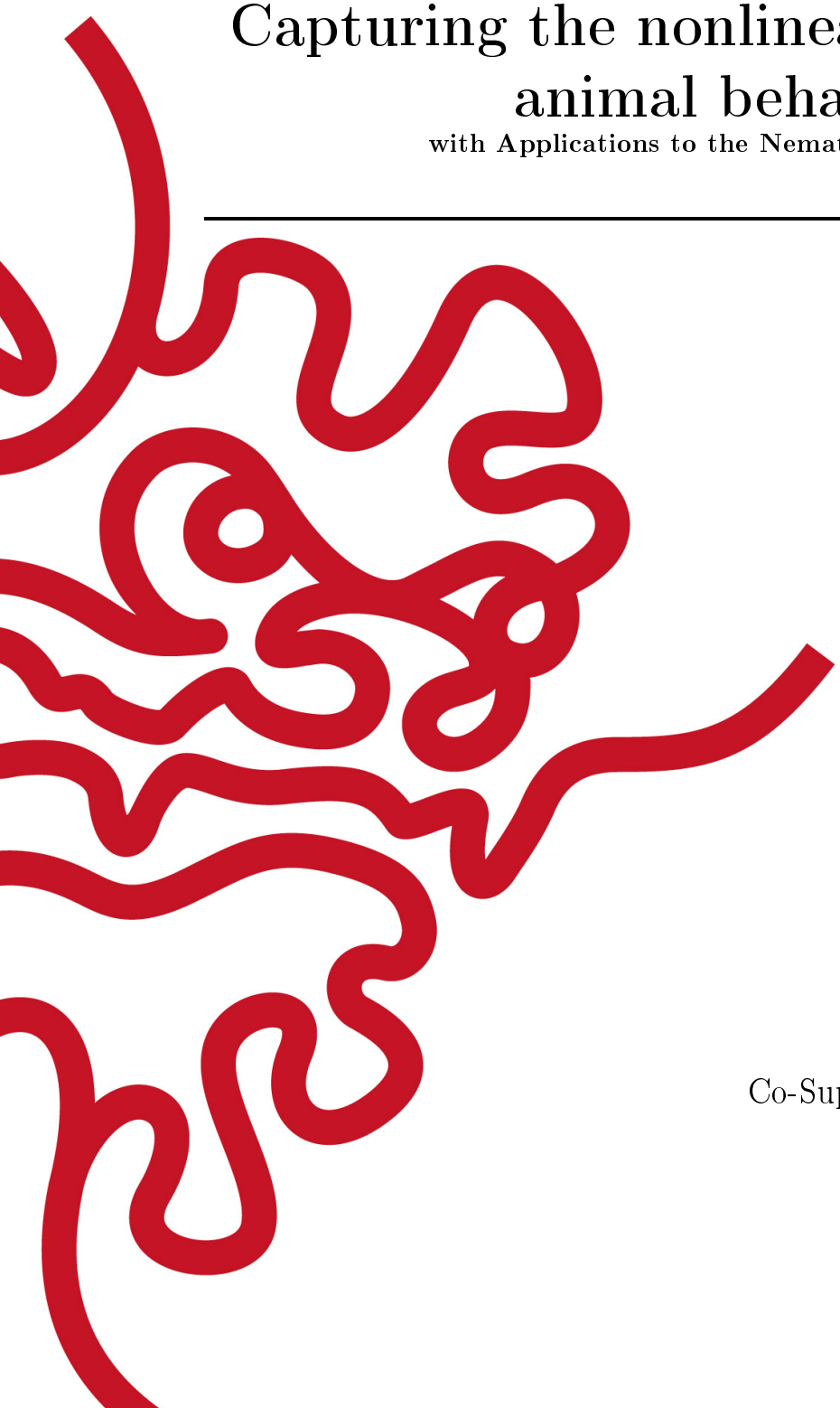
Okinawa Institute of Science and Technology
Graduate University

Thesis submitted for the degree

Doctor of Philosophy

Capturing the nonlinear dynamics of animal behavior

with Applications to the Nematode *C. elegans*



by

Tosif Ahamed

Supervisor: **Greg Stephens**
Co-Supervisor: **Ichiro Maruyama**

December, 2018

Declaration of Original and Sole Authorship

I, Tosif Ahamed, declare that this thesis entitled *Capturing the nonlinear dynamics of animal behavior*
with Applications to the Nematode C. elegans and the data presented in it are original and my own work.

I confirm that:

- No part of this work has previously been submitted for a degree at this or any other university.
- References to the work of others have been clearly acknowledged. Quotations from the work of others have been clearly indicated, and attributed to them.
- In cases where others have contributed to part of this work, such contribution has been clearly acknowledged and distinguished from my own work.
- None of this work has been previously published elsewhere, with the exception of the following: (provide list of publications or presentations, or delete this part). (If the work of any co-authors appears in this thesis, authorization such as a release or signed waiver from all affected co-authors must be obtained prior to publishing the thesis. If so, attach copies of this authorization to your initial and final submitted versions, as a separate document for retention by the Graduate School, and indicate on this page that such authorization has been obtained).

Date: December, 2018
Signature:

Abstract

Capturing the nonlinear dynamics of animal behavior
with Applications to the Nematode *C. elegans*

From microorganisms to humans, animals behave by making complex changes in their shape and posture over time with remarkable flexibility. To deal with the complexity of animal behavior existing analysis methods view it as a discrete time process, which is composed of transitions between a finite number of stereotyped motifs, such as walking or reaching. This viewpoint, however, ignores the fact that most behavior is not stereotyped. There is, therefore, a need for a perspective that captures the continuous complexity of animal behavior and offers detailed insights into general principles underlying its generation and control.

In my Ph.D. thesis, I propose a new approach of analyzing animal behavior, based on the idea that it is fundamentally a continuous time spatiotemporal dynamical system. I develop methods to transform behavioral recordings into a geometrical object called the "behavioral state space". As an organism moves, the corresponding behavioral state traces out a continuous trajectory in the state space, such that the geometry and topology of the trajectories encode quantitative and qualitative properties of behavior. Finally, I characterize an organism's behavioral dynamics in terms of the topological invariants estimated from the local Jacobians of the state space trajectories. The invariants capture essential aspects of a dynamical system, such as the number of degrees of freedom, symmetries in the governing equations of motion, and measures of predictability and variability.

I use the tools and concepts developed the above to perform a detailed characterization of continuous dynamics of freely behaving *C. elegans* worms.

Acknowledgment

I would first like to thank Prof. Ichiro Maruyama and his lab at OIST for not only hosting but teaching a physicist important lessons in Biology. I would especially like to thank fellow students Kazuto Kawamura, and Viktoras Liscicovas,who have taught me most of what I know about C. elegans. My fascination with the worm owes a great deal to these guys.

This work, would not have been possible without the understanding and patience of my primary supervisor, mentor, and advisor Prof. Greg Stephens, who has tolerated a stubborn, off the beaten track graduate student with great skill. I would like to thank him for having confidence in me when I did not. Most importantly, he is to blame for setting me down the path of nonlinear dynamical systems.

My thanks as well to the OIST institution. There are not many places around where a work, as interdisciplinary as this could have taken place. May it prosper and continue to bring misfits together.

Finally, special thanks to my family who have always believed in me, and to my wife Yoriko Yamamura for correcting me every time I was wrong and did not make much sense.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

Animals move and behave in a wide variety of ways - they run, crawl, fly and swim. The complex, nonlinear postural dynamics generating these behaviors evolve over multiple spatiotemporal scales and show a conflicting coexistence of regularity and variability. At large scales, behavior is structured, organized into stereotyped motifs such as walking or running, but the small scale dynamics within each motif are usually highly variable. Additionally, behavioral dynamics also show a large degree of intermittency as a result of sudden transients or jumps between different stereotyped motifs. Despite this complexity, animals have fine control over their movements. They can make rapid maneuvers to escape a predator, or to catch prey, and they can easily adapt to any new situation. This remarkable flexibility and efficiency in response to unpredictable environmental perturbations is rarely observed in robots and other man-made control systems, reflecting our lack of understanding of mechanisms underlying generation and control of animal behavior. Consequently, understanding animal behavior has the potential to drive future progress in several related fields, such as neuroscience [1], ethology [2, 3], physics of complex systems [4], control theory, and robotics and artificial intelligence [5].

Recent technological advances in the ability to gather high-resolution movement data make it possible for us to ask fundamental questions about how animals generate and control complex behaviors [3, 6]. However, the complexity of behavioral dynamics has presented unique challenges in quantification and analysis of these datasets. The main challenge is to find a quantitative representation of behavior that maps high-resolution recordings of animal behavior into a small set of interpretable numbers without losing any information about the dynamics. Formally, any representation that assigns numbers to instantaneous behavior of an animal must satisfy the following conditions. First, it must not map distinct behaviors to the same set of numbers, in other words, it must be one-to-one, or as close to it as possible. Additionally, in a good representation, the numerical difference between two sets of numbers should to some degree reflect how similar or dissimilar two behaviors are. Finally, a good representation must capture the dynamics of animal behavior, which is to say that it must allow us to predict what the animal will do in the near future. In addition to the formal requirements mentioned above, to be practical, a representation of animal behavior

should be low dimensional and interpretable. Thus, each behavior should map to a small set of numbers, which together provide insight into the biophysical principles of how an animal generates and controls its behavior.

Clearly, when available, a representation that satisfies all the conditions above can be invaluable in the study of animal behavior and its control. However, the conditions mentioned above do not uniquely specify a quantitative behavioral representation. Indeed, for most complex dynamical systems there can be several equivalent representations, such as the state space (or phase space), normal modes, or the collection of periodic orbits. Given that there is probably no unique way of representing animal behavior, the search for the best representation might seem like a pointless exercise. Fortunately, there are a set of measures that are independent of the particular representation used to describe a dynamical system. They capture only the essential, most universal aspects of the dynamics. Examples of these so-called dynamical or topological invariants include Lyapunov exponents, Kolmogorov-Sinai (or metric) entropy, and information dimension. These invariants pack a tremendous amount of information about qualitative and quantitative aspects of dynamical systems, such as the number of degrees of freedom, the symmetries of underlying dynamics, measures of predictability and variability and so on. Because of their universal properties, seeking these invariants from observations of behaving animals can lead us to general principles of how animal behavior is organized and controlled.

In this thesis, I have taken a two-pronged approach. Unable to give up entirely on the advantages of a good representation when it can be found. I first use ideas from state space reconstruction and embedding theory in nonlinear dynamics, see for example [7] to develop an interpretable and dynamic representation for *C. elegans* behavior. Second, I estimate dynamical invariants from data of freely moving worms. Both these approaches together provide unprecedented insights into the organization and control of the behavior of freely moving worms. This task could not be done directly using existing tools for state space reconstruction. Consequently, along the way, I have also made contributions to the field of nonlinear time-series analysis by developing tools and techniques that allow for robust estimation of an embedding and of the dynamical invariants. Specifically, the main contributions of this thesis are as follows

1. Estimation of Optimal Embedding Parameters

    (a) In chapter 2, I lay out the key ideas behind state space reconstruction followed by a discussion of popular approaches and their drawbacks. Subsequently, **I describe a new method of reconstructing the state space from time-series data based on maximizing predictability**, and apply it to known systems.

2. Robust Estimation of Lyapunov Exponents and Local Jacobians in State Space.

    (a) In chapter 3 I detail the theory behind Lyapunov exponents and describe common methods used to estimate them. I end by describing **a new method for robustly estimating local Jacobians in a state-space embedding**, which is demonstrated on known model systems.

3. A new, dynamic behavioral representation of *C. elegans* behavior

   (a) In chapter 4 I apply the methods developed in chapter 2 to reconstruct the state space for *C. elegans* behavior from behavioral recordings. **This is a new method of representing and measuring *C. elegans* behavior that captures the nonlinear continuous dynamics in a principled, model independent manner**. I then use the representation for measuring *C. elegans* behavior at multiple scales, from instantaneous posture dynamics to long time-scale foraging dynamics.

4. Estimation of the all the Lyapunov exponents and related invariants of freely moving worms

   (a) In chapter 5 I estimate the entire spectrum of Lyapunov exponents of freely moving worms and estimate other related dynamical invariants. To the best of my knowledge, **this is the first for the first attempt at estimating the Lyapunov exponents for an entire animal**.

In addition to above, I end with preliminary results suggesting further use of local Lyapunov exponents in understanding the control system underlying worm behavior. I give evidence for and put forward a new hypothesis regarding how the worm maintains a steady locomotory state by using an energy based feedback.

Before beginning I will briefly describe some model dynamical systems used in this thesis.

1.2 Example Dynamical Systems Used

Throughout the thesis I will use the following dynamical systems to clarify concepts and as a testbed for the methods to be developed. All simulations are performed with Matlab's ode45 adaptive Runge-Kutta solver with a time-step of 0.001s and error tolerances of 10^{-6}. When observation noise is added to simulations, it is added at 2 percent the standard deviation of the time-series.

1.2.1 Lorenz

First is the following canonical model of chaotic dynamics, first proposed by Ed Lorenz in 1963 [8] as a simple model of atmospheric convection rolls.

$$\begin{aligned} \dot{s_1} &= 10(s_2 - s_1) \\ \dot{s_2} &= s_1(28 - s_3) - s_2 \\ \dot{s_3} &= s_1 s_2 - \frac{8}{3} s_3 \end{aligned} \tag{1.1}$$

The state space is three dimensional, $\boldsymbol{s} = (s_1, s_2, s_3)$, and the evolution is described by three coupled ordinary differential equations with two quadratic nonlinearities. Fig. 1.1A illustrates the simulated trajectories of this system.

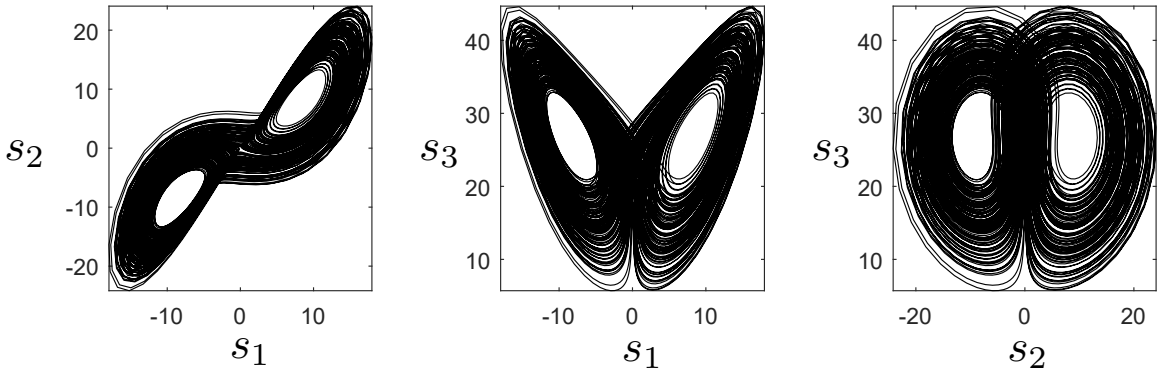


Figure 1.1: A. The phase portrait of the Lorenz system described by eqn 1.1.

1.2.2 Linear Coupled Oscillators and Normal Modes

Coupled oscillators play an important role in Neuroscience, whether it's central pattern generators in locomotion, or neural rhythm generators in other cognitive and memory related processes. To explore how state space reconstruction methods work with data obtained from coupled oscillators I will use the system consisting of two particles coupled together by three springs as shown in Fig. 1.2. The equations of motion of the system can be obtained by applying Newton's laws to give the following coupled second order differential equations:

$$\begin{aligned} m\ddot{q_1} &= -2kq_1 + kq_2 \\ m\ddot{q_2} &= kq_1 - 2kq_2 \end{aligned} \tag{1.2}$$

Where q_1, q_2 are the displacements of the two particles from their equilibrium positions, while k,and m refer to spring constant and the masses of particles respectively. Both these parameters will be set to unity, i.e. $k = 1, m = 1$ for the rest of the thesis.

The coupled mass spring system above makes a four dimensional dynamical system. We can represent the two second order differential equations in eqn 1.2, as four coupled first order differential equations as follows

$$\begin{aligned} \dot{q_1} &= \frac{p_1}{m} \\ \dot{p_1} &= -2kq_1 + kq_2 \\ \dot{q_2} &= \frac{p_1}{m} \\ \dot{p_2} &= kq_1 - 2kq_2 \end{aligned} \tag{1.3}$$

Where, p_1, p_2 are the momenta of the particles corresponding to q_1, q_2. The phase space of the system in (q_1, p_1) and (q_2, p_2) coordinates is shown in Fig. 1.3A along with the corresponding power spectra on the right. As can be seen, both springs oscillate with a mixture of frequencies $\omega_1 = 1$Hz and $\omega_2 = \sqrt{3}$Hz. Even though the system in eqn 1.3 is linear, the mixture of frequencies makes the phase portraits, and time-series look complex.

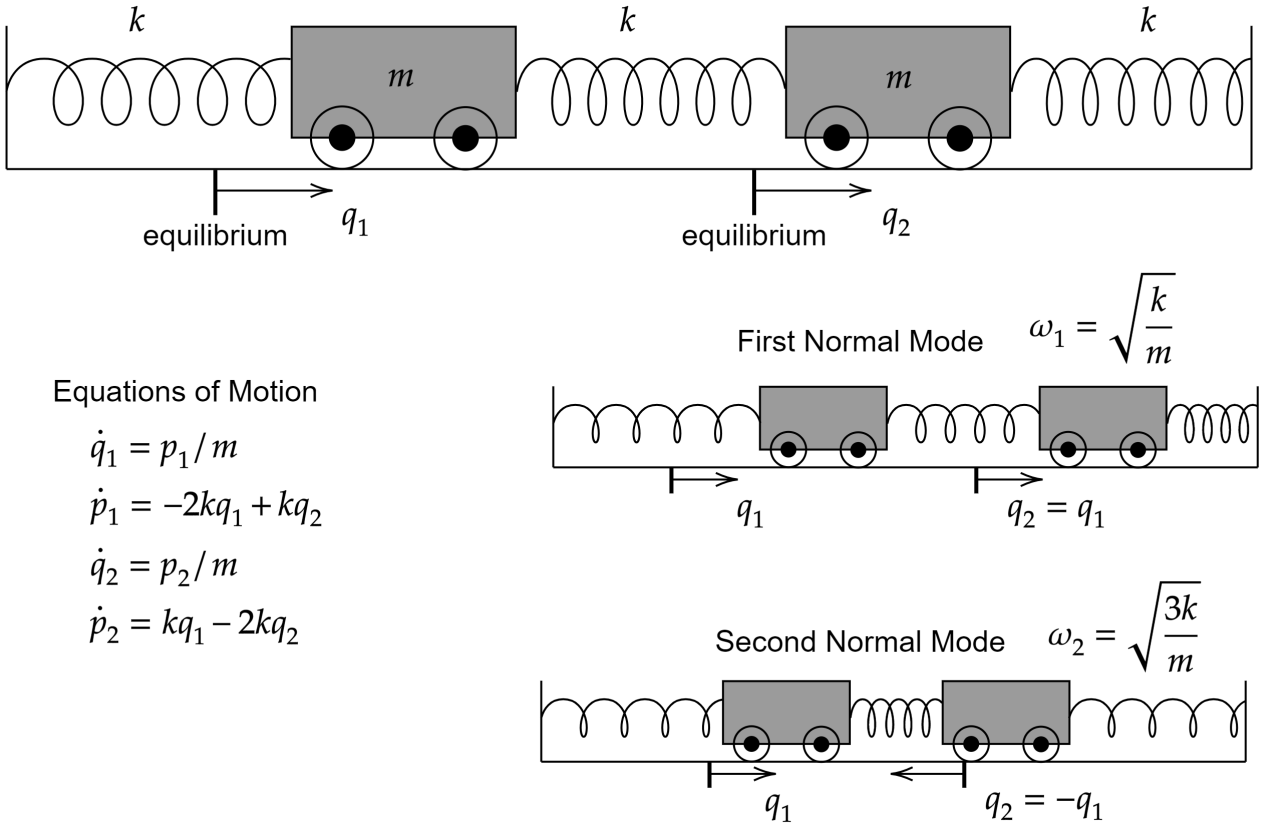


Figure 1.2: Illustration of the spring mass system and it's two normal modes.

Fortunately, for coupled linear systems, there is an alternate coordinate system known as the normal mode coordinates in which the oscillations are decoupled, and system behaves as a system of uncoupled oscillators. To see this for the spring mass system above, we can add and subtract the two equations in eqn 1.2 to give

$$\begin{aligned} m(\ddot{q}_1 + \ddot{q}_2) &= -k(q_1 + q_2) \\ m(\ddot{q}_1 - \ddot{q}_2) &= -3k(q_1 - q_2) \end{aligned} \tag{1.4}$$

If we now relabel the variables as $n_1 = q_1 + q_2$, and $n_2 = q_1 - q_2$, the above equations simplify to

$$\begin{aligned} m\ddot{n_1} &= -kn_1 \\ m\ddot{n_2} &= -3kn_2 \end{aligned} \tag{1.5}$$

The two equations above described by the normal modes n_1 and n_2 are *uncoupled*. Each normal mode behaves like a spring oscillating at a single frequency: n_1 at $\omega_1 = \sqrt{k/m}$ and n_2 at $\omega_2 = \sqrt{3k/m}$. It is in this sense that we can say that the normal mode coordinates behave like the coordinates of two uncoupled oscillators. By transforming to the normal mode coordinates, we have *uncoupled* the oscillations.

Admittedly, the normal mode coordinates are an abstraction of the physical coordinates (q_1, q_2). However, when a linear system is expressed in these coordinate, the motion simplifies considerably. For the spring mass system, the first normal mode n_1, corresponds to both springs moving in unison oscillating with $\omega_1 = 1$Hz. While, in the second mode n_2, springs oscillate exactly out of phase with each other at a frequency of $\omega_2 = \sqrt{3}$Hz. These modes are illustrated pictorially in 1.2. Any motion of the original coupled system can be expressed as a superposition of the normal modes n_1, n_2.

The phase portraits in the normal mode coordinate system are shown in Fig. 1.3B along with the power spectra. As can be seen that the same system as Fig. 1.3A, looks much simpler. Each mode has a single peak in the power spectrum and the phase portrait is a circle, corresponding to a regular oscillation.

1.2.3 Nonlinear Coupled Oscillators and Normal Modes

In the nonlinear version of coupled oscillators in shown in Fig. 1.2, we replace the springs by nonlinear springs. Specifically, instead of Hookean springs with a potential equal to $V(x) = \frac{1}{2}kx^2$ which cause a linear restoring force of $F = -kx$. I will use the Duffing potential equal to $V(x) = \frac{1}{2}ax^2 + \frac{1}{2}bx^4$, which leads to a nonlinear restoring force of $F = -ax - bx^3$ on the springs.

In addition to the spring restoring force, we can make the system more realistic by adding a small viscous damping and a periodic driving. The equations of motion for the damped driven, nonlinear spring oscillator system with the same configuration as in Fig. 1.2 and with the masses set to 1, is as follows. In this thesis I have set $a = 1$ and $b = -1$.

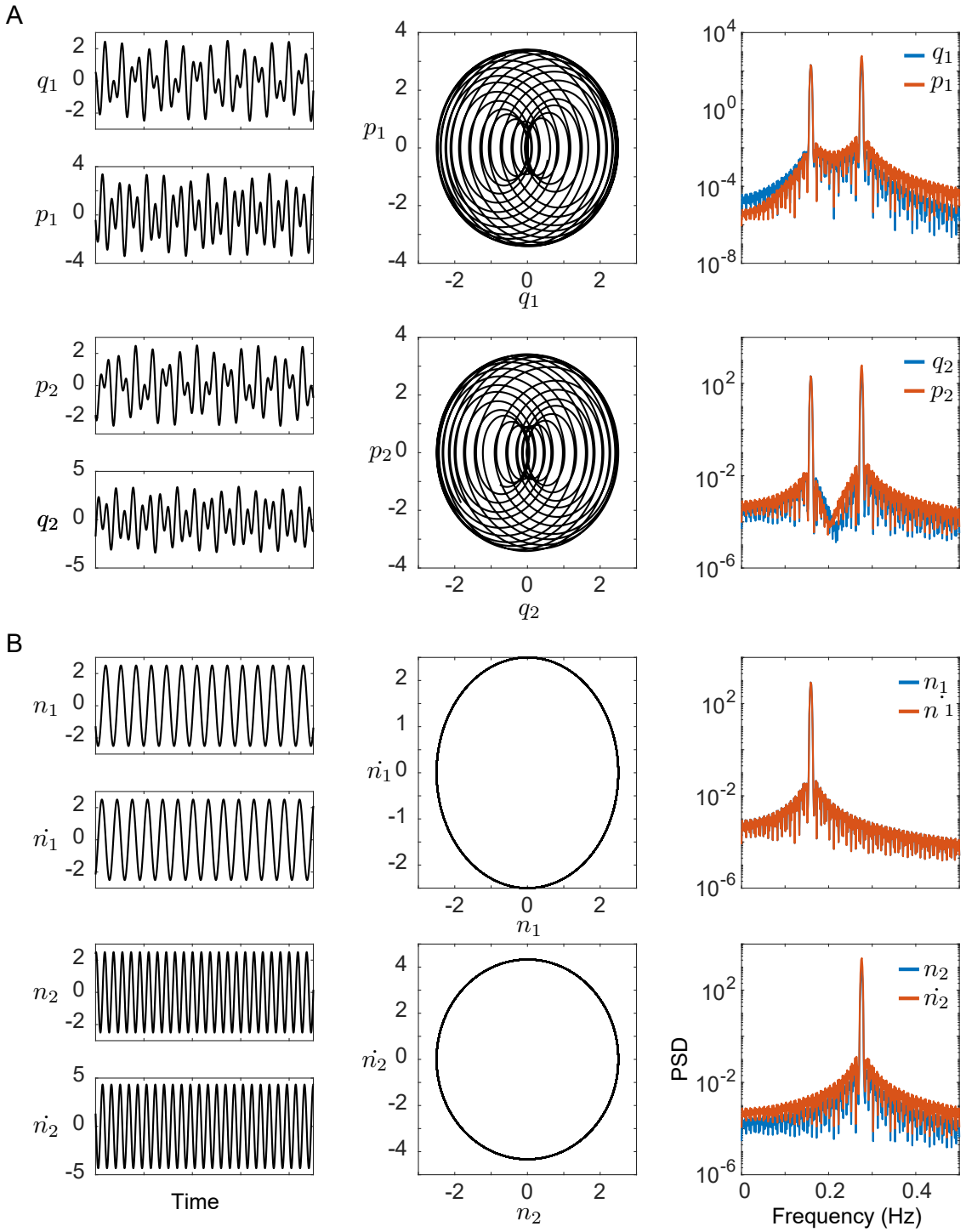


Figure 1.3: Comparison of real space coordinate system (**A.**), and the normal mode coordinate system (**B.**). Left column shows the time traces of the coordinates, middle column shows the trajectories in the phase space and right column shows the power spectral density (PSD) estimate of the variables. **A.** The behavior of the system in the original coordinates is significantly more complex, as a result of each mass oscillating with a combination of two frequencies as shown in the PSD on the right. **B.** Transformation to normal mode coordinates simplifies the motion considerably, with the system appearing as two uncoupled oscillators with different frequencies.

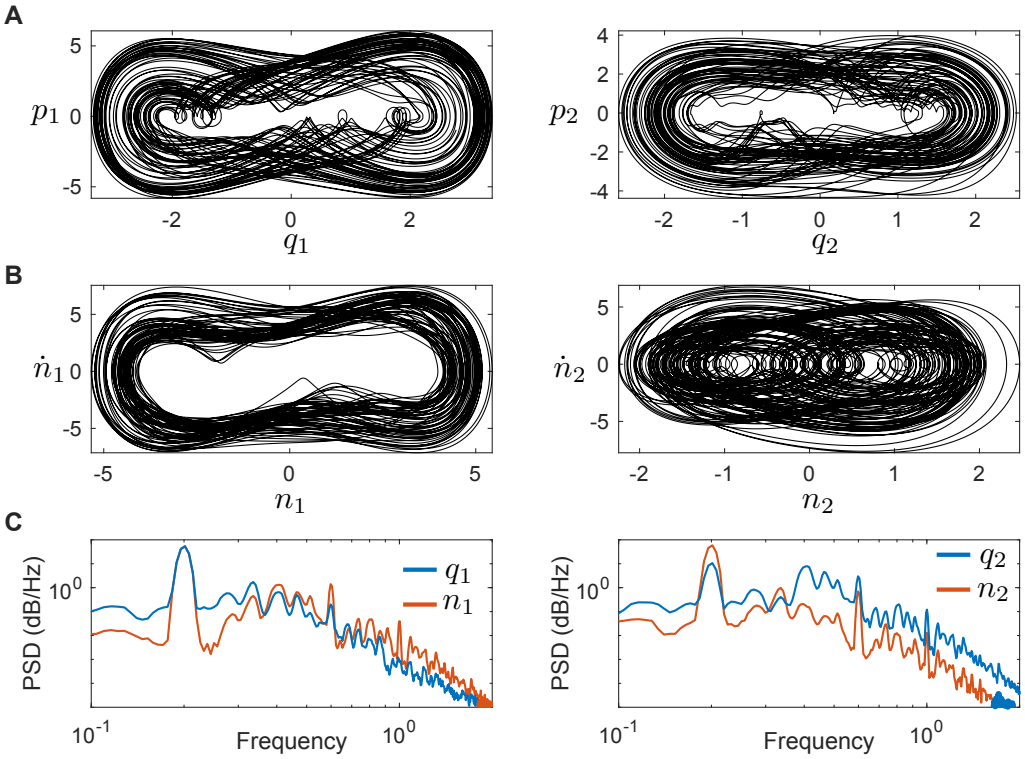


Figure 1.4: **A.** Phase portraits in the original coordinates of the duffing oscillator system look chaotic. **B.** A transformation to normal mode coordinates doesn't simplify the system significantly. **C.** The PSD estimates of original normal mode coordinate system also show that transformation to normal mode coordinates is not very helpful in simplifying the dynamics.

$$
\begin{aligned}
\dot{q_1} &= p_1 \\
\dot{p_1} &= -dp_1 - 2aq_1 + aq_2 - bq_1^3 + b(q_2 - q_1)^3 + fq_3 \\
\dot{q_2} &= p_1 \\
\dot{p_2} &= -dp_2 + aq_1 - 2aq_2 - bq_2^3 - b(q_2 - q_1)^3 \\
\dot{q_3} &= p_3 \\
\dot{p_3} &= -kq_3
\end{aligned}
\tag{1.6}
$$

The system in eqn 1.6 is ostensibly a 6D system as (q_3, p_3) make a linear spring with a spring constant of k which drives the spring q_1 with an amplitude of f. In addition, d is the viscous damping due to the environment. For the simulations in this thesis, the spring parameters are arbitrarily set to $a = -1, b = 1$, $d = 0.5$ and $f = 7.3$. The spring constant for the driving spring is set so that its oscillation frequency is equal to $0.4Hz$.

The phase portraits of the coupled oscillator system is shown in Fig. 1.4A. I also plotted the projections in the normal mode coordinate system in Fig. 1.4B. It can be seen that in this case the normal modes are not necessarily a good coordinate system

as they are not able to disassociate the two frequencies.

In the next chapter I will set up the main problem of this thesis, which is estimate the state space, from noisy measurements of a given dynamical system.

Chapter 2

Phase Space Reconstruction: Theory and Practice

2.1 Dynamical Systems and Topological Equivalence

A dynamical system in continuous time is formally represented as

$$\frac{d\boldsymbol{s}}{dt} = \boldsymbol{F}(\boldsymbol{s}) \tag{2.1}$$

while in discrete time it is

$$\boldsymbol{s}(t+1) = \boldsymbol{F}(\boldsymbol{s}(t)) \tag{2.2}$$

where, $\boldsymbol{s} = (s_1, s_2, \dots)$, called the state of the system evolves in time according to $\boldsymbol{F}$. Thus, the state variable encapsulates all the information that is necessary to determine the future of the system. In dynamical systems theory, $\boldsymbol{s}$ is usually thought of as a point in a d-dimensional *state space* S, while $\boldsymbol{F}(\boldsymbol{s})$ is generally a nonlinear vector field in S that determines how $\boldsymbol{s}$ changes in time.

Given an initial condition $\boldsymbol{s}(0)$, the solution of eqn 2.1 and eqn 2.1 can be written as $\boldsymbol{s}(t) = f^t(\boldsymbol{s}(0))$, where f^t is the t-time map of the state space to itself. It maps the initial condition $\boldsymbol{s}(0)$ t-steps forward in time. More generally, we can write $\boldsymbol{s}(t+\tau) = f^\tau(\boldsymbol{s}(t))$, where f^τ maps state at time t to state at time $t+\tau$ for some $\tau \geq 0$.

As illustrated in Fig. 2.1, geometrically, $f^t(\boldsymbol{s}(0))$ can be interpreted as a trajectory or orbit in the state space starting at $\boldsymbol{s}(0)$ and ending in $\boldsymbol{s}(t)$. The solution to all possible initial conditions, $f^t(S)$, describes a bundle of trajectories called the phase portrait and can be interpreted as a flow of points in the state space. In most systems of interest, the trajectories $f^t(S)$ are bounded, i.e., the distance between $\boldsymbol{s}(t)$ and $\boldsymbol{s}(0)$ does not grow to infinity. This can happen for example due to the presence of dissipative forces, because of energy conservation, or additional constraints on $\boldsymbol{s}$ and $\boldsymbol{F}$. For dissipative systems, in fact, the flow $f^t(S)$ generally contracts to a set of dimension lower than that of S, called an *attractor*. The attractor typically does not have a global Euclidean structure, but exists within a smooth manifold M with dimension $d_a < d$. For physical systems, the dimensionality d of the original state space is related to the number of degrees of freedom of the system, but once the state lies on the attractor, the dynamics have fewer *active degrees of freedom* and thus it requires less information to specify its state.

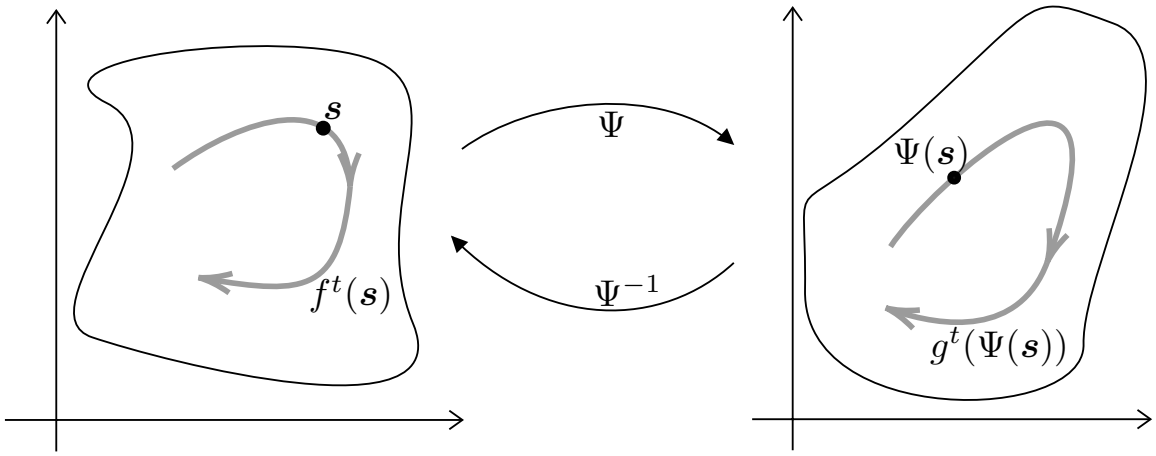


Figure 2.1: A schematic showing topological equivalence. If an invertible, and one-to-one map Ψ exists between two state spaces then they are to be topologically equivalent.

The solution f^t of eqn 2.1, is in general impossible to obtain when $\boldsymbol{F}(\boldsymbol{s})$ is a nonlinear vector field. However, it is still possible to characterize the geometrical and topological properties of the state space. An important concept which enables comparison and classification of different dynamical systems is the idea of topological equivalence. Two vector fields, F and G are said to be topologically equivalent to each other if there is a continuous and one-to-one mapping, Ψ, between them that preserves the orientation of their corresponding orbits $f^t(\boldsymbol{s})$ and $g^t(\Psi(\boldsymbol{s}))$ [9]. Formally, we can write

$$\Psi \circ f^t(\boldsymbol{s}) = g^t \circ \Psi(\boldsymbol{s}) \tag{2.3}$$

This concept is illustrated schematically in Fig. 2.1. The mapping Ψ can be thought of as a coordinate transformation which induces a smooth distortion in the state space flow without messing the order in which the points on the orbits are visited. Implying that the causal structure of the dynamics is preserved under the coordinate transformation.

As a consequence of topological equivalence, any fixed points, closed orbits, limit sets such as attractors, and other important topological details of F are preserved [9]. Moreover, if Ψ is differentiable as well as being continuous, then the stability properties of fixed points and closed orbits are also preserved by the equivalence [9].

A remarkable practical consequence of topological equivalence is that if it's possible to guarantee the existence of the mapping Ψ between a vector field from an unobserved system of interest, and another vector field extracted from some observed dynamical system, then we can use the latter to gain nontrivial insights about the behavior of the unobserved system. In the next section, I will describe how this idea can be applied to the study of complex dynamical systems in the real world.

2.2 The reconstruction problem

Dynamical systems describe real world phenomena evolving in time, such as motion of fluids, population dynamics, animal locomotion and neural dynamics. The underlying state $\boldsymbol{s}$, and evolution law f^t of the dynamical system $\boldsymbol{s}(t) = f^t(\boldsymbol{s}(0))$ governing these phenomena is rarely available explicitly. Instead experimenters measure a D dimensional time-series $\boldsymbol{y}(t) = (y_1(t), y_2(t), \ldots, y_D(t))$ of quantities that they think are most relevant to the phenomena. For example, in the study of animal movement, the underlying state is the neuromuscular dynamics that drive an animal, along with sensory input, and the physics of the interaction with the environment. In practice though, experimenters might track the position of a point on the body; or they might take video recordings and extract some measure of the instantaneous posture of the animal. Another example is fluid phenomena, where the underlying state is governed by the Navier-Stokes equations. In this case, experimenters usually measure the velocity or pressure in a region of space.

We say that the measured time-series $\boldsymbol{y}(t)$ is related to the original dynamical system by $\boldsymbol{y}(t) = h(\boldsymbol{s}(t)) + \boldsymbol{\xi}(t)$, where h is typically an unknown *measurement function*, which maps a point $\boldsymbol{s}(t)$ in the original d-dimensional state space onto the D-dimensional time-series $\boldsymbol{y}(t)$. In addition, $\boldsymbol{y}(t)$ is subjected to observation noise associated with each measurement $\boldsymbol{\xi}(t) = (\xi_1(t), \xi_2(t), \ldots, \xi_D(t))$.

In the absence of explicit knowledge of f and h, it is difficult, if not impossible to uncover the true state space S from noisy measurements alone. However, given, the setup above, one can ask how much useful information is it possible to extract about the underlying state space from the recorded time-series? *This is the state space reconstruction problem, where the goal is to construct a map Ψ that transforms the time-series data to an alternative state space of dimensionality m, with state vectors $\boldsymbol{x} = \Psi(\boldsymbol{y})$, that is topologically equivalent to the original state space in the sense described above.* This is schematically illustrated in Fig. 2.2.

Constructing Ψ from time series data is also known as embedding a time-series, because to be successful, the map Ψ has to be topologically an embedding [10]. To the extent it can be done, embedding an experimental time-series captures as much information about the underlying dynamics as possible. This is because, when Ψ is an embedding, the dynamics in the reconstructed space g^t, are related to the dynamics in the original state space by the following continuous transformation

$$g^t(\boldsymbol{x}) = \Psi \circ f^t \circ \Psi^{-1}\boldsymbol{x} \tag{2.4}$$

Which follows from the condition of topological equivalence stated in eqn 2.3. Equation 2.4 simply says that when Ψ is an embedding, the reconstructed state space vector $\boldsymbol{x}$ after t-steps, which is generated by $g^t(\boldsymbol{x})$, can be obtained by smoothly mapping $\boldsymbol{x}$ to the original state space by $\Psi^{-1}\boldsymbol{x}$, propagating it forward t steps by applying the original dynamics f^t, and then mapping it back to the reconstructed state space by Ψ.

Equation 2.4 is significant, because it means that *an embedding not only gives us a one-to-one copy of the original state space, but also of the dynamics*! This means that estimating g from the reconstructed state space might allow us to gain nontrivial insights about the underlying dynamics, such as their symmetries, conservation laws

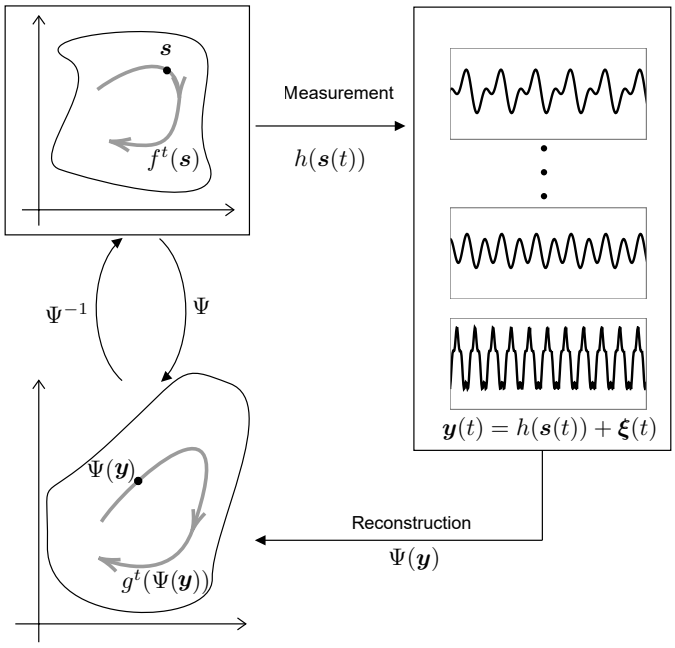


Figure 2.2: The reconstruction problem. True state $\boldsymbol{s}$, and dynamics f are hidden. We only have access to noisy observations in the form of D separate $\boldsymbol{y}$ time-series measurements from the system. The problem is to reconstruct a state space $\boldsymbol{x} = \Psi(\boldsymbol{y})$, that is topologically equivalent to the original state space, using just the time-series data.

and other dynamical properties. We can also use the estimated model for things like noise reduction, prediction, and comparison and classification of observed time-series. Finally, even if the dynamics are non-stationary, say due to a time-dependent control signal; g will still capture the stationary part of the dynamics, meaning that any discrepancy between model predictions and observed time-series allows us to obtain insights about the time-dependent control signal. In the following section I will describe some of the methods that allow the construction of en embedding from the time-series.

In the thesis I will refer to Ψ as the *total reconstruction map;* Ψ is a composition of all the transformations we do as part of reconstructing the state space, starting from the process of measurement. As an example, if before reconstruction we perform a noise reduction step, then Ψ is the composition of the measurement function, noise reduction map, and another state-space reconstruction map. These ideas will be developed in detail in the following sections.

2.3 Background to Takens' Embedding Theorem

The fact that it is possible to reconstruct an embedding from a time-series that is topologically equivalent to the underlying state space was first demonstrated numerically by Packard et al [11], they used numerical derivatives estimated from a single time-series to form an embedding and showed that topological invariants such as dimensions and Lyapunov exponents could be estimated from it. About the same time, Floris Takens independently considered an alternative approach [10]; the *delay reconstruction map* Φ, which amounts to using delays of a time-series as the embedding coordinates $\boldsymbol{x}(t)$, i.e.

$$\boldsymbol{x}(t) = \big(\boldsymbol{y}(t), \boldsymbol{y}(t-\tau), \ldots, \boldsymbol{y}(t-(m-1)\tau)\big)$$

where τ is the time lag, and m is the dimension of the reconstructed space. By studying the properties of Φ, he famously proved what is now known as Takens' embedding theorem, which states that in the noiseless case, the delay reconstruction map Φ is generically an embedding for $m \geqslant 2d+1$, where d is the dimension of the original state space S [10]. In the same work he also proved that the derivative coordinates, which can be seen as linear transformations of delay vectors $\boldsymbol{x}(t)$, also form an embedding for $m \geqslant 2d+1$ in the noiseless case. Simply put, Takens theorem states that the historical record of a time-series can be leveraged to gain information about the underlying system. This is perhaps not surprising, indeed, statisticians had been using this idea under the framework of Autoregressive modeling for a while [12]. Takens' contribution was to show that temporal history can also recover geometrical and topological properties of the dynamical system. Thus in some sense, the information contained in the temporal history can be complete, it contains all there is to know about the system.

Takens' original embedding theorem had several theoretical [10] (e.g. it didn't apply to dynamics on compact sets, did not apply to stochastic or multi-deimensional systems etc). However, over the years it was generalized significantly. First significant improvement was done in [13], where authors removed the limitation of compact sets and filtered time-series. Subsequent generalizations were made to include a broad class of stochastic systems [14, 15], systems driven by deterministic forcing [16], time-series

with uneven sampling [17], filtered time-series [13], and most recently to the embedding of multivariate signals [18].

Takens' original theorem and its generalizations provide important guarantees and justifications for state space reconstruction. Using Takens' embedding theorem as a general strategy for geometrical and topological characterization of time-series is applicable to a wide variety of dynamical phenomena. However, the theorems are of little practical use in guiding towards generating good embeddings from noisy real world data. Specifically, they provide no insight about the role of the measurement function $h(\boldsymbol{s}(t))$ on state space reconstructions, which is a key choice an experimenter has to make when faced with an unknown dynamical system. They also do not discuss what constitutes an optimal embedding, neither do they talk about how to choose key embedding parameters such as the lag time τ, dimension of embedding space m. Finally, most generalizations do not mention how to best combine multivariate time-series to form a single embedding.

Efforts to close the gap between the theory and practice of state space reconstructions have been in two general directions. First, to develop an understanding of the effect of various parameters on the quality of an embedding [19, 20], which has led to several heuristics towards choosing good values for different embedding parameters [21, 22]. While, the second line of work has used tools from statistics to find good transformations of the delay coordinates. The most famous example of this is Broomhead and King's singular spectrum analysis (SSA) [23] which uses Singular Value Decomposition (SVD) on time delays to reconstruct a low dimensional state space. Although, much progress has been made since Packard et al, and Takens' pioneering work. Key issues still remain open. The biggest issue in my opinion is that there are no agreed upon quantitative measures for the quality of an embedding. For low-dimensional systems, a visual inspection works best to choose parameters leading to a good embedding. However, visual inspection becomes impractical as soon as we go up in dimensionality. This is perhaps part of the reason why state space reconstruction methods haven't been as successful in the study of high dimensional dynamical systems as they have been for low dimensional systems. In the next section I review popular reconstruction methods in greater detail and discuss common problems that all methods encounter.

2.4 Review of Popular Reconstruction Methods

To make the ideas in this section more concrete, let's set up the following measurement problem. We will receive simulated observations from the Lorenz system described in eqn 1.1. Specifically, let's now suppose that we measure N samples of the first variable s_1, i.e. the measurement function is $h(\boldsymbol{s}(t)) = s_1(t)$, and our observed time-series is $y(t) = s_1(t) + \xi(t),\ t \in [1, N]$. The problem is to reconstruct a mapping to the original state space using the information is y(t). Fig. 2.3A shows projections of the original state space coordinates show, while Fig. 2.3B, shows a sample of the measured signal. We describe how different methods achieve this and the problems they encounter.

2.5 Method of Delays

The most straightforward implementation of Takens idea is to simply construct the following delay vector from the above time-series, i.e.

$$\boldsymbol{x}(t) = \big(y(t), y(t-\tau), y(t-2\tau), \ldots, y(t-(m-1)\tau)\big)$$

for some lag time τ and, and embedding dimension m. The convention in this thesis will be to treat$\boldsymbol{x}(t)$ as a row vector. The first delay vector is $\boldsymbol{x}(t_0)$, where $t_0 = (m-1)\tau + 1$, and the last is $\boldsymbol{x}(N)$, thus the total number of delay vectors will be $N' = N - (m-1)\tau$. As an example, if we had $N = 8$ samples $\big(y(1), \ldots, y(8)\big)$, and we set $m = 3$, $\tau = 2$, then we can arrange the $N' = 4$ delay vectors into the following 4×3 delay matrix X

$$X = \begin{bmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_0+1) \\ \boldsymbol{x}(t_0+2) \\ \boldsymbol{x}(t_0+3) \end{bmatrix} = \begin{bmatrix} y(5) & y(3) & y(1) \\ y(6) & y(4) & y(2) \\ y(7) & y(5) & y(3) \\ y(8) & y(6) & y(4) \end{bmatrix}$$

In general, a time-series of N samples will give us the following delay matrix $N' \times m$ delay matrix

$$X = \begin{bmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_0+1) \\ \vdots \\ \boldsymbol{x}(N) \end{bmatrix} = \begin{bmatrix} y((m-1)\tau+1) & y((m-2)\tau+1) & \ldots & y(1) \\ y((m-1)\tau+2) & y((m-2)\tau+2) & \ldots & y(2) \\ \vdots & \vdots & \ddots & \vdots \\ y(N) & y(N-\tau) & \ldots & y(N-(m-1)\tau) \end{bmatrix}$$

The delay reconstruction as set up above, requires two parameters, the lag time τ, and the embedding dimension m. For a fixed embedding dimension, changing τ changes the length of the time window that defines a state. In the example above, for $m = 3$ and $\tau = 2$, the window length denoted by K is 5. This means that the state at time t contains information from the current time point, and previous 4 time points. In general the window length is $K = (m-1)\tau + 1$.

For systems that decorelate, either because of noise or because of chaotic dynamics, the window length K is a crucial parameter. If K is too small, then we might hit the noise time-scale or the sampling time-scale; in either case there will not be enough samples to see any nontrivial structure. On the other hand, if K is too large (compared to, say the autocorrelation time), then the points included in the definition of the state might not be related to each other at all, and the state space trajectories will look random. This idea is implicit in Takens' theorem which states that the state space points must lie on the same trajectory. In the continuum limit which Takens studied, it is easy to ensure that points lie on the same trajectory. However, for real, noisy data, or even numerical integration schemes, it is much harder to enforce the constraint. In fact the right choice of K might depend on several different factors such as noise, sampling time, numerical precision, and even integration scheme used, as has been discussed earlier [24, 25].

The effect of window size on the reconstructions using the method of delays is shown in the figure below. We have kept the embedding dimension $m = 3$ fixed

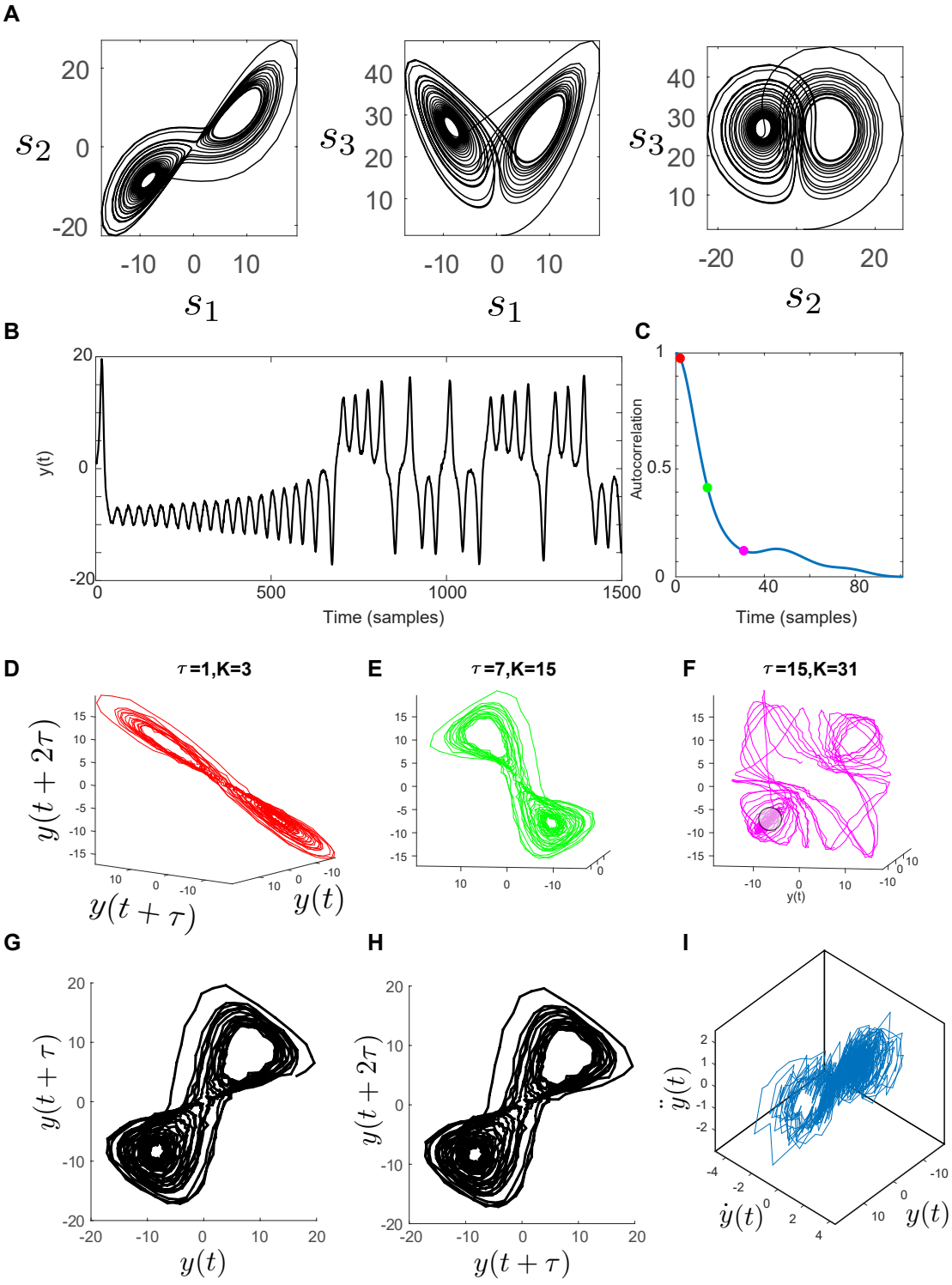


Figure 2.3: Delay embedding on the Lorenz attractor. **A.** The original state space variables of eqn 1.1. **B.** A sample of the noisy observation which is used for state space reconstruction. **C.** Autocorrelation function of y(t) with the dots corresponding to $K = 3, K = 15$ and $K = 31$. **D-F.** Delay embedding reconstructions for the parameters listed on top. **G-H.** Two delay coordinate projections that look identical. **I.** Derivative coordinates using backward finite difference approximation.

and plotted reconstructions for different lag times τ, corresponding to window sizes of $K = 3, 15, 31$ respectively. Fig. 2.3C shows the relationship between the window sizes and the autocorrelation function. When the window size is, $K = 3$, the reconstructed attractor is stretched out along diagonal in state space, and squeezed in the orthogonal directions. This happens generically for small windows, as the points in a delay vector are nearly identical to each other (up to observation noise). Next, for $K = 15$, we are able to resolve most of the details, the attractor looks more uniform, and visually at least, it is a good embedding. Finally, for $K = 31$, the geometry is significantly more complex; in fact, there is barely any resemblance left to the original state space of the Lorenz attractor. Importantly, there are regions in the reconstructed attractor, such as the one marked by the black circle, where the state space points are densely packed and difficult to resolve.

The window size $K = 31$ lies at the first minimum of the autocorrelation function, possibly suggesting that there is a relationship between K and the correlation structure of signal time points. Indeed, several heuristics exist in literature which suggest setting K to quantities derived from the autocorrelation or automutual information function. Examples of these suggestions include, the autocorrelation time [26], the first minimum of the automutual information [21], and the first zero crossing or the first minimum of the autocorrelation function [21]. For systems where there is a single underlying period, most of the above heuristics give a similar estimate of K. However, when there is no periodic structure, or there are multiple coexisting periods, the estimates generally differ, and there is no well agreed upon heuristic [26]. As we will see, finding an optimal value for K is a first step in all state space reconstruction techniques. Consequently, it could be argued that the lack of reliable guides to optimize K is a major limitation of state space reconstruction methods.

The effect of embedding dimension m, is in some sense less severe, and better understood. Until the attractor is not fully embedded, and the dimension is insufficiently low, it consists of self-intersections where the state space cannot be fully resolved. However, once, the attractor is fully embedded, the effect depends on the application. For example, estimates that are based on the local density of points, such as the correlation dimension, or the metric entropy are not fundamentally affected by the value of m once the attractor is embedded. On the other hand, estimates that are based on approximating the local Jacobians, such as the spectrum of Lyapunov exponents or the Kaplan-Yorke dimension, are affected. This is because, in m dimensions, the local Jacobians are $m \times m$ matrices, and have m, usually distinct eigenvalues. But if the attractor has dimension $d < m$, then we have $m - d$ spurious eigenvalues which are typically not easy to weed out [7]. Consequently, knowing the right embedding dimension is crucial for the estimate of the Lyapunov spectrum and related quantities.

We can use results from Fig. 2.3, to conceptually understand the distinction between a good embedding and a bad embedding. Formally, an embedding enables a one-to-one mapping from the reconstructed state space to the underlying state space using the inverse of the total reconstruction map Ψ^{-1}. As Takens showed, in the continuum limit, this only fails when there are self intersections. However, for real world data, there is a limit to our resolution, denoted by ϵ. If distance between two points is of the order ϵ, then it becomes impossible to tell them apart. This can happen when two points are actually on an intersection, or it can happen because the complex

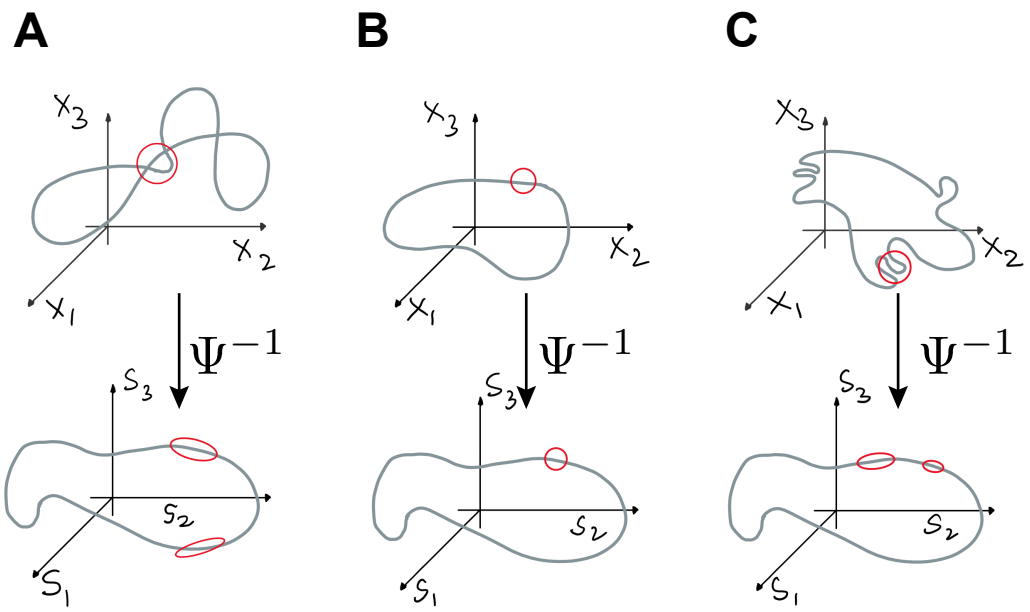


Figure 2.4: Good vs Bad embeddings **A.** A real self-intersection implies that points in the neighborhood cannot be mapped uniquely back to the underlying state space. **B.** A good embedding has no self-intersections and a simple geometry, as a result, points in the reconstructed space can be mapped back to their original states. **C.** Complex geometry can also lead to effective self-intersections when the attractor gets so curved that two points that are not neighbors come within minimum resolution of the dataset. Denoted here by the red circle.

geometry of the reconstructed attractor brings them within a distance ϵ of each other. Schematically, this is illustrated in Fig. 2.4. If the radius of the red ball is taken to be ϵ, then in Fig. 2.4A we see that a self intersection will result in the mapping that is not one-to-one to the original state space. In Fig. 2.4B, we have a "good" embedding, where a simple geometry of the attractor leads to a successful one-to-one mapping of all the states. Finally, in Fig. 2.4C, we see that when the reconstructed attractor has a complex geometry, we can again fail to get a one-to-one mapping of all states, as two points might come within ϵ of each other due to strong curvature effects. The difficulty of formalizing this notion, in order to develop a measure of a good embedding is that, without access to the underlying state space it is not possible to know whether a reconstruction is one-to-one. Nevertheless, as we will see in a later section, the effects of complex geometry can be measured, which allows us to find an optimal value of the embedding window K.

Finally, to conclude this section, we would like to point out some shortcomings of the method of delays as it's implemented above. First, the basis that describes the canonical delay embeddings is arbitrary, there is no reason to use integer multiples of τ as the coordinate axes. Indeed, the most general case would be to use $\boldsymbol{x}(t) = \big(y(t), y(t-\tau_1), \ldots, y(t-\tau_m)\big)$ to define an m dimensional embedding. However, this would introduce m different parameters, $(\tau_1 \ldots \tau_m)$, with no clear way of determining them. A consequence of the choice of basis vectors for Takens embedding is that it introduces artificial symmetries. This the projections difficult to interpret, and results in a confusing property of the delay coordinates where many projections are identical. This is shown in Fig. 2.3G and 2.3H which show the projection of the delay embedding for $tau = 7\, m = 3$ on two different planes. Further, the method does not deal with observation noise in any way. And finally, the method is designed for scalar time-series. It is not immediately clear how one might include information from a multivariate time-series into the framework. In the following sections, we generalize the framework of delay embedding by abstracting out the choice of basis from constructing delays and then use it to review two other reconstruction methods.

2.6 An Abstraction

It is useful to introduce the following abstraction in the delay embedding framework we have described above. Consider the general case of D dimensional time-series $\boldsymbol{y}$, and let's introduce the following matrix.

$$\overline{Y} = \begin{bmatrix} \boldsymbol{y}(K) & \boldsymbol{y}(K-1) & \ldots & \boldsymbol{y}(1) \\ \boldsymbol{y}(K+2) & \boldsymbol{y}(K) & \ldots & \boldsymbol{y}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{y}(N) & \boldsymbol{y}(N-1) & \ldots & \boldsymbol{y}(N-K+1) \end{bmatrix} \tag{2.5}$$

Where, as before, K is the window length, i.e. the total number of past samples that go into the definition of a state, and N is the number of samples in our time-series. Clearly, the size of $\overline{Y}$ is $N' \times KD$, where $N' = N - K + 1$. For a scalar time-series, $\overline{Y}$ is the simply the delay matrix introduced above, with $\tau = 1$, and $m = K$. However,

the usefulness of $\overline{Y}$, comes from the fact that *all* state space reconstructions can be seen as linear (or nonlinear) transformations of $\overline{Y}$. Thus, it is the starting point, and the fundamental object of interest in any state space reconstruction method. This also makes it explicit that the window length K is the key parameter that determines the success of an embedding.

We will call $\overline{Y}$ the *state matrix* of a time-series because it contains all the information that goes into reconstructing a state from a time-series. To make this idea more concrete, let's consider the above example again, where we observe $N = 8$ samples $\big(y(1), \ldots, y(8)\big)$ and $K = 5$. The state matrix is then

$$\overline{Y} = \begin{bmatrix} y(5) & y(4) & y(3) & y(2) & y(1) \\ y(6) & y(5) & y(4) & y(3) & y(2) \\ y(7) & y(6) & y(5) & y(4) & y(3) \\ y(8) & y(7) & y(6) & y(5) & y(4) \end{bmatrix}$$

To obtain Takens' delay reconstruction with $m = 3$ and $\tau = 2$, we simply project $\overline{Y}$ onto the matrix V_{Takens}, as shown below

$$X = \overline{Y}V_{Takens} = \begin{bmatrix} y(5) & y(4) & y(3) & y(2) & y(1) \\ y(6) & y(5) & y(4) & y(3) & y(2) \\ y(7) & y(6) & y(5) & y(4) & y(3) \\ y(8) & y(7) & y(6) & y(5) & y(4) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} y(5) & y(3) & y(1) \\ y(6) & y(4) & y(2) \\ y(7) & y(5) & y(3) \\ y(8) & y(6) & y(4) \end{bmatrix}$$

Thus, in general, Takens' delay reconstruction corresponds to a linear transformation $\overline{Y}$ by a $K \times m$ matrix V_{Takens}, where each column of V_{Takens} is a basis vector with 1 at the $(i-1)\tau + 1$ entry and 0 otherwise.

Numerical derivatives are another common embedding method that can be seen as a linear transformation of $\overline{Y}$. For example, for window length $K = 3$, a simple backward finite difference approximation of the first and second order derivatives of the data above would be,

$$X = \overline{Y}V_{fdiff} = \begin{bmatrix} y(3) & y(2) & y(1) \\ y(4) & y(3) & y(2) \\ y(5) & y(4) & y(3) \\ y(6) & y(5) & y(4) \\ y(7) & y(6) & y(5) \\ y(8) & y(7) & y(6) \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} y(3) & y(3) - y(2) & y(3) - 2y(2) + y(1) \\ y(4) & y(4) - y(3) & y(4) - 2y(3) + y(2) \\ y(5) & y(5) - y(4) & y(5) - 2y(4) + y(3) \\ y(6) & y(6) - y(5) & y(6) - 2y(5) + y(4) \\ y(7) & y(7) - y(6) & y(7) - 2y(6) + y(5) \\ y(8) & y(8) - y(7) & y(8) - 2y(7) + y(6) \end{bmatrix}$$

A projection onto the derivative coordinates is shown in Fig. 2.3I, and illustrates that in the presence of noise, finite difference based derivatives result in significantly worse embedding than delays.

Multidimensional Observations

When we are dealing with a multidimensional observations only the construction of the state matrix $\overline{Y}$ needs to change. Specifically, we treat the multidimensional time-series

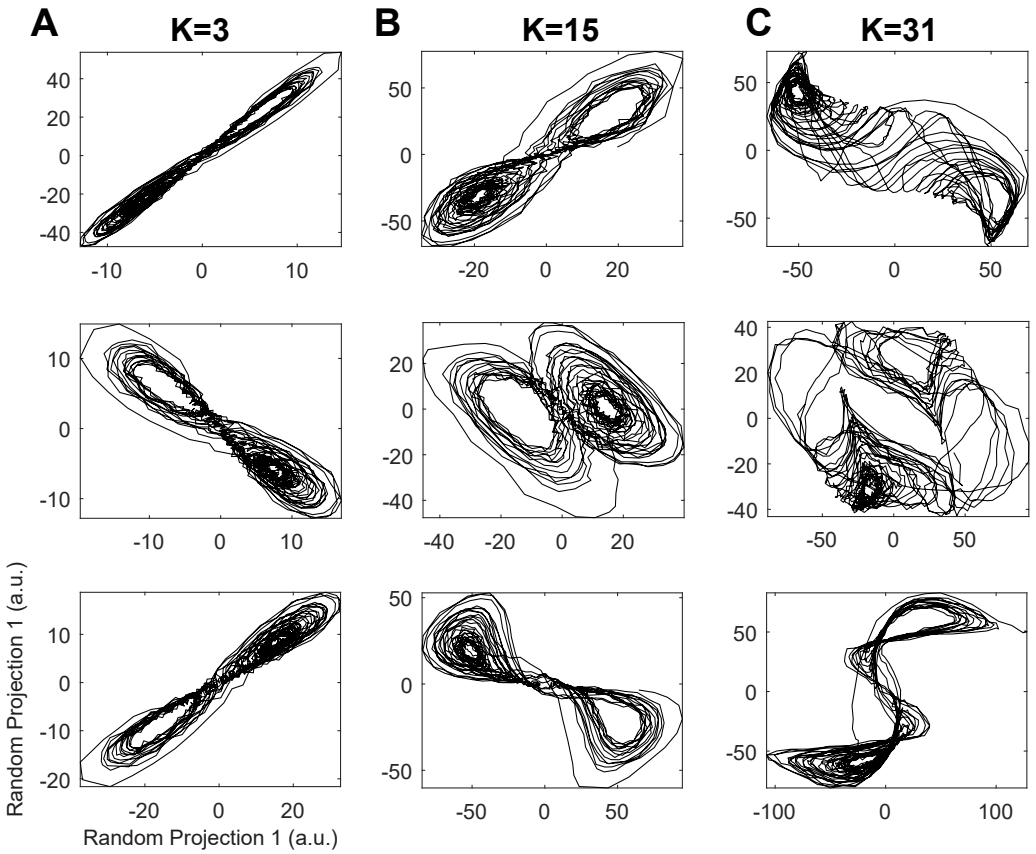


Figure 2.5: Random projections can also result in valid reconstructions **A.** $\overline{Y}$ with $K = 3$ is projected onto a random matrix with unit Gaussian entries. Qualitatively similar features to delay reconstruction are observed; attractor is stretched along the diagonal line and squeezed perpendicular. **B.** Random projection of $\overline{Y}$ with $K = 15$, here the attractor is more uniform, and more details are visible. **C.** Random projection of $\overline{Y}$ with $K = 31$, once again we see that increasing the window length leads to a complex geometry of the attractor with little resemblance to the original state space.

$\boldsymbol{y}(t)$ as a single object and construct $\overline{Y}$ according to eqn 2.5. For example, let's say we observe $N = 8$ samples of a 2−dimensional time-series, $\boldsymbol{y}(t) =$ as shown below

$$\boldsymbol{y} = \begin{bmatrix} y_1(1) & y_2(1) \\ y_1(2) & y_2(2) \\ y_1(3) & y_2(3) \\ y_1(4) & y_2(4) \\ y_1(5) & y_2(5) \\ y_1(6) & y_2(6) \\ y_1(7) & y_2(7) \\ y_1(8) & y_2(8) \end{bmatrix}$$

In this case $\overline{Y}$ has the following form

$$\overline{Y} = \begin{bmatrix} y_1(5) & y_2(5) & y_1(4) & y_2(4) & y_1(3) & y_2(3) & y_1(2) & y_2(2) & y_1(1) & y_2(1) \\ y_1(6) & y_2(6) & y_1(5) & y_2(5) & y_1(4) & y_2(4) & y_1(3) & y_2(3) & y_1(2) & y_2(2) \\ y_1(7) & y_2(7) & y_1(6) & y_2(6) & y_1(5) & y_2(5) & y_1(4) & y_2(4) & y_1(3) & y_2(3) \\ y_1(8) & y_2(8) & y_1(7) & y_2(7) & y_1(6) & y_2(6) & y_2(5) & y_2(5) & y_2(4) & y_2(4) \end{bmatrix}$$

After construction of $\overline{Y}$, rest of the pipeline is identical to what is described above. This framework allows us to split process of reconstruction into two parts, first is the construction of the state matrix $\overline{Y}$, which is essentially the choice of the window length K. Second, is the choice of a transformation P, which maps KD dimensional vectors in the state matrix, to an m dimensional state space. Here, the choice of embedding dimension m is made explicit. In principle, P can be any linear or nonlinear transformation that results in an embedding. Indeed, as shown in Fig. 2.5, even random matrices, i.e. setting P equal to a random matrix V with normally distributed entries (or any other distribution) result in perfectly reasonable embeddings, that are significantly better than derivative coordinates [27]. Fig. 2.5 emphasizes the dependence of attractor geometry on the widow size K. It also suggests that the window length's role on attractor geometry is independent of the transformation V.

If P can be any transformation of the state matrix. Then a natural question to ask is: does an optimal transformation P exists? And if so, is there a straightforward way to find it? In the next section we will see that if we restrict ourselves to the space of linear transformations, then the answer is yes to both questions.

2.7 Singular Spectrum Analysis

As we saw in the previous section, the transformations giving rise to both delay embedding, and derivate embeddings were in some sense arbitrary. In fact the derivative embeddings actually performed worse than simply projecting $\overline{Y}$ on to a random set of basis vectors. The question then is can we do better? In 1986, Broomhead and King showed that if we take into account the correlational structure of the points in $\overline{Y}$, then we can [23]. Their idea was to perform a singular value decomposition on $\overline{Y}$, using the first m right singular vectors as the basis vectors for their embedding space, a method they called Singular System Analysis. Soon, Vautard and Ghil [28] working in the

context of climate sciences extended Broomhead and King's work, calling it Singular Spectrum Analysis (SSA). They also connected it to older results in stochastic processes literature in the name of Karhunen-Loeve decomposition [29, 30, 31]. Since then SSA has become an important tool in fields dealing with spatiotemporal dynamical systems such as climate and geo sciences [32, 33].

The basic idea underlying SSA is that although, each point in the state matrix $\overline{Y}$ is of dimension KD, typically however, $KD \gg d$. This means that the points in $\overline{Y}$ do not span the entire KD dimensional space; but are well approximated by a lower-dimensional subspace. Thus, the best linear transformation of the state matrix $\overline{Y}$ is the one that identifies the linear subspace along which the data points in $\overline{Y}$ mostly lie. This is because any other transformation will contain directions which do not explain most of the data points and will therefore be redundant [19, 23].

The problem then is to find a set of orthogonal basis vectors $\boldsymbol{v}_i$, along which most of the data lies. This is done by a singular value decomposition (SVD) of the state matrix $\overline{Y}$ [19, 23]. Briefly, SVD finds the matrix V of orthogonal basis vectors $\boldsymbol{v}_i$, such that the projections onto $\boldsymbol{v}_i$ are maximally spread apart. Formally, we want the sum of squared distances, σ_i^2, of the projections $\overline{Y}\boldsymbol{v}_i$, onto each of the basis vectors to be maximum. Statistically, this amounts to maximizing the variance of the data along each basis vector $\boldsymbol{v}_i$. Indeed, the basis vectors $\boldsymbol{v}_i$ are nothing but the eigenvectors of the covariance matrix $\overline{Y}^T\overline{Y}$, and σ_i^2 are its eigenvalues.

In matrix notation, the above can be written as $\overline{Y}V = U\Sigma$, or the way its more popularly written, $\overline{Y} = U\Sigma V^T$. Where, V is a $KD \times KD$ orthogonal matrix with each column being the basis vector $\boldsymbol{v}_i$. Σ is a $KD \times KD$ diagonal matrix consisting of singular values σ_i, and U is the matrix of projections onto the singular vectors scaled by the singular values, i.e., $U = \overline{Y}V\Sigma^{-1}$. Simply speaking, we are describing the points in $\overline{Y}$, as an ellipsoid; $\boldsymbol{v}_i$ are the directions of the principal axes of the ellipsoids, and σ_i are their lengths.

The usefulness of SSA comes from the fact that the singular values are ordered in decreasing order by the variance explained by the projection onto each $\boldsymbol{v}_i$ (or the lengths of principal axis of the ellipse), i.e. $\sigma_1 \geqslant \ldots \geqslant \sigma_{KD}$. This singular value spectrum provides important information about the noise and precision limits of the data. Typically, the singular values decrease until they hit a plateau, beyond which they're roughly equal [19, 23]. The height of the plateau is related to the noise variance $\langle \boldsymbol{\xi}^T\boldsymbol{\xi} \rangle$ in the observations $\boldsymbol{y}(t)$, and any singular value which is in the plateau is dominated by noise [19, 23]. Thus, in the case of noisy observations, if the points approximately lie on a p dimensional subspace, then $\sigma_1 \geqslant \ldots \sigma_p \gg \sigma_{p+1} \ldots \sigma_{KD} \approx \langle \boldsymbol{\xi}^T\boldsymbol{\xi} \rangle$, i.e. the first p singular values are significantly above the plateau and the remaining are roughly equal to the noise variance.

One might wonder if the number of significant singular values is related to the dimensionality of the attractor in a trivial way. Unfortunately, it has been shown analytically, and numerically, that the number of significant singular values p, are not related to the embedding dimension [19, 26]. Instead, they reflect the number of linear modes in the dynamics that are above the noise resolution [19, 23]. In fact, it can be shown that for nonlinear dynamics, even in infinite precision, no singular value will be identically 0 [19]. *This means that a separate method of estimating the embedding*

dimension is required once have projected $\overline{Y}$ *onto the top p singular vectors.*

An important advantage of SSA is that it is theoretically well understood. It can be shown that in the large window limit of K tends to infinity, SSA is equivalent to discrete Fourier analysis [23, 28]. Further it was shown in [19], that the small window limit of K tending to zero corresponds to estimating numerical derivatives of the time-series. Several other theoretical properties of SSA and its connections with other methods in signal processing can be found in the review [32].

Finally, Fig. 2.6 shows the results of state space reconstruction using SVD. The projections in first three plots of Fig. 2.6A shows qualitatively similar results for small K as before. The reconstructions are dominated by noise. Additionally the singular value spectrum is decreasing for all three dimensions, and no cutoff is observed, meaning that there is not enough data to obtain a separation between signal and noise. For $K = 15$ in Fig. 2.6B, we observe our best reconstructions yet. The trajectories are noise free, and geometrically simple. The singular value spectrum shows a clear separation between signal and noise, with $p = 6$ singular values above the noise floor. Finally, the singular vectors resemble Legendre polynomials, as was proven by authors in [19] for small K. In Fig. 2.6C we show results for $K = 31$. Once again, we see that for long window size, geometry of the attractor becomes complex, trajectories appear random, and the reconstruction loses resemblance to the original state space. The singular spectrum shows $p = 10$ significant singular values, and the singular vectors appear more sinusoidal consistent with the long window limit approaching fourier basis [19]. It should be emphasized again that the number of significant singular values do not signal the dimensionality of the attractor or the embedding space.

2.7.1 SSA and Linear Modes

SSA is able to easily deal with multidimensional time-series. Importantly, for linear dynamical systems SSA is able to reconstruct the normal modes from time-series of the configuration variables alone (depending on the window length K)). To illustrate this point, I applied SSA on a time-series measurement of q_1, q_2 from the linear coupled oscillators described in eqn 1.3. To reconstruct the state space, I first constructed $\overline{Y}$ with K corresponding to a quarter of ω_2 and then applied SVD to $\overline{Y}$. The results are shown in Fig. 2.7. As can be seen the reconstructed coordinates are composed of two regular oscillations corresponding to the two normal modes of this system. The phase space shows two circles corresponding to the regular oscillations. The phase portrait is simply a scaled version of the phase portrait of the normal modes shown in Fig. 1.3B.

2.7.2 Independent Component Analysis and Additional Linear Transformations

As mentioned above SSA finds the optimal subspace, in a least square sense, in which most of the attractor lies. However, because of the SVD the basis vectors are constrained to be orthogonal. In practice, it might not be valid to assume that the basis vectors describing the reconstructed state space be orthogonal. Ideally, we want the state space coordinates to be *independent*, and it could be desirable to perform a further optimization within the SVD subspace for a set of independent basis vectors.

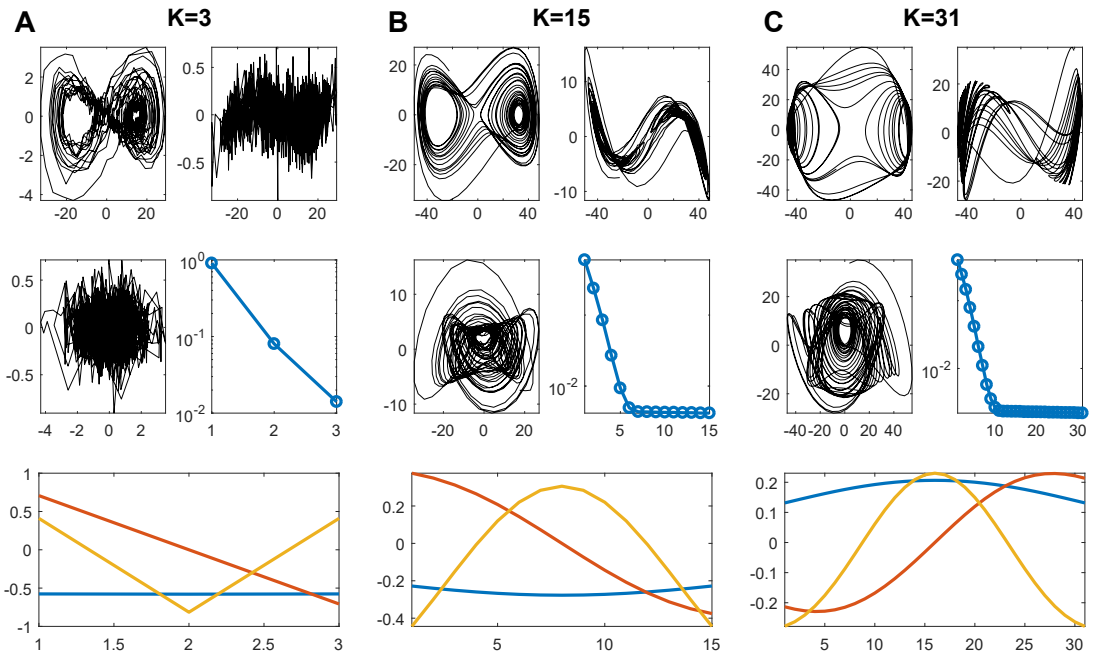


Figure 2.6: Results from application of SVD on $\overline{Y}$ with different window sizes. The first three plots in **A.**,**B.**,**C.** show projections of the attractor onto the three pairs of singular vectors. Fourth plot shows normalized singular values, $\sigma_i / \sum \sigma_i$ on a log scale. And the bottom plot shows the three singular vectors. **A.** SVD on $\overline{Y}$ with $K = 3$. As before, we see that for small K, the projections are dominated by noise. Additionally, the singular values do not saturate and the singular vectors are piecewise linear functions. **B.** SVD on $\overline{Y}$ with $K = 15$ gives us the cleanest embedding so far, the trajectories are smooth, and the attractor has a relatively simple geometry. The singular values, hit the noise floor after about $p = 6$ dimensions, implying that higher dimensions are dominated by noise. **C.** SVD on $\overline{Y}$ with $K = 31$, again results in a complex shaped attractor with no resemblance to the original state space. This suggests that SVD can also not deal with the effects of large embedding windows. The singular spectrum drops to the noise floor after about $p = 10$ dimensions, meaning that the number of significant singular vectors are dependent on the window length. Finally, the singular vectors are starting to look more sinusoidal, consistent with the large window limit.

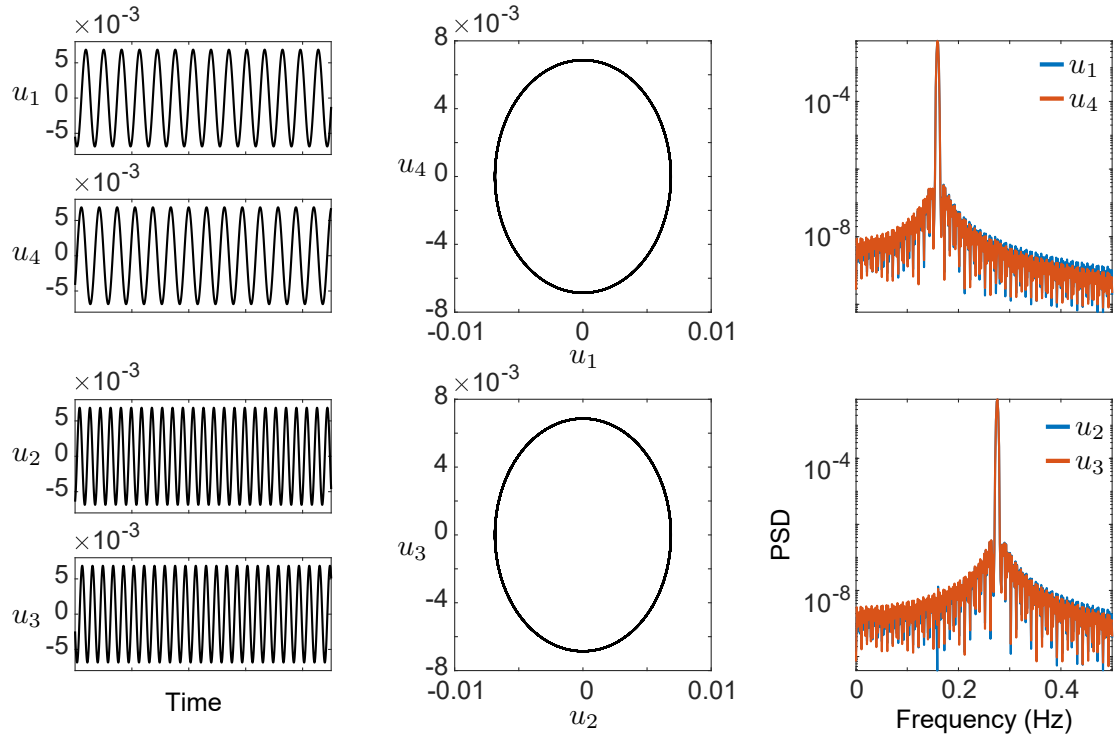


Figure 2.7: The four coordinates derived by SSA correspond to the normal modes of the linear oscillator described in eqn 1.3. Notice the similarity with Fig. 1.3B.

Independent component analysis (ICA) [34] is a popular method that achieves the goal mentioned above. In particular, ICA attempts to find basis vectors that are independent, instead of being just orthogonal. Specifically, the ICA basis can be found by an orthogonal transformation of the projections in the normalized or whitened SVD basis. The procedure corresponds to simply a rotation of the whitened SVD subspace. The orthogonal transformation (rotation) is estimated by maximizing the non-Gaussianity of the marginal densities along each axes [34]. The justification for this comes from the central limit theorem, which states that a linear combination of independent variables is more Gaussian than the original variables. Thus ICA makes an equivalence between independence and non-Gaussianity, and uses the latter to estimate an independent set of basis vectors describing the data. In several applications the additional linear transformation given by ICA can be useful for the purposes of interpretation. Indeed, in the latter chapters I will use ICA to perform reconstruction of *C. elegans* crawling dynamics.

Along with ICA several other linear transformations can be used depending on the assumptions one wants to make. For example sparse PCA [35], non-negative matrix factorization [36] and sparse autoencoder neural networks [37].

2.8 Choosing The Right Embedding By Maximizing Predictability

The full reconstruction pipeline is shown in Fig. 2.8. To summarize, we first lift a D dimensional experimental time-series into a KD dimensional space by stacking delayed copies of the time-series up to a window length K. This gives us the state matrix $\overline{Y}_K = \Phi_K(y)$, where Φ_K represents the delay map with window length parameter K. Next, we find a transformation P_m to project $\overline{Y}_K$ on to an m dimensional subspace of $\mathbb{R}^{KD}$, giving us a candidate reconstructed state space $X_{K,m} = P_m \circ \Phi_K(\boldsymbol{y})$. Subscripts in Φ_K and P_m make their parameter dependence explicit. The entire process is therefore represented by the total reconstruction map $\Psi_{K,m} = P_m \circ \Phi_K \circ h$. For any choice of K and m we have a candidate state space reconstruction, the problem then is to choose those reconstructions that are optimal in some sense. We have seen already in the previous section that not any choice of parameters gives us a reasonable reconstruction. Moreover, it is not immediately clear from Takens' theorem and its generalizations which values of K and m should result in a good state-space reconstruction, or what a good state space reconstruction even is. In the following, I use the condition of topological equivalence discussed earlier to present arguments that suggest a good state-space reconstruction allows us to make good predictive models of the time-series. Consequently, optimal embedding parameters can be identified by maximizing an appropriately defined measure of predictability. Although the following discussion is applicable to any reconstruction method, I will focus on SSA (i.e. P_m is given by the right singular vectors V_m of $\overline{Y}_K$), as it is the main method that will be used throughout this thesis.

As mentioned in section 1, for $\Psi_{K,m}$ to result in an embedding, it must be topologically equivalent to the underlying state space. Thus, if our candidate reconstruction is an embedding then it must satisfy the condition of topological equivalence in eqn 2.3.

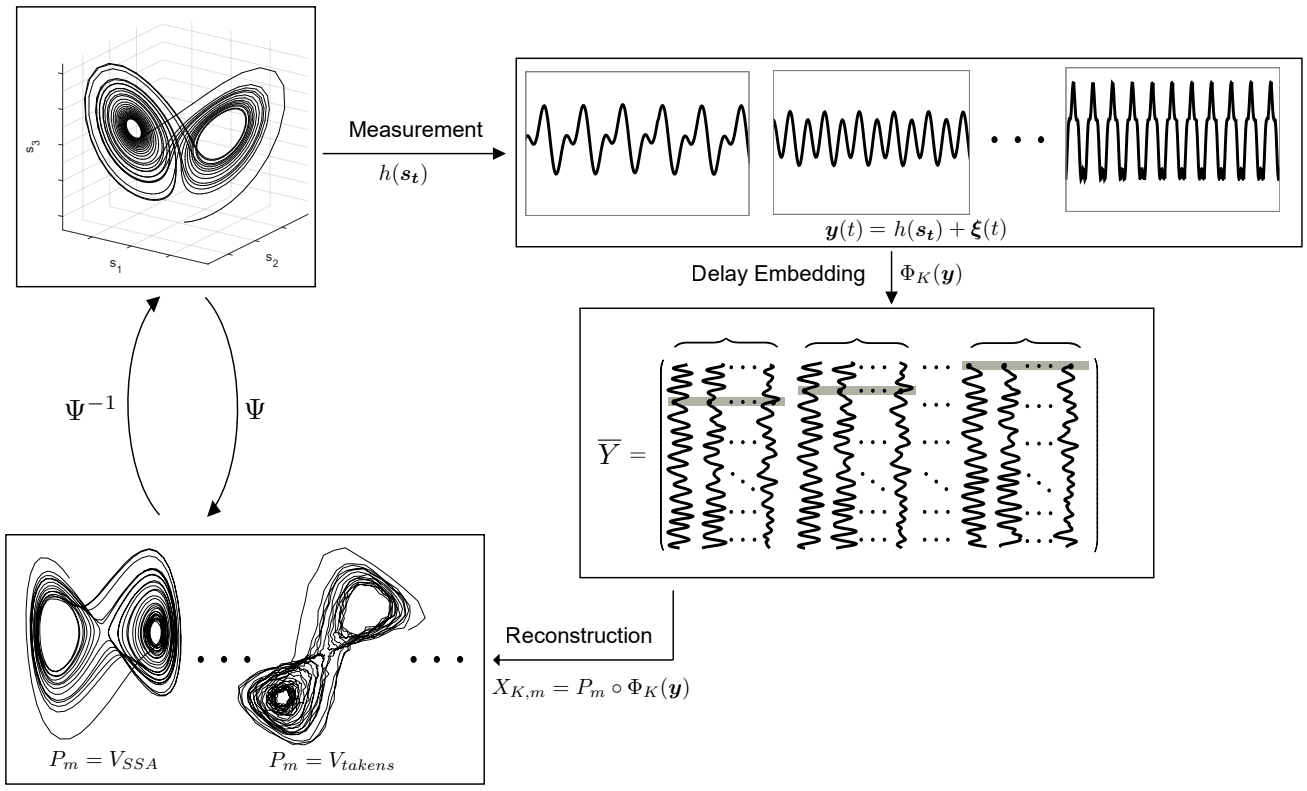


Figure 2.8: The entire reconstruction pipeline summarized. Starting from some measurements, we construct the state matrix by choosing the window length K and then project to an m-dimensional subspace by applying the transformation P_m. Examples of P_m include, Takens' method of delays, singular spectrum analysis, numerical derivatives and random projections.

Specifically, we can write

$$\Psi_{K,m} \circ f^{\tau}(\boldsymbol{s}_t) = g^{\tau}_{K,m} \circ \Psi_{K,m}(\boldsymbol{s}_t) \tag{2.6}$$

where, as before, f represents the evolution law in the underlying state space S. It maps the current state $\boldsymbol{s_t}$, τ steps forward, i.e. $\boldsymbol{s}(t+\tau) = f^{\tau}(\boldsymbol{s}_t)$. While, $g_{K,m}$ is the corresponding evolution law in the candidate state space reconstruction $X_{K,m}$. Equation 2.6, is just the statement that when $\Psi_{K,m}$ results in an embedding, f and $g_{K,m}$ are related to each other by a continuous transformation. We can expand $\Psi_{K,m}$ on the LHS of 2.6 and substitute $\Psi_{K,m} = P_m \circ \Phi_K \circ h$. Additionally, we know that $\Psi_{K,m}(\boldsymbol{s}_t)$ is just the current state in the reconstructed state space, so we can substitute $\boldsymbol{x}_t = \Psi_{K,m}(\boldsymbol{s}_t)$ on the RHS. Equation 2.6 then becomes,

$$P_m \circ \Phi_K \circ h \circ f^{\tau}(\boldsymbol{s_t}) = g^{\tau}_{K,m}(\boldsymbol{x}_t) \tag{2.7}$$

Assuming for the moment that there is no observation noise, $h \circ f^{\tau}(\boldsymbol{s_t})$ is the measurement applied to state $\boldsymbol{s}(t+\tau)$, and corresponds to the time-series at time $t+\tau$. Thus, we can substitute $\boldsymbol{y}(t+\tau) = h \circ f^{\tau}(\boldsymbol{s_t})$ on the LHS of eqn 2.7 above. Next, we recognize that $P_m \circ \Phi_K$ is invertible, meaning that we can map a point in the reconstructed state space $\boldsymbol{x}_t$ back on to the time-series by $(P_m \circ \Phi_K)^{-1}\boldsymbol{x}_t$. This allow us to finally write

$$\begin{aligned} \boldsymbol{y}(t+\tau) &= (P_m \circ \Phi_K)^{-1} \circ g^{\tau}_{K,m}(\boldsymbol{x}_t) \\ &= \left(\Phi_K^{-1} \circ P_m^{-1}\right) \circ g^{\tau}_{K,m}(\boldsymbol{x}_t) \end{aligned} \tag{2.8}$$

Equation 2.8 is the main idea of this section, and states that for a candidate state space reconstruction to be an embedding, the observed value of the time-series at a future time $t+\tau$, must be equal to the value obtained by propagating the current state in a reconstruction $\boldsymbol{x}_t$ forward τsteps by $g^{\tau}_{K,m}$, and mapping it back to the time-series by $\left(\Phi_K^{-1} \circ P_m^{-1}\right)$. Simply put, eqn 2.8 says that when a candidate reconstruction is an embedding, it allows us to accurately predict the future of the time-series. This gives us a way to measure the "goodness" of a candidate state space without having access to the underlying state space. For parameter values where the reconstructed state space $X_{K,m}$ is not an embedding, say due to self-intersections or complex geometry, eqn 2.8 will not hold, and we will not be able to predict $\boldsymbol{y}(t+\tau)$ accurately.

Arguments presented above give us a way to find optimal values of K, and m simultaneously by minimizing the discrepancy between $\boldsymbol{y}(t+\tau)$ and $\left(\Phi_K^{-1} \circ P_m^{-1}\right) \circ g^{\tau}_{K,m}(\boldsymbol{x}_t)$. A commonly used measure of this discrepancy is the root mean squared error $E(\tau)$. If we let $\hat{\boldsymbol{y}}_{K,m}(t+\tau) = \left(\Phi_K^{-1} \circ P_m^{-1}\right) \circ g^{\tau}_{K,m}(\boldsymbol{x}_t)$, then the root mean squared error is written as $E_{K,m}(\tau) = \langle \|\boldsymbol{y}(t+\tau) - \hat{\boldsymbol{y}}_{K,m}(t+\tau)\|^2 \rangle_t^{1/2}$, where $\|\cdot\|$ is the Euclidean norm, and $\langle\cdot\rangle_t$ denotes averaging over different time instances.

We can relax the assumption of noise free measurements by considering the probability density function $p(\boldsymbol{y}(t+\tau)|\boldsymbol{x}_t)$, which is the object of direct relevance for noisy measurements. Under the assumption of noise distributed according to an isotropic Gaussian with variance η^2, $p(\boldsymbol{y}(t+\tau)|\boldsymbol{x}_t)$ is given by the following

$$p(\boldsymbol{y}(t+\tau)|\boldsymbol{x}_t) = \frac{1}{\left(\sqrt{2\pi}\right)^D \eta} \exp\left\{-\frac{1}{2\eta^2}\|\boldsymbol{y}(t+\tau) - \hat{\boldsymbol{y}}_{K,m}(t+\tau)\|^2\right\}$$

It can be shown that the maximum likelihood estimate of $\boldsymbol{y}(t+\tau)$ distributed as above is given by parameters that also minimize the root mean square, $E_{K,m}(\tau) = \langle \|\boldsymbol{y}(t+\tau) - \hat{\boldsymbol{y}}_{K,m}(t+\tau)\|^2 \rangle_t^{1/2}$. Thus, even in the presence of observation noise, the optimal strategy for identifying a good reconstruction is to find coordinates that enable best predictability as measured by $E_{K,m}(\tau)$.

Arguments similar to above, along with analytical calculations can be found Casdagli et al [38]. Others such as Judd and Mees have also argued for prediction as a method to find optimal embedding parameters [39, 40]. Indeed, even ostensibly geometric methods of parameter optimization such as false nearest neighbors can be interpreted as optimizing predictability [22].

To turn the above ideas into an operational algorithm for finding optimal embedding parameters, two key ingredients are necessary. First is the estimation of the evolution law of the reconstruction, which is used to map $\boldsymbol{x}_t$ to $\boldsymbol{x}_{t+\tau}$. Specifically, we want an approximation of $g_{K,m}^{\tau}$. While, second is the fact that $E_{K,m}(\tau)$ is not a number, but a function of prediction or forecast time τ. If we interpret $E_{K,m}(\tau)$ as a number, then it would seem that optimal embedding parameters depend on our choice of a particular time-scale τ. Indeed, others have argued that optimal embedding parameters depend on the intended application [39]. This is especially true when several different time-scales are present simultaneously. For example, an embedding that is optimal for modeling daily changes in weather might not be optimal for modeling seasonal or yearly variations. Similarly, in animal behavior, an embedding that is optimal for modeling instantaneous changes in posture might not be optimal for modeling large time-scale foraging dynamics. However, for attractors and other bounded systems, there is a natural time-scale τ_s beyond which the error $E_{k,m}$ stops changing. In the following I will show that a direct consequence of the existence of an upper bound on the value of $E_{k,m}$ allows us to estimate the predictability time, i.e. the time after which our predictions become useless. We will see that good embeddings maximize this predictability time, meaning that a good coordinate system is one which allows us to predict for as long as it is possible to predict. Before that however, I will address how to approximate the dynamics $g_{K,m}^{\tau}$ in the reconstructed state space.

2.8.1 Simple nearest neighbor predictor to predict the time-series

In principle there are several ways to model the dynamics $g_{K,m}^{\tau}$ with varying levels of complexity. However, because our goal here is to measure predictability rather than generating accurate predictions, it makes sense to choose the simplest method. Specifically, if $g_{K,m}^{\tau}$ exists and is continuous, then we can use the future of the nearest neighbor of $\boldsymbol{x}(t)$ to predict $\boldsymbol{x}(t+\tau)$. This is the nearest neighbor predictor, also known as Lorenz's "method of analogs" [41], and has been used previously to choose embedding parameters [39, 40]. We are essentially approximating $g_{K,m}^{\tau}$ by a locally constant model, as opposed to a local linear, or local polynomial model. Specifically, to predict $\boldsymbol{x}(t+\tau)$ we find the nearest neighbor of $\boldsymbol{x}(t)$ in state space, say $\boldsymbol{x}(t')$, and assume that $\boldsymbol{x}(t'+\tau)$ is the predicted value for $\boldsymbol{x}(t+\tau)$. To ensure that we only use

past points to predict the future state we enforce that $t' < t$. Formally, we write

$$\begin{aligned} \hat{\boldsymbol{y}}_{K,m}(t+\tau) &= \left(\Phi_K^{-1} \circ P_m^{-1}\right) \boldsymbol{x}\left(t'+\tau\right), \text{ where} \\ t' &= \underset{T}{\operatorname{argmin}}\{\|\boldsymbol{x}(t) - \boldsymbol{x}(T)\| : T < t\} \end{aligned} \tag{2.9}$$

Intuitively, this means that if $g^{\tau}_{K,m}$ is continuous, then two neighboring points in the state space are also close to each other when mapped back on the time-series, ensuring that $E_{K,m}(\tau)$ is small. The schematic in Fig. 2.9 illustrates these points further.

The nearest neighbor predictor provides an upper bound to the predictability of an embedding. In an embedding where the nearest neighbor prediction error is small, better modeling methods must also give small errors. But the biggest advantages of this method is it's simplicity, computational efficiency and the fact that there are no parameters to estimate. Once we have selected embedding then we can use other more advanced methods to generate better predictions.

We address the fact that $E_{K,m}(\tau)$ depends on forecast time τ. Most authors set $\tau = 1$ and minimize the one-step prediction error [39, 40, 42]. However, good one-step predictions do not necessarily imply good multi-step predictions. Indeed, for real systems $E_{K,m}(\tau)$ is a complex consequence of all the different factors that affect the predictability of a time-series, be it noise, amount of data, dynamical complexity such as Lyapunov exponents or mixing and ergodic properties. In the following I will define a new quantity that captures all the different affects and measures the overall predictive performance of a model.

2.8.2 A new measure of overall predictability

The first thing to note is that for a stationary system $E_{K,m}(\tau)$ is bounded from both below and above. For finite data, no prediction can be better than the minimum resolution of our data, denoted from here on by ϵ. Clearly, for a completely predictable, noiseless time-series, with a perfect model $E_{K,m}(\tau) = 0$. On the other hand, the worse we can do is guess randomly, i.e. set $\hat{\boldsymbol{y}}_{K,m}(t+\tau)$ to a random time point from past. If we let e_s denote the average error obtained by predicting randomly, then for a completely random sequence $E_{K,m}(\tau) = e_s$, and we have, $\epsilon \leqslant E_{K,m}(\tau) \leqslant e_s$. Geometrically, ϵ is related to the minimum distance between two points in the state space, while e_s is related to the size of the attractor. Thus, the bounds on the error are simply bounds on the pairwise distances in the reconstructed state space.

Broadly speaking, for most real systems which get more unpredictable with time, $E_{K,m}(\tau)$, starts with some minimum value, denoted by e_1 and grows with τ according to a nontrivial growth process, possibly involving multiple growth rates [41, 43]. For attractors and bounded systems $E_{K,m}(\tau)$ finally saturates to a value e_s. Saturation typically happens due to dissipation or energy constraints on the underlying dynamical system. For finite, noisy data, saturation means that error gets within ϵ of the saturation value e_s for some time τ_s and remains there for all $\tau > \tau_s$. Formally, the following is true $e_s - E_{K,m}(\tau) < \epsilon$, for all $\tau > \tau_s$. Fig. 2.10A shows a schematic error curve along with the bounds. The existence of a saturation level signifies the presence of an invariant geometrical object on which the trajectories are forced to lie for all

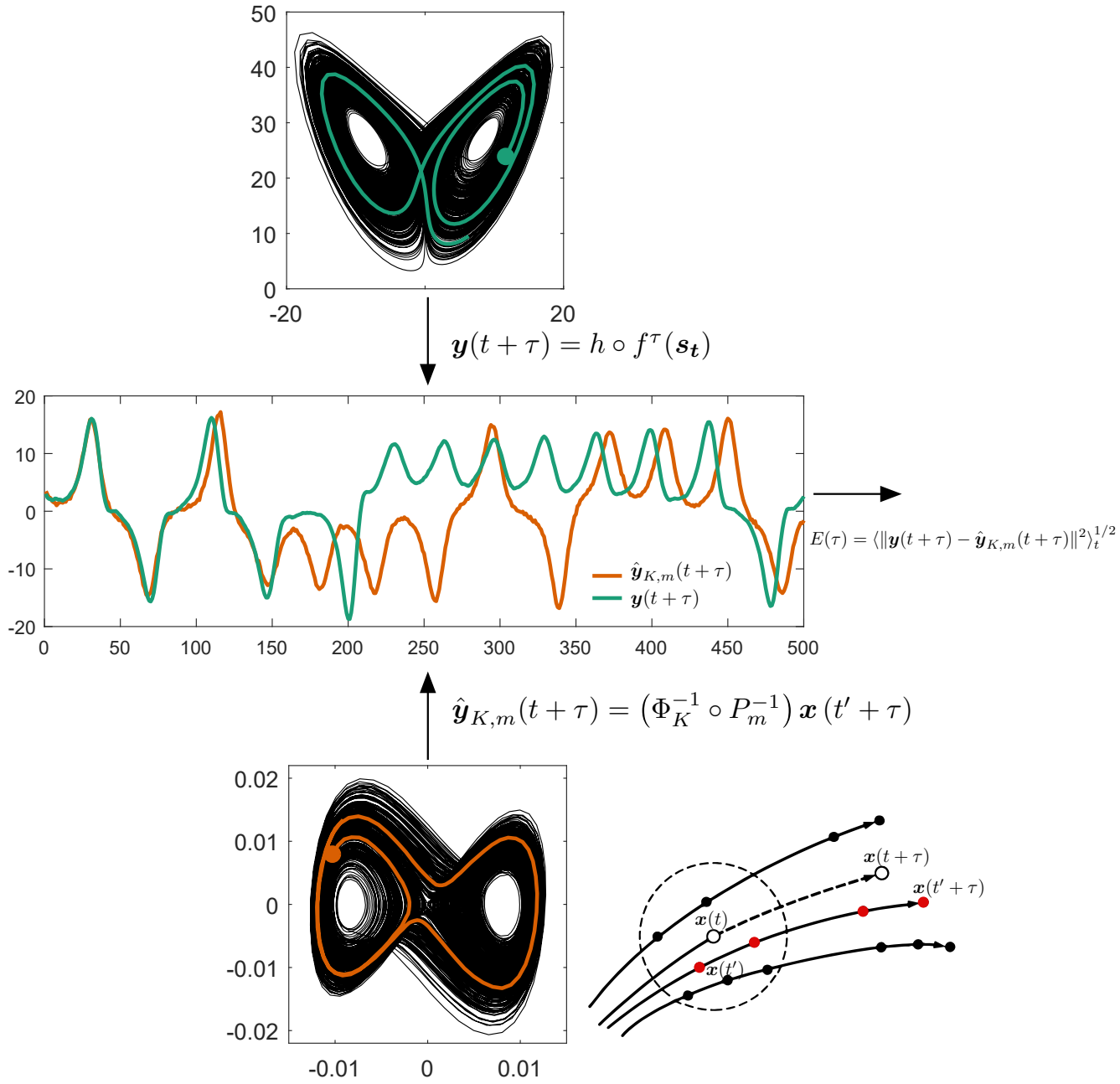
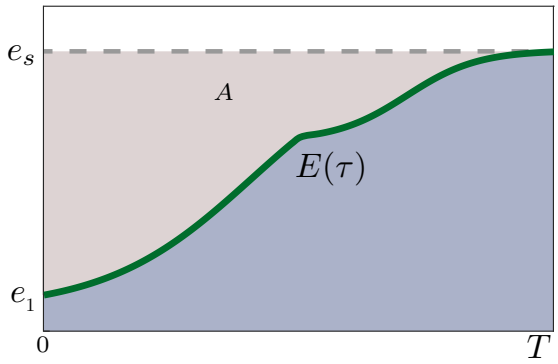


Figure 2.9: Predictability as criterion for choosing optimal embedding parameters. To test whether a set of embedding parameters lead to faithful copy of the dynamics, we would ideally like to compare it to the original state space. However, the true state of the dynamical system is only available to us in the form of a noisy time-series. In this case, the best we can do is use the observations as a proxy to underlying state space, and generate predictions of the time-series from the dynamics of different candidate state space reconstructions. Optimal embedding parameters will necessarily allow us to generate good predictions. We use the simplest possible predictor, which is to use the future of nearest neighbor as an estimate of the future value of the state. The current state $x(t)$ and its future $x(t+\tau)$ is denoted by open circles, and neighboring trajectory is denoted by red circles. The estimated value of $x(t+\tau)$ is the set to $x(t'+\tau)$



$$T_{pred} = \frac{A}{e_s} = \frac{1}{e_s}\int_0^{\infty}(e_s - E(\tau))d\tau$$

Figure 2.10: Measuring overall prediction performance. **A** For many real systems $E(\tau)$ starts from some small value, and grows with τ accordingto some growth process (possible with multiple scale dependent growth rates), before saturating to a level e_s. The growth process can in general be quite complicated and might not be simply characterized by a single number. I propose that the normalized area A bounded between e_u and $E(\tau)$ for $0 \leq \tau \leq \tau_s$ as a measure of the overall predictability.

$\tau > \tau_s$. For example, a purely stochastic process like Brownian motion is not bounded, consequently we will not expect $E_{K,m}(\tau)$ to saturate.

The one step error, e_1, measures the short term predictability and is determined by noise, amount of data, prediction method used and the distortion introduced by the embedding coordinates. Next, as mentioned earlier, the growth of $E_{K,m}(\tau)$ can be quite complicated and carries important information about the dynamics. For chaotic systems with one positive Lyapunov exponent, short term error growth is known to be exponential with a rate equal to the Lyapunov exponent [41, 44]. However, over long times it has been hypothesized by Lorenz and others that a logistic process describes the error growth best [41]. Indeed, as we will see later a logistic function fits $E_{K,m}(\tau)$ very well for systems with a single positive Lyapunov exponents. For chaotic systems with multiple Lyapunov exponents, $E_{K,m}(\tau)$ grows with different exponential rates in a scale dependent fashion [43, 45] before it saturates to e_s. Additionally, damped driven stochastic dynamics have been known to show power law growth [45]. Thus, in general the growth of $E_{K,m}(\tau)$ is largely determined by the nature of the underlying dynamics. A proper characterization of of the growth process can tell us a great deal about the underlying cause of unpredictability in a dynamical system. Finally, the saturation level e_s measures long time predictability and it's largely determined by the probability density of the attractor (also known as the invariant density for stationary or nearly stationary systems).

Thus, the precise shape of $E_{K,m}(\tau)$ is a result all the different factors that affect the predictability of a time-series. Despite the complicated nature of $E_{K,m}(\tau)$, most studies only look at the short time error properties by either studying e_1 or the local

error growth. However, an embedding should preserve both local and global properties. What is required then is an overall measure of predictability that takes into account various factors outlined above. Here, I claim that the following quantity

$$T_{pred} = \frac{A}{e_s} = \frac{1}{e_s}\int_0^\infty (e_s - E(\tau))d\tau \tag{2.10}$$

which is the normalized area bounded by e_s; (see Fig. 2.10) measures the predictability time of a dynamical system. The argument for introducing this quantity is the following. The total prediction error up to some time T is given by $\int_0^T E(\tau)d\tau$ or the area under the curve $E(\tau)$. We can see geometrically that for $T > \tau_s$, the integral is equal to the area of the rectangle e_sT minus the area A, which is given by $\int_0^\infty (e_s - E(\tau))d\tau$. Thus, the total prediction error is composed of an extensive term e_sT that scales with time T and a non-extensive term which equals the area bounded between the curve $E(\tau)$ and the level e_s. As the prediction time T gets larger, the contribution of the extensive term e_sT increases, hence e_s is a measure of the long term predictability. On the other hand, the non-extensive term determined by the area A governs short term predictability. The time at which these two are equal is the predictability time T_{pred} defined in eqn 2.10. It is important to point out that as defined, T_{pred} doesn't make any assumption about the nature of the dynamics (other than boundedness). It is fundamentally a measure of predictability of dynamics, and it can characterize chaotic, stochastic, limit cycle, or any combination of different effects that lead to unpredictability in a dynamical system. As we will see later, for chaotic systems with a single positive Lyapunov exponent λ, T_{pred} is trivially related to λ. In the following I will show the error curves for known dynamical systems, and then illustrate how different embedding parameters change the form of $E_{K,m}(\tau)$. In all the cases, T_{pred} as defined in eqn 2.10 correctly captures the overall behavior of $E_{K,m}(\tau)$.

2.8.3 $E(\tau)$ for canonical systems

In Fig. 2.11, I have plotted $E(\tau)$ for the Lorenz systems in its original coordinates without embedding any particular time-series. For clarity I have removed the subscripts K, m as we're dealing with the original state space here. It can be seen that the error curve $E(\tau)$ is well fit by the logistic function of the following form

$$E(\tau) = \epsilon + \frac{e_s - \epsilon}{1 + c\exp(-\lambda\tau)}$$

where $E(-\infty) = \epsilon$ is the lower asymptote corresponding to the minimum resolution, and $E(\infty) = e_s$ is saturation level. The saturation time τ_s is technically ∞. However, in practical settings we are interested in the time it takes for $E(\tau)$ to get within ϵ of e_s, which is well defined. The parameter c determines the value of e_1, and λ is the Lyapunov exponent. The logistic fit is consistent with previous claim by Lorenz that for simple chaotic systems with a single positive exponent the error growth is a logistic. Consequently, the Lyapunov exponent (and technically e_1) is all that's required to measure their predictability [41]. Next, I will describe, how different embedding parameters affect $E(\tau)$ for the Lorenz system. We will see that bad embedding param-

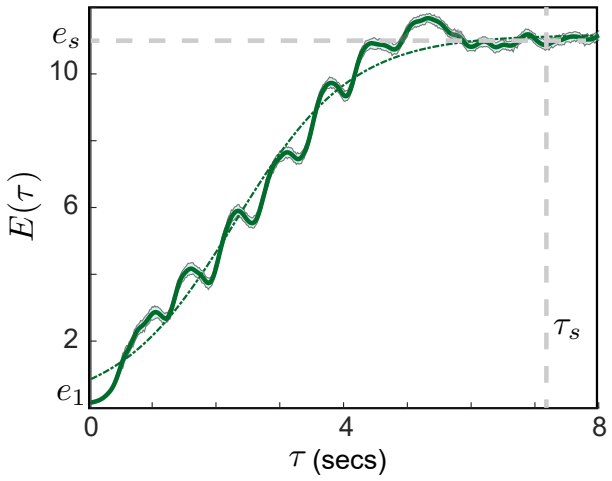


Figure 2.11: Prediction error for Lorenz system. $E(\tau)$ for the Lorenz system in original coordinates (s_1, s_2, s_3) has a logistic form with a growth rate given by the maximum exponent, which is consistent with reported literature.

eters lead to a loss in predictability and again T_{pred} provides an adequate measure of predictive performance.

2.8.4 Dependence of $E(\tau)$ on embedding parameters K and m

To see how different window lengths affect predictability I first embedded the time-series derived from the first component s_1 of the Lorenz system (with noise variance ξ^2=0.04) using three embedding windows $K = 3, K = 15, K = 13$, while keeping the embedding dimension fixed at $m = 3$. To study the effect of embedding dimension, I kept $K = 15$ and tested the following embedding dimension $m = 1, 2, 3, 6$. The estimated values of $E_{K,m}(\tau)$ are plotted in Fig. 2.12.

In Fig. 2.12 A, we see that for small window sizes (green curve), the loss of predictability happens by the way of rapid growth of the error. This means that the time it takes before predictions become comparable to random is significantly shorter compared to other window sizes. This is presumably because, as we saw earlier, embeddings with shorter window lengths are strongly affected by noise which leads to a rapid error. While, longer windows (blue curve) lead to a sharp increase in the short term error, including the one-step error e_1, but a slight decrease in the long term error as seen by the longer saturation time. The rate of growth for long windows is also smaller than that of small windows. All these observations imply that longer windows average over small scale dynamics important for short term prediction, with the upside of better long term prediction. Finally, moderately sized windows around $K = 15$ provide the best compromise, as they have best overall predictability. It can be visually seen that the area bounded from above captures all these effects, we shall soon see this reflected in the numerical estimates of T_{pred} across different window sizes.

The effect of embedding dimension is simpler compared to embedding window.

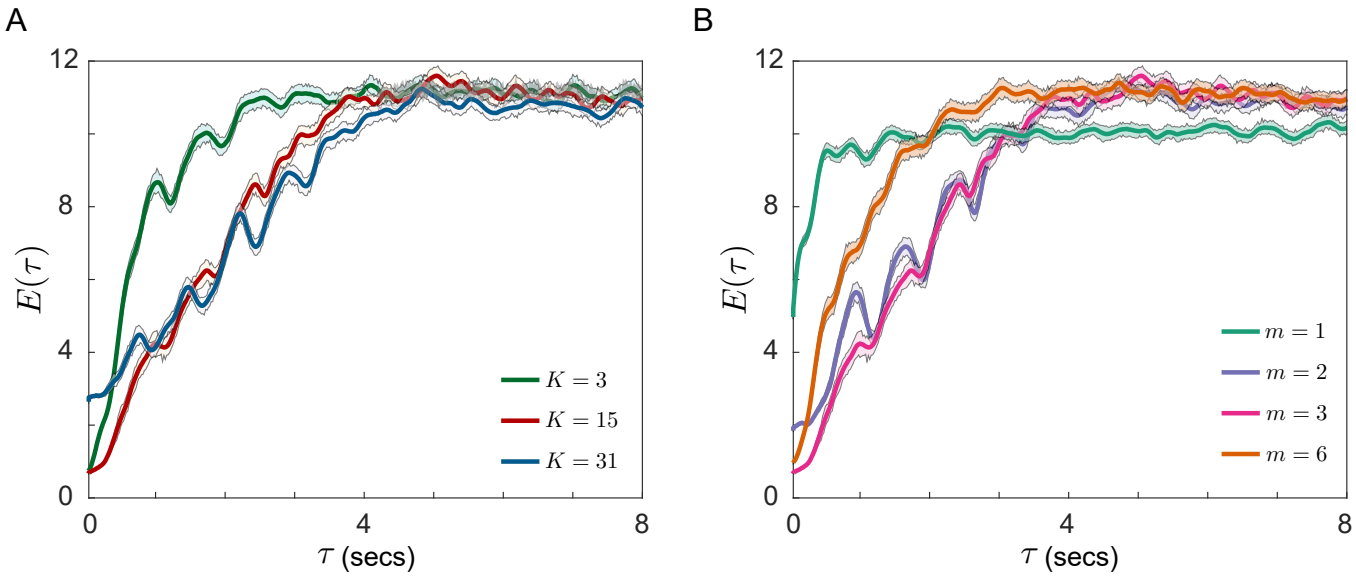


Figure 2.12: Suboptimal embedding parameters lead to loss in predictability **A.** $E(\tau)$ for the Lorenz system embedded by SSA with $K = 3, 15, 31$ and $m = 3$. For $K = 3$ the error starts small but grows at a rapid rate and saturates faster than for larger windows. For $K = 15$, e_1 remains small but the growth is more steady. While for $K = 31$, the one step error e_1 is large, but the long term performance is slightly better and the saturation time τ_s is larger compared to others. Thus larger windows allow better long term prediction at the cost of worse short term prediction. These differences are captured in the normalized area. **B.** $E(\tau)$ shown for $K = 15$ and $m = 2, 4, 6$. $m = 3$ gives the best overall predictive performance. In $m = 1$ dimensions, the attractor is folded tightly which causes a significant loss of predictability. While, once the attractor is embedded for $m \geq 3$, SSA coordinates are dominated by noise and lead to a loss of predictive performance.

Fig. 2.12B shows that the predictability drastically increases between $m = 1$ and $m = 2$ (green and blue curves respectively). It increases by a small amount for $m = 3$ (pink curve), and then decreases again for $m = 6$ (orange curve). This indicates that the best predictive performance is reached (for $K = 15$) in $m = 3$ dimensions. For $m = 1$ and $m = 2$ we note that most of the differences lie in the short term predictability, this is because when the system is not fully embedded, a fraction of nearest neighbors are not true neighbors but are a result of self intersections due to insufficient embedding [22]. The nearest neighbor predictor is sensitive to these so called false nearest neighbors, which leads to bad short term predictability. For large embedding dimensions such as $m = 6$, the error curve looks similar to that of small embedding windows in Fig. 2.12B ($K = 3$), suggesting that the loss of predictability happens for a similar reason. Indeed, that seems to be the case, as SSA coordinates in higher dimensions are dominated by noise, which leads to faster error growth.

2.8.5 Numerical estimation of $T_{pred|}$ to find optimal embedding parameters

The examples above hopefully make it clear the optimal embedding parameters lead to good overall predictability, and the area bounded between $E(\tau)$ and e_s which determines the predictability time T_{pred} is a valid characterization of overall predictive performance. In this section, I will estimate T_{pred} numerically to obtain optimal embedding parameters for SSA embeddings of the Lorenz and Duffing oscillator systems. The numerical estimation of T_{pred} is relatively straight forward as compared to other measures derived from state space reconstructions, such as Lyapunov exponents, dimensions and entropies. I exploit the fact that for $\tau > \tau_s$, $\int_0^T E(\tau)d\tau = e_s T - A$ to estimate both e_s and A at the same time. First $E(\tau)$ is estimated by the nearest neighbor method outlined above for a reasonably large value of τ (say, twice the autocorrelation time). I then numerically integrate $E(\tau)$. For $\tau > \tau_s$ this integral is given by the linear function $e_s T - A$, consequently, I fit a line to the last one tenth of the numerically estimated integral. The slope of the line estimates e_s while the intercept estimates the bounded area A. T_{pred} is then estimated by $\frac{A}{e_s}$.

Once we have a way to estimate T_{pred}, then the optimization of embedding parameters proceeds in two steps. First, I optimize the embedding window K by estimating T_{pred} for a range of K values in the embedding space described by $\overline{Y}_K$. Typically, there is a range of K values for which T_{pred} is maximized, any value of K within this range is appropriate. After choosing optimal K, denoted by K^* that maximizes T_{pred}, in the next step I perform SSA on $\overline{Y}^*_K$ and again estimate T_{pred} in the SSA coordinates as a function of embedding dimension. Once the attractor is fully embedded at the optimal embedding dimension m^*, T_{pred} either saturates or decreases on increasing the embedding dimension. It is possible to perform this optimization by genetic programming or some other nonlinear optimization method. This was done previously to optimize other predictive measures for choosing embedding parameters [39, 40]. However, in practice, the range of values for K and m is not prohibitively large, so that a brute force search is a reasonable approach.

I estimated T_{pred} for a range of values of K and m for the Lorenz. The result is shown in Fig. 2.13. In 2.13A, we see that T_{pred} rises with K reaching it's maximum value at $K = 20$, after which it declines gradually. The fact that there is a maximum for the window length means that there is an optimal range of values for K, beyond which the predictability gets worse. To estimate the embedding dimension I performed SSA with $K = 20$ and plotted T_{pred} as a function of m in 2.13B. It can be seen that for $m > 3$ there is a drastic reduction in predictability, implying that this a 3 dimensional system. It is notable that in terms of predictability $m = 2$ is also a good embedding, this is because it is known that the dimension of the Lorenz system is approximately 2.

2.9 Conclusion

In this chapter I discussed the state space reconstruction problem in detail. I described the general mathematical ideas behind the process, previous attempts at reconstruction and their drawbacks. Finally, I proposed a new method of estimating optimal reconstruction parameters that were shown to result in a good reconstruction of model systems. In chapter 4 I will show the utility of these methods in the reconstruction of *C. elegans* locomotory dynamics.

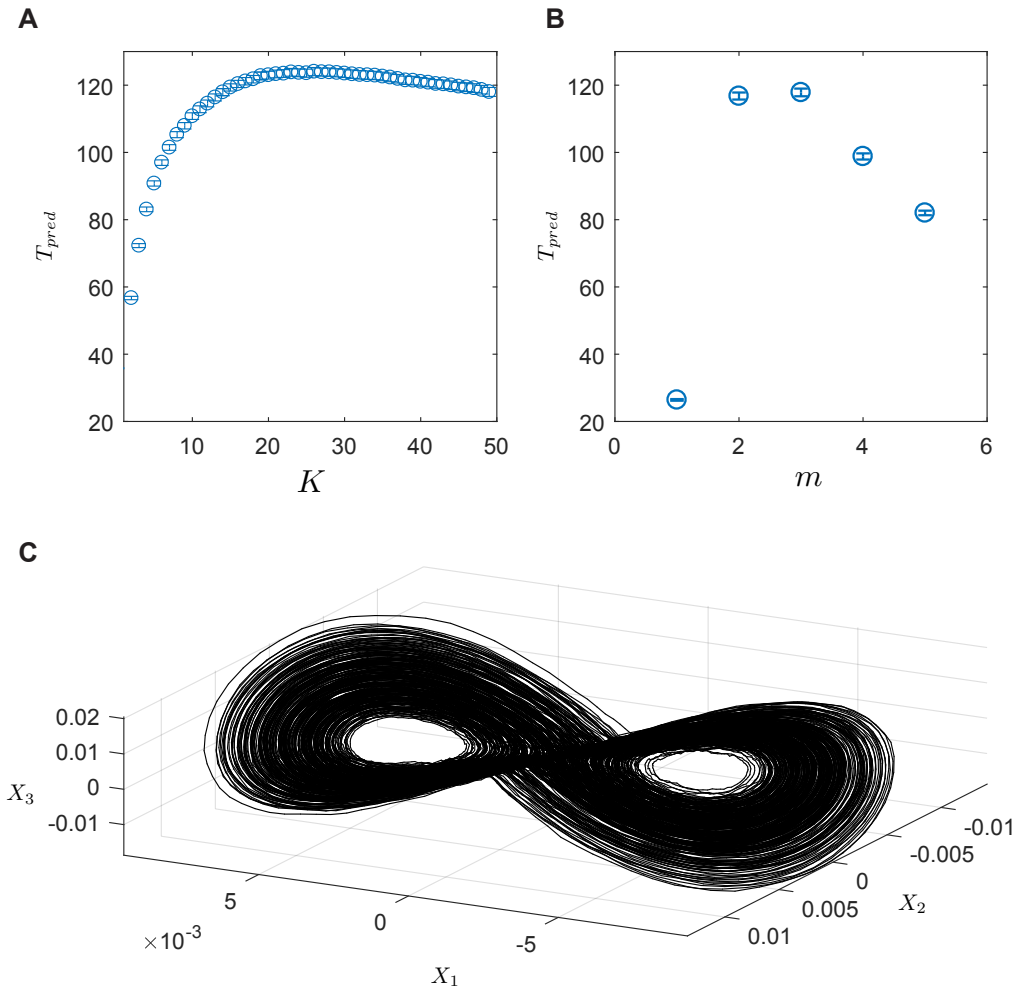


Figure 2.13: Optimization of reconstruction parameters for the Lorenz system. **A.** Estimated T_{pred} for different values of K in the embedding $\overline{Y}_K$. We see that at about 20 samples T_{pred} reaches its maximum value, after which it gradually decreases. **B.** T_{pred} in the SSA embedding at $K = 20$ for different values of the embedding dimension shows that a 3 dimensional embedding optimizes predictability of the Lorenz system. It is notable that $m = 2$ also seems a good embedding. This is perhaps because even though it takes $m = 3$ Euclidean dimensions to embed the Lorenz attractor, it is mostly a 2D object. **B.** The attractor reconstruction for $K = 20$ and $m = 3$ shows a good reconstruction, balanced between the effects of noise and strong curvature.

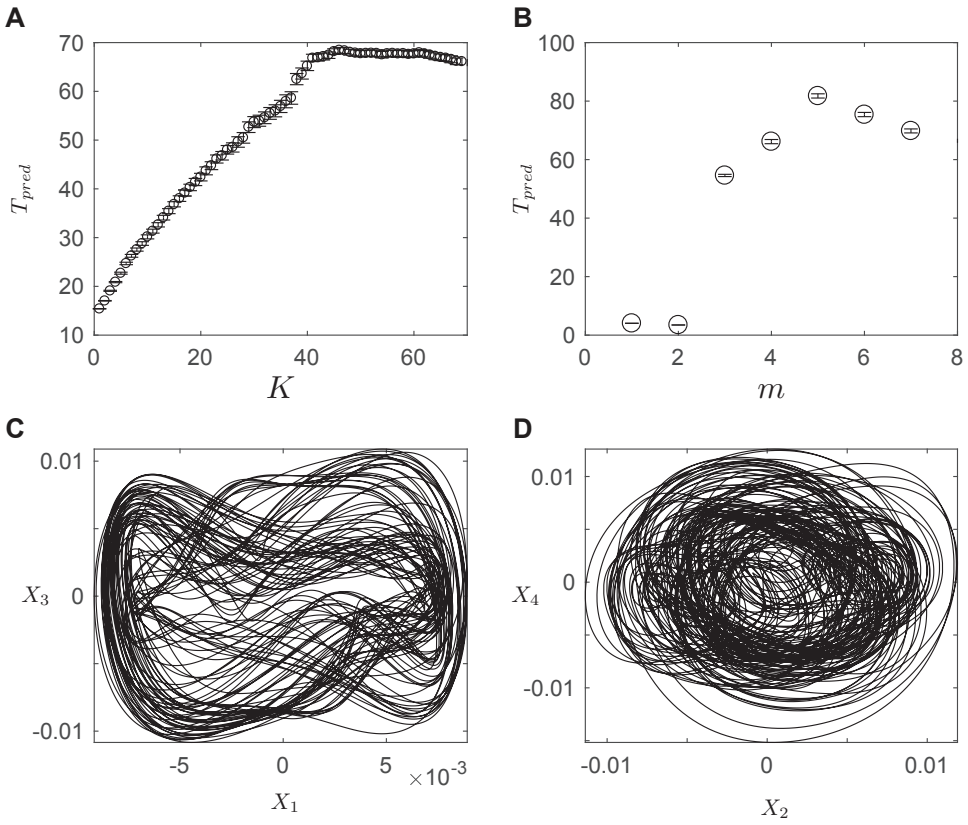


Figure 2.14: Optimization of reconstruction parameters for the coupled Duffing oscillators. **A.** Estimated T_{pred} for different values of K in the embedding $\overline{Y}_K$ suggests that K between 40 to 60 samples would make an optimal embedding. **B.** Estimates of T_{pred} as a function of embedding dimension show that $m = 5$ dimensions make an optimal embedding. **C,D.** 2D projections of the reconstructed attractor with $K = 40$ and $m = 5$ (X_5 is not shown).

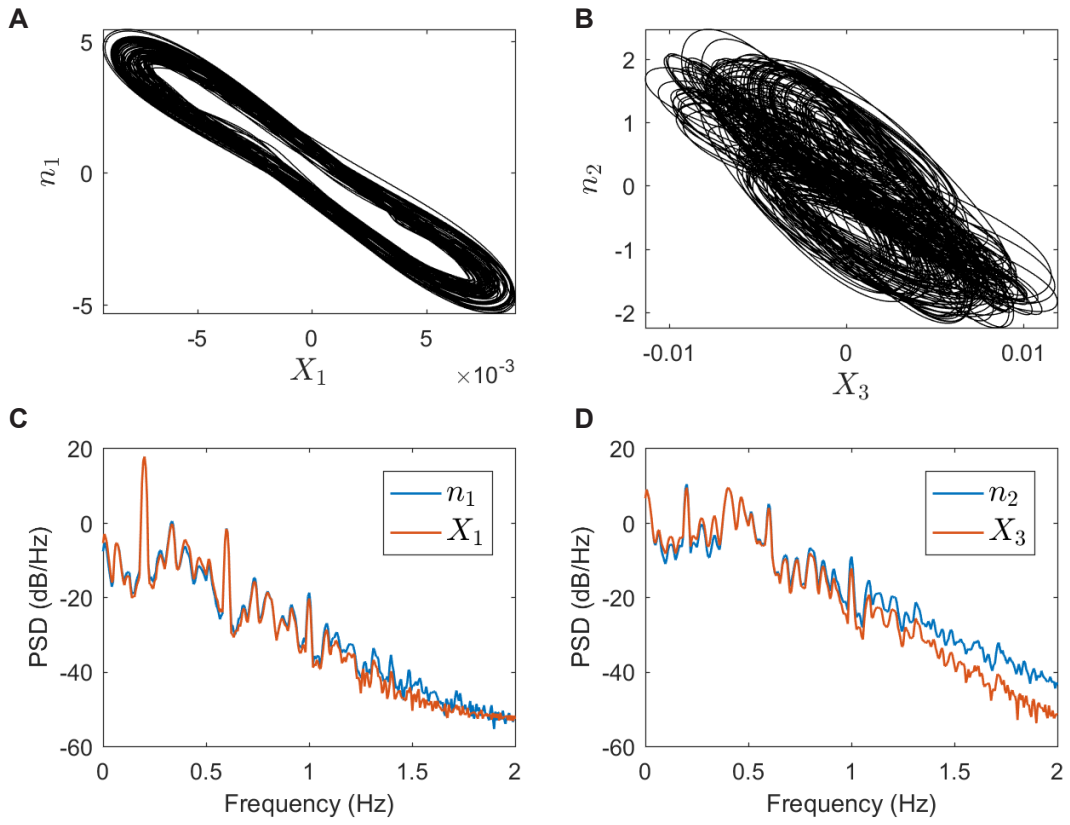


Figure 2.15: SSA embedding method approximates the linear normal modes for a nonlinear system. **A.,B.** The reconstructed coordinates X_1 and X_3 compared with the linear normal modes $n_1 = q_1 + q_3$ and $n_2 = q_1 - q_3$. There is a strong linear correlation between the two implying that SSA reconstruction approximate the linear modes of a nonlinear dynamical system. **C.,D.** Comparison of the PSD of the reconstructed coordinates X_1 and X_3 vs the linear normal modes. The relationship between SSA reconstruction and linear modes is even more clearly visible here.

Chapter 3

Lyapunov Exponents: Theory and Practice

Lyapunov exponents are a generalization of eigenvalues for non-linear dynamical systems. Unsurprisingly, they are as central to the study of non-linear dynamics as eigenvalues are for linear dynamical systems. As any central concept, Lyapunov exponents can be motivated in several different ways. A simple way of seeing where Lyapunov exponents come from is to ask, is the evolution law F in $x(t+1) = F(x(t))$ differentiable? And if so what are its consequences? The most natural consequence of differentiable dynamics is that no matter how complicated the nonlinearity in F, the dynamics in any neighborhood of $x(t)$ are described by a linear dynamical system described by the derivative of F at $x(t)$ given by the Jacobian matrix. As a consequence of chain rule, as we move along the trajectory, each of the local Jacobians are multiplied, their product describing the entire trajectory. This fact directly leads to two fundamental questions: Can we define a (geometric) average of the product of Jacobians? And, is it well behaved, in particular, under what conditions does it converge as we consider longer and longer trajectories? Answers to these questions lead to some deep insights, and naturally define the concept of Lyapunov exponents (among others). We will see that the Lyapunov exponents reveal a lot more about the nature of the dynamics than we could've asked for. They measure things like predictability time, entropy rate, dimension, dissipation rate, symmetry properties, ergodic properties, and a lot more. It is somewhat of a miracle that something as simple as a differential structure on the dynamics can give us so much! In the following I have tried to simplify the theory and the algorithms to estimate Lyapunov exponents as much as possible.

In this chapter I will present the theory and practice of the estimation of Lyapunov exponents. I will propose a new algorithm for stable estimation of all the exponents and apply it to known systems.

3.1 Lyapunov Exponents: Theory

To start things simply, I will first present a geometric description of the Lyapunov exponents and discuss the idea that they can be obtained from a geometric mean of the derivatives of F along a trajectory. Next, I will show how in 1D this definition

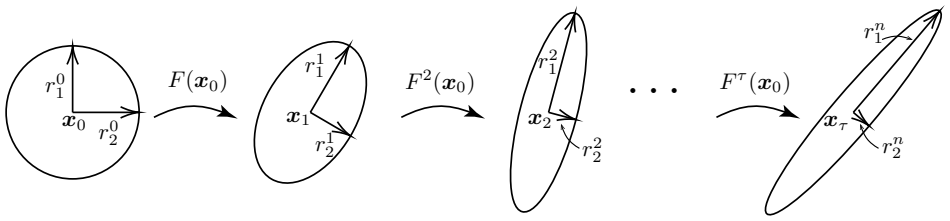


Figure 3.1: A unit ball gets sheared under the action of the linearization of F^τ

leads to first converting the geometric average to an arithmetic average by application of logarithms and then applying the additive ergodic theorem. In the section after that I will show how to make the geometric mean work in dimensions more than 1, despite the complications that arise due to the non-commuting nature of matrix multiplication. The logical structure and organization of the following presentation here is mine, but the details follow closely the review by Eckmann and Ruelle [46] which is still a standard text in the field. The algorithmic details are integrated from [46], [47], [48] and [49].

3.1.1 Geometric definition

Consider an infinitesimal ball of radius $r_1^0 = r_2^0 = \epsilon$, centered at $\boldsymbol{x}_0$ as depicted schematically in Fig. 3.1 above. After one step, the ball gets distorted under the action of F. However, if the ball is infinitesimal, and the F is differentiable, then the distortion is governed by the linearization (derivative/Jacobian) of F at $\boldsymbol{x}_0$. Consequently, under the linear dynamics, the ball is distorted to an ellipsoid. Accordingly, it expands along the long axes of the ellipsoid and contracts along the short axes. After τ steps, the starting ball evolves into a longer and thinner ellipsoid. The per step change of the lengths of the principle axes of the ellipsoid are called (local) Lyapunov exponents for the trajectory starting at x_0 and going for *tau* steps. They measure the average rates of expansion or contraction along a trajectory.

For an m dimensional state space, there are clearly m exponents measuring the rates of separation along m orthogonal directions, denoted by $\Lambda_k(x_0, \tau)$ for $1 \leqslant k \leqslant m$. There is a natural ordering to the exponents, such that the first exponent ($k = 2$) corresponds to the longest principle axis of the final ellipsoid, i.e. direction along which the separation between nearby points is the greatest. The second exponent ($k = 1$) corresponds to the second longest principle axes and so on. The ordered collection is known as a spectrum of local Lyapunov exponents, and each exponent is defined according to

$$\Lambda_k(x_0, \tau) = \lim_{\epsilon \to 0} \frac{1}{\tau} \log \frac{r_k^\tau}{\epsilon} \tag{3.1}$$

The exponents above are defined locally, i.e. they depend on the state space point x_0, and the duration of the trajectory τ. When the system governed by F is stationary on a limit set with an invariant measure, i.e. it has settled on an attractor, then the

infinite time limit

$$\Lambda_k = \lim_{\tau\to\infty} \lim_{\epsilon\to 0} \frac{1}{\tau} \log \frac{r_k^\tau}{\epsilon} \tag{3.2}$$

exists [50], and is independent of the initial condition x_0 on the attractor. It is important to note that this independence is with respect to the attractor, and not the entire state space. Indeed, when there are several co-existing attractors in the same state space, then generically speaking, each attractor has a different set of Lyapunov exponents. The limiting exponents Λ_k are known as global Lyapunov exponents. In this way the Lyapunov exponents measure the long term expansion and contraction rates of a dynamical system.

It is common to interpret the Lyapunov exponents as characterizing the stability of the system. By stability, we usually mean the response of the system to an infinitesimal perturbation of size ϵ. When the Lyapunov exponent is positive in a particular direction, it means that the system is unstable, causing any perturbations to grow exponentially at the rate given by the maximum Lyapunov exponent. On the other hand, a negative exponent means that the system is stable and perturbations decay.

Stability is not the only interpretation of Lyapunov exponents, as they are also related to the average predictability of the system. If we consider the initial ball around $\boldsymbol{x}_0$ in Fig. 3.1 as possible predictions for the initial state $\boldsymbol{x}_0$, with the size corresponding to the initial error in prediction, then the rate of expansion and contraction is related directly to how the error in our predictions change under the action of the dynamics. When the dynamics are predictable, the sphere would either remain the same size as τ increases (errors remain constant under the action of F), or it will shrink (errors would decrease with time, like in a limit cycle). However, for unpredictable systems, there will at least be one direction along which the ellipse will expand, and the prediction errors will get worse with time. It can be shown that the sum of all expansion rates (i.e. the sum of positive Lyapunov exponents) is related to the entropy rate of the system which measures the overall unpredictability. Another related quantity, the rate of dissipation is given by the rate of change of phase space volume is given by the sum of all Lyapunov exponents. The global Lyapunov exponents are also related to the dimension of the attractor, which give information about the number of active degrees of freedom in the physical system under study. Finally, the global Lyapunov exponents are independent of the initial condition $\boldsymbol{x}_0$, and are also topological invariants. Meaning that they are independent of any smooth coordinate transformation and can be measured in any (smooth) coordinate transformation of the original dynamics.

In the next section, I will describe the ideas in more detail in the setting of a 1D dynamical system where the concept is remarkably straightforward.

3.1.2 Lyapunov exponents in 1D

Dynamical systems in discrete time are known as maps, they are described by recurrence relations of the form $x_{t+1} = F(x_t)$. This is in contrast to continuous time dynamical systems, called flows, which are described by ODEs or PDEs. A major difference between flows and maps is that unlike flows which need a 3D state space to show chaos, maps can be chaotic even in a 1D state space.

Figure 3.2: **A.** Schematic of iterates of a chaotic dynamical system in a 1D state space. **B.** One step evolution of the displacement between two neighbors in state space.

Consider a 1D differentiable dynamical system in discrete time described by $x_{t+1} = F(x_t)$ with state space being the real line bounded between some limits $[a, b]$. Example iterations are of such a system are shown in Fig. 3.2.

When the system is chaotic, the points hop around in the 1D state space and never repeat themselves as shown schematically in Fig. 3.2A. Now consider a point x_t, and another point, $\tilde{x}_t$, in its neighborhood at a distance δx_t, i.e. $\tilde{x}_t = x_t + \delta x_t$ (Fig. 3.2B). The next iterate of $\tilde{x}_t$ is then given by

$$\tilde{x}_{t+1} = F\left(\tilde{x}_t\right) = F(x_t + \delta x_t)$$

Since F is differentiable, we can perform a Taylor expansion of $F(x_t + \delta x_t)$ about x_t as follows

$$\begin{aligned} \tilde{x}_{t+1} &\approx F(x_t) + F'(x_t)\ \delta x_t \\ \tilde{x}_{t+1} - x_{t+1} &= F'(x_t)\ \delta x_t \\ \delta x_{t+1} &= F'(x_t)\ \delta x_t \end{aligned} \tag{3.3}$$

where $F'(x_t) = \left.\dfrac{dF}{dx}\right|_{x_t}$ is the derivative of the map evaluated at x_t. The final equation of eqn 3.3 above says that the derivative of the map governs how small perturbations in the neighborhood of x_t change over time. Say we start from an initial condition x_0 and a neighboring point $\tilde{x}_0 = x_0 + \delta x_0$, then the displacement will grow as

$$\begin{aligned} \delta x_1 &= F'(x_0)\ \delta x_0 \\ \delta x_2 &= F'(x_1)\ \delta x_1 = F'(x_1)F'(x_0)\ \delta x_0 \\ \delta x_3 &= F'(x_2)\ \delta x_2 = F'(x_2)F'(x_1)F'(x_0)\ \delta x_0 \\ &\vdots \\ \delta x_t &= F'(x_t)\ \delta x_t = F'(x_{t-1})F'(x_{t-1})\ldots F'(x_0)\ \delta x_0 \\ &= \prod_{k=0}^{t-1} F'(x_t)\ \delta x_0 \end{aligned}$$

Thus, in order to study the properties of a long trajectory, it is necessary to study the product of derivatives along the trajectory, $\prod_{k=0}^{t-1} F'(x_t)$. Because of the multiplicative structure, it is appropriate to consider the geometric mean of the magnitude of the

derivatives along the trajectory. The geometric mean is called the local Lyapunov number $L(x_0, t)$ of the trajectory and is defined as

$$L(x_0, t) = \left(\prod_{k=0}^{t-1} |F'(x_t)| \right)^{1/t} \tag{3.4}$$

The absolute value is taken because the derivatives can change sign along the trajectory, however, we're only interested in their magnitude. The local Lyapunov exponents $\lambda(x_0, t)$, are simply the logarithms of the Lyapunov numbers, i.e.

$$\begin{aligned} \lambda(x_0, t) &= \log L(x_0, t) \\ &= \frac{1}{t} \log \prod_{k=0}^{t-1} |F'(x_t)| \\ &= \frac{1}{t} \sum_{k=0}^{t-1} \log \ |F'(x_t)| \end{aligned}$$

it can be easily verified that

$$\lambda(x_0, t) = \frac{1}{t} \log \ \left| \frac{\delta x_t}{\delta x_0} \right|$$

implying that $|\delta x_t| = |\delta x_0| \ \exp(t\lambda(x_0, t))$. Thus, the local Lyapunov exponents measure the average rate of expansion along a trajectory starting at x_0, and this definition is consistent with the geometric definition given above.

As before we're also interested in the long term properties of the trajectory, so we can consider the limit of $\lambda(x_0, t)$ as $t \to \infty$, i.e. $\lim_{t\to\infty} \lambda(x_0, t)$. A natural question then is when does this limit exist? This seemingly benign question is related to several deep issues that lead to major developments in the theory of dynamical systems, and resulted in the birth of at least two scientific disciplines, that of statistical mechanics in physics and ergodic theory in mathematics. The short answer that is relevant here comes from GD Birkhoff's celebrated ergodic theorem (also called additive ergodic theorem) [51, 52] which asserts that when $x(t)$ is a stationary process, i.e. the dynamics settle on an invariant set as $t \to \infty$, with an invariant probability distribution; then not only does the limit exists, but it can be replaced by an average over the probability distribution on the invariant set (so called invariant measure), i.e.

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t-1} \log \ |F'(x_t)| = \langle \log |F'(x)| \rangle$$

where the expectation is taken over the invariant measure. In this way the additive ergodic theorem generalizes the (strong) law of large numbers for non i.i.d. (independent and identically distributed) random numbers. One of the remarkable consequences of above is that the global Lyapunov exponent λ is a topological invariant, meaning that it does not depend on our choice of coordinate system and is available from an experimental reconstruction of the dynamical system.

3.1.3 Lyapunov exponents in higher dimensions

The basic idea of Lyapunov exponents in higher dimensions is the same as in 1D: they are related to the multiplicative average of the derivatives along a trajectory. However, there are several difficulties before we can make this idea workable. For a multidimensional system, $\boldsymbol{x}_{t+1} = F(\boldsymbol{x}_t)$, derivatives are matrices containing all the partial derivatives of each component of F with respect to each component of $\boldsymbol{x}_t$. As a result of matrix multiplication being non-commuting, the straightforward arguments employed above to define the Lyapunov exponent in 1D cannot be used directly. Substantial amount of work is required first to define multivariate Lyapunov exponents and then prove that they are well defined as an infinite time average is taken. Additional mathematical machinery is then required to develop an algorithm to estimate them from known dynamical systems, and finally, even more work is needed to estimate them in a reconstructed state space from experimental data. In the following, I will go over the definitions of local and global Lyapunov exponents in more than 1 dimensions, describe their properties (with proofs wherever possible), and finally detail the algorithms for computing them from data.

As an example to help illustrate the ideas, consider a 2D state space with the current state $\boldsymbol{x}_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix}$ and the evolution law as $F(\boldsymbol{x}_t) = \begin{pmatrix} f(u_t, v_t) \\ g(u_t, v_t) \end{pmatrix}$. The next state is then determined by,

$$\begin{aligned} \boldsymbol{x}_{t+1} &= F(\boldsymbol{x}_t) \\ \begin{pmatrix} u_{t+1} \\ v_{t+1} \end{pmatrix} &= \begin{pmatrix} f(u_t, v_t) \\ g(u_t, v_t) \end{pmatrix} \end{aligned}$$

Consider now a point in the neighborhood of $\boldsymbol{x}_t$, $\tilde{\boldsymbol{x}}_t = \boldsymbol{x}_t + \boldsymbol{\delta x}_t = \begin{pmatrix} u_t + \delta u_t \\ v_t + \delta v_t \end{pmatrix}$, where $\boldsymbol{\delta x}_t = \begin{pmatrix} \delta u_t \\ \delta v_t \end{pmatrix}$ is a displacement vector, more commonly known as a tangent vector at $\boldsymbol{x}_t$. The space of all possible infinitesimal displacements is called the tangent space $\boldsymbol{x}_t$. The next iterate of the neighboring state$\tilde{\boldsymbol{x}}_t$ is

$$\begin{aligned} \tilde{\boldsymbol{x}}_{t+1} &= F\left(\tilde{\boldsymbol{x}}_t\right) = F(\boldsymbol{x}_t + \boldsymbol{\delta x}_t) \\ \begin{pmatrix} u_{t+1} + \delta u_{t+1} \\ v_{t+1} + \delta v_{t+1} \end{pmatrix} &= \begin{pmatrix} f(u_t + \delta u_t, v_t + \delta v_t) \\ g(u_t + \delta u_t, v_t + \delta v_t) \end{pmatrix} \end{aligned}$$

performing a Taylor expansion as before

$$\begin{aligned} \begin{pmatrix} u_{t+1} + \delta u_{t+1} \\ v_{t+1} + \delta v_{t+1} \end{pmatrix} &= \begin{pmatrix} f(u_t, v_t) + \left.\frac{\partial f}{\partial u}\right|_{\boldsymbol{x}_t} \delta u_t + \left.\frac{\partial f}{\partial v}\right|_{\boldsymbol{x}_t} \delta v_t \\ g(u_t, v_t) + \left.\frac{\partial g}{\partial u}\right|_{\boldsymbol{x}_t} \delta u_t + \left.\frac{\partial g}{\partial v}\right|_{\boldsymbol{x}_t} \delta v_t \end{pmatrix} \\ \begin{pmatrix} u_{t+1} + \delta u_{t+1} \\ v_{t+1} + \delta v_{t+1} \end{pmatrix} &= \begin{pmatrix} u_{t+1} \\ v_{t+1} \end{pmatrix} + \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix}_{\boldsymbol{x}_t} \begin{pmatrix} \delta u_t \\ \delta v_t \end{pmatrix} \\ \begin{pmatrix} \delta u_{t+1} \\ \delta v_{t+1} \end{pmatrix} &= \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix}_{\boldsymbol{x}_t} \begin{pmatrix} \delta u_t \\ \delta v_t \end{pmatrix} \\ \boldsymbol{\delta x}_{t+1} &= \mathbf{J}(\boldsymbol{x}_t) \boldsymbol{\delta x}_t \end{aligned} \tag{3.5}$$

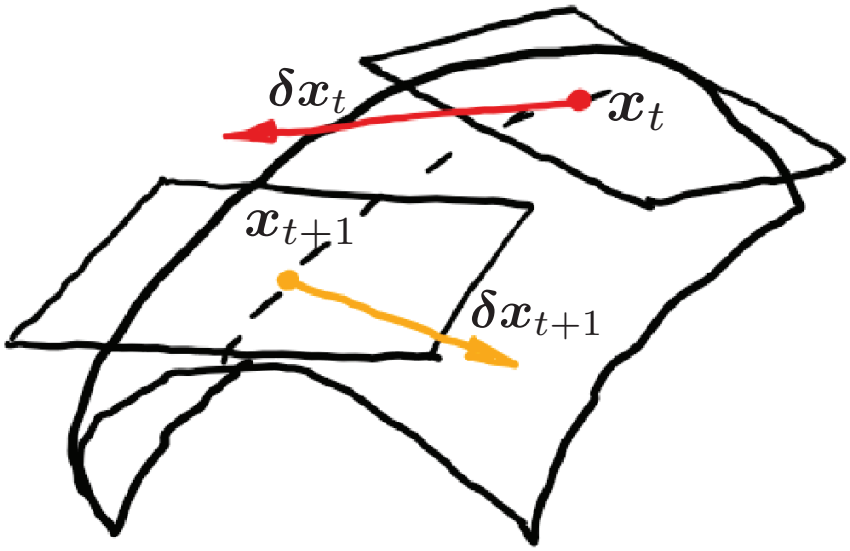


Figure 3.3: A. Schematic showing the evolution of a tangent vector under the action of linear dynamics around $\boldsymbol{x}_t$.

, where $\mathbf{J}(\boldsymbol{x}_t)$ is the matrix of all partial derivatives evaluated at $\boldsymbol{x}_t$, and is known as the Jacobian matrix it is synonymous with the derivative for multivariate functions. The last equation of eqn 3.5 although derived for 2D is valid in any dimensions, and is known as the tangent space dynamics (see Fig. 3.3). From here on we will leave the 2D example and work in an m dimensional state space. The tangent space dynamics in eqn above describes how small displacements, or tangent vectors at $\boldsymbol{x}_t$ evolve under the action of the Jacobian at $\boldsymbol{x}_t$. When the equations of motion are known, the Jacobian is easily estimated from the partial derivatives of F. When we do not know the dynamics, it can be estimated by inverting the equation above as we will discuss later.

Just as derivatives along a trajectory in 1D had multiplicative structure, the Jacobians along a trajectory in higher dimensions also have a multiplicative structure. To see this, we start with a tangent vector $\boldsymbol{\delta}_0$ separating $\boldsymbol{x}_0$ and a neighboring point $\tilde{\boldsymbol{x}}_0$. Following the tangent space dynamics, we have,

$$
\begin{aligned}
\boldsymbol{\delta}_1 &= \mathbf{J}(\boldsymbol{x}_0)\boldsymbol{\delta}_0 \\
\boldsymbol{\delta}_2 &= \mathbf{J}(\boldsymbol{x}_1)\boldsymbol{\delta}_1 = \mathbf{J}(\boldsymbol{x}_1)\mathbf{J}(\boldsymbol{x}_0)\boldsymbol{\delta}_0 \\
\boldsymbol{\delta}_3 &= \mathbf{J}(\boldsymbol{x}_2)\boldsymbol{\delta}_2 = \mathbf{J}(\boldsymbol{x}_2)\mathbf{J}\ (\boldsymbol{x}_1)\ \mathbf{J}(\boldsymbol{x}_0)\boldsymbol{\delta}_0 \\
&\vdots \\
\boldsymbol{\delta}_t &= \mathbf{J}(\boldsymbol{x}_{t-1})\boldsymbol{\delta}_{t-1} = \mathbf{J}(\boldsymbol{x}_{t-1})\mathbf{J}\ (\boldsymbol{x}_{t-2}) \ldots \mathbf{J}(\boldsymbol{x}_0)\boldsymbol{\delta}_0 \\
&= \prod_{i=0}^{t-1} \mathbf{J}(\boldsymbol{x}_i)\boldsymbol{\delta}_0 \\
&= \mathbf{J}^t(\boldsymbol{x}_0)\boldsymbol{\delta}_0
\end{aligned}
$$

, where we have denoted by $\mathbf{J}^t(\boldsymbol{x}_0)$, the product of local Jacobians along the trajectory

starting from $\boldsymbol{x}_0$ and going for t steps in future. In case of 1D, the sign of the displacement δ_t changed irregularly, and we considered its absolute value. Similarly, in higher dimensions, the direction of the displacement vector $\boldsymbol{\delta}_t$ changes irregularly under the action of the local Jacobian, so we need to consider its magnitude given by the squared length of the vector $\boldsymbol{\delta}_t$, i.e.

$$\|\boldsymbol{\delta}_\tau\|^2 = \boldsymbol{\delta}_\tau^T \cdot \left[\mathbf{J}^t(\boldsymbol{x}_0)\right]^T \cdot \mathbf{J}^t(\boldsymbol{x}_0) \cdot \boldsymbol{\delta}_t$$

where $[\cdot]^T$ denotes matrix transpose. The key object that governs the magnitude of displacement is the matrix $\left[\mathbf{J}^t(\boldsymbol{x}_0)\right]^T \cdot \mathbf{J}^t(\boldsymbol{x}_0)$, which encodes the correlations of the derivatives along a trajectory in phase space. We can now define the following geometric mean composed of a product of $2t$ matrices, to study the average per-step change in the magnitude along a trajectory,

$$\mathbf{OSL}(\boldsymbol{x}_0, t) = \left(\left[\mathbf{J}^t(\boldsymbol{x}_0)\right]^T \cdot \mathbf{J}^t(\boldsymbol{x}_0)\right)^{\frac{1}{2t}}$$

In an m dimensional state space, $\mathbf{OSL}(\boldsymbol{x}_0, t)$ has m positive eigenvalues $L_i(\boldsymbol{x}_0, t)$ for $1 \leqslant i \leqslant m$, and are called the local Lyapunov numbers of the trajectory. The logarithms of the Lyapunov numbers,

$$\Lambda_i(\boldsymbol{x}_0, t) = \log L_i(\boldsymbol{x}_0, t)$$

are known as the local Lyapunov exponents and their ordered set, $\Lambda_1(\boldsymbol{x}_0, t) \geqslant \Lambda_2(\boldsymbol{x}_0, t) \geqslant \ldots \Lambda_m(\boldsymbol{x}_0, t)$ is called the local Lyapunov spectrum. Note that this definition corresponds exactly to the geometric definition of the stretching rate, as the eigenvalues of the matrix $\left[\mathbf{J}^t(\boldsymbol{x}_0)\right]^T \cdot \mathbf{J}^t(\boldsymbol{x}_0)$ are the singular values of $\mathbf{J}^t(\boldsymbol{x}_0)$, and measure the lengths of the principle axis of the ellipsoid described by the action of $\mathbf{J}^t(\boldsymbol{x}_0)$ on a unit sphere.

To make the transition from a local to a global description, we need to consider the infinite time limit of $\mathbf{OSL}(\boldsymbol{x}_0, t)$, i.e.

$$\lim_{t\to\infty} \mathbf{OSL}(\boldsymbol{x}_0, t) = \lim_{t\to\infty} \left(\left[\mathbf{J}^t(\boldsymbol{x}_0)\right]^T \cdot \mathbf{J}^t(\boldsymbol{x}_0)\right)^{\frac{1}{2t}}$$

In 1D, derivatives are scalar, and we dealt with a product of real numbers by applying logarithms turn into a sum. Consequently, the geometric mean was replaced by an arithmetic mean of the logarithms of the derivatives along a trajectory. As a result, we were able to apply Birkhoff's additive ergodic theorem, which under the condition of stationary dynamics ensured that the infinite time average was well defined and equals the average over the probability distribution in state space. Unfortunately, the non-commutativity of matrix multiplication means that we can't use logarithms to turn the product into a sum in the equation above. Accordingly, Birkhoff's additive ergodic theorem cannot be applied, and we have no obvious way of knowing the properties of the infinite time average in the product above.

The situation concerning the product in equation above was resolved by Russian mathematician Valery Oseledets, who in mid 1960s was the first one to define the matrix $\mathbf{OSL}(\boldsymbol{x}_0, t)$, and studied its asymptotic properties (the matrix OSL stands for Oseledets matrix). He proved a seminal theorem now known as Oseledets multiplicative ergodic theorem which states that the limit

$$\lim_{t\to\infty} \mathbf{OSL}(\boldsymbol{x}_0, t) = \mathbf{P}$$

exists for a stationary process evolving on an invariant set with an invariant probability distribution, and is independent of the initial condition $\boldsymbol{x}_0$ for (almost) all $\boldsymbol{x}_0$ on the attractor. Consequently, the m global Lyapunov exponents λ_k obtained by the logarithms of the eigenvalues of $\mathbf{P}$ are well defined quantities for stationary dynamics on an invariant set. The theorem also proved that the ordered set of global Lyapunov exponents $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_m$, called the global Lyapunov spectrum is a topological invariant of the system on account of $\mathbf{P}$ being independent of $\boldsymbol{x}$ and t. Referring to our previous example, the multiplicative ergodic theorem is a generalization of the law of large numbers to a multiplicative average. Unlike the 1D case however, the multiplicative ergodic theorem doesn't allow us to represent the infinite time average in eqn 3.1.3 by an average over the state space probability distribution (in a straightforward way). A practical consequence of this is that most algorithms for estimating the Lyapunov spectrum have to deal with an infinitely long product leading to obvious numerical difficulties (more on this later). Irrespective of this fact however, the multiplicative ergodic theorem has led to a lot of progress in the ergodic theory of multiplicative processes. [50, 53].

3.2 Properties of Lyapunov Exponents

The Lyapunov exponents of a deterministic system allow quantifying not only its stability but also the predictability of the resulting dynamics, and other important properties such as the fractal dimension of the underlying attractor, its dynamical entropy, the rate at which volumes in state space dissipate, and symmetry properties associated with the evolution law F. The local Lyapunov exponents also carry substantial information about the time and state dependence of the dynamics.

3.2.1 Predictability and Entropy

One important notion should emerge from these examples: We are really concerned with changes in entropy, not in the entropy value itself. For example, if we choose a set of initial conditions in two cells and if the motion is regular, then the entropy value would not be 0, but it would remain constant as the system evolves. The change in entropy is characterized by the Kolmogorov-Sinai entropy rate (sometimes called the KS entropy), which describes the rate of change of entropy as the system evolves. First we shall give a rough definition of the K-S entropy (rate).

The key property that governs the rate at which chaotic systems become unpredictable is the Kolmogorov-Sinai entropy (KS entropy), also known as metric entropy [44, 54]. The KS entropy measures the rate of change of entropy as a dynamical system evolves. Thus, it's technically an entropy rate [55].

Like the Lyapunov exponents, the KS entropy is a topological invariant. In fact, KS entropy is intimately connected to Lyapunov exponents, as it can be shown that it measures the growth rate of the number of different trajectories that can be generated by a dynamical system [56, 57]. Although, the precise nature of the relationship between Lyapunov exponents and KS entropy isn't generally known. It is known that the KS entropy is bounded by the sum of positive exponents [57, 58], and in special cases

it is exactly equal to the sum of positive Lyapunov exponents, a fact known as Pesin's theorem [56]. For convenience it is common to assume Pesin theorem holds, so that the sum of positive exponents form a first estimate of the KS entropy $\hat{h}_\mu$. KS entropy determines the total predictability time, or the duration for which our predictions remain valid. Consequently, a positive KS entropy implies chaotic dynamics.

3.2.2 Volume Contraction Rate

The sum $S_m = \sum_{i=1}^{m} \lambda_i$ of all Lyapunov exponents of an m-dimensional system measures the contraction rate of volumes in the state space on the invariant set and gives a measure of the overall dissipation rate. For mechanical systems this quantity is closely connected to the rate of change of thermodynamic entropy. We will see in chapter 5, that under some conditions, the sum of local Lyapunov exponents can also be used to estimate the effective control signal to a system.

3.2.3 Symmetries and Zero Exponents

It shouldn't be surprising that if there are underlying symmetries and conservation laws in the evolution law, then they play an important role in determining the spectrum of Lyapunov exponents. For example, it can be shown that all continuous symmetries lead to zero exponents [59]. The most common of which is the time translation symmetry, which states that the evolution law is invariant under the substitution $t \to t + \tau$. This result provides us a handy way to verify time translation invariance in data. When the underlying dynamics are a result of a differential or partial differential equation then there must be at least one zero exponent [60]. There is a simple idea for why this must be true: The substitution $t \to t + \tau$ corresponds to a perturbation along the trajectory in state space where it can neither shrink or expand for a system with time translation invariance.

Hamiltonian systems are another source of a large number of zeros and symmetries. First, every integral of motion in a Hamiltonian system corresponds to a zero Lyapunov exponent. Another interesting property of a Hamiltonian system is that because of their time-reversal invariance (a discrete symmetry related to the change of sign, i.e. $t \to -t$) the Lyapunov exponents come in conjugate pairs that sum to zero. Even more interestingly, this symmetry is only slightly changed when the Hamiltonian structure is destroyed by the addition of viscous damping [61, 62]. Specifically, if a Hamiltonian system is subjected to a damping α, then the Lyapunov exponents still show conjugate pairing, but instead sum to $-\alpha$, i.e. $\lambda_i + \lambda_{m-i+1} = -\alpha$.

3.2.4 Local Exponents

Global Lyapunov exponents characterize the dynamics on the attractor as a whole. However, the fluctuations in the exponents over time and in the state space are important. They can allow us to measure state dependent predictability, i.e., how well can can predict t steps ahead of the current state. For example, the Lorenz system is difficult to predict near the transition region [63]. Clearly, characterizing this state and

time-dependent variation in predictability is important. It can reveal long time-scale changes over which the probability density on the attractor is likely not invariant. Related quantities such as local entropy rate, and local volume contraction rate also can prove useful in measuring local variations in the dynamics [64].

Finally, local Lyapunov exponents can be used to map out geometrical structures that shape the geometry of the state space, such as stable/unstable manifolds, and other general kinds of coherent structures [65]. These objects too have direct implications regarding control, e.g. many nonlinear (chaos) control algorithms operate by putting the system near the stable manifold of an unstable periodic orbit [66].

3.3 Numerical Estimation

The determination of the eigenvalues of the matrix $\mathbf{OSL}(\boldsymbol{x}_0, t)$ is not numerically trivial even in low dimensional dynamics when the equations of motion are known in advance. This is a key reason why estimation of the entire spectrum is typically avoided. The biggest difficulty is that the product that determines $\mathbf{OSL}(\boldsymbol{x}_0, t)$ gets ill conditioned rapidly as t increases [63]. The condition number, which is the ratio of the largest to the smallest singular values, and measures how ill conditioned a transformation is grows exponentially as to $e^{t(\lambda_1 - \lambda_m)}$ [63]. Geometrically speaking, the ellipsoids in Fig. 3.1 get thinner and thinner under the action of the local Jacobians until they collapse onto a line eventually, losing numerical resolution. To see this, consider two orthogonal vectors forming the basis of the tangent space at $\boldsymbol{x}_0$ shown in Fig. 3.4A. In general, the local Jacobian $\mathbf{J}(\boldsymbol{x}_0)$ will rotate and stretch the orthogonal basis vectors to a non-orthogonal set of tangent vectors at $\boldsymbol{x}_1$. In the next step $\mathbf{J}^2(\boldsymbol{x}_0)$ will shear the original basis vectors even more. As we continue this, and build up the product, the basis vectors will collapse onto the direction corresponding the maximal exponent leading to an ill conditioned behavior of the product $\mathbf{J}^t(\boldsymbol{x}_0)$.

To get around this issue, the method of recursive QR decomposition was proposed by Eckmann and Ruelle [46] based on previous work by Johnson et al who used it in a simplified proof of the multiplicative ergodic theorem [67]. The key idea behind recursive QR decomposition is to decompose each matrix $\mathbf{A}$ featuring in the product inside $\mathbf{OSL}(\boldsymbol{x}_0, t)$ into $\mathbf{A} = \mathbf{Q} \cdot \mathbf{R}$, where $\mathbf{Q}$ is an orthogonal $m \times m$ matrix, and $\mathbf{R}$ is an $m \times m$ upper triangular matrix with non-negative diagonal elements. The main advantage is that even when $\mathbf{A}$ is ill conditioned, the individual upper-triangular factors in recursive QR decomposition are well behaved and not badly conditioned. Geometrically, as shown in Fig. 3.4B, this corresponds to performing a Gram-Schmidt orthogonalization at each step and keeping track of the $\mathbf{R}$ matrices, the diagonals of which are non-negative and correspond to the lengths of the principle axes of the ellipsoid $\mathbf{J}^t(\boldsymbol{x}_0)$. The method of QR factorization is slightly different for global and local exponents as we detail below.

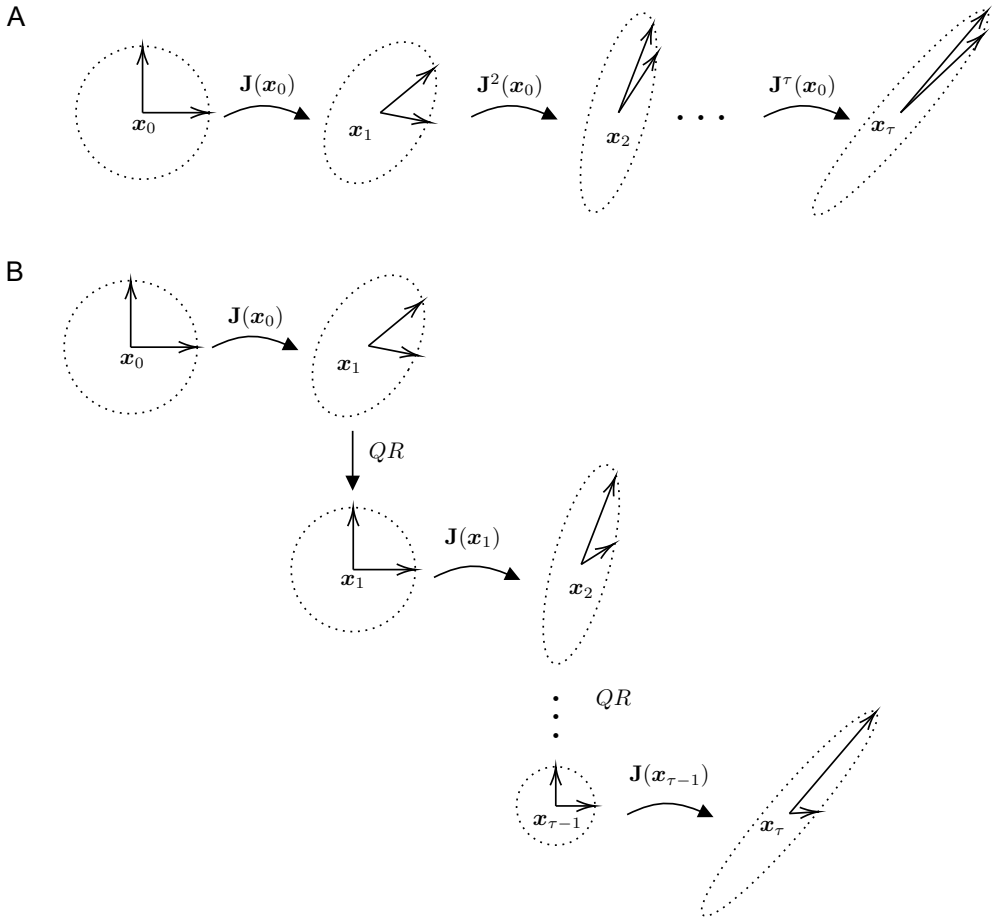


Figure 3.4: A. Two orthogonal vectors get closer and closer together as the number of steps increases until they get identified. This means that the long product of matrices is ill-conditioned **B.** The geometric idea behind recursive QR decomposition.

3.3.1 Recursive QR decomposition for global Lyapunov Exponents

As described above, the matrix $\mathbf{OSL}(\boldsymbol{x}_0, t)$ is a product of $2t$ matrices, to declutter the notation we denote them as following: $\mathbf{A}(2t).\mathbf{A}(2t-1)\cdots\mathbf{A}(1)$. The idea is to start with the $m \times m$ identity matrix $\mathbf{Q}(0) = \mathbf{I}_m$, and decompose each $\mathbf{A}(j) \cdot \mathbf{Q}(j-1)$ as

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

If we expand this out, we get

$$\begin{aligned} \mathbf{A}(1) &= \mathbf{Q}(1) \cdot \mathbf{R}(1) \\ \mathbf{A}(2) \cdot \mathbf{Q}(1) &= \mathbf{Q}(2) \cdot \mathbf{R}(2) \end{aligned}$$

the orthogonality means that $\mathbf{Q}^{-1}(1) = \mathbf{Q}^T(1)$ allows us to write

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

if we now consider the product we get

$$\begin{aligned} \mathbf{A}(2) \cdot \mathbf{A}(1) &= \mathbf{Q}(2) \cdot \mathbf{R}(2) \cdot \mathbf{Q}^T(1) \cdot \mathbf{Q}(1) \cdot \mathbf{R}(1) \\ &= \mathbf{Q}(2) \cdot \mathbf{R}(2) \cdot \mathbf{R}(1) \end{aligned}$$

where we have used the orthogonality of Q again in the last equation. Continuing like this the next step will give us

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

and the full product of $2t$ matrices will be

$$\mathbf{A}(2t) \cdot \mathbf{A}(2t-1) \cdots \mathbf{A}(1) = \mathbf{Q}(2t) \cdot \mathbf{R}(2t) \cdot \mathbf{R}(2t-1) \cdots \mathbf{R}(1)$$

This can be easily diagonalized, as the product of upper right triangular matrices as an upper right triangular matrix, and the eigenvalues of such a matrix are numbers along the diagonal. The Lyapunov exponents are then simply

$$\lambda_i = \lim_{t \to \infty} \frac{1}{2t} \sum_{\tau=1}^{2t} \log \left[R_{ii}(\tau) \right]$$

where $R_{ii}(\tau)$ is the i^{th} diagonal entry of $\mathbf{R}(\tau)$. Intuitively, the repeated use of orthogonalization procedure has resulted in a sequence of bases for the tangent spaces such that the action of DF on a general delta can be represented as an action on a set of independent scalar quantities - the projections of delta onto the basis vectors. As a consequence the limit above has the form of an additive average and thus Birkhoff's theorem guarantees the above limit when the dynamics are stationary.

3.3.2 Iterated Recursive QR decomposition for Local Lyapunov exponents

In the case treated above, the limit $t \to \infty$ makes the average converge, importantly the $\mathbf{Q}$ matrices which describe the orientation of the axes after each iteration settle down to the Lyapunov vectors and thus $\mathbf{Q}(t)$ tends to the identity matrix $\mathbf{I}_m$ for long t. This means that the eigenvalues all lie in the product of the upper triangular $\mathbf{R}$ matrices. However, same cannot be said about the finite time product, specifically, the eigenvalues of the matrix $\mathbf{Q}(2t)$ accounts for some contribution, if $\mathbf{Q}(2t)$ was identity, then we would be okay. Abarbanel et al. used a result from numerical linear algebra to provide an iterative scheme to address this issue and estimate the eigenvalues [49]. The idea is to start with the matrix

$$\mathbf{M}_1 = \mathbf{Q}_1(2t) \cdot \mathbf{R}_1(2t) \cdot \mathbf{R}_1(2t-1) \cdots \mathbf{R}_1(1)$$

and move $\mathbf{Q}_1(2t)$ to the right giving us the matrix (the reason to subscripts will become clear shortly)

$$\mathbf{M}_2 = \mathbf{R}_1(2t) \cdot \mathbf{R}_1(2t-1) \cdots \mathbf{R}_1(1) \cdot \mathbf{Q}_1(2t)$$

It is easy to see from above that

$$\mathbf{M}_2 = \mathbf{Q}_1^T(2t) \cdot \mathbf{M}_1 \cdot \mathbf{Q}(2t)$$

and because of the orthogonality, $\mathbf{M}_1$ and $\mathbf{M}_2$ both have the same eigenvalues. Next, we perform QR decomposition of $\mathbf{M}_2$

$$\mathbf{M}_2 = \mathbf{Q}_2(2t) \cdot \mathbf{R}_2(2t) \cdot \mathbf{R}_2(2t-1) \cdots \mathbf{R}_2(1)$$

and shift $\mathbf{Q}_2(2t)$ to the right again to define $\mathbf{M}_3$ as

$$\mathbf{M}_3 = \mathbf{R}_2(2t) \cdot \mathbf{R}_2(2t-1) \cdots \mathbf{R}_2(1) \cdot \mathbf{Q}_2(2t)$$

performing QR decomposition of M_3 and continuing the sequence for K steps we have

$$\mathbf{M}_K = \mathbf{Q}_K(2t) \cdot \mathbf{R}_K(2t) \cdot \mathbf{R}_K(2t-1) \cdots \mathbf{R}_K(1)$$

A theorem from numerical linear algebra guarantees that $\mathbf{Q}_K(2t)$ tends to the identity matrix as K increases. In practice, $\mathbf{Q}_K(2t)$ coverages rapidly and it is not required to take K past double digits. Once $\mathbf{Q}_K(2t)$ is sufficiently close to identity, then the local Lyapunov exponents are determined by eigenvalues of $\mathbf{M}_K$ (which are the same as $\mathbf{M}_1$) and can be estimated from the diagonal elements of $\mathbf{R}_K$ as follows

$$\Lambda_i(\boldsymbol{x}_0, \tau) = \frac{1}{2t} \sum_{\tau=1}^{2t} \log\left[R_K(\tau)_{ii}\right]$$

3.3.3 Lyapunov Spectrum Estimation of Model Systems Using Analytical Jacobians

When we have access to the underlying equations of motion, we can use them to estimate the Jacobian matrices $\mathbf{J}(\boldsymbol{x}_t)$ at each point $\boldsymbol{x}_t$ along the trajectory, and perform

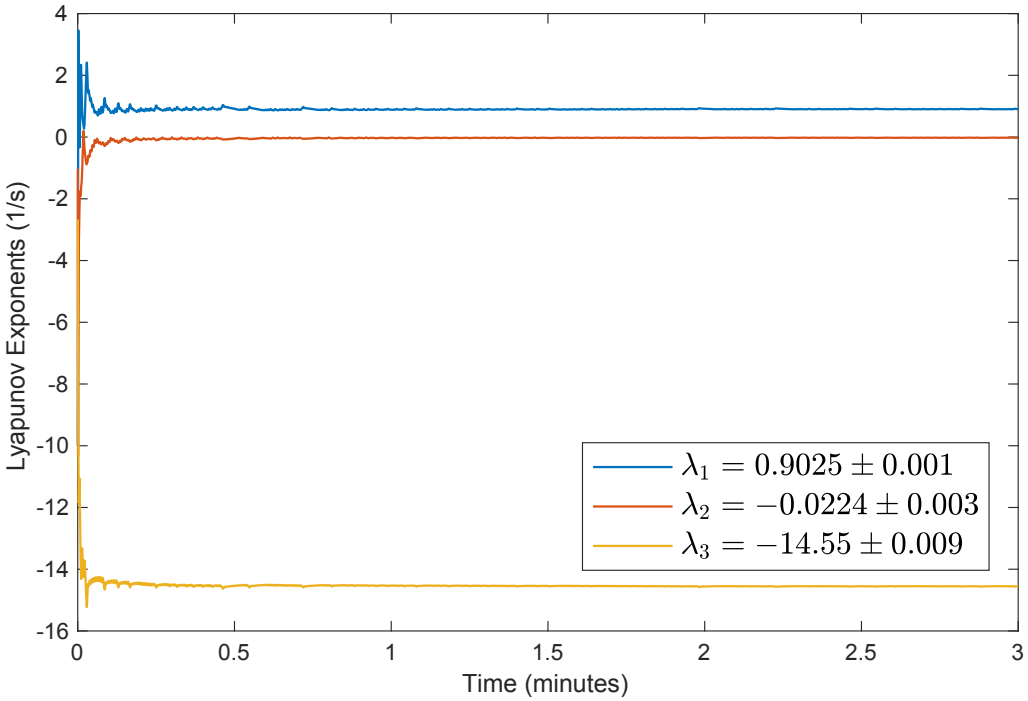


Figure 3.5: Running average of the three Lyapunov exponents for the Lorenz system estimated using analytical Jacobians.

recursive QR decomposition as described above. In this section I show the estimation of Lyapunov spectrum for the Lorenz and Duffing oscillator systems using their analytical Jacobians. The running average of the Lyapunov spectrum for the Lorenz system is shown in Fig. 3.5. The estimated values are close to the reported values of $(0.906, 0, -14.572)$ in literature [68]. However, there are still errors in estimation even though I am using the analytical Jacobians for the system. Notably, λ_2 should be identically equal to zero but my estimates show $\lambda_2 = -0.02 \pm 0.003$. These estimates do get better with longer runs, but converge very slowly.

Next, in Fig. 3.6 I show estimates of the Lyapunov spectrum for the coupled Duffing oscillators for different values of the drive coupling f. The difficulty of estimation here is even more pronounced as can be seen by the noisy structure of the estimates. In addition to previously mentioned difficulties in estimation, here we also have issues arising due to additional symmetries in the equations of motion. Specifically, there are up to three zero exponents for this system. Several different zero exponents, and degeneracies in general lead to difficulty in estimation. Regardless, we can see in Fig. 3.6, that the Lyapunov spectrum is symmetric about -0.1. This symmetry results from the damped driven Hamiltonian structure of this system.

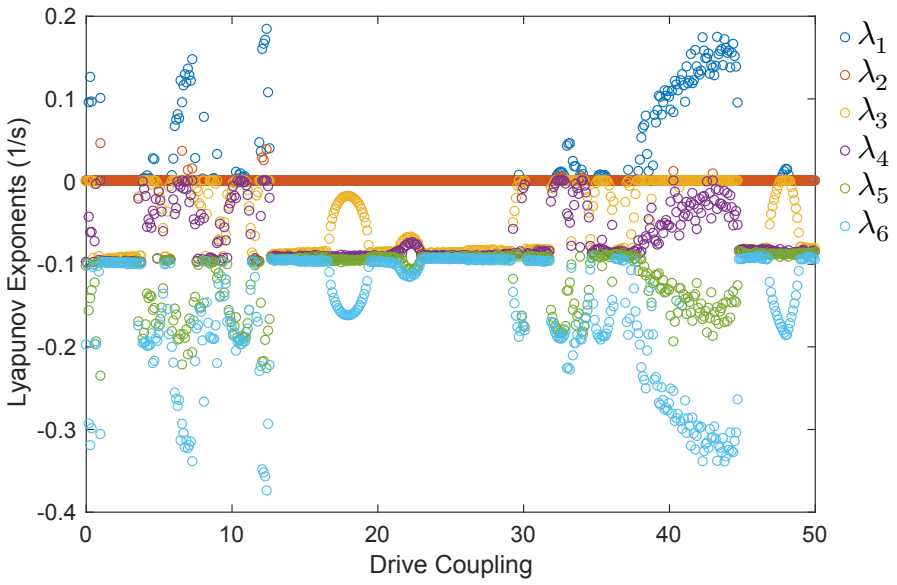


Figure 3.6: Lyapunov exponents estimated from analytical Jacobian of damped driven coupled Duffing oscillators. The spectrum shows a symmetry about a point related to the overall damping.

3.3.4 Background to Estimation of Jacobians from experimental data

The task of Lyapunov spectrum estimation becomes significantly more challenging in a reconstructed state space when we do not have access to the equations, or perhaps are not even sure whether the dynamics are described by a differentiable function! In such a case we have to estimate the Jacobian matrices by modeling the dynamics in the embedding coordinates. If we're successful then it is safe to assume that the dynamics are differentiable and we can use the estimated matrices to estimate the local and global Lyapunov exponents.

We can either use a global model where we fit a single nonlinear predictive model for all the data and then differentiate it to estimate $\mathbf{J}(\boldsymbol{x}_t)$, or we can use local models where we estimate different models for each neighborhood of the state space. We already saw an example of local modeling when we used the nearest neighbor predictor to optimize the embedding parameters. In this case however, we are interested in estimating local derivatives, accordingly, it is appropriate to consider local linear models. Specifically, we begin with an approximation to the tangent space dynamics in eqn 3.5

$$\boldsymbol{\delta x}_{t+1} \approx \mathbf{J}(\boldsymbol{x}_t)\ \boldsymbol{\delta x}_t$$

where $\boldsymbol{\delta x}_t$ now is an estimate of the tangent vector at $\boldsymbol{x}_t$. It is obtained by finding another state space point $\boldsymbol{x}_{t'}$ in the neighborhood of $\boldsymbol{x}_t$ and forming the difference vector $\boldsymbol{\delta x}_t = \boldsymbol{x}_{t'} - \boldsymbol{x}_t$. The difference vector τ steps ahead $\boldsymbol{\delta x}_{t+1} = \boldsymbol{x}_{t'+1} - \boldsymbol{x}_{t+1}$ is a tangent vector at $\boldsymbol{x}_{t+1}$ and approximates the image of $\boldsymbol{\delta x}_t$ under the action of $\mathbf{J}(\boldsymbol{x}_t)$. Thus, our goal is to invert equation above to estimate the Jacobian matrix at $\boldsymbol{x}_t$, $\mathbf{J}(\boldsymbol{x}_t)$.

Clearly the equation above is under-determined, we therefore consider a collection of nearest neighbors of $\boldsymbol{x}_t$. This can be done by either picking a fixed number of N_b neighbors $\boldsymbol{x}_{t'}$, for each $\boldsymbol{x}_t$. Or by choosing all points within a fixed distance ϵ of $\boldsymbol{x}_t$, i.e. by picking all $\boldsymbol{x}_{t'}$, such that$\| \boldsymbol{x}_{t'} - \boldsymbol{x}_t \| \leqslant \epsilon$ for some fixed ϵ. In the former case N_b is a parameter and the size of each neighborhood changes ϵ for each state space point considered. While in the latter case, the size of the neighborhood ϵ is a parameter and number of neighbors N_b changes for each state space point. In all the analysis in this thesis, we choose the first approach and fix the number of neighbors. After choosing the neighbors, we form all the N_b displacement vectors $\boldsymbol{\delta x}_t$ and stack them into a $N_b \times m$ neighborhood matrix $\mathbf{B}_{\boldsymbol{x}_t}$. Next, we calculate the displacement vectors $\boldsymbol{\delta x}_{t+1}$ at the next time step and stack them into the evolved neighborhood matrix $\mathbf{B}^1_{\boldsymbol{x}_t}$, also of size $N_b \times m$. For a small neighborhood, if the underlying dynamics are differentiable, then the two neighborhood matrices are related by the following set of linear equations [48]

$$\mathbf{B}^1_{\boldsymbol{x}_t} = \mathbf{B}_{\boldsymbol{x}_t} \cdot \mathbf{J}(\boldsymbol{x}_t) \tag{3.6}$$

the usual way of solving the inverse problem in eqn 3.6 is to use the Moore-Penrose pseudo-inverse of $\mathbf{B}_{\boldsymbol{x}_t}$, which is denoted by $\mathbf{B}^\dagger_{\boldsymbol{x}_t}$, and is a $d \times N_b$ matrix calculated from the SVD of $\mathbf{B}_{\boldsymbol{x}_t}$. The estimated 1-step Jacobian matrix is then given by

$$\mathbf{J}(\boldsymbol{x}_t) = \mathbf{B}^\dagger_{\boldsymbol{x}_t} \cdot \mathbf{B}^1_{\boldsymbol{x}_t} \tag{3.7}$$

Repeating the fitting process for all $\boldsymbol{x}_t$ gives us a collection of estimated Jacobian matrices at each point in state space. The recursive QR procedure described above can now be used to estimate the local or global Lyapunov exponents for the dynamics.

3.3.5 A New Algorithm for Estimation of Jacobians from experimental data

Unsurprisingly, the procedure described above doesn't always work. In particular, the algorithm is extremely sensitive to the geometry of the reconstructed attractor. When there are sharp twists and high-curvature regions in the original attractor, the density of local neighborhood keeps changing meaning that the distance threshold ϵ is not a constant but can vary significantly over the attractor. I propose a method inspired by a somewhat old prediction schemed proposed in [69] to account for this variation of local attractor geometry. The basic idea is to use weighted linear regression by assigning greater weight to points on the attractor that are near the current state $\boldsymbol{x}_t$. The exact weight given to an observation is given by

$$w_{t'} = \exp \frac{-\theta \|\boldsymbol{x}_{t'} - \boldsymbol{x}_t\|}{r}$$

where r is the average distance between all neighbors,i.e.

$$r = \frac{1}{n} \sum_{t'}^{N_b} \|\boldsymbol{x}_{t'} - \boldsymbol{x}_t\|$$

.

Thus, we still solve the least square problem in eqn 3.6. The only difference is that each entry in the matrices $\mathbf{B}_{\boldsymbol{x}_t}$ and $\mathbf{B}^1_{\boldsymbol{x}_t}$ is multiplied by the corresponding weight $w_{t'}$.

The scale parameter $\theta \geq 0$ determines how strongly the regression is localized in the neighborhood of each state point $\boldsymbol{x}_t$. For $\theta = 0$, we're essentially fitting a global linear model to the dynamics. As θ increases the regression gets increasingly local, until in the extreme case most of the weight is given to only the current state $\boldsymbol{x}_t$ and is therefore more sensitive to noise. Optimal values of θ lead to stable estimates of the Jacobian $\mathbf{J}(\boldsymbol{x}_t)$ and correspond to small prediction errors. Next, I will use this algorithm to estimate the Lyapunov exponents of Lorenz and Duffing systems.

Estimating Lyapunov Spectrum for Model Systems Using Estimated Jacobians

First I attempt to estimate the Lyapunov spectrum of the Lorenz system, which is shown in Fig. 3.7. In Fig. 3.7A (left), we see the dependence of the exponents on the scale parameter θ. As can be seen that both λ_1 and λ_2 are relatively stable, but the negative exponent λ_3 is only available accurately for a narrow range of θ. The range of θ for which accurate estimates are available correspond to the minimum of the mean squared error in estimating the system in eqn 3.6, as is shown in Fig. 3.7A (right). The running average of the Lyapunov spectrum estimates for $\theta = 9$ is plotted in Fig. 3.7B and shows that the exponents converge to stable values over time. Finally, in Fig. 3.7C, I test the robustness of the algorithm developed here as a function of noise amplitude. The algorithm is remarkably robust against observation noise. The only problem occurs at extremely low values of noise for which the negative exponent isn't accurate. This presumably happens of collinearity in least square problem and can be easily removed by using a regularized least square. Adding noise essentially serves as a good regularizer and that is why the estimates actually get better as I increase the noise.

Next, I estimate the spectrum of the more challenging case of coupled Duffing oscillators in a 5D embedding built in the previous chapter. The results are shown in Fig. 3.8. We can see that despite the challenging nature of this problem, we recover all the exponents approximately correctly Fig. 3.8C. The variation due to the scale parameter is strong for this case, as seen in Fig. 3.8A, but the mean squared error plot can be used to pick the scale parameter θ, Fig. 3.8B. Finally, in Fig. 3.8D, we are also able to recover the fact that the symmetry of the Lyapunov spectrum confirming that the coupled oscillator system is a damped driven Hamiltonian system.

3.4 Conclusion

In this chapter I reviewed the mathematical ideas behind Lyapunov exponents and their estimation. I discussed the several useful insights we can derive from a calculation of all the exponents. Next, I elaborated on the numerous difficulties with the estimation of Lyapunov exponents from observations and proposed a new algorithm which is more robust to noise and effects of attractor geometry. Finally, I applied the proposed algorithm to model systems and verified its applicability. In chapter 5 I will apply

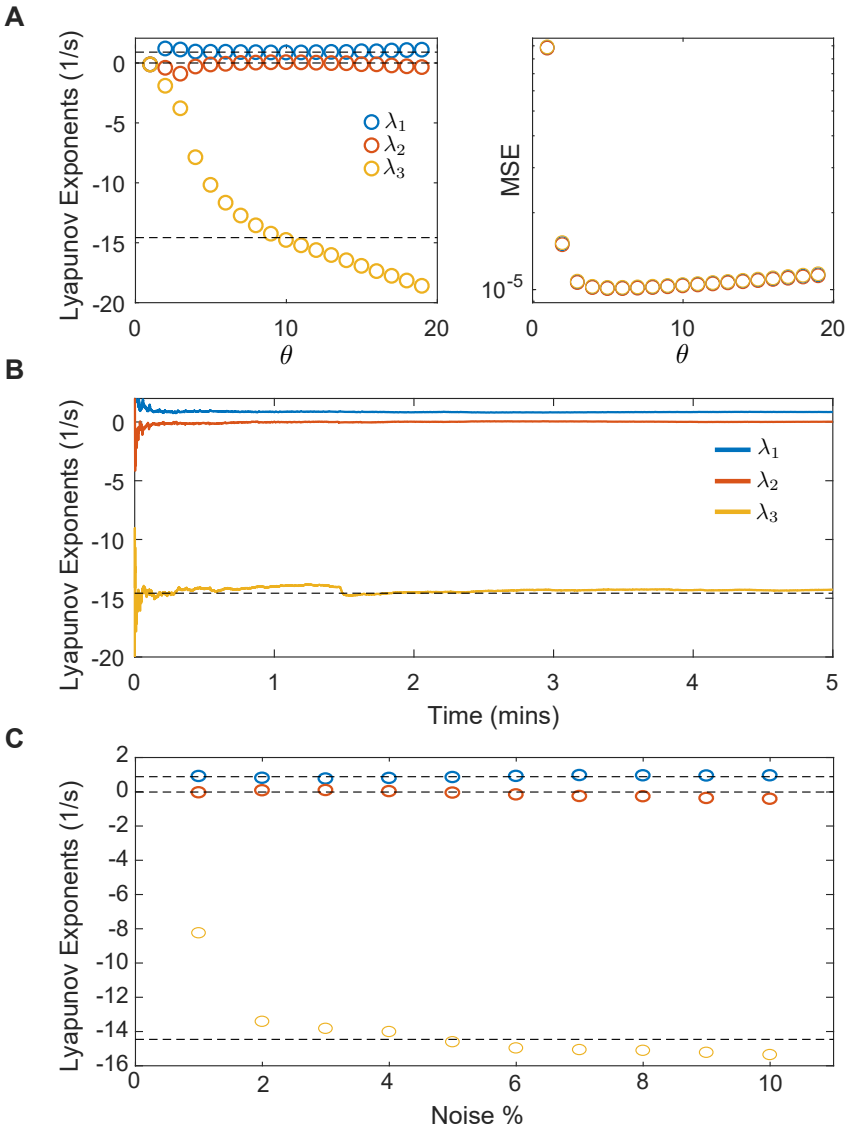


Figure 3.7: Lyapunov exponents estimation using learned Jacobians for the Lorenz system. **A.** Exponents plotted as a function of the scale parameter θ (left), and MSE also plotted as a function of θ (right). **B.** Running average of the estimated exponents shows that they converge to stable values. **C.** Estimated Lyapunov exponents plotted as a function of observation noise shows that the algorithm is robust to a wide range of noise amplitude.

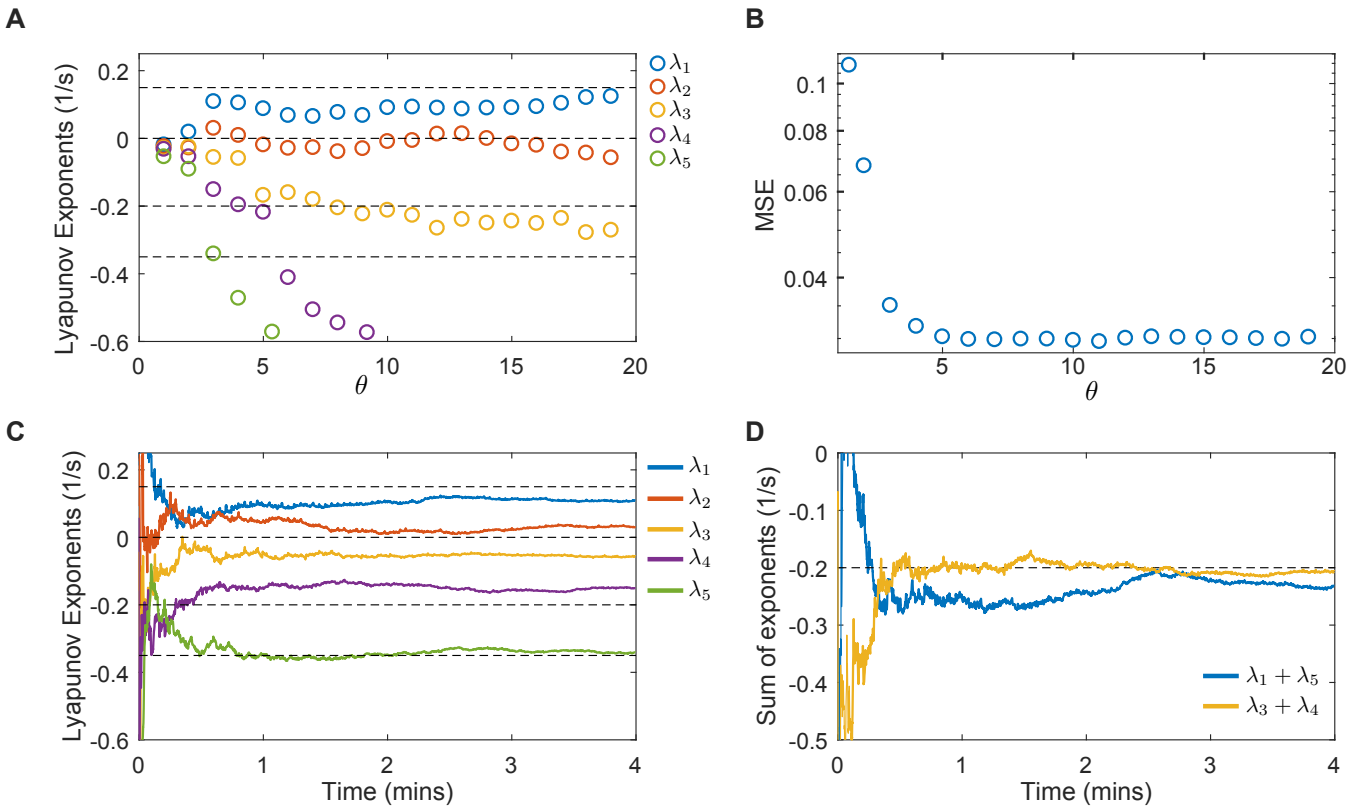


Figure 3.8: Lyapunov exponents of the coupled duffing oscillators from Jacobians estimated by weighted least squares regression in a 5D embedding. **A.** Lyapunov exponents plotted as a function of the scale parameter θ. **B.** The mean squared error in prediction plotted as a function of θ. **C.** Running average of the 5 exponents plotted for $\theta = 3$. **C.** Sum of the conjugate exponents shows that we can also recover the damping and the symmetry corresponding to damped driven Hamiltonian systems.

these ideas to estimate the entire spectrum of Lyapunov exponents for the movement of *C. elegans*.

Chapter 4

Phase Space Reconstruction for *C. elegans* Behavioral Dynamics

4.1 Introduction

In this chapter I will list the results of applying the reconstruction pipeline developed in chapter 2 to the study of worm's locomotory dynamics. As we will see, mapping the complexity of *C. elegans* locomotion onto a geometric structure, even a high dimensional one, has the effect of making it more approachable and amenable to different kinds of analysis.

4.2 Experimental Details and Preprocessing

I analyze three datasets of worm locomotion. The first one is composed of $N = 12$ L4 stage N2 worms foraging freely on a 2D agar plate. It was first published in [70] and has been since made publicly available by [71]. Briefly, the foraging dataset is recorded at 32 Hz with high resolution tracking microscopy. For the analysis it was downsampled to a frame rate of 16Hz. Worms were cultivated under standard conditions at 20°C [72]. Before beginning the assay, worms were rinsed off *E.coli* bacteria by making them swim in NGM buffer for 1 minute, they were then placed on a 9.1cm assay plate (Petri-Dish) with a 5cm copper ring pressed into the agar surface to prevent them from climbing the edge. The assay started 5 minutes after the transfer and lasted 35 minutes.

In the second dataset, I analyze the *C.elegans* escape response described first in [73], and also made publicly available by [71]. In these experiments, $N = 92$ worms were targeted on the head with a 100ms, 75mA IR laser pulse from a diode laser ($\lambda = 1440$nm). Images were recorded at 20Hz for 30s (10s before stimulation and 20s after stimulation). To prevent adaptation each worm was only assayed once. I downsampled the data to $16Hz$ (after interpolation), so that it can be easily compared with the foraging dataset.

Finally, the third dataset is composed of $N = 30$ adult N2 worms swimming in M9 buffer solution. This dataset was requested from David A. Gagnon and Paulo E. Arratia of University of Pennsylvania, and was first published in [74]. Each recording here consists of 6-12 cycles of forward movement.

4.2.1 Image Analysis and Posture Space Estimation

The tracking and posture space estimation pipeline is identical to the one in [71]. Briefly, we parameterize the shape of a worm by tangent angles calculated at 100 points along its backbone. This gives us a matrix $\theta \in \mathbb{R}^{T\times 100}$, containing the shape information for each uncrossed frame where the worm's body doesn't intersect itself. Next, a 5 dimensional approximation of the 100 dimensional time series is calculated by projecting the elements of θ onto the basis given by the first 5 singular vectors (or eigenworms) of θ. To process frames where a worm's body crosses itself, we use the inverse tracking algorithm [71] which searches the 5 dimensional shape space for a point p which generates a frame image that best matches the actual worm image in a given crossed frame. The search is done by optimizing an error function f_{err} which quantifies the similarity between a reconstructed worm image $\tilde{W}(p)$ and the movie frame image W. At the end of the optimization we get a $T \times 5$ vector time series $\mathbf{A} = [a_{1:5}^{1:T}]$, where T is the total number of frames.

4.3 Results

A 6D Embedding Captures Worm Locomotory Dynamics

In the lab, worms move in a 2D environment by propagating sinusoidal body bends along their body. Their behavior can be broadly categorized into three categories, forward movement, backward movement, and turns which come in two types, gradual shallow turns, or sharp turns resulting from deep body bends [75, 76]. It was previously shown that almost all the postures that worms make on a 2D plate can be obtained by combining four basic shapes called eigenworms in different amounts [77] Fig. 4.1A. The coefficients of the eigenworms describe only the shape of the worm at an instant, not its dynamics. However the authors in [77] used the fact that the coefficients varied continuously in time to estimate the derivatives of the eigenworm coefficients and approximate the postural dynamics of forward and backward crawling.

To reconstruct the dynamics I use the concepts and pipeline developed in chapter 1. First, I construct the delay matrix $\overline{Y}$ by stacking delayed copies of the eigenmode projections upto some time K, which is shown schematically in Fig. 4.1B. Next, For a range of values of K, I estimate the predictability time T_{pred} and pick K such that predictability time is maximized. As a reminder, for complex dynamics the prediction error grows with time until it saturates when the errors reach the size of the attractor. T_{pred} is the average time required for the error to grow to the size of the attractor.

There is a large degree of variability in T_{pred} across different worms, so I center them by removing the mean $\langle T_{pred}\rangle$ for each worm. These $T_{pred} - \langle T_{pred}\rangle$ plots averaged over all $N = 12$ different worms are shown in Fig. 4.1C. The plot shows that predictability, time averaged over all worms in the dataset is maximized between $K = 5$ and $K = 11$. In general, larger values of K have an effect of smoothing the data, while smaller values of K lead to more noisy looking reconstructions. Thus, although in principle, any value of K between $K = 5$ and $K = 11$ is valid in terms of predictability, I choose the upper limit of the range, i.e. $K = 11$, because it results in smoother trajectories without affecting the predictability. This choice of K corresponds to $0.7s$ or about a quarter of

the body wave.

Once the choice of K is fixed, the next step is to perform a singular value decomposition on the resulting $T \times KD$ matrix and find the smallest number of basis vectors m that result in a good embedding. This is done by evaluating T_{pred} as a function of m. In Fig. 4.1D we can see the $T_{pred} - \langle T_{pred} \rangle$ values plotted for different embedding dimensions. The predictability rises gradually until 6 dimensions after which it roughly stays constant. We can conclude from this, that the 6 dimensional phase space, almost completely captures the dynamics of *C. elegans* movement. An example of the error curves used to estimate K and m is shown in supplementary figure 4.11.

The Reconstructed Phase Space Decomposes Worm Locomotion into Three Behavioral Modes

Next, I perform a singular value decomposition of $\overline{Y}$ for $K = 11$, followed by independent component analysis of the first $m = 6$ right singular vectors. This results in a 6-dimensional state space which decomposes *C. elegans* behavioral dynamics into linear combinations of 6 independent posture sequences, each of length $K = 11$ frames.

The resulting six components are surprisingly interpretable and come in three pairs, each pair corresponding to forward, backward or turning locomotion Fig. 4.2. It should be emphasized that this happens in an entirely unsupervised manner, as the input to the pipeline doesn't contain any information about the global organization of worm behavior into discrete categories. In analogy to normal modes for linear dynamical systems I call these three pairs behavioral modes.

Left column of Fig. 4.2A shows the curvature kymograph of the forward mode components Γ_{f1} and Γ_{f2}, which are composed of two waves in quadrature going from the head to tail. They represent the ventrally and dorsally initiated anterior-posterior body waves that worms make during forward locomotion. Middle column shows the curvature kymographs for the reversal mode components. Γ_{r1} corresponds to the initial part of the backward movement where worms swing their head backwards, as suggested in [78]. while Γ_{r2} captures the subsequent posterior-anterior body waves as the animal moves backwards. Finally, the turning mode components Γ_{t1} and Γ_{t2} correspond to the beginning and end of the deep body bends such as omega and delta turns worms use to rapidly change direction during navigation.

The geometry and topology of trajectory in behavioral phase space contains important qualitative and quantitative information about worm's behavior. An example trajectory corresponding to 15 minutes, or $N = 14400$ samples is visualized in Fig. 4.2B as planar projections onto three mode combinations. For the forward and backward modes, the projections are color-coded by the tangential velocity of the centroid, which is negative when the worm is moving backward and positive for forward locomotion. While, for the turning mode, the trajectory is color coded by the turning rate. Globally, the state space is composed of three qualitatively distinct sets of cyclic trajectories corresponding to the different body waves generated during forward, backward and turning locomotion. Locally, however, the trajectories are bundled together according to the continuous behavioral measures of worm's speed and turning rate. Thus, the reconstructed phase space naturally captures both discrete and continuous *C. elegans*

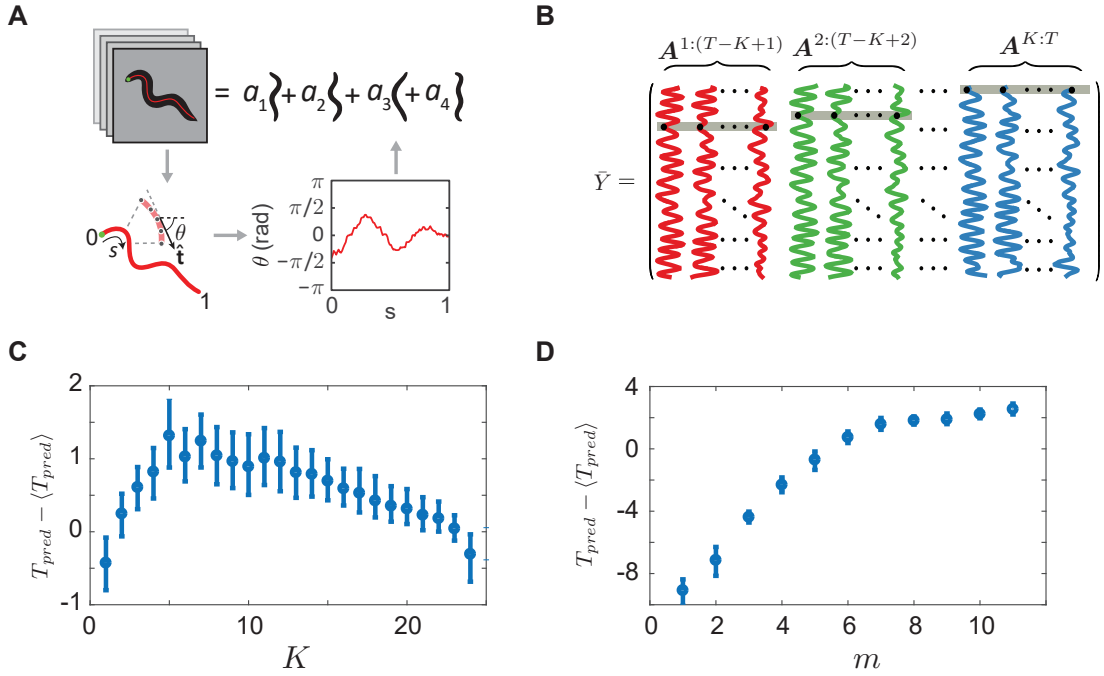


Figure 4.1: **A.** The instantaneous posture of the worm can be decomposed into a basis of representative shapes called eigenworms. First 4 eigenworms are shown here along with a schematic of the image processing pipeline used to obtain the eigenworms projections. The pipeline is identical to the one used in [71]. I used first 5 eigenworms projections for this analysis, consequently each frame in a video recording is represented by a $T \times 5$ matrix. **B.** Delay embedding matrix $\overline{Y}$ is made by stacking delayed copies of the eigenworms projections in **A** upto a delay time K, hence it has a dimensionality of $(T - K + 1 \times 5K)$. **C.** Predictability as a function of K. Error bars show bootstrapped 95% CIs across different worms. The predictability increases gradually, reaching a plateu between $K = 5$ and $K = 11$. I choose the upper limit $K = 11 (\approx 0.7s)$ as it results in smoother trajectories compared to $K = 5$. **D.** Predictability as a function of embedding dimension m. The centered predictability times $T_{pred} - \langle T_{pred} \rangle$ are plotted as a function of m. Our ability to predict the future locomotory behavior increases gradually until $m = 6$ dimensions after which it roughly saturates. I choose $m = 6$. Notice that the variability in the embedding dimension as across different worms is much smaller compared to that in estimates of K.

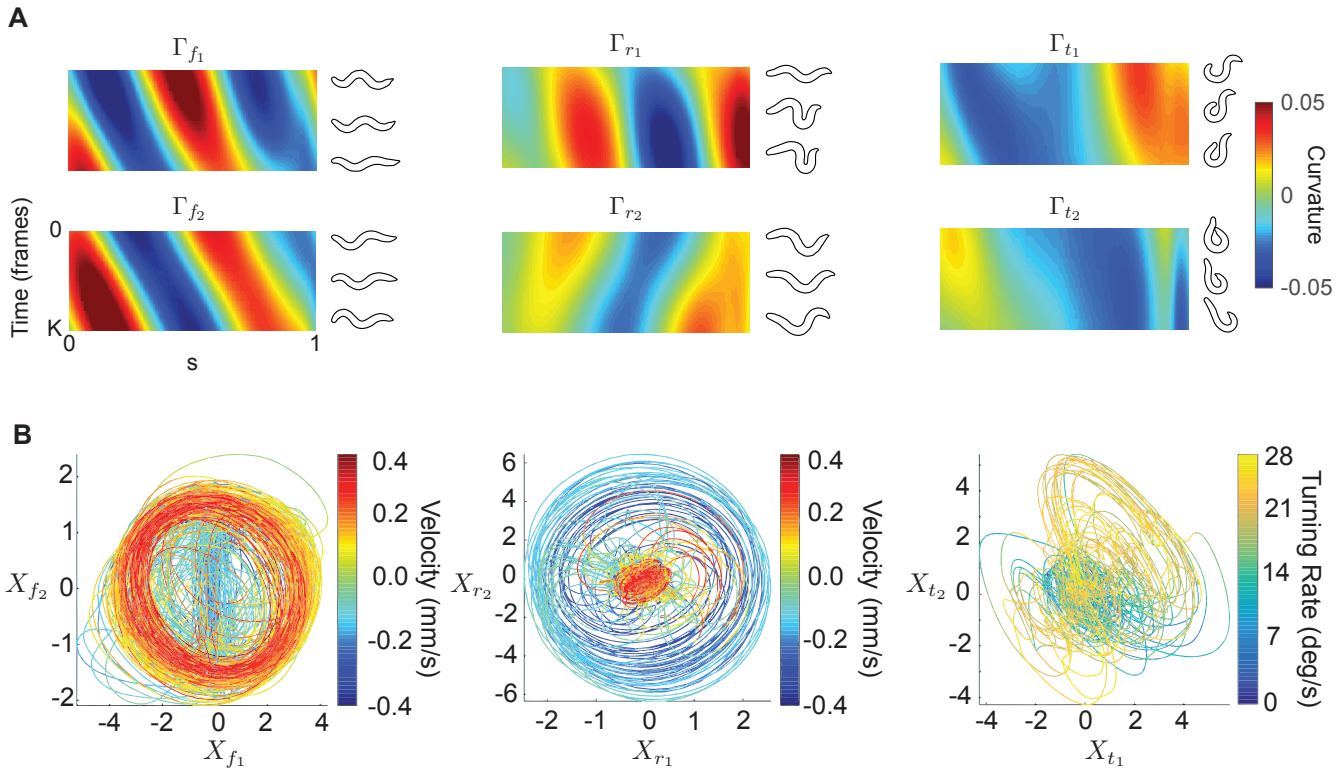


Figure 4.2: Modes and Projections of the Phase Space of Foraging Worms. **A.** Left column shows the forward mode components Γ_{f1} and Γ_{f2}. Middle column shows the reversal mode components, Γ_{r1} and Γ_{r2}. Finally, the turning mode components Γ_{t1} and Γ_{t2} are shown in the right column. **B** An example state space trajectory visualized here as planar projections. Projections in the forward and backward modes are color-coded by the tangential velocity of the centroid, while in turning mode the color code is according to the centroid turning rate.

behaviors.

Normalized Mode Energies Allow a Simple Way to Categorize Worm Behavior

In principle, the state space trajectories represent the instantaneous state of worm locomotion. However, the trajectories in Fig. 4.2B also capture longer time-scale behavior such as discrete behavioral states. This suggests that we can use the state space to categorize worm behavior into distinct behavioral categories.

I introduce following quantities, which are normalized mode energies as

$$
\begin{aligned}
E_f &= \frac{X_{f1}^2 + X_{f2}^2}{\sum_i X_i^2} \\
E_r &= \frac{X_{r1}^2 + X_{r2}^2}{\sum_i X_i^2} \\
E_t &= \frac{X_{t1}^2 + X_{t2}^2}{\sum_i X_i^2}
\end{aligned}
\tag{4.1}
$$

The mode energies range from 0 to 1 and measure the activity balance of different modes. For example when E_r is close to 1 for a duration of time, it means that the reversal modes are active, while the others suppressed. In general a stereotyped behavior leads to most of the energy being concentrated in that mode. While behaviors that cannot be neatly categorized into forward, reversal or turn are ones where energy is shared roughly equally amongst the three modes. These behaviors are labeled as transient behaviors.

Fig. 4.3A shows the normalized mode energies for a duration of 30 seconds. We can see for the first few seconds E_r is close to 1, meaning that most of the energy is concentrated in the reversal modes. The energy then transfers to the turning modes, followed by a brief transient period where it is shared roughly equally between the three modes. Another transient period happens at around 13 seconds.

The normalized mode energies provide a natural way to categorize worm's behavior into discrete categories. If for a duration, a single mode accounts for more than 55% of the total energy, then I label that behavioral sequence by the behavior corresponding to that mode. In addition, the sequences where the energy is shared roughly equally between all modes are labeled as transient behaviors as they cannot be neatly characterized into forward, backward or turn. An example ethogram corresponding to the behavioral sequence in Fig. 4.3A is shown in Fig. 4.3B.

It is also possible to order the state space trajectories into distinct categories. Fig. 4.3C shows 3D projections of the two forward and one of the reversal coordinates color coded according to the inferred behavioral states. We can see that the energy based discretization neatly orders trajectories into coherent bundles. Additionally, it shows that trajectories corresponding to forward and reversal are orthogonal cycles implying that forward and reversals can be treated as independent behaviors. Next, Fig. 4.3D shows color coded 3D projections of the reversal and turning planes. We can see that that many reversal trajectories are connected to deep body bends seen as a large excitation in the X_{t2} axis. This is consistent with the well observed phenomenon that reversals and deep body bends such as omega turns are sequentially linked [79, 80]. Interestingly, most transient states occur during transition behaviors, especially the transition between reversal and turning.

Based on the discretization proposed above we can estimate various statistics of the behavioral states. First I find that foraging worms spend about half the time in the forward state, while the remaining is split equally between reversal, turn and transient states Fig. 4.4A. I can also estimate the average duration of each behavioral bout, the distributions of which are shown in Fig. 4.4B. I estimate the average forward duration to be $1.84(1.79, 1.90)s$. The average reversal duration to be $1.35(1.30, 1.39)s$, average turn duration as $1.22(1.18, 1.26)s$ and the average duration of transient behaviors as

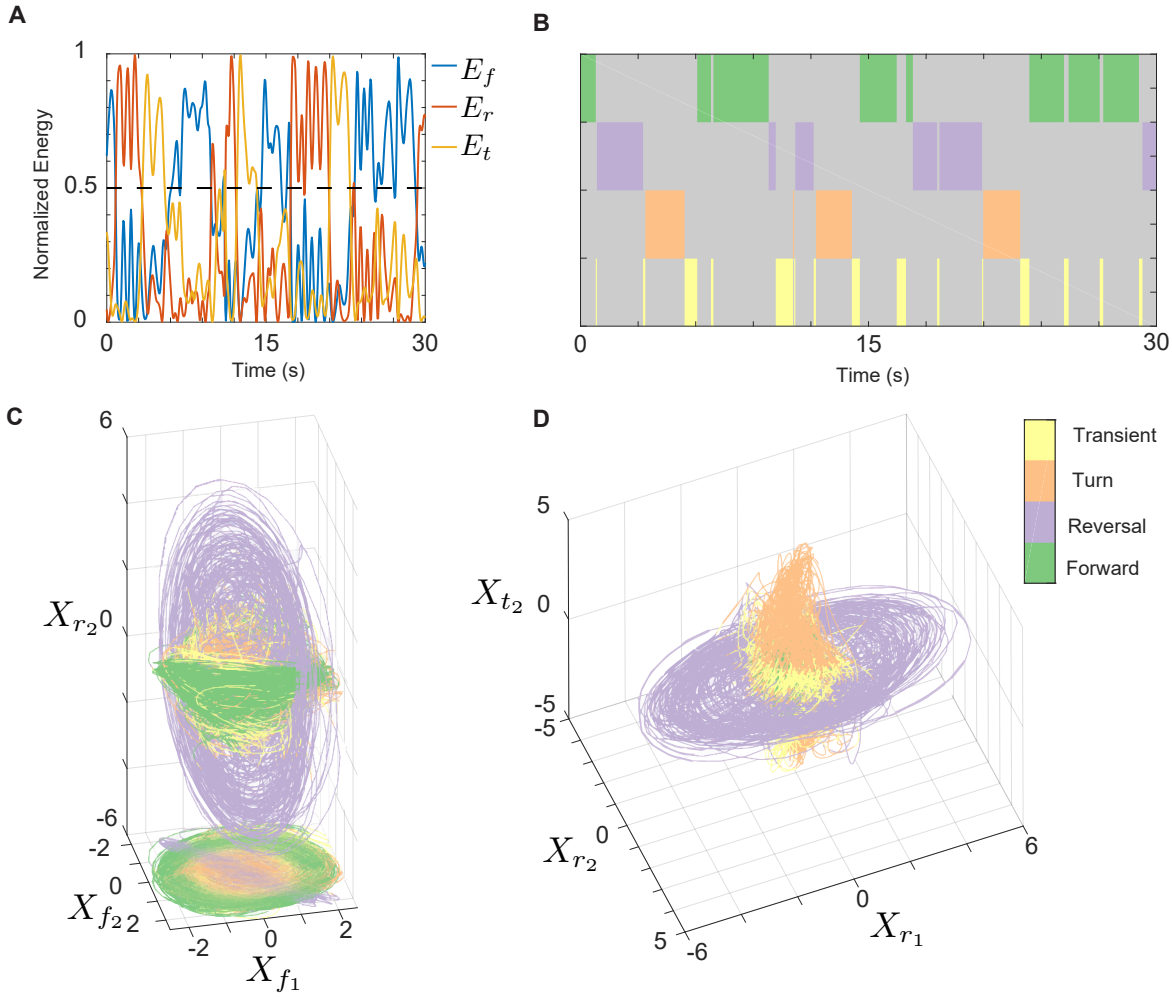


Figure 4.3: Discretizing Behavior Using Mode Energies. **A.** A sample trace of the normalized mode energies used to discretize worm behavior. **B.** Example ethogram corresponding to the behavioral sequence in A. **C.** 3D projections of the two forward and one reversal coordinates of the state space color coded according to the inferred behavioral states. **D.** Color coded 3D projections of the reversal and turning planes.

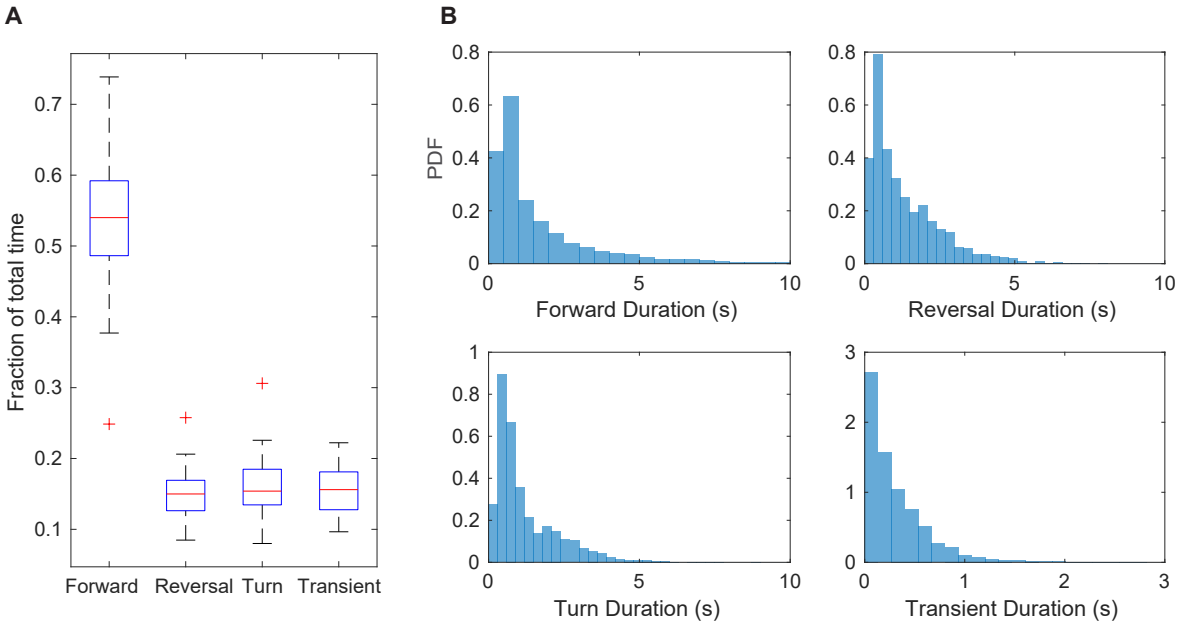


Figure 4.4: A. Fraction of total time spent in different behaviors. **B.** Distribution of the duration of each behavioral bout.

$0.332(0.326, 0.338)s$. The parenthesis are bootstrapped 95% upper and lower confidence intervals. These estimates are broadly consistent with those previously reported in literature [81, 82, 83]. The main difference is in the estimates of the average duration of forward behavior, which appears to be much smaller here. The estimates in the literature range from 5 to 35 seconds [70, 81, 83, 84, 85]. A possibility for this large discrepancy is the difficulty of defining the forward state and distinguishing it from turning and transient behaviors. Indeed, if I redefine the forward state as being composed of turning and transient behaviors then the revised estimate of average forward bout duration is $7.30(6.94, 7.66)s$, well within the range reported in literature so far.

Finally, easy access to transient behaviors is an interesting outcome of state space based discretization of behavior, as they're usually difficult to isolate. An analysis of distribution of transient states shows that the transient states are not randomly placed on the attractor but form a geometrically coherent structure. In Fig. 4.5 I compare the density of states of forward and transient behaviors estimated using a 2d kernel density estimate. As can be seen in Fig. 4.5B, states corresponding to forward behavior are mostly concentrated near zero in the turning and reversal projections, and concentrated strongly on a periodic orbit in the forward projection. On the other hand states corresponding to transient behaviors in Fig. 4.5A do not have a zero value in any of the projections. Importantly, they form a regular geometric structures in all projections, orbits in forward and turning plane, and two symmetrically placed points in the reversal plane. This is reminiscent of invariant manifolds such homoclinic orbits and coherent structures from the theory of dynamical systems which connect different dynamically distinct regions together [66]. Additionally, the transient behaviors show

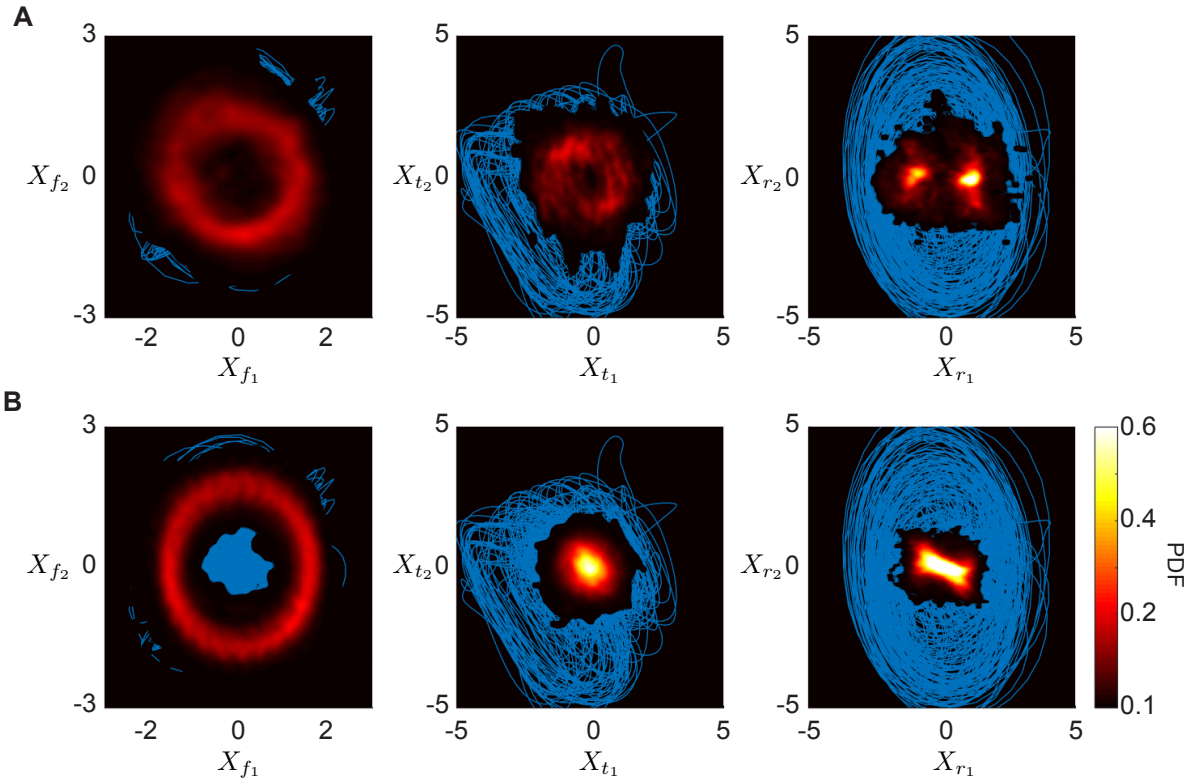


Figure 4.5: Transition zones form a coherent structure in state space. **A.** The density of points in state space corresponding to transient behaviors, i.e., the states where energy is roughly equally distributed amongst the three behavioral modes. The state space trajectories are overlayed for reference. **B.** Same as above, but for states corresponding to forward behavior.

increased density at the transition regions between behaviors instead of being diffuse over a region in phase space. Consequently, they allow a window into how worms transition between different behaviors. I conjecture that the these transition regions are indeed coherent structures in the form of homoclinic or heteroclinic orbits. Based on this it is suitable to hypothesize that at least in principle some of the behavioral transitions can take place spontaneously. without any outside influence. At the very least the topology of these structures means that the effort required to transition is small.

Time Averaged Mode Energies Allow Measurement of Long Timescale Behaviors

Freely foraging worms show a long term behavior lasting up to $15-20$ minutes where after leaving a food filled plate they search an area densely, increasing their frequency of reversals and omega turns [86]. Over time there is a transition to making fewer reversals and turns, and increased forward movement leading to a more ballistic search. In some sense then, the average bout duration estimated above is arguably not the complete picture, as the bout durations are themselves changing slowly over time. A natural question then is can we get to these long timescale behaviors lasting minutes, or possibly longer. To capture long timescale behavior, I define a running time average of the mode energies as follows

$$\overline{E_i}(T) = \frac{1}{T}\int_0^T E_i(t)dt \tag{4.2}$$

Where i can be f for forward, r for reversal and t for turning mode. The time averaged energies average over small scale fluctuations in mode energies seen in Fig. 4.3A, but preserve long time scale structure. In Fig. 4.6 I plot the time averaged normalized energy of different modes against each other. We see that between 5-10 minutes, during the area restricted search phase, all modes have equal energy. Thus, it appears that equipartition of average energy in state space corresponds to a a diffusive search which high frequencies of reversals and turning. Over time we can see that the energy balance shifts such that forward mode is mostly active after about 15 minutes, leading to the ballistic search part of the search strategy.

The datasets I analyzed do not have truly long time scale behavior, such as roaming dwelling, or aging related behaviors and thus do not allow a proper test of this method. Nevertheless, the ability show long term patterns illuminating search behaviors in foraging worms gives me confidence that the basic idea should work for arbitrarily long time scale behaviors.

Embedding of Escape Behaviors is Similar to Foraging Worms

Next, I attempt to form an embedding of the escape response dataset. This dataset is qualitatively different from the foraging dataset. First, the behaviors are different. Unlike escape behavior, foraging isn't driven by a direct sensory stimulus. Second, instead of one long recording in the case of foraging, for escape sequence we have

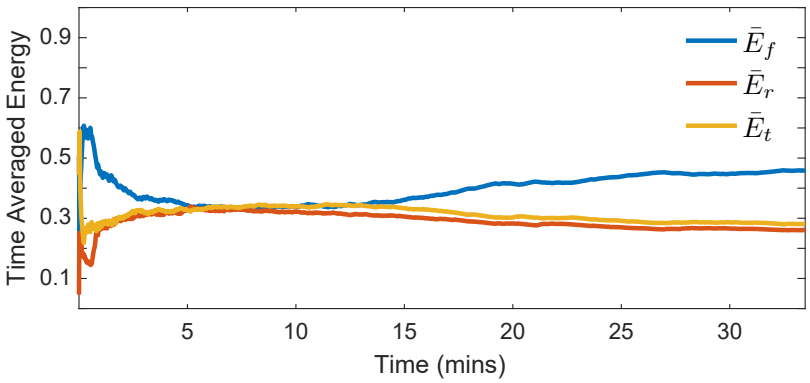


Figure 4.6: The time average of the normalized mode energies captures the long time scale non-stationaries of foraging worms. Shown here is an example trace of a worm for a 30 minute recording. We first see that the contribution of different modes is not static but keeps changing over the time-scale of minutes. In this example The forward mode dominates for the first 3-4 minutes, after which all modes have roughly equal energy. This equipartition phase corresponds to the area restricted search behavior which makes the initial part of their search strategy. After about 15 minutes, the forward mode again dominates as worms transition to a more ballistic search strategy.

$N = 92$ short recordings. This means that the variability we are sampling in the escape response dataset is over the population of N2 worms, while the variability in foraging dataset comes from a single worm's behavior over time. Given these differences, we do not expect a strong relationship between the two embeddings. Surprisingly, however, the reconstructed state space for escape response dynamics is nearly identical to that of foraging worms.

Fig. 4.7 shows the reconstruction for the escape response dataset. Similar to crawling we find that an embedding window between $K = 10$ to $K = 15$ maximize predictability, Fig. 4.7A. I choose $K = 11$ so that the results are comparable to crawling dataset. The embedding dimension is also equal to $m = 6$, as beyond $m = 6$, there is no significant gain in predictability Fig. 4.7B. The state space trajectories with a sample escape sequence overlayed is shown in Fig. 4.7C. The state space trajectories are nearly identical to the ones constructed from foraging behavior in Fig. 4.2B. These observations suggest that the embedding constructed by the pipeline described in this chapter is capturing essential aspects of worm's locomotion that are the same across different experimental conditions and across different worms.

We can also estimate normalize mode energies averaged over all the $N = 92$ worms. This is shown in Fig. 4.8. The mode energies allow us to clearly see the dynamics involved during an escape response. The initiation of an escape response involves a simultaneous excitation of reversal modes, and a suppression of turning modes. They essentially mirror each other, suggesting a common mechanism underlying the excitation of reversal and suppression of body bending. This is followed by a slow decay of reversal modes to baseline, while the turning modes rise until they peak at about 8 seconds, which corresponds to the turning part of the response. Finally, after about

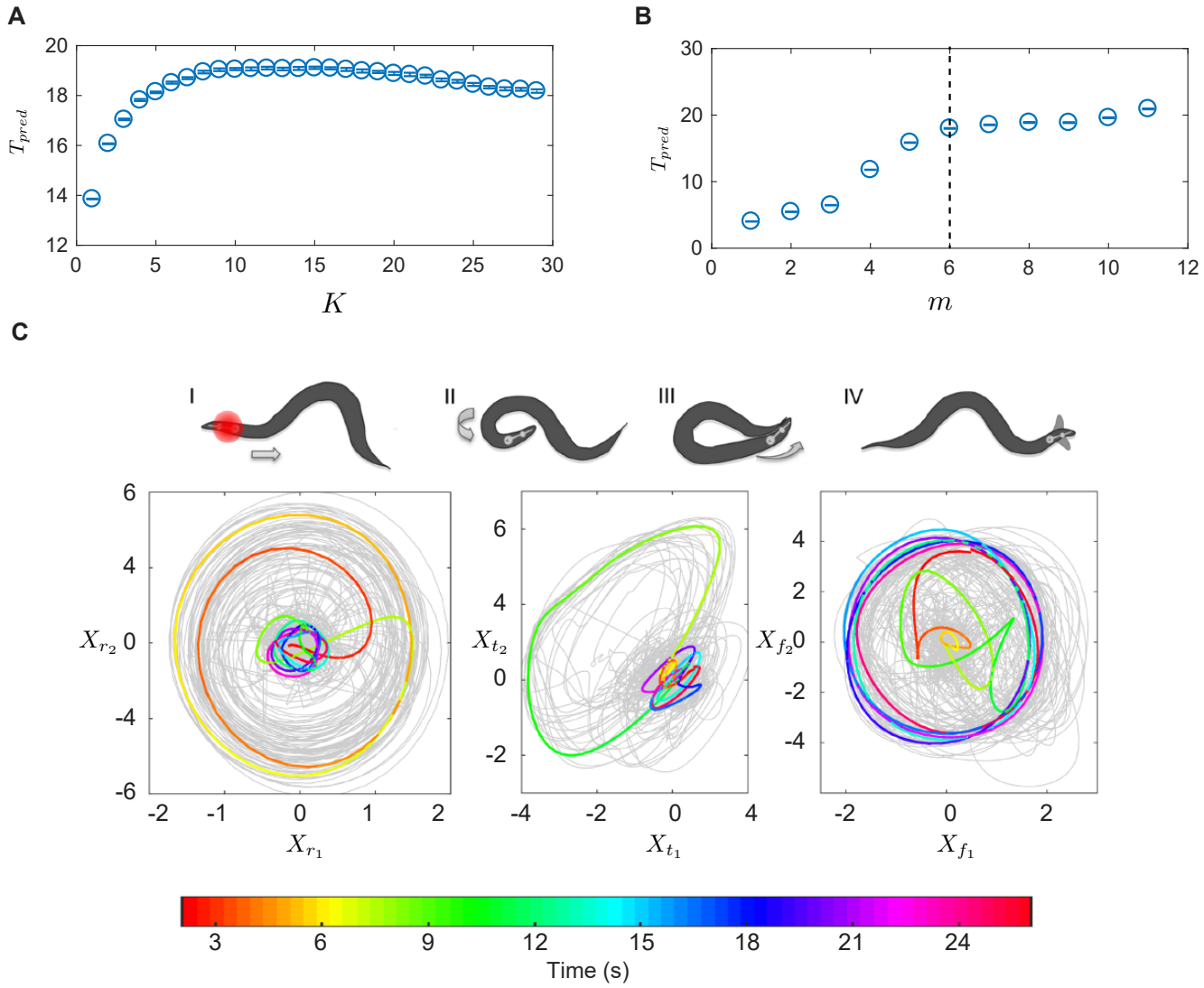


Figure 4.7: Embedding of Escape Response **A.** Predictability time T_{pred} plotted as a function of window length K for the escape response dataset. The trend is broadly similar to the crawling dataset, with an optimum window between $K = 10$ and $K = 15$ frames. I choose $K = 11$ so that the embeddings are comparable to the crawling dataset. **B.** T_{pred} plotted as a function of embedding dimension. Similar to crawling we see that the predictive performance saturates after $m = 6$ dimensions. **C.** State space trajectories of a single escape event. The reconstructed state space is similar to the one constructed from foraging dataset.

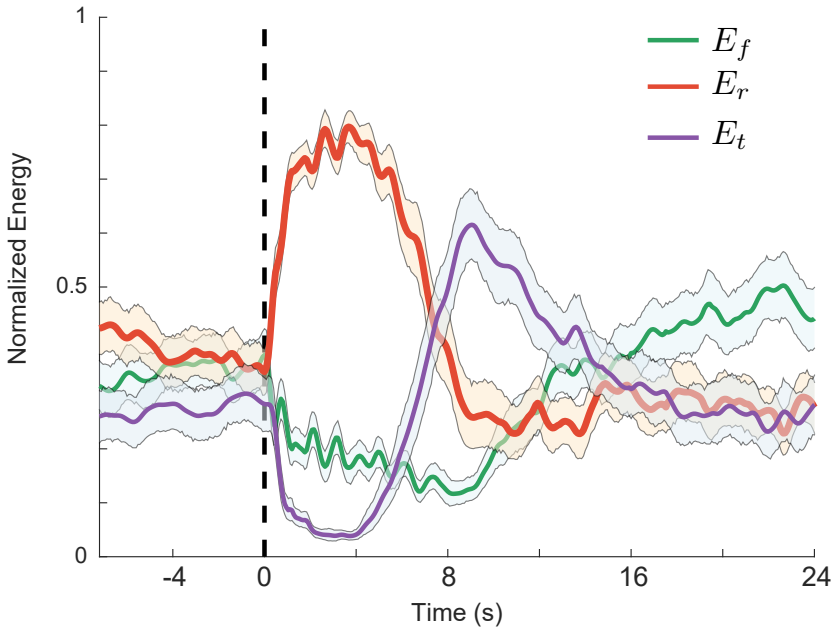


Figure 4.8: Normalized mode energies for the escape response averaged across all $N = 92$ worms.

16 seconds, the worms resume forward movement in the opposite direction.

Embedding of Swimming Gait is a Regular Oscillation with Odd Harmonics

The swimming gait of *C. elegans* is strikingly different from the crawling gait. Worms make faster, long wavelength C-shaped body bends when swimming in a low viscosity medium. This is shown schematically in Fig. 4.9A. The predictability time T_{pred} varies sinusoidally as a function of K (Fig. 4.9B), peaking at exactly one full period, which happens at $K = 50$ corresponding to an oscillation frequency of 2.5 Hz. Once again, we see that T_{pred} as a function of embedding dimension stops increasing after $m = 6$, implying that 6D are enough to capture the dynamics of swimming gait. In Fig. 4.9C, I plot the three dimensional reconstruction of the swimming state space, color coded by the 4 dimension. They phase space is essentially a thick bundle of periodic orbits. Frequency analysis reveals more information about the 6 components. The PSD of all the 6 components is plotted in Fig. 4.9D, and shows that the first two components (top panel) oscillate at a frequency of 2.5Hz. The 3rd and 4th components (middle panel) show a mixture of the primary frequency and it's third harmonic equal to 7.5 Hz. The remaining components (bottom panel) show a peak only at the 3rd harmonic.

It is not clear in advance why the second harmonic is absent from the swimming gait. In the final chapter I put forward a qualitative hypothesis that a feedback control signal which keeps the average mode energies constant, can lead to oscillations without even harmonics in the normal modes.

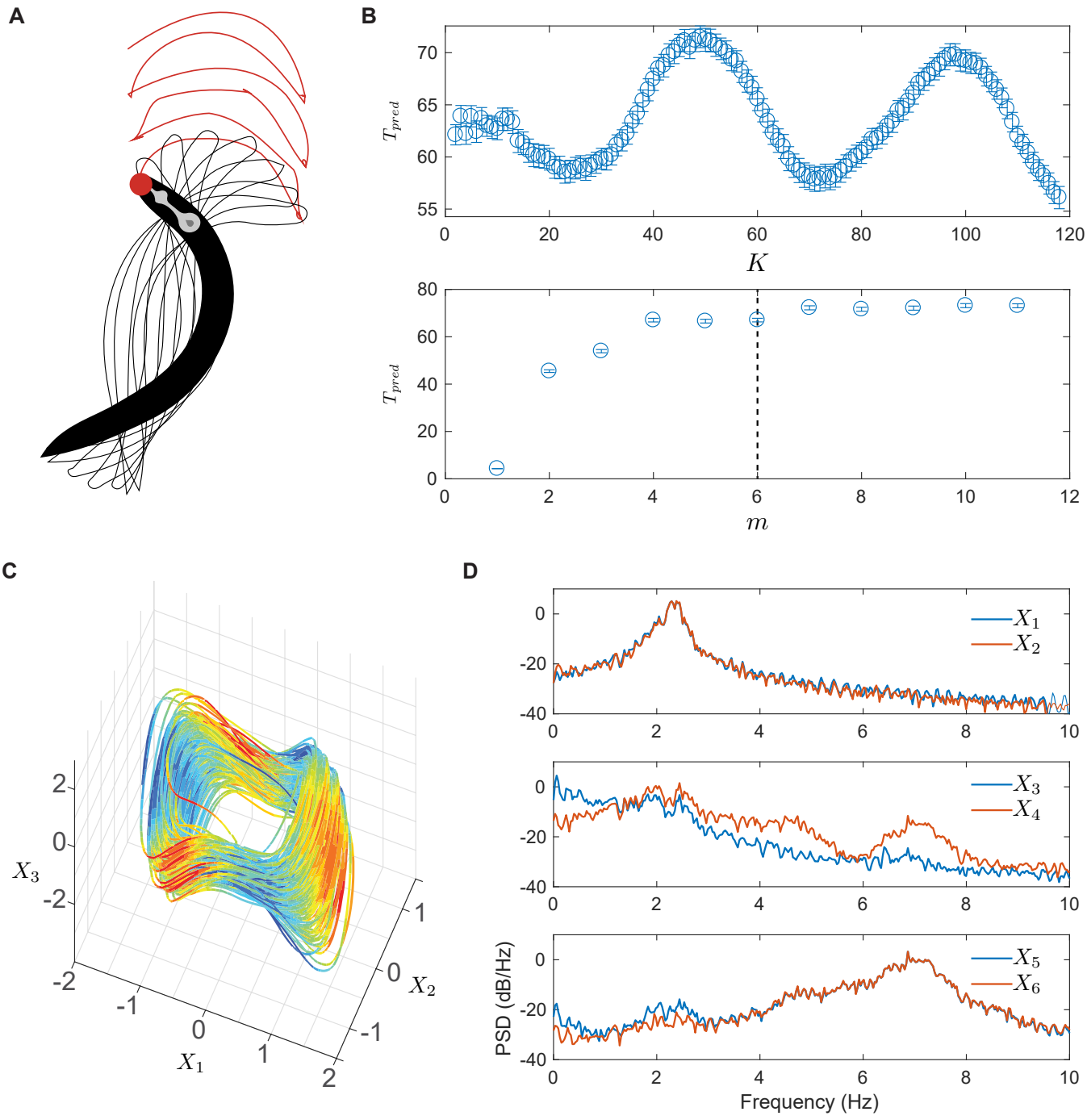


Figure 4.9: Embedding of a swimming worm **A.** A schematic of a swimming gait, which is different from the crawling gait. **B.** T_{pred} varies sinusoidally as a function of K, peaking at exactly one full period, which happens here at $K = 50$ corresponding to an oscillation frequency of 2.5 Hz. Bottom panel shows T_{pred} as a function of embedding dimension. Once again we see that predictability stops increasing after $m = 6$. **C.** Three dimensional reconstruction of the swimming state space, with color being the 4 the dimension. **D.** PSD of all the 6 components shows that the first two components (top panel) oscillate at a frequency of 2.5 Hz. The 3rd and 4th components (middle panel) show a mixture of the primary frequency and it's third harmonic equal to 7.5 Hz. The remaining components (bottom panel) show a peak only at the 3rd harmonic.

4.4 Conclusion

In this chapter I have described a new method of analyzing *C. elegans* behavior. The reconstructed state space is highly interpretable, and allows us to capture behavior across different time-scales, from instantaneous posture dynamics, to discrete behavioral categories over time-scale of a body wave, and to even longer time-scales of minutes to hours. Additionally, the state space trajectories allow a deeper look into how worms transition to different behaviors, a process that is usually difficult to study because of the lack of a clear definition of a "transient behavior".

The ability of the proposed framework to capture detailed continuous dynamics will be important for interpreting neuroimaging data. Typically, continuous time neural activities are correlated with discrete, coarse grained behavioral measures such as centroid velocity, or transition rates [87]. However, postural dynamics are more closely related to neural activities. Indeed, a recent study correlated whole brain neural activities to postural dynamics obtained from phase velocity in the (a_1, a_2) eigenworm planes. They showed that direct connections between continuous time postural dynamics and neural activities [88]. The framework proposed in this thesis is ideally suited for interpreting neural recordings. Mutant analysis meant to understand the genetic correlates of worm behavior as done in [89, 90], has also mostly focused on discrete time postural representations. A continuous time framework like the one proposed in this thesis will also be beneficial for such questions.

4.5 Supplementary Images

Fig. 4.10 shows the temporal sequence of modes for the crawling dataset.

Fig. 4.11 shows prediction error $E(\tau)$ for different embedding dimensions.

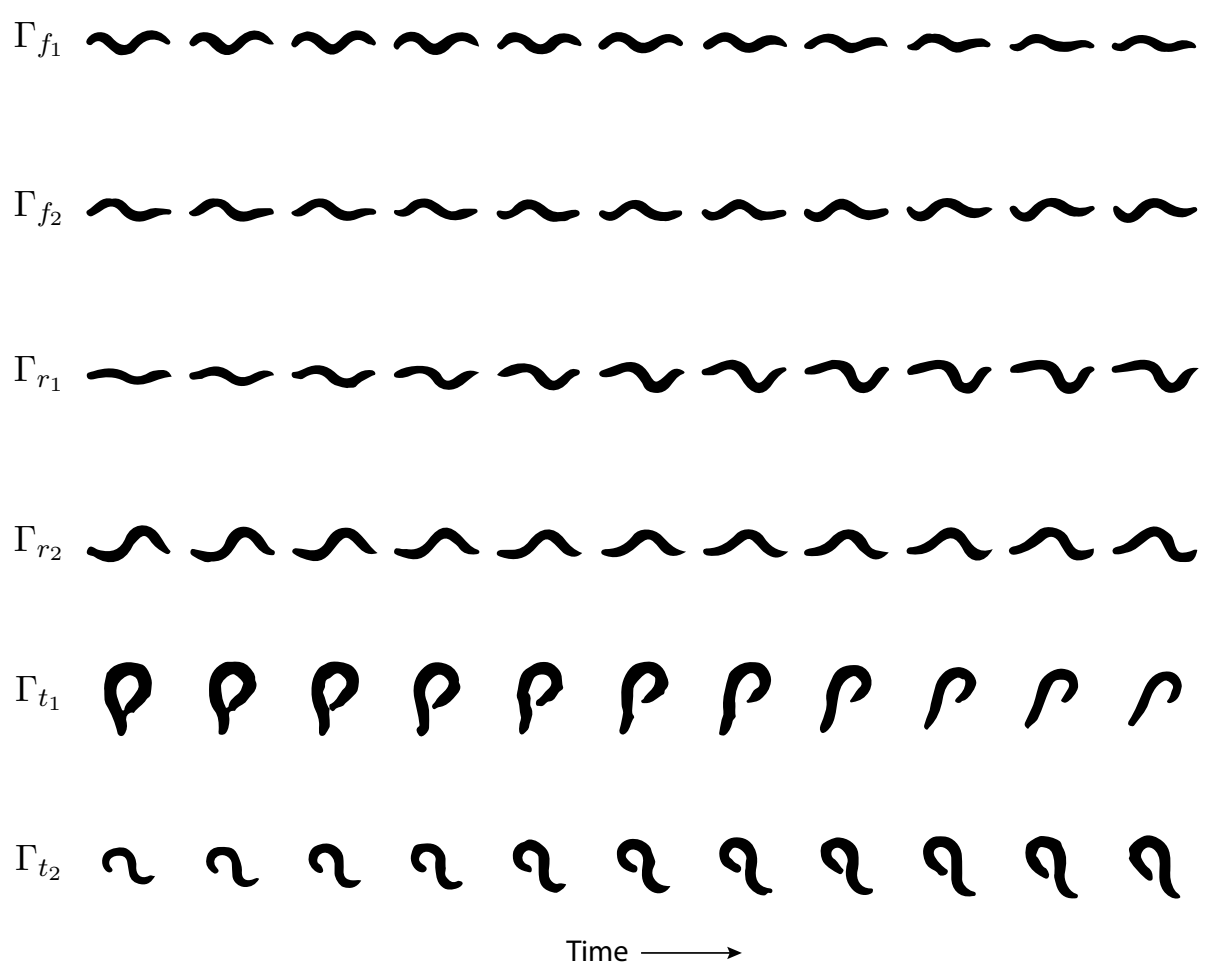


Figure 4.10: Shown here is the temporal evolution of all the 11 frames for 6 modes in Fig. 4.2.

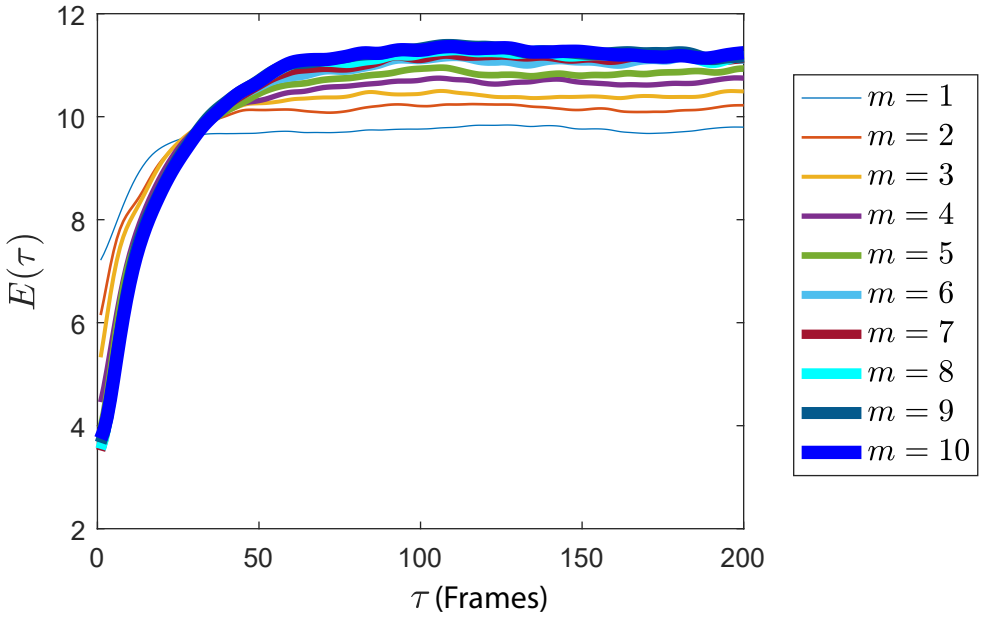


Figure 4.11: $E(\tau)$ curves for a single worm for different embedding dimensions. The error curves collapse onto each other for $m \geq 6$ indicating that $m = 6$ is a good estimate for the embedding dimension

.

Chapter 5

Estimation of Lyapunov Exponents for *C. elegans* Behavioral Dynamics

5.1 Introduction

In this chapter I will utilize all the methods developed in this thesis thus far to estimate the spectrum of all six Lyapunov exponents. The estimated exponents allow an unprecedented look into the complexity of natural behavior in *C. elegans*. Local Lyapunov exponents allow a measurement of instantaneous variability of worm behavior. While global exponents allow a characterization of average variability along with important symmetries and invariances involved with worm locomotion. The key finding is an observation that the Lyapunov spectrum of worm behavior belongs to the class of damped driven Hamiltonian dynamics, such as the damped driven Duffing oscillator discussed in chapter 3. The damped driven Hamiltonian nature, allows a natural decomposition of the equations of motion into worm's biomechanics, environmental effects and neuromuscular control. In the final section of this chapter I exploit the damped driven Hamiltonian structure to propose a way to measure effective control signals from behavioral measurements alone. Based on these estimates, I put forward a conjecture that part of the feedback control signal in the worm serves to keep the average mode energies constant.

5.2 Measurement of the Lyapunov Spectrum of *C. elegans* Locomotory Dynamics

As we saw in chapter 3, chaotic dynamics are characterized by the fact that on average, the distance between any two neighboring points, δ_t grows exponentially with time as $\delta_t = \delta_0 e^{(\lambda_1 t)}$, where λ_1, or the maximal Lyapunov exponent (MLE) is the average rate of exponential separation. Positive MLE is a signature of chaotic dynamics and is known to underlie the exponential loss of predictability in chaotic systems [44, 69]. More generally, points in a local neighborhood of a chaotic phase space are sheared by the flow, and get simultaneously stretched and squeezed along different directions in phase space. Which means that for a complete characterization of an m dimensional

chaotic phase space, we need the spectrum of m Lyapunov exponents, $\lambda_i, 1 \leq i \leq m$ measuring the rates of separation along the m orthogonal dimensions.

In the absence of equations of motions, the Lyapunov exponents carry invaluable information about the dynamics of a system. In particular, the sum of all the exponents measures the average dissipation rate of a dynamical system, being zero for systems that conserve energy and negative for systems with dissipation. Then, the sum of the positive exponents gives the average rate of information generation in a dynamical system, also known as the Kolmogorov-Sinai (KS) entropy and is a measure of the unpredictability or complexity of the dynamics (it is what determines T_{pred}). Another measure of dynamical complexity, the dimension of the manifold on which the phase space trajectories lie can be estimated from the Lyapunov exponents as per the Kaplan-Yorke relation. And finally, the Lyapunov exponents can be very useful in determining the underlying symmetries of the dynamical system.

To estimate the Lyapunov spectrum from data, I follow the algorithm developed in chapter 3. Briefly, I first approximate the local Jacobians, $\mathbf{J}(\boldsymbol{x}_t)$ by a weighted least square fit to the local dynamics. The weighting parameter θ is such that nearby points are given a larger weight compared to far away points. The Lyapunov exponents are then obtained by the long time average of the logarithm of eigenvalues of the product of the $\mathbf{J}(\boldsymbol{x}_t)^T\mathbf{J}(\boldsymbol{x}_t)$ matrices, which for the reasons of numerical stability are estimated by a repeated orthonormalization of the $\mathbf{J}(\boldsymbol{x}_t)$ matrices done using a recursive QR decomposition. I also estimate the local Lyapunov spectrum, by taking the logarithm of the square root of the eigenvalues of individual $\mathbf{J}(\boldsymbol{x}_t)^T\mathbf{J}(\boldsymbol{x}_t)$ matrices for each point in the phase space. These local exponents carry important information about the fluctuations in local predictability of the dynamics. In particular, the local KS entropy estimated from the sum of the positive local Lyapunov exponents correlates well with the complexity of the centroid track made by foraging worms on the agar plate, Fig. 5.1. It is maximum when the worms are being unpredictable, for instance when they make a complicated sequence of sharp turns, while it's minimum when worms are engaged in highly stereotyped behaviors such as simple straight line motion. Thus the local KS entropy forms a principled measure of instantaneous variability and unpredictability of worm behavior.

A bootstrap estimate of the probability distributions of the global Lyapunov exponents from worm's behavioral phase space are shown in Fig. 5.2A. I find that there are two positive exponents implying that there are two directions in phase space along which points in a neighborhood get stretched ($\lambda_1 = 0.64\,(0.61, 0.68)s^{-1}$ and $\lambda_2 = 0.21\,(0.19, 0.23)s^{-1}$). Where the numbers in parenthesis indicate bootstrap estimates for 95% CI lower and upper bounds. The sum of positive exponents which bounds the KS entropy was estimated to be $\hat{h}_\mu = 0.86\,(0.8, 0.92)s^{-1}$. Thus, the positive exponents set the predictability time scale at about 3/4th a body wave.

Next, I find that $\lambda_3 = -0.02(-0.04, 0)s^{-1}$, and can be treated as being equal to zero within the estimation error. As mentioned earlier zero exponents usually signal the presence of continuous symmetries in the underlying equations of motion and it can be shown that all systems resulting from continuous (deterministic) dynamical systems must have at least one zero exponent corresponding to the time translation invariance [60]. In these systems the zero exponent corresponds to the direction of flow in phase space, along which no stretching or expansion can take place [60]. This again suggests

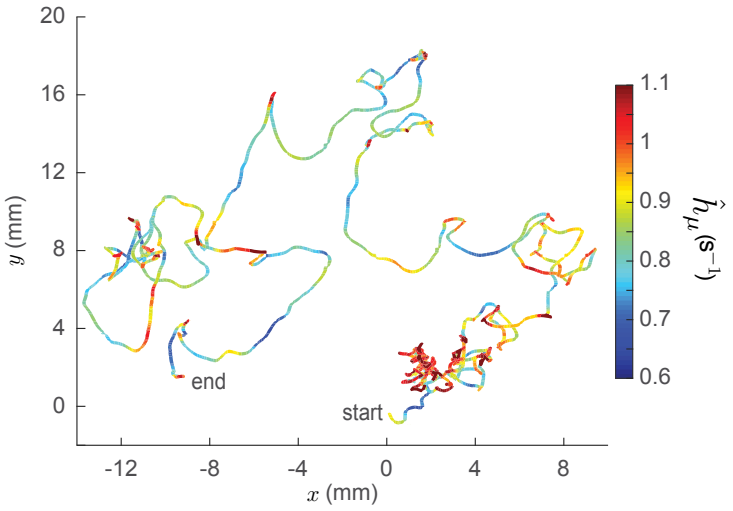


Figure 5.1: Local KS entropy $\hat{h}_\mu$ estimated from local Lyapunov exponents show significant variation across the worm's foraging behavior that correlates with the complexity of the track.

that worm behavior is governed by underlying set of deterministic differential equations. The remaining exponents are all negative ($\lambda_4 = -0.24(-0.26, -0.22)s^{-1}, \lambda_5 = -0.50(-0.54, -0.47)s^{-1}$ and $\lambda_6 = -0.94(-1.00, -0.88)s^{-1}$). These exponents counter the expansion by the positive exponents.

I also find that the sum of all the Lyapunov exponents, which is equal to the volume expansion rate in phase space is negative, implying that worm dynamics are dissipative on average. This means that even though trajectory bundles expand locally in phase space, the dissipation causes the phase volumes to contract as a whole and relax to an attracting set. These conflicting requirements of local expansion and global contraction are essence of behavioral complexity and cause the attractor to have a complex fractal structure with a non-integer dimension (also known as a strange attractor). I estimate the fractal dimension of the attractor by the KY relation to be $D_{KY} = 5.1(5.04, 5.25)$, which is only slightly less than the dimension of the full phase.

Curiously, I observe observe a striking symmetry in the Lyapunov spectrum, such that the exponents come in conjugate pairs that sum to the same number $\alpha = -0.28\,(-0.3, -0.26)s^{-1}$ Fig. 5.2B,C. This results in a striking symmetry of the entire spectrum about $\frac{\alpha}{2}$, Fig. 5.2A,B. Lyapunov spectra that are symmetric about zero, such that conjugate pairs sum to $\alpha = 0$ are a hallmark of conservative or Hamiltonian dynamical systems. In these systems the stretching in phase space is exactly balanced by shrinking resulting in the sum of the exponents being zero, as a consequence the phase space volume remains constant. This pairing of exponents, also known as the symplectic structure of the phase space is preserved when a viscous damping term is added to the system, but instead of summing to zero, the exponents sum to a constant value equal to the coefficient of damping α. Such a symmetry has been previously observed in two classes of systems, both derived from Hamiltonian dynamics. Coupled oscillators with viscous damping where α is simply the dissipation per degree of free-

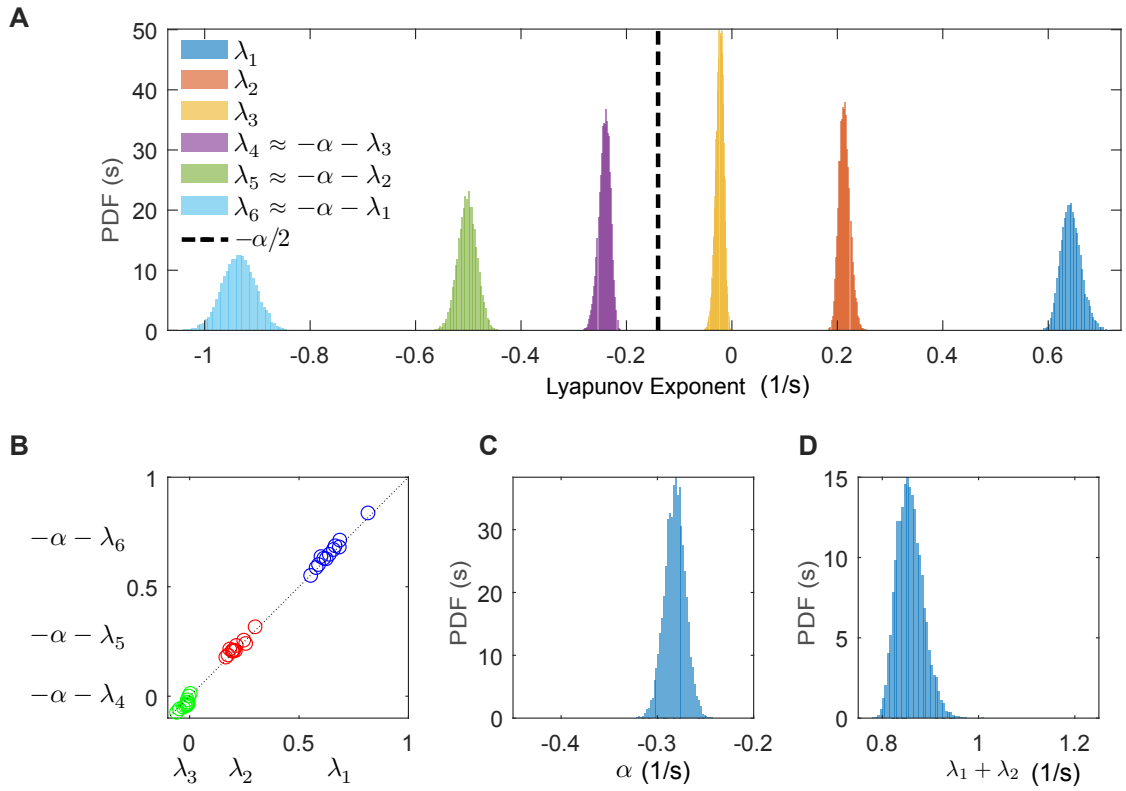


Figure 5.2: The Lyapunov spectrum for the reconstructed phase space of foraging worms. **A.** A bootstrapped probability distribution across all $N = 12$ worms is shown here. There are two positive exponents, one zero exponent indicating the direction of continuous flow and three negative (dissipative) exponents. The spectrum is symmetric about the point $\alpha/2$ and this symmetry is an indication of an underlying Hamiltonian structure in the dynamics. The small negative value of α indicates that worm's dynamics are surprisingly only mildly dissipative. **B.** A more direct measurement of the symmetry of the spectrum shown by plotting λ_i and $-\alpha - \lambda_{m-i+1}$ against each other. Their equality is shown by the fact they cluster around the $x = y$ line. **C.** Distribution of the dissipation constant α. **D.** Distribution of the sum of positive exponents which estimate the Kolmogorov-Sinai entropy $\hat{h}_\mu$.

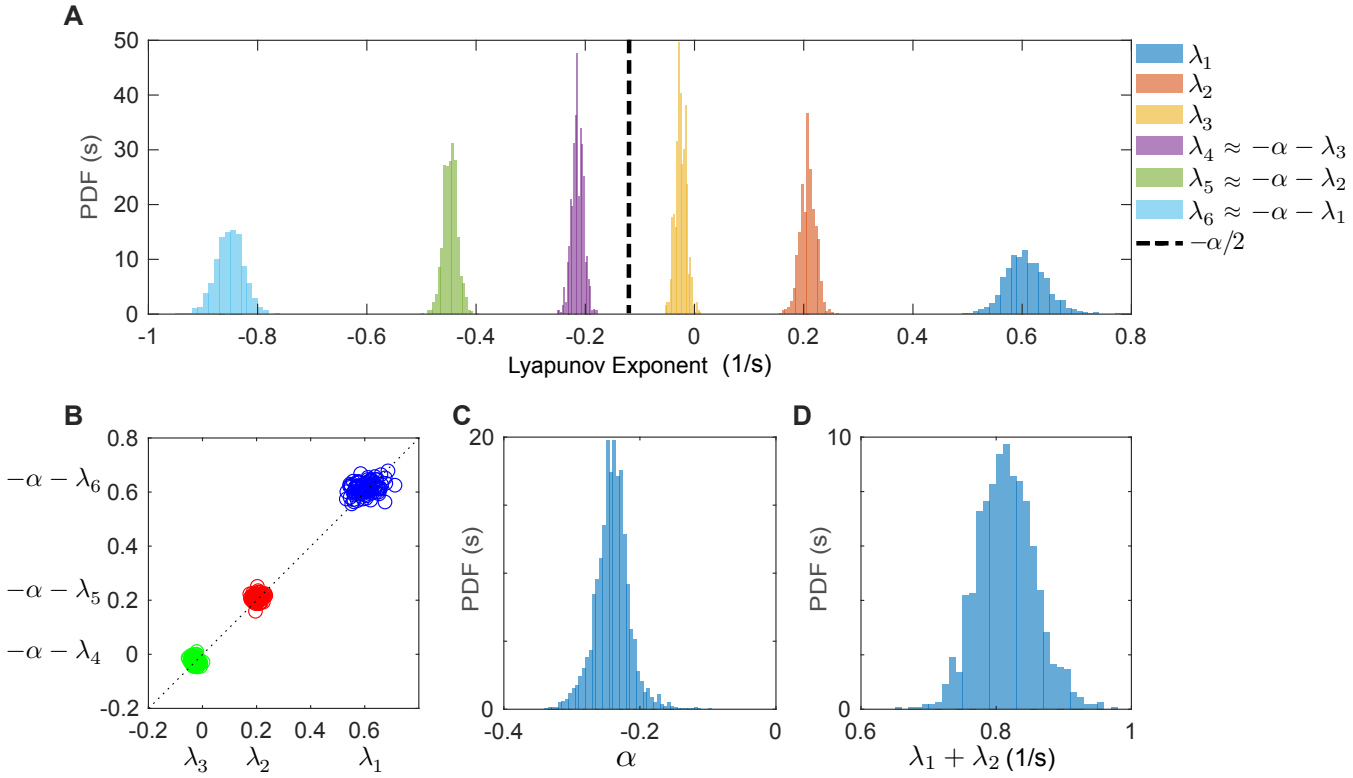


Figure 5.3: Lyapunov Spectrum for escape response dataset is nearly identical to that of foraging worms. **A.** Bootstrapped probability distribution across all $N = 92$ worms. **B.** Conjugate exponents plotted against each other to measure the symmetry. **C.** Distribution of the dissipation constant α. **D.** Distribution of the sum of positive exponents estimating the Kolmogorov-Sinai entropy $\hat{h}_\mu$.

dom [61], and thermostatted molecular dynamic simulations where the α is a feedback friction force that acts to maintain a dynamic equilibrium by either keeping the total internal energy or the kinetic energy of the particles constant [62]. In either case these observations lead us to conclude that worm dynamics belong the family of damped driven Hamiltonian systems.

A natural question is whether the damped driven Hamiltonian structure we observe is specific to the crawling worm or does it generalize across worm behaviors. To test, I repeat the analysis pipeline first on worms executing an escape response Fig. 5.3, and worms swimming in a liquid Fig. 5.4. Escaping worms test whether the symmetry is a consequence of stationary motion in the absence of direct sensory stimulation. While swimming worms are an important test case because of their qualitatively different gait. In both cases we observe that the estimated Lyapunov spectrum has the same qualitative structure with 2 positive exponents, one zero and three negative with the sum being negative. We again see a symmetric Lyapunov spectrum in both cases, suggesting that the damped driven Hamiltonian structure we observe is not specific to the crawling gait and might be a general feature of worm behavior.

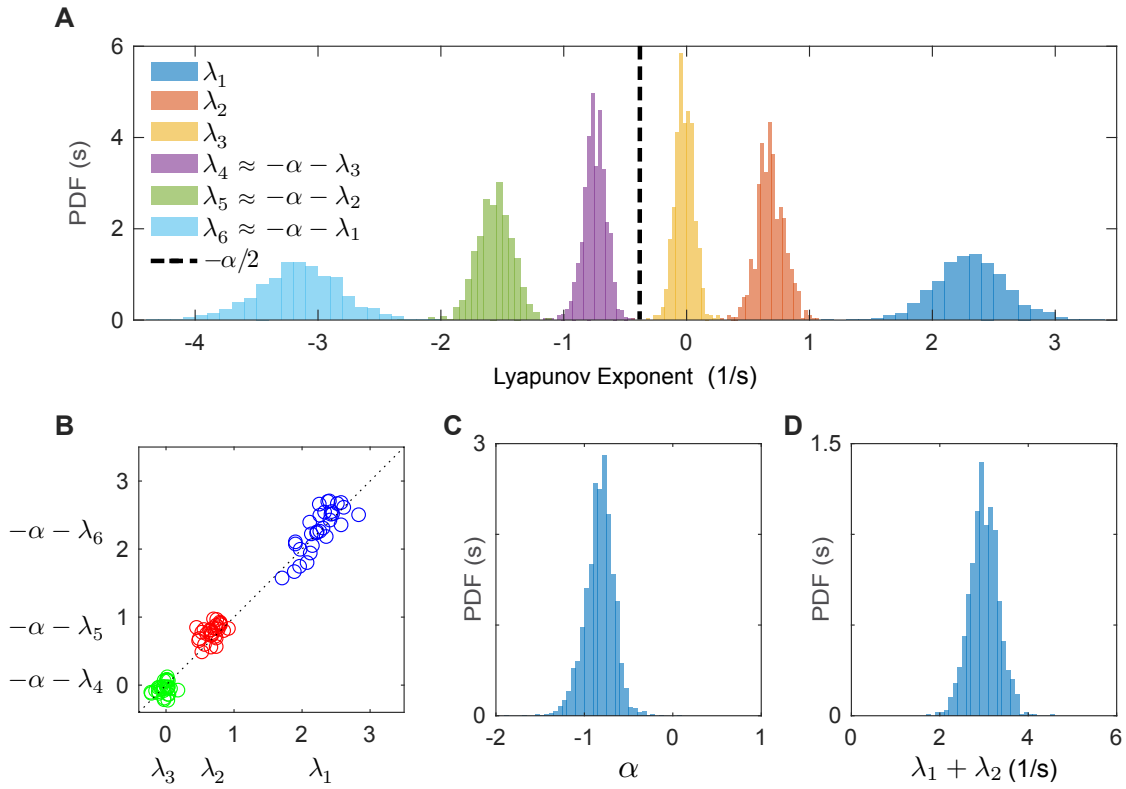


Figure 5.4: Lyapunov Spectrum for swimming dataset is quantitatively different but qualitatively similar. **A.** Bootstrapped probability distribution across all $N = 30$ worms in the swimming dataset. **B.** Conjugate exponents plotted against each other to measure the symmetry. **C.** Distribution of the dissipation constant α. **D.** Distribution of the sum of positive exponents estimating the Kolmogorov-Sinai entropy $\hat{h}_\mu$.

Estimating the Effective Control Signal Using Local Lyapunov Exponents

The finding that damped driven Hamiltonian dynamics underlie worm locomotion allows to interpret the normal modes as coupled damped driven Hamiltonian oscillators. Furthermore, we can propose that the three oscillators obey the following equations of motion,

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} - \eta p_i \ + f_i \Omega_i(t) + C(p_i, y(t)) \end{aligned} \tag{5.1}$$

Where (q_i, p_i) are coordinates for i_{th} oscillator. The function $H(q_i, p_i)$ is the total energy of the oscillators, η is the environmental damping, $f_i\Omega_i(t)$ is the oscillatory driving force to each oscillator and $C(p_i, y(t))$ is the neuromuscular feedback signal which depends on the sensory stimuli $y(t)$ and momentum variable p_i. The above equation enables a decomposition of worm behavior into biomechanics of worm's body, dissipative forces in the environment and neuromuscular feedback control forces due to neural processing of various sensory stimuli.

A potentially useful consequence of damped driven Hamiltonian dynamics is that it allows us to measure parts of the control signal of the worm from behavior alone. The sum of the Lyapunov exponents, which is equal to the total volume dissipation rate of the dynamics is formally given by the divergence of eqn 5.1, as follows

$$\sum_{i}^{6} \lambda_i \approx \nabla.V = -3\eta + \sum_{i}^{3} \frac{\partial C(p_i, y(t))}{\partial p_i} \tag{5.2}$$

Thus, remarkably, the local Lyapunov exponents can be used to extract information about the control signal in the worm from behavioral measurements alone!

Energy Based Feedback Control

To test the hypothesis mentioned above, I estimated the sum of local Lyapunov exponents for swimming data first. A short sequence is shown in Fig. 5.5A. The estimated signal is noisy, but it has a clear oscillatory structure. The PSD of the sum of Lyapunov exponents shown in Fig. 5.5B, shows that it has a peak at about 5Hz, or double oscillation frequency of swimming worms which was estimated in Fig. 4.9D.

This is an interesting observation because as we saw in Fig. 4.9D, swimming worms only displayed odd harmonics. Therefore it is not immediately clear why the sum of local Lyapunov exponents would show a peak at double the oscillation frequency. A possibility inspired by energy based control in thermostatted systems [62], is that at least part of the feedback signal $C(p_i, y(t))$ acts to keep the energy of worm's movement constant. The simplest form of such an energy based control involves setting $C(p_i, y(t)) = -\gamma(p_i^2 - E_0)p_i$ [91]. This means that the feedback is proportional to difference between the energy of a mode measured by p_i^2 and some reference energy

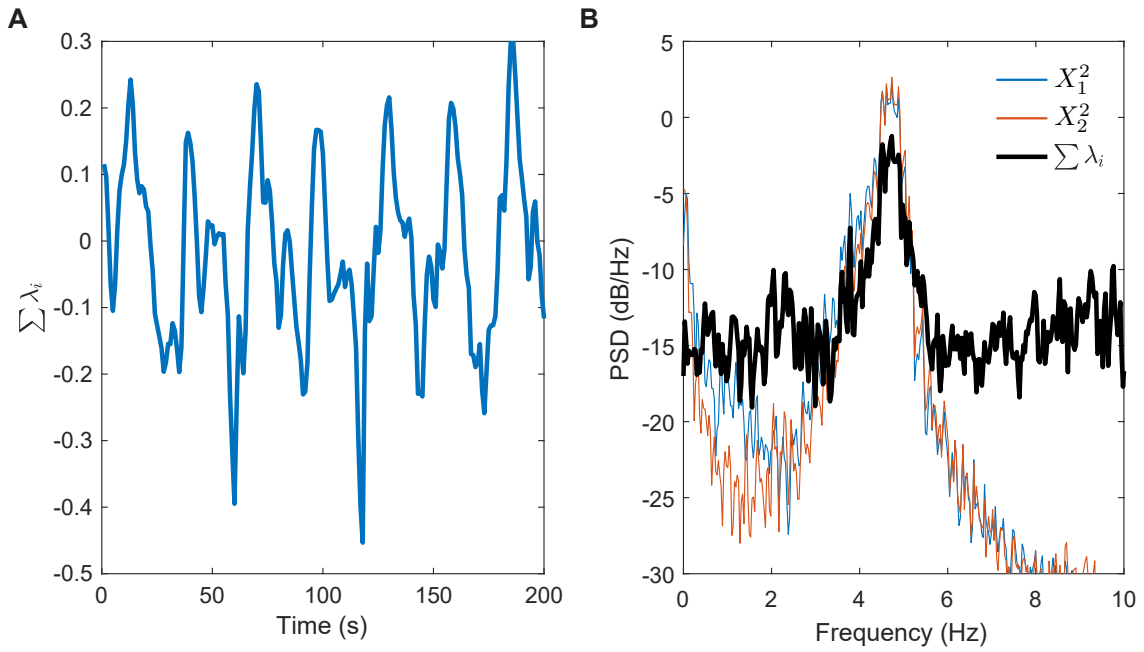


Figure 5.5: A. A short time-sequence of the sum of Lyapunov exponents for swimming worms shows an oscillatory structure. **B.** PSD estimate shows that the signal in A. has peak at double the frequency of oscillation of swimming worms. The structure of the power spectrum is matched well by the PSD of the square of the swimming components X_1^2 and X_2^2

E_0. The control gain γ parameterizes the strength of the feedback. Substituting this feedback term into eqn 5.1 we obtain

$$\begin{aligned} \dot{q_i} &= \frac{\partial H}{\partial p_i} \\ \dot{p_i} &= -\frac{\partial H}{\partial q_i} - \alpha p_i \ + f_i \Omega_i(t) - \gamma(p_i^2 - E_0)p_i \end{aligned} \tag{5.3}$$

If we had a system in the form of eqn 5.3, then we would expect the sum of the Lyapunov exponents (or the divergence of the equation of motion) to be,

$$\sum_i^6 \lambda_i \approx \nabla.V = -3\eta + \gamma E_0 - \gamma p_i^2 \tag{5.4}$$

There are two things to note about the terms on the RHS of eqn 5.4. First the dissipation from the environment η is being countered by a positive term γE_0. This shows one possible way in which a feedback can counter the effects of environmental forces. Second, there is a term proportional to p_i^2. In the case of a sinusoidal motion, if p_i oscillates at frequency f, then p_i^2 oscillates at $2f$, i.e. at precisely double the frequency. Thus, a possible explanation of the doubling of frequency in PSD of sum of local Lyapunov exponents is that the signal shown in Fig. 5.5A is proportional to the square of at least some of the components. To see if this is the case, I also plotted the PSD of X_1^2 and X_2^2 in Fig. 5.5B. It can be clearly seen that there is a close match between the PSD of X_1^2 and X_2^2, and that of the sum of Lyapunov exponents.

The observations in Fig. 5.5 are consistent with the hypothesis that at least part of the feedback signal is proportional to the square (or the amplitude) of the modes. To see if we recover these observations in a known system, I simulated first the linear coupled oscillators described in eqn 1.3, with the energy based feedback as follows

$$\begin{aligned} \dot{q_1} &= p_1 \\ \dot{p_1} &= -2q_1 + q_2 - \gamma(p_1^2 - E_0)p_1 \\ \dot{q_2} &= p_2 \\ \dot{p_2} &= q_1 - 2q_2 - \gamma(p_2^2 - E_0)p_2 \end{aligned} \tag{5.5}$$

The results of this simulation (with $E_0 = 0.6$ and $\gamma = 0.5$ set arbitrarily) are shown in Fig. 5.6A-D. The simulations show several interesting outcomes. First giving a feedback term to keep the energy constant results in mode 1 completely dissipating away as seen in Fig. 5.6A, while mode 2 remains active Fig. 5.6C. In this way, even though we are in principle dealing with a 4D coupled oscillator system, the actual dynamics are only 2D because mode 1 contains negligible energy compared to the second mode n_2. Thus, there is a natural reduction in the dimensionality of the system. Second, the PSD of mode 2 oscillations shows odd harmonics as seen in Fig. 5.6D. This is reminiscent of the odd harmonics oscillations we observed in swimming modes (Fig. 4.9D). Finally, in Fig. 5.6B, we see that PSD of the divergence of eqn 5.5 also shows a peak at double the frequency of mode 2 oscillations (Fig. 5.6D).

Thus a simple model of linear coupled oscillators is consistent with several non-trivial observations. Namely, we observe that the model dynamics become effectively

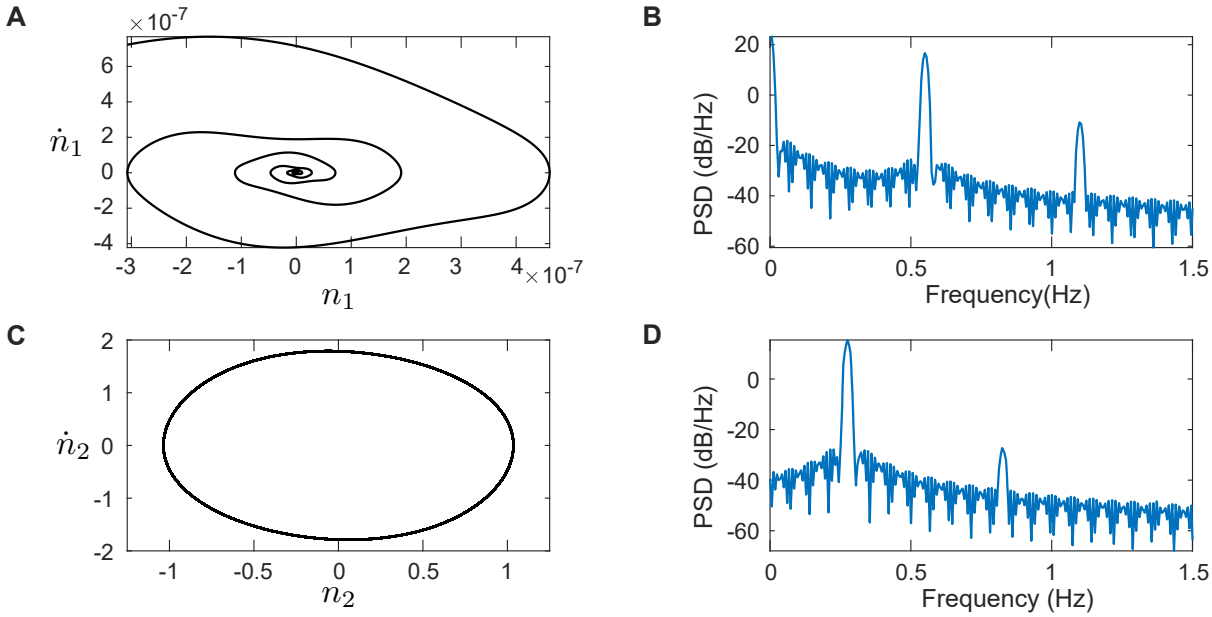


Figure 5.6: A.,C. First mode of the simulated systems shows rapidly decaying oscillations. While the second mode remains active. **B.** PSD of the divergence term in eqn 5.4. **D.** PSD of the mode 2 oscillations. The peak frequency is half the frequency of the divergence signal in eqn 5.4

low dimensional. The feedback results in oscillations containing odd harmonics, and finally the oscillations of the divergence of the equation of motion happen at double the frequency of mode observations. Additionally, we also know from published literature that an energy based feedback can result in a symmetric Lyapunov spectrum.

Goal Driven Control

To test if I can also study goal driven behavior by using the sum of local Lyapunov exponents as a proxy for control signal I considered the transition between forward and reversals during spontaneous locomotion and during an escape response. Worms make a forward to reversal transition as part of the escape behavior when they receive an aversive stimulus, but foraging worms can also make the same transition spontaneously without an apparent sensory stimulus. I expect that only for the case of the escape reversal, control signal estimated by sum of Lyapunov exponents will show a sharp impulsive rise as the worm rapidly changes course from its current state to a reversal on the onset of an aversive stimulus. On the other hand, I do not expect a strong phase locking of the sum of Lyapunov exponents to the onset of spontaneous reversals.

To test these predictions, I compared the sum of local Lyapunov exponents for spontaneous reversals extracted and aligned from the foraging dataset and a set of escape reversals from the escape response dataset. As shown in Fig. 5.7, the sum of Lyapunov exponents for escape reversals shows a strong impulse post stimulus onset, after which it gradually saturates to baseline over the course of next 2-3 seconds. It is worth noting that the actual escape sequence is much longer and lasts for more than 10

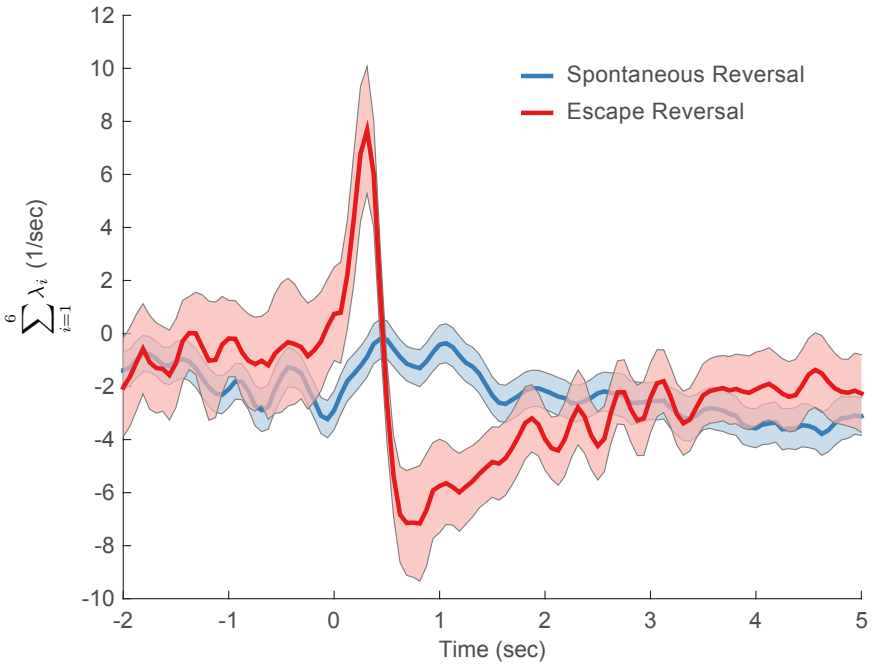


Figure 5.7: Estimated sum of all local Lyapunov exponents for aligned escape (red) and spontaneous reversals (blue).

seconds. Spontaneous reversals on the other hand do not show a significant deviation from baseline in this signal, which again aligns with our expectations and suggests that spontaneous reversals do not necessarily need a strong driving signal.

The observations and experiments in this section not only display the utility of local Lyapunov exponents in studying the control logic of the worm, but also suggest a possibility that worms might keep their movement relatively steady as a result of a feedback that keeps their energy constant. The results are preliminary, but suggest intriguing avenues for further research on account of being consistent with several different observations.

5.3 Discussion

In this chapter I used the methods developed in chapter 3 for robust estimation of local and global Lyapunov exponents to estimate these quantities for all the three datasets, crawling, escape response, and swimming. Global Lyapunov exponents allow us to extract the symmetries of the equations of motion governing *C. elegans* locomotion. The zero Lyapunov exponent suggests that the equations of motion have are time-translation invariant, strengthening the interpretation that the dynamics are mostly deterministic. Surprisingly, I also observe another symmetry which suggests that worm's locomotion is closely related to Hamiltonian dynamics. Specifically, the symmetry is such that the 6 lyapunov exponents come in 3 conjugate pairs that sum to a small negative value. This implies that *C. elegans* locomotory dynamics can be considered as a damped driven Hamiltonian system.

Next, I suggest a useful consequence of the Hamiltonian structure is that the sum of local Lyapunov exponents reflect the non-autonomous terms, and therefore contain information about the control neuromuscular control signals. I then estimate the sum of local Lyapunov exponents for swimming worms, which show an oscillatory structure with double the frequency of the body wave. Implying that the control signal has a frequency that is double the body wave frequency. This can be explained if we assume that part of the control signal is forcing the dynamics so that the average mode energies are constant. I also use sum of local Lyapunov exponents to study the difference between spontaneous reversals and stimulus induced escape reversals. Here, I find that during spontaneous reversals the sum of the local Lyapunov exponents is roughly constant suggesting that they don't require a strong control input. On the other hand, escape reversals show a strong peak briefly after the moment of impulse, reflecting a controlled behavior.

Estimation Issues

In this section I discuss the robustness of the estimate as a function of the scale parameter θ in Fig. 5.8. Most exponents grow as a function of scale, except λ_3 which remains stable around zero Fig. 5.8A. This increases our confidence in the fact that worm dynamics are governed by a deterministic dynamical system. I utilize different measures to judge the quality of Jacobian estimates, as shown in Fig. 5.8A. First, the mean squared error of the prediction of a linear model parameterized by θ is plotted as a function of θ. The values of θ for which the prediction error takes a small value correspond to Jacobian estimates that can be trusted. As can be seen in Fig. 5.8B, there is narrow range of θ where local Jacobians are correctly estimated. The same values of θ also correspond to where λ_3 reaches a value closest to zero (Fig. 5.8B middle panel), and the conjugate exponents sum to the same value implying that the spectrum is symmetric (Fig. 5.8B lower panel). These observations serve to significantly increase the confidence of Jacobian estimates. Finally, there is a question about whether the Lyapunov exponents converge to a stable value, or do they keep changing over time? In Fig. 5.8C I have plotted the running average of various exponents as a function of time. It can be seen that the estimates are quite stable.

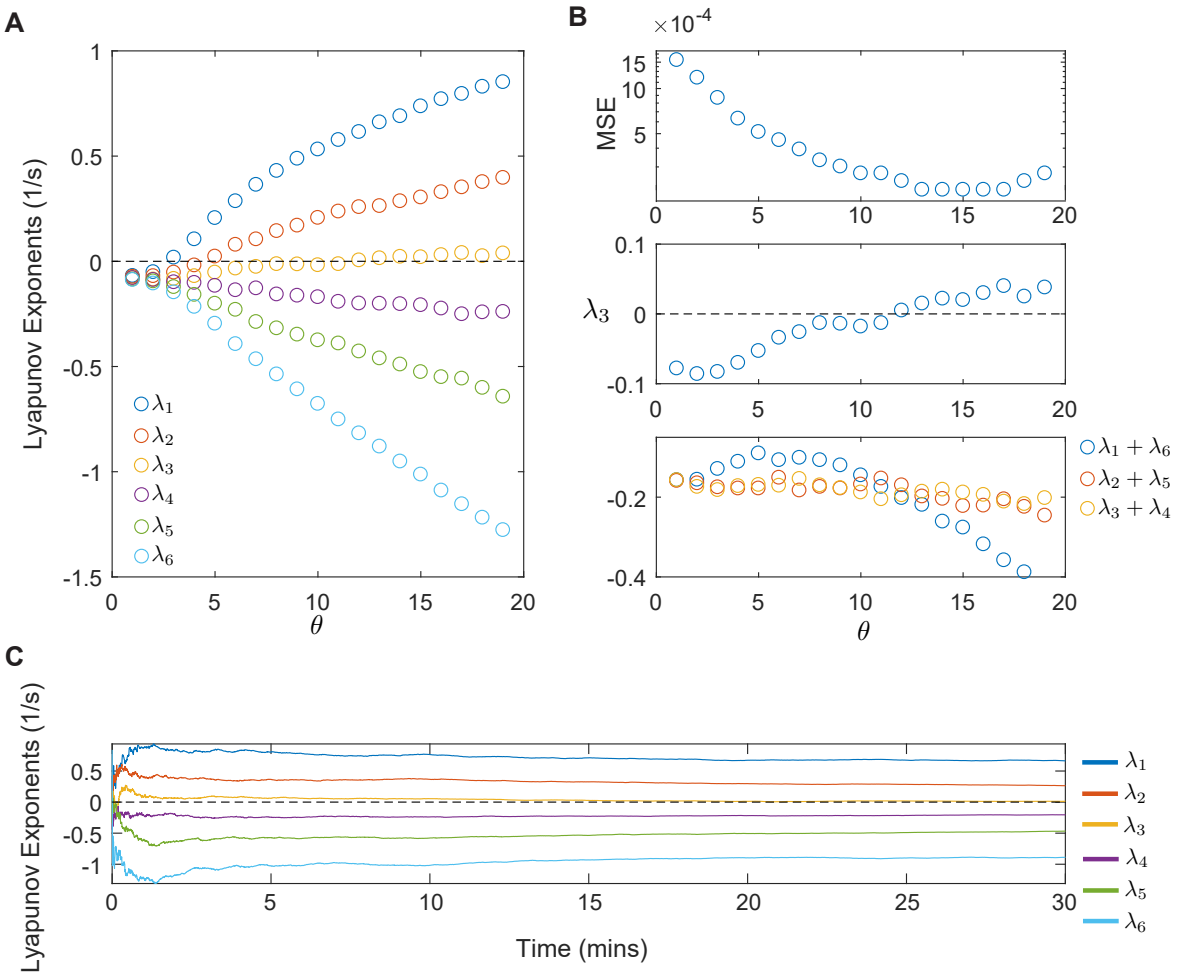


Figure 5.8: Lyapunov spectrum estimation for a worm. **A.** Lyapunov exponents are plotted against the scale parameter θ. Most exponents grow with θ except λ_3 which roughly stays around 0. **B.** Prediction error of linear model also plotted as a function of θ (top panel). The minimum of the prediction error occurs at a value of θ between 10 to 15. These are also the values of θ at which λ_3 is closest to 0 (middle panel), and at which the spectrum is most symmetric, as measured by the sum of the conjugate exponents (bottom panel). **C.** The convergence of the Lyapunov exponents shown over for $\theta = 13$. The Lyapunov exponents converge to a steady value over the course of a 30 minute recording.

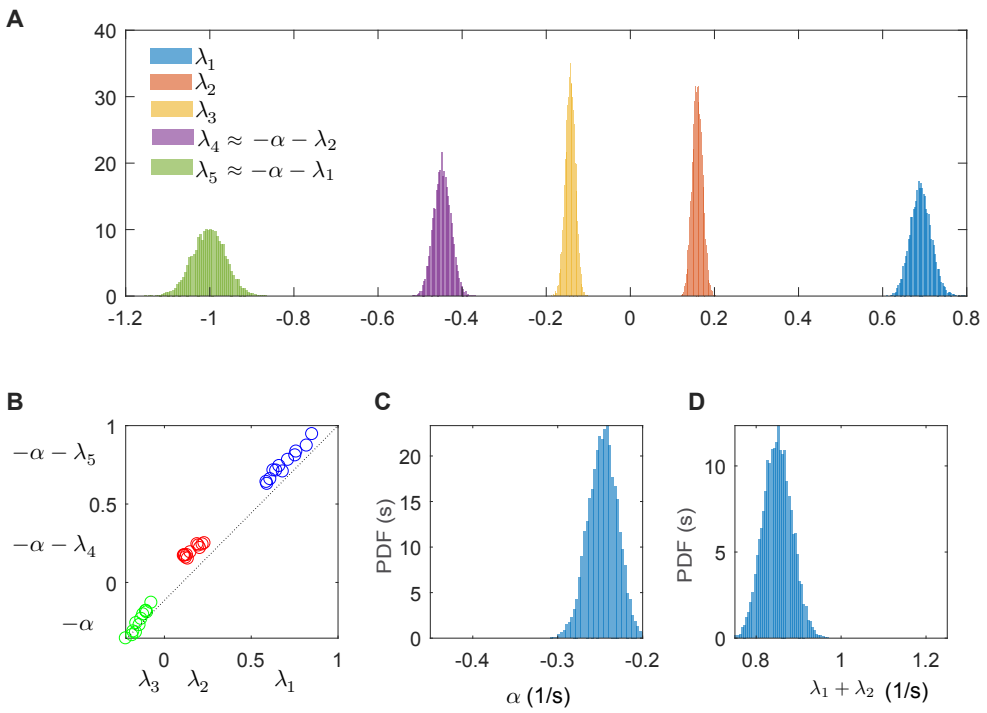


Figure 5.9: Lyapunov spectrum for foraging worms in a 5D embedding. **A.** Bootstrapped probability distribution across all $N = 12$ worms. The key difference here is the loss of the zero exponent. **B.** Conjugate exponents plotted against each other show that spectrum is less symmetric compared to 6D. **C.** Distribution of the dissipation constant α. **D.** Distribution of the sum of positive exponents estimating the Kolmogorov-Sinai entropy $\hat{h}_\mu$.

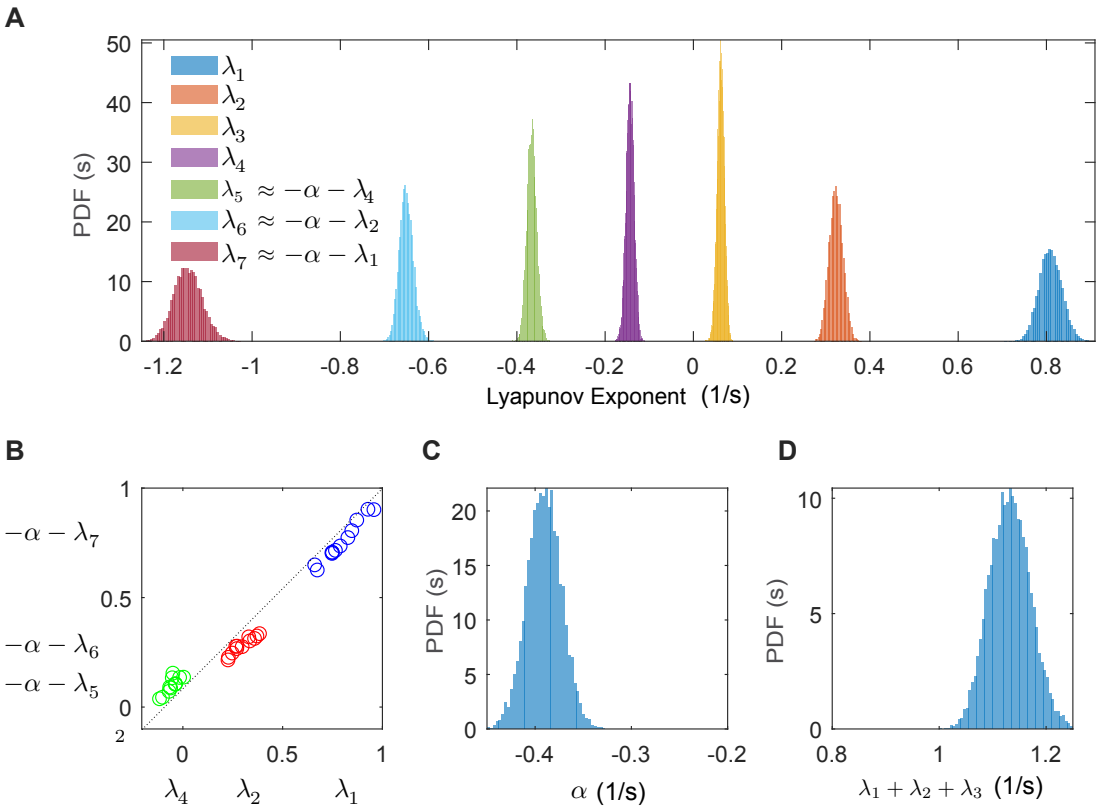


Figure 5.10: Lyapunov spectrum for foraging worms in a 7D embedding. **A.** Bootstrapped probability distribution across all $N = 12$ worms. Similar to 5D, the main difference is the loss of the zero exponent. **B.** Conjugate exponents plotted against each other again show that the 7D spectrum is less symmetric than the 6D spectrum. **C.** Distribution of the dissipation constant α shows that the dissipation is larger in 7D as compared to 6D. **D.** Distribution of the sum of positive exponents estimating the Kolmogorov-Sinai entropy $\hat{h}_\mu$ showing that $\hat{h}_\mu$ is larger in 7D.

Conclusion

In this thesis I have made an attempt to build a framework to capture the nonlinear dynamics of animal behavior in a model independent manner, and in as much details as possible. At the most basic level, the construction of this framework required non-trivial contributions to the field of nonlinear time-series analysis, in the form of robust estimation of embedding parameters, and local Jacobians.

Using the framework developed in this thesis. The first contribution is a new behavioral analysis pipeline which allows one to go from videos of freely moving worms to a phase space of their body wave dynamics. The phase space maps almost all the information about *C. elegans* locomotion onto highly interpretable geometrical representation. In particular, it organizes *C. elegans* behavior as a hierarchical set of cycles or periodic orbits. Globally, on large scales, we see three sets of periodic orbits, corresponding the three broadest classes of *C. elegans* behavior, forward crawling, backward crawling and turning. However, within each of these classes of stereotyped behaviors there is tremendous variability. The variability is such that it isn't possible to predict worm behavior too far in the future.

The next contribution is an unprecedented characterization of the observed behavioral variability. I use the local Jacobian estimates to estimate the spectrum of local and global Lyapunov exponents which measure the rates of expansion and contraction along different directions in state space. Along with positive exponents, which signify chaotic dynamics, I find surprisingly, that the spectrum of Lyapunov exponents is symmetric. In particular, the 6 exponents come in 3 pairs that sum to a constant negative value. This structure is a hallmark of damped driven Hamiltonian dynamics suggesting that worm locomotion can be decomposed into coupled damped driven Hamiltonian oscillators.

The final contribution is the observation that the damped driven Hamiltonian structure can be exploited to extract information about the neuromuscular control signals by estimating the sum of local Lyapunov exponents. Remarkably, this is possible from behavioral observations alone, without measuring any neural activity! As a preliminary test of this idea, I show that the sum of local Lyapunov exponents is oscillatory with a frequency which is double the frequency of the body wave. This observation is consistent with the hypothesis that the equations governing worm locomotion have a feedback that maintains the average mode energies constant. Next, I study the difference between spontaneous and stimulus driven transitions between forward and backward motion in the worm. Remarkably, the sum of local Lyapunov exponents show a clear distinction between spontaneous and stimulus driven reversals. This suggests that spontaneous reversals can be executed without a strong neuromuscular control

input, while stimulus induced reversals occur by means of a neuromuscular input.

The significance of these results comes from the fact that they characterize the complex, nonlinear dynamics of animal behavior in unprecedented detail. This detailed characterization of behavioral dynamics will be crucial in interpreting large scale neural recordings that are starting to become feasible [87, 92, 93]. The continuous nature of these recordings means that discrete behavioral representations will not be able to capture all the information in the neural signals. On the other hand an interpretable continuous representation can be directly correlated with the neural signals. Genetic perturbations will benefit from this characterization as its possible to measure subtle properties of the dynamics, such as unpredictability as measured by local and KS entropy, or aspects of control input as measured by local Lyapunov exponents.

Finally, although, most of the tools are applied to *C. elegans*, they are not specific to it. The framework and tools developed in this thesis are broadly applicable.

Bibliography

[1] John W Krakauer, Asif A Ghazanfar, Alex Gomez-Marin, Malcolm A MacIver, and David Poeppel. Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, 93:480–490, February 2017. ISSN 1097-4199. doi: 10.1016/j.neuron.2016.12.041.

[2] André EX Brown and Benjamin de Bivort. Ethology as a physical science. *Nature Physics*, page 1, 2018.

[3] Gordon J Berman. Measuring behavior across scales. *BMC biology*, 16:23, February 2018. ISSN 1741-7007. doi: 10.1186/s12915-018-0494-7.

[4] Sriram Ramaswamy. Active matter. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(5):054002, 2017.

[5] Jeffrey Aguilar, Tingnan Zhang, Feifei Qian, Mark Kingsbury, Benjamin McInroe, Nicole Mazouchova, Chen Li, Ryan Maladen, Chaohui Gong, Matt Travers, Ross L Hatton, Howie Choset, Paul B Umbanhowar, and Daniel I Goldman. A review on locomotion robophysics: the study of movement at the intersection of robotics, soft matter and dynamical systems. *Reports on progress in physics. Physical Society (Great Britain)*, 79:110001, November 2016. ISSN 1361-6633. doi: 10.1088/0034-4885/79/11/110001.

[6] Alex Gomez-Marin, Joseph J Paton, Adam R Kampff, Rui M Costa, and Zachary F Mainen. Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nature neuroscience*, 17:1455–1462, November 2014. ISSN 1546-1726. doi: 10.1038/nn.3812.

[7] Rainer Hegger, Holger Kantz, and Thomas Schreiber. Practical implementation of nonlinear time series methods: The tisean package. *Chaos (Woodbury, N.Y.)*, 9:413–435, June 1999. ISSN 1089-7682. doi: 10.1063/1.166424.

[8] E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of Atmospheric Sciences*, 20:130–148, March 1963. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

[9] Jacob Palis and Welington de Melo. *Geometric Theory of Dynamical Systems*. Springer US, 1982. doi: 10.1007/978-1-4612-5703-5. URL `https://doi.org/10.1007%2F978-1-4612-5703-5`.

[10] Floris Takens. Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence*, 898:366–381, 1981. doi: 10.1007/bfb0091924. URL `http://dx.doi.org/10.1007/bfb0091924`.

[11] Norman H Packard, James P Crutchfield, J Doyne Farmer, and Robert S Shaw. Geometry from a time series. *Physical review letters*, 45(9):712, 1980.

[12] George Udny Yule. Vii. on a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226(636-646):267–298, 1927.

[13] Martin Casdagli, Tim Sauer, and James A Yorke. Embedology. *J. Stat. Phys*, 65 (3/4):579–616, 1991.

[14] J Stark, David S Broomhead, ME Davies, and J Huke. Delay embeddings for forced systems. ii. stochastic forcing. *Journal of Nonlinear Science*, 13(6):519–577, 2003.

[15] M. R. Muldoon, D. S. Broomhead, J. P. Huke, and R. Hegger. Delay embedding in the presence of dynamical noise. *Dynamics and Stability of Systems*, 13, 1998. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.56.8139`.

[16] Jaroslav Stark. Delay embeddings for forced systems. i. deterministic forcing. *Journal of Nonlinear Science*, 9(3):255–332, 1999.

[17] Jeremy P Huke and David S Broomhead. Embedding theorems for non-uniformly sampled dynamical systems. *Nonlinearity*, 20(9):2205, 2007.

[18] Ethan R Deyle and George Sugihara. Generalized theorems for nonlinear state space reconstruction. *PLoS One*, 6(3):e18295, 2011.

[19] John F Gibson, J Doyne Farmer, Martin Casdagli, and Stephen Eubank. An analytic approach to practical state space reconstruction. *Physica. D, Nonlinear phenomena*, 57(1-2):1–30, 1992.

[20] Martin Casdagli. Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, 35(3):335–356, 1989.

[21] Andrew M Fraser and Harry L Swinney. Independent coordinates for strange attractors from mutual information. *Physical review A*, 33(2):1134, 1986.

[22] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6):3403–3411, March 1992. doi: 10.1103/physreva.45.3403. URL `http://dx.doi.org/10.1103/physreva.45.3403`.

[23] D. S. Broomhead and Gregory P. King. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2-3):217–236, June 1986. ISSN 01672789. doi: 10.1016/0167-2789(86)90031-x. URL `http://dx.doi.org/10.1016/0167-2789(86)90031-x`.

[24] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp. Singular-value decomposition and the grassberger-procaccia algorithm. *Physical Review A*, 38(6):3017–3026, sep 1988. doi: 10.1103/physreva.38.3017. URL `https://doi.org/10.1103%2Fphysreva.38.3017`.

[25] J. M. Martinerie, A. M. Albano, A. I. Mees, and P. E. Rapp. Mutual information, strange attractors, and the optimal estimation of dimension. *Physical Review A*, 45(10):7058–7064, may 1992. doi: 10.1103/physreva.45.7058. URL `https://doi.org/10.1103%2Fphysreva.45.7058`.

[26] AI Mees, PE Rapp, and LS Jennings. Singular-value decomposition and embedding dimension. *Physical Review A*, 36(1):340, 1987.

[27] Satohiro Tajima, Toru Yanagawa, Naotaka Fujii, and Taro Toyoizumi. Untangling brain-wide dynamics in consciousness by cross-embedding. *PLOS Computational Biology*, 11(11):e1004537, nov 2015. doi: 10.1371/journal.pcbi.1004537. URL `https://doi.org/10.1371%2Fjournal.pcbi.1004537`.

[28] Robert Vautard and Michael Ghil. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenomena*, 35(3):395–424, 1989.

[29] Keinosuke Fukunaga. *Introduction to statistical pattern recognition*. Elsevier, 2013.

[30] K Karhunen. Zur spektraltheorie stochastischer prozesse. suomalaisen tiedeakatemian toimituksia: Ser. a: 1, mathematica. *Physica*, 34, 1946.

[31] M Loeve. Fonctions aleatoires du second ordre, supplement to p. *Lévy, Processus Stochastiques et Mouvement Brownien, Gauthier-Villars, Paris*, 1948.

[32] Michael Ghil, MR Allen, MD Dettinger, K Ide, D Kondrashov, ME Mann, Andrew W Robertson, A Saunders, Y Tian, F Varadi, et al. Advanced spectral methods for climatic time series. *Reviews of geophysics*, 40(1):3–1, 2002.

[33] James B Elsner and Anastasios A Tsonis. *Singular spectrum analysis: a new tool in time series analysis*. Springer Science & Business Media, 2013.

[34] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13:411–430, 2000. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.79.7003`.

[35] Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse principal component analysis. *Journal of computational and graphical statistics*, 15(2):265–286, 2006.

[36] Daniel D Lee and H Sebastian Seung. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562, 2001.

[37] Adam Coates, Andrew Ng, and Honglak Lee. An analysis of single-layer networks in unsupervised feature learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 215–223, 2011.

[38] Martin Casdagli, Stephen Eubank, J. Doyne Farmer, and John Gibson. A theory of state space reconstruction in the presence of noise. In *Information Dynamics*, pages 61–96. Springer US, 1991. doi: 10.1007/978-1-4899-2305-9_5. URL https://doi.org/10.1007%2F978-1-4899-2305-9_5.

[39] Kevin Judd and Alistair Mees. Embedding as a modeling problem. *Physica D: Nonlinear Phenomena*, 120(3-4):273–286, sep 1998. doi: 10.1016/s0167-2789(98)00089-x. URL https://doi.org/10.1016%2Fs0167-2789%2898%2900089-x.

[40] Liangyue Cao, Alistair Mees, and Kevin Judd. Dynamics from multivariate time series. *Physica D: Nonlinear Phenomena*, 121(1-2):75–88, oct 1998. doi: 10.1016/s0167-2789(98)00151-1. URL https://doi.org/10.1016%2Fs0167-2789%2898%2900151-1.

[41] Edward N Lorenz. Predictability: A problem partly solved. In *Proc. Seminar on predictability*, volume 1, 1996.

[42] Martin Casdagli, Stephen Eubank, J Doyne Farmer, and John Gibson. State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, 51(1-3): 52–98, 1991.

[43] E Aurell, G Boffetta, A Crisanti, G Paladin, and A Vulpiani. Predictability in the large: an extension of the concept of lyapunov exponent. *Journal of Physics A: Mathematical and General*, 30(1):1–26, jan 1997. doi: 10.1088/0305-4470/30/1/003. URL https://doi.org/10.1088%2F0305-4470%2F30%2F1%2F003.

[44] J Doyne Farmer and John J Sidorowich. Predicting chaotic time series. *Physical review letters*, 59(8):845, 1987.

[45] J. B. Gao, J. Hu, W. W. Tung, and Y. H. Cao. Distinguishing chaos from noise by scale-dependent lyapunov exponent. *Physical Review E*, 74(6), dec 2006. doi: 10.1103/physreve.74.066204. URL https://doi.org/10.1103%2Fphysreve.74.066204.

[46] J. P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617–656, jul 1985. doi: 10.1103/revmodphys.57.617. URL https://doi.org/10.1103%2Frevmodphys.57.617.

[47] Reggie Brown, Paul Bryant, and Henry D. I. Abarbanel. Computing the lyapunov spectrum of a dynamical system from an observed time series. *Phys. Rev. A*, 43 (6):2787–2806, Mar 1991. doi: 10.1103/PhysRevA.43.2787. URL https://link.aps.org/doi/10.1103/PhysRevA.43.2787.

[48] A.G. Darbyshire and D.S. Broomhead. Robust estimation of tangent maps and liapunov spectra. *Physica D: Nonlinear Phenomena*, 89(3-4):287–305, jan 1996. doi: 10.1016/0167-2789(95)00246-4. URL https://doi.org/10.1016%2F0167-2789%2895%2900246-4.

[49] H. D. I. Abarbanel, R. Brown, and M. B. Kennel. Local lyapunov exponents computed from observed data. *Journal of Nonlinear Science*, 2(3):343–365, sep 1992. doi: 10.1007/bf01208929. URL `https://doi.org/10.1007%2Fbf01208929`.

[50] Valery Iustinovich Oseledec. A multiplicative ergodic theorem. liapunov characteristic number for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.

[51] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[52] George David Birkhoff. What is the ergodic theorem? *The American Mathematical Monthly*, 49(4):222–226, 1942.

[53] David Ruelle. Ergodic theory of differentiable dynamical systems. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 50(1):27–58, 1979.

[54] J Doyne Farmer. Information dimension and the probabilistic structure of chaos. *Zeitschrift für Naturforschung A*, 37(11):1304–1326, 1982.

[55] Andrei Nikolaevitch Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. In *Dokl. Akad. Nauk SSSR*, volume 124, pages 754–755, 1959.

[56] Yakov B Pesin. *Dimension theory in dynamical systems: contemporary views and applications*. University of Chicago Press, 2008.

[57] Lai-Sang Young. Dimension, entropy and lyapunov exponents. *Ergodic theory and dynamical systems*, 2(1):109–124, 1982.

[58] F Ledrappier and L-S Young. Stability of lyapunov exponents. *Ergodic Theory and Dynamical Systems*, 11(3):469–484, 1991.

[59] Arkady Pikovsky and Antonio Politi. *Lyapunov exponents: a tool to explore complex dynamics*. Cambridge University Press, 2016.

[60] H Haken. At least one lyapunov exponent vanishes if the trajectory of an attractor does not contain a fixed point. *Physics Letters A*, 94(2):71–72, 1983.

[61] Ute Dressler. Symmetry property of the lyapunov spectra of a class of dissipative dynamical systems with viscous damping. *Physical Review A*, 38(4):2103, 1988.

[62] D Gupalo, AS Kaganovich, and EGD Cohen. Symmetry of lyapunov spectrum. *Journal of statistical physics*, 74(5-6):1145–1159, 1994.

[63] H. D. I. Abarbanel, R. Brown, and M. B. Kennel. Variation of Lyapunov exponents on a strange attractor. *Journal of Nonlinear Science*, 1(2):175–199, June 1991. doi: 10.1007/bf01209065. URL `http://dx.doi.org/10.1007/bf01209065`.

[64] Mozheng Wei. Quantifying local instability and predictability of chaotic dynamical systems by means of local metric entropy. *International Journal of Bifurcation and Chaos*, 10(01):135–154, 2000.

[65] Hanqi Guo, Wenbin He, Tom Peterka, Han-Wei Shen, Scott M Collis, and Jonathan J Helmus. Finite-time lyapunov exponents and lagrangian coherent structures in uncertain unsteady flows. *IEEE transactions on visualization and computer graphics*, 22(6):1672–1682, 2016.

[66] Edward Ott. *Chaos in dynamical systems*. Cambridge university press, 2002.

[67] Russell A Johnson, Kenneth J Palmer, and George R Sell. Ergodic properties of linear dynamical systems. *SIAM journal on mathematical analysis*, 18(1):1–33, 1987.

[68] Julien Clinton Sprott. *Elegant Chaos*. World Scientific Publishing Co. Pte. Ltd., 2010. doi: 10.1142/9789812838827. URL `https://doi.org/10.1142%2F9789812838827`.

[69] George Sugihara and Robert M May. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344(6268):734, 1990.

[70] Greg J. Stephens, Matthew Bueno de Mesquita, William S. Ryu, and William Bialek. Emergence of long timescales and stereotyped behaviors in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7286–7289, May 2011. ISSN 1091-6490. doi: 10.1073/pnas.1007868108. URL `http://dx.doi.org/10.1073/pnas.1007868108`.

[71] Onno D Broekmans, Jarlath B Rodgers, William S Ryu, and Greg J Stephens. Resolving coiled shapes reveals new reorientation behaviors in c. elegans. *eLife*, 5: e17227, 2016.

[72] J. E. Sulston and S. Brenner. The DNA of C. elegans. *Genetics*, 77:95–104, 1974. URL `http://view.ncbi.nlm.nih.gov/pubmed/4858229`.

[73] Aylia Mohammadi, Jarlath Byrne Rodgers, Ippei Kotera, and William S Ryu. Behavioral response of caenorhabditis elegans to localized thermal stimuli. *BMC neuroscience*, 14(1):66, 2013.

[74] David A Gagnon and Paulo E Arratia. The cost of swimming in generalized newtonian fluids: experiments with c. elegans. *Journal of Fluid Mechanics*, 800: 753–765, 2016.

[75] N A Croll. Behavioural analysis of nematode movement. *Advances in parasitology*, 13:71–122, 1975. ISSN 0065-308X.

[76] Neil A Croll. Components and patterns in the behaviour of the nematode caenorhabditis elegans. *Journal of zoology*, 176(2):159–176, 1975.

[77] Greg J. Stephens, Bethany Johnson-Kerner, William Bialek, and William S. Ryu. Dimensionality and Dynamics in the Behavior of C. elegans. *PLoS Computational Biology*, 4(4):e1000028+, April 2008. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000028. URL `http://dx.doi.org/10.1371/journal.pcbi.1000028`.

[78] Jennifer K Pirri, Adam D McPherson, Jamie L Donnelly, Michael M Francis, and Mark J Alkema. A tyramine-gated chloride channel coordinates distinct motor programs of a caenorhabditis elegans escape response. *Neuron*, 62(4):526–538, 2009.

[79] Jesse M Gray, Joseph J Hill, and Cornelia I Bargmann. A circuit for navigation in caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, 102:3184–3191, March 2005. ISSN 0027-8424. doi: 10.1073/pnas.0409009101.

[80] J T Pierce-Shimomura, T M Morse, and S R Lockery. The fundamental role of pirouettes in caenorhabditis elegans chemotaxis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19:9557–9569, November 1999. ISSN 1529-2401.

[81] Yi Zheng, Penelope J Brockie, Jerry E Mellem, David M Madsen, and Andres V Maricq. Neuronal control of locomotion in c. elegans is modified by a dominant mutation in the glr-1 ionotropic glutamate receptor. *Neuron*, 24(2):347–361, 1999.

[82] Taizo Kawano, Michelle D Po, Shangbang Gao, George Leung, William S Ryu, and Mei Zhen. An imbalancing act: gap junctions reduce the backward motor circuit activity to bias c. elegans for forward locomotion. *Neuron*, 72(4):572–586, 2011.

[83] William M Roberts, Steven B Augustine, Kristy J Lawton, Theodore H Lindsay, Tod R Thiele, Eduardo J Izquierdo, Serge Faumont, Rebecca A Lindsay, Matthew Cale Britton, Navin Pokala, Cornelia I Bargmann, and Shawn R Lockery. A stochastic neuronal model predicts random search behaviors at multiple spatial scales in *C. elegans*. *eLife*, 5:e12572, jan 2016. ISSN 2050-084X. doi: 10.7554/eLife.12572. URL `https://dx.doi.org/10.7554/eLife.12572`.

[84] William S Ryu and Aravinthan DT Samuel. Thermotaxis in caenorhabditis elegans analyzed by measuring responses to defined thermal stimuli. *Journal of Neuroscience*, 22(13):5727–5733, 2002.

[85] Penelope J Brockie, Jerry E Mellem, Thomas Hills, David M Madsen, and Andres V Maricq. The c. elegans glutamate receptor subunit nmr-1 is required for slow nmda-activated currents that regulate reversal frequency during locomotion. *Neuron*, 31(4):617–630, 2001.

[86] Thomas Hills, Penelope J Brockie, and Andres V Maricq. Dopamine and glutamate control area-restricted search behavior in caenorhabditis elegans. *Journal of Neuroscience*, 24(5):1217–1225, 2004.

[87] Saul Kato, Harris S. Kaplan, Tina Schrödel, Susanne Skora, Theodore H. Lindsay, Eviatar Yemini, Shawn Lockery, and Manuel Zimmer. Global Brain Dynamics Embed the Motor Command Sequence of Caenorhabditis elegans. *Cell*, 163(3): 656–669, October 2015. ISSN 00928674. doi: 10.1016/j.cell.2015.09.034. URL `http://dx.doi.org/10.1016/j.cell.2015.09.034`.

[88] Monika Scholz, Ashley N Linder, Francesco Randi, Anuj K Sharma, Xinwei Yu, Joshua W Shaevitz, and Andrew Leifer. Predicting natural behavior from whole-brain neural dynamics. *bioRxiv*, 2018. doi: 10.1101/445643. URL `https://www.biorxiv.org/content/early/2018/10/17/445643`.

[89] Eviatar Yemini, Tadas Jucikas, Laura J Grundy, André E X Brown, and William R Schafer. A database of caenorhabditis elegans behavioral phenotypes. *Nature methods*, 10:877–879, September 2013. ISSN 1548-7105. doi: 10.1038/nmeth.2560.

[90] Roland F Schwarz, Robyn Branicky, Laura J Grundy, William R Schafer, and André E X Brown. Changes in postural syntax characterize sensory modulation and natural variation of c. elegans locomotion. *PLoS computational biology*, 11: e1004322, August 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004322.

[91] Denis J Evans and WG Hoover. Flows far from equilibrium via molecular dynamics. *Annual review of fluid mechanics*, 18(1):243–264, 1986.

[92] Jeffrey P. Nguyen, Frederick B. Shipley, Ashley N. Linder, George S. Plummer, Mochi Liu, Sagar U. Setru, Joshua W. Shaevitz, and Andrew M. Leifer. Whole-brain calcium imaging with cellular resolution in freely behaving caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, 113(8):E1074–E1081, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1507110112. URL `https://www.pnas.org/content/113/8/E1074`.

[93] Vivek Venkatachalam, Ni Ji, Xian Wang, Christopher Clark, James Kameron Mitchell, Mason Klein, Christopher J. Tabone, Jeremy Florman, Hongfei Ji, Joel Greenwood, Andrew D. Chisholm, Jagan Srinivasan, Mark Alkema, Mei Zhen, and Aravinthan D. T. Samuel. Pan-neuronal imaging in roaming caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, 113(8):E1082–E1088, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1507109113. URL `https://www.pnas.org/content/113/8/E1082`.