Convergence rates and data requirements for Jacobian-based estimates of lyapunov exponents from data. *Physics Letters A 153*, 6 (1991), 357–363.

[45] Ellner, S., and Turchin, P. Chaos in a noisy world: new methods and evidence from time-series analysis. *American Naturalist* (1995), 343–375.

[46] Esling, P., and Agon, C. Time-series data mining. *ACM Computing Surveys (CSUR) 45*, 1 (2012), 12.

[47] Farmer, J. D., and Sidorowich, J. J. Optimal shadowing and noise reduction. *Physica D: Nonlinear Phenomena 47*, 3 (1991), 373–392.

[48] Franca, L. F. P., and Savi, M. A. Distinguishing periodic and chaotic time series obtained from an experimental nonlinear pendulum. *Nonlinear Dynamics 26*, 3 (2001), 255–273.

[49] Fraser, A. M. Information and entropy in strange attractors. *Information Theory, IEEE Transactions on 35*, 2 (1989), 245–262.

[50] Gao, J., Chen, C., Hwang, S., and Liu, J. Noise-induced chaos. *International Journal of Modern Physics B 13*, 28 (1999), 3283–3305.

[51] Gao, J., Hu, J., Tung, W., and Cao, Y. Distinguishing chaos from noise by scale-dependent Lyapunov exponent. *Physical Review E 74*, 6 (2006), 066204.

[52] Gao, J., Hwang, S., and Liu, J. When can noise induce chaos? *Physical Review Letters 82*, 6 (1999), 1132.

[53] Garcin, M., and Guegan, D. Probability density of the wavelet coefficients of a noisy chaos.

[54] Garcin, M., and Guegan, D. Optimal wavelet shrinkage of a noisy dynamical system with non-linear noise impact.

[55] Geist, K., Parlitz, U., and Lauterborn, W. Comparison of different methods for computing Lyapunov exponents. *Progress of Theoretical Physics 83*, 5 (1990), 875–893.

[56] Gilmore, C. G. An examination of nonlinear dependence in exchange rates, using recent methods from chaos theory. *Global Finance Journal 12*, 1 (2001), 139–151.

[57] Gilmore, R. Topological analysis of chaotic dynamical systems. *Reviews of Modern Physics 70*, 4 (1998), 1455.

[58] Glass, L., and Mackey, M. C. *From clocks to chaos: the rhythms of life*. Princeton University Press, 1988.

[59] Gleick, J. *Chaos: Making a new science*. Open Road Media, 2011.

[60] Grassberger, P., Hegger, R., Kantz, H., Schaffrath, C., and Schreiber, T. On noise reduction methods for chaotic data. *Chaos: An Interdisciplinary Journal of Nonlinear Science 3*, 2 (1993), 127–141.

[61] Grebogi, C., Ott, E., and Yorke, J. A. Unstable periodic orbits and the dimension of chaotic attractors. *Physical Review A 36*, 7 (1987), 3522.

[62] Grebogi, C., Ott, E., and Yorke, J. A. Unstable periodic orbits and the dimensions of multifractal chaotic attractors. *Physical Review A 37*, 5 (1988), 1711.

[63] Hamilton, F., Berry, T., and Sauer, T. Ensemble Kalman filtering without a model. *Physical Review X 6*, 1 (2016), 011021.

[64] Hamilton, F., Berry, T., and Sauer, T. Kalman-Takens filtering in the presence of dynamical noise. *The European Physical Journal Special Topics 226*, 15 (2017), 3239–3250.

[65] Hammel, S. M. A noise reduction method for chaotic systems. *Physics Letters A 148*, 8 (1990), 421–428.

[66] Hegger, R., and Schreiber, T. A noise reduction method for multivariate time series. *Physics Letters A 170*, 4 (1992), 305–310.

[67] Holzfuss, J., and Kadtke, J. Global nonlinear noise reduction using radial basis functions. *International Journal of Bifurcation and Chaos 3*, 03 (1993), 589–596.

[68] Hu, J., Gao, J., and White, K. Estimating measurement noise in a time series by exploiting nonstationarity. *Chaos, Solitons & Fractals 22*, 4 (2004), 807–819.

[69] Hwang, S., Gao, J., and Liu, J. Noise-induced chaos in an optically injected semiconductor laser model. *Physical Review E 61*, 5 (2000), 5162.

[70] Iwanski, J. S., and Bradley, E. Recurrence plots of experimental data: To embed or not to embed? *Chaos: An Interdisciplinary Journal of Nonlinear Science 8*, 4 (1998), 861–871.

[71] Jafari, S., Golpayegani, S. H., and Jafari, A. A novel noise reduction method based on geometrical properties of continuous chaotic signals. *Scientia Iranica 19*, 6 (2012), 1837–1842.

[72] Kadtke, J. B., Brush, J., and Holzfuss, J. Global dynamical equations and Lyapunov exponents from noisy chaotic time series. *International Journal of Bifurcation and Chaos 3*, 03 (1993), 607–616.

[73] Kantz, H. A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics letters A 185*, 1 (1994), 77–87.

[74] Kantz, H., and Schreiber, T. *Nonlinear time series analysis*, vol. 7. Cambridge university press, 2004.

[75] Kantz, H., Schreiber, T., Hoffmann, I., Buzug, T., Pfister, G., Flepp, L., Simonet, J., Badii, R., and Brun, E. Nonlinear noise reduction: A case study on experimental data. *Physical Review E 48*, 2 (1993), 1529–1538.

[76] Kato, S., and Yamada, M. Unstable periodic solutions embedded in a shell model turbulence. *Physical Review E 68*, 2 (2003), 025302.

[77] Kawaguchi, A., Yonemoto, K., and Yanagawa, T. Estimating the correlation dimension from a chaotic system with dynamic noise. *Journal of the Japan Statistical Society 35*, 2 (2005), 287–302.

[78] Kawahara, G., and Kida, S. Periodic motion embedded in plane Couette turbulence: regeneration cycle and burst. *Journal of Fluid Mechanics 449* (2001), 291–300.

[79] Kazantsev, E. Unstable periodic orbits and attractor of the barotropic ocean model. *Nonlinear processes in Geophysics 5*, 4 (1998), 193–208.

[80] Kleinhans, D., Friedrich, R., Wächter, M., and Peinke, J. Markov properties in presence of measurement noise. *Physical Review E 76*, 4 (2007), 041109.

[81] Knudsen, C. Chaos without nonperiodicity. *The American Mathematical Monthly 101*, 6 (1994), 563–565.

[82] Kostelich, E. J. Problems in estimating dynamics from data. *Physica D: Nonlinear Phenomena 58*, 1 (1992), 138–152.

[83] Kostelich, E. J., and Schreiber, T. Noise reduction in chaotic time-series data: a survey of common methods. *Physical Review E 48*, 3 (1993), 1752.

[84] Kostelich, E. J., and Yorke, J. A. Noise reduction in dynamical systems. *Physical Review A 38*, 3 (1988), 1649.

[85] Kostelich, E. J., and Yorke, J. A. Noise reduction: Finding the simplest dynamical system consistent with the data. *Physica D: Nonlinear Phenomena 41*, 2 (1990), 183–196.

[86] Kruel, T.-M., Eiswirth, M., and Schneider, F. Computation of Lyapunov spectra: effect of interactive noise and application to a chemical oscillator. *Physica D: Nonlinear Phenomena 63*, 1 (1993), 117–137.

[87] Kuznetsov, N., and Leonov, G. Stability by the first approximation for discrete systems. *Vestnik - St Petersburg University Mathematics 36* (2003), 21–27.

[88] Lai, Y.-C., Grebogi, C., Yorke, J., and Kan, I. How often are chaotic saddles nonhyperbolic? *Nonlinearity 6*, 5 (1993), 779.

[89] Lathrop, D. P., and Kostelich, E. J. Characterization of an experimental strange attractor by periodic orbits. *Physical Review A 40*, 7 (1989), 4028–4031.

[90] Leonov, G. A., and Kuznetsov, N. V. Time-varying linearization and the perron effects. *International Journal of Bifurcation and Chaos 17*, 04 (2007), 1079–1107.

[91] Leontitsis, A., Bountis, T., and Pagge, J. An adaptive way for improving noise reduction using local geometric projection. *Chaos: An Interdisciplinary Journal of Nonlinear Science 14*, 1 (2004), 106–110.

[92] Li, T.-Y., and Yorke, J. A. Period three implies chaos. *American Mathematical Monthly* (1975), 985–992.

[93] Liebert, W., Pawelzik, K., and Schuster, H. Optimal embeddings of chaotic attractors from topological considerations. *EPL (Europhysics Letters) 14*, 6 (1991), 521.

[94] Lorenz, E. *Predictability: does the flap of a butterfly's wing in Brazil set off a tornado in Texas?* na, 1972.

[95] Lorenz, E. N. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences 20*, 2 (1963), 130–141.

[96] Lü, J., and Chen, G. A new chaotic attractor coined. *International Journal of Bifurcation and chaos 12*, 03 (2002), 659–661.

[97] Lü, J., Chen, G., and Zhang, S. The compound structure of a new chaotic attractor. *Chaos, Solitons and Fractals 14*, 5 (2002), 669–672.

[98] Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J. Recurrence plots for the analysis of complex systems. *Physics Reports 438*, 5 (2007), 237–329.

[99] Marwan, N., Romano, M. C., Thiel, M., and Kurths, J. Recurrence plots for the analysis of complex systems. *Physics Reports 438*, 5-6 (2007), 237–329.

[100] Matsumoto, T. A chaotic attractor from Chua's circuit. *IEEE Transactions on Circuits and Systems 31*, 12 (1984), 1055–1058.

[101] McKenzie, M. D. Chaotic behavior in national stock market indices: New evidence from the close returns test. *Global Finance Journal 12*, 1 (2001), 35–53.

[102] Mees, A. I. *Dynamics of feedback systems*. John Wiley & Sons, Inc., 1981.

[103] Mera, M., and Morán, M. Measurement noise reduction by local projection methods.

[104] Mindlin, G. M., and Gilmore, R. Topological analysis and synthesis of chaotic time series. *Physica D: Nonlinear Phenomena 58*, 1 (1992), 229–242.

[105] Mitschke, F., Möller, M., and Lange, W. Measuring filtered chaotic signals. *Physical Review A 37*, 11 (1988), 4518.

[106] Moore, J. M., Small, M., and Karrech, A. Improvements to local projective noise reduction through higher order and multiscale refinements. *Chaos: An Interdisciplinary Journal of Nonlinear Science 25*, 6 (2015), 063114.

[107] Mori, U., Mendiburu, A., and Lozano, J. A. Distance measures for time series in r: The tsdist package. *R J. 8* (2016), 451–459.

[108] Oseledec, V. I. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc 19*, 2 (1968), 197–231.

[109] Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. Geometry from a time series. *Physical Review Letters 45*, 9 (1980), 712–716.

[110] Paluš, M. *Chaotic measures and real-world systems*. Springer, 1998, pp. 49–66.

[111] Poincaré, H. Mémoire sur les courbes définies par une équation différentielle (ii). *Journal de Mathématiques Pures et Appliquées* (1882), 251–296.

[112] Rabinovich, M., and Fabrikant, A. Stochastic self-modulation of waves in nonequilibrium media. *Journal of Experimental and Theoretical Physics 77* (1979), 617–629.

[113] Rempel, E. L., and Chian, A. C.-L. Space plasma dynamics: Alfvén intermittent chaos. *Advances in Space Research 35*, 5 (2005), 951–960.

[114] Rosenstein, M. T., Collins, J. J., and De Luca, C. J. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena 65*, 1 (1993), 117–134.

[115] Rössler, O. E. An equation for continuous chaos. *Physics Letters A 57*, 5 (1976), 397–398.

[116] Saiki, Y. Numerical detection of unstable periodic orbits in continuous-time dynamical systems with chaotic behaviors. *Nonlinear Processes in Geophysics 14*, 5 (2007), 615–620.

[117] Sano, M., and Sawada, Y. Measurement of the Lyapunov spectrum from a chaotic time series. *Physical Review Letters 55*, 10 (1985), 1082.

[118] Sauer, T. A noise reduction method for signals from nonlinear systems. *Physica D: Nonlinear Phenomena 58*, 1 (1992), 193–201.

[119] Savi, M., and Pacheco, P. Chaos and hyperchaos in shape memory systems. *International Journal of Bifurcation and Chaos 12*, 03 (2002), 645–657.

[120] Sayers, C. L. Statistical inference based upon non-linear science. *European Economic Review 35*, 2 (1991), 306–312.

[121] Schmelcher, P., and Diakonos, F. Detecting unstable periodic orbits of chaotic dynamical systems. *Physical Review Letters 78*, 25 (1997), 4733.

[122] Schmelcher, P., and Diakonos, F. General approach to the localization of unstable periodic orbits in chaotic dynamical systems. *Physical Review E 57*, 3 (1998), 2739.

[123] Schouten, J. C., Takens, F., and van den Bleek, C. M. Estimation of the dimension of a noisy attractor. *Physical Review E 50*, 3 (1994), 1851.

[124] Schreiber, T. Extremely simple nonlinear noise-reduction method. *Physical Review E 47* (1993), 2401–2404.

[125] Schreiber, T., and Grassberger, P. A simple noise-reduction method for real data. *Physics letters A 160*, 5 (1991), 411–418.

[126] Schreiber, T., and Kantz, H. Noise in chaotic data: diagnosis and treatment. *Chaos: An Interdisciplinary Journal of Nonlinear Science 5*, 1 (1995), 133–142.

[127] Shimada, I., and Nagashima, T. A numerical approach to ergodic problem of dissipative dynamical systems. *Progress of Theoretical Physics 61*, 6 (1979), 1605–1616.

[128] Shin, K., Hammond, J., and White, P. Iterative SVD method for noise reduction of low-dimensional chaotic time series. *Mechanical Systems and Signal Processing 13*, 1 (1999), 115–124.

[129] Siefert, M., Kittel, A., Friedrich, R., and Peinke, J. On a quantitative method to analyze dynamical and measurement noise. *EPL (Europhysics Letters) 61*, 4 (2003), 466.

[130] Smirnov, D. A., Vlaskin, V. S., and Ponomarenko, V. I. Estimation of parameters in one-dimensional maps from noisy chaotic time series. *Physics Letters A 336*, 6 (2005), 448–458.

[131] Smith, S. W., et al. The scientist and engineer's guide to digital signal processing.

[132] So, P., Ott, E., Sauer, T., Gluckman, B. J., Grebogi, C., and Schiff, S. J. Extracting unstable periodic orbits from chaotic time series data. *Physical Review E 55*, 5 (1997), 5398.

[133] So, P., Schi, S., Ott, E., Kaplan, D. T., Sauer, T., and Grebogi, C. Detecting unstable periodic orbits in chaotic experimental data. *Physical Review Letters 76*, 25 (1996), 4705–4708.

[134] Stark, J., Broomhead, D., Davies, M., and Huke, J. Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis: Theory, Methods Applications 30*, 8 (1997), 5303–5314.

[135] Stark, J., Broomhead, D. S., Davies, M., and Huke, J. Delay embeddings for forced systems. ii. stochastic forcing. *Journal of Nonlinear Science 13*, 6 (2003), 519–577.

[136] Stewart, I. *Does God play dice?: The new mathematics of chaos*. Penguin UK, 1997.

[137] Takens, F. *Detecting strange attractors in turbulence*. Springer, 1981, pp. 366–381.

[138] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena 58*, 1 (1992), 77–94.

[139] Thiel, M., Romano, M. C., Kurths, J., Meucci, R., Allaria, E., Arecchi, F. T., and Agarwal, A. Influence of observational noise on the recurrence quantification analysis. *Physica D: Nonlinear Phenomena 171*, 3 (2002), 138–152.

[140] Touhey, P. Yet another definition of chaos. *The American Mathematical Monthly 104*, 5 (1997), 411–414.

[141] Trulla, L., Giuliani, A., Zbilut, J., and Webber Jr, C. Recurrence quantification analysis of the logistic equation with transients. *Physics Letters A 223*, 4 (1996), 255–260.

[142] Vellekoop, M., and Berglund, R. On intervals, transitivity= chaos. *The American Mathematical Monthly 101*, 4 (1994), 353–355.

[143] Webber Jr, C. L., and Zbilut, J. P. Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences* (2005), 26–94.

[144] Wolf, A. Quantifying chaos with Lyapunov exponents. *Chaos* (1986), 273–290.

[145] Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena 16*, 3 (1985), 285–317.

[146] Wu, X., and Kapral, R. Effects of molecular fluctuations on chemical oscillations and chaos. *The Journal of Chemical Physics 100*, 8 (1994), 5936–5948.

[147] Yanagita, T., and Iba, Y. Exploration of order in chaos using the replica exchange Monte Carlo method. *Journal of Statistical Mechanics: Theory and Experiment 2009*, 02 (2009), P02043.

[148] Yonemoto, K., and Yanagawa, T. Estimating the Lyapunov exponent from chaotic time series with dynamic noise. *Statistical Methodology 4*, 4 (2007), 461–480.

[149] Zhang, D., Györgyi, L., and Peltier, W. R. Deterministic chaos in the Belousov-Zhabotinsky reaction: Experiments and simulations. *Chaos: An Interdisciplinary Journal of Nonlinear Science 3*, 4 (1993), 723–745.