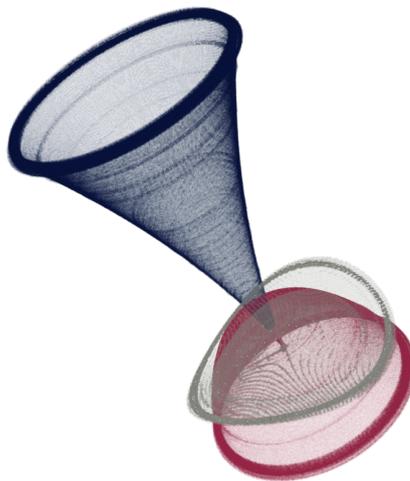




UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

THE COMPUTATION AND ANALYSIS OF
INVARIANT SETS OF
INFINITE-DIMENSIONAL SYSTEMS



Von der Fakultät für
Elektrotechnik, Informatik und Mathematik
der Universität Paderborn
zur Erlangung des akademischen Grades

DOKTOR DER NATURWISSENSCHAFTEN
– DR. RER. NAT. –

genehmigte Dissertation

von

Raphael Gerlach

Paderborn 2021

Gutachter: Prof. Dr. Michael Dellnitz
Prof. Dr. Péter Koltai

Tag der mündlichen Prüfung: 03.12.2021

Acknowledgments

I am grateful for the constant support of all the people involved that made this dissertation possible. Thank you for all given opportunities, ideas and encouragement.

First of all, I would like to thank my advisor *Prof. Dr. Michael Dellnitz* for his guidance, support and ideas, for sharing his enthusiasm for dynamical systems with me, and for the freedom he has given me during my research.

I am grateful to *Prof. Dr. Péter Koltai* for all the interesting talks we had together. Thank you for always asking the right (research) questions and also helping me find the answers to those.

I acknowledge the support I have received within the *DFG Priority Programme 1881 – Turbulent Superstructures*. Thank you for introducing me to the field of fluid dynamics. Special thanks go to *Prof. Dr. Bruno Eckhardt, Prof. Dr. Kathrin Padberg-Gehle, Prof. Dr. Jörg Schumacher, Dr. Daniel Feldman, Dr. Daniel Karrasch, Dr. Moritz Linkman, Dr. Christiane Schneide, Alvaro de Diego, Anna Klünker, Nathanael Schilling* and *Martin Stahn*.

Many thanks to my past and present colleagues at Paderborn University: *Jun.-Prof. Dr. Sebastian Peitz, Dr. Mirko Hessel-von Molo, Dr. Karin Mora, Dr. Feliks Nüske, Dr. Adrian Ziessler, Manuel Berkemeier, Katharina Bieker* and *Bennet Gebken*. Working with you was and still is a pleasure. Thank you for creating a very pleasant and enjoyable atmosphere.

I would like to express my gratitude to my family for caring for me in every situation. In particular, I praise *Miriam Sophia Czichy* for her persistent encouragement, for showing me the world, and for her everlasting love. I am looking forward to spending my life with her and our daughter *Maivi Leona Sophie*.

*Like what you do,
and then you will do your best.*

KATHERINE JOHNSON

Abstract

One central aspect in the analysis of dynamical systems is the characterization of their long-term behavior. Thus, we are interested in accurately describing the so-called *global attractor* of the dynamical system, which is an invariant set that attracts all the trajectories. By definition, the global attractor contains every steady state, every periodic orbit and, in particular, their related unstable manifolds. Moreover, when the underlying system depends on a parameter it is crucial to analyze the change of the attractor with respect to this parameter. For the study of invariant sets such as the attractor or an unstable manifold we rely on numerical algorithms for their approximation in general. However, typically those tools can only be applied to finite-dimensional dynamical systems, described for instance by ordinary differential equations or discrete maps.

In the first part of this thesis we present a framework for the global dynamical analysis of infinite-dimensional systems. We will utilize embedding techniques for the definition of a dynamically equivalent finite-dimensional system, the so-called *core dynamical system* (CDS). This system is then used for the approximation of corresponding embedded invariant sets, i.e., one-to-one images of invariant sets in the infinite-dimensional state space. Here, we focus on set-oriented numerical tools that generate coverings of the set of interest and adapt them to the CDS. In particular, we extend the subdivision scheme to parameter-dependent systems which allows us to efficiently track the corresponding parameter-dependent attractor. To this end, we will numerically realize a set-valued linearization that serves as an initial guess for a consecutive corrector step.

For the construction of the CDS it is crucial to choose an appropriate *observation map*. Moreover, its corresponding inverse is in general not known and has to be numerically designed. Hence, in the second part of this thesis we will present suitable numerical realizations of the CDS for delay differential equations and partial differential equations for the approximation of embedded invariant sets via set-oriented algorithms. For a subsequent geometric analysis we will consider a learning technique called *diffusion maps*. This nonlinear dimensionality reduction method applied to the generated coverings reveals its intrinsic geometry. In this context we present a set-oriented *landmark* selection scheme, that computes points that sample the generated covering sufficiently well, and an intrinsic dimension estimator.

Finally, in the last part of this thesis, we apply the developed numerical tools for the global dynamical analysis of some well-known (infinite-dimensional) dynamical systems. We will compute the global attractor of a reduced order model for shear flows and of the Mackey-Glass delay differential equation for a range of delay times. Furthermore, we approximate the unstable manifold of the one-dimensional Kuramoto-Sivashinsky equation as well as the unstable manifold of an *edge state* in a plane Poiseuille flow.

Zusammenfassung

Ein zentraler Aspekt in der Analyse von dynamischen Systemen ist die Charakterisierung ihres Langzeitverhaltens. Wir sind daher daran interessiert, den sogenannten *globalen Attraktor* des dynamischen Systems, welcher eine invariante Menge ist, die alle Trajektorien anzieht, genau zu beschreiben. Per Definition enthält der globale Attraktor jeden stationären Zustand, jeden periodischen Orbit und insbesondere die damit verbundenen instabilen Mannigfaltigkeiten. Wenn das zugrundeliegende System von einem Parameter abhängt, ist es außerdem wichtig, die Änderung des Attraktors in Bezug auf diesen Parameter zu analysieren. Für die Untersuchung invarianter Mengen wie des Attraktors oder einer instabilen Mannigfaltigkeit stützen wir uns im Allgemeinen auf numerische Algorithmen, um sie zu approximieren. Typischerweise können diese Methoden jedoch nur auf endlichdimensionale dynamische Systeme angewendet werden, die zum Beispiel durch gewöhnliche Differentialgleichungen oder diskrete Abbildungen beschrieben werden.

Im ersten Teil dieser Arbeit präsentieren wir ein Werkzeug für die globale dynamische Analyse unendlichdimensionaler Systeme. Wir werden Einbettungstechniken zur Definition des sogenannten *core dynamical system* (CDS) verwenden, welches ein dynamisch äquivalentes, aber endlichdimensionales System ist. Dieses System wird dann zur Approximation entsprechender eingebetteter invarianter Mengen, also Eins-zu-eins-Bildern von invarianten Mengen im unendlichdimensionalen Zustandsraum, benutzt. Hier konzentrieren wir uns auf mengenorientierte numerische Methoden, die Überdeckungen der invarianten Menge erzeugen, und passen diese an das CDS an. Insbesondere erweitern wir den Unterteilungsalgorithmus auf parameterabhängige Systeme, mit dem wir den entsprechenden parameterabhängigen Attraktor effizient verfolgen können. Zu diesem Zweck werden wir numerisch eine mengenwertige Linearisierung realisieren, die als erste Schätzung für einen danach folgenden Korrekturschritt dient.

Für die Konstruktion des CDS ist es wichtig, eine geeignete *Beobachtung* zu wählen. Darüber hinaus ist deren entsprechende Umkehrung im Allgemeinen nicht bekannt und muss numerisch entworfen werden. Daher werden im zweiten Teil der Arbeit geeignete numerische Realisierungen des CDS für retardierte und für partielle Differentialgleichungen für die Approximation eingebetteter invarianter Mengen mittels mengenorientierter Algorithmen vorgestellt. Für eine nachfolgende geometrische Analyse betrachten wir eine Lerntechnik namens *diffusion maps*. Diese nichtlineare Dimensionsreduktionsmethode, die auf die erzeugten Überdeckungen angewendet wird, offenbart ihre intrinsische Geometrie. In diesem Zusammenhang präsentieren wir ein mengenorientiertes Auswahlschema für *landmarks*, das Punkte berechnet, die die erzeugte Überdeckung ausreichend gut abtasten, sowie einen intrinsischen Dimensionsschätzer.

Schließlich nutzen wir im letzten Teil dieser Arbeit die entwickelten numerischen Methoden für die globale dynamische Analyse einiger bekannter (unendlichdimensionaler) dynamischer Systeme. Wir werden den globalen Attraktor eines reduzierten Modells für Scherströmung sowie der Mackey-Glass-Gleichung für einen Bereich von Verzögerungszeiten berechnen. Darüber hinaus approximieren wir die instabile Mannigfaltigkeit der eindimensionalen Kuramoto-Sivashinsky-Gleichung sowie die instabile Mannigfaltigkeit eines *edge states* in der ebenen Poiseuille Strömung.

Contents

1	Introduction	1
2	Embedding Theory	9
2.1	Inertial Manifolds	10
2.2	Embedding Finite-dimensional Sets into \mathbb{R}^k	12
2.2.1	Finite-dimensional Manifolds	12
2.2.2	Fractal Sets in \mathbb{R}^N	13
2.2.3	Infinite-dimensional State Spaces	18
2.3	Delay Coordinate Embeddings	28
2.4	The Core Dynamical System (CDS)	30
3	Set-Oriented Numerics for Infinite-Dimensional Systems	35
3.1	The Subdivision Scheme	36
3.1.1	Approximation of Attracting Sets	41
3.2	A Path Following Method	43
3.2.1	A Set-Oriented Predictor-Corrector Method	49
3.3	The Continuation Method	54
4	The Choice of the Observation Map	61
4.1	Linear Operators	62
4.1.1	Point Evaluation for DDEs	62
4.1.2	Coordinate Representation for PDEs	65
4.2	Diffusion Maps	69
4.2.1	An Intrinsic Dimension Estimator	75
4.2.2	A Set-Oriented Landmark Selection Scheme	78
5	Applications	83
5.1	A Four-Dimensional Model of Self-Sustained Flows	83
5.2	The Mackey-Glass Equation	86
5.3	The Kuramoto-Sivashinsky Equation	89
5.3.1	The Traveling Wave	90
5.3.2	The Stable Heteroclinic Cycle	94
5.3.3	Further Intrinsic Geometric Analysis	96
5.4	Plane Poiseuille Flow	98
6	Conclusion and Outlook	101
6.1	Embedding Techniques	101
6.1.1	The Core Dynamical System (CDS)	102
6.2	The Choice of the Observation Map	102
6.2.1	Diffusion Maps	103
	Bibliography	105

1 Introduction

Behind every process that changes in time hides a *dynamical system*. May it be a simple problem such as a pendulum and a mass-spring-damper or even the complex motion of the bodies in the solar system. In addition to processes in solid mechanics, dynamical systems also have their "raison d'être" in other scientific areas such as physics (e.g. climate models [MB20]), biology [Fre80], epidemiology [BCC12], chemistry (e.g. molecular dynamics [PWS⁺11]), economics [Gan71] and medicine [JR15]. They are also utilized in non-STEM fields such as history [Tur08] and social science [Con13]. In this thesis we will mostly consider problems from physics, especially from fluid dynamics.

Mathematically, a dynamical system is given by three components: the *time*, the *state space* and an *evolution law*. The time is either *discrete* (e.g. \mathbb{N} or \mathbb{Z}) or *continuous* (e.g. $\mathbb{R}_{\geq 0}$ or \mathbb{R}) and we note that one can always *discretize* the time of a time-continuous system which is especially helpful for the numerical analysis. Often the state space is a *finite-dimensional* Euclidean space such as \mathbb{R}^n (e.g. consisting of positions and velocity components in mechanical systems) but in this thesis we focus our attention on *infinite-dimensional* spaces. More precisely, the state space will be a space of functions on some domain, e.g., velocity fields of a fluid in a channel. The evolution is a map on the state space that models the underlying dynamical process. In continuous time the evolution is often given by a set of *differential equations*. Typically, one distinguishes between *ordinary differential equations* (ODEs), *partial differential equations* (PDEs) and *delay differential equations* (DDEs). Observe that the latter two induce an infinite-dimensional state space and hence will be considered in this thesis. For the sake of completeness we note that there are also mixed systems such as *hybrid* systems which consist of time-discrete and time-continuous subsystems and *algebraic-differential equations* where an algebraic constraint is added to the differential equations. In any case, under appropriate assumptions, the set of differential equations generates a *flow* that maps a given initial condition in state space onto its corresponding evolved state at some prescribed time. Hence, it generates the so-called *trajectory* of an initial state. A dynamical system where the temporal evolution does not depend on the initial time but only on the initial state is called *autonomous* and we restrict ourselves to the analysis of such systems. However, some results which will be presented in this thesis may be extended to the *non-autonomous* case. In particular, we will only consider *deterministic* systems and leave *stochastic* processes for future research. For an introduction in the theory of dynamical systems and also for further reading we refer the reader to [KH95, GH13].

The central questions in the context of dynamical systems primarily concern their long-term behavior (e.g. stability, chaos and ergodicity). Thus, typically the object of interest is the so-called *global attractor* of the dynamical system. The global attractor is an *invariant* set in state space that attracts every trajectory under the temporal evolution. That is, every state inside the global attractor stays in that set for all times and every initial condition outside of the attractor eventually converges to the global attractor. More precisely, the distance between the evolved states and the attractor gets arbitrarily close to zero. Thus, by definition, the global attractor contains every equilibrium, i.e., a steady state under the dynamics, and every periodic orbit. Both objects are crucial for the dynamical behavior of trajectories. Other important sets for the long-term analysis of dynamical system are the *invariant stable* and *unstable manifolds* of a steady state or

periodic orbit. These sets consist of all states that converge to the corresponding object in forward respectively backward time. In particular, the unstable manifold is also contained in the global attractor. In analogy to the eigenspace decomposition in a linear system such invariant manifolds generalize this concept and decompose the state space. For instance, they can be used for efficiently solving optimal control problems [FOBK12].

For a general nonlinear dynamical system the computation of the (un)stable manifolds or the global attractor by direct numerical simulation of (some) initial values is not feasible or even possible. That is why efficient numerical algorithms for the approximation of such invariant sets are required. For an overview of methods that allow the computation of invariant manifolds we refer to [KOD⁺05]. For instance, two dimensional manifolds can be approximated by geodesic level sets obtained by solving a boundary value problem [KO03, KO07]. There are also several approaches for the computation of attractors. In [JK17] a variational approach has been developed where the distance between a set of states and their image under the dynamics is minimized. Furthermore, a linear program can be formulated such that its feasible solutions are real-valued functions whose preimage contain the attractor and for the optimum this approximation gets tight [SK20]. Another class of algorithms that allow the computation of invariant sets is given by so-called *set-oriented* numerical methods [DH96, DH97, DFJ01, DJ02]. The basic idea of such tools is to generate coverings of the object of interest such as the global attractor or invariant manifolds by outer approximations which are created by *subdivision techniques*. Moreover, such methods allow the approximation of the *Perron-Frobenius* operator by *Ulam's method* [Ula60, BM01] which can be used for the computation of the *invariant measure* of the underlying dynamical system, i.e., the fixed point of the operator (see [DJ99, DFS00, KKS16]). In particular, eigenvalues close to one can be utilized for the identification of *almost invariant* sets, i.e., sets in which typical trajectories stay inside on average for a quite long time [FD03, Fro05]. We note that these methods also have been extended to non-autonomous systems for the approximation of *coherent sets* [FP09, FLS10, DH12] and the analysis of transport [DFHP09] as well as to random dynamical systems [KO99]. These set-oriented techniques have been successfully applied in several different areas such as molecular dynamics [DDJS99, SHD01, DGM⁺05], astrophysics [DJK⁺05, DJL⁺05, DJ06] and ocean dynamics [FPET07, DFH⁺09, SFM10, FHR⁺12]. They can also be utilized for uncertainty quantification from a global point of view [DKZ17]. In the software package *GAIO* (*Global Analysis of Invariant Objects*) [DFJ01] available on [GitHub](#) all set-oriented methods are implemented.

Until recently, the set-oriented algorithms have been restricted to dynamical systems with a finite-dimensional phase space. Even though an infinite-dimensional system can be numerically approximated by a (possibly very) high but finite-dimensional system using an appropriate discretization of each state, the classical set-oriented methods remain infeasible. In order to gain an idea of the dynamics of an infinite-dimensional system, one typically performs long-term simulations of arbitrary initial states. However, this is not sufficient for the analysis of *global* long-term behavior. To this end, a novel set-oriented framework for the computation of finite-dimensional invariant sets of infinite-dimensional systems has been developed [DHZ16, ZDG19]. Dynamical systems with an infinite-dimensional state space, but finite-dimensional attractors arise in particular in two areas of applied mathematics, namely in certain types of dissipative PDEs [CFT85, Ha10] and in DDEs with small constant delay [Dri68, Chi03, CMRV05]. Examples of such PDEs

include the Kuramoto-Sivashinsky equation (KSE) [KT76, Siv77, JKT90, Rob94], the Ginzburg-Landau equation [DGHN88] and scalar reaction-diffusion equations with cubic nonlinearity [Jol89].

Both types of systems may possess a so-called *inertial manifold*, which is a finite-dimensional invariant manifold that attracts all trajectories exponentially fast [Tem97, CFNT89, FJK⁺88]. In particular, the global attractor lies inside the inertial manifold. Furthermore, the inertial manifold allows to set up an appropriate finite-dimensional *reduced order model* (ROM), e.g., by Galerkin expansion - that can be used for the analysis of the long-term behavior instead. The dimension of such a ROM is then given by the dimension of the manifold, which can be much larger than the dimension of the attractor itself (see, e.g., [CEES93, Rob94, TW94, Tem97]). Moreover, the conditions known to be sufficient for proving the existence of an inertial manifold are very restrictive [FST88] and there are cases in which one can prove the existence of a finite-dimensional attractor but not (yet) of an inertial manifold. Arguably, the most interesting example that falls into this category might be the Navier-Stokes equation [Lad85, MN96, MW96, Rob13]. Thus, it is desirable to construct a ROM that has a dimension comparable to the dimension of the attractor and possesses the same dynamics as the underlying infinite-dimensional system on the attractor. To this end, we employ the idea of *observing* the system by real-valued functions that are one-to-one on the invariant set under appropriate assumptions. In other words, we reconstruct the underlying dynamics by *embedding* the invariant set into a finite-dimensional space.

As a first fundamental result in the theory of embeddings, Hassler Whitney proved that every finite-dimensional manifold \mathcal{M} can be embedded into some \mathbb{R}^k , provided $k \in \mathbb{N}$ is large enough [Whi36, Whi44]. In particular, such an embedding is injective and its image is one-to-one on the given manifold \mathcal{M} . For the needed *embedding dimension* $k \in \mathbb{N}$ it turns out that $k > 2d$ is sufficient, where d is the dimension of the manifold \mathcal{M} . Moreover, the set of such embeddings is *generic* (with respect to a particular topology), i.e., open and dense. Thus, given any (smooth) map $R : \mathcal{M} \rightarrow \mathbb{R}^k$ there is, in particular, an embedding arbitrary close to such a map R . Additionally, every arbitrary small perturbation of an embedding remains an embedding. In this context R could be given by k independent measurements, also called *observations*, of an experiment. Hence, we will call R *observation map* and the corresponding embedding space \mathbb{R}^k *observation space*. Later, inspired by Whitney's work, Floris Takens strengthened this result to the case where there is an underlying dynamical system defined on the finite-dimensional manifold [Tak81]. Instead of making k independent observations it is enough to choose one real-valued *observable* and observe the system at k delayed time instances, that is, k consecutive snapshots in time. More precisely, his theorem states that for *pairs* (Φ, f) of a dynamical system $\Phi : \mathcal{M} \rightarrow \mathcal{M}$ and an observable $f : \mathcal{M} \rightarrow \mathbb{R}$ it is a generic property that the *delay coordinate map* $D_k[f, \Phi] : \mathcal{M} \rightarrow \mathbb{R}^k$ given by

$$D_k[f, \Phi](x) = (f(x), f(\Phi(x)), \dots, f(\Phi^{k-1}(x)))$$

is an embedding. Note that in this form of the theorem the dynamical system is not fixed. In order to formulate an analogous result for a given (fixed) dynamical system Φ , additional assumptions on the underlying system concerning the number of periodic points and the eigenspectrum of the linearization at those periodic points have to be made (see,

e.g., [Huk06]). With his celebrated theorem Takens laid the theoretical foundation for the dynamical analysis of complex systems and experiments (see [BK86] and [BSS⁺83, BS87] for fluid flow problems) and the prediction of chaotic time series [FS87, Cas89]. However, a generic property is not immediately a property with high (Lebesgue) probability. In particular, from a numerical and experimentalist's point of view, it is desirable to know if a given observation or delay coordinate map is an embedding with probability one. The problem with such a statement is that the space of observation maps is an *infinite-dimensional* function space and there is no obvious generalization to such infinite-dimensional spaces from finite-dimensional spaces. In fact, for an arbitrary Banach space there is not even a measure that corresponds to the Lebesgue measure on finite-dimensional subspaces [HSY92]. Thus, in [HSY92] the measure-theoretic term *prevalence* is introduced which extends the condition for a property to hold "almost-everywhere" in a finite-dimensional space. This notion even turns out to be a more appropriate condition than the topological concepts of "open and dense" in the context of probabilistic results on the likelihood of a given property as desired. The mentioned embedding theorems are not directly applicable to infinite-dimensional systems such as DDEs and PDEs since for the reconstruction of its attractor it has to be contained in a finite-dimensional manifold which is in general not the case. In order to directly embed the attractor instead, two additional ingredients are needed. First, we replace the finite-dimensional manifold with a (fractal) set with *box-counting dimension* $d_{\text{box}} \in \mathbb{R}_{\geq 0}$. For the definition of this fractal dimension the set is covered by a minimal number N of balls with radius $\varepsilon > 0$ and then d_{box} describes the scaling law $N \sim \varepsilon^{-d_{\text{box}}}$ for $\varepsilon \rightarrow 0$. Utilizing this notion Takens' theorem was extended in [SYC91] to dynamical systems on \mathbb{R}^N where a compact finite-dimensional invariant set is embedded into \mathbb{R}^k . Additionally, in the case of an infinite-dimensional state space it is necessary to know how well the invariant set can be approximated by finite-dimensional subspaces. This information is given by the *thickness exponent* τ introduced in [HK99] and the authors showed that "almost every" (in the sense of prevalence) Lipschitz map $R : H \rightarrow \mathbb{R}^k$ is one-to-one on a prescribed compact d_{box} -dimensional set \mathcal{A} contained in a Hilbert space H , provided $k > 2d_{\text{box}}$. Moreover, its inverse on \mathcal{A} is Hölder continuous with exponent θ which depends on the embedding dimension k , the box-counting dimension d_{box} and the thickness exponent τ . In particular, this result extends Whitney's theorem to the infinite-dimensional setting. We note that this result can also be formulated for finite-dimensional spaces but without the use of the thickness exponent since in this case $\tau = 0$ [Rob10]. Finally, Robinson combined the work of [SYC91] and [HK99] and proved a delay embedding theorem for infinite-dimensional dynamical systems on a Hilbert space. Unfortunately, the analogous result where the ambient space is a general Banach space given in [HK99] is not true (see [Rob09]). In this case one has to consider the *dual thickness exponent* [Rob09] or assume that the thickness exponent is bounded by $\tau < 1$ [MR19]. However, many of the attractors arising in mathematical physics have a thickness exponent of zero [OHK06]. For instance, by [FR99] the thickness exponent is in some sense inverse proportional to smoothness and hence zero for smooth systems. We note that there are several further extensions of Takens' delay coordinate theorem, e.g., for forced systems [Sta99, SBDH03] or for the general infinite-dimensional non-autonomous case [Rob08].

The main goal in this thesis is the development of a set-oriented numerical framework for the (global) analysis of (parameter-dependent) infinite-dimensional dynamical systems regarding their invariant sets and their (geometric) structure. Given an infinite-dimensional dynamical system Φ with compact finite-dimensional invariant set \mathcal{A} we will utilize the infinite-dimensional embedding results of [HK99] and [MR19] for the construction of the so-called *core dynamical system* (CDS) φ . The CDS is a finite-dimensional dynamical system that reconstructs the original infinite-dimensional dynamics on \mathcal{A} . More precisely, we will embed \mathcal{A} into the observation space \mathbb{R}^k for some $k \geq 2d_{\text{box}}(\mathcal{A})$ using an observation map R . In particular, this embedded invariant set $A_k = R(\mathcal{A})$ will be one-to-one to the original set \mathcal{A} and hence can be considered for the analysis instead. Moreover, there is an inverse \tilde{E} of $R|_{\mathcal{A}}$ defined on A_k that can be extended to a map E on the entire space \mathbb{R}^k . Finally, the CDS is defined by the composition of the three maps R, Φ and E (see Figure 1.1 for an illustration), i.e.,

$$\varphi = R \circ \Phi \circ E.$$

This CDS then serves as a surrogate model defined on the observation space for the approximation of embedded attractors and embedded invariant manifolds of the infinite-dimensional system Φ . In this thesis we will particularly consider parameter-dependent infinite-dimensional dynamical systems Φ_λ , where a one-dimensional real-valued parameter λ influences the complexity of the dynamics. To this end, we will construct a *parameter-dependent core dynamical system* (pCDS) φ_λ as follows. We choose a dense subset of parameter values λ_n and utilize an observation map that is independent of all λ_n . Then it turns out that we can embed the parameter-dependent attractors \mathcal{A}^{λ_n} with a uniform observation map and proceed as before. The pCDS is then defined by

$$\varphi_\lambda = R \circ \Phi_\lambda \circ E$$

for all λ . In this form φ_λ only reproduces the dynamics of Φ_{λ_n} on \mathcal{A}^{λ_n} . For the remaining parameter values the observation map may not be one-to-one but close to it, provided the invariant set \mathcal{A}^λ behaves well under parameter perturbation. Thus, we can use φ_λ for the numerical analysis of the dynamics of Φ_λ for every λ .

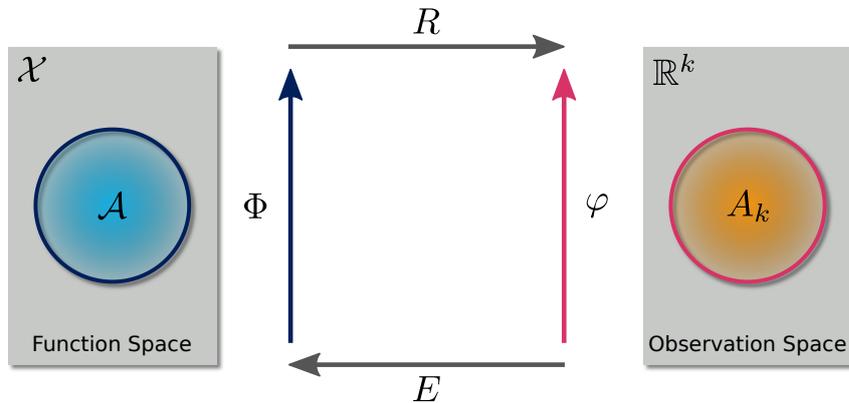


Figure 1.1: Illustration of the definition of the CDS φ .

As a starting point for the presented numerical methods for the analysis of infinite-dimensional systems in this thesis, we will review an extension of the classical subdivision scheme [DH97] developed in [DHZ16]. This tool allows us to compute the *relative global attractor* which is part of the embedded (global) attractor. Starting with an initial covering of the embedded invariant set in observation space we successively subdivide all sets in the covering and delete those sets which do not contain any part of the relative global attractor [DHZ16]. Repeating these two steps eventually leads to convergence to the relative global attractor. Hence, we can stop after a prescribed number of iterations or after the desired accuracy of the approximation is obtained. This procedure can be seen as a combination of cell mapping techniques [Hsu87] with a subdivision scheme.

Next, we will extend this subdivision algorithm to the efficient numerical approximation of attractors of parameter-dependent (infinite-dimensional) dynamical systems. Inspired by the numerical path following of steady states [AG03] we develop two *set-oriented path following methods* for the approximation of parameter-dependent attractors [GZED20, ZGD20] by reusing previously computed approximations. These algorithms allow us to track the attractor with respect to a one-dimensional parameter without restarting the whole subdivision scheme anew. In particular, we significantly reduce the overall computational cost since one computation of an attracting set might be very expensive. Note that these methods indirectly allow the numerical analysis of global bifurcations [Wig88, GS02]. For example, these types of bifurcations occur when invariant sets, such as periodic orbits, collide with equilibria. Therefore, typical examples are formations of homoclinic and heteroclinic orbits. We start with the approximation of the attractor $A_k^{\lambda_0}$ for a parameter value λ_0 using the (classical) subdivision method. Then we choose a computed covering of $A_k^{\lambda_0}$ as an initial guess for the approximation of the attractor for a parameter $\lambda_1 > \lambda_0$ sufficiently close to λ_0 . Finally, we employ the subdivision method on this initial computational domain and obtain an approximation of the attractor $A_k^{\lambda_1}$. Hence, repeating these two steps yields approximations of the attractor for discrete values in parameter space. We improve this scheme by developing an efficient *set-oriented predictor-corrector* method which will enhance the performance significantly [ZGD20]. That is, we use two previously computed coverings of the attractor for two consecutive parameter values in order to predict an initial overset of the attractor corresponding to a third parameter. This prediction step can be viewed as numerical realization of a *set-valued Taylor expansion*. Then a modification of the selection step in the subdivision algorithm serves as a corrector step such that ultimately the attractor is approximated.

Inspired by the classical continuation algorithm introduced in [DH96] the second method developed in this thesis aims at approximating embedded unstable invariant manifolds by set-oriented numerics [ZDG19]. Here we first employ the (extended) subdivision scheme on a small neighborhood of an unstable fixed point in observation space which yields a covering of its local unstable manifold. In a second step this covering is then globalized by a continuation procedure, that is, we successively compute parts of the embedded unstable manifold in a prescribed computational domain.

For the subsequent geometric analysis of the obtained approximation of some embedded invariant set of an infinite-dimensional dynamical system we will utilize manifold learning techniques [TDSL00, RS00, DG03, BN03, ZZ04]. In particular, we will focus on a tool called *diffusion maps* introduced in [CL06a]. Different from linear dimensionality reduction methods, such as principal component analysis (POD, cf. Section 4.1.2), this

nonlinear feature extraction method allows the identification of intrinsic coordinates of a given data set and hence revealing its intrinsic geometry. These coordinates are given by the eigenvectors and eigenvalues of a diffusion operator on the data. We note that the diffusion maps algorithm is robust to noise perturbation such that it can deal with the outer approximations of the set of interest generated by the proposed set-oriented schemes. In order to apply diffusion maps to such a covering, we will develop a novel *landmark selection* scheme that generates points that sample the embedded set sufficiently well. This scheme is inspired by 3d point cloud simplification methods (see, e.g., [PGK02] and the reference therein). Given an initial point cloud approximating the covering, we will iteratively move the points away from each other according to an repelling force that decays linearly with distance and becomes zero at a fixed radius $r > 0$. If during this procedure a point leaves the covering we project it back onto the covering. In the end, we obtain a point cloud such that the pairwise intersection of the r -balls around these points is (close to) empty. These landmarks will then be used as the underlying data for the construction of the diffusion maps embedding. In order to embed additional *out-of-sample* points, we extend the diffusion map algorithm inspired by the Nyström method [BPV⁺04, CL06b]. This allows us to compute the diffusion coordinates of (dynamically) import features contained in the invariant set such as equilibria, periodic orbits or long-term simulations which further reveal the intrinsic structure of the underlying set. For the construction of the diffusion maps a crucial parameter – the bandwidth of a kernel – has to be carefully chosen. To this end, we will employ an intrinsic dimension estimator that approximates the intrinsic dimension of the underlying data set and – in doing so – also automatically extracts an appropriate bandwidth [CSSS08, BH16]. Originally, this tool was designed for the analysis of uniformly sampled data which we will extend to the non-uniform case in this thesis.

This thesis is organized as follows. In [Chapter 2](#) we give a detailed review of embedding theorems in the context of finite-dimensional manifolds and fractal sets. We will employ the ideas and arguments used for a finite-dimensional ambient space for the treatment of sets contained in general Banach or Hilbert spaces. This allows us to state and prove the main results of [HK99] and [MR19], where in this thesis the latter is slightly extended to a prevalent version. Moreover, we briefly review some results in the context of delay coordinate embeddings. We note that in this thesis the original construction of the CDS given in [DHZ16] is improved such that the (p)CDS is not only continuous but even Hölder continuous. Moreover, we reformulate [DHZ16, Proposition 1] and prove that a prevalent set of Lipschitz maps generates a CDS.

In [Chapter 3](#) we present the numerical set-oriented framework for the approximation of embedded invariant sets. We start by reviewing the extension of the original subdivision scheme [DH97] such that it is applicable to the CDS [DHZ16]. Based on this work we will develop a path following method for the approximation of parameter-dependent attractors [GZED20, ZGD20]. Finally, we present an extension of the continuation method introduced in [DH96] for the approximation of embedded invariant manifolds [ZDG19].

In principle these algorithms are applicable to arbitrary infinite-dimensional dynamical systems. However, in this thesis we will restrict our attention to DDEs and PDEs. Therefore, in [Chapter 4](#) we will propose a suitable numerical realization of the CDS for both types of systems. On the one hand, we will utilize the delay coordinate map where the observable is the evaluation of the function at a prescribed point for the analysis of

DDEs. On the other hand, for PDEs, we decompose the state into modes and observe its corresponding coefficients. At the end of this chapter we introduce the diffusion map framework for the subsequent geometric analysis. In particular, we present the landmark selection scheme and the intrinsic dimension estimator.

In [Chapter 5](#) we apply the developed numerical algorithms to several examples. We will approximate the attractor of a reduced order model for turbulent shear flows for a range of Reynolds numbers. Moreover, we will employ the path following method for the approximation of the attractor of the Mackey-Glass equation for a range of delay times showing its bifurcation behavior on a global level. For a parameter value that presumably shows chaotic behavior we will investigate the generated covering using diffusion maps. Furthermore, the embedded unstable manifold of the one-dimensional Kuramoto-Sivashinsky equation in different regimes will be computed and its intrinsic structure will be analyzed by the diffusion map framework. Finally, we approximate the unstable manifold of the *edge state* in a plane Poiseuille flow.

We conclude this thesis with [Chapter 6](#), where we summarize the findings and give potential directions for future research.

Parts of this thesis grew out of publications to which the author has made substantial contributions. They are referenced at the beginning of the respective chapters and sections.

2 Embedding Theory

In this chapter we will lay the theoretical foundation for the following numerical analysis of infinite-dimensional systems such as partial differential equations (PDEs) and delay differential equations (DDEs). Following [Rob99] we will consider as a leading example for the discussion a dissipative PDE on a Hilbert space H of the form

$$\frac{\partial}{\partial t}u(t, x) = \mathcal{F}(u(t, x)), \quad u \in H, \quad (2.1)$$

that generates a semi-flow $S(t)$ for $t \geq 0$, i.e., a strongly continuous semigroup

$$S(0) = \mathbf{Id}, \quad S(s)S(t) = S(s+t), \quad S(t)u_0 \text{ continuous in } t \text{ and } u_0, \quad (2.2)$$

such that a unique solution $S(t)u_0 = u(t; u_0)$ exists for every initial condition $u_0 \in H$. The corresponding infinite-dimensional system in continuous time is then given by

$$u(t, x) = S(t)u(0, x), \quad t \geq 0. \quad (2.3)$$

In addition to that we will also consider the time- T -map Φ of (2.3) as the corresponding dynamical system in discrete time, i.e.,

$$u_{j+1}(x) = \Phi u_j(x), \quad j = 0, 1, \dots, \quad (2.4)$$

where the time $T > 0$ is fixed.

As in the study of finite-dimensional systems such as ordinary differential equations, the asymptotic behavior of solutions and properties of globally attracting sets are of particular interest. Thus, we further assume that (2.1) possesses a global attractor $\mathcal{A} \subseteq H$, i.e., a maximal compact invariant set which attracts all bounded sets:

$$S(t)\mathcal{A} = \mathcal{A}, \quad \text{dist}(S(t)B, \mathcal{A}) \rightarrow 0 \text{ as } t \rightarrow \infty \text{ for all bounded sets } B \subseteq H, \quad (2.5)$$

where the distance for two subsets B, C of H is given by

$$\text{dist}(B, C) = \sup_{u \in B} \inf_{v \in C} \|u - v\|, \quad (2.6)$$

where $\|\cdot\|$ denotes the norm on H . It has been shown that many dissipative PDEs possess (nontrivial) globally attracting sets which are finite-dimensional, even though the underlying state space is a suitable infinite-dimensional space. We will later define an appropriate measure for the dimension of arbitrary subsets of a Banach space. Equations that fall into this category are, for instance, reaction-diffusion equations [Jol89], the 2D Navier-Stokes equation, the Kuramoto-Sivashinsky equation [KT76, Siv77], the Ginzburg-Landau equation and other examples (see [Tem97, Hal10]).

Since the attractor is finite-dimensional, a natural question that arises in this context is whether there exists a finite-dimensional dynamical system that possesses the asymptotic dynamics on the attractor of the original flow (2.3). That is, can we *embed* the dynamics of S (or Φ) on \mathcal{A} into some \mathbb{R}^k , where $k \in \mathbb{N}$ is chosen appropriately? More precisely, the task is to find an *embedding* $R : H \rightarrow \mathbb{R}^k$ of \mathcal{A} into \mathbb{R}^k , i.e., a homeomorphism

$R : \mathcal{A} \rightarrow R(\mathcal{A}) \subseteq \mathbb{R}^k$ and, for the continuous time case, a system of ordinary differential equations (ODEs)

$$\dot{x}(t) = f(x(t)), \quad x \in \mathbb{R}^k, \quad (2.7)$$

with unique solutions and global attractor $A_k = R(\mathcal{A})$ on which the dynamics is conjugated to those on \mathcal{A} via R . That is, if $T(t)$ is the flow of (2.7), we have

$$T(t) \circ R = R \circ S(t) \text{ on } \mathcal{A}.$$

Accordingly, for a discrete dynamical system $\Phi : H \rightarrow H$ we search for a continuous map $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ such that

$$\varphi \circ R = R \circ \Phi \text{ on } \mathcal{A}.$$

However, proving the mere existence of such an embedding R is not sufficient for our purpose since we want to numerically analyze the finite-dimensional systems (2.7) and φ . To this end, the map R has to be explicitly known or constructed such that it can be implemented numerically. If this is not possible, we at least desire that the class of such embeddings R is in some sense “large”, e.g., it is a *generic* property that a map R is an embedding. In the following section we discuss various solutions to this problem.

2.1 Inertial Manifolds

One approach to tackle the problem at hand is to prove the existence of a so-called *inertial manifold*, a finite-dimensional, positively invariant manifold \mathcal{M} that attracts all trajectories exponentially fast [CFNT89, FJK⁺88, Tem97], i.e.,

$$S(t)\mathcal{M} \subseteq \mathcal{M} \text{ and } \text{dist}(S(t)u_0, \mathcal{M}) \leq C_{u_0} e^{-\kappa t}, \quad \forall t \geq 0,$$

for some $\kappa > 0$ and constant $C_{u_0} > 0$ depending only on the initial condition $u_0 \in H$. In this case the attractor $\mathcal{A} \subseteq H$ lies inside the manifold \mathcal{M} , that is, $\mathcal{A} \subseteq \mathcal{M}$. Thus, the problem can be simplified to embedding the inertial manifold instead. However, there are some drawbacks that will be discussed later.

For the sake of completeness we summarize how the existence of such inertial manifolds can be proven (see [FST88] for details). Typically, they have been studied for systems of the form

$$\frac{\partial}{\partial t} u(t, x) + Au(t, x) = f(u(t, x)),$$

where A is a positive, linear and self-adjoint operator with compact inverse and f is a Lipschitz function from the domain $D(A^\alpha)$ into H for some $0 \leq \alpha < 1$. According to [RR06] such operators have a set of orthonormal eigenfunctions $\{v_k\}_k$ with corresponding (ordered) eigenvalues $\{\lambda_k\}_k$ such that

$$Av_k = \lambda_k v_k, \quad \lambda_{k+1} \geq \lambda_k, \quad \lambda_k \rightarrow \infty.$$

In this situation one can define finite-dimensional projection operators P_k and their orthogonal complements Q_k by

$$P_k u = \sum_{i=1}^k \langle u, v_i \rangle v_i, \quad Q_k u = \sum_{i=k+1}^{\infty} \langle u, v_i \rangle v_i,$$

where $\langle \cdot, \cdot \rangle$ is the scalar product on H . If there is a certain sufficiently large spectral gap in the eigenvalues λ_k , one can construct a (at least) Lipschitz function $\phi : P_k H \rightarrow Q_k H$ that allows one to define the manifold \mathcal{M} as a graph over one of the finite-dimensional subspaces $P_k H$, i.e.,

$$\mathcal{M} = \{ p + \phi(p) \mid p \in P_k H \}.$$

The ordinary differential equation (2.7) is then given by restricting the flow (2.2) to the manifold \mathcal{M} . For $p = P_k u$ this yields

$$\dot{p} + Ap = P_k(f(p + \phi(p))), \quad (2.8)$$

where the fact that P_k commutes with A is used. Since $p \in P_k H \simeq \mathbb{R}^k$ and ϕ is Lipschitz, (2.8) has unique solutions by the Picard-Lindelöf theorem (see, e.g., [Tes12, Theorem 2.2]). On $P_k \mathcal{A}$ the solutions are precisely those projected down from \mathcal{A} such that

$$p(t) = P_k S(t)(p(0) + \phi(p(0))).$$

It follows that $P_k \mathcal{A}$ is the global attractor of (2.8) because \mathcal{M} is invariant. Finally, with $P_k^{-1}x = x + \phi(x)$ we precisely obtain a finite-dimensional system of the form (2.7).

At a first glance, one might think that this approach can not be used directly for numerical applications since the map ϕ , and thus the embedding R , is not explicitly known. However, its construction is based on a fixed point problem that is solved using the Banach fixed point theorem (see, e.g., [For13]). Consequently, ϕ could at least be approximated and there are also error bounds. If one neglects the construction of \mathcal{M} , embedding techniques like Whitney's theorem (see Section 2.2.1) still allow the definition of an equivalent system on \mathbb{R}^{2k+1} . But this is not preferable due to the increase of the system's dimension. In fact, since $\mathcal{A} \subseteq \mathcal{M}$ the dimension $k \in \mathbb{N}$ of the manifold, and hence of the derived system of differential equation (2.8), is much larger than the dimension of the attractor itself (see for example [CEES93, Rob94, TW94, Tem97] for the Kuramoto-Sivashinsky equation). Thus, especially from the numerical point of view, the goal is to produce a system of ordinary differential equations of dimension which is comparable to the dimension of the attractor. That is, with the idea of Whitney's theorem in mind, we aim to embed \mathcal{A} into \mathbb{R}^k for some $k \in \mathbb{N}$.

In addition to that, the conditions known to be sufficient for proving the existence of an inertial manifold are very restrictive. While they are satisfied for those examples mentioned above there are cases in which one can prove the existence of a finite-dimensional attractor but not (yet) of an inertial manifold. Arguably, the most interesting example that falls into this category might be the $2d$ Navier-Stokes equation [Rob13].

This is why in the thesis we are going to directly work with the attractor and embed its dynamics into \mathbb{R}^k as discussed in the following sections.

2.2 Embedding Finite-dimensional Sets into \mathbb{R}^k

A different solution to the question stated at the beginning of this chapter is given by utilizing embedding techniques. More precisely, we show that so-called *observations* enable us to embed the attractor \mathcal{A} into a finite-dimensional Euclidean space of sufficiently high dimension in a generic way.

However, if one embeds the dynamics only on the attractor, one neglects the transient behavior (which is included to some extent for the inertial manifolds). That is why we have to make sure that the asymptotic behavior of solutions of the infinite-dimensional systems is determined to some extent by the dynamics on the global attractor in order to justify the restriction to the attractor. Proposition 2.1 solves this issue for our guiding example. It shows that solutions $u(t)$ are followed by trajectories on the attractor more and more closely for longer and longer time.

Proposition 2.1 ([LR99, Corollary 3.2]). *Given a solution $u(t)$ of (2.1) there exists a sequence $(\varepsilon_n)_n$,*

$$\varepsilon_n > 0, \quad \varepsilon_n \rightarrow 0 \text{ as } n \rightarrow \infty,$$

a sequence of times $(t_n)_n$,

$$t_{n+1} > t_n \geq 0, \quad t_{n+1} - t_n \rightarrow \infty \text{ as } n \rightarrow \infty$$

and a sequence of points $(v_n)_n \subseteq \mathcal{A}$ such that

$$\|u(t) - S(t - t_n)v_n\| \leq \varepsilon_n \quad \text{for all } t_n \leq t \leq t_{n+1}.$$

Furthermore, the jumps $|v_{n+1} - S(t_{n+1} - t_n)v_n|$ decrease to zero as $n \rightarrow \infty$.

We note that the corresponding proof only relies on the attraction property of \mathcal{A} and the continuity of solutions with respect to initial conditions. This allows us to use this result for more general flows and, in particular, it covers delay differential equations as well.

2.2.1 Finite-dimensional Manifolds

To motivate the idea of embedding \mathcal{A} into a finite-dimensional space, we will summarize some results on embeddings in the context of finite-dimensional manifolds. For simplicity we will consider a smooth compact manifold \mathcal{M} of dimension $d \in \mathbb{N}$ and show that \mathcal{M} is generically embedded in \mathbb{R}^{2d+1} , that is, the set of all smooth embeddings from \mathcal{M} to \mathbb{R}^{2d+1} is open and dense. To begin with, let us state the celebrated result of Hassler Whitney [Whi36].

Theorem 2.2 (Whitney's Embedding Theorem, [Lee13, Theorem 6.15]). *Every smooth manifold of dimension $d \in \mathbb{N}$ admits a smooth embedding into \mathbb{R}^{2d+1} .*

A few years later in [Whi44], this result was improved to the so-called *strong Whitney embedding theorem*, reducing the required dimension to $2d$. But for reasons of comparison to the results in the infinite-dimensional context, Theorem 2.2 is sufficient for our purpose.

With this at hand, the existence of an embedding is guaranteed. But in particular for numerical applications and for experiments, a stronger result is required. To this end, we introduce a topology on all smooth maps between two smooth manifolds in order to improve the mere existence to openness and denseness of embeddings.

Definition 2.3 ((Weak) Topology on $C^\infty(\mathcal{M}, \mathcal{N})$). Let \mathcal{M}, \mathcal{N} be smooth manifolds and $f \in C^\infty(\mathcal{M}, \mathcal{N})$. Given charts $(\phi, U), (\psi, V)$ on \mathcal{M}, \mathcal{N} , a compact set $K \subseteq U$ such that $f(K) \subseteq V$ and $\varepsilon > 0$, we define a *basis neighborhood*

$$\mathcal{U}(f; (\phi, U), (\psi, V), K, \varepsilon) \tag{2.9}$$

to be the set of smooth maps $g : \mathcal{M} \rightarrow \mathcal{N}$ such that $g(K) \subseteq V$ and

$$\sup_{x \in \phi(K), r \in \mathbb{N}_0} \left\| D^r(\psi f \phi^{-1})(x) - D^r(\psi g \phi^{-1})(x) \right\| < \varepsilon.$$

This means that the local representation of f and g , together with all their derivatives, have a distance less than ε at each point in K . The (*weak*) *topology* on $C^\infty(\mathcal{M}, \mathcal{N})$ is then generated by these sets (2.9) and it defines the topological space $C_W^\infty(\mathcal{M}, \mathcal{N})$.

We remark that there is also the *strong* or *Whitney* topology, but for a compact manifold \mathcal{M} the resulting topological spaces are the same. Endowing $C^\infty(\mathcal{M}, \mathbb{R}^{2d+1})$ with that topology allows us to state that embeddings are open and dense.

Theorem 2.4 ([Hir12, Proposition 1.0 & Theorem 1.4]). *Let \mathcal{M} be a compact smooth manifold of dimension $d \in \mathbb{N}$. Then the set of embeddings $\mathcal{M} \rightarrow \mathbb{R}^{2d+1}$ is open and dense in $C_W^\infty(\mathcal{M}, \mathbb{R}^{2k+1})$.*

Thus, given an embedding, every small perturbation remains an embedding and there exists an embedding arbitrary close to every smooth map, regardless whether it is an embedding or not. This (partially) justifies to assume that an arbitrary smooth map $R : \mathcal{M} \rightarrow \mathbb{R}^{2d+1}$ is, in particular, one-to-one.

2.2.2 Fractal Sets in \mathbb{R}^N

For the moment let us consider a finite-dimensional subset A of \mathbb{R}^N for which we will derive a corresponding embedding result before moving on to infinite-dimensional spaces. We will see that the underlying ideas will be helpful for proving the desired infinite-dimensional embedding results.

First of all, we formally introduce the measure of dimensionality for a set $A \subseteq \mathbb{R}^N$ we consider in this thesis.

Definition 2.5 (Box-counting dimension, see, e.g., [Fal14]). Let \mathcal{X} be a metric space and $A \subseteq \mathcal{X}$. For $\varepsilon > 0$, denote by $N(A, \varepsilon)$ the minimal number of closed balls of radius ε with centers in A necessary to cover A . The (*upper*) *box-counting dimension* of A is defined by

$$d_{\text{box}}(A) = \limsup_{\varepsilon \rightarrow 0} \frac{\log N(A, \varepsilon)}{-\log \varepsilon},$$

essentially $N(A, \varepsilon) \sim \varepsilon^{-d_{\text{box}}(A)}$ as $\varepsilon \rightarrow 0$.

Clearly, by definition, if $d > d_{\text{box}}(A)$ then there exist $\varepsilon_0 > 0$ and $C_{A,d} > 0$ such that

$$N(A, \varepsilon) < C_{A,d} \cdot \varepsilon^{-d} \quad \text{for all } \varepsilon < \varepsilon_0, \quad (2.10)$$

which will take an important role in the proof of the upcoming embedding results. In addition to that, we note that there are several equivalent definitions, e.g., using the diameter instead of the radius, and refer to [Fal14] for further reading.

Remark 2.6. In this thesis we will utilize set-oriented numerical tools that generate successively finer coverings Q_ℓ of a set of interest $A \subseteq \mathcal{X}$. Hence, we can estimate an upper bound for the box-counting dimension by approximating $N(A, \varepsilon)$ by the number of sets forming the covering (see [ST06] for corresponding error estimates).

In the following proposition we summarize some key properties of the box-counting dimension that will be used later on.

Proposition 2.7. *Let $(\mathcal{X}, \rho_{\mathcal{X}})$ and $(\mathcal{Y}, \rho_{\mathcal{Y}})$ be metric spaces.*

(a) *If $A \subseteq B \subseteq \mathcal{X}$, then $d_{\text{box}}(A) \leq d_{\text{box}}(B)$.*

(b) *Suppose $\mathcal{X} \times \mathcal{Y}$ is equipped with the “ p -metric” ρ_p , $p \geq 1$, i.e.,*

$$\rho_p((x_1, y_1), (x_2, y_2)) = (\rho_{\mathcal{X}}(x_1, x_2)^p + \rho_{\mathcal{Y}}(y_1, y_2)^p)^{\frac{1}{p}},$$

or the “maximum metric” ρ_∞ , where

$$\rho_\infty((x_1, y_1), (x_2, y_2)) = \max \{ \rho_{\mathcal{X}}(x_1, x_2), \rho_{\mathcal{Y}}(y_1, y_2) \}.$$

Then for $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$ it follows that

$$d_{\text{box}}(A \times B) \leq d_{\text{box}}(A) + d_{\text{box}}(B).$$

In particular, we can apply the above proposition to the case where $Z \subseteq A \times A$ and $d > d_{\text{box}}(A)$. That is,

$$d_{\text{box}}(Z) \leq d_{\text{box}}(A \times A) \leq 2d_{\text{box}}(A) \leq 2d. \quad (2.11)$$

This equation will also be important in the upcoming proofs.

According to [Theorem 2.4](#) embeddings on a manifold are open and dense. One might conclude that open and dense sets are large in terms of probability, but this is not necessarily the case as the following example shows.

Example 2.8 (cf. [Arn12, SYC91]). Consider the family of circle diffeomorphisms $g_{\omega,k} : S^1 \rightarrow S^1$ given by

$$g_{\omega,k}(x) = x + \omega + k \sin(x) \quad \text{mod } 2\pi,$$

where $0 \leq \omega \leq 2\pi$ and $0 \leq k < 1$ and define for each k the set

$$\text{Stab}(k) = \{ 0 \leq \omega \leq 2\pi \mid g_{\omega,k} \text{ has a stable periodic orbit} \}.$$

Then the set $\text{Stab}(k)$ is a countable union of disjoint open intervals of positive Lebesgue measure for $0 < k < 1$ and, additionally, is an open dense subset of $[0, 2\pi]$. But the Lebesgue measure of $\text{Stab}(k)$ approaches zero for $k \rightarrow 0$. Thus, the Lebesgue measure of an open and dense set can be arbitrary small.

With that in mind we would like to make a stronger statement. Given a particular map f we would like to show that almost every map near f is an embedding. More precisely, the central statement of the main theorem in this section is to show that μ -almost every linear perturbation of a smooth map $f : \mathbb{R}^N \rightarrow \mathbb{R}^k$ is one-to-one on $A \subseteq \mathbb{R}^N$ and has Hölder continuous inverse on A , provided $k \in \mathbb{N}$ is large enough. Here μ is a probability measure on $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ with compact support defined as follows.

Any linear map $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ is given by k linear maps $\mathbb{R}^N \rightarrow \mathbb{R}$ and each of them corresponds to taking the scalar product with some $l_j \in \mathbb{R}^N$. We write l_j^* for the corresponding linear map from \mathbb{R}^N into \mathbb{R} given by $x \mapsto l_j^T x$. This construction might be unnecessary and complicated since L is just a matrix, but we explicitly use this notation to establish a connection to infinite-dimensional Hilbert spaces H discussed later, where the corresponding dual space H^* along with the Riesz–Fréchet representation theorem is used. The support of μ is then defined as all linear maps with $\|l_j\| \leq 1, j = 1, \dots, k$, i.e.,

$$E = \{ L = (l_1^*, \dots, l_k^*) \mid l_j \in B_N \},$$

where $B_N = B(0, 1) = \{x \in \mathbb{R}^N \mid \|x\| \leq 1\} \subseteq \mathbb{R}^N$ is the unit ball in \mathbb{R}^N . Now we identify E with $(B_N)^k$ and equip E with the product measure $\mu = \bigotimes_{j=1}^k \lambda_N$, where λ_N is the normalized Lebesgue measure on $B_N \subseteq \mathbb{R}^N$. Note that μ is a probability measure with compact support E . Furthermore, the following estimate is essential for the proof of the embedding theorem in this section.

Lemma 2.9 ([Rob10, Lemma 4.1]). *For any $\alpha \in \mathbb{R}^k$ and $x \in \mathbb{R}^N$ we have*

$$\mu \{ L \in E \mid |\alpha + Lx| \leq \varepsilon \} \leq cN^{k/2} \left(\frac{\varepsilon}{|x|} \right)^k, \quad (2.12)$$

where $c > 0$ is an absolute constant.

The embedding theorem then states that there is a “large” set of maps that embed A into \mathbb{R}^k with Hölder continuous inverse on A provided $k > 2d_{\text{box}}(A)$.

Theorem 2.10 ([HK99, Theorem 3.1]). *Let $A \subseteq \mathbb{R}^N$ be a compact set with box-counting dimension $d_{\text{box}}(A) = d$. Let $k > 2d$ be an integer and $\theta \in \mathbb{R}$ with*

$$0 < \theta < \frac{k - 2d}{k}.$$

Then, given any Lipschitz function $f : \mathbb{R}^N \rightarrow \mathbb{R}^k$, for μ -almost every map $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ there exists $C = C_L > 0$ such that

$$|x - y| \leq C |(f + L)(x) - (f + L)(y)|^\theta \quad \text{for all } x, y \in A. \quad (2.13)$$

In particular, $f + L$ is one-to-one on A with a θ -Hölder continuous inverse on A .

Note that for $f = 0$ it follows that μ -almost every $L \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^k)$ satisfies (2.13) and is one-to-one on $A \subseteq \mathbb{R}^N$. But keeping f arbitrary is the key idea in the notion of *prevalence* which allows for similar results in the infinite-dimensional context and will be discussed later on. Also observe that the smoothness of the inverse depends on the embedding dimension $k \in \mathbb{N}$. In particular, in the limit $k \rightarrow \infty$ the bound on the Hölder exponent θ converges to one, i.e., $f + L$ is a bi-Lipschitz map on $A \subseteq \mathbb{R}^N$ in this limit.

Proof. We follow the proof given in [Rob10, Theorem 4.3]. First, we aim to bound the measure of linear maps $L \in E$ that do not satisfy

$$|(f(x) + L)(x) - (f + L)(y)|^\theta > |x - y| \text{ for some pair } (x, y) \in A \times A.$$

Thus, we define

$$Z_n = \{ (x, y) \in A \times A \mid |x - y| \geq 2^{-n\theta} \} \subseteq A \times A$$

and set

$$Q_n = \{ L \in E \mid |(f + L)(x) - (f + L)(y)| \leq 2^{-n} \text{ for some } (x, y) \in Z_n \}.$$

Essentially, Q_n represents the set of “bad” maps for which (2.13) fails for some (x, y) with $|x - y| \geq 2^{-n\theta}$. According to Proposition 2.7 or (2.11) we estimate the box-counting dimension of Z_n by $d_{\text{box}}(Z_n) \leq d_{\text{box}}(A \times A) \leq 2d_{\text{box}}(A)$, where we endowed $\mathbb{R}^N \times \mathbb{R}^N$ with the 1-metric, that is,

$$|(x_1, y_1), (x_2, y_2)|_1 = |x_1 - x_2| + |y_1 - y_2|.$$

Thus, after choosing and fixing $d > d_{\text{box}}(A)$, we can cover Z_n by no more than $C_{A,d} \cdot 2^{2nd}$ balls $B((x_j, y_j), 2^{-n})$ of radius 2^{-n} whose centers (x_j, y_j) lie in Z_n . That is, we consider $\varepsilon = 2^{-n}$ in (2.10). Since E is compact, $\|L\|$ is uniformly bounded. Hence, there is a Lipschitz constant $M > 0$ that is uniformly valid for $f + L$ for all $L \in E$. Suppose there is $(x, y) \in Y_j = Z_n \cap B((x_j, y_j), 2^{-n})$ with $|(f + L)(x) - (f + L)(y)| \leq 2^{-n}$. Then

$$\begin{aligned} |(f + L)(x_j) - (f + L)(y_j)| &\leq M|x_j - x| + M|y_j - y| + |(f + L)(x) - (f + L)(y)| \\ &\leq M|(x_j, y_j), (x, y)|_1 + 2^{-n} \\ &\leq (M + 1)2^{-n} \\ &= C_f \cdot 2^{-n}, \end{aligned}$$

where the constant $C_f > 0$ only depends on f . Finally, using Lemma 2.9 we can estimate

$$\begin{aligned} &\mu \{ L \in E \mid |(f + L)(x) - (f + L)(y)| \leq 2^{-n} \text{ for some } (x, y) \in Y_j \} \\ &\leq \mu \{ L \in E \mid |(f + L)(x_j) - (f + L)(y_j)| \leq C_f \cdot 2^{-n} \} \\ &= \mu \{ L \in E \mid |(f(x_j) - f(y_j)) + L(x_j - y_j)| \leq C_f \cdot 2^{-n} \} \\ &\leq cN^{k/2} \left(\frac{C_f \cdot 2^{-n}}{|x_j - y_j|} \right)^k \leq C_{f,N,k} \cdot \left(\frac{2^{-n}}{2^{-n\theta}} \right)^k = C_{f,N,k} \cdot 2^{(\theta k - k)n}, \end{aligned} \tag{2.14}$$

where $C_{N,k,f} > 0$ only depends of N, k and f .

Thus, the total measure Q_n is bounded by

$$\mu(Q_n) \leq C_{A,d} \cdot 2^{2nd} \cdot C_{f,N,k} 2^{(\theta k - k)n} = C_{A,d,f,N,k} \cdot 2^{(2d - k + \theta k)n}$$

and we have $2d - k + \theta k < 0$ by assumption. Hence, the series $\sum_{n=1}^{\infty} \mu(Q_n) < \infty$ converges and the Borel-Cantelli lemma (see, e.g., [Kle13, Theorem 2.7]) implies that μ -almost every $L \in E$ belongs to only finitely many Q_n . For such an L , there exist a n_L such that $L \notin Q_n$ for all $n \geq n_L$, i.e.,

$$|x - y| \geq 2^{-n\theta} \implies |(f + L)(x) - (f + L)(y)| \geq 2^{-n} \quad \text{for all } n \geq n_L.$$

Since A is bounded, so is the set $A - A = \{x - y \mid x, y \in A\}$ and thus there exist $R > 0$ such that $|x - y| \leq R$ for all $x, y \in A$. We conclude that for $|x - y| > 2^{-n_L\theta}$ we have

$$|(f + L)(x) - (f + L)(y)| \geq 2^{-n_L} \geq \frac{2^{-n_L}}{R^{1/\theta}} |x - y|^{1/\theta}$$

and if $|x - y| \leq 2^{-n_L\theta}$, then there is $n \geq n_L$ such that $2^{-(n+1)\theta} \leq |x - y| < 2^{-n\theta}$. Thus,

$$|(f + L)(x) - (f + L)(y)| \geq 2^{-(n+1)} \geq \frac{1}{2} |x - y|^{1/\theta},$$

which implies the conclusion as

$$|(f + L)(x) - (f + L)(y)| \geq \max \left\{ \frac{1}{2}, \frac{2^{-n_L}}{R^{1/\theta}} \right\} |x - y|^{1/\theta}.$$

□

Remark 2.11.

- (a) In 1991 Sauer et al. already proved a similar result but without Hölder continuity of the inverse [SYC91, Theorem 2.3], i.e., $f + L$ is one-to-one on A . However, they additionally added that $f + L$ is an immersion on each compact subset of a smooth manifold contained in A . This has great importance when computing and analyzing invariant manifolds that lie inside the attractor.
- (b) As seen in the proof, a feasible embedding dimension $k \in \mathbb{N}$ is connected to the dimension of $A \times A$, that is,

$$k > 2d > 2d_{\text{box}}(A) \geq d_{\text{box}}(A \times A) \geq d_{\text{box}}(Z_n).$$

That is why the Hausdorff-dimension (see [Fal14] for a definition) can not be used in this theorem, since only the reverse inequality holds, i.e.,

$$d_H(A \times A) \geq 2d_H(A),$$

even though d_H behaves similar to d_{box} under Lipschitz- or Hölder maps (cf. Proposition 2.7).

2.2.3 Infinite-dimensional State Spaces

In order to prove a similar result to [Theorem 2.10](#) for a subset of an infinite-dimensional space, we have to adapt a number of ingredients. Recall that the measure μ constructed in the last section is given as the product of normalized Lebesgue measures. In particular, for $k = 1$, the space of linear maps $\mathbb{R}^N \rightarrow \mathbb{R}$ is identified with the finite-dimensional space \mathbb{R}^N itself. Analogously, if the underlying space is an infinite-dimensional Hilbert space H , we could identify the dual space H^* , i.e., the space of (bounded) linear maps $H \rightarrow \mathbb{R}$, with H itself using Fréchet-Riesz representation theorem (see, e.g., [[Wer11](#), Theorem V.3.6]). However, since H is infinite-dimensional, the Lebesgue measure can not be used anymore and we have to adapt the construction of μ . In fact, there is no measure on a Banach space that corresponds to the Lebesgue measure on finite-dimensional subspaces. This is why we first have to extend the notion of “Lebesgue almost every” linear map to an appropriate concept. To this end, the term *prevalence* was coined in [[HSY92](#)] (see also [[OY05](#)]) which, additionally, turns out to be a more appropriate condition than the topological concepts of “open and dense” in the context of probabilistic results on the likelihood of a given property in a function space.

Definition 2.12 (Prevalence, [[HSY92](#)]). A measure μ is *transverse* to a Borel set $S \subseteq V$ of a vector space V if the two following conditions hold:

- (i) There is a compact set $U \subseteq V$ for which $0 < \mu(U) < \infty$.
- (ii) $\mu(S + v) = 0$ for every $v \in V$.

In this case S will be called *shy* (with respect to μ). If S is not a Borel subset, we say that S is *shy* if it is contained in a shy Borel subset. The complement of a shy set is *prevalent*.

By condition (i) a transverse measure μ can be restricted to a probability measure on the compact set U . In fact, if S is shy with respect to μ , S is also shy with respect to the restricted measure. This is why we can replace condition (i) if needed with

- (i*) μ is a probability measure with compact support $E \subseteq V$.

In this case, for a more intuitive understanding of prevalence, one can think of E as the *probe space* of allowable perturbations. In this sense, $S \subseteq V$ is *prevalent* if for every $v \in V$, $v + e \in S$ for μ -almost every $e \in E$. This means that μ -almost every perturbation of any point in the ambient space V lies inside S . With that in mind, the construction of a the set E and a corresponding measure μ on E is critical for proving that a set is prevalent.

Additionally, it turns out that the support of the restricted measure can be taken to have arbitrary small support. To see this, cover the compact set U by a finite number of balls of (arbitrary small) radius $\varepsilon > 0$. Then at least one of those balls must intersect U in set of positive μ measure. The intersection of U with the closure of that ball is compact and condition (ii) also holds for the restriction of μ to this set.

Nonetheless, for applications it is often useful to consider measures which are neither finite nor have compact support such as the Lebesgue measure. Roughly speaking, the less concentrated the measure μ is, the more sets it is transverse to. In particular, for a Dirac measure only the empty set is shy and thus only the entire space V is prevalent. In order to show that prevalence extends the notion of “Lebesgue almost all” to infinite-dimensional spaces, we summarize some of its properties.

Proposition 2.13. *Let V be a vector space.*

- (a) *If $S \subseteq V$ is shy (resp. prevalent), then the translated set $S + v = \{s + v \in V \mid s \in S\}$ for $v \in V$ is also shy (resp. prevalent).*
- (b) *The countable union of shy sets in V is itself shy. Therefore, a countable intersection of prevalent sets in V is itself prevalent.*
- (c) *Suppose V is finite-dimensional. Then $S \subseteq V$ is shy if and only if S has Lebesgue measure zero.*
- (d) *A prevalent set $S \subseteq V$ is dense.*

Proof.

- (a) This is an immediate consequence of the definitions of shyness and prevalence in Definition 2.12.
- (b) For the proof we restrict ourselves to considering only finite collections of shy sets and refer to [HSY92] for the general case. We only consider two shy sets $S, T \subseteq V$ with transverse probability measures μ and ν . The result for arbitrary finite collections follows then by induction. A measure which is transverse to both S and T is then given by the convolution $\mu * \nu$ of μ and ν defined for a Borel subset $B \subseteq V$ by

$$\mu * \nu(B) = \mu \times \nu(B^\Sigma),$$

where $\mu \times \nu$ is the product measure on $V \times V$ and $B^\Sigma = \{(x, y) \in V \times V \mid x + y \in B\}$ is a Borel subset of $V \times V$. Since μ and ν are finite, so is $\mu \times \nu$ and by Fubini's theorem (see, e.g., [DS88]) it follows

$$\mu * \nu(B) = \int_V \mu(B - y) d\nu(y) = \int_V \nu(B - x) d\mu(x).$$

Note that the characteristic function of B^Σ is integrable with respect to $\mu \times \nu$. Thus, $\mu * \nu$ is transverse to S and T and for $v \in V$ it follows

$$\begin{aligned} \mu * \nu([S \cup T] + v) &\leq \mu * \nu(S + v) + \mu * \nu(T) \\ &= \int_V \mu(S + v - y) d\nu(y) + \int_V \nu(S - x) d\mu(x) \\ &= 0, \end{aligned}$$

i.e., $S \cup T$ is shy with respect to $\mu * \nu$.

- (c) Without loss of generality we can assume that $V = \mathbb{R}^n$ and $S \subseteq \mathbb{R}^n$ is a Borel set, since we are only interested in whether or not sets have measure zero and subsets of Borel sets with Lebesgue measure zero also have Lebesgue measure zero. If S has Lebesgue measure zero, then certainly the Lebesgue measure itself is transverse to S and thus S is shy. On the other hand, if S is shy, then there is probability measure μ on \mathbb{R}^n with compact support which is transverse to S . It follows by Tonelli's theorem (see, e.g., [DS88]) that

$$0 = \int_{\mathbb{R}^n} \mu(S - y) d\lambda(y) = \int_{\mathbb{R}^n} \lambda(S - x) d\mu(x) = \mu(\mathbb{R}^n) \lambda(S) = \lambda(S),$$

where λ is the Lebesgue measure on \mathbb{R}^n .

- (d) As already mentioned above, for a shy Borel set S the support of its corresponding measure can be taken to have arbitrary small diameter. Hence, S has no interior and the same is true for every shy set, since those are contained in a shy Borel set. It follows that a prevalent set is dense. □

In this work, the notion of prevalence will be used in the context where the underlying vector space V is a function space. From now on we will say that “almost every” map in a function space V satisfies a certain property if the set of such maps is prevalent, even in the infinite-dimensional case. In this case this property will also be called *generic* (in the sense of prevalence).

Remark 2.14.

- (a) In [SYC91] Whitney’s embedding Theorem 2.4 is strengthened using the notion of prevalence. Let \mathcal{M} be a compact smooth manifold contained in \mathbb{R}^N of dimension $d \in \mathbb{N}$. Then almost every (in the sense of prevalence) smooth map $\mathbb{R}^N \rightarrow \mathbb{R}^{2d+1}$ is an embedding of $\mathcal{M} \subseteq \mathbb{R}^N$.
- (b) Theorem 2.10 can be reformulated in terms of the notion prevalence, that is, under the assumptions made there, almost every (in the sense of prevalence) Lipschitz map $\mathbb{R}^N \rightarrow \mathbb{R}^k$ satisfies (2.13).

Hilbert Spaces

As already discussed in the beginning of this section our goal is to extend the finite-dimensional embedding Theorem 2.10 to an infinite-dimension ambient space. To this end, we first consider a Hilbert space H and use the ideas in Section 2.2.2 and construct a probe space E and a measure μ with support E on the set of bounded linear maps $H \rightarrow \mathbb{R}^k$ as follows.

In the finite-dimensional setting we essentially defined E to be k -fold Cartesian product of the unit ball B_N . However, for a infinite-dimensional Hilbert space this idea has to be slightly adjusted since E would not be compact. To begin with, we consider $k = 1$, i.e., the dual space H^* , and define a probe space E_0 that afterwards serves as the basis for the general case $k \in \mathbb{N}$.

Let $\mathcal{V} = \{V_n\}_{n=1}^\infty$ be a sequence of subspaces $V_n \subseteq H$ with dimension $d_n < \infty$ and let S_n be the unit ball in V_n . As in the finite-dimensional case, for $l \in H$ we denote by l^* the element in H^* given by $x \mapsto \langle l, x \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the scalar product in H . Note that $l \mapsto l^*$ is in fact an isometric (anti-)isomorphism by Fréchet-Riesz representation theorem. Then we define the probe space E_0 by

$$E_0(\mathcal{V}) = \left\{ \left(\sum_{n=1}^\infty n^{-2} \phi_n \right)^* \mid \phi_n \in S_n \right\} \subseteq H^*.$$

Since V_n is finite-dimensional, it is straightforward to show that E_0 is indeed compact by Tychonoff’s theorem (see, e.g., [Jä05]). Note that S_n can be identified with the unit ball $B_{d_n} \subseteq \mathbb{R}^{d_n}$ by choosing an orthonormal basis of V_n . Thus, the uniform probability

measure on B_{d_n} induces a measure λ_n on S_n and we obtain a probability measure μ_0 on E_0 by

$$\mu_0 = \bigotimes_{n=1}^{\infty} \lambda_n.$$

For arbitrary $k \in \mathbb{N}$ we again take the k -fold Cartesian product of E_0 and define E as

$$E(\mathcal{V}) = E_0(\mathcal{V})^k = \left\{ L = (l_1^*, \dots, l_k^*) \mid l_j = \sum_{n=1}^{\infty} n^{-2} \phi_{j,n}, \phi_{j,n} \in S_n \right\}. \quad (2.15)$$

Again by Tychonoff's theorem E is also compact and we define a probability measure μ as the k -fold product measure

$$\mu = \bigotimes_{j=1}^k \mu_0. \quad (2.16)$$

Rephrasing [Rob10, Lemma 5.6] we can bound the measure μ as follows.

Lemma 2.15 (cf. [Rob10, Lemma 5.6]). *For any $n \in \mathbb{N}$, $\alpha \in \mathbb{R}^k$ and $x \in H$ we have*

$$\mu \{ L \in E \mid |\alpha + Lx| \leq \varepsilon \} \leq c \left(n^2 d_n^{1/2} \frac{\varepsilon}{\|P_n x\|} \right)^k, \quad (2.17)$$

where $c > 0$ is a constant independent of n and α , and P_n is the orthogonal projection onto the d_n -dimensional subspace $V_n \subseteq H$.

In the proof of the finite-dimensional embedding result we applied the key inequality (2.12) with $\varepsilon \sim 2^{-n}$ and $|x| \geq 2^{-n\theta}$ (see (2.14)). In order to make a similar argument using the estimate made in Lemma 2.15, we have to pay special attention to the choice of the subspaces V_n and their dimensions d_n . Ideally, for $\|x\| \geq 2^{-n\theta}$ we would like to have $\|P_n x\| \geq c 2^{-n\theta}$, where $c > 0$ is some constant. This is indeed possible if V_n is sufficiently close to the set $\mathcal{A} \subseteq H$. Suppose $\text{dist}(\mathcal{A}, V_n) \leq \frac{2^{-n\theta}}{3}$ and x is of the form $x = x_j - y_j$ with $\|x_j - y_j\| \geq 2^{-n\theta}$ (as in the proof), then we have

$$\|P_n x\| = \|P_n(x_j - y_j)\| \geq \|x_j - y_j\| - \|x_j - P_n x_j\| - \|y_j - P_n y_j\| \geq \frac{2^{-n\theta}}{3}. \quad (2.18)$$

Since the dimension d_n also occurs in the estimate (2.17), we need to control how the dimension d_n of V_n grows with n , or in other words, as $\text{dist}(\mathcal{A}, V_n)$ decreases. This can be done via the *thickness exponent* also introduced in [HK99].

Definition 2.16 (Thickness exponent, [HK99]). Let \mathcal{X} be a Banach space and $\mathcal{A} \subseteq \mathcal{X}$. For $\varepsilon > 0$, denote by $d(\mathcal{A}, \varepsilon)$ the minimum dimension of all finite-dimensional subspaces $V \subseteq \mathcal{X}$ such that $\text{dist}(\mathcal{A}, V) \leq \varepsilon$. The *thickness exponent* of \mathcal{A} is defined by

$$\tau(\mathcal{A}) = \limsup_{\varepsilon \rightarrow 0} \frac{\log d(\mathcal{A}, \varepsilon)}{-\log \varepsilon},$$

essentially $d(\mathcal{A}, \varepsilon) \sim \varepsilon^{-\tau(\mathcal{A})}$ as $\varepsilon \rightarrow 0$.

Roughly speaking, $\tau(\mathcal{A})$ is a measure for how well \mathcal{A} can be approximated by finite-dimensional subspaces of the ambient space \mathcal{X} . In particular, if $\tau > \tau(\mathcal{A})$, then we can find $\varepsilon_0 > 0$ and finite-dimensional spaces $V_n \subseteq \mathcal{X}$ with

$$\dim(V_n) \leq C_{\mathcal{A},\tau} \cdot \varepsilon^{-\tau} < \infty \text{ and } \text{dist}(\mathcal{A}, V_n) \leq \varepsilon \quad \text{for all } \varepsilon < \varepsilon_0. \quad (2.19)$$

Thus, the thickness exponent provides the existence of subspaces that can be used in the construction of a transverse measure μ as described above.

In order to further understand this notion, we list some key properties of it.

Proposition 2.17. *Let \mathcal{X} be a Banach space and $\mathcal{A} \subseteq \mathcal{X}$.*

- (a) *The thickness exponent of \mathcal{A} is bounded by the box-counting dimension $d_{\text{box}}(\mathcal{A})$, i.e., $\tau(\mathcal{A}) \leq d_{\text{box}}(\mathcal{A})$.*
- (b) *If $\mathcal{A} \subseteq \mathcal{M}$ for a finite-dimensional C^s -manifold, then $\tau(\mathcal{A}) \leq \frac{d_{\text{box}}(\mathcal{A})}{s}$.*
- (c) *Suppose $\mathcal{X} = L^2(\Omega)$ for some minimally smooth bounded region $\Omega \subseteq \mathbb{R}^m$ (see [EE18] for a definition of minimally smooth) and \mathcal{A} is compact. If \mathcal{A} is uniformly bounded in the sobolev space $H^s(\Omega)$ for some $s > 0$, then $\tau(\mathcal{A}) \leq \frac{m}{s}$. In particular, if \mathcal{A} consists of “smooth functions”, i.e., \mathcal{A} is uniformly bounded in $H^s(\Omega)$ for every $s \in \mathbb{N}$, then $\tau(\mathcal{A}) = 0$.*

Proof.

- (a) This follows directly from the definition. Suppose \mathcal{A} is covered by $N(\mathcal{A}, \varepsilon)$ balls of radius $\varepsilon > 0$. Then \mathcal{A} lies within ε distance of the space spanned by the centers of these balls. Thus, $d(\mathcal{A}, \varepsilon) \leq N(\mathcal{A}, \varepsilon)$ which completes the proof.
- (b) As in [HK99] we only sketch the proof. Let $\varepsilon > 0$ and cover \mathcal{A} with the minimal number $N(\mathcal{A}, \varepsilon)$ of balls of radius ε that is needed. Then one can approximate the manifold \mathcal{M} within each ball to within ε^r distance by a Taylor polynomial. It follows that $d(\mathcal{A}, \varepsilon^s) \leq C_{s, \dim(\mathcal{M})} \cdot N(\mathcal{A}, \varepsilon)$ as desired. Here $C_{s, \dim(\mathcal{M})}$ is the number of terms in the Taylor polynomial that only depends on s and the dimension of \mathcal{M} .
- (c) This statement is the main theorem in [FR99] and we refer to that article for the corresponding proof. However, as pointed out in [Rob05] there is a small gap in the proof that is fixed in the footnote of [Rob05, Proposition 3.2].

□

We remark that part (b) and (c) of the above proposition roughly says that $\tau(\mathcal{A})$ is inversely proportional to “smoothness” of \mathcal{A} . Hence, in our later numerical studies it is reasonable to assume that \mathcal{A} is smooth, i.e., $\tau(\mathcal{A}) = 0$.

The notion of the thickness of \mathcal{A} allows us to state and prove the embedding result by Hunt and Kaloshin for the case where \mathcal{A} is a subset of a possible infinite-dimensional Hilbert space.

Theorem 2.18 ([HK99, Theorem 3.6]). *Let $\mathcal{A} \subseteq H$ be a compact set with box-counting dimension $d_{\text{box}}(\mathcal{A}) = d$ and thickness exponent $\tau(\mathcal{A}) = \tau$. Let $k > 2d$ be an integer and $\theta \in \mathbb{R}$ with*

$$0 < \theta < \frac{k - 2d}{k(1 + \tau/2)}. \quad (2.20)$$

Then for almost every (in the sense of prevalence) Lipschitz function $f : H \rightarrow \mathbb{R}^k$ there exists $C > 0$ such that

$$\|x - y\| \leq C |f(x) - f(y)|^\theta \quad \text{for all } x, y \in \mathcal{A}. \quad (2.21)$$

In particular, f is one-to-one on \mathcal{A} with θ -Hölder continuous inverse on \mathcal{A} .

Note that for $\tau = 0$ the bound for the Hölder exponent θ is the same as in the finite-dimensional embedding [Theorem 2.10](#). In particular, for $k \rightarrow \infty$ the (sharp) upper bound $\frac{2}{2+\tau}$ becomes one in this case. But in general the thickness exponent further controls the smoothness of the embedding.

Proof. To incorporate the ideas of [Section 2.2.2](#) we give an in-depth proof of this result. At first let us define the corresponding transverse measure μ to test for prevalence. Choose and fix $\tau > \tau(\mathcal{A})$ such that still

$$0 < \theta < \frac{k - 2d}{k(1 + \tau/2)}. \quad (2.22)$$

Then by definition of the thickness exponent it follows that there are finite-dimensional subspaces $V_n \subseteq H$ such that

$$d_n := \dim(V_n) \leq C_{\mathcal{A},\tau} \cdot 2^{n\theta\tau} < \infty \text{ and } \text{dist}(\mathcal{A}, V_n) \leq \frac{2^{-n\theta}}{3}. \quad (2.23)$$

That is, we consider $\varepsilon = \frac{2^{-n\theta}}{3}$ in [\(2.19\)](#) and merge the arising additional factor 3^τ in the constant $C_{\mathcal{A},\tau}$. Let $\mathcal{V} = \{V_n\}_{n=1}^\infty$, $E = E(\mathcal{V})$ and μ be the probability measure on E defined in [\(2.15\)](#) and [\(2.16\)](#).

With that at hand the remainder of the proof follows closely to the proof of [Theorem 2.10](#). Thus, we recall the definitions and results made there. Let $f : H \rightarrow \mathbb{R}^k$ be a Lipschitz function. Define

$$Z_n = \{ (x, y) \in \mathcal{A} \times \mathcal{A} \mid \|x - y\| \geq 2^{-n\theta} \} \subseteq \mathcal{A} \times \mathcal{A}$$

and set

$$Q_n = \{ L \in E \mid |(f + L)(x) - (f + L)(y)| \leq 2^{-n} \text{ for some } (x, y) \in Z_n \}.$$

Again, for a fixed $d > d_{\text{box}}(\mathcal{A})$, we can cover Z_n by no more than $C_{\mathcal{A},d} \cdot 2^{2nd}$ balls $B((x_j, y_j), 2^{-n})$, where we use the 1-metric on the product space $H \times H$.

Now $\|L\|$ is uniformly bounded by

$$\|L\|^2 \leq \sum_{j=1}^k |l_j^*|^2 = \sum_{j=1}^k |l_j|^2,$$

since the map $l \mapsto l^*$ is an isometry, and

$$|l_j|^2 = \left| \sum_{n=1}^{\infty} n^{-2} \phi_{j,n} \right| \leq \sum_{n=1}^{\infty} n^{-4} < \infty.$$

Thus, there is a Lipschitz constant $M > 0$ that is valid for $f + L$ for all $L \in E$. Applying the same arguments as in the previous proof it follows

$$\begin{aligned} & \mu \{ L \in E \mid |(f + L)(x) - (f + L)(y)| \leq 2^{-n} \text{ for some } (x, y) \in Y_j \} \\ & \leq \mu \{ L \in E \mid |(f + L)(x_j) - (f + L)(y_j)| \leq C_f \cdot 2^{-n} \} \\ & = \mu \{ L \in E \mid |(f(x_j) - f(y_j)) + L(x_j - y_j)| \leq C_f \cdot 2^{-n} \}. \end{aligned}$$

Now Lemma 2.15 allows us to further bound the measure, i.e.,

$$\begin{aligned} & \mu \{ L \in E \mid |(f(x_j) - f(y_j)) + L(x_j - y_j)| \leq C_f \cdot 2^{-n} \} \\ & \leq c \left(n^2 d_n^{1/2} \frac{C_f \cdot 2^{-n}}{\|P_n(x_j - y_j)\|} \right)^k \end{aligned}$$

and by (2.23) and (2.18) we have

$$\begin{aligned} & c \left(n^2 d_n^{1/2} \frac{C_f \cdot 2^{-n}}{\|P_n(x_j - y_j)\|} \right)^k \\ & \leq c \left(n^2 \cdot C_{\mathcal{A},\tau}^{1/2} \cdot 2^{n\theta\tau/2} \cdot C_f \cdot 2^{-n} \cdot 2^{n\theta} \cdot 3 \right)^k \\ & = C_{\mathcal{A},\tau,f,k} \cdot n^{2k} \cdot 2^{[\theta k(1+\tau/2)-k]n}. \end{aligned}$$

Thus, we can estimate the measure of Q_n by

$$\mu(Q_n) \leq C_{\mathcal{A},d} \cdot 2^{2nd} \cdot C_{\mathcal{A},\tau,f,k} \cdot n^{2k} \cdot 2^{[\theta k(1+\tau/2)-k]n} = C_{\mathcal{A},d,\tau,f,k} \cdot n^{2k} \cdot 2^{[2d-k+\theta k(1+\tau/2)]n}$$

which yields a converging series $\sum_{n=1}^{\infty} \mu(Q_n) < \infty$ by (2.22). Finally, the rest of the proof follows as in Theorem 2.10 using the Borel-Cantelli lemma. \square

Banach Spaces

The authors of [HK99] tried to extend their result for subsets of a Banach space \mathcal{X} . However, in their proof it is claimed that there exists a linear isometry from the dual of any finite-dimensional subspace of Y to a subspace of the dual of Y . This is only true if (and only if by Riesz–Fréchet representation theorem) Y is a Hilbert space [Kak40, KM44]. Thus, we can not directly extend the construction for Hilbert spaces, where we identified elements $\phi_n \in S_n$ via the Riesz’ mapping $x \mapsto \langle \cdot, x \rangle$, to Banach spaces. That is why we will directly start with a sequence of subspaces in the dual space \mathcal{X}^* .

So let $\mathcal{V} = \{V_n\}_{n=1}^\infty$ be a sequence of subspaces $V_n \subseteq \mathcal{X}^*$ with dimension $d_n < \infty$ and let S_n be the unit ball in V_n . In analogy to the Hilbert space case we define the (compact) probe space E_0 for $k = 1$ by

$$E_0(\mathcal{V}) = \left\{ \sum_{n=1}^{\infty} n^{-2} \phi_n \mid \phi_n \in S_n \right\} \subseteq \mathcal{X}^*.$$

Now we choose a basis for V_n , so that we can identify S_n with a symmetric convex set $U_n \subseteq \mathbb{R}^{d_n}$ via the coordinate representation. Again, the uniform probability measure on U_n induces a measure λ_n on S_n and we obtain a probability measure μ_0 on E_0 by

$$\mu_0 = \bigotimes_{n=1}^{\infty} \lambda_n.$$

For arbitrary $k \in \mathbb{N}$ we again take the k -fold Cartesian product of E_0 and define E as

$$E(\mathcal{V}) = \left\{ L = (L_1, \dots, L_k) \mid L_j = \sum_{n=1}^{\infty} n^{-2} \phi_{j,n}, \phi_{j,n} \in S_n \right\}. \quad (2.24)$$

The corresponding probability measure μ is then the k -fold product measure

$$\mu = \bigotimes_{j=1}^k \mu_0. \quad (2.25)$$

This is, indeed, an extension since the V_n of a Hilbert space H are simply replaced by the subspaces V_n^* obtained by the isometry $x \mapsto \langle \cdot, x \rangle$, the unit ball S_n in V_n corresponds to the unit ball in V_n under the same mapping and $U_n = B_{d_n}$.

This measure can also be estimated as follows by rephrasing [Rob10, Lemma 5.10].

Lemma 2.19 (cf. [Rob10, Lemma 5.10]). *If $\alpha \in \mathbb{R}^k$ and $x \in \mathcal{X}$, then for every $n \in \mathbb{N}$ we have*

$$\mu \{ L \in E \mid |\alpha + Lx| < \varepsilon \} \leq \left(n^2 d_n \frac{\varepsilon}{|g(x)|} \right)^k$$

for any $g \in S_n$.

The last ingredient needed is the definition of an Auerbach basis for a finite-dimensional Banach space.

Definition 2.20. Let W be a d -dimensional Banach space. An *Auerbach basis* for W is formed by a basis $\{w_1, \dots, w_d\}$ of W and a corresponding basis $\{w_1^*, \dots, w_d^*\}$ of the dual space W^* that satisfies $\|w_i\| = \|w_i^*\|$ and $w_i^*(w_j) = \delta_{i,j}$ for all $i, j = 1, \dots, d$. For the existence of such a basis see, for instance, [Wer11, Satz II.2.6].

With that at hand we are in the position to formulate the embedding result for a subset of a Banach space.

Theorem 2.21 (cf. [MR19, Theorem 3.2]). *Let $\mathcal{A} \subseteq \mathcal{X}$ be a compact set with box-counting dimension $d_{\text{box}}(\mathcal{A}) = d$ and thickness exponent $\tau(\mathcal{A}) = \tau < 1$. Let $k > 2d$ be an integer and $\theta \in \mathbb{R}$ with*

$$0 < \theta < (1 - \tau) \frac{k - 2d}{k(1 + \tau)}. \quad (2.26)$$

Then for almost every (in the sense of prevalence) Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}^k$ there exists $C > 0$ such that

$$\|x - y\| \leq C |f(x) - f(y)|^\theta \quad \text{for all } x, y \in \mathcal{A}. \quad (2.27)$$

In particular, f is one-to-one on \mathcal{A} with θ -Hölder continuous inverse on \mathcal{A} .

Note that in contrast to Theorem 2.18, for $k \rightarrow \infty$ the bound of the thickness exponent is in (2.26) given by $\frac{1-\tau}{1+\tau}$ which is less than or equal to the bound $\frac{2}{2+\tau}$ (cf. (2.21)), allowing for less smooth embeddings. But for zero thickness we still obtain a maximal bound of one.

Proof. We will proceed as in the proof of Theorem 2.18 and also reuse some arguments. Take $1 > \tau > \tau(\mathcal{A})$ and $d > d_{\text{box}}(\mathcal{A})$ such that still

$$0 < \theta < (1 - \tau) \frac{k - 2d}{k(1 + \tau)}. \quad (2.28)$$

By definition of the thickness exponent there are finite-dimensional subspaces $W_n \subseteq \mathcal{X}$ such that

$$\dim(W_n) = d_n \leq C_{\mathcal{A}, \tau, \beta} \cdot 2^{n\theta\beta\tau} < \infty \text{ and } \text{dist}(\mathcal{A}, W_n) \leq \frac{2^{-n\theta\beta}}{3}, \quad (2.29)$$

where $\beta > 1$ will be specifically chosen later. In order to construct the sequence of subspaces V_n of \mathcal{X}^* , we will consider an Auerbach basis for W_n . Let $\{e_1^n, \dots, e_{d_n}^n\}$ be a basis for W_n and $\{r_1^n, \dots, r_{d_n}^n\}$ be the corresponding basis for W_n^* which satisfies

$$\|r_i^n\| = 1 \quad \forall i \text{ and } r_i^n(e_j^n) = \delta_{i,j}, \quad \forall i \neq j.$$

Applying the Hahn-Banach theorem (see, e.g., [Wer11, Theorem III.1.5]) we extend the elements of $\{r_1^n, \dots, r_{d_n}^n\}$ to a set $\{f_1^n, \dots, f_{d_n}^n\} \subseteq \mathcal{X}^*$ and set $V_n = \text{span}\{f_1^n, \dots, f_{d_n}^n\}$,

which is an at most d_n -dimensional subspace of \mathcal{X}^* . Let $\mathcal{V} = \{V_n\}_{n=1}^\infty$, $E = E(\mathcal{V})$ and μ be the probability measure on E defined in (2.24) and (2.25).

Refer to the proofs of [Theorem 2.10](#) and [Theorem 2.18](#) for the definitions of Z_n, Q_n , the covering of Z_n and Y_j . Again $\|L\|$ is uniformly bounded and we obtain

$$\begin{aligned} & \mu \left\{ L \in E \mid |(f+L)(x) - (f+L)(y)| \leq 2^{-n} \text{ for some } (x, y) \in Y_j \right\} \\ & \leq \mu \left\{ L \in E \mid |(f(x_j) - f(y_j)) + L(x_j - y_j)| \leq C_f \cdot 2^{-n} \right\} \end{aligned}$$

Now [Lemma 2.19](#) allows us to further bound the measure, i.e.,

$$\begin{aligned} & \mu \left\{ L \in E \mid |(f(x_j) - f(y_j)) + L(x_j - y_j)| \leq C_f \cdot 2^{-n} \right\} \\ & \leq \left(n^2 d_n \frac{C_f \cdot 2^{-n}}{|g_n(x_j - y_j)|} \right)^k \end{aligned}$$

for all $g_n \in S_n$. For the explicit choice of g_n we refer to [\[MR19\]](#) and only sketch the ideas. Consider $(x^n, y^n) \in W_n \times W_n$ such that $\|(x^n, y^n) - (x_j, y_j)\|_1 \leq \frac{2^{n\theta\beta}}{3}$. Setting $\beta = \frac{1}{1-\tau}$ one finds $g_n \in S_n$ such that

$$|g_n(x^n - y^n)| \geq d_n^{-1} \|x^n - y^n\| \quad \text{and} \quad |g_n(x_j - y_j)| \geq C \cdot 2^{-n\theta\beta}$$

with some constant $C > 0$. Thus, (2.29) and [Lemma 2.19](#) yield

$$\begin{aligned} \mu(Q_n) & \leq C_{\mathcal{A},\tau} \cdot 2^{2nd} \cdot \left(n^2 d_n \frac{C_f \cdot 2^{-n}}{C \cdot 2^{-n\theta\beta}} \right)^k \\ & \leq C_{\mathcal{A},\tau,f,k} \cdot n^{2k} \cdot 2^{[2d-k+\theta\beta\tau k+\theta\beta k]n}. \end{aligned}$$

Finally, by (2.28) the exponent $2d - k + \theta\beta k + \theta\beta\tau k = 2d - k + \theta\beta k(\tau + 1) < 0$ is negative and the series $\sum_{n=1}^\infty \mu(Q_n)$ converges as desired. Using the Borel-Cantelli lemma as in [Theorem 2.10](#) finishes the proof. \square

Remark 2.22.

- (a) [Theorem 2.21](#) was originally proven in [\[MR19\]](#) for μ -almost every linear map. However, it turns out that slightly adjusting the proof allows us to extend the result to incorporate the notion of prevalence.
- (b) Based on the approximation required in the course of the arguments in [Theorem 2.18](#), the *dual thickness exponent* $\tau^*(\mathcal{A})$ was introduced in [\[Rob09\]](#). Its definition allows for the same result as in [Theorem 2.18](#), where $\tau/2$ is replaced by τ^* . Unfortunately, there is no known general relationship between both exponents in the context of Banach spaces, yet. However, one can prove that zero thickness implies zero dual thickness [\[Rob10\]](#), which fits perfectly for our standing assumption $\tau(\mathcal{A}) = 0$ in the upcoming numerical analysis. In addition to that, τ^* is bounded by twice the box-counting dimension of \mathcal{A} [\[MR19\]](#) (cf. [Proposition 2.17](#)).

2.3 Delay Coordinate Embeddings

In the previous sections we considered finite-dimensional compact sets in an (possibly infinite-dimensional) ambient space. We did not take into account that this set is invariant under the dynamics of an underlying dynamical system. That is why we will now summarize some embedding results in which the dynamics is *observed* by a real-valued map using delay coordinates.

As before we start with the case where the set is a smooth manifold. Based on the results of Whitney (see Section 2.2.1), Floris Takens was able to prove that delay coordinates of a generic discrete dynamical system Φ on \mathcal{M} can also be used to generate an embedding [Tak81]. Here, we will fix Φ which entails some additional assumptions.

Theorem 2.23 (Takens' (unstated) Theorem, [Huk06]). *Let \mathcal{M} be a compact smooth manifold of dimension $d \in \mathbb{N}$. Let $\Phi : \mathcal{M} \rightarrow \mathcal{M}$ be a diffeomorphism with the following properties:*

- 1) Φ has finitely many periodic points of period less or equal to $2d$.
- 2) If $x \in \mathcal{M}$ is a periodic point with period $p \leq 2d$, then the eigenvalues of the linearization $D\Phi^p(x)$ are distinct.

Then for generic $f \in C^2(\mathcal{M}, \mathbb{R})$ the delay coordinate map $D_{2d+1}[f, \Phi] : \mathcal{M} \rightarrow \mathbb{R}^{2d+1}$ defined by

$$D_{2d+1}[f, \Phi](x) = (f(x), f(\Phi(x)), \dots, f(\Phi^{2d}(x)))$$

is an embedding.

Here, generic means “open and dense” in the C^1 -topology, even though f is C^2 . Sometimes the *delay coordinate map* $D_k[f, \Phi]$ is also called *k-fold observation map*. Condition 1) is responsible for the fact that $D_{2d+1}[f, \Phi]$ is injective whereas condition 2) is a needed property to show that it is, indeed, an embedding. By replacing genericity with prevalence Sauer et al. [SYC91] proved an analogous result for sets in \mathbb{R}^N with finite box-counting dimension. In the following we will denote the set of p -periodic points of a map Φ by $A_p(\Phi)$, that is, $A_p(\Phi) = \{x \in A \mid \Phi^p(x) = x\}$.

Theorem 2.24 ([SYC91, Theorem 2.7.]). *Let $A \subseteq \mathbb{R}^N$ be compact with box-counting dimension $d_{\text{box}}(A) = d$. Let $k > 2d$ be an integer and let $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^N$ be a diffeomorphism with the following properties:*

- 1) $d_{\text{box}}(A_p(\Phi)) < p/2$ for all $p \leq k$.
- 2) If $x \in A$ is a periodic point with period $p \leq k$, then the eigenvalues of the linearization $D\Phi^p(x)$ are distinct.

Then for almost every (in the sense of prevalence) smooth map $f : \mathbb{R}^N \rightarrow \mathbb{R}$ the delay coordinate map $D_k[f, \Phi] : \mathbb{R}^N \rightarrow \mathbb{R}^k$ is one-to-one on $A \subseteq \mathbb{R}^N$ and an immersion on each compact subset of a smooth manifold contained in A .

Observe that the first assumption in this theorem is clearly satisfied if Φ has finitely many periodic points of period $p \leq k$ since the box-counting dimension is zero in this case. Thus, the assumptions made here are weaker compared to the ones in Theorem 2.23. We note that condition 2) is only needed for proving that $D_k[f, \Phi]$ is an immersion.

Finally, Robinson [Rob05] combined this result with the Hilbert space embedding [Theorem 2.18](#). That is, first $\mathcal{A} \subseteq H$ is embedded by a map L via [Theorem 2.18](#) in \mathbb{R}^N for a large enough $N \in \mathbb{N}$ and then a ‘‘Hölder version’’ of [Theorem 2.24](#) provides a delay coordinate embedding of $L(\mathcal{A}) \subseteq \mathbb{R}^N$ using the composition $L \circ \Phi \circ L^{-1}$.

Theorem 2.25 ([Rob05]). *Let $\Phi : H \rightarrow H$ be a Lipschitz map, where H is a Hilbert space. Let $\mathcal{A} \subseteq H$ be a Φ -invariant compact subset with box-counting dimension $d_{\text{box}}(\mathcal{A}) = d$ and thickness exponent $\tau(\mathcal{A}) = \tau$. Let $k > 2(1 + \tau/2)d$ be an integer and further suppose $d_{\text{box}}(\mathcal{A}_p(\Phi)) < \frac{p}{2+\tau}$ for all $p \leq k$. Then for almost every (in the sense of prevalence) Lipschitz map $f : H \rightarrow \mathbb{R}$ the delay coordinate map $D_k[f, \Phi] : H \rightarrow \mathbb{R}^k$ is one-to-one on $\mathcal{A} \subseteq H$.*

Again the condition on $d_{\text{box}}(\mathcal{A}_p(\Phi))$ is satisfied if Φ has finitely many periodic points of period $p \leq k$. Observe that Φ is not differentiable and therefore the immersion property of $D_k[f, \Phi]$ provided by condition 2) in [Theorem 2.24](#) is not valid. Even considering a smoother dynamical system Φ would not yield an analogous result since the delay coordinate map defined on \mathbb{R}^N is ‘‘lifted’’ to a map on \mathcal{A} via a map that is only Hölder continuous (cf. (2.21)).

If the ambient space is a Banach space, it is straightforward to show that the above theorem is also valid provided $\tau(\mathcal{A}) < 1$, using [Theorem 2.21](#) instead of [Theorem 2.18](#). Note that in this case the embedding dimension and the bound on the box-counting dimension of the periodic points have to be adapted as well.

Theorem 2.26. *Let $\Phi : \mathcal{X} \rightarrow \mathcal{X}$ be a Lipschitz map, where \mathcal{X} is a Banach space. Let $\mathcal{A} \subseteq \mathcal{X}$ be a Φ -invariant compact subset with box-counting dimension $d_{\text{box}}(\mathcal{A}) = d$ and thickness exponent $\tau(\mathcal{A}) = \tau < 1$. Let $k > 2\frac{1+\tau}{1-\tau}d$ be an integer and further suppose $d_{\text{box}}(\mathcal{A}_p(\Phi)) < \frac{p(1-\tau)}{2(1+\tau)}$ for all $p \leq k$. Then for almost every (in the sense of prevalence) Lipschitz map $f : \mathcal{X} \rightarrow \mathbb{R}$ the delay coordinate map $D_k[f, \Phi] : \mathcal{X} \rightarrow \mathbb{R}^k$ is one-to-one on $\mathcal{A} \subseteq \mathcal{X}$.*

2.4 The Core Dynamical System (CDS)

Given a discrete dynamical system of the form

$$u_{j+1} = \Phi(u_j), \quad j = 0, 1, \dots, \quad (2.30)$$

where $\Phi : \mathcal{X} \rightarrow \mathcal{X}$ is a Lipschitz continuous map on some (possibly) infinite-dimensional Banach space \mathcal{X} , we will employ the embedding results in order to construct the so-called *core dynamical system* (CDS). To this end, we assume that Φ has an invariant compact set $\mathcal{A} \subseteq \mathcal{X}$ with finite box-counting dimension $d_{\text{box}}(\mathcal{A})$ and thickness exponent $\tau(\mathcal{A})$. A typical example for such a map Φ is the time- T -map of a partial differential equation (cf. (2.4)) or a delay differential equation. In this setting the set \mathcal{A} could be the global attractor of the system.

Originally defined in [DHZ16] the CDS is a finite-dimensional continuous dynamical system φ that possesses an invariant set on which the dynamics are topologically conjugated to those of Φ on \mathcal{A} . Thus, the CDS allows for the (numerical) analysis of the infinite-dimensional system Φ . More precisely, the CDS is constructed utilizing an embedding R (cf. Chapter 2). In particular, every such embedding R generates a different CDS. We will include this observation and prove that a prevalent set of Lipschitz maps generates a Hölder continuous map φ that essentially has the same dynamics as Φ on the attractor. Also note that in contrast to [DHZ16] the dynamical system φ is not only continuous but even Hölder. In order to use the embedding Theorem 2.21, we additionally have to assume that $\tau(\mathcal{A}) < 1$. Choose and fix an integer $k \in \mathbb{N}$ with $k > 2d_{\text{box}}(\mathcal{A})$ such that the corresponding result is valid. Then there is a prevalent set of Lipschitz maps $R : \mathcal{X} \rightarrow \mathbb{R}^k$ that are one-to-one on \mathcal{A} . In the following we will call such a map R an *observation map*. Denote by A_k the image of $\mathcal{A} \subseteq \mathcal{X}$ under a chosen observation map R , that is,

$$A_k = R(\mathcal{A}) \subseteq \mathbb{R}^k.$$

By Theorem 2.21 this observation map R is invertible as a mapping from \mathcal{A} to A_k which implies the existence of a unique map $\tilde{E} : A_k \rightarrow \mathcal{X}$ satisfying

$$(\tilde{E} \circ R)(u) = u \quad \forall u \in \mathcal{A} \quad \text{and} \quad (R \circ \tilde{E})(x) = x \quad \forall x \in A_k. \quad (2.31)$$

In addition to that, for some $\theta \in \mathbb{R}$ such that (2.26) holds, there is a constant $C > 0$ such that

$$\|u - v\| \leq C |R(u) - R(v)|^\theta \quad \text{for all } u, v \in \mathcal{A}.$$

In particular, \tilde{E} is Hölder continuous with Hölder exponent θ . On $A_k \subseteq \mathbb{R}^k$ we define the CDS φ as the composition of the three maps R , Φ and \tilde{E} , that is,

$$\varphi = R \circ \Phi \circ \tilde{E}.$$

By construction φ is θ -Hölder continuous and $A_k \subseteq \mathbb{R}^k$ is an invariant set of φ . In order to define $\varphi(x)$ for $x \in \mathbb{R}^k$, it remains to extend the map \tilde{E} to a map $E : \mathbb{R}^k \rightarrow \mathcal{X}$ with $E|_{A_k} = \tilde{E}$. To this end, we employ the following extension theorem of Whitney type by Stein [Ste70].

Theorem 2.27 ([MS03, Theorem 2.2]). *Let $A \subseteq \mathbb{R}^k$ be a non-empty set, \mathcal{X} be a real Banach space and $g : A \rightarrow \mathcal{X}$ be a Hölder continuous map with exponent $\theta \in (0, 1]$. Then there exists an extension $\bar{g} : \mathbb{R}^k \rightarrow \mathcal{X}$ of g that is Hölder continuous with the same exponent.*

Utilizing this theorem we can extend \tilde{E} to a map $E : \mathbb{R}^k \rightarrow \mathcal{X}$ that is still θ -Hölder continuous and, in the end, the CDS is defined on the *observation space* \mathbb{R}^k by

$$x_{j+1} = \varphi(x_j) = (R \circ \Phi \circ E)(x_j), \quad j = 0, 1, \dots \quad (2.32)$$

In Figure 2.1 we illustrate the construction of the CDS.

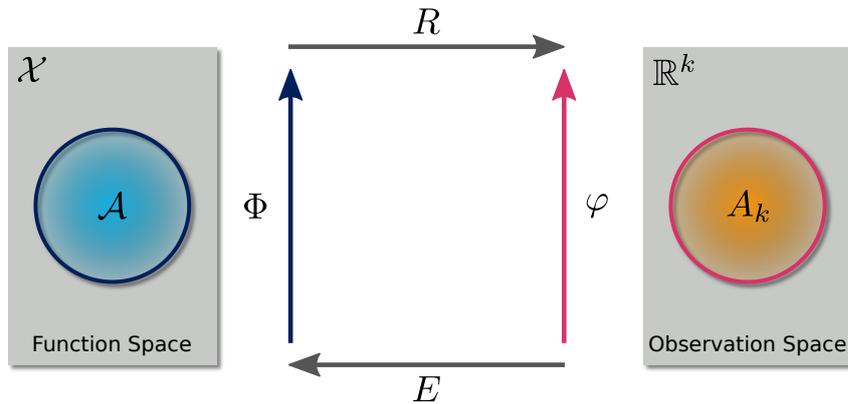


Figure 2.1: Definition of the CDS φ .

Note that after extending \tilde{E} the set A_k remains invariant under φ , i.e., $\varphi(A_k) = A_k$. Finally, we summarize the arguments made above in the following proposition.

Proposition 2.28. *Let \mathcal{X} be a Banach space and $\Phi : \mathcal{X} \rightarrow \mathcal{X}$ be a Lipschitz map that has an invariant compact set $\mathcal{A} \subseteq \mathcal{X}$ with finite box-counting dimension $d_{\text{box}}(\mathcal{A}) = d$ and thickness exponent $\tau(\mathcal{A}) = \tau < 1$. Let $k > 2d$ be an integer and $\theta \in \mathbb{R}$ such that (2.26) holds. Then for almost every (in the sense of prevalence) Lipschitz function $R : \mathcal{X} \rightarrow \mathbb{R}^k$ there is a θ -Hölder continuous map $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ satisfying*

$$\varphi(R(u)) = R(\Phi(u)) \text{ for all } u \in \mathcal{A}.$$

Remark 2.29.

- (a) In [DHZ16] the map \tilde{E} is only continuously extended using a generalization of Tietze's extension theorem [DS88, I.5.3] proven by Dugundji [Dug51, Theorem 4.1] which leads to a continuous core dynamical system φ .
- (b) If \mathcal{X} is in fact a Hilbert space, then the assumption $\tau(\mathcal{A}) < 1$ can be dropped and $\theta \in \mathbb{R}$ has to satisfy (2.20) instead. In this case Theorem 2.18 provides the existence of the prevalent set of Lipschitz functions generating φ .
- (c) By adjusting the assumptions on Φ and the embedding dimension $k \in \mathbb{N}$ one can also utilize the delay embedding Theorem 2.26 or 2.25 for the construction of the CDS. In this case the term prevalence is with respect to the space of Lipschitz maps $\mathcal{X} \rightarrow \mathbb{R}$.

- (d) A crucial drawback in the arguments made above is that the map E is unknown. Even though the extension procedure E of \tilde{E} by [Theorem 2.27](#) is partly constructive, the embedding result only guarantees the existence of \tilde{E} . That is why the particular realization of E will depend on the application at hand and we will discuss possible choices for DDEs and PDEs in [Chapter 4](#).

In particular, the set of Lipschitz functions generating a CDS φ is non-empty and dense (see [Proposition 2.13](#)). After choosing a specific observation map R we obtain a finite-dimensional dynamical system φ whose dynamics on the *embedded* invariant set $A_k = R(\mathcal{A})$ is topologically conjugated to that of the infinite-dimensional system Φ on \mathcal{A} . Thus, we reduced the problem of computing \mathcal{A} (or subsets of it) to the approximation of corresponding *embedded* sets.

The Parameter-Dependent CDS

In this section we will introduce a one-dimensional real-valued parameter to the system (2.30). More precisely, we now consider a dynamical system of the form

$$u_{j+1} = \Phi(u_j, \lambda), \quad j = 0, 1, \dots, \quad (2.33)$$

where $\Phi : \mathcal{X} \times \Lambda \rightarrow \mathcal{X}$ is Lipschitz continuous in $u \in \mathcal{X}$ for each parameter value λ in a compact interval $\Lambda \subseteq \mathbb{R}$ and uniformly continuous in λ on bounded sets. For brevity, we will write Φ_λ for $\Phi(\cdot, \lambda)$. Moreover, we assume that Φ_λ has a compact invariant set $\mathcal{A}^\lambda \subseteq \mathcal{X}$ with finite box-counting dimension $d_{\text{box}}(\mathcal{A}^\lambda)$ and thickness exponent $\tau(\mathcal{A}^\lambda)$ for every $\lambda \in \Lambda$. The aim in this section is to construct corresponding *parameter-dependent core dynamical system* (pCDS) φ_λ that reconstructs the dynamics of Φ_λ on \mathcal{A}^λ using an observation map that is independent of the parameter λ . Note that this construction goes beyond the current definition in [\[DHZ16\]](#) since the same arguments as before can not be directly applied.

A straightforward approach for the definition of the pCDS would be to embed all the sets \mathcal{A}^λ simultaneously. More precisely, ideally we would like to have a prevalent set of Lipschitz maps $R : \mathcal{X} \rightarrow \mathbb{R}^k$ that are one-to-one on \mathcal{A}^λ for every $\lambda \in \Lambda$. In particular, such a map R is independent of λ and R would be one-to-one on the union

$$\mathcal{A}^\Lambda = \bigcup_{\lambda \in \Lambda} \mathcal{A}^\lambda.$$

However, this approach causes some technical issues that we will discuss in the following.

In order to prove the existence of such a prevalent set, a first idea would be to use the embedding [Theorem 2.18](#) or [2.21](#) on each \mathcal{A}^λ individually. In fact, this yields for every $\lambda \in \Lambda$ the existence of a prevalent set L_λ of Lipschitz maps $R_\lambda : \mathcal{X} \rightarrow \mathbb{R}^k$ that are one-to-one on \mathcal{A}^λ with θ -Hölder inverse. In general, this set L_λ depends on $\lambda \in \Lambda$, which is emphasized by the subscripts of L_λ and R_λ , respectively. Note that the embedding dimension $k \in \mathbb{N}$ and the Hölder exponent θ can be made independent of λ by choosing $k \in \mathbb{N}$ sufficiently large, i.e.,

$$2 \max_{\lambda \in \Lambda} d_{\text{box}}(\mathcal{A}^\lambda) < k < \infty.$$

Additionally, the Hölder exponent can be made independent of λ by choosing θ sufficiently small, i.e.,

$$\begin{aligned} \theta &< \frac{k - 2 \max_{\lambda \in \Lambda} d_{\text{box}}(\mathcal{A}^\lambda)}{k(1 + \max_{\lambda \in \Lambda} \tau(\mathcal{A}^\lambda)/2)}, & \text{if } \mathcal{X} \text{ is a Hilbert space,} \\ \theta &< \left(1 - \max_{\lambda \in \Lambda} \tau(\mathcal{A}^\lambda)\right) \frac{k - 2 \max_{\lambda \in \Lambda} d_{\text{box}}(\mathcal{A}^\lambda)}{k(1 + \max_{\lambda \in \Lambda} \tau(\mathcal{A}^\lambda))}, & \text{if } \mathcal{X} \text{ is a Banach space,} \end{aligned}$$

(cf. (2.21) and (2.27)). Then the intersection of all these sets L_λ , that is,

$$L_\Lambda = \bigcap_{\lambda \in \Lambda} L_\lambda$$

will be a set of Lipschitz maps $R : \mathcal{X} \rightarrow \mathbb{R}^k$ that are one-to-one on the union \mathcal{A}^Λ with θ -Hölder inverse. In particular, L_Λ is independent of λ as desired. Unfortunately, L_Λ is an *uncountable* intersection and hence not prevalent in general. Observe that directly embedding \mathcal{A}^Λ using Theorem 2.18 or 2.21 is not possible, since the *uncountable* union \mathcal{A}^Λ is, in general, not compact even though each individual \mathcal{A}^λ is.

In order to circumvent the problem that $\Lambda \subseteq \mathbb{R}$ is too “large”, we will consider a countable enumeration $\Lambda_{\mathbb{N}} = \{\lambda_n\}_{n \in \mathbb{N}}$ of $\mathbb{Q} \cap \Lambda$. Then the corresponding sets

$$L_{\Lambda_{\mathbb{N}}} = \bigcap_{n \in \mathbb{N}} L_{\lambda_n} \text{ and } \mathcal{A}^{\Lambda_{\mathbb{N}}} = \bigcup_{n \in \mathbb{N}} \mathcal{A}^{\lambda_n}$$

reduce to a countable intersection and a countable union, respectively. Note that in general $\mathcal{A}^{\Lambda_{\mathbb{N}}}$ still may be not compact and the embedding Theorem 2.18 or 2.21 can not be applied on $\mathcal{A}^{\Lambda_{\mathbb{N}}}$. To do so one has to consider a finite subset in parameter space. However, $L_{\Lambda_{\mathbb{N}}}$ will be a prevalent set by Proposition 2.13 (b) and thus almost every (in the sense of prevalence) Lipschitz map $R_{\Lambda_{\mathbb{N}}} : \mathcal{X} \rightarrow \mathbb{R}^k$ will be one-to-one on $\mathcal{A}^{\Lambda_{\mathbb{N}}} \subseteq \mathcal{X}$. In particular, there is a unique θ -Hölder continuous map $\tilde{E}_{\Lambda_{\mathbb{N}}} = R_{\Lambda_{\mathbb{N}}}^{-1}|_{\mathcal{A}^{\Lambda_{\mathbb{N}}}}$ which can be extended to a map $E_{\Lambda_{\mathbb{N}}} : \mathbb{R}^k \rightarrow \mathcal{X}$ by Theorem 2.27 such that

$$(E_{\Lambda_{\mathbb{N}}} \circ R_{\Lambda_{\mathbb{N}}})(u) = u \quad \forall u \in \mathcal{A}^{\Lambda_{\mathbb{N}}} \quad \text{and} \quad (R_{\Lambda_{\mathbb{N}}} \circ E_{\Lambda_{\mathbb{N}}})(x) = x \quad \forall x \in A_k^{\Lambda_{\mathbb{N}}},$$

where $A_k^{\Lambda_{\mathbb{N}}} = R_{\Lambda_{\mathbb{N}}}(\mathcal{A}^{\Lambda_{\mathbb{N}}})$. The *parameter-dependent core dynamical system* (pCDS) is then defined by

$$x_{j+1} = \varphi_{\lambda_n}(x_j) = (R_{\Lambda_{\mathbb{N}}} \circ \Phi_{\lambda_n} \circ E_{\Lambda_{\mathbb{N}}})(x_j), \quad j = 0, 1, \dots,$$

where the parameter value λ_n can be taken from the dense subset $\Lambda_{\mathbb{N}} \subseteq \Lambda$. By construction $A_k^{\lambda_n}$ is an invariant set of φ_{λ_n} and the dynamics of φ_{λ_n} on $A_k^{\lambda_n}$ are topologically conjugated to those of Φ_{λ_n} on \mathcal{A}^{λ_n} for every $\lambda_n \in \Lambda_{\mathbb{N}}$.

In the numerical realization of the pCDS we will design an observation map R that at most depends on a chosen fixed $\hat{\lambda}$. The corresponding map E is then constructed in such a way that $(R \circ E)(x) = x$ is in fact satisfied for every $x \in \mathbb{R}^k$. The second condition $(E \circ R)(u) = u$ is then enforced at least approximately by a bootstrapping

method. Therefore, we actually define the pCDS for every $\lambda \in \Lambda$ by

$$x_{j+1} = \varphi_\lambda(x_j) = (R \circ \Phi_\lambda \circ E)(x_j), \quad j = 0, 1, \dots \quad (2.34)$$

According to the above discussion, φ_λ reproduces the dynamics of Φ_λ on \mathcal{A}^λ for a dense subset $\Lambda_{\mathbb{N}} \subseteq \Lambda$. For the remaining parameter values the map R may not be one-to-one but close to it, provided the invariant set \mathcal{A}^λ behaves well under parameter perturbation (cf. Definition 3.15). With that in mind we will use φ_λ (2.34) for the numerical analysis of the dynamics of Φ_λ for every $\lambda \in \Lambda$.

Remark 2.30.

- (a) If there is a (large) compact set $\mathcal{A} \subseteq \mathcal{X}$ with finite box-counting dimension that contains \mathcal{A}^λ for all $\lambda \in \Lambda$, one can avoid the discretization of Λ and instead embed \mathcal{A} using Theorem 2.18 or 2.21.
- (b) Typically, for the numerical analysis of parameter-dependent systems such as Φ_λ , one discretizes the parameter space by some finite set $\Lambda_N = \{\lambda_0, \lambda_1, \dots, \lambda_N\} \subseteq \Lambda$ instead of choosing a countable subset and in the following we proceed in this way. However, choosing a dense subset of the parameter space Λ allows for a “smoother” (theoretical) definition of the pCDS, that is, it actually reconstructs the dynamics of Φ_λ for almost all (in the sense of denseness) parameter values.

3 Set-Oriented Numerics for Infinite-Dimensional Systems

In this chapter we present a family of tools, so-called *set-oriented methods*, for the approximation of invariant sets. The basic idea of such tools is to construct coverings of the object of interest, e.g., *attractors* and *(un-)stable manifolds* of a steady state, by outer approximations which are generated via subdivision and continuation techniques. Those methods have been successfully applied in several areas such as molecular dynamics [DDJS99, SHD01, DGM⁺05], astrophysics [DJK⁺05, DJL⁺05, DJ06] and ocean dynamics [FPET07, DFH⁺09, SFM10, FHR⁺12]. The purpose of this chapter is to extend the *subdivision scheme* [DH97] for the approximation of attractors and the *continuation method* [DH96] for the computation of invariant manifolds to infinite-dimensional systems. To this end, we will employ embedding techniques (cf. Sections 2.2.3 and 2.3) which allow the construction of the finite-dimensional CDS (see Section 2.4). Since the CDS is in general not a homeomorphism, the original subdivision scheme [DH97] has to be adapted to continuous dynamical systems which then enables us to approximate the *embedded attractor* [DHZ16]. Furthermore, we will present a path following method for the approximation of parameter-dependent attractors. That is, we will introduce a one-dimensional real parameter to the system and aim to compute the corresponding attractors by reusing previously computed coverings. In doing so, we will develop a set-oriented predictor-corrector technique. Finally, applying the subdivision scheme to the (embedded) local unstable manifold and using the continuation technique developed in [DH96] afterwards leads to the computation of *embedded unstable manifolds* [ZDG19]. We note that large parts of this chapter are also contained in [ZDG19], [GZED20], [ZGD20] and [GZ20] to which the author has made substantial contributions.

Throughout the remainder of this chapter we will consider a dynamical system of the form

$$x_{j+1} = \varphi(x_j), \quad j = 0, 1, \dots, \quad (3.1)$$

where $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is continuous. Moreover, we assume that φ has a compact invariant set $A_k \subseteq \mathbb{R}^k$. In particular, the CDS (2.32) is of such a type. In this case, there is an underlying (infinite)-dimensional Lipschitz continuous system Φ on a Banach space \mathcal{X} , that possesses a finite-dimensional compact invariant set $\mathcal{A} \subseteq \mathcal{X}$ and $A_k = R(\mathcal{A})$ for the embedding R chosen according to Theorem 2.21 or 2.26 (cf. (2.30)).

3.1 The Subdivision Scheme

In this section we review the contents of [DHZ16] for the approximation of the set $A_k \subseteq \mathbb{R}^k$. To this end, we will employ a subdivision scheme as defined in [DH97] that has been extended to the CDS in [DHZ16]. Let $Q \subseteq \mathbb{R}^k$ be a (large) compact set that contains A_k . The *global attractor relative to Q* is then defined by

$$A_Q = \bigcap_{j \geq 0} \varphi^j(Q). \quad (3.2)$$

By construction $A_Q \subseteq Q$ and A_Q is compact since φ is continuous. Also observe that A_Q is a subset of the global attractor A of (3.1). However, in general we have $A_Q \neq A \cap Q$. In particular, every backward invariant set is contained in A_Q .

Lemma 3.1 ([DH97, Lemma 3.3]). *Let $B \subseteq Q$ such that $\varphi^{-1}(B) \subseteq B$. Then $B \subseteq A_Q$.*

Proof. By $\varphi^{-1}(B) \subseteq B$ it immediately follows that $B \subseteq \varphi^j(B)$ for all $j \geq 0$. Hence,

$$B \subseteq \bigcap_{j \geq 0} \varphi^j(B) \subseteq \bigcap_{j \geq 0} \varphi^j(Q) = A_Q$$

as desired. □

Applying this lemma to the φ -invariant set $A_k \subseteq Q \subseteq \mathbb{R}^k$ we obtain the following result.

Proposition 3.2 ([DHZ16, Proposition 2]). *Let A_Q be the global attractor relative to the compact set Q and suppose that $A_k \subseteq Q$. Then*

$$A_k \subseteq A_Q.$$

Unfortunately, we can not expect that $A_k = A_Q$ in general as desired. In fact, A_Q may contain several invariant sets and related heteroclinic connections besides A_k . However, if $A_k = R(\mathcal{A})$ for some embedding R and some attractive, Φ -invariant set \mathcal{A} , utilizing sufficiently high powers of Φ “tightens” this inclusion $A_k \subseteq A_Q$ (see Section 3.1.1 for details). In particular, this will become important for the approximation of the local unstable manifold (cf. Proposition 3.21). With Proposition 3.2 in mind we will present a method for the approximation of A_Q in the remainder of this section. Roughly speaking, the idea is as follows.

Suppose we have a finite collection \mathcal{B}_0 of compact subsets of \mathbb{R}^k that cover Q , that is,

$$Q = \bigcup_{B \in \mathcal{B}_0} B.$$

Then one recursively obtains new collections \mathcal{B}_ℓ , from $\mathcal{B}_{\ell-1}$, for $\ell = 1, 2, \dots$, by first subdividing each of the sets $B \in \mathcal{B}_\ell$ into smaller ones and afterwards throwing away those subsets that do not contain any part of the relative global attractor A_Q . Clearly, continuing this process leads to successively better approximations. More precisely, we

first construct a refined box collection $\widehat{\mathcal{B}}_\ell$ such that

$$\bigcup_{B \in \widehat{\mathcal{B}}_\ell} B = \bigcup_{B \in \mathcal{B}_{\ell-1}} B$$

and $\text{diam}(\widehat{\mathcal{B}}_\ell) < \text{diam}(\mathcal{B}_{\ell-1})$, where the diameter is given by $\text{diam}(\mathcal{B}_\ell) = \max_{B \in \mathcal{B}_\ell} \text{diam}(B)$.

This step guarantees that the collections \mathcal{B}_ℓ consist of successively finer sets for increasing ℓ . In fact, by construction the diameter converges to zero for $\ell \rightarrow \infty$. We then define the new collection \mathcal{B}_ℓ by removing each subset in $\widehat{\mathcal{B}}_\ell$ whose preimage does neither intersect itself nor any other subset in $\widehat{\mathcal{B}}_\ell$. This step is responsible for the fact that the unions

$$Q_\ell = \bigcup_{B \in \mathcal{B}_\ell} B$$

approach the relative global attractor. This process will be stopped when a predefined lower bound on the diameter of \mathcal{B}_ℓ relative to Q is reached. We summarize the obtained scheme in [Algorithm 1](#).

Algorithm 1: The Subdivision Scheme

Initialization: Choose a compact set $Q \subseteq \mathbb{R}^k$ such that $A_k \subseteq Q$ and set $\mathcal{B}_0 = \{Q\}$. Fix $\varepsilon > 0$ and $0 < \theta_{\min} \leq \theta_{\max} < 1$.

Repeat the following two steps until $\text{diam}(\mathcal{B}_\ell) < \varepsilon \text{diam}(Q)$.

1) **Subdivision Step:** Construct a new collection $\widehat{\mathcal{B}}_\ell$ such that

$$\bigcup_{B \in \widehat{\mathcal{B}}_\ell} B = \bigcup_{B \in \mathcal{B}_{\ell-1}} B$$

and

$$\text{diam}(\widehat{\mathcal{B}}_\ell) = \theta_\ell \text{diam}(\mathcal{B}_{\ell-1}), \quad (3.3)$$

where $0 < \theta_{\min} \leq \theta_\ell \leq \theta_{\max} < 1$.

2) **Selection Step:** Define the new collection \mathcal{B}_ℓ by

$$\mathcal{B}_\ell = \left\{ B \in \widehat{\mathcal{B}}_\ell \mid \exists \widehat{B} \in \widehat{\mathcal{B}}_\ell \text{ such that } \varphi^{-1}(B) \cap \widehat{B} \neq \emptyset \right\}. \quad (3.4)$$

Remark 3.3.

- (a) Algorithm 1 is included in the software package **GAIO** (*Global Analysis of Invariant Objects*) [DFJ01]. There, the compact sets $B \in \mathcal{B}$ are given by generalized k -dimensional rectangles, hence called “boxes”, of the form

$$B(c, r) = \{ y \in \mathbb{R}^k \mid |y_i - c_i| \leq r_i \text{ for } i = 1, \dots, k \},$$

where $c \in \mathbb{R}^k$ is the center and $r \in \mathbb{R}^k$ contains the radii $r_i > 0$ for $i = 1, \dots, k$. In each subdivision step every box B of the current collection is subdivided by bisection with respect to the j -th coordinate, where j is varied cyclically. This procedure leads to two new rectangles $B_-(c^-, \hat{r})$ and $B_+(c^+, \hat{r})$ where

$$\hat{r}_i = \begin{cases} r_i & \text{for } i \neq j \\ r_i/2 & \text{for } i = j \end{cases}, \quad c_i^\pm = \begin{cases} c_i & \text{for } i \neq j \\ c_i \pm r_i/2 & \text{for } i = j. \end{cases}$$

This construction allows the storage of the boxes in a binary tree which enables fast search algorithms for the selection step (and the continuation step later on, see (3.17)).

- (b) In the selection step (3.4) we have to decide whether or not a preimage of a set $B \in \hat{\mathcal{B}}_\ell$ has a non-empty intersection with another set $\hat{B} \in \hat{\mathcal{B}}_\ell$. For simple problems this may be done analytically, but in general we have to discretize each set $\hat{B} \in \hat{\mathcal{B}}_\ell$ by a finite set of test points $x \in \hat{B}$ and replace (3.4) by

$$\exists x \in \hat{B} \text{ such that } \varphi(x) \in B.$$

Thus, a box B is kept in the collection \mathcal{B}_ℓ if there is a least one test point x such that $\varphi(x) \in B$. Obviously, this is a weaker condition and we make the algorithm more robust by reintroducing boxes if the image $\varphi(x)$ of a test point x is not contained in any $B \in \mathcal{B}_\ell$.

- (c) There are several strategies for the choice of the test points. For instance, one can choose them from a regular grid within each box or sample them at random with respect to the uniform distribution. In this thesis, we will take the k -dimensional Halton sequence which is a quasi-random sequence [Hal64]. Also note that there even is a rigorous discretization possible that reduces the numerical effort, provided that local Lipschitz constants for φ are known [Jun00].

Because the unions Q_ℓ define a nested sequence of compact sets, that is, $Q_{\ell+1} \subseteq Q_\ell$, we have, for each m ,

$$Q_m = \bigcap_{\ell=1}^m Q_\ell. \tag{3.5}$$

Additionally, we can also consider the limit $m \rightarrow \infty$, i.e.,

$$Q_\infty = \bigcap_{\ell=1}^{\infty} Q_\ell,$$

which is the set one obtains after performing infinitely many subdivision and selection steps.

In order to prove the convergence of Algorithm 1, we start by showing that Q_∞ is contained in the relative global attractor A_Q .

Lemma 3.4 ([DHZ16, Lemma 4.2]). *Let A_Q be the global attractor relative to the compact set Q and let Q_∞ be the set obtained after infinitely many subdivision and selection steps. Then $Q_\infty \subseteq A_Q$.*

Proof. Let $x \in Q_\infty$, i.e., $x \in Q_\ell$ for all $\ell \in \mathbb{N}$. Thus, for every $\ell \in \mathbb{N}$ there is a set $B_\ell(x) \in \mathcal{B}_\ell$ such that $x \in B_\ell(x)$. Due to the selection step (3.4) there are points $y_\ell \in Q_\ell$ with $\varphi(y_\ell) \in B_\ell(x)$ and we may assume that $y = \lim_{\ell \rightarrow \infty} y_\ell$ by choosing a convergent subsequence of $(y_\ell)_\ell$ since $(y_\ell)_\ell$ is contained in the compact set Q . By construction we have $y \in Q_\infty$ and since the diameter of $B_\ell(x)$ converges to zero for $\ell \rightarrow \infty$ we conclude with $\lim_{\ell \rightarrow \infty} \varphi(y_\ell) = x$. Finally, the continuity of φ yields $x = \varphi(y) \in \varphi(Q_\infty)$. Hence, $Q_\infty \subseteq \varphi(Q_\infty)$ and Lemma 3.1 yields $Q_\infty \subseteq A_Q$ as desired. \square

Remark 3.5. The sequence $(y_\ell)_\ell$ is a so-called $\text{diam}(\mathcal{B}_\ell)$ -pseudo orbit, i.e.,

$$y_\ell \in Q_\ell \text{ and } \varphi(y_\ell) \in B_{\ell+1}(y_{\ell+1}) \quad \forall \ell \in \mathbb{N},$$

where $B_{\ell+1}(y_{\ell+1})$ denotes a set in $\mathcal{B}_{\ell+1}$ such that $y_{\ell+1} \in B_{\ell+1}(y_{\ell+1})$. Observe that the sets B_ℓ do not have to be unique since sets in \mathcal{B}_ℓ may have non-empty intersections. It turns out that $(y_\ell)_\ell$ converges (after taking a subsequence if necessary) and then applying φ to the limit y yields the point $x \in Q_\infty$ since $\varphi(y_\ell) \in B_\ell(x)$. The concept of such a pseudo orbit will also be useful later on (see the proof of Proposition 3.24).

In order to further show that A_Q is always contained in Q_ℓ for every $\ell \in \mathbb{N}$ and, in particular, $A_Q \subseteq Q_\infty$, we have to introduce an additional assumption, namely

$$\varphi^{-1}(A_Q) \subseteq A_Q.$$

This would be automatically satisfied if φ is in fact a homeomorphism. Alternatively, it also follows if A_k is attracting and $A_k = A_Q$. These observations justify this assumption.

Lemma 3.6 (cf. [DH97, Lemma 3.2]). *Let A_Q be the global attractor relative to the compact set Q and \mathcal{B}_0 be a finite collection of compact sets such that $Q_0 = \bigcup_{B \in \mathcal{B}_0} B = Q$. Suppose A_Q satisfies $\varphi^{-1}(A_Q) \subseteq A_Q$. Then $A_Q \subseteq Q_\ell$ for all $\ell \in \mathbb{N}$. In particular, $A_Q \subseteq Q_\infty$.*

Proof. We will prove this result by induction. By definition $A_Q \subseteq Q = Q_0$. Now suppose $A_Q \subseteq Q_{\ell-1}$ and there is a point $x \in A_Q \subseteq Q_{\ell-1}$ such that $x \notin Q_\ell$ for some $\ell \in \mathbb{N}$. Then there is a set $B \subseteq \widehat{\mathcal{B}}_\ell$ containing x that is removed in the selection step (3.4), i.e., $\varphi^{-1}(B) \cap Q_{\ell-1} = \emptyset$. Hence, $\varphi^{-1}(x) \notin Q_{\ell-1}$ which contradicts the standing assumptions $\varphi^{-1}(A_Q) \subseteq A_Q$. \square

Combining both lemmas finally yields the convergence result of Algorithm 1.

Proposition 3.7 ([DHZ16, Proposition 2]). *Suppose the relative global attractor A_Q satisfies $\varphi^{-1}(A_Q) \subseteq A_Q$. Then*

$$A_Q = Q_\infty.$$

Since Q_∞ may be seen as the limit of the Q_ℓ , Proposition 3.7 can be restated as

$$\lim_{\ell \rightarrow \infty} h(A_Q, Q_\ell) = 0,$$

where $h(B, C)$ is the Hausdorff distance between two compact subsets $B, C \subseteq \mathbb{R}^k$.

Remark 3.8. In order to guarantee that the approximation of A_k is a one-to-one image of \mathcal{A} , the embedding dimension $k \in \mathbb{N}$ has to be chosen sufficiently large (see Sections 2.2 and 2.3). If an estimate of the dimension of \mathcal{A} is not available or where known bounds are extremely large, a sequential procedure has been developed to adaptively increase k without restarting the subdivision scheme 1 [Zie18].

We finish this section by illustrating the proposed scheme 1 on an example.

Example 3.9 (Lorenz attractor). We apply Algorithm 1 on the well-known Lorenz system [Lor63] given by

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{aligned} \tag{3.6}$$

where we use the standard parameters $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$. For these values the system possesses a chaotic attractor [Tuc99, Ste00] which we will approximate using Algorithm 1. To this end, we consider the time- T -map of (3.6) with $T = 0.2$ as the dynamical system (3.1). In Figure 3.1 we show successively finer box coverings Q_ℓ of A_Q for $Q = [-30, 30] \times [-30, 30] \times [-13, 67]$.

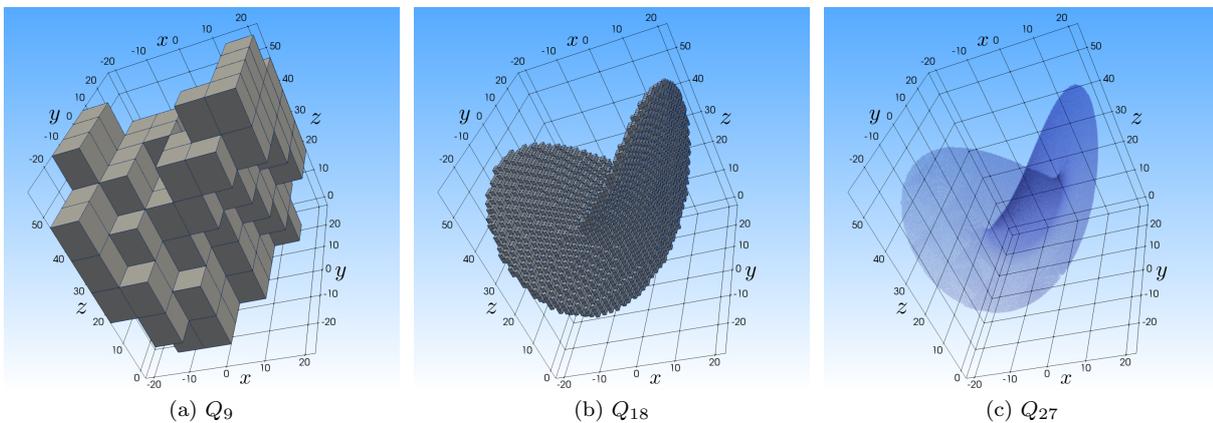


Figure 3.1: (a)-(c) Successively finer coverings Q_ℓ of the relative global attractor A_Q of the Lorenz system (3.6) obtained by Algorithm 1. In (c) we show transparent boxes in order to illustrate the wing shape of A_Q .

3.1.1 Approximation of Attracting Sets

Let us now specifically consider the CDS (2.32) as the dynamical system in (3.1). In this situation the relative global attractor A_Q contains a one-to-one image $A_k = R(\mathcal{A})$ of the invariant set $\mathcal{A} \subseteq \mathcal{X}$ of Φ according to Proposition 3.2. We will show now that the subdivision scheme 1 actually computes A_k if \mathcal{A} is attracting using sufficiently high powers of Φ .

Thus, suppose in this subsection that \mathcal{A} is an attractive set with fundamental neighborhood $\mathcal{U} \subseteq \mathcal{X}$, that is, \mathcal{A} attracts all bounded sets within \mathcal{U} (cf. (2.5)). For instance, \mathcal{A} could be the global attractor of the time- T -map Φ of a DDE or PDE. Moreover, we assume that the computational domain $Q \in \mathbb{R}^k$ is chosen such that

$$A_k \subseteq Q \text{ and } E(Q) \subseteq \mathcal{U}, \quad (3.7)$$

where E is the extended inverse map of the chosen embedding R (see Section 2.4). Hence, for every $x \in Q$, the trajectories $\Phi^j(E(x))$, $j \in \mathbb{N}$, will eventually approach the attracting set \mathcal{A} . Unfortunately, these assumptions do not guarantee that A_k is also an attracting set for the CDS φ . Due to the fact that the observation map R is only one-to-one on \mathcal{A} there may exist “spurious fixed points” $\bar{x} \in Q$ in the sense that

$$\bar{x} = \varphi(\bar{x})$$

even though $\Phi(E(\bar{x})) \neq E(\bar{x})$ may be closer to \mathcal{A} than $E(\bar{x})$. To overcome this problem we replace the map Φ in the definition of the CDS with higher powers Φ^m for $m \in \mathbb{N}$, i.e., we define a family of continuous maps by

$$\varphi_m = R \circ \Phi^m \circ E. \quad (3.8)$$

Keep in mind that in general $\varphi_m \neq \varphi^m$ since equality can only be guaranteed on the embedded set $A_k \subseteq \mathbb{R}^k$. Nevertheless, A_k remains an invariant set for φ_m since \mathcal{A} is an invariant set for Φ^m for every $m \in \mathbb{N}$. Thus, the embedding results in Sections 2.2 and 2.3 are still valid and the chosen observation map R remains one-to-one on \mathcal{A} .

We will denote the corresponding global attractors relative to Q with respect to φ_m by A_Q^m and see that A_Q^m covers A_k by (3.7) and Lemma 3.1 since A_k is φ_m -invariant.

Lemma 3.10 ([DHZ16, Lemma 4.4]). *Let A_Q^m be the global attractors relative to Q with respect to φ_m . Then*

$$A_k \subseteq A_Q^m \text{ for all } m \in \mathbb{N}.$$

Finally, by considering

$$A_Q^\infty = \bigcap_{m \geq 1} A_Q^m \quad (3.9)$$

as the “limit” of A_Q^m for $m \rightarrow \infty$ we have $A_k \subseteq A_Q^\infty$ by the previous lemma and we can even prove equality.

Proposition 3.11 ([DHZ16, Proposition 4]). *Let A_Q^m be the global attractors relative to Q with respect to φ_m and $A_Q^\infty = \bigcap_{m \geq 1} A_Q^m$. Then*

$$A_k = A_Q^\infty.$$

Proof. Suppose that there is a point $x \in A_Q^\infty \setminus A_k$. As A_k is compact, this implies $\text{dist}(x, A_k) = \epsilon > 0$. In addition, since \mathcal{A} is compact and R is continuous, there is $\delta > 0$ such that for $u \in \mathcal{X}$:

$$\text{dist}(u, \mathcal{A}) < \delta \quad \Rightarrow \quad \text{dist}(R(u), A_k) < \frac{\epsilon}{2}.$$

Now for $\mathcal{V} = E(Q)$ we have $\mathcal{V} \subseteq \mathcal{U}$ by assumption (see (3.7)) and since \mathcal{A} is attracting and \mathcal{V} is bounded within \mathcal{U} (E is continuous, Q is compact) there is $m \in \mathbb{N}$ such that

$$\text{dist}(\Phi^m(\mathcal{V}), \mathcal{A}) < \delta.$$

By the choice of δ it follows that

$$\text{dist}(\varphi_m(Q), A_k) = \text{dist}(R(\Phi^m(\mathcal{V})), A_k) < \frac{\epsilon}{2}.$$

This implies $x \notin \varphi_m(Q)$ since $\text{dist}(x, A_k) = \epsilon > 0$. Thus, $x \notin A_Q^m$ and, in particular, $x \notin A_Q^\infty$ which contradicts the initial assumption. \square

Proposition 3.11 roughly says that it is possible to approximate a one-to-one image of an attracting set \mathcal{A} of Φ if we perform the computations with appropriately high iterates of Φ . In this sense we can close the “gap” in the inclusion $A_k \subseteq A_Q$ (cf. Proposition 3.2).

3.2 A Path Following Method

As discussed in the last section [Algorithm 1](#) allows the approximation of the (relative) global attractor which we will now extend to parameter-dependent systems. Therefore, we introduce a real-valued parameter to the system (3.1) and develop a path following method for the approximation of corresponding parameter-dependent invariant sets that is based on the ideas of [Algorithm 1](#). That is, we will consider a dynamical system of the form

$$x_{j+1} = \varphi(x_j, \lambda), \quad j = 0, 1, \dots, \quad (3.10)$$

where $x_j \in \mathbb{R}^k$ and $\varphi : \mathbb{R}^k \times \Lambda \rightarrow \mathbb{R}^k$ is continuous for each parameter λ in a compact interval $\Lambda \subseteq \mathbb{R}$ and uniformly continuous in λ on bounded subsets of \mathbb{R}^k . In what follows, for fixed $\lambda \in \Lambda$, we use the abbreviation $\varphi_\lambda(x) := \varphi(x, \lambda)$. Moreover, we assume that φ_λ has a compact global attractor $A_k^\lambda \subseteq \mathbb{R}^k$ for each $\lambda \in \Lambda$. Later on, we will additionally assume that A_k^λ is upper semi-continuous in λ .

Remark 3.12. For a parameter-dependent infinite-dimensional system $\Phi : \mathcal{X} \times \Lambda \rightarrow \mathcal{X}$ (2.33) with finite-dimensional compact global attractor $\mathcal{A}^\lambda \subseteq \mathcal{X}$ the corresponding pCDS $\varphi_\lambda : \mathbb{R}^k \rightarrow \mathbb{R}^k$ (2.34) does not directly fall into the class of systems we are considering here. In fact, the set $A_k^\lambda := R(\mathcal{A}^\lambda)$ may not be attractive for φ_λ even though \mathcal{A}^λ is the global attractor of $\Phi_\lambda = \Phi(\cdot, \lambda)$ (see “spurious fixed points” in [Section 3.1.1](#)). However, in this thesis we still assume that A_k^λ is indeed the compact global attractor for φ_λ . This allows us to actually compute the one-to-one image A_k^λ of the attractor \mathcal{A}^λ using the subdivision scheme 1 for some finite number of parameter values λ . Otherwise, we will only compute the parameter-dependent relative global attractor A_Q^λ which contains A_k^λ (see [Proposition 3.2](#)).

In this section we develop two path following algorithms that allow us to compute the relative global attractor for various parameter values $\lambda \in \Lambda$ of (3.10) by reusing previously obtained coverings. For the first method we approximate the relative global attractor $A_Q^{\lambda_0}$ for some $\lambda_0 \in \Lambda$ and then use a covering of this set, denoted by $Q_\ell^{\lambda_0}$, as a starting point to compute the relative global attractor $A_Q^{\lambda_1}$ for $\lambda_1 \in \Lambda$ sufficiently close to λ_0 . The second algorithm improves this idea by *predicting* a good initial covering from the approximation of two previously computed attractors using a *set-valued Taylor expansion*. Afterwards a *corrector* step actually computes the invariant set for the current parameter value. Large parts of this section are also contained in [[GZED20](#)] and [[ZGD20](#)] to which the author has made substantial contributions.

We start with the following observation that an attractor A_k can be approximated using different initial sets $Q_i \subseteq \mathbb{R}^k$, provided they each cover A_k .

Lemma 3.13 (cf. [[GZED20](#), Lemma 2.3]). *Let $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be continuous with compact global attractor $A_k \subseteq \mathbb{R}^k$. Furthermore, let $Q^1, Q^2 \subseteq \mathbb{R}^k$ be compact sets such that $A_k \subseteq Q^i$, $i = 1, 2$. Then*

$$A_{Q^1} = A_k = A_{Q^2},$$

where A_{Q^i} is the global attractor relative to Q^i for $i = 1, 2$.

Proof. This statement follows directly from the definition of the global attractor relative to a compact set Q (cf. (3.2)) and the fact that $A_k \subseteq Q^i$ for $i = 1, 2$. \square

Combining this lemma with the convergence result 3.7 allows the computation of A using Algorithm 1. As a direct consequence of Lemma 3.13 we obtain the following.

Proposition 3.14 (cf. [GZED20, Proposition 2.4]). *Let $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be continuous with compact global attractor $A_k \subseteq \mathbb{R}^k$. Let $Q^1, Q^2 \subseteq \mathbb{R}^k$ be compact sets such that $A_k \subseteq Q_i$ and $\varphi^{-1}(A_{Q_i}) \subseteq A_{Q_i}$ for $i = 1, 2$. Denote by Q_ℓ^i , $i = 1, 2$, the corresponding approximations of A_{Q_i} obtained after ℓ subdivision and selection steps (see (3.5)). Then*

$$\lim_{\ell \rightarrow \infty} h(Q_\ell^1, Q_\ell^2) = 0,$$

where $h(B, C)$ is the Hausdorff distance between two subsets $B, C \subseteq \mathbb{R}^k$.

Proof. By Proposition 3.7 and Lemma 3.13 we have, for $i = 1, 2$,

$$\lim_{\ell \rightarrow \infty} h(Q_\ell^i, A_{Q_i}) = \lim_{\ell \rightarrow \infty} h(Q_\ell^i, A_k) = 0$$

and the triangle inequality finishes the proof. \square

This result says that by performing sufficiently many subdivision and selection steps we obtain a good approximation of the same set, which is indeed the attracting set A_k , no matter how we choose the computational domain Q as it contains A_k . Note that in contrast to the original Proposition 2.4 in [GZED20] we additionally assume $\varphi^{-1}(A_{Q_i}) \subseteq A_{Q_i}$ which is required in Proposition 3.7 since φ is just continuous.

In what follows, we assume that the computational domain $Q \subseteq \mathbb{R}^k$ is large enough such that it contains every attractor A_k^λ for all $\lambda \in \Lambda$. For a number of iterations $m \in \mathbb{N}$ of Algorithm 1 we denote by $\mathcal{Q}(\lambda_0)$ the generated family of compact sets $Q_\ell^{\lambda_0}$, $\ell = 0, 1, \dots, m$, using the map φ_{λ_0} . Note that every such $Q_\ell^{\lambda_0}$ contains the global attractor $A_Q^{\lambda_0}$ relative to Q of the map φ_{λ_0} .

In the next step we want to approximate $A_k^{\lambda_1}$ for a parameter value $\lambda_1 \in \Lambda$ in a neighborhood of λ_0 . To this end, we either fix $\lambda_1 \in \Lambda$ and choose $L \leq m$ sufficiently small such that $A_k^{\lambda_1} \subseteq Q_L^{\lambda_0}$ or we fix $L \leq m$ and choose $\lambda_1 \in \Lambda$ sufficiently close to λ_0 such that $A_k^{\lambda_1} \subseteq Q_L^{\lambda_0}$. In both scenarios Lemma 3.13 and Proposition 3.14 allow us to approximate the relative global attractor $A_Q^{\lambda_1}$ of the map φ_{λ_1} using the initial covering $Q_L^{\lambda_0}$ with the subdivision Algorithm 1. Note that $L = 0$ is always a *feasible* choice for every λ_1 , since $Q_0^{\lambda_0} = Q$ and $A_Q^{\lambda_1} \subseteq Q$ for $\lambda_1 \in \Lambda$ by assumption. Note that in order to achieve the same approximation quality, we only have to perform $m - L$ subdivision and selection steps which yields a smaller computational effort compared to starting Algorithm 1 with the initial set $Q \subseteq \mathbb{R}^k$.

To discuss the choice of λ_1 for fixed L we have to make sure that A_k^λ behaves well under small perturbations of λ in a certain sense. To this end, we choose the distance between two sets (see (2.6)) in order to compare the attractors. This allows us to characterize *upper semi-continuity* of attractors as follows.

Definition 3.15. A family of attractors $A^\lambda \subseteq \mathcal{X}$ in some metric space \mathcal{X} is *upper semi-continuous* in $\lambda_0 \in \Lambda$ if

$$\lim_{\lambda \rightarrow \lambda_0} \text{dist}(A^\lambda, A^{\lambda_0}) = 0. \quad (3.11)$$

The attractor A^λ is called *upper semi-continuous* if A^λ is upper semi-continuous for each $\lambda_0 \in \Lambda$.

Observe that upper semi-continuity at λ_0 implies that for every $\varepsilon > 0$ there exists a neighborhood $U_\delta(\lambda_0) \subseteq \Lambda$ of λ_0 such that $A^\lambda \subseteq U_\varepsilon(A^{\lambda_0})$ for all $\lambda \in U_\delta(\lambda_0)$, where $U_\varepsilon(A^{\lambda_0}) \subseteq \mathbb{R}^k$ denotes the ε -neighborhood of A^{λ_0} . Therefore, the attractor can not suddenly “explode” by slightly varying λ which is a naturally needed property for our proposed path following scheme. Otherwise, if the attracting set grows too much we can not assure that any previously computed covering $Q_L^{\lambda_0}$ besides $L = 0$ still covers $A_k^{\lambda_1}$. In this case the subdivision scheme 1 fails to approximate the global attractor $A_k^{\lambda_1}$ but still generates a covering of the global attractor $A_{Q_k^{\lambda_0}}$ relative to $Q_L^{\lambda_0}$ of the map φ_{λ_1} . This set may miss some parts of the whole attractor $A_k^{\lambda_1}$ and to counteract this phenomenon we will use the recovering technique described in Remark 3.3 (b).

For the sake of completeness we note that *lower semi-continuity* of A^λ , i.e.,

$$\lim_{\lambda \rightarrow \lambda_0} \text{dist}(A^{\lambda_0}, A^\lambda) = 0 \quad \forall \lambda_0 \in \Lambda$$

prevents a sudden “shrinking” of A^λ which holds for instance for gradient systems with hyperbolic fixed points [HR89]. But this is, in general, no issue for the proposed algorithm.

Under the assumption that A_k^λ is upper semi-continuous we can actually prove that for every $L \in \{0, \dots, m\}$ there is indeed a range of parameter values $\lambda \in \Lambda$ such that the attractor A_k^λ can be approximated by the subdivision scheme 1 with initial compact set $Q_L^{\lambda_0}$.

Proposition 3.16 ([GZED20, Proposition 2.6]). *Let $\mathcal{Q}(\lambda_0) = \{Q_0^{\lambda_0}, \dots, Q_m^{\lambda_0}\}$ be the family of sets generated by m steps of the subdivision scheme 1. Suppose that there is $\varepsilon > 0$ such that $U_\varepsilon(A_k^{\lambda_0}) \subseteq Q_\ell^{\lambda_0}$ for all $\ell = 0, \dots, m$ and A_k^λ is upper semi-continuous at λ_0 . Then for every $\ell = 0, \dots, m$ there exists $\delta = \delta(\ell) > 0$ such that $A_k^\lambda \subseteq Q_\ell^{\lambda_0}$ for all $\lambda \in U_\delta(\lambda_0)$. In particular, A_k^λ can be approximated by using the initial compact set $Q_\ell^{\lambda_0}$.*

Proof. Let $\ell \in \{0, \dots, m\}$ be fixed. Due to the upper semi-continuity of A_k^λ in λ_0 there is $\delta > 0$ such that

$$\text{dist}(A_k^\lambda, A_k^{\lambda_0}) < \varepsilon \text{ for all } \lambda \in U_\delta(\lambda_0).$$

Thus, we conclude by assumption that

$$A_k^\lambda \subseteq U_\varepsilon(A_k^{\lambda_0}) \subseteq Q_\ell^{\lambda_0} \text{ for all } \lambda \in U_\delta(\lambda_0).$$

□

The assumption that a small neighborhood of $A_k^{\lambda_0}$ is still contained in the approximation $Q_\ell^{\lambda_0}$ is only technical and always satisfied in practice. It is only invalid when their

boundaries have non-empty intersection. In the following we will call $\lambda \in U_\delta(\lambda_0)$ *feasible* and Proposition 3.16 guarantees the existence of feasible $\lambda \in \Lambda$. However, the size $\delta(\ell)$ of the neighborhood $U_\delta(\lambda_0)$ is not known in practice and we will discuss this issue later on (see Section 3.2.1).

In order to utilize Proposition 3.16 for the pCDS (2.34), we make use of the following observation.

Lemma 3.17. *Let $\mathcal{A}^\lambda \subseteq \mathcal{X}$ be upper semi-continuous in $\lambda_0 \in \Lambda$ and $R : \mathcal{X} \rightarrow \mathbb{R}^k$ be a Lipschitz continuous map that is one-to-one on \mathcal{A}^λ for all $\lambda \in \Lambda$. Then $A_k^\lambda = R(\mathcal{A}^\lambda) \subseteq \mathbb{R}^k$ is upper semi-continuous in $\lambda_0 \in \Lambda$. In particular, if $\mathcal{A}^\lambda \subseteq \mathcal{X}$ is upper semi-continuous then $A_k^\lambda \subseteq \mathbb{R}^k$ is also upper semi-continuous.*

Proof. Let $L_R > 0$ be the Lipschitz constant of R . By assumption R is one-to-one on \mathcal{A}^λ and it immediately follows that

$$\begin{aligned} \lim_{\lambda \rightarrow \lambda_0} \text{dist}(A_k^\lambda, A_k^{\lambda_0}) &= \lim_{\lambda \rightarrow \lambda_0} \sup_{x \in A_k^\lambda} \inf_{y \in A_k^{\lambda_0}} |x - y| \\ &= \lim_{\lambda \rightarrow \lambda_0} \sup_{u \in \mathcal{A}^\lambda} \inf_{v \in \mathcal{A}^{\lambda_0}} |R(u) - R(v)| \\ &\leq L_R \lim_{\lambda \rightarrow \lambda_0} \sup_{u \in \mathcal{A}^\lambda} \inf_{v \in \mathcal{A}^{\lambda_0}} \|u - v\| \\ &= L_R \lim_{\lambda \rightarrow \lambda_0} \text{dist}(\mathcal{A}^\lambda, \mathcal{A}^{\lambda_0}) = 0, \end{aligned}$$

which proves the claimed upper semi-continuity of A_k^λ in λ_0 . □

Thus, provided that R is one-to-one on \mathcal{A}^λ for all $\lambda \in \Lambda$ it is sufficient to assume that the attractor $\mathcal{A}^\lambda \subseteq \mathcal{X}$ of Φ_λ is upper semi-continuous since it carries over to the attractors $A_k^\lambda \subseteq \mathbb{R}^k$ of φ_λ . In fact, this assumption is justified by several classical results, e.g., [Klo06] for delay differential equations or [Hal80] for dissipative systems on metric spaces. However, we recall that the observation map R is only one-to-one for a dense subset $\Lambda_{\mathbb{N}} \subseteq \Lambda$ as discussed in Section 2.4. Throughout the remainder of this section we now suppose that A_k^λ is upper semi-continuous and the additional assumptions in Proposition 3.16 are satisfied.

After discussing the choice of $\lambda_1 \in \Lambda$ for fixed $L \in \{0, \dots, m\}$ the following proposition tells us, how to choose $L \in \{0, \dots, m\}$ for a fixed step size $\delta > 0$ in the parameter space Λ . To this end, we additionally assume that the distance $\text{dist}(A_k^{\lambda_1}, A_k^{\lambda_0})$ depends on δ .

Proposition 3.18. *Let $\delta > 0$ and $\lambda_1 \in \Lambda$ with $|\lambda_1 - \lambda_0| \leq \delta$. Suppose there is a constant $C > 0$ such that*

$$\text{dist}(A_k^{\lambda_1}, A_k^{\lambda_0}) \leq C\delta \tag{3.12}$$

and $U_{C\delta}(A_k^{\lambda_0}) \subseteq Q_L^{\lambda_0}$ for some $L \in \{0, \dots, m\}$. Then $A_k^{\lambda_1} \subseteq Q_L^{\lambda_0}$ and, in particular, λ_1 is feasible for the approximation of $A_k^{\lambda_1}$ with initial set $Q_L^{\lambda_0}$.

Proof. By (3.12) the $C\delta$ -neighborhood $U_{C\delta}(A_k^{\lambda_0})$ of $A_k^{\lambda_0}$ contains $A_k^{\lambda_1}$ and by assumption we immediately obtain

$$A_k^{\lambda_1} \subseteq U_{C\delta}(A_k^{\lambda_0}) \subseteq Q_L^{\lambda_0}$$

as claimed. □

In practice, however, checking upper semi-continuity of attractors is hard for a nonlinear dynamical system. In particular, proving the assumption made in Proposition 3.18 is not possible in general and the result has only theoretical significance. We summarize the proposed path following method in Algorithm 2.

Algorithm 2: Set-Oriented Path Following Method

Initialization: Let $Q \subseteq \mathbb{R}^k$ be compact such that $A_k^\lambda \subseteq Q \forall \lambda \in \Lambda$ and $\lambda_0 \in \Lambda$. Apply m iterations of the subdivision scheme 1 on Q with respect to φ_{λ_0} to obtain a family $\mathcal{Q}(\lambda_0) = \{Q_0^{\lambda_0}, \dots, Q_m^{\lambda_0}\}$ of approximations of $A_k^{\lambda_0}$.

Path Following: For $j = 0, 1, \dots$

- 1) Choose $L_{j+1} \in \{0, \dots, m\}$ and take a feasible $\lambda_{j+1} > \lambda_j$ such that $A_k^{\lambda_{j+1}} \subseteq Q_{L_{j+1}}^{\lambda_j}$.
- 2) Perform $m - L_{j+1}$ subdivision and selection steps starting on $Q_{L_{j+1}}^{\lambda_j}$ with respect to $\varphi_{\lambda_{j+1}}$ in order to generate a new family

$$\mathcal{Q}(\lambda_{j+1}) = \{Q_0^{\lambda_{j+1}}, \dots, Q_m^{\lambda_{j+1}}\}$$

of approximations of $A_k^{\lambda_{j+1}}$, where $Q_\ell^{\lambda_{j+1}} = Q_\ell^{\lambda_j}$ for $\ell = 0, \dots, L_{j+1}$.

Remark 3.19.

- (a) In step 1) of Algorithm 2 one can also always take $\lambda_{j+1} < \lambda_j$ and follow the path in the negative direction.
- (b) Intuitively, choosing a larger L_{j+1} decreases the range of feasible parameters λ_{j+1} since the set $Q_{L_{j+1}}^{\lambda_j}$ gets smaller. However, one has to perform less subdivision and selection steps to reach the same approximation quality of the final box covering of $A_k^{\lambda_{j+1}}$. With more knowledge on the upper semi-continuity property (3.11) this can be made more precise (see Proposition 3.18).
- (c) If λ_{j+1} is not feasible, i.e., $\lambda_{j+1} \notin U_\delta(\lambda_j)$, $Q_{L_{j+1}}^{\lambda_j}$ might not contain all of $A_k^{\lambda_{j+1}}$ and thus the subdivision algorithm does not approximate the whole attractor $A_k^{\lambda_{j+1}}$ but only parts of it. Therefore, to make the algorithm more robust, we reintroduce sets if the image $\varphi_{\lambda_{j+1}}(x)$ of a test point $x \in \mathbb{R}^n$ is not contained in any set of the current collection $\mathcal{B}_k(\lambda_j)$ (cf. Remark 3.3 (b)).

Finally, in Figure 3.2 we illustrate two steps of Algorithm 2 applied to the Lorenz system (3.6) where we use β as our parameter of interest (cf. Example 3.9). We note that the derivation of the Lorenz model shows that β is related to the aspect ratio of the convection cells.

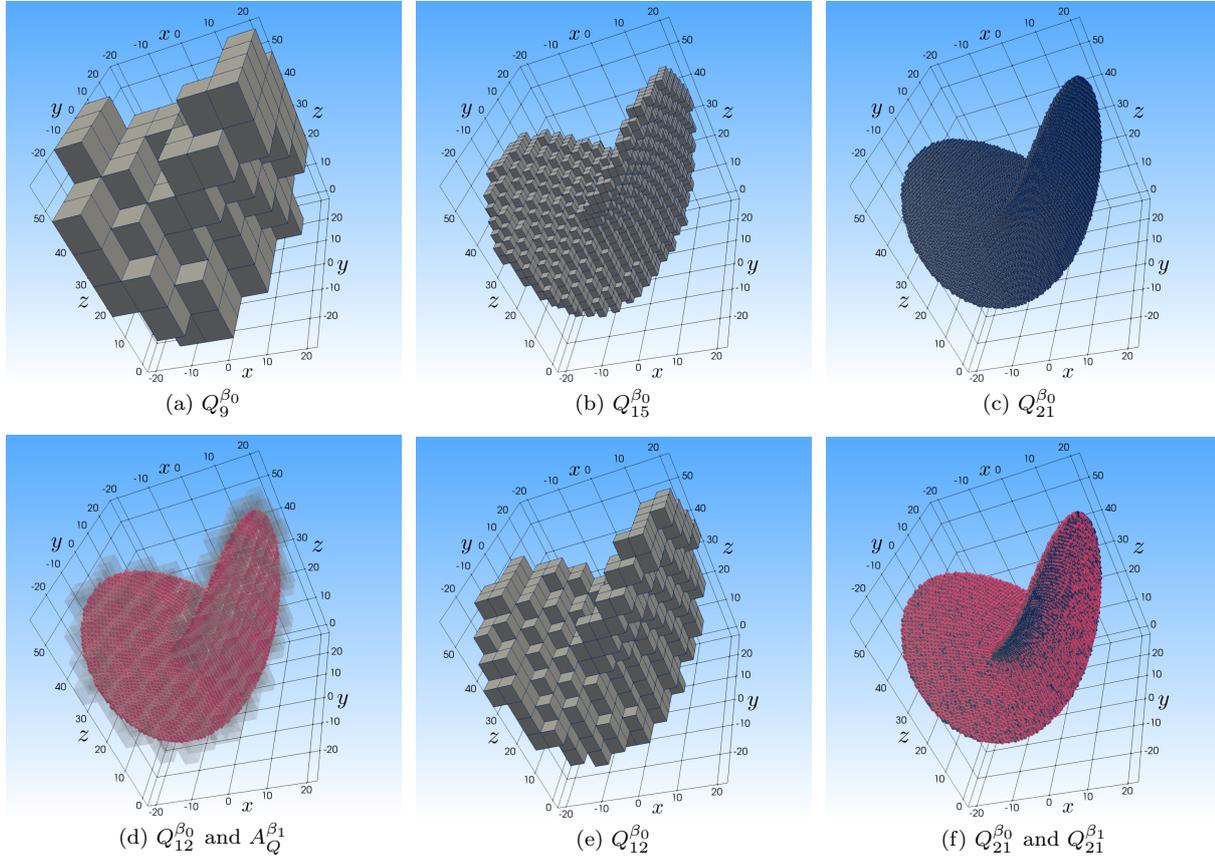


Figure 3.2: Illustration of Algorithm 2 for the Lorenz system (3.6) with $\beta_0 = 8/3$ and $\beta_1 = 2.5$. (a)-(c) Successively finer box coverings $Q_\ell^{\beta_0}$ of the attractor $A_Q^{\beta_0}$. In (c) the final approximation $m = 21$ is shown. (d) Neighborhood $Q_{12}^{\beta_0}$ (transparent boxes) such that $A_Q^{\beta_1} \subseteq Q_{12}^{\beta_0}$ (red). (e) Initial box covering $Q_{12}^{\beta_0}$ for the approximation of $A_Q^{\beta_1}$. (f) Final box covering $Q_{21}^{\beta_1}$ of $A_Q^{\beta_1}$ after 9 subdivision and selection steps. Here, the blue boxes are the previously computed approximation $Q_{21}^{\beta_0}$ illustrating the slight change of the attracting set.

3.2.1 A Set-Oriented Predictor-Corrector Method

Depending on the underlying dynamics a small change of the parameter of interest typically leads to a small change of the attracting set (cf. Definition 3.15). Hence, Algorithm 2 is able to follow the attracting set very efficiently. However, if changing the parameter results in a rapid growth of the attractor, we either have to choose a significantly smaller parameter value or a very coarse covering for the next parameter, which makes Algorithm 2 inefficient. For instance, this can be seen in Figure 3.2 (d) where one has to go $m - L = 9$ steps back in order to find a suitable covering. This occurs in particular in higher dimensions which we illustrate by the Mackey-Glass delay differential equation [MG77] where the parameter of interest is the delay time τ . For $\tau \in [5, 13]$, this system possesses a stable limit cycle that grows and rotates in state space for increasing time delays τ . Thus, an embedding dimension of $k = 5$ suffices for the approximation of a one-to-one image of the global attractor via the pCDS (2.34). A detailed discussion of this particular dynamical system can be found in Section 5.2. In Figure 3.3 (a) we show a projection of the embedded attractor which is indeed a (perturbed) circle. Changing the parameter to τ_1 leads to a larger orbit which is slightly rotated in the embedding space. Therefore, we have to use a very coarse covering of $A_k^{\tau_0}$ such that $A_k^{\tau_1}$ is covered (see Figure 3.3 (b)). Moreover, it is not clear a priori how coarse this covering has to be.

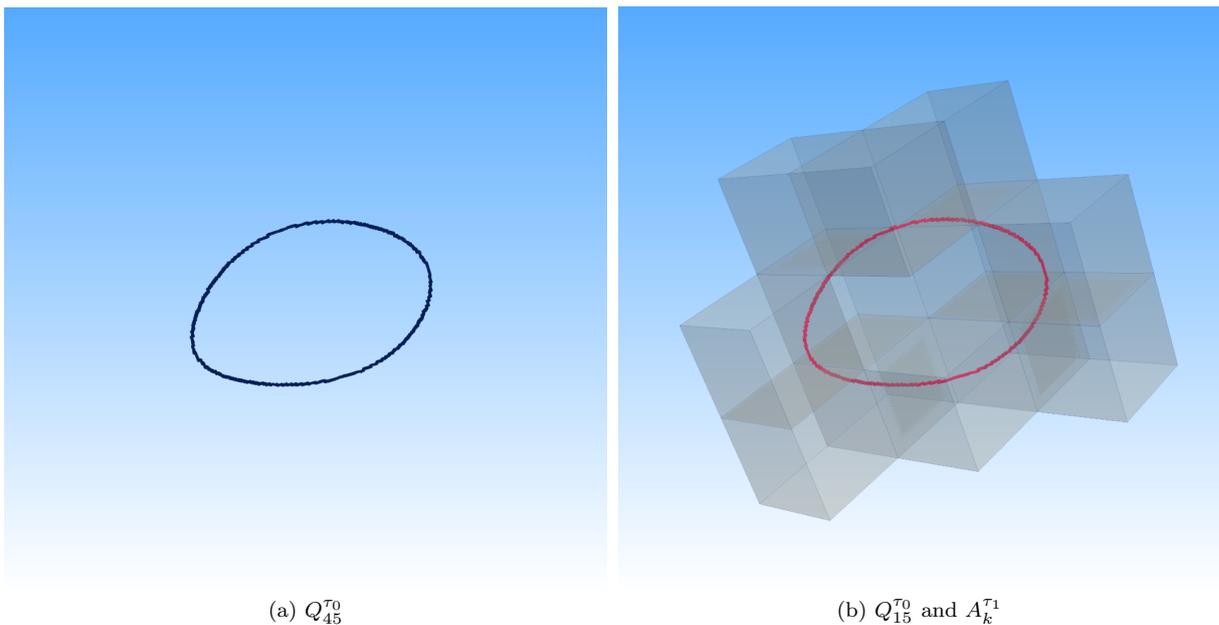


Figure 3.3: Illustration of a possible inefficiency of Algorithm 2 by means of the Mackey-Glass equation for $\tau_0 = 6.00$ and $\tau_1 = 6.05$. (a) Approximation $Q_{45}^{\tau_0}$ (blue) of $A_k^{\tau_0}$ obtained by Algorithm 1. (b) Best initial covering $Q_{15}^{\tau_0}$ for $A_k^{\tau_0}$ such that $A_k^{\tau_1} \subseteq Q_{15}^{\tau_0}$. In this case, we have to perform $L = 30$ subdivision and selection steps in order to obtain a covering $Q_{45}^{\tau_1}$ (red) of $A_k^{\tau_1}$ of the same approximation quality as the covering of $A_k^{\tau_0}$.

Motivated by this observation, we propose an improved method for the computation of parameter-dependent attractors. In view of the discussion in the last section we want to utilize previous computations of the attracting sets in order to accelerate further approx-

imations under slight variations of the parameter of interest. To this end, we propose a *set-oriented predictor-corrector method* (cf. [AG93] for standard approaches of numerical continuation methods). More concretely, let us consider the map

$$\iota : \Lambda \rightarrow \mathcal{P}(\mathbb{R}^k), \quad \lambda \mapsto A_k^\lambda,$$

which maps a parameter value $\lambda \in \Lambda$ onto the attractor $A_k^\lambda \subseteq \mathbb{R}^k$. Then the aim in this section can be rephrased by numerically evaluating this map for some discrete subset of Λ . Now given three parameters values $\lambda_{j-1} < \lambda_j < \lambda_{j+1}$ *formal* Taylor expansion of this map ι yields

$$\begin{aligned} \iota(\lambda_{j+1}) &\approx \iota(\lambda_j) + \iota'(\lambda_j)(\lambda_{j+1} - \lambda_j) \\ &\approx \iota(\lambda_j) + \frac{\lambda_{j+1} - \lambda_j}{\lambda_j - \lambda_{j-1}}(\iota(\lambda_j) - \iota(\lambda_{j-1})), \end{aligned}$$

where we *formally* approximated the derivative $\iota'(\lambda_j)$ by finite differences. Again we want to emphasize that this equation is only formally correct since there is no well-defined set-valued difference operator on arbitrary subsets of \mathbb{R}^k let alone a well-defined derivative. But note that there are several concepts for compact convex subsets (see [BJ70, CCRFJG11, KTZ15] and the references therein). Thus, $A_k^{\lambda_{j+1}}$ is *formally* approximated as

$$A_k^{\lambda_{j+1}} \approx A_k^{\lambda_j} + h_j(A_k^{\lambda_j} - A_k^{\lambda_{j-1}}), \quad (3.13)$$

where $h_j = \frac{\lambda_{j+1} - \lambda_j}{\lambda_j - \lambda_{j-1}}$. Now suppose we approximated $A_k^{\lambda_{j-1}}$ and $A_k^{\lambda_j}$ by $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$, respectively, using Algorithm 1. Then in the spirit of (3.13) we *predict* an initial guess for the approximation of $A_k^{\lambda_{j+1}}$ by

$$Q_L^{\lambda_{j+1}} \approx Q_m^{\lambda_j} + h_j(Q_m^{\lambda_j} - Q_m^{\lambda_{j-1}}). \quad (3.14)$$

Here, we choose a level L for some $L \leq m$ to compensate for possible errors and a consecutive *corrector* step will achieve the same approximation quality as before, i.e, we compute $Q_m^{\lambda_{j+1}}$ by using $Q_L^{\lambda_{j+1}}$. In the following we will develop a numerical realization of (3.14).

Since, in general, there is no one-to-one correspondence between the boxes contained in the coverings $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$ or corresponding discretizations $X = \{x_p^{j-1}\}_{p=1}^{n_1} \subseteq Q_m^{\lambda_{j-1}}$ and $Y = \{y_q^j\}_{q=1}^{n_2} \subseteq Q_m^{\lambda_j}$, it is not immediately clear how to define a box $B \subseteq Q_L^{\lambda_{j+1}}$. In our numerical realization we consider every point $y_q^j \in Y$ that lies inside a predefined neighborhood U of $x_p^{j-1} \in X$ as a potential image point under the underlying map $X \rightarrow Y$. Then, with (3.14) in mind, for every $y_q^j \in Y$ we compute $z^{j+1} = y_q^j + h_j(y_q^j - x_p^{j-1})$ for every $x_p^{j-1} \in U$. The initial guess $Q_L^{\lambda_{j+1}}$ will then be constructed by all boxes $B \in \mathcal{B}_L$ that contain some z^{j+1} for a predefined partition \mathcal{B}_L of the computational domain $Q \subseteq \mathbb{R}^k$.

Roughly speaking, this procedure can be interpreted as a set-valued finite difference approach and we summarize the scheme in Algorithm 3. We note that a similar predictor method has been developed in [Pei17] for parameter-dependent multi-objective optimization problems.

Algorithm 3: Set-Oriented Predictor Step

Initialization: Let $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$ be approximations of $A_k^{\lambda_{j-1}}$ and $A_k^{\lambda_j}$, respectively, and $0 \leq L \leq m$ and $K \in \mathbb{N}$. Let $\mathcal{B}_0 = \{Q\}$ and set $h_j = \frac{\lambda_{j+1} - \lambda_j}{\lambda_j - \lambda_{j-1}}$.

- 1) Discretize $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$ by finite sets of points $X = \{x_p^{j-1}\}_{p=1}^{n_1}$ and $Y = \{y_q^j\}_{q=1}^{n_2}$.
- 2) **Search Radius:** For every $x_p^{j-1} \in X$ find the K nearest neighbors in Y and denote by $I_p \subseteq \{1, \dots, n_2\}$ the corresponding indices and their distances by $d_{p,i}$ for $i = 1, \dots, K$. Set

$$r = \max_{p \in \{1, \dots, n_1\}} d_{p,2}$$

and remove in each index set I_p those indices $\hat{i} \in I_p$ such that $d_{p,\hat{i}} > r$.

- 3) Construct a partition \mathcal{B}_L^{j+1} of Q using only step 1) of Algorithm 1. Set $Q_L^{\lambda_{j+1}} = \emptyset$.
- 4) **Prediction Step:** For $p = 1, \dots, n_1$:
For each $i \in I_p$ compute the point

$$z^{j+1} = y_i^j + h_j (y_i^j - x_p^{j-1})$$

and add the box $B \in \mathcal{B}_L^{j+1}$ such that $z^{j+1} \in B$ to the covering $Q_L^{\lambda_{j+1}}$, i.e.,

$$Q_L^{\lambda_{j+1}} = Q_L^{\lambda_{j+1}} \cup B.$$

Remark 3.20.

- (a) In order to discretize $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$ in step 1) of Algorithm 3, one can simply choose the midpoints of the boxes of the coverings which generate a grid. Sometimes more sophisticated choices can be useful (cf. Algorithm 9) but, in general, those require a longer computational time.
- (b) A good choice for the number of neighbors is $K = 2k + 1$ since a k -dimensional box has $2k$ hyper surfaces. This corresponds to a $2k + 1$ -point stencil in \mathbb{R}^k for the finite difference approximation of the derivative.
- (c) By defining a cut-off radius $r > 0$ we neglect points that are too far away for the ensuing computation. In fact, we only want to consider those points that are within r distance, where r is the largest distance between all first nearest neighbors. Note that in practice we compute r with respect to the distance to the second nearest neighbor since often times the first nearest neighbor will be the point itself due to the grid structure chosen in (a). In particular, if $X \subseteq Y$ then r would be zero, if it had been computed with respect to the first nearest neighbor. We also note that the nearest neighbor search has to be performed only once which can be done fast using a k -d tree [FBF77].
- (d) The partition \mathcal{B}_L^{j+1} does not have to be stored in memory. In fact, in GAIO the boxes $B \in \mathcal{B}_L^{j+1}$ will be added as leaves at level L in the binary tree.

The predictor step introduced in Algorithm 3 can significantly improve the initial guess for the next parameter value and we illustrate this using the Mackey-Glass example discussed at the beginning of this section (cf. Figure 3.3). After computing the attracting sets $A_k^{\tau_0}$ and $A_k^{\tau_1}$ we use Algorithm 3 in order to create an initial box covering for the attracting set $A_k^{\tau_2}$. In Figure 3.4 (a) we show the covering $Q_{45}^{\tau_0}$ (blue) of $A_k^{\tau_0}$ and the covering $Q_{45}^{\tau_1}$ (red) of $A_k^{\tau_1}$, both obtained by Algorithm 1. Using these coverings, we compute an initial covering $Q_{40}^{\tau_2}$ obtained via Algorithm 3 (gray). This covering is an acceptable initial guess for $A_k^{\tau_2}$ (green) which is illustrated in Figure 3.4 (b).



Figure 3.4: Illustration of the set-oriented predictor step by means of the Mackey-Glass equation for $\tau_0 = 6.000$, $\tau_1 = 6.05$ and $\tau_2 = 6.10$. (a) Using the coverings $Q_{45}^{\tau_0}$ (blue) of $A_k^{\tau_0}$ and $Q_{45}^{\tau_1}$ (red) of $A_k^{\tau_1}$ we obtain an initial guess $Q_{40}^{\tau_2}$ (gray) for $A_k^{\tau_2}$ by Algorithm 3. (b) In addition to the initial guess $Q_{40}^{\tau_2}$ we show the box covering $Q_{45}^{\tau_2}$ (green) obtained by Algorithm 1. Observe that our initial guess perfectly covers the approximation of $A_k^{\tau_2}$.

Now let us assume we have obtained an initial guess $Q_L^{\lambda_{j+1}}$ of $A_k^{\lambda_{j+1}}$ by Algorithm 3. Then it remains to define a set-oriented corrector step, which takes the dynamics depending on λ_{j+1} into account. In Algorithm 2 we applied $m - L$ steps of the original scheme 1 in order to achieve the same approximation quality. Here, however, we want to accelerate the computation by proposing a modified selection step (cf. (3.4)). Instead of performing $m - L$ subdivision and selection steps we subdivide the current box collection $m - L$ times and then do *one* selection step with an appropriate power of $\varphi_{\lambda_{j+1}}$, that is, we remove those boxes whose preimage under $\varphi_{\lambda_{j+1}}^{m-L}$ does neither intersect itself nor any other box. Analogously to the original selection step (3.4), we will numerically realize this modified step using test points (see Remark 3.3 (b)). We summarize this idea in Algorithm 4.

Algorithm 4: Set-Oriented Corrector Step

Initialization: Let

$$Q_L^{\lambda_{j+1}} = \bigcup_{B \in \mathcal{B}_L^{j+1}} B$$

be an initial guess for the approximation of $A_k^{\lambda_{j+1}}$.

- 1) **Refinement:** Define $\widehat{\mathcal{B}}_L^{j+1} = \mathcal{B}_L^{j+1}$ and make $m - L$ subdivision steps on $\widehat{\mathcal{B}}_L^{j+1}$ in order to construct a collection $\widehat{\mathcal{B}}_m^{j+1}$ such that

$$\bigcup_{B \in \widehat{\mathcal{B}}_m^{j+1}} B = \bigcup_{B \in \mathcal{B}_L^{j+1}} B$$

and

$$\text{diam}(\widehat{B}_m^{j+1}) = \theta^{m-L} \text{diam}(B_L^{j+1}), \quad \text{for some } 0 < \theta < 1.$$

- 2) **Modified Selection Step:** Define the new collection \mathcal{B}_m^{j+1} by

$$\mathcal{B}_m^{j+1} = \left\{ B \in \widehat{\mathcal{B}}_m^{j+1} \mid \exists \widehat{B} \in \widehat{\mathcal{B}}_m^{j+1} \text{ such that } \varphi_{\lambda_{j+1}}^{-(m-L)}(B) \cap \widehat{B} \neq \emptyset \right\} \quad (3.15)$$

and set $Q_m^{\lambda_{j+1}} = \bigcup_{B \in \mathcal{B}_m^{j+1}} B$.

Finally, alternating between both Algorithms 3 and 4 results into a novel *set-oriented predictor-corrector method* (SOPCM) which we summarize in Algorithm 5.

Algorithm 5: Set-Oriented Predictor-Corrector Method (SOPCM)

Initialization: Let $\lambda_0 < \lambda_1 < \dots < \lambda_N$ be a discretization of Λ and let $m, L \in \mathbb{N}$ such that $L \leq m$. Approximate $A_k^{\lambda_0}$ and $A_k^{\lambda_1}$ by $Q_m^{\lambda_0}$ and $Q_m^{\lambda_1}$, respectively, using Algorithm 1.

SOPCM: For $j = 1, \dots, N - 1$ repeat the following two steps.

- 1) **Predictor Step:** Apply Algorithm 3 on $Q_m^{\lambda_{j-1}}$ and $Q_m^{\lambda_j}$ in order to obtain an initial guess $Q_L^{\lambda_{j+1}}$ for the approximation of $A_k^{\lambda_{j+1}}$.
 - 2) **Corrector Step:** Use Algorithm 4 on $Q_L^{\lambda_{j+1}}$ to get an approximation $Q_m^{\lambda_j}$ of $A_k^{\lambda_{j+1}}$.
-

3.3 The Continuation Method

In the last sections we presented an algorithm for the approximation of (one-to-one images of) the (relative) global attractor of a (infinite-dimensional) dynamical system. Applying this method to a small neighborhood of an unstable steady state allows the approximation of the corresponding local unstable manifold. Hence, combining the subdivision scheme 1 with a *continuation* step [DH96] extends this local manifold to larger parts of the (global) unstable manifold. Moreover, we will prove that the developed algorithm actually approximates the closure of the (global) unstable manifold in the limit of infinitely small boxes, provided that the manifold is attractive. We note that large parts of this section are also contained in [ZDG19] to which the author has made substantial contributions.

For the sake of clarity we will only consider the CDS φ (2.32) as the underlying dynamical system and present an algorithm for the approximation of so-called *embedded* invariant manifolds. In the setting where φ is an arbitrary diffeomorphism on \mathbb{R}^k we refer the reader to [DH96]. However, the following underlying ideas are based on this work as well.

Given an unstable steady state $u^* \in \mathcal{A}$ of Φ (2.30) we denote by

$$\mathcal{W}_{\Phi}^u(u^*) \subseteq \mathcal{A} \tag{3.16}$$

the unstable manifold of $u^* \in \mathcal{A}$ and we define the *embedded unstable manifold* $W^u(p)$ as the image of $\mathcal{W}_{\Phi}^u(u^*)$ under the observation map R , that is,

$$W^u(p) = R(\overline{\mathcal{W}_{\Phi}^u(u^*)}) \subseteq A_k,$$

where $p = R(u^*) \in \mathbb{R}^k$. Note that we assume $\mathcal{W}_{\Phi}^u(u^*) \subseteq \mathcal{A}$ which implies $W^u(p) \subseteq A_k$. In particular, this implies that $W^u(p)$ is an invariant set of φ by construction (cf. (2.31)). The assumption made in (3.16) is automatically satisfied if $\mathcal{A} \subseteq \mathcal{X}$ is the compact global attractor of Φ .

The aim in this section is to develop an algorithm for the approximation of compact subsets of $W^u(p)$ or even the entire closure $\overline{W^u(p)}$ via subdivision and continuation. To this end, we denote by $\mathcal{W}_{\Phi,loc}^u(u^*) \subseteq \mathcal{A}$ the local unstable manifold of the steady state u^* and choose a compact neighborhood $C \subseteq A_k$ such that

$$W_{loc}^u(p) = R(\overline{\mathcal{W}_{\Phi,loc}^u(u^*)}) \subseteq C.$$

We will call $W_{loc}^u(p)$ *embedded local unstable manifold*. It turns out that $W_{loc}^u(p)$ is compact since R is continuous and $\overline{\mathcal{W}_{\Phi,loc}^u(u^*)}$ is compact as a closed subset of a compact set $C \subseteq \mathbb{R}^k$.

Proposition 3.21 ([ZDG19, Proposition 3.1]).

(a) Let A_C be the global attractor relative to C . Then

$$W_{loc}^u(p) \subseteq A_C.$$

(b) Suppose $\overline{\mathcal{W}_{\Phi,loc}^u(u^*)}$ is a compact attracting set with basin of attraction $\mathcal{U} \supset \overline{\mathcal{W}_{\Phi,loc}^u(u^*)}$. If $C \subseteq A_k$ is chosen such that $W_{loc}^u(p) \subseteq C \subseteq A_k$ and $E(C) \subseteq \mathcal{U}$, then

$$W_{loc}^u(p) = A_C.$$

Proof.

- (a) By Lemma 3.1 it suffices to show that $W_{loc}^u(p) \subseteq \varphi(W_{loc}^u(p))$ which is true by the construction of the CDS (cf. (2.31)). In fact, by continuity of R , Φ and E we have

$$W_{loc}^u(p) = R(\overline{\mathcal{W}_{\Phi,loc}^u(u^*)}) \subseteq R(\Phi(\overline{\mathcal{W}_{\Phi,loc}^u(u^*)})) = \varphi(W_{loc}^u(p)).$$

- (b) Recall the definition of φ_m and A_C^∞ (see (3.8) and (3.9) where $C = Q$). Then by Proposition 3.11 (b) we obtain $W_{loc}^u(p) = A_C^\infty$ and it remains to show that $A_C^\infty = A_C$. Since $C \subseteq A_k$ and (2.31), it follows that

$$\varphi_m(C) = \varphi^m(C) \quad \text{for all } m \in \mathbb{N}$$

and relabeling the indices in definition of A_C^∞ and A_C yields the desired statement, that is,

$$A_C^\infty = \bigcap_{m \geq 1} \bigcap_{j \geq 0} \varphi_m^j(C) = \bigcap_{m \geq 1} \bigcap_{j \geq 0} \varphi^{jm}(C) = \bigcap_{i \geq 0} \varphi^i(C) = A_C.$$

□

Remark 3.22.

- (a) Observe that Proposition 3.21 (b) states that the embedding of the local unstable manifold is identical to the local embedded unstable manifold, i.e., the global attractor relative to C . In particular, this implies that W_{loc}^u can be approximated by the subdivision scheme 1.
- (b) The assumed properties on the compact set $C \subseteq A_k$ are mild in the sense that, if the steady state $u^* \in \mathcal{A}$ is hyperbolic, then $\overline{\mathcal{W}_{\Phi,loc}^u(u^*)}$ is attractive since by assumption its dimension is finite (cf. (3.16)).

In the following we assume that the assumptions of Proposition 3.21 (b) are satisfied. Hence, Algorithm 1 generates a covering of $W_{loc}^u(p)$ which will be globalized in the upcoming continuation algorithm in order to obtain an approximation of the entire embedded unstable manifold $W^u(p)$.

Let $Q \subseteq \mathbb{R}^k$ be (large) compact set containing $p = R(u^*) \subseteq \mathbb{R}^k$ and we assume $\overline{W^u(p)} \subseteq Q$ for simplicity. If this is not satisfied, then it can in general not be guaranteed that we approximate the entire set $W^u(p)$ or even $\overline{W^u(p)} \cap Q$ (see [ZDG19, Remark 3.5 (a)] for details). The purpose of the continuation method is to approximate subsets $W_j \subseteq W^u(p)$ where $W_0 = W_{loc}^u(p)$ and

$$W_{j+1} = \varphi(W_j) \quad \text{for } j = 0, 1, 2, \dots$$

Thus, in a first step we approximate W_0 using Algorithm 1 where we realize the subdivision process using a family \mathcal{P}_n of partitions of Q . Here, a partition \mathcal{P} of Q is a finite family of compact subsets of Q such that

$$\bigcup_{B \in \mathcal{P}} B = Q \quad \text{and} \quad \text{int } B \cap \text{int } B' = \emptyset \quad \text{for all } B, B' \in \mathcal{P}, B \neq B',$$

and we consider a nested sequence \mathcal{P}_n of successively finer partitions of Q , such that for

all $B \in \mathcal{P}_n$ there exist $B_1, \dots, B_m \in \mathcal{P}_{n+1}$ such that

$$B = \bigcup_{i=1}^m B_i \text{ and } \text{diam}(B_i) \leq \theta \text{diam}(B)$$

for some $0 < \theta < 1$. Moreover, we denote by $\mathcal{P}_n(x) \in \mathcal{P}_n$ the element of \mathcal{P}_n containing $x \in Q$.

In what follows, we assume that $C = \mathcal{P}_s(p) \subseteq A_k$ for s sufficiently large such that $p \in \text{int } C$. Hence, performing $\ell \in \mathbb{N}_0$ steps of [Algorithm 1](#) gives an approximation $C_0^{(s,\ell)}$ of $W_0 = W_{\text{loc}}^u(p)$ by [Proposition 3.21](#) (b), where we assume that [Algorithm 1](#) constructs box collections that are contained in the partitions \mathcal{P}_n , $n > s$. Then in order to compute W_j we introduce a continuation step that iteratively adds those boxes $B \in \mathcal{P}_{s+\ell}$ to $C_j^{(s,\ell)}$ which satisfy $B \cap \varphi(\widehat{B}) \neq \emptyset$ for some $\widehat{B} \in \mathcal{P}_{s+\ell}$, starting with $C_0^{(s,\ell)}$.

We summarize the proposed method for the approximation of embedded unstable manifolds in [Algorithm 6](#).

Algorithm 6: The continuation method for embedded unstable manifolds

Initialization: Choose a compact set $Q \subseteq \mathbb{R}^k$ such that $\overline{W^u(p)} \subseteq Q$ and $p \in Q$. Let \mathcal{P}_n be a sequence of partitions of Q and choose $s \in \mathbb{N}$ and $C \in \mathcal{P}_s$ such that $p \in C$. Fix $\ell \in \mathbb{N}_0$.

- 1) **Subdivision Step:** Perform ℓ steps of [Algorithm 1](#) on $\mathcal{B}_0 = \{C\}$ to obtain a collection $\mathcal{B}_\ell \subseteq \mathcal{P}_{s+\ell}$ that covers W_{loc}^u and set $C_0^{(s,\ell)} = \mathcal{B}_\ell$.
- 2) **Continuation Step:** For $j = 0, 1, 2, \dots$ define

$$C_{j+1}^{(s,\ell)} = \left\{ B \in \mathcal{P}_{s+\ell} \mid \exists \widehat{B} \in C_j^{(s,\ell)} \text{ such that } B \cap \varphi(\widehat{B}) \neq \emptyset \right\}. \quad (3.17)$$

Remark 3.23.

- (a) In practice, it is not necessary to perform the subdivision step 1 in [Algorithm 6](#). If we desire a better approximation of the unstable manifold it is sufficient to increase s , i.e., to perform the computation in a finer partition, while keeping $\ell = 0$ fixed.
- (b) Again we have to decide whether the intersection of two sets is non-empty which is replaced by

$$\varphi(x) \in B \text{ for at least one test point } x \in \widehat{B}$$

in the numerical realization of the continuation step (3.17) (cf. [Remark 3.3](#) (b)).

- (c) If φ (resp. Φ) is a time- T -map of a continuous time system the continuation step (3.17) in Algorithm 6 can be modified in order to avoid “holes” in the approximation of the unstable manifold. To this end, one defines a finite time grid $\{t_1, \dots, t_N\}$ with $t_N = T$ and gathers all the boxes that are visited at each time instance t_i to the collection $\mathcal{C}_{j+1}^{(s,\ell)}$. Note that this procedure does not increase the numerical effort since, typically, a time- T -map is realized using a numerical integrator with some step size. In addition to that, it is possible to approximate the manifold by performing only one continuation step but with a (very) large number of test points and a (very) large integration time T .

Intuitively, it is clear that

$$C_j^{(s,\ell)} = \bigcup_{B \in \mathcal{C}_j^{(s,\ell)}} B$$

will cover W_j and the larger s and ℓ are chosen the better the approximation should be. Observe that $C_j^{(s,\ell)}$ forms a nested sequence in s and ℓ , i.e.,

$$C_j^{(0,\ell)} \supset C_j^{(1,\ell)} \supset \dots \supset C_j^{(s,\ell)} \dots \quad \text{and} \quad C_j^{(s,0)} \supset C_j^{(s,1)} \supset \dots \supset C_j^{(s,\ell)} \dots \quad (3.18)$$

In fact, by the continuity of φ it is also a nested sequence in j , i.e.,

$$C_0^{(s,\ell)} \subseteq C_1^{(s,\ell)} \dots \subseteq C_j^{(s,\ell)} \dots$$

Due to the compactness of Q this procedure will terminate after finitely many, say J_ℓ , steps and we denote the final covering by

$$G^{(s,\ell)} = \bigcup_{j=0}^{J_\ell} C_j^{(s,\ell)} = C_{J_\ell}^{(s,\ell)}$$

which covers $\overline{W^u(p)} \subseteq A_k$. In the following proposition we prove that the continuation method 6 converges.

Proposition 3.24. *Let $s \in \mathbb{N}$ be fixed.*

- (a) *The sets $C_j^{(s,\ell)}$ cover W_j for all $j, \ell = 0, 1, \dots$. Moreover, for fixed j , we have*

$$\bigcap_{\ell=0}^{\infty} C_j^{(s,\ell)} = W_j.$$

- (b) *Suppose that $\overline{W^u(p)}$ is linearly attractive, i.e., there is a $\lambda \in (0, 1)$ and a neighborhood $U \supset Q \supset \overline{W^u(p)}$ such that*

$$\text{dist}(\varphi(y), \overline{W^u(p)}) \leq \lambda \text{dist}(y, \overline{W^u(p)}) \quad \forall y \in U. \quad (3.19)$$

Then the final covering obtained by Algorithm 6 converges to the closure of the embedded unstable manifold $\overline{W^u(p)}$ for $\ell \rightarrow \infty$. That is,

$$\bigcap_{\ell=0}^{\infty} G^{(s,\ell)} = \overline{W^u(p)}.$$

Proof.

(a) In step 2) of Algorithm 1 we obtain a collection $\mathcal{B}_\ell \subseteq \mathcal{P}_{s+\ell}$ such that

$$A_C \subseteq Q_\ell = \bigcup_{B \in \mathcal{B}_\ell} B,$$

which yields the first statement by definition of the continuation step (3.17). For the second statement we note that by Proposition 3.21 (b) $A_C = W_{loc}^u(p) = W_0$ and by Proposition 3.7 the covering Q_ℓ converges to A_C for $\ell \rightarrow \infty$. Finally, since j is fixed a continuity argument shows that the sets $C_j^{(s,\ell)}$ converge to W_j for $\ell \rightarrow \infty$, i.e.,

$$\bigcap_{\ell=0}^{\infty} C_j^{(s,\ell)} = W_j.$$

(b) For every $\ell \in \mathbb{N}_0$ Algorithm 6 yields a covering $G^{(s,\ell)}$ of $\overline{W^u(p)}$ and therefore

$$\bigcap_{\ell=0}^{\infty} G^{(s,\ell)} \supset \overline{W^u(p)}.$$

For the other inclusion suppose there is $x \in \bigcap_{\ell=0}^{\infty} G^{(s,\ell)} \setminus \overline{W^u(p)}$. Since $\overline{W^u(p)}$ is compact it follows that $\text{dist}(x, \overline{W^u(p)}) > 0$. By definition of x , for every $\ell \geq 0$ Algorithm 6 generates a $\text{diam}(\mathcal{B}_\ell)$ -pseudo orbit $\{x_0, \dots, x_{j(\ell)}\}$, where $x_{j(\ell)} = x$ (cf. Remark 3.5). That is,

$$x_j \in C_j^{(s,\ell)} \text{ and } \varphi(x_j) \in \mathcal{P}_{s+\ell}(x_{j+1}) \quad \forall j \in \{0, \dots, j(\ell) - 1\}.$$

Here $\mathcal{P}_{s+\ell}(x_{j+1}) \subseteq C_{j+1}^{(s,\ell)}$ denotes the unique element of $\mathcal{P}_{s+\ell}$ containing $x_{j+1} \in C_{j+1}^{(s,\ell)}$ and $j(\ell) = \min \left\{ j \in \{0, \dots, J_\ell\} \mid x \in C_j^{(s,\ell)} \right\}$, i.e., $x \in C_j^{(s,\ell)}$ for $j \geq j(\ell)$ continuation steps. Observe that the sequence $j(\ell)$ is monotonically increasing in ℓ (cf. (3.17) and (3.18)) and

$$\|x_j - \varphi(x_{j-1})\| \leq \text{diam}(B_\ell) \quad \forall j \in \{0, \dots, j(\ell) - 1\}. \quad (3.20)$$

We first show by contradiction that $j(\ell)$ is unbounded. Suppose that $j(\ell)$ is bounded by some $J \in \mathbb{N}_0$, i.e., $\max j(\ell) = J$. Hence, by monotony of $j(\ell)$ there is $\ell_0 \in \mathbb{N}_0$ such that $j(\ell) = J$ for all $\ell \geq \ell_0$ and by (3.18) we have

$$x \in \bigcap_{\ell=0}^{\infty} C_{j(\ell)}^{(s,\ell)} = \left(\bigcap_{\ell=0}^{\ell_0-1} C_{j(\ell)}^{(s,\ell)} \right) \cap \left(\bigcap_{\ell=\ell_0}^{\infty} C_J^{(s,\ell)} \right) \subseteq \bigcap_{\ell=\ell_0}^{\infty} C_J^{(s,\ell)} = \bigcap_{\ell=0}^{\infty} C_J^{(s,\ell)}.$$

However, by Proposition 3.24 (a) it follows that $x \in W_J \subseteq \overline{W^u(p)}$ which is a contradiction to $\text{dist}(x, \overline{W^u(p)}) > 0$. Thus, $j(\ell)$ must be unbounded. By assumption $\overline{W^u(p)}$ is linearly attractive in a neighborhood U . Hence, we can use (3.19) and (3.20) on the $\text{diam}(\mathcal{B}_\ell)$ -pseudo orbit $\{x_0, \dots, x_{j(\ell)}\}$ in combination with the triangle

inequality to obtain

$$\begin{aligned}
 \text{dist}(x, \overline{W^u(p)}) &\leq \text{dist}(\varphi(x_{j(\ell)-1}), \overline{W^u(p)}) + \text{diam}(\mathcal{B}_\ell) \\
 &\leq \lambda \text{dist}(x_{j(\ell)-1}, \overline{W^u(p)}) + \text{diam}(\mathcal{B}_\ell) \\
 &\vdots \\
 &\leq \lambda^{j(\ell)} \text{dist}(x_0, \overline{W^u(p)}) + \text{diam}(\mathcal{B}_\ell) \sum_{i=0}^{j(\ell)-1} \lambda^i \\
 &\leq \lambda^{j(\ell)} \text{dist}(x_0, \overline{W^u(p)}) + \frac{\text{diam}(\mathcal{B}_\ell)}{1-\lambda} \longrightarrow 0 \text{ for } \ell \rightarrow \infty.
 \end{aligned}$$

Here, the last expression converges to zero because $\lambda \in (0, 1)$ and $\text{diam}(\mathcal{B}_\ell)$ converges to zero for $\ell \rightarrow \infty$ (cf. (3.3)). Again we have a contradiction to $\text{dist}(x, \overline{W^u(p)}) > 0$ and it follows that

$$\bigcap_{\ell=0}^{\infty} G^{(s,\ell)} \subseteq \overline{W^u(p)},$$

which yields the desired statement. □

Remark 3.25.

- (a) The assumption in Proposition 3.24 (b) is not satisfied if, for instance, $\overline{W^u(p)}$ forms a heteroclinic connection between the steady state p and another unstable hyperbolic steady state q . In this case the algorithm would also generate a covering of the embedded unstable manifold of q since its embedded local unstable manifold $W_{\text{loc}}^u(q)$ will eventually be covered by the continuation step (3.17).
- (b) If (3.19) is not satisfied, but $\overline{\mathcal{W}_\Phi^u(u^*)}$ is attractive, one can apply the subdivision scheme 1 to G_ℓ in order to approximate $\overline{W^u(p)}$ more accurately (cf. Proposition 3.11).

Finally, we illustrate the continuation method 6 in Figure 3.5 on the Lorenz system (cf. Example 3.9) and compute the unstable manifold of

$$x_0 = (\sqrt{\beta(\rho - 1)}, \sqrt{\beta(\rho - 1)}, \rho - 1) \quad (3.21)$$

using the time- T -map of (3.6) for $T = 0.2$. As the initial level of the the partition we choose $s = 27$ and skip step 1) of Algorithm 6, i.e., we perform $\ell = 0$ steps of Algorithm 1.

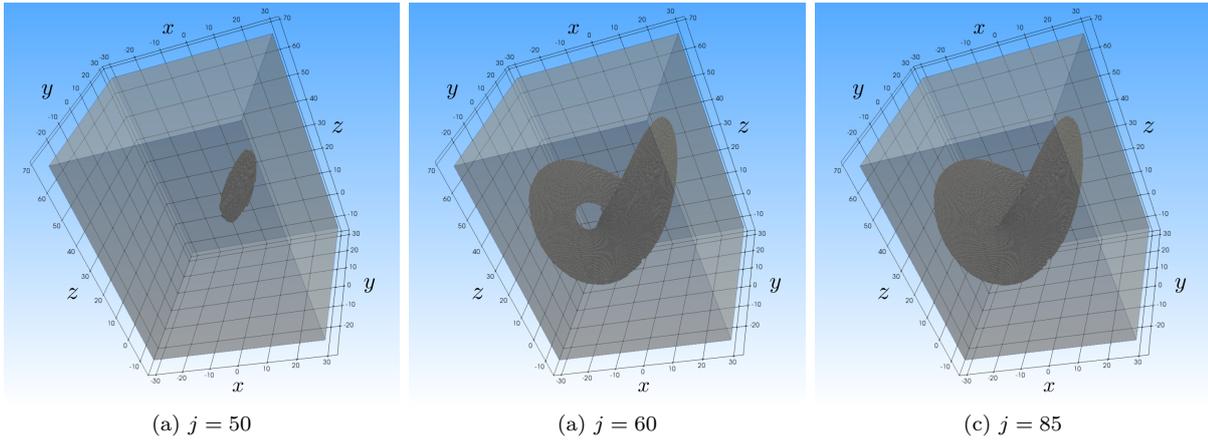


Figure 3.5: Computation of the unstable manifold of the unstable equilibrium (3.21) of the Lorenz system (3.6) using Algorithm 6. (a) For up to $j = 10$ steps the unstable manifold grows like a disc. (b) After $j = 60$ the other “wing” is reached. (d) The computation is finished after $j = 85$ steps. Observe that the obtained covering is the same as for the subdivision scheme in Figure 3.1 (c).

4 The Choice of the Observation Map

As mentioned in Section 2.4 the CDS heavily depends on the particular choice of the observation map R or the observable f when using delay coordinates. In addition to that, the map E is not explicitly known and has to be designed according to the problem at hand. This is why in the first part of this chapter we will present possible (linear) choices for delay differential equations [DHZ16] and partial differential equations, where the latter is also contained in [ZDG19] to which the author has made substantial contributions.

From now on we assume that upper bounds for both the box-counting dimension $d_{\text{box}}(\mathcal{A})$ and the thickness exponent $\tau(\mathcal{A}) < 1$ are available. This allows us to fix $k \in \mathbb{N}$ according to Theorem 2.21 or 2.26. In order to numerically realize the CDS $\varphi = R \circ \Phi \circ E$ as described in Proposition 2.28, we have to choose a map R , design an appropriate map E such that (2.31) holds at least approximately and numerically implement the time- T -map Φ . For the latter we rely on standard tools for the forward time integration of DDEs [BZ13] and PDEs, e.g., a fourth-order time stepping method for the one-dimensional Kuramoto-Sivashinsky equation [KT05]. The map R will be linear and bounded, and for the numerical construction of the map E we will employ a bootstrapping method that re-uses results of previous computations. In this way we will in particular ensure that the identities in (2.31) are at least approximately satisfied.

We will see that with a linear observation map R it is easy to define a map E such that the second required condition $(R \circ E)(x) = x$ is not only satisfied for every point $x \in R(\mathcal{A})$ (cf. (2.31)) but also for every $x \in \mathbb{R}^k$. In contrast, the first condition $(E \circ R)(u) = u$ for all $u \in \mathcal{A}$ does, in general, not hold. However, since we want to apply the CDS to the numerical schemes described in Chapter 3, we will enforce this identity utilizing the following observation (cf. Remark 3.5).

Remark 4.1.

- (a) Taking a closer look at the selection step (3.4) we see that, if a box $B \in \mathcal{B}_\ell$ then there must have been a box $\widehat{B} \in \mathcal{B}_{\ell-1}$ such that $\bar{x} = R(\Phi(E(\widehat{x}))) \in B$ for at least one point $\widehat{x} \in \widehat{B}$.
- (b) Analogously, according to the continuation step (3.17) we can conclude that, if a box $B \in \mathcal{C}_{j+1}^{(s,\ell)}$ then there must have been a box $\widehat{B} \in \mathcal{C}_j^{(s,\ell)}$ such that $\bar{x} = R(\Phi(E(\widehat{x}))) \in B$ for at least one point $\widehat{x} \in \widehat{B}$.

Therefore, we can use the information on the trajectories of $\Phi(E(\widehat{x}))$ that were mapped into B by R in the previous step for a good approximation of $(E \circ R)$ on \mathcal{A} .

In the second part of this chapter we turn our attention to one particular nonlinear observation map. Driven by the desire to obtain further intuitive understanding of the geometric structure of the (embedded) invariant set (and thus hopefully also of the corresponding dynamics on them), the aim will be to identify nonlinear coordinates revealing their intrinsic geometry in the embedding space. To this end, we will discuss a tool called *diffusion maps* that allows us to obtain the geometric and dynamical structure of the covering of an embedded invariant set. Diffusion maps are one among many data-driven manifold learning techniques that find intrinsic coordinates of a data set [TDSL00, RS00, DG03, BN03, ZZ04]. First introduced by Coifman and Lafon [CL06a, CLL⁺05], diffusion maps is a nonlinear feature extraction algorithm that

computes a family of embeddings of a (possibly) high-dimensional data set into a low-dimensional space, whose coordinates are given by the eigenvectors and eigenvalues of a diffusion operator on the data. Different from linear dimensionality reductions methods, such as principal component analysis (POD, cf. Section 4.1.2), diffusion maps focus on discovering the underlying manifold from which the data set is sampled. For instance, this tool has been successfully applied to data from turbulent Rayleigh-Bénard convection revealing the large-scale flow structure [KW20]. Moreover, the algorithm is robust to noise perturbation such that it can deal with the outer approximations that cover the set of interest generated by the set-oriented numerical methods.

Ideally, one would like to approximate the set of interest in its intrinsic coordinate right away using the diffusion maps as the observation map R . However, at the current state this is not possible and we leave this to future research. Some ideas towards this will be given in the outlook in Chapter 6. Thus, in this thesis we will compute the (embedded) invariant set using a linear observation map and afterwards apply diffusion maps on the generated data.

4.1 Linear Operators

To begin with we will start with linear observations since they can quite easily be realized numerically and there is a natural approach for finding an inverse. Typically, in order to solve an infinite-dimensional system such as a PDE or DDE, the underlying state space is discretized, i.e., the Banach space \mathcal{X} is replaced by \mathbb{R}^N for some (large) $N \in \mathbb{N}$. Thus, for now let us consider a linear map $R : \mathbb{R}^N \rightarrow \mathbb{R}^k$, that is, by exploiting the notation, R can be seen as a matrix in $\mathbb{R}^{k \times N}$ where $N \gg k$. Then the following remark will be useful in the design of an appropriate map $E : \mathbb{R}^k \rightarrow \mathbb{R}^N$, respectively matrix $E \in \mathbb{R}^{N \times k}$.

Remark 4.2. Suppose $R \in \mathbb{R}^{k \times N}$ has full rank, then the *pseudo inverse*

$$E := R^+ = R^T(RR^T)^{-1}$$

satisfies $RE = \mathbf{I}_k$. In particular, E as a mapping $\mathbb{R}^k \rightarrow \mathbb{R}^N$ fulfills $(R \circ E)(x) = x$ for all $x \in \mathbb{R}^k$.

4.1.1 Point Evaluation for DDEs

A simple, yet effective, observable is the evaluation of a function at prescribed points in space and/or time. It turns out that this choice perfectly fits in the framework of delay differential equations (DDEs). In contrast to ordinary differential equations, for DDEs (also called time-delay systems) the derivative with respect to time of the unknown function depends on the state at the current and, additionally, at previous times [Kua93]. Thus, for numerically solving a DDE an initial history over a time interval, which serves as an initial function, has to be given. Hence, the solution operator will be defined on a function space generating an infinite-dimensional dynamical system. Typically, DDEs are used for modeling dynamical phenomena that require time-delayed aftereffects, e.g., for applications in population dynamics, epidemiology and mechanics [Kua93, NP02, Del20, AHD07, KM13].

In this section we consider a DDE with a (small) delay of the form

$$\begin{aligned} \dot{y}(t) &= g(y(t), y(t - \tau)), & 0 \leq t \leq t_f, \\ y(t) &= y_0(t), & t \leq 0, \end{aligned} \tag{4.1}$$

where $y(t) \in \mathbb{R}^n$, $\tau > 0$ is a constant time delay and $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a smooth map. For the case where the time delay depends on the state we refer to [Zie18]. For simplicity, we will only consider the one-dimensional case $n = 1$ and refer to [DHZ16] for $n > 1$.

Following [HL93] we denote by $\mathcal{X} = C([- \tau, 0], \mathbb{R}^n)$ the (infinite-dimensional) state space of (4.1) which is indeed a Banach space when equipped with the maximum norm. For a given initial condition $u \in \mathcal{X}$ let $y_u(t)$ be the trajectory generated by (4.1). Then the flow $\Phi^t : \mathcal{X} \rightarrow \mathcal{X}$ of (4.1) is given by

$$u \mapsto \Phi^t(u), \text{ where } \Phi^t(u)(s) = y_u(t + s) \text{ for } s \in [-\tau, 0].$$

Next, we choose and fix $T > 0$ as a natural fraction of τ , that is,

$$T = \frac{\tau}{K} \text{ for } K \in \mathbb{N}.$$

and consider the corresponding time- T -map $\Phi^T : \mathcal{X} \rightarrow \mathcal{X}$ as the infinite-dimensional dynamical system (2.30) with invariant set $\mathcal{A} \subseteq \mathcal{X}$.

Numerical Realization of R

For the definition of R we will make use of part (c) of Remark 2.29 and consider a delay coordinate embedding $R = D_k[f, \Phi]$. Thus, we have to specify an observable f and simply choose the point evaluation of $u \in \mathcal{X}$ at $-\tau$, that is,

$$f(u) = u(-\tau).$$

Hence, in this case the delay coordinate map $D_k[f, \Phi]$ is

$$D_k[f, \Phi](u) = (u(-\tau), \Phi(u)(-\tau), \dots, \Phi^{k-1}(u)(-\tau))$$

and, provided $k \in \mathbb{N}$ is large enough, $D_k[f, \Phi]$ is one-to-one on \mathcal{A} (cf. Theorem 2.26).

Remark 4.3.

- (a) A quite natural choice for K would be $K = k - 1$. In this case, the delay coordinate map extracts the point values of a function $u \in \mathcal{X}$ at k equally distributed time steps within the interval $[-\tau, 0]$. Thus, $D_k[f, \Phi]$ coincides with the observation map

$$R(u) = \left(u(-\tau), u\left(-\tau + \frac{\tau}{k-1}\right), u\left(-\tau + \frac{2\tau}{k-1}\right), \dots, u(0) \right), \tag{4.2}$$

which allows one to apply Theorem 2.21 instead and R would be a matrix whose k rows are appropriate canonical vectors in \mathbb{R}^N . We illustrate this observation map in Figure 4.1 (a).

- (b) Later on we will consider the time delay τ as a bifurcation parameter and vary $\tau \in \Lambda$ for some compact interval $\Lambda \subseteq \mathbb{R}_{\geq 0}$. In order to properly compare the numerical results, we need a universal state space as well as a common observation map for all $\tau \in \Lambda$. To this end, we fix $\bar{\tau} = \max \Lambda$ and set $\mathcal{X} = C([-\bar{\tau}, 0], \mathbb{R}^n)$. The observation map is then given by (4.2) with $\tau = \bar{\tau}$.

Numerical Realization of E

Given a point $x \in \mathbb{R}^k$ the image $E(x)$ has to be an adequate initial condition for the forward integration of the DDE (4.1). In addition to that we have to enforce the identities (2.31) (at least approximately), i.e., E has to be designed in such a way that it is the inverse of R on \mathcal{A} . Thus, according to the choice of R in Section 4.1.1 and Remark 4.2 a first idea for a function $u = E(x)$ is $R^+(x)$, that is,

$$u(t) = \begin{cases} x_i, & t = t_i \\ 0, & \text{otherwise} \end{cases} \quad (4.3)$$

for $t_i = -\tau + i \cdot T$, $i = 0, \dots, k-1$. However, we can improve this by constructing a piecewise linear or spline function $u = E(x)$ with $u(t_i) = x_i$ instead. By this choice the second condition $(R \circ E)(x) = x$ is also not only satisfied for every point $x \in R(\mathcal{A})$ (cf. (2.31)) but for every $x \in \mathbb{R}^k$. We show these three numerical realizations of the map E in Figure 4.1 (b).

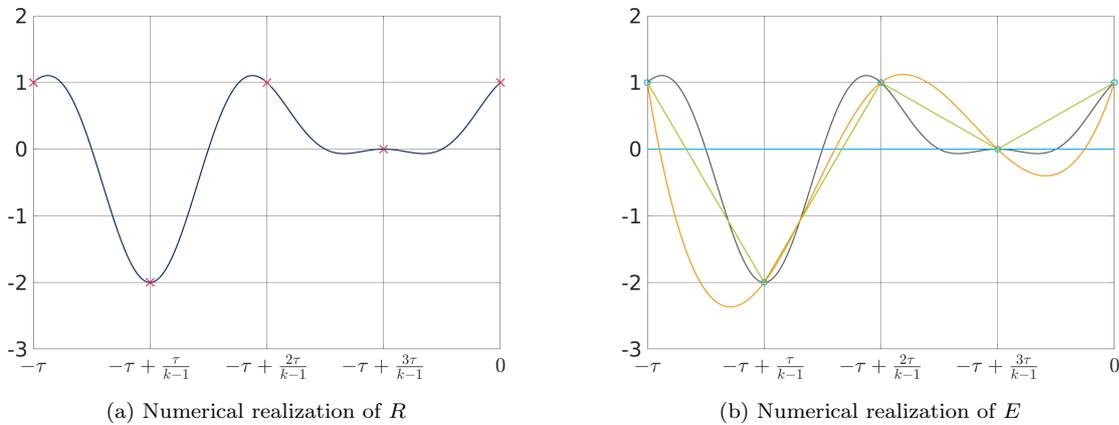


Figure 4.1: Illustration of the numerical realization of R and E for $k = 5$. (a) The observation map R extracts the value x_i of $u \in \mathcal{X}$ (dark blue curve) at $k = 5$ equally distributed time instances t_i (red crosses). (b) The original function $u \in \mathcal{X}$ is approximated by $R^+(x)$ (see (4.3)) as the light blue discontinuous function, by a green linear curve and a spline interpolation shown in orange. Note that all three realizations fail to reconstruct $u \in \mathcal{X}$ between the observed point values x_i .

In order to enforce the first condition $(E \circ R)(u) = u$ for all $u \in \mathcal{A}$ at least approximately, we utilize [Remark 4.1](#) and in every step of both procedures, for every box $B \in \mathcal{B}_\ell$ (resp. $B \in \mathcal{C}_{j+1}^{(s,\ell)}$) we save $k_i \geq 1$ additional equally distributed function values of $\Phi(E(\hat{x}))$ for each interval $(-\tau + (i-1)T, -\tau + iT)$, $i = 1, \dots, k-1$. This allows us to improve the construction of E for exactly those test points. When $\varphi(B)$ needs to be evaluated using test points in B , we first use the points in B for which additional information is available and generate the corresponding initial value functions via linear or spline interpolation. Note that the more information we store, i.e., the larger k_i , the smaller the error $\|\Phi(E(\hat{x})) - E(x)\|$ becomes for $x = R(\Phi(E(\hat{x})))$. That is, we enforce an approximation of the identity $(E \circ R)(u) = u$ for all $u \in \mathcal{A}$ (see (2.31)). If the additional information is available only for a few points in B , we generate new test points in B at random and construct the corresponding trajectories by interpolation as described before. In particular, this is the case in the first step of [Algorithms 1](#) and [6](#).

4.1.2 Coordinate Representation for PDEs

Another useful observation is the representation of a function with respect to a prescribed basis which for instance works for partial differential equations of the form

$$\frac{\partial}{\partial t}u(y, t) = \mathcal{F}(y, u), \quad u(y, 0) = u_0(y) \quad (4.4)$$

with suitable boundary conditions, where $u \in \mathcal{X}$ for some Banach space \mathcal{X} and \mathcal{F} is a (nonlinear) differential operator. Assuming that the PDE (4.4) has a well-defined semi-flow on \mathcal{X} we will consider the time- T -map of (4.4) as the dynamical system Φ in (2.30).

Numerical Realization of R

A widely used approach for constructing reduced order models is Galerkin projection [[KV99](#), [SV10](#)]. The general concept of such a projection is to find a finite-dimensional representation of an unknown function $u \in \mathcal{X}$ (see, e.g., [[HLBR12](#)]), i.e., we assume in what follows that the function $u \in \mathcal{X}$ can be represented in terms of an (ordered) basis $\{\psi_i\}_{i=1}^\infty$, i.e.,

$$u(y, t) = \sum_{i=1}^{\infty} x_i(t) \psi_i(y), \quad (4.5)$$

where the ψ_i are elements of a suitable space. Then the proposed observation is a function that simply maps $u \in \mathcal{X}$ onto the truncated *coordinate vector*, i.e.,

$$R(u) = (x_1, \dots, x_k) \in \mathbb{R}^k. \quad (4.6)$$

Observe that R is linear and bounded and hence, for k sufficiently large, [Theorem 2.21](#) guarantees that generically (in the sense of prevalence) R will be a one-to-one map on \mathcal{A} .

In order to find an adequate basis $\{\psi_i\}_{i=1}^\infty$, we summarize some key aspects that are required. First of all, every function $u \in \mathcal{X}$ should be represented exactly by $\{\psi_i\}_{i=1}^\infty$, i.e., the system must be complete. This representation should be unique, which follows from

linearly independence. In practice, an orthogonal or orthonormal system is beneficial. Every element ψ_i must satisfy the boundary conditions of (4.4). We note that these three properties have been discussed in the context of reduced order models for PDEs [Pei17]. Finally, we want a basis as small as possible, i.e., we desire to find a truncation of (4.5) to a small number S of basis function:

$$u(y, t) \approx \sum_{i=1}^S x_i(t) \psi_i(y).$$

Obviously, this contradicts the assumption that $\{\psi_i\}_{i=1}^\infty$ is complete. On the one hand we search for a basis that is capable of representing $u \in \mathcal{X}$ with an error as small as possible and on the other hand is of small size S . For the Hilbert space $\mathcal{X} = L^2(\Omega)$ for some spatial domain $\Omega \subseteq \mathbb{R}^n$ this problem is addressed by the *proper orthogonal decomposition* (POD) (cf. [Sir87, BHL93, Cha00, HLBR12]), also known as the *principal component analysis* or the *Karhunen-Loève transformation* and can be formulated as an optimization problem ([Row05, Vol11, Fah00]):

$$\begin{aligned} \min_{\psi_1, \dots, \psi_S \in L^2} \int_0^T \left\| u(\cdot, t) - \sum_{i=1}^S \langle u(\cdot, t), \psi_i \rangle_{L^2} \psi_i \right\|_{L^2}^2 dt \\ \text{s.t. } \langle \psi_i, \psi_j \rangle_{L^2} = \delta_{i,j}, \quad 1 \leq i, j, \leq S. \end{aligned}$$

That is, the average (squared) error between $u \in \mathcal{X}$ and its projection onto the space spanned by the basis function $\{\psi_i\}_{i=1}^S$ is minimized. In this sense, POD modes contain the “*most characteristic*” data from an ensemble of functions. This optimization problem is practically realized by discretizing time in $0 = t_0 < t_1 < \dots < t_r < T$ and taking r snapshots $u(\cdot, t_j)$ at these time instances [Sir87]. Hence, this approach is also called *method of snapshots*. Thereby we can equivalently transform the optimization problem (4.1.2) into

$$\begin{aligned} \max_{\psi_1, \dots, \psi_S \in L^2} \sum_{i=1}^S \frac{1}{r} \sum_{j=1}^r \langle u(\cdot, t_j), \psi_i \rangle_{L^2}^2 \\ \text{s.t. } \langle \psi_i, \psi_j \rangle_{L^2} = \delta_{i,j}, \quad 1 \leq i, j, \leq S. \end{aligned} \tag{4.7}$$

see, e.g., [HLBR12, Pei17]. In particular, for $S = 1$ this yields

$$\frac{1}{r} \langle u(\cdot, t_j), \psi_1 \rangle_{L^2}^2 = \langle \psi_1, \mathcal{R} \psi_1 \rangle_{L^2},$$

where the operator

$$\mathcal{R} \psi_1 = \int_{\Omega} \left(\frac{1}{r} \sum_{j=1}^r u(\cdot, t_j) u(y, t_j) \right) \psi_1(y) dy$$

is linear and self-adjoint. Thus, the operator \mathcal{R} , also called *two-point correlation* [Pop00], possesses a set of orthonormal eigenfunctions $\{\psi_i\}_{i=1}^r$ with associated positive eigenvalues $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$. It turns out that the $S < r$ eigenfunctions corresponding to the S largest eigenvalues of \mathcal{R} are precisely the solution of the optimization problem (4.7) (see

[BHL93] or [Fah00]), that is, $\{\psi_i\}_{i=1}^S$ forms a basis of size S which optimally represents the kinetic energy within the snapshots [HLBR12]. Moreover, the eigenvalues contain the amount of information that is neglected by truncating the basis to the size $S < r$ [Sir87]:

$$\varepsilon(S) := \frac{\sum_{i=1}^S \sigma_i}{\sum_{j=1}^r \sigma_j}. \quad (4.8)$$

What remains to discuss is how we can compute such a POD basis. To this end, we first generate r snapshots from a long-time simulation for some set of initial conditions of the underlying PDE. More precisely, we denote by $u_h(t_i) \in \mathbb{R}^{n_x}$, $t_i \in \{t_1, \dots, t_r\}$, the numerical solution of (4.4) defined on a finite-dimensional grid at n_x nodes obtained at r time instances. Then we arrange the data in the so-called *snapshot matrix*

$$S_M = \begin{pmatrix} | & & | \\ u_h(t_1) & \cdots & u_h(t_r) \\ | & & | \end{pmatrix} \in \mathbb{R}^{n_x \times r}.$$

Since there is a close relationship between the proper orthogonal decomposition (POD) and the singular values decomposition (SVD) [Cha00, LLL⁺02, Vol11] we perform a SVD of the matrix S_M and obtain

$$S_M = U \Sigma V^\top \quad (4.9)$$

where $U \in \mathbb{R}^{n_x \times n_x}$, $\Sigma \in \mathbb{R}^{n_x \times r}$ and $V \in \mathbb{R}^{r \times r}$. The columns of U give us a discrete representation of the POD modes ψ_i , whereas the diagonal elements σ_i of Σ can be utilized to determine the amount of information that is neglected by truncating the basis to size $S < r$, i.e., $\varepsilon(S)$ is close to one (cf. (4.8)). For many applications the eigenvalues decay fast such that a truncation to a small basis is possible.

Remark 4.4.

- (a) Since the POD basis is orthogonal we can compute the POD coefficients x_i for the observation map R in (4.6) by taking the scalar product of u and ψ_i , i.e.,

$$R(u) = (\langle u, \psi_1 \rangle, \dots, \langle u, \psi_k \rangle)$$

- (b) The matrix R is given by $\widehat{U} \in \mathbb{R}^{k \times n_x}$, where the i -th row of \widehat{U} is the i -th column of U in the SVD (4.9) for $i = 1, \dots, k$.

Numerical Realization of E

By Section 4.1.2 the state space of the CDS φ is given by the first k POD coefficients $x \in \mathbb{R}^k$. Analogous to the numerical realization of E for DDEs in Section 4.1.1, we have to define a function $u = E(x) \in \mathcal{X}$ and according to Remark 4.2 when no additional information (cf. Remark 4.1) is available we simply construct u by $u = R^+(x)$, i.e.,

$$E(x) = \sum_{i=1}^k x_i \psi_i.$$

Observe that in the language of matrices the rows of R are orthonormal and hence $R^+ = R^T$. By this choice we have $(R \circ E)(x) = x$ for all $x \in \mathbb{R}^k$. To enforce $(E \circ R)$ on \mathcal{A} we extend this expansion and augment the state space to \mathbb{R}^S using the remaining $S - k$ POD coefficients that we can compute. More precisely, we construct an initial function by

$$E(x) = \sum_{i=1}^k x_i \psi_i + \sum_{l=k+1}^S x_l \psi_l. \quad (4.10)$$

Since only the first k POD coefficients are given by the coordinates of points inside a box $B \subseteq \mathbb{R}^k$, it remains to discuss how to choose the POD coefficients x_{k+1}, \dots, x_S . The idea is to use a heuristic strategy that utilizes statistical information provided by earlier numerical integration.

Suppose we want to evaluate φ for a large number of test points in a box $B \in \mathbb{R}^k$. By Remark 4.1 we can compute the POD coefficients $\bar{x}_{k+1}, \dots, \bar{x}_S$ of $\bar{x} = R(\Phi(E(\hat{x}))) \in B$ by

$$\bar{x}_i = \langle \Phi(E(\hat{x})), \Psi_i \rangle, \quad i = k + 1, \dots, S.$$

Then we sample the box B with all points \bar{x} for which additional information is available and compute $E(\bar{x})$ according to (4.10). However, the number of these points \bar{x} might be too small, such that B is not discretized sufficiently well and we have to sample additional test points. To this end, we choose a certain number of points $\tilde{x} \in B \subseteq \mathbb{R}^k$ at random and extend these points to elements in \mathbb{R}^S as follows: We compute the componentwise mean value μ_i and the variance σ_i^2 of all POD coefficients \bar{x}_i , for $i = k + 1, \dots, S$ which allows us to make a Monte Carlo sampling for the additional coefficients of \tilde{x}_i for $i = k + 1, \dots, S$, i.e.,

$$\tilde{x}_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad \text{for } i = k + 1, \dots, S.$$

Thus, we obtain S POD coefficients that generate an initial function of the form in (4.10). By this construction we expect in each step to generate initial functions that satisfy an approximation of the identity $(E \circ R)(u) = u$ for all $u \in \mathcal{A}$.

4.2 Diffusion Maps

In the last section we presented two linear observation maps suitable for the construction of the CDS φ (2.32). On the one hand we considered a delay coordinate map where the observable is a point evaluation for DDEs and on the other hand we projected a solution of a PDE onto an appropriate basis and then observed the corresponding coefficients. Thus, combining embedding techniques with set-oriented numerical methods allows the computation of one-to-one images of attractors and manifolds of the infinite-dimensional dynamical system (2.30). However, the embedding can still be high-dimensional, even though the box-counting dimension is low ($k > 2d_{\text{box}}$). Thus, the embedded set is topologically uninformative and it is hard to identify geometrical features of the underlying attractor or manifold. To highlight these important features and possibly further decrease the embedding dimension we rely on feature extraction methods that are typically nonlinear. One particular example of such a nonlinear dimensionality reduction method is *diffusion maps* which we will consider in detail in the following. We note that some parts in this section are contained in [GKD19] to which the author has made substantial contributions.

First, we will review the diffusion maps method introduced in [CL06a] for our purposes. Suppose $X \subseteq \mathbb{R}^k$ is an approximation of the embedded attractor $A_k \subseteq \mathbb{R}^k$ or the embedded unstable manifold $W^u(p) \subseteq A_k$ and let $k : X \times X \rightarrow \mathbb{R}$ be a symmetric and positivity preserving *kernel*, that is,

$$k(x, y) = k(y, x) \text{ and } k(x, y) \geq 0 \quad \text{for all } x, y \in X. \quad (4.11)$$

For the construction of the diffusion maps any kernel k that satisfies (4.11) works, but since $X \subseteq \mathbb{R}^k$ it is tempting to use an isotropic (rotation-invariant) kernel, e.g.,

$$k(x, y) = h\left(\frac{\|x - y\|_2^2}{\varepsilon}\right) \quad (4.12)$$

for some carefully chosen bandwidth $\varepsilon > 0$. In our application we will use the Gaussian kernel, that is, $h(z) = \exp(-z)$, and we will discuss the choice of ε in Section 4.2.1. This kernel represents the similarity between points in X and one can think of them as being nodes in a symmetric graph (X, k) whose weight function is given by the kernel k . Hence, we can define a reversible Markov chain on X by the normalized graph Laplacian construction [CG97] given by

$$d(x) = \int_X k(x, y) dy \quad \text{and} \quad p(x, y) = \frac{k(x, y)}{d(x)}.$$

Then the new kernel p still is positivity preserving but no longer symmetric. However, it is now a transition kernel of a Markov chain on X since

$$\int_X p(x, y) dy = 1.$$

Hence, the operator P defined by

$$Pf(x) = \int_X p(x, y)f(y)dy$$

preserves constant functions $f : X \rightarrow \mathbb{R}$. It is easy to see that the Markov chain has a stationary distribution given by

$$\pi(y) = \frac{d(y)}{\int_X d(z)dz}.$$

If the graph is connected, which we from now on assume, it follows that the stationary distribution is unique. Furthermore, it is reasonable to assume that $k(x, x) > 0$ as this represents the affinity of x with itself and as a consequence the Markov chain is aperiodic. Since the graph is connected by assumption, it is also irreducible. In particular, it follows that the chain is ergodic. In addition to that, the chain is reversible and it follows the *detailed balance condition*

$$\pi(x)p(x, y) = \pi(y)p(y, x). \tag{4.13}$$

It turns out that the operator P contains geometric information about the underlying set X and one of the main ideas of the diffusion framework is advancing the Markov chain forward in time, that is, taking larger powers of P . Thus, for $t \geq 0$ we will consider the kernel of P^t which is given by $p_t(x, y)$ describing the probability of transition from x to y in t time steps. A classical way to describe the powers of P is to employ spectral theory, namely eigenvectors and eigenvalues. By (4.13) and a mild additional assumption (see (4.15)) we can prove that P has a discrete sequence of eigenvalues λ_i and eigenfunction ϕ_i , $i \geq 0$, such that

$$1 = \lambda_0 > |\lambda_1| \geq |\lambda_2| \geq \dots \tag{4.14}$$

In fact, conjugating p with $\sqrt{\pi}$ yields a kernel

$$a(x, y) = \frac{\sqrt{\pi(x)}}{\sqrt{\pi(y)}}p(x, y) = \frac{k(x, y)}{\sqrt{\pi(x)}\sqrt{\pi(y)}}$$

that is symmetric. The corresponding integral operator A with kernel a is therefore self-adjoint in $L^2(X)$ and also compact if we assume

$$\int_X \int_X a(x, y)^2 dy dx < \infty \tag{4.15}$$

(see, e.g., [Ped99]). Observe that this condition is always satisfied if X is finite, that is, X is a finite discretization of the embedded invariant set, as A is just a matrix. By compactness A possesses a discrete set of eigenvalues $\{\lambda_i\}_{i \geq 0}$ and an orthonormal set of eigenfunction $\{\phi_i\}_{i \geq 0}$ forming a basis of $L^2(X)$, where $\phi_0 = \sqrt{\pi}$. Thus,

$$a(x, y) = \sum_{i \geq 0} \lambda_i \phi_i(x) \phi_i(y)$$

and defining $\psi_i(x) = \phi_i(x)/\sqrt{\pi(x)}$, in particular $\psi_0(x) = 1$, and $\varphi_i(y) = \phi_i(y)\sqrt{\pi(y)}$ we have

$$p(x, y) = \sum_{i \geq 0} \lambda_i \psi_i(x) \varphi_i(y).$$

For the kernel of the powers P^t of P this yields

$$p_t(x, y) = \sum_{i \geq 0} \lambda_i^t \psi_i(x) \varphi_i(y). \quad (4.16)$$

Furthermore, by [Ped99] $\lambda_0 = 1$ and since the chain is ergodic, apart from λ_0 all other eigenvalues have a magnitude strictly less than 1.

What remains to discuss is the relationship between the spectral properties of P and the geometry of the set X . To this end, we introduce a family of *diffusion distances* $\{D_t\}_{t \in \mathbb{N}}$ given by

$$D_t(x, y)^2 = \|p_t(x, \cdot) - p_t(y, \cdot)\|_{L^2(X, dz/\pi(z))}^2 = \int_X (p_t(x, z) - p_t(y, z))^2 \frac{dz}{\pi(z)}.$$

That is, $D_t(x, y)$ is a weighted L^2 distance between the distributions $z \mapsto p_t(x, z)$ and $z \mapsto p_t(y, z)$. In particular, for a fixed t , D_t defines a distance on X and $D_t(x, y)$ will be small if there is a large number of short paths connecting x and y and vice versa. Thus, it reflects the connectivity of X at a given scale t . Moreover, $D_t(x, y)$ is very robust to noise perturbation as it involves summing over all path of length t connection x and y a vice versa.

It turn outs that $D_t(x, y)$ can be computed using the eigenvector and eigenvalues of P (cf. (4.16)) by

$$D_t(x, y)^2 = \sum_{i \geq 1} \lambda_i^{2t} (\psi_i(x) - \psi_i(y))^2,$$

where $i = 0$ is neglected because ψ_0 is constant. Given a prescribed accuracy $\delta > 0$ we can truncate this sum according to (4.14). To this end, define

$$s(\delta, t) = \max \{ i \in \mathbb{N} \mid |\lambda_i|^t > \delta |\lambda_1|^t \}$$

and the family of *diffusion maps* $\{\Psi^t\}_{t \in \mathbb{N}}$ is then given by

$$\Psi^t(x) = (\lambda_1^t \psi_1(x), \dots, \lambda_{s(\delta, t)}^t \psi_{s(\delta, t)}(x)).$$

The components of Ψ^t are called *diffusion coordinates* and the relationship between Ψ^t and D_t can be summarized as follows.

Proposition 4.5 ([CL06a, Proposition 1]). *The diffusion map Ψ^t embeds Y into $\mathbb{R}^{s(\delta, t)}$, such that, the Euclidean distance in $\mathbb{R}^{s(\delta, t)}$ is equal to the diffusion distance up to relative accuracy δ :*

$$\|\Psi^t(x) - \Psi^t(y)\|_2 = D_t(x, y).$$

In particular, Proposition 4.5 states that Ψ^t reorganizes the points in X according to their mutual diffusion distances.

Now in order to utilize diffusion maps numerically, we consider X to be a finite set of sample points of A_k or $W^u(p)$, respectively, and from now on we call X *landmarks*. For instance, X can be the midpoints of the boxes $B \in \mathcal{B}$ that cover the set of interest. With that choice the embedded invariant set is discretized by a grid, although the number of landmarks is very large in general. To decrease the number and reduce the numerical effort we will later develop a novel landmark selection scheme that generates points that sample the underlying set sufficiently well (see Section 4.2.2). Since the sampling of X is generally not related to the geometry of the underlying embedded set, one would like to recover the structure regardless of the distribution of the points. That is why we adjust the procedure described above and introduce a family of *anisotropic diffusion maps* parameterized by a real-valued parameter α which specifies the amount of influence of the density. Instead of applying the graph Laplacian normalization on a graph with isotropic weights we first renormalize the kernel into an anisotropic one. We formulate the idea in Algorithm 7.

Algorithm 7: Anisotropic Diffusion Maps

Initialization: Let $X = \{x_i\}_{i=1}^m \subseteq \mathbb{R}^k$ be a finite set of landmarks and let $k_\varepsilon(x, y)$ be a rotation-invariant kernel of the form (4.12). Choose $\alpha \in \mathbb{R}$ and $s \in \mathbb{N}$.

1) **Renormalization:** Form a new kernel $k_\varepsilon^{(\alpha)}$ on X by

$$k_\varepsilon^{(\alpha)}(x_i, x_j) = \frac{k_\varepsilon(x_i, x_j)}{q_i^\alpha q_j^\alpha}, \text{ where } q_i = \sum_{j=1}^m k_\varepsilon(x_i, x_j). \quad (4.17)$$

2) **Graph Laplacian Normalization:** Employ the graph Laplacian normalization to $k_\varepsilon^{(\alpha)}$ in order to obtain a transition matrix P , that is,

$$P_\varepsilon^{(\alpha)}(x_i, x_j) = \frac{k_\varepsilon^{(\alpha)}(x_i, x_j)}{d_i}, \text{ where } d_i = \sum_{j=1}^m k_\varepsilon^{(\alpha)}(x_i, x_j).$$

3) **Diffusion Maps:** Compute the first $s + 1$ eigenvalues $\{\lambda_i\}_{i=0}^s$ and eigenvectors $\{\psi_i\}_{i=0}^s$ of $P_\varepsilon^{(\alpha)}$. Then the anisotropic diffusion maps $\{\Psi^t\}_{t \in \mathbb{N}}$ are given by

$$\Psi^t(x_i) = (\lambda_1^t \psi_1(x_i), \dots, \lambda_s^t \psi_s(x_i)).$$

Remark 4.6.

- (a) In [Algorithm 7](#) the quantities $k_\varepsilon^{(\alpha)}$ and $P_\varepsilon^{(\alpha)}$ are $m \times m$ matrices and the argument (x_i, x_j) corresponds to the entry at the i -th row and j -th column.
- (b) In order to increase the sparsity of P and reduce the numerical effort, a Gaussian kernel with a cut-off is often used. More precisely, h has the form $h(z) = c_r \exp(-z) \mathbf{1}_{z \leq r}$ with some cutoff radius $r > 0$ and constant c_r such that $\int h(\|z\|^2) dz = 1$. For simplicity one can choose $r = \sqrt{2\varepsilon}$, to assure that interaction between data points further apart than r is sufficiently small. This allows to employ a range search algorithm on a k -d tree [[FBF77](#)] for fast distance computations.

Observe that the diffusion maps generated by [Algorithm 7](#) are only defined on the landmarks X . Thus, in order to embed some *out-of-sample* point $x \in \mathbb{R}^k$ of the approximation of the invariant set, we have to extend this scheme. To this end, we interpolate the diffusion coordinates inspired by the Nyström method [[BPV⁺04](#), [CL06b](#)]. The reason why we present this extension method is that we want to apply diffusion maps not only to the given data points X but also to new data points without the costly recomputation of the entire diffusion maps. Therefore, we can easily embed trajectories of the underlying dynamical system to reveal the dynamics in diffusion coordinates or add additional data points to obtain a finer discretization. We summarize the obtained method in [Algorithm 8](#).

Algorithm 8: Nyström Interpolation

Initialization: Let $X = \{x_i\}_{i=1}^m \subseteq \mathbb{R}^k$ be a finite set of landmarks and $x \in \mathbb{R}^k$ an out-of-sample point. Let $k_\varepsilon(x, y)$, q_i and Ψ^t as in [Algorithm 7](#).

- 1) **Renormalization:** Evaluate the kernel $k_\varepsilon^{(\alpha)}$ for all pairs (x, x_j) , that is,

$$k_\varepsilon^{(\alpha)}(x, x_j) = \frac{k_\varepsilon(x, x_j)}{\tilde{q}^\alpha q_j^\alpha}, \text{ where } \tilde{q} = \sum_{j=1}^m k_\varepsilon(x, x_j).$$

Note that q_j was computed in (4.17).

- 2) **Graph Laplacian Normalization:** Normalize $k_\varepsilon^{(\alpha)}(x, x_j)$ by

$$p_\varepsilon^{(\alpha)}(x, x_j) = \frac{k_\varepsilon^{(\alpha)}(x, x_j)}{\tilde{d}_i}, \text{ where } \tilde{d}_i = \sum_{j=1}^m k_\varepsilon^{(\alpha)}(x, x_j).$$

- 3) **Interpolation:** Then the i -th diffusion coordinate of $x \in \mathbb{R}^k$ is given by

$$\tilde{\Psi}_i^t(x) = \lambda_i^{t-2} p_\varepsilon^{(\alpha)} \Psi_i^t(x_1, \dots, x_m) = \lambda_i^{t-1} \sum_{j=1}^m p_\varepsilon^{(\alpha)}(x_j) \psi_i(x_j).$$

Remark 4.7.

- (a) Similar to Algorithm 7 the quantities $k_\varepsilon^{(\alpha)}$ and $p_\varepsilon^{(\alpha)}$ in Algorithm 8 are just $1 \times m$ row vectors and the argument (x, x_j) corresponds to its i -th entry (cf. Remark 4.6 (a)). In particular, we note that this method can naturally be vectorized.
- (b) Note that this construction is consistent with the definition on the data set X , i.e., for every landmark $x_i \in X$ we have $\tilde{\Psi}^t(x_i) = \Psi^t(x_i)$.
- (c) In the case where a kernel with some cut-off radius $r > 0$ (cf. Remark 4.6 (b)) is used, it is possible that there is no landmark in the r -ball of $x \in \mathbb{R}^k$ and thus x will be mapped to the origin ($p_\varepsilon^{(\alpha)} = 0$). To prevent this phenomenon we adapt the extension method as follows: We successively increase r by 10% until there are k neighbors without changing ε to obtain a coefficient vector $p_\varepsilon^{(\alpha)}$ that has at least k non-vanishing entries. However, this approach is not optimal. In fact, the proposed extension method is only accurate for points within the kernel bandwidth [LF17].

Concerning the choice of $\alpha \in \mathbb{R}$ we refer to the result of [CL06a]. Let \mathcal{M} be a compact C^∞ submanifold of \mathbb{R}^k and let $q(x)$ be the density of the points on \mathcal{M} . Denote by Δ the (positive semi-definite) Laplace–Beltrami operator on \mathcal{M} . Then Δ has eigenfunctions that satisfy the Neumann condition at the boundary $\partial\mathcal{M}$ and form a Hilbert basis of $L^2(\mathcal{M})$. Let E_s be the linear span of the first $s + 1$ Neumann eigenfunctions of Δ .

Proposition 4.8 ([CL06a, Theorem 2 & Proposition 3]). *Let $L_\varepsilon^{(\alpha)} = \frac{1 - P_\varepsilon^{(\alpha)}}{\varepsilon}$ be the (discrete-time) infinitesimal generator of the Markov chain. Then for a fixed $s > 0$, we have for $f \in E_s$*

$$\lim_{\varepsilon \rightarrow 0} L_\varepsilon^{(\alpha)} f = \frac{\Delta(fq^{1-\alpha})}{q^{1-\alpha}} - \frac{\Delta(q^{1-\alpha})}{q^{1-\alpha}} f.$$

In other words, the eigenfunctions of $P_{\varepsilon,\alpha}$ can be used to approximate those of the following symmetric Schrödinger operator:

$$\Delta\phi - \frac{\Delta(q^{1-\alpha})}{q^{1-\alpha}}\phi,$$

where $\phi = fq^{1-\alpha}$. In particular, for $\alpha = 1$ we have

$$\lim_{\varepsilon \rightarrow 0} L_\varepsilon^{(1)} = \Delta$$

and for any $t > 0$ the Neumann heat kernel $e^{-t\Delta}$ can be approximated on $L^2(\mathcal{M})$ by $\left(P_\varepsilon^{(1)}\right)^{\frac{t}{\varepsilon}}$:

$$\lim_{\varepsilon \rightarrow 0} P_{\varepsilon,1}^{\frac{t}{\varepsilon}} = e^{-t\Delta}.$$

Thus, for $\alpha = 1$ the infinitesimal generator is simply the Laplace–Beltrami operator Δ and the Markov chain converges to the Brownian motion on \mathcal{M} . Consequently the normalization removes the influence of the density and we recover the Riemannian geometry of the data set as desired. Hence, we can separate distribution of the data from the geometry of the underlying manifold \mathcal{M} and from now on we choose and fix $\alpha = 1$. It remains to discuss the choice of the bandwidth ε and the landmarks X which is done in the next sections.

4.2.1 An Intrinsic Dimension Estimator

In this section we present a method for the approximation of the *intrinsic* dimension. Moreover, this tool finds a good choice for the bandwidth ε . To this end, we will consider a d_{int} -dimensional manifold \mathcal{M} as the underlying set. According to an observation made in [CSSS08] the kernel localizes the data set when the data is drawn according to a uniform distribution on the manifold and ε is well tuned such that

$$S(\varepsilon) = \frac{1}{m^2} \sum_{i,j} k_\varepsilon(x_i, x_j) \approx \frac{(\pi\varepsilon)^{d_{\text{int}}/2}}{\text{vol}(\mathcal{M})}, \quad (4.18)$$

where d_{int} is the intrinsic dimension of the \mathcal{M} . Note that $S(\varepsilon) \rightarrow 1$ for $\varepsilon \rightarrow \infty$ and $S(\varepsilon) \rightarrow \frac{1}{m}$ for $\varepsilon \rightarrow 0$. Moreover, $S(\varepsilon)$ should be locally well approximated by a power law $S(\varepsilon) \sim \varepsilon^{d_{\text{int}}/2}$, where

$$d_{\text{int}} = 2 \frac{d(\log S)}{d(\log \varepsilon)} \quad (4.19)$$

is the local slope at appropriate values ε of the $\log S$ versus $\log \varepsilon$ curve. Thus, following [BH16] we evaluate S for a large range of ε_i and approximate the derivative by finite differences

$$d_i = 2 \frac{\log S(\varepsilon_{i+1}) - \log S(\varepsilon_i)}{\log \varepsilon_{i+1} - \log \varepsilon_i}. \quad (4.20)$$

A good choice for ε would be a value in the linear region in the $\log \varepsilon$ vs. $\log S$ plot, i.e., the $\log \varepsilon$ vs. d_{int} curve is constant. If there is no obvious region we choose a value near the maximizer ε^* of d_{int} . Observe that this machinery allows in particular the approximation of the dimension d_{int} of \mathcal{M} .

However, this estimation only works if the data is uniformly sampled which is in practice never satisfied. In order to extend this tool for data drawn according to an arbitrary distribution, say q , it is necessary to remove the influence of q in this estimation. To this end, let us recall that for $\alpha = 1$ the underlying distribution q is separated from the geometry by Proposition 4.8. Hence, considering (4.17) for $\alpha = 1$, i.e.,

$$k_\varepsilon^{(1)}(x_i, x_j) = \frac{k_\varepsilon(x_i, x_j)}{q_i q_j}, \quad \text{where } q_i = \sum_{j=1}^m k_\varepsilon(x_i, x_j),$$

instead of $k_\varepsilon^{(0)}$ in (4.18) should give us the desired result. In fact, let us assume q to be smooth and, for simplicity, take the kernel function $k_\varepsilon(x, y) = \exp\left(-\frac{\|x-y\|_2^2}{\varepsilon}\right)$. If ε is sufficiently small in comparison to the scale on which q varies (essentially governed by the gradient of q), then by interpreting the following sum in the Monte Carlo sense, we obtain

$$\frac{1}{m} q_i \rightarrow \int_{\mathcal{M}} k_\varepsilon(x_i, y) q(y) dy \approx C q(x_i)$$

for $m \rightarrow \infty$ almost surely, where

$$C = \int_{T_{x_i} \mathcal{M}} \exp\left(-\frac{\|x_i - y\|_2^2}{\varepsilon}\right) dy$$

is independent of x_i . Now we can identify $T_{x_i} \mathcal{M}$ with $\mathbb{R}^{d_{\text{int}}}$ and it follows that $C = (\pi\varepsilon)^{d_{\text{int}}/2}$. Since the x_i are drawn i.i.d., the pair (x_i, x_j) is for $i \neq j$ distributed according to $q \otimes q$ and we obtain in the Monte Carlo sense

$$\begin{aligned} \sum_{i,j} k_\varepsilon^{(1)}(x_i, x_j) &\approx \int_{\mathcal{M}} \int_{\mathcal{M}} \frac{k_\varepsilon(x, y)}{C^2 q(x)q(y)} (q \otimes q)(x, y) dx dy \\ &\approx \frac{1}{C^2} \int_{\mathcal{M}} \int_{\mathbb{R}^{d_{\text{int}}}} \exp\left(-\frac{\|x - y\|_2^2}{\varepsilon}\right) dx dy \\ &= \frac{\text{vol}(\mathcal{M})}{(\pi\varepsilon)^{d_{\text{int}}/2}}. \end{aligned}$$

Here the first approximation works if ε is sufficiently large, such that the point cloud $\{x_i\}_{i=1}^m$ “resolves” the functions $\exp\left(-\frac{\|\cdot - x_j\|_2^2}{\varepsilon}\right)$ properly, such that the Monte Carlo estimation is valid. The second approximation is valid if ε is sufficiently small, such that the integral $\int_{\mathcal{M}} \exp\left(-\frac{\|\cdot - y\|_2^2}{\varepsilon}\right) dx$ is well approximated by the same integral on the tangent space $\mathbb{R}^{d_{\text{int}}}$ of \mathcal{M} at y . To this end, the “Gaussian bell” should be sufficiently localized, i.e., ε is small enough. It is assumed that in between there is a sweet spot for the values of ε that both approximations hold. We emphasize that this argumentation is also used in [KW20, Appendix A.2] for uniformly sampled data.

Thus, for non-uniformly sampled data one should use $k_\varepsilon^{(1)}$ in (4.18) instead of k_ε in (4.18) to form a sum over all pairs of data points. With that choice, we have $S^{(1)}(\varepsilon) := \frac{1}{m^2} \sum_{i,j} k_\varepsilon^{(1)}(x_i, x_j) \rightarrow -\frac{1}{m^2}$ for $\varepsilon \rightarrow \infty$ and $S^{(1)}(\varepsilon) \rightarrow -\frac{1}{m}$ for $\varepsilon \rightarrow 0$. Also note that in this case we have to add a minus sign in (4.19) and (4.20), respectively, since $S^{(1)}(\varepsilon) \sim \varepsilon^{-d_{\text{int}}/2}$, i.e., $\varepsilon^{d_{\text{int}}/2}$ is now in the denominator. Even though this estimation is theoretically only valid when the underlying set is a manifold we will also employ this tool if the data is sampled from an arbitrary (fractal) set.

Remark 4.9. For the approximation of the dimension d_{int} the kernel $k_\varepsilon^{(1)}$ has to be evaluated for all data points and a lot of bandwidths ε_i . Thus, in order to keep the numerical effort within reasonable limits, we first compute the pairwise distances $\|x_i - x_j\|_2 \leq r$ for some cut-off radius $r > 0$ and store $\tilde{d}_{i,j} = -\|x_i - x_j\|_2^2$ in memory. Here, we emphasize that the cut-off radius has to be chosen sufficiently large, e.g., $r = \sqrt{2 \max_i \varepsilon_i}$, such that it is still appropriate for larger bandwidths. Afterwards for each prescribed ε_i we evaluate $k_\varepsilon(x_i, x_j) = \exp\left(\frac{\tilde{d}_{i,j}}{\varepsilon_i}\right)$ and compute q_i . Note that for the computation of $k_\varepsilon^{(1)}(x_i, x_j)$ one does not have to store $k_\varepsilon^{(1)}(x_i, x_j)$ as a $m \times m$ matrix. Finally, we build $S^{(1)}(\varepsilon)$ by summing over all entries and approximate the derivative of the $\log(\varepsilon)$ versus $\log(S^{(1)})$ curve by finite differences (cf. (4.19)).

In order to illustrate the diffusion maps embedding as well as the automated choice for the bandwidth ε , we will consider a popular example given by the *swiss roll*. This set is a rolled up 2-dimensional rectangle embedded in \mathbb{R}^3 (see Figure 4.3 (a)). We choose $m = 10^4$ random landmarks and compute $S^{(1)}(\varepsilon)$ for $\varepsilon \in [10^{-4}, 10^5]$ and the quantities d_i . The corresponding curves are shown in Figure 4.2. Unfortunately, we do not see a clear (large) linear region in (a), or a constant region in (b), respectively. However, at around $\varepsilon \approx 10^{0.5}$ we see a slight hint of a plateau whose level is close to the exact value of $d_{\text{int}} = 2$. Note that for a larger bandwidth we overestimate the dimension. This is due to the fact, that the ‘‘Gaussian bell’’ is too large such that it does not represent a local intrinsic neighborhood of a point anymore.

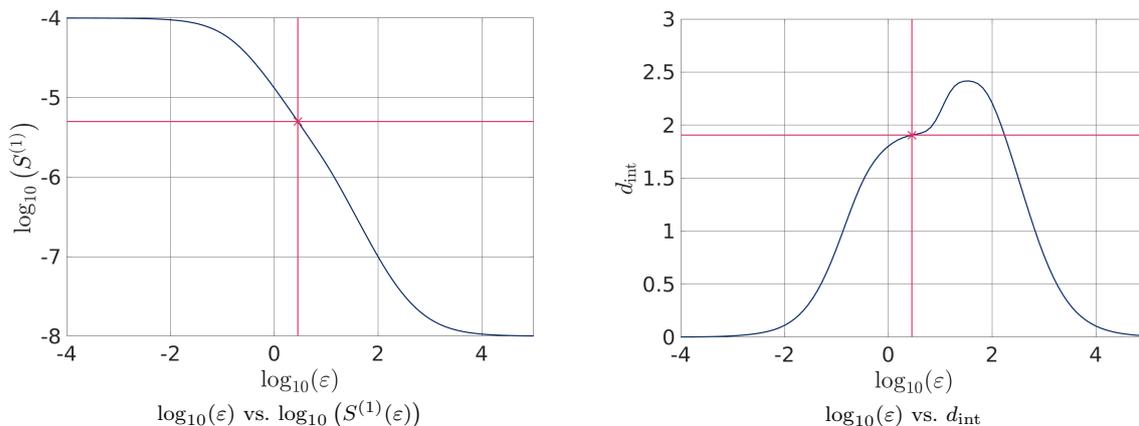


Figure 4.2: Illustration of the estimation of the intrinsic dimension. The red cross depicts the appropriate bandwidth $\varepsilon = 10^{0.5}$ and its corresponding values $\log_{10}(S^{(1)}(\varepsilon))$ and $d_{\text{int}}(\varepsilon)$, respectively.

In Figure 4.3 (b) we show the set in diffusion coordinates scaled by the constant function ψ_0 , where we additionally embed 10^6 additional out-of-sample points using Algorithm 8. Note that the diffusion coordinates unroll the rectangle as desired. The first coordinate corresponds to the direction in which the rectangle is rolled up, whereas the second coordinate corresponds to the y direction in Cartesian coordinates.

Now having a tool for estimating the intrinsic dimension as well as finding a suitable bandwidth ε it remains to discuss how to discretize an embedded invariant set sufficiently well.

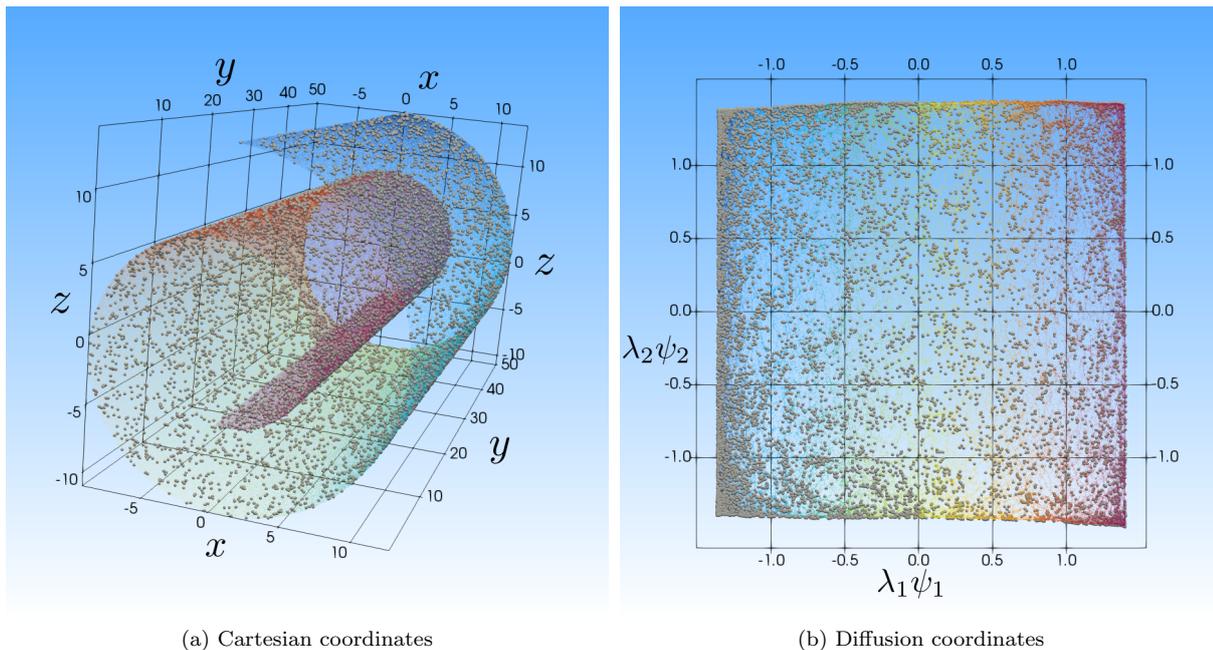


Figure 4.3: Illustration of the diffusion maps embedding 7 and 8 on the swiss roll. In gray we show the randomly chosen landmarks and their corresponding embedding, whereas the out-of-sample points are colored according to their first diffusion coordinate.

4.2.2 A Set-Oriented Landmark Selection Scheme

In order to employ diffusion maps in our framework, we have to choose a set of landmarks that discretize the embedded set of interest Q given by a box collection \mathcal{B} . The simplest strategy to discretize such a collection by points $X = \{x_i\}_{i=1}^m \subseteq \mathbb{R}^k$ is to take the midpoints of each box $B \in \mathcal{B}$, that is, $x_i = c_i$ according to Remark 3.3 (a). However, in general the number $M = |\mathcal{B}|$ of boxes in \mathcal{B} will be very large and hence not suitable for the diffusion map Algorithm 7. More precisely, the computation of the pairwise distance and the eigendecomposition will be computationally infeasible. Thus, the purpose of this section is to present a novel algorithm that generates $m \ll M$ landmarks that are well-sampled in the sense that the r -balls for some $r \geq 0$ around the points cover the embedded set and its pairwise intersection is small, ideally empty.

Starting with a random subset $X = \{x_i\}_{i=1}^m \subseteq \{c_i\}_{i=1}^M$ we employ a point cloud simplification method (see [PGK02] for an overview and the references therein) for the computation of the landmarks. We note that the developed method is inspired by the works of Turk [Tur91, Tur92]. There, surfaces discretized by polygons are re-sampled by first seeding points on the surfaces and then moving them around by repulsion between neighboring points. If during this procedure a point is pushed off one polygon, it is moved onto an adjacent polygon. Thus, the idea is to push the points $x_i \in X$ that (coarsely) approximate the embedded set of interest around by a repelling force that decays linearly with distance and becomes zero at a fixed radius $r > 0$. If during that procedure a point leaves the covering, we project it onto the nearest box in the box collection. The proposed scheme is formulated in Algorithm 9.

Algorithm 9: Set-Oriented Landmark Selection Scheme

Initialization: Let \mathcal{B} be a box collection and $X = \{x_i\}_{i=1}^m$ a finite discretization of the corresponding covering $Q = \bigcup_{B \in \mathcal{B}}$. Fix a *repulsion* radius $r > 0$.

Particle Simulation: Repeat the following steps.

- 1) For each $x_i \in X$ compute all neighbors $\{y_j\}_{j=1}^{n_i} \subseteq X$ of x_i such that $\|x_i - y_j\| \leq r$. Set

$$F_i = \sum_{j=1}^{n_i} (r - \|x_i - y_j\|) \frac{x_i - y_j}{\|x_i - y_j\|}.$$

- 2) Update x_i by $x_i = x_i + \kappa_i F_i$ for some $\kappa_i > 0$.
 - 3) If $x_i \notin B$ for all $B \in \mathcal{B}$, project x_i onto the boundary of the nearest box $B \in \mathcal{B}$.
-

In what follows we will discuss the choice of the repulsion radius $r > 0$, the factor $\kappa_i > 0$ and how to realize the projection step 3) in [Algorithm 9](#). For a good choice of $r > 0$ let us recall a helpful property of the box-counting dimension (cf. [Definition 2.5](#)), that is,

$$m(Q, r) \sim r^{-d_{\text{box}}(Q)},$$

where $Q = \bigcup_{B \in \mathcal{B}}$ is a cover of the embedded set and $m(Q, r)$ is the minimal number required to cover Q with balls of radius r . Hence, in order to find a good covering by a fixed number m of r -balls, the radius should be of the form

$$r \sim m^{-1/d_{\text{box}}(Q)}.$$

However, for the final definition of r we have to take the size of Q into account. To this end, let us consider the intervals $Q_1 = [0, 1] \subseteq \mathbb{R}$ and $Q_2 = [0, 2] \subseteq \mathbb{R}$. Observe that the box-counting dimension of both sets coincides and is one. Then for a fixed number m of balls that should cover Q_1 one has to choose a radius $r_1 \geq 1/(m+1)$, whereas for Q_2 it is $r_2 \geq 2/(m+1)$. That is why for general $Q \subseteq \mathbb{R}^k$ we define $r > 0$ as

$$r = \text{diam}(Q) m^{-1/d_{\text{box}}(Q)}.$$

For the choice of $\kappa_i > 0$ let us consider the graph that is implicitly generated in step 1) of [Algorithm 9](#), i.e., we connect x_i with x_j if their pairwise distance is less or equal than r . Suppose x_i is only connected to x_j before the update in step 2) and vice versa. Then the corresponding lengths of F_i and F_j coincide but they point in opposite directions, i.e., $F_i = -F_j$. In order to keep x_i and x_j connected after the update in step 2), κ_i and κ_j must be smaller or equal than $1/2$. In the case where there are more connections this translates to

$$\kappa_i \leq \frac{1}{n_i + 1}$$

by the triangle inequality, where $n_i \in \mathbb{N}_0$ is the number of neighbors of x_i , that is, the degree of the node x_i in the graph. Thus, in the hope of a fast convergence of Algorithm 9 while preserving existing edges we choose the maximal allowed value

$$\kappa_i = \frac{1}{n_i + 1}. \quad (4.21)$$

In step 3) in Algorithm 9 we want to project a point x onto the boundary of the nearest box which we realize as follows. Let $c \in B \in \mathcal{B}$ be the closest midpoint to x . Then we reduce the problem to finding a point on the intersection of the line between x and c and the boundary of the box B . Thus, let $d = x - c \in \mathbb{R}^k$ be the direction vector from c to x and $r \in \mathbb{R}^k$ the generalized radius of B (see Remark 3.3). The task is then to find the interval $I = [L_0, L_1]$ such that $c + ld \in B$ if and only if $l \in I$. In other words,

$$ld_i \leq r_i \text{ and } -ld_i \leq r_i \quad (4.22)$$

for all $i = 1, \dots, k$. We can rewrite (4.22) in a vector valued constraint, that is,

$$l \begin{pmatrix} d \\ -d \end{pmatrix} \leq \begin{pmatrix} r \\ r \end{pmatrix}.$$

Since $r_i > 0$, positive entries u_i give the upper bound L_1 and negative u_i the lower bound L_0 . If $u_i = 0$ the direction is parallel to the i -th box constraint and provides no information on the bounds. Thus, L_0 and L_1 are given by

$$L_0 = \max_{i:u_i < 0} \frac{r_i}{u_i} \text{ and } L_1 = \min_{i:u_i > 0} \frac{r_i}{u_i}$$

(also see [TvBP13, Section 3.2]). Since we want the closest intersection of that line and the boundary, we choose $l = L_1$ and define the projection of x as

$$x^{(p)} = c + L_1 d = L_1 x + (1 - L_1) c.$$

Remark 4.10.

- (a) As in Algorithm 3 the neighbor search step 1) of Algorithm 9 can be numerically realized by a k -d tree which accelerates the computation. Here we have to build a new tree or rebalance the previous one in each iteration since the points $x_i \in X$ change their location. Hence, the neighbor search can be computationally expensive, in particular, if $m < M$ is large.
- (b) Instead of considering the r -ball around each point x_i one can also only compute the nearest neighbor and define F_i analogously. In this case, κ_i should be less or equal to $1/2$ according to (4.21). However, in practice this modification yields that the points “jump around” and the range search proposed in step 1) is much smoother.

- (c) A stopping criterion for Algorithm 9 can be to perform a prescribed number of iterations or to compute the minimal distance between the current points and the updated points (after step 3), for instance. If this distance is then lower than a predefined value Algorithm 9 ends. With our choice of $r > 0$ and κ_i one can also design a stopping criterion based on the number of neighbors remaining in each r -ball and the corresponding distances. In the following we will simply stop after a fixed number of iterations.

Finally, we illustrate Algorithm 9 on a simple example in Figure 4.4.

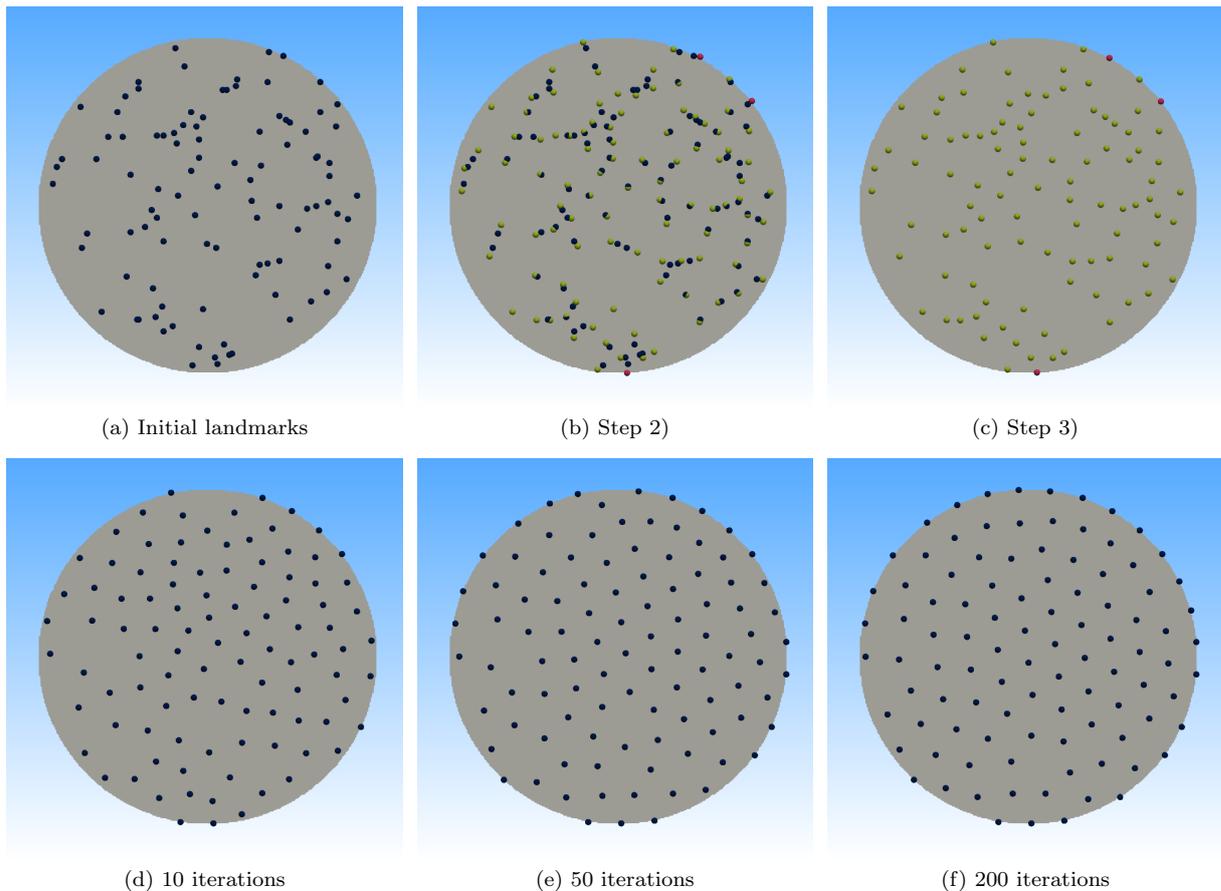


Figure 4.4: Illustration of Algorithm 9 on a covering of the unit disc. (a) $m = 100$ random initial landmarks X . (b) The points were moved according to step 2). Green points are still inside the set, while red ones moved slightly outside the disc. (c) The red points are projected onto the boundary. (d-f) Further iteration of the algorithm. We see that the distribution of the points is forming a grid like structure and we also observe that there are more points on the boundary in each iteration.

5 Applications

In this chapter we apply the developed numerical tools for the numerical analysis of (infinite-dimensional) systems on several examples. First, we will employ the set-oriented path following method 2 for the approximation of the global attractor of a reduced order model for turbulent shear flows. Moreover, we will use the set-oriented predictor-corrector method 5 for the approximation of the attracting set of the Mackey-Glass delay differential equations for a range of delay times. In the chaotic parameter regime we investigate the geometry of the generated covering using the diffusion map scheme (see Section 4.2). For the computation of an embedded unstable manifold we will consider the one-dimensional Kuramoto-Sivashinsky equation, where we will also reveal its intrinsic structure using diffusion maps. At the end of this chapter we will approximate the embedded unstable manifold of the *edge state* in a plane Poiseuille flow.

We note that some parts of this chapter are also contained in [ZDG19, GKD19, GZED20] and [ZGD20] to which the author has made substantial contributions.

5.1 A Four-Dimensional Model of Self-Sustained Flows

We start with using the set-oriented path following method 2 in order to explore the dynamics of shear flows during the transition to turbulence. Analyzing the dynamics in the full phase space is very complicated because of the high-dimensionality. In order to prepare for the application to such computationally expensive examples such as plane Poiseuille flow considered later in Section 5.4, we begin here with an analysis of a simple model that capture much of the phenomenology. Thus, in this thesis, as a model of shear flows where the modes have a physical interpretation, we will consider the four-dimensional model proposed by Fabian Waleffe [Wal95a, Wal95b]: the four components represent the transversal velocity components, the vortices, the streak, and the mean velocity. We also note that in [GZED20] the proposed method has been applied successfully to a Fourier based nine-dimensional model for turbulent shear flows with free slip boundaries and a sinusoidal base profile [MFE04, MFE05].

The corresponding four-dimensional model is given by the following nonlinear ordinary differential equation

$$\frac{d}{dt} \begin{pmatrix} u \\ v \\ w \\ m \end{pmatrix} = \frac{1}{R} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sigma \end{pmatrix} - \frac{1}{R} \begin{pmatrix} \lambda u \\ \mu v \\ \nu w \\ \sigma m \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\gamma w & v \\ 0 & 0 & \delta w & 0 \\ \gamma w & -\delta w & 0 & 0 \\ -v & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ m \end{pmatrix}, \quad (5.1)$$

where $R > 0$ is the Reynolds number and $\lambda, \mu, \nu, \sigma, \gamma,$ and δ are positive parameters. For the following computations we fix $\lambda = \mu = \sigma = 10, \nu = 15, \delta = 1$ and $\gamma = 0.5$. In this case the attractor of (5.1) does not contain any invariant structures besides the laminar profile $(u, v, w, m) = (0, 0, 0, 1)$ up to the critical Reynolds number $R_c = 98.6325$, where a saddle-node bifurcation occurs. The aim in this section is to numerically analyze the change of the attractor for the Reynolds number $R \in [99, 400]$. To this end, we consider the time- T -map of (5.1) for $T = 20$ as the dynamical system φ_R (cf. (3.10)) where $R > 0$ is the parameter of interest. Furthermore, we choose $Q = [-0.9, 1.1] \times [-0.8, 1.2] \times [-1, 1] \times [-0.8, 1.2] \subseteq \mathbb{R}^4$

as the initial box in which we want to approximate the relative global attractor A_Q^R using Algorithm 2 where we choose $m = 36$ and fix $L = 32$. This yields a box diameter of 0.004 of $B \in \mathcal{B}_m$ in the final approximation Q_m^R . Following Remark 3.19 (a) we start the scheme at $R_0 = 400$ and define $R_{j+1} = R_j - 1$ for $j = 0, \dots, 302$ such that the interval $\Lambda = [98, 400]$ is discretized equidistantly.

In what follows we will analyze the attractor from a global point of view. To this end, we show in Figure 5.1 and Figure 5.2 three-dimensional projections (u, v, m) of the relative global attractor obtained by Algorithm 2 for different Reynolds numbers. For R just above the saddle node bifurcation, i.e., for $R_c = 98.6325$, Figure 5.1 (a) shows that the attractor now contains an upper branch steady solution, as well as the lower branch saddle state, also called *edge state*. This situation persists up to $R = 100.0232$, after which the upper branch undergoes a supercritical Hopf bifurcation. The emerging limit cycle is stable (see Figure 5.1 (b)) until $R = 101.0311$, where it disappears in a homoclinic bifurcation. We observe that the attracting set does not change significantly, although the dynamics does: in the homoclinic bifurcation the attractor rips open and becomes a transient saddle. Hence, a trajectory initialized near the unstable upper branch eventually converges to the laminar solution again (see Figure 5.1 (c)). In fact, for $R = 101$ there is a strict separation in the attractor by the edge state which breaks up for $R = 102$. Note that the obtained box coverings do not only cover the laminar profile but also the unstable manifolds since they are backward invariant (see Lemma 3.1). To further demonstrate this phenomenon the average lifetimes for each box in the attractor have been computed in [GZED20].

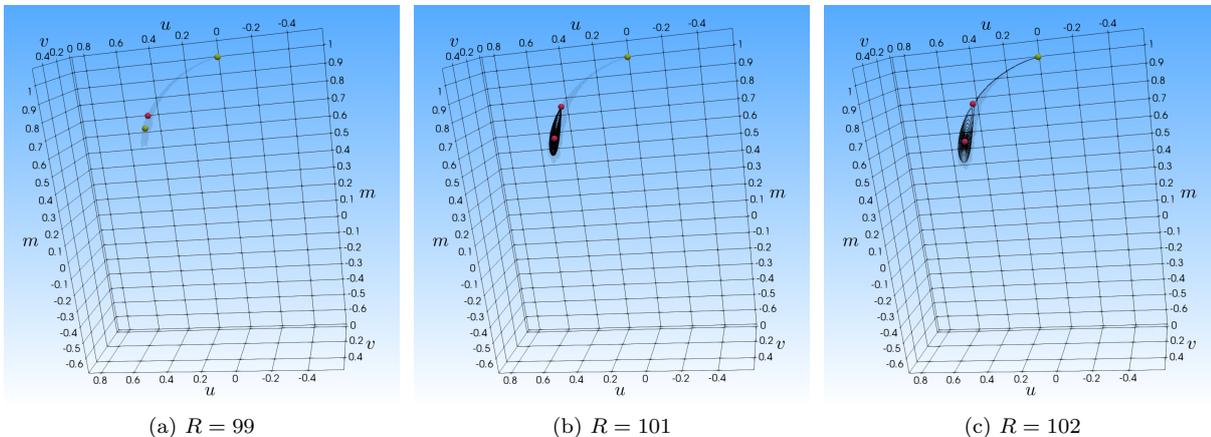


Figure 5.1: Three-dimensional projections of A_Q^R at subdivision level $m = 36$ for the low-dimensional model of a self-sustained flow (5.1). Stable and unstable equilibria are shown in green and red, respectively. The green equilibrium at the top of each figure depicts the laminar solution $(u, v, w, m) = (0, 0, 0, 1)$ which remains stable for all values of R [Wal95b]. In (b) and (c) we show a long-term simulation initialized near the unstable upper branch state.

For $102 < R \leq 356$ the attractor develops a more complicated structure with folds and other features (see Figure 5.2), that arise from the projection from the four-dimensional state space to the three-dimensional image plane. Finally, for $356 < R < 435$ another stable limit cycle appears (see Figure 5.2 (e) and (f)) and the edge state separates the attractor again.

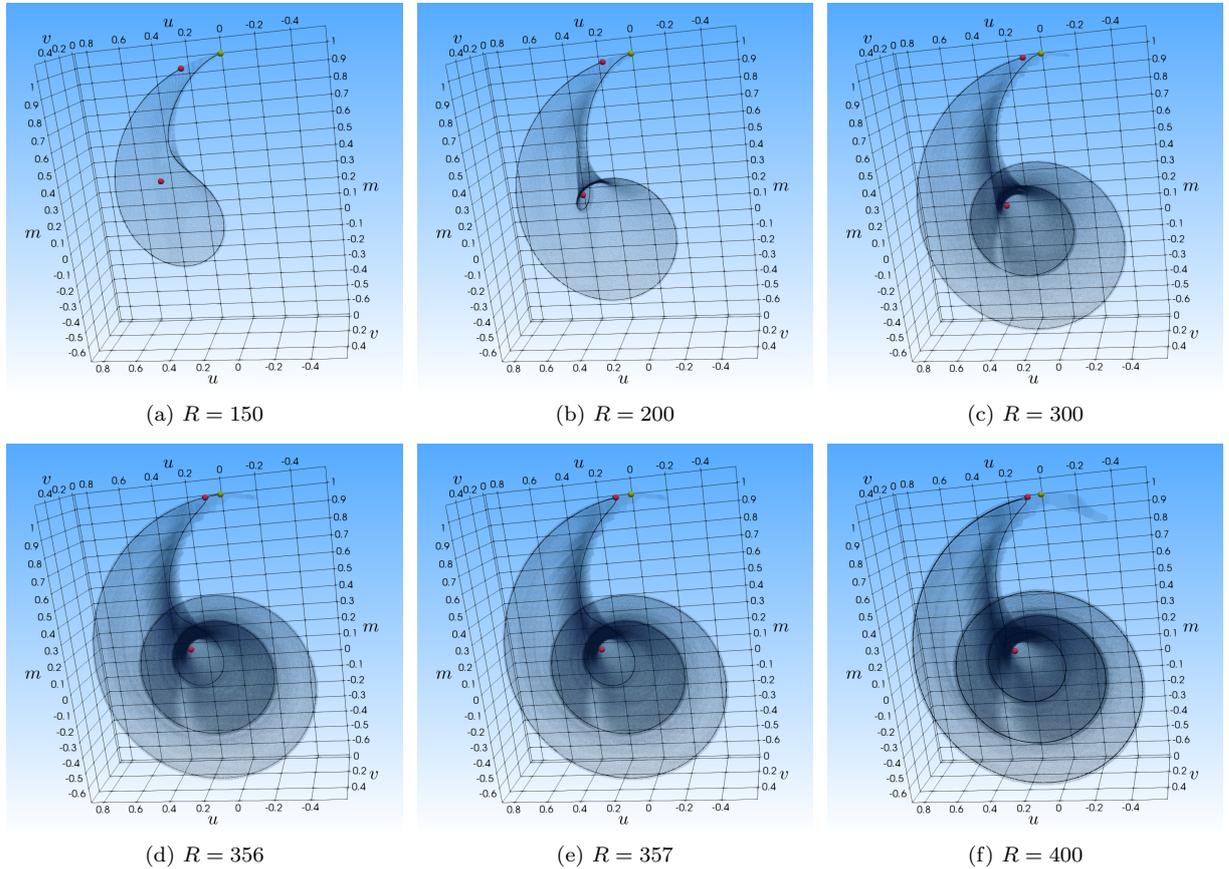


Figure 5.2: Three-dimensional projections of A_Q^R at subdivision level $m = 36$ for the low-dimensional model of a self-sustained flow (5.1). Stable and unstable fixed points are shown in green and red, respectively. The green fixed point at the top of each figure depicts the laminar solution $(u, v, w, m) = (0, 0, 0, 1)$ which remains stable for all values of R [Wal95b]. Additionally, we show a long-term simulation initialized near the edge state.

5.2 The Mackey-Glass Equation

As a first example where the underlying system is infinite-dimensional we consider the well-known delay differential equation introduced by Mackey and Glass in 1977 [MG77] defined by

$$\dot{u}(t) = \beta \frac{u(t - \tau)}{1 + u(t - \tau)^\eta} - \gamma u(t), \quad (5.2)$$

where $\beta = 0.2, \gamma = 0.1, \eta = 10$ are fixed parameters. This equation models the concentration of blood production at a specific time t which also depends on the concentration at an earlier time $t - \tau$, i.e., when the request for more blood is made. Here, we choose the constant time-delay $\tau \in \Lambda \subseteq \mathbb{R}_{\geq 0}$ as our parameter of interest and consider $\Lambda = [6, 20]$. For this parameter regime a detailed analysis can be found in [Far82]. Given the parameters β, γ and η as above, this system possesses the equilibria $u_1(t) = 0$ and $u_2(t) = 1$, where u_2 is a stable fixed point for $\tau < 4.53$. At $\tau \approx 4.53$ the equilibrium u_2 undergoes a supercritical Hopf bifurcation which results in a stable limit cycle attractor.

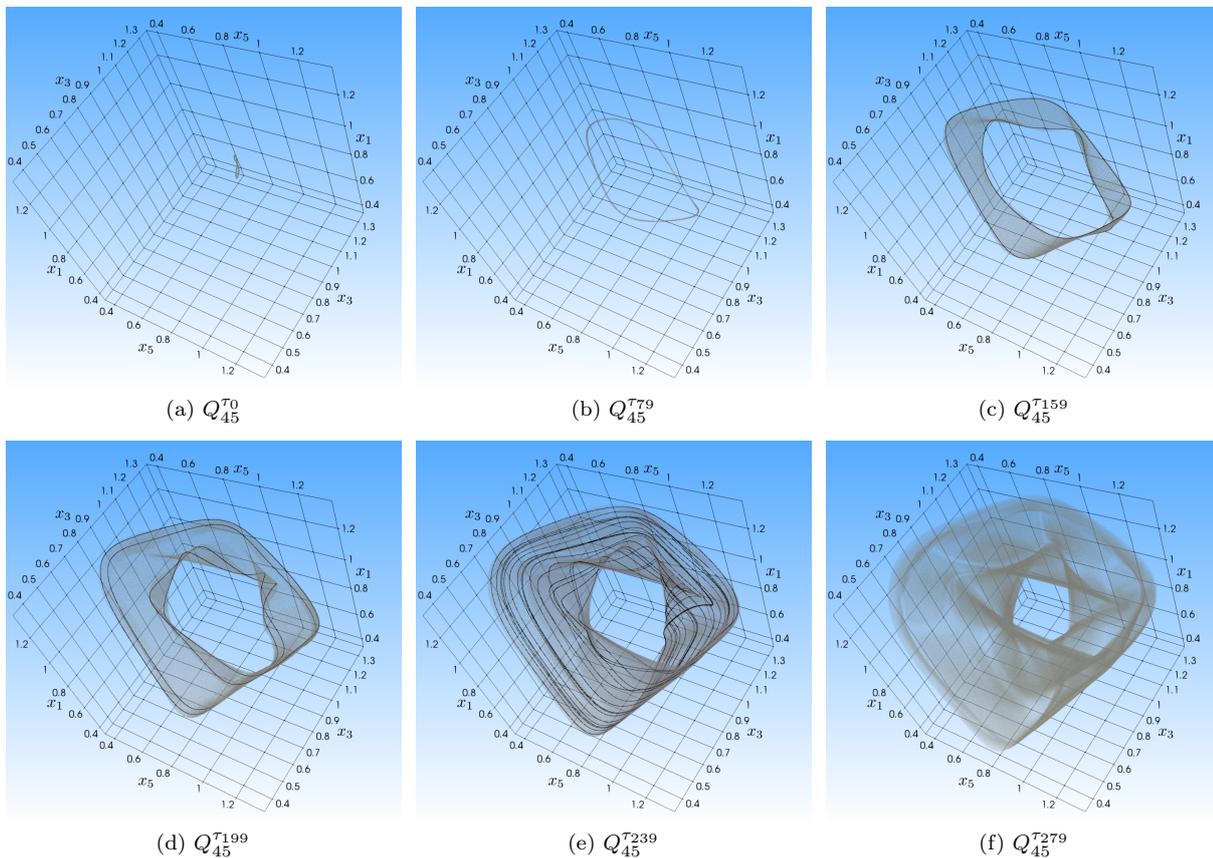


Figure 5.3: Three-dimensional projections of the box covering of the embedded attracting set for the Mackey-Glass delay differential equation (5.2). (a) $Q_{45}^{\tau_0}$ for $\tau = 6$ is obtained by Algorithm 1 whereas the rest is computed by Algorithm 5. (b) Stable periodic orbit for $\tau = 10$. (c) Period-doubled orbit for $\tau = 14$. (d) Periodic-doubling bifurcation sequence for $\tau = 16$. (e) Chaotic behavior for $\tau = 18$. (f) Transparent boxes depicting the chaotic attractor for $\tau = 20$.

For $\tau > 4.53$ the stable limit cycle grows with increasing values of τ and loses its stability in a period-doubling bifurcation at $\tau \approx 13.3$. This initiates a period-doubling bifurcation sequence that reaches its accumulation parameter at $\tau = 16.8$. For $\tau > 16.8$ numerical simulations show chaotic attractors.

In what follows we want to reproduce the results discussed in [Far82] from a global point of view, i.e., we want to show how the attracting set changes under variations of the time delay τ . To this end, we choose the embedding dimension $k = 5$ which should be large enough according to Theorem 2.21 and the dimension analysis made in [Far82]. Since a solution of (5.2) represents the concentration of blood, we are only interested in positive solutions. Direct numerical simulations then indicate that it is sufficient to choose the initial box $Q = [0, 1.5]^5$. Furthermore, we set $T = 5\bar{\tau}$ for $\bar{\tau} = 20$, $\tau_0 = 6$ and define $\tau_{j+1} = \tau_j + 0.05$ for $j = 0, \dots, 279$. Moreover, we set $m = 45$ and $L = 40$. The initial step of Algorithm 5 computes the box coverings $Q_{45}^{\tau_0}$ and $Q_{45}^{\tau_1}$ of the corresponding attracting sets $A_k^{\tau_0}$ and $A_k^{\tau_1}$, respectively. For both parameter values the sets consist of two-dimensional unstable manifolds of $u_2(t) = 1$ which accumulate in a stable periodic orbit at their boundaries. However, since we are only interested in the stable limit cycle attractor, we remove a small open neighborhood U of the embedded equilibrium $R(u_2)$ from the initial box $Q = [0, 1.5]^5$ and approximate $A_k^{\tau_0}$ and $A_k^{\tau_1}$ using $\tilde{Q} = Q \setminus U$.

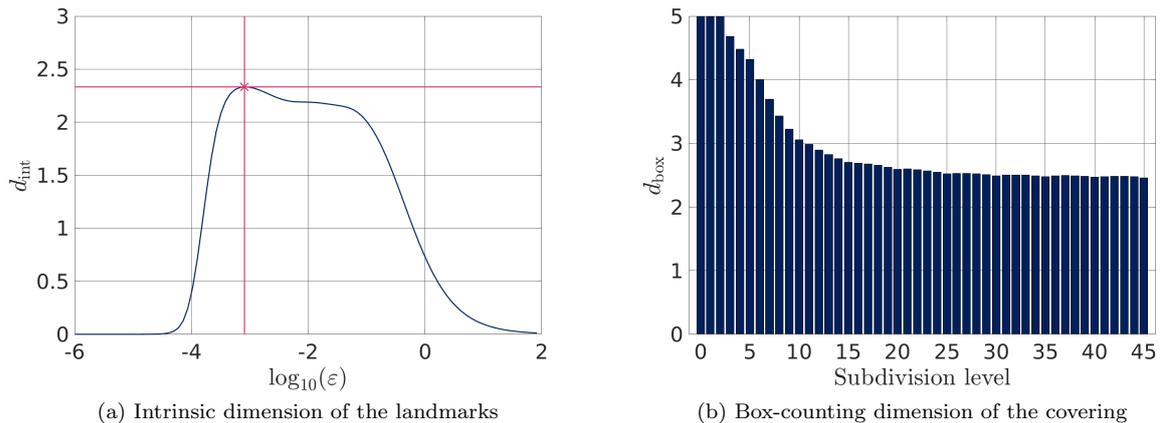


Figure 5.4: Estimation of the dimension of the covering for $\tau = 20$. (a) Intrinsic dimension according to Section 4.2.1. (b) Box-counting dimension (see Definition 2.5). Note that the estimated dimensions coincide with the results made in [Far82].

A three-dimensional projection of $Q_{45}^{\tau_0}$ is shown in Figure 5.3 (a). Then Algorithm 5 computes successively the box coverings $Q_{45}^{\tau_j}$ for $j = 2, \dots, 279$. In Figure 5.3 (b) we show the box covering of Q_{45}^{τ} for $\tau = 10$ obtained after $j = 79$ steps. As expected, the stable limit cycle grows in diameter with increasing values of the time delay τ . Using direct numerical simulation we observe that the stable limit cycle loses its stability in a period-doubling bifurcation at $\tau \approx 13.3$. In Figure 5.3 (c) and (d) we show the attracting sets obtained for $\tau = 14$ and $\tau = 16$ using Algorithm 5 and the periodic solutions obtained by direct numerical simulation. For $\tau = 14$ we see a Möbius strip like structure with the period-doubled periodic solution at its boundaries. Further increasing τ induces a period-doubling bifurcation sequence (cf. Figure 5.3 (d)) and this Möbius strip like set

grows and develops additional loops which increases the complexity of the attracting set. In Figure 5.3 (e), i.e., for $\tau = 18$, a direct numerical simulation of a random initial value depicts a chaotic behavior of (5.2). This set even becomes more complex for increasing values of τ (cf. Figure 5.3 (f)).

For a further insight in the intrinsic geometry of the chaotic attractor for $\tau = 20$ we employ the framework of diffusion maps presented in Section 4.2. We sample $m = 10^4$ landmarks out of $M \approx 4.6 \cdot 10^6$ box midpoints after 50 iterations. According to Section 4.2.1 the intrinsic dimension is estimated by $d_{\text{int}} \approx 2.33$ for an optimal value $\varepsilon \approx 10^{-3.1}$ (see Figure 5.4 (a)). In Figure 5.4 (b) we estimate the corresponding box-counting dimension d_{box} by approximately 2.46 (cf. Remark 2.6).

The discretized embedded attracting set and its corresponding diffusion coordinates scaled by the constant function ψ_0 are shown in Figure 5.5. Even for the chaotic parameter regime of $\tau = 20$ we reveal a Möbius strip like structure that is already seen in delay coordinates for $\tau = 14$ (cf. Figure 5.3 (c)).

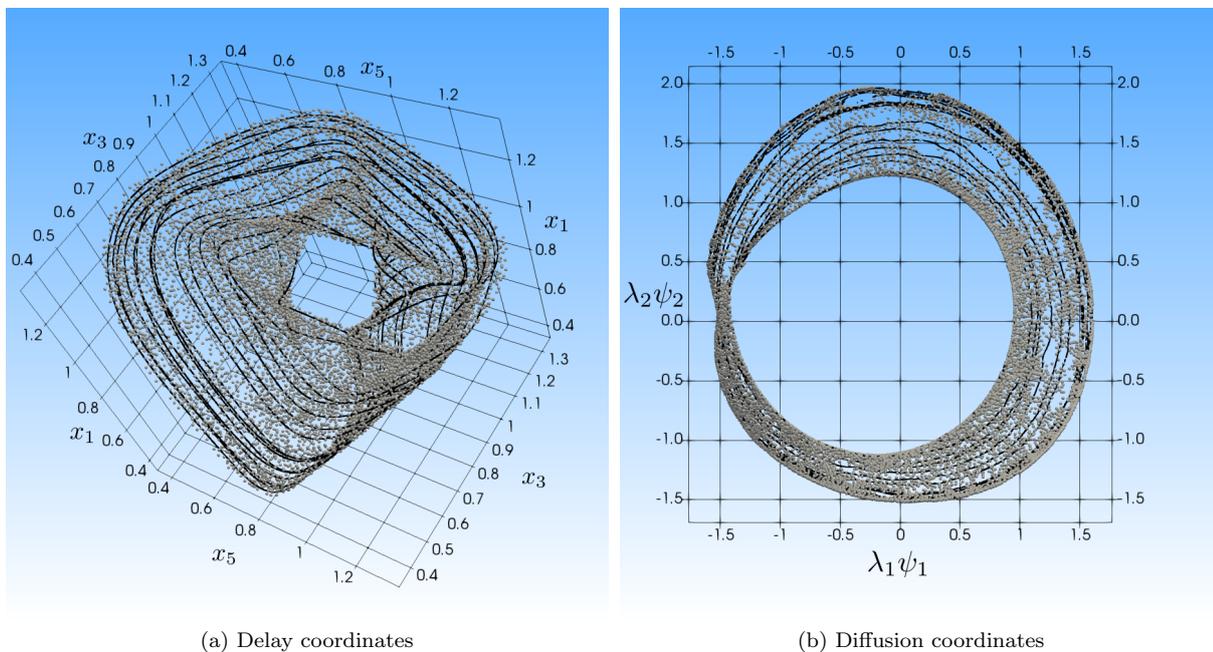


Figure 5.5: Employing the diffusion maps technique to the covering of the embedded attracting set for $\tau = 20$. In gray we show the selected landmarks generated by Algorithm 9, whereas a long-term simulation (embedded by Algorithm 8) is shown in black.

5.3 The Kuramoto-Sivashinsky Equation

For the computation of the embedded unstable manifold of a PDE we consider the well-known Kuramoto-Sivashinsky equation (KSE) in one spatial dimension which is given by

$$\begin{aligned} u_t + \nu u_{yyyy} + u_{yy} + \frac{1}{2}(u^2)_y &= 0, \quad 0 \leq y \leq L, \\ u(y, 0) = u_0(y), \quad u(y + L, t) &= u(y, t), \end{aligned} \quad (5.3)$$

where $\nu > 0$ is a fixed damping parameter and $L > 0$ denotes the size of a pattern. This equation has been studied extensively over the past 40 years. For instance, it has been used to model phase dynamics in reaction-diffusion systems [KT76] or small thermal diffusive instabilities in laminar flame fronts [Siv77]. It turns out that increasing L leads to more complicated spatio-temporal dynamics. Following [HNZ86, KNS90] we normalize the KSE to an interval length of 2π and set the damping parameter to the original value derived by Sivashinsky, i.e., $\nu = 4$. Then equation (5.3) can be written as

$$\begin{aligned} u_t + 4u_{yyyy} + \mu \left[u_{yy} + \frac{1}{2}(u^2)_y \right] &= 0, \quad 0 \leq y \leq 2\pi, \\ u(y, 0) = u_0(y), \quad u(y + 2\pi, t) &= u(y, t), \end{aligned} \quad (5.4)$$

where, additionally, the state u and the time t has to be rescaled. The newly introduced parameter $\mu = L^2/4\pi^2$ then directly relates to the size of a typical pattern scale $L > 0$. In [HNZ86, KNS90] numerical and analytical studies were made by varying μ over a finite interval, showing the complex hierarchy of bifurcations. We note that (5.4) is equivariant with respect to the action of $\mathbf{O}(2)$ defined by

$$R_\alpha u(y, t) = u(y + \alpha, t) \text{ and } Su(y, t) = -u(-y, t). \quad (5.5)$$

due to the periodic boundary condition.

In this section we are interested in computing and analyzing the unstable manifold of the trivial unstable steady state $u^* = 0$ for parameter values $\mu \in [15, 18]$. For the computation we will employ the continuation method 6 on the pCDS φ_μ defined in (2.34). Thus, for the construction of the CDS we need a good estimate of the box-counting dimension of the invariant set \mathcal{A}^μ . In [Rob94] it has been shown that the dimension of the inertial manifold of (5.3) for $\nu = 1$ is $d \leq L^{2.46}$, i.e., each invariant set has finite-dimension. However, these estimates are very pessimistic and we expect that we will obtain one-to-one images of the unstable manifold for smaller related embedding dimensions $k \in \mathbb{N}$. Thus, in what follows we choose $k = 7$ which should suffice according to the discussion on the dimensions later on.

According to Section 4.1.2 the observation space is defined by the first k POD coefficients and thus $p = R(u^*) = 0 \in \mathbb{R}^k$. We compute the POD basis by using the snapshot-matrix obtained through a long-time integration with the initial conditions

$$u_0(y) = 0.0001 \cos(y) \cdot (1 + \sin(y)).$$

Furthermore, we choose S such that $\varepsilon(S) > 0.99999$ (cf. (4.8)). For the purpose of

comparing the parameter-dependent manifolds, we fix the POD basis that is computed for $\mu = 15$ if not said otherwise. Hence, the observation space \mathbb{R}^k given by the corresponding coordinate representation in that basis is fixed as well.

For the upcoming computations we choose $Q = [-50, 50]^7$ and initialize a fine partition \mathcal{P}_s of Q for $s = 28, 42, 56$. Next we set $T = 2$ and define a finite time grid $\{t_0, \dots, t_N\}$, where $t_i = ih$ for $h = 0.001$ according to Remark 3.23 (c) and integrate $128 \cdot 10^3$ points in the small initial box.

Remark 5.1. For the computations and simulations made in [ZDG19] the normalized KSE (5.4) was considered but the state u and time t was mistakenly not rescaled. Hence, in this thesis related quantities such as T in the time- T -map of (5.4) and the magnitude of states u or the corresponding POD coefficients x differ from [ZDG19]. However, qualitatively the results will be the same.

5.3.1 The Traveling Wave

For the parameter value $\mu = 15$ the KSE possesses two stable traveling waves (limit cycles) traveling in opposite directions due to the $\mathbf{O}(2)$ -symmetry (5.5) imposed by the periodic boundary conditions [KNS90]. In addition to that, we numerically find an unstable steady state $v^* \neq 0$ by long-term simulation for the construction of the POD basis. It turns out that a long-term simulation starting near $u^* = 0$ first approaches the point v^* and then eventually converges to one of the traveling waves. We illustrate this behavior in Figure 5.6.

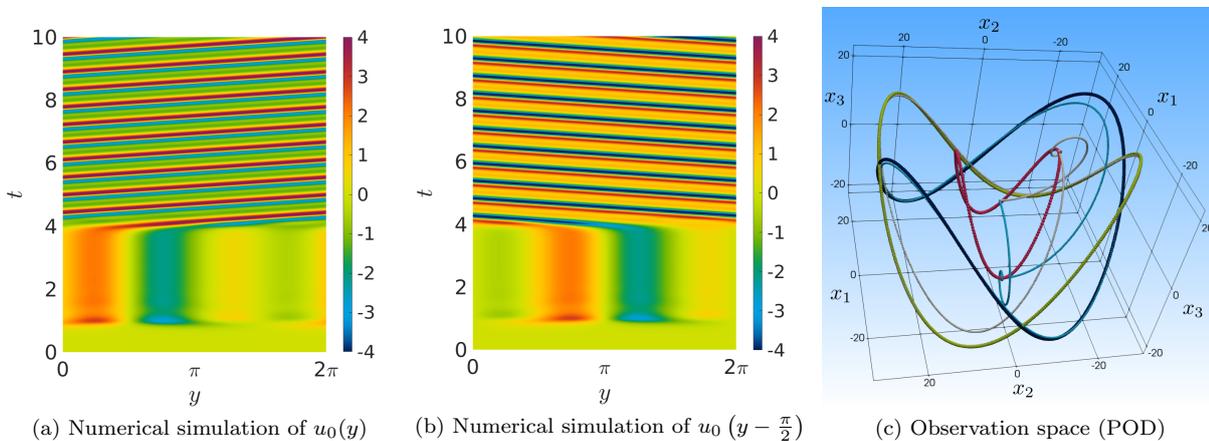


Figure 5.6: Direct simulation of the Kuramoto-Sivashinsky equation (5.4) for $\mu = 15$.

(a) The initial condition u_0 is first attracted to another (unstable) steady state v^* , then leaves this state and converges to a traveling wave. (b) Shifting the initial condition by $\frac{\pi}{2}$ leads to convergence to the wave traveling in opposite direction. (c) Corresponding embedding in observation space of both simulations (light blue and gray). Here, the red curve depicts the loop of unstable steady states (v^* and its translates) to which $u_0(x)$, respectively $u_0(x - \frac{\pi}{2})$, is attracted first. The corresponding stable limit cycles are shown in dark blue and green.

Thus, due to the $O(2)$ symmetry there is an entire (topological) circle of unstable steady states that surrounds $u^* = 0$. Note that according to [KNS90] the point v^* emerges in a pitchfork bifurcation already at $\mu = 4$ as a stable steady state but loses its stability at $\mu = 13.005$ to the attracting limit cycles. Numerical analysis of the stability of u^* and v^* shows that u^* has two unstable directions (with the same eigenvalue) whereas v^* has only one. Therefore, according to the long-term simulation analysis (cf. Figure 5.6) we are in the situation of Remark 3.25 (a). More precisely, we expect that Algorithm 6 initialized at $u^* = 0$ first approximates the 2-dimensional unstable manifold of the origin and afterwards continues to cover all the 1-dimensional unstable manifolds of all translations of v^* , i.e., a 2-dimensional manifold. We conclude that we will in fact cover two connected 2-dimensional manifolds. Therefore, assuming that the thickness exponent is zero, we have to choose $k \geq 5$ in order to obtain a one-to-one image and our choice $k = 7$ should be large enough.

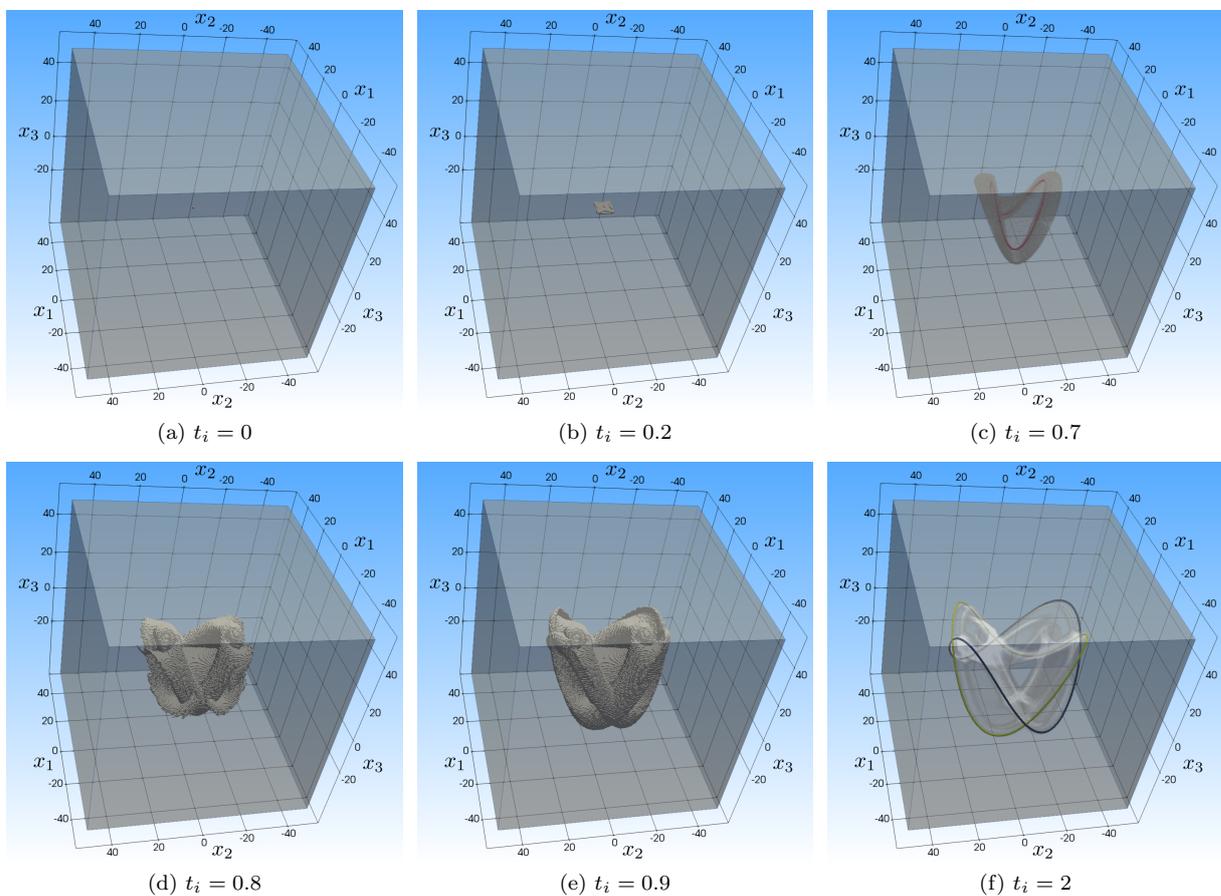


Figure 5.7: Three-dimensional projection of the box covering at level $\ell = 56$ of parts of the embedded unstable manifold of the Kuramoto-Sivashinsky equation (5.4) for several integration times t_i . If necessary we show transparent boxes such that the internal structure is visible. (a) Initial neighborhood of $p = 0$. (a) Points were integrated until $t_i = 0.2$. (c) At $t_i = 0.6$ the manifold accumulates in a loop of unstable equilibria (red). (d-e) The algorithm continues and approximates the unstable manifold of the loop for $t_i = 0.8$ and $t_i = 0.9$, respectively. (f) The manifold accumulates in two periodic orbits (blue and green) at $t_i = 2$.

In Figure 5.7 we show how the embedded manifold(s) grows with respect to the integration time t_i . As expected Algorithm 6 first covers the embedded unstable manifold of the origin (see (a) and (b)), that accumulates in a whole loop of other unstable states (see (c)). Hence, in (d) and (e) we also compute the embedded unstable manifolds of each of those unstable states which eventually accumulate in the two limit cycles (see (f)).

In order to reveal the internal geometric structure of the approximated object, we employ the framework of diffusion maps presented in Section 4.2. Using Algorithm 9 we sample $m = 10^4$ landmarks out of $M \approx 2.5 \cdot 10^6$ box midpoints after 50 iterations. With the help of the machinery developed in Section 4.2.1 we find an optimal value of $\varepsilon \approx 13.3$ (see Figure 5.8 (a)) and an intrinsic dimension of $d_{\text{int}} \approx 2.56$. We note that we slightly overestimate the dimension of the two connected 2-dimensional manifolds which is due to the outer approximation of the object. In fact, the 2-dimensional manifold of the origin spirals into the loop of unstable states v^* and its translates. Hence, our box covering approximates a solid 3-dimensional torus at the boundary of the manifold of the origin, that is, the loop is covered by a solid 3-dimensional torus. To support this observation we estimate the corresponding box-counting dimension of the covering by approximately $d_{\text{box}} \approx 2.66$ (see Figure 5.8 (b)).

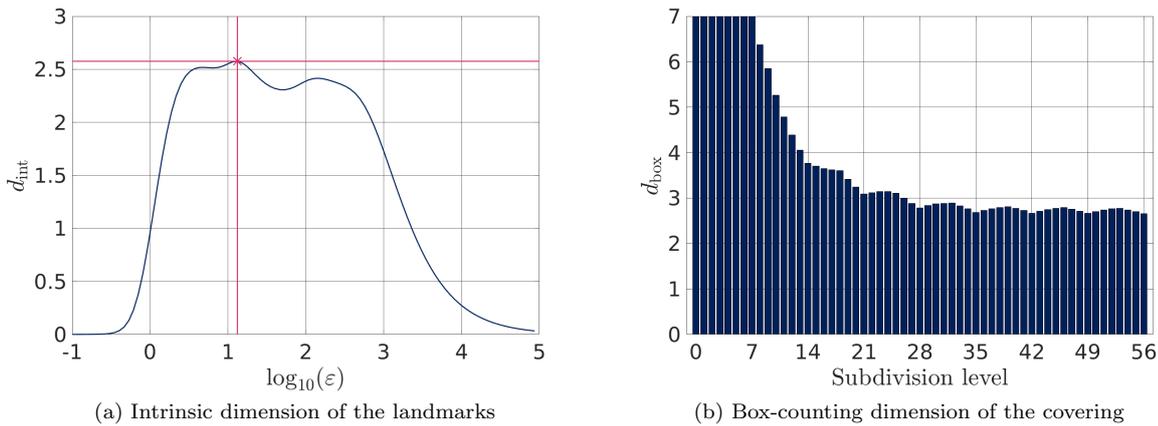


Figure 5.8: Estimation of the dimension of the covering for $\mu = 15$. (a) Intrinsic dimension according to Section 4.2.1. (b) Box-counting dimension (see Definition 2.5). Note that we overestimate the assumed dimension of 2 in both cases which is due to our outer approximation of the set.

In Figure 5.9 we show the discretized embedded manifold(s) and its diffusion coordinates scaled by the constant function ψ_0 . We see that the first two diffusion coordinates like the first two POD coordinates form a circular disc as expected. However, the third diffusion coordinate reveals more structure than the third POD coefficient. In fact, it distinguishes between both limit cycles: $\psi_3 < 0$ represents convergence to the first limit cycle where analogously $\psi_3 > 0$ shows the convergence to the second limit cycle. In addition to that, $\psi_3 = 0$ marks the inner part of the manifold which connects the unstable steady state $u^* = 0$ with the entire loop of unstable steady states (plotted in red), that lie at the boundary of the disk. We observed that the higher order coordinates are so-called higher harmonics, i.e., functions of the first three diffusion coordinates and thus not giving any additional topological information. In conclusion, the shape of the manifold(s) can be described as a cylinder that has a disk inside it cutting it perpendicularly to its cylindrical axis. A corresponding long-term simulation starting near u^* in these coordinates first radially leaves the origin to the boundary of the disc and then spirals up- or downwards on the cylinder to the limit cycles.

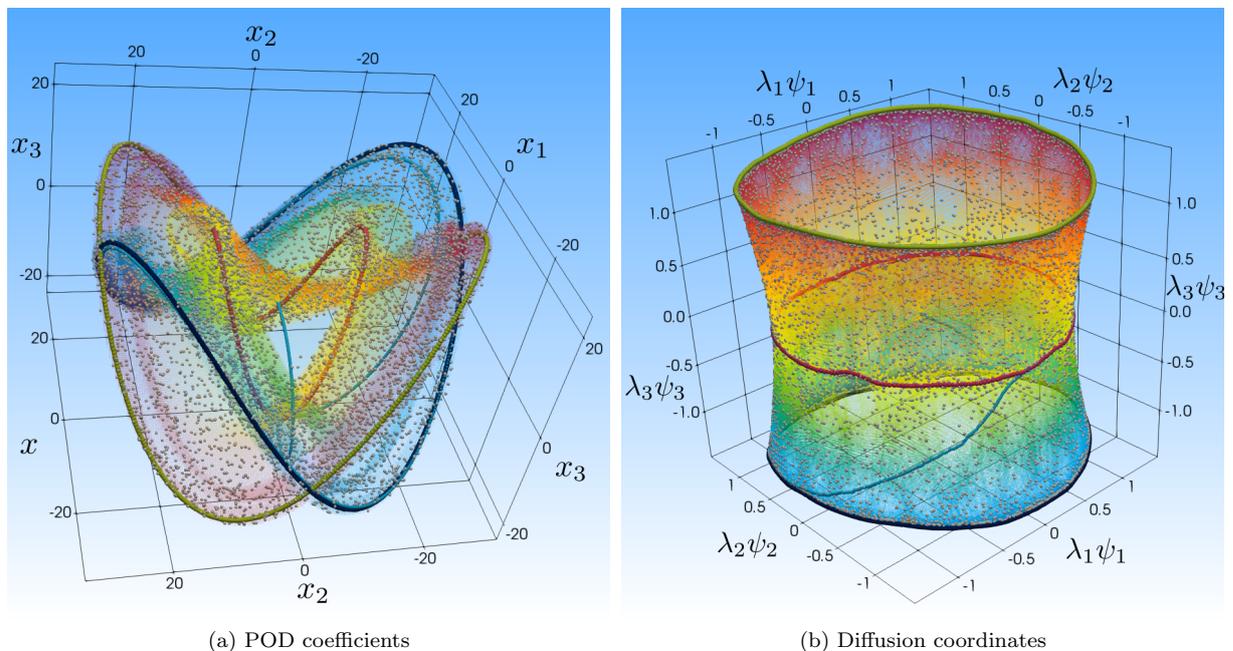


Figure 5.9: Employing the diffusion maps technique to the covering of the embedded manifold(s) for $\mu = 15$. In gray we show the selected landmarks generated by Algorithm 9. The loop of unstable steady states is illustrated in red, whereas a long-term simulation is shown in light blue. (a) POD coordinates of the box midpoints colored in the third diffusion coordinate distinguishing between both limit cycles (blue and green). (b) Diffusion maps embedding of the landmarks and (subsequently) box midpoints by Algorithm 8. The coloring is according to the third diffusion coordinate as well.

5.3.2 The Stable Heteroclinic Cycle

For $\mu = 18$ the observed long-term behavior consists of a pulsation between two states w_1^* and w_2^* , which are $\pi/2$ -translations of each other, that is, $w_2^*(x) = w_1^*(x - \frac{\pi}{2}) = w_2^*(x - \pi)$. In particular, they are symmetric with respect to the reflection (cf. (5.5)). The transients linger close to one of these states for a comparatively long time before they pulse back to the other (cf. Figure 5.10 (a) and (b)). Those states are bimodal and thus the pulsation projected onto the $\cos(2x)$ and $\sin(2x)$ coefficient plane, respectively, appears as a straight line passing through the origin [KNS90]. In addition to that, different pulsations, resulting from different initial conditions, give straight lines that are rotations of each other about the origin. By projecting the pulsation onto the first three POD coefficients, we observe a similar behavior in observation space which is shown in Figure 5.10 (c). Thus, we expect that the unstable manifold will be of dimension at least three. In fact, numerical analysis of the eigenspectrum of the linearization at the origin u^* shows two pairs of unstable eigenvalues. Hence, our chosen embedding dimension $k = 7$ might be too small and we expect to approximate just a projection of the unstable manifold. For related discussion in the finite-dimensional context we refer the reader to [SYC91]. Additionally, we note that according to [KNS90] the unstable states w_i^* are born in a pitchfork bifurcation from the origin u^* at $\mu = 16$ and merge with the unstable states v^* (see the discussion for $\mu = 15$) at $\mu = 16.1399$ in another pitchfork bifurcation. Observe that this bifurcation is invisible to simple simulations since both branches participating are unstable.

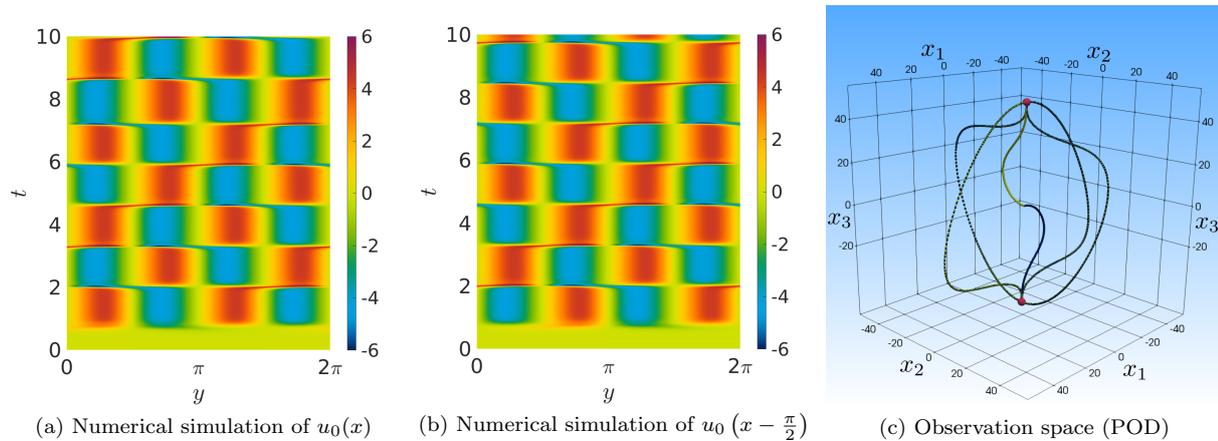


Figure 5.10: Numerical simulation of the Kuramoto-Sivashinsky equation (5.4) for $\mu = 18$. (a) The initial condition u_0^+ pulsates between the (unstable) states w_1^* and $w_2^*(x) = w_1^*(x - \frac{\pi}{2})$ starting with w_1^* . (b) Shifting the initial condition by $\frac{\pi}{2}$ we first approach w_2^* . (c) Corresponding embedding in the observation space of both simulations (blue and green). Here, the red dots depict the states w_1^* and w_2^* .

In Figure 5.11 we show projections of three box coverings obtained by the continuation method 6 for different values $s \in \mathbb{N}$ of the partition \mathcal{P}_s of Q . We see that the object is bounded on the one hand by the long-term simulation (cf. Figure 5.10 (c)) and on the other hand by the $\mathbf{O}(2)$ orbit of w_i^* , which corresponds to a circle in the x_3 - x_4 plane (see Figure 5.11 (d)-(f)). As expected the embedded unstable manifold appears to be a solid three-dimensional object in this projection. This coincides with the observation mentioned above. In fact, the analysis made in the following section will yield a dimension larger than three.

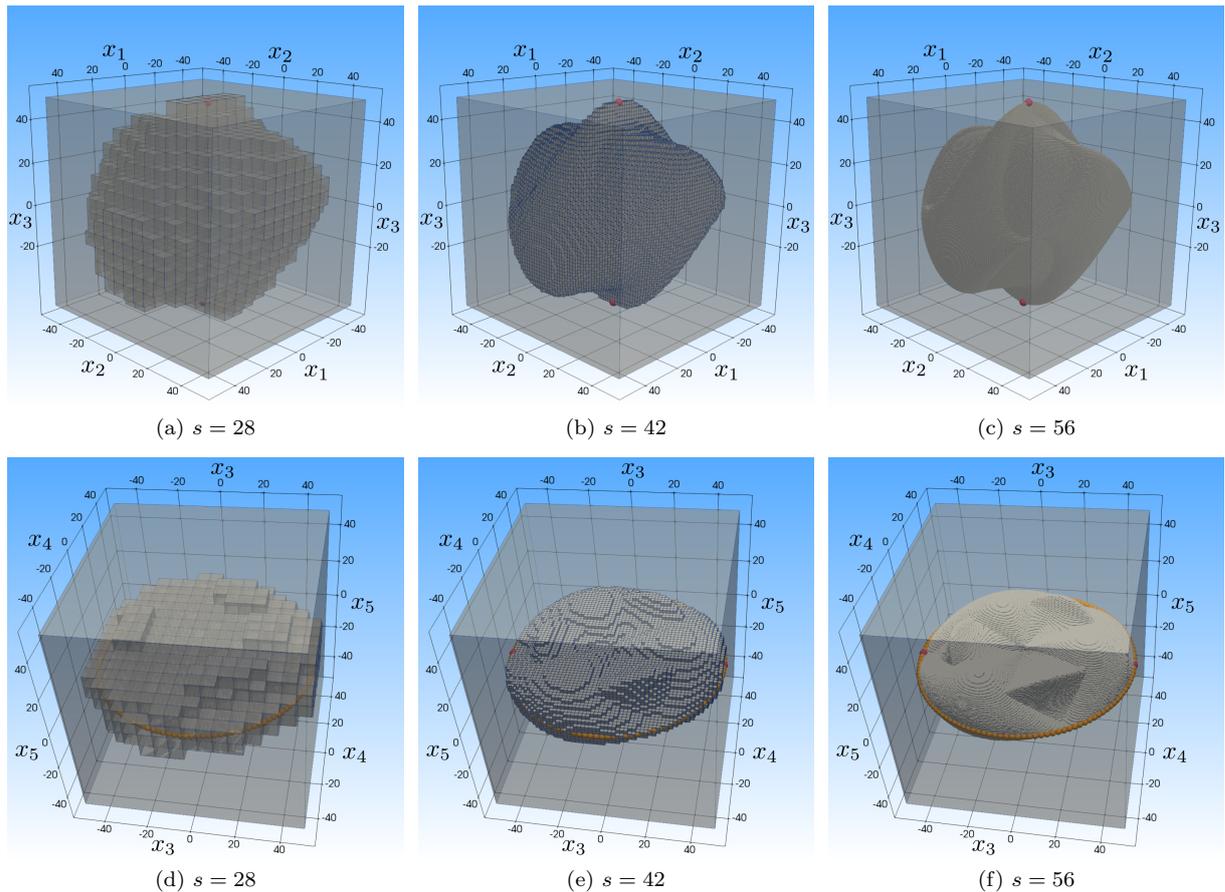


Figure 5.11: Three-dimensional projections of the successively finer box coverings of the embedded unstable manifold of the Kuramoto-Sivashinsky equation (5.4) for $\mu = 18$. The red dots depict the (unstable) states w_i^* , $i = 1, 2$, between which a long-term simulation pulsates. First row (a)-(c): projection onto the first three POD coefficients. Second row (d)-(f): projection onto the coefficients x_3, x_4 and x_5 . The translates of w_i^* are illustrated in orange.

5.3.3 Further Intrinsic Geometric Analysis

Previous research [HNZ86] shows that the unstable manifold strongly changes its structure depending on the parameter μ . To further investigate this behavior we will analyze how the cylindrical shape that is revealed in diffusion coordinates for $\mu = 15$ changes by increasing the parameter. Thus, now our focus lies in trying to follow these three coordinates as long as possible before considering newly arising diffusion coordinates. To this end, we employ Algorithm 6 in order to approximate the embedded invariant manifold(s) of the origin for several parameter value $\mu \in [15, 18]$. Afterwards, we utilize the selection scheme 9 to sample $m = 10^4$ landmarks after 50 iterations. We note that this is a very coarse sampling for larger values of μ since it grows in diameter. For instance, for the final value $\mu = 18$ the covering of the embedded manifold has approximately $M \approx 7 \cdot 10^7$ boxes. However, with this choice we do not rely on a cut-off radius for the kernel and we can consider all pairwise distances, e.g., for the approximation of the intrinsic dimensions. In Figure 5.12 we show the behavior of the estimated intrinsic and box-counting dimensions of the computed embedded invariant manifold for $\mu \in [15, 18]$. We see a clear increase in dimension when the parameter value μ grows. Starting at $d_{\text{int}} \approx 2.58$ and $d_{\text{box}} \approx 2.66$ for $\mu = 15$ the dimensions reach values of $d_{\text{int}} \approx d_{\text{box}} \approx 3$ at around $\mu = 16$ which further goes up to $d_{\text{int}} \approx 3.34$ and $d_{\text{box}} \approx 3.26$ for $\mu = 18$. Note that both dimensions behave similarly as expected. Indeed, we obtain a dimension larger than three for $\mu = 18$ (cf. Section 5.3.2).

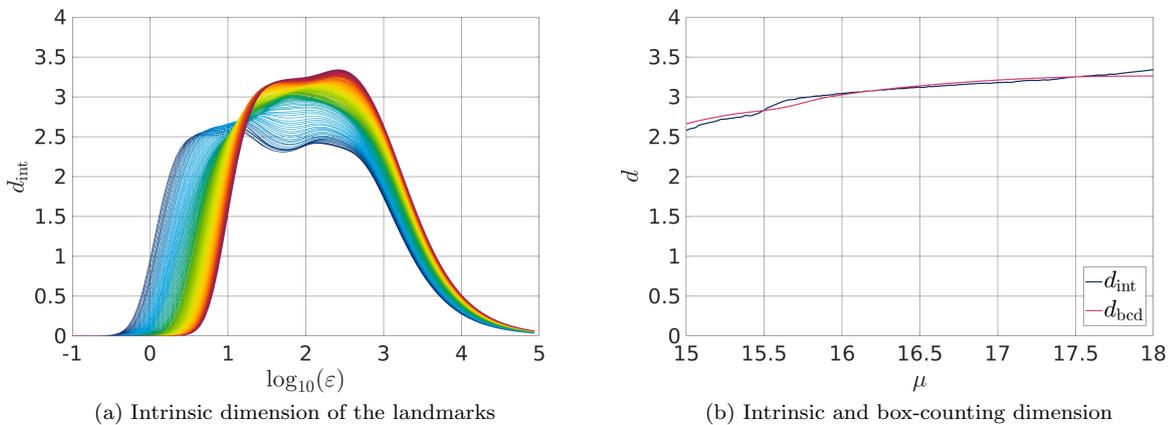


Figure 5.12: Estimation of the dimension of the coverings for $\mu \in [15, 18]$. (a) Local slope approximation of the intrinsic dimension. The coloring is from $\mu = 15$ (blue) to $\mu = 18$ (red). (b) Estimated intrinsic and box-counting dimension.

For the corresponding diffusion maps embedding we compute the first $s + 1 = 10$ dominant eigenvalues of the matrix $P_\varepsilon^{(1)}$ (see step 3) in Algorithm 7, where ε is chosen according to Figure 5.12 (a). In Figure 5.13 we show the embedding for specifically chosen parameter values in the interval of interest. It turns out that the cylindrical shape can only be revealed up to $\mu = 15.46$, where we even have to take larger eigenvalues into account. For larger parameter values the clear separation of the two limit cycles starts to fade. However, another structure that is comparable to the geometry in POD coordinates arises (cf. Figure 5.11).

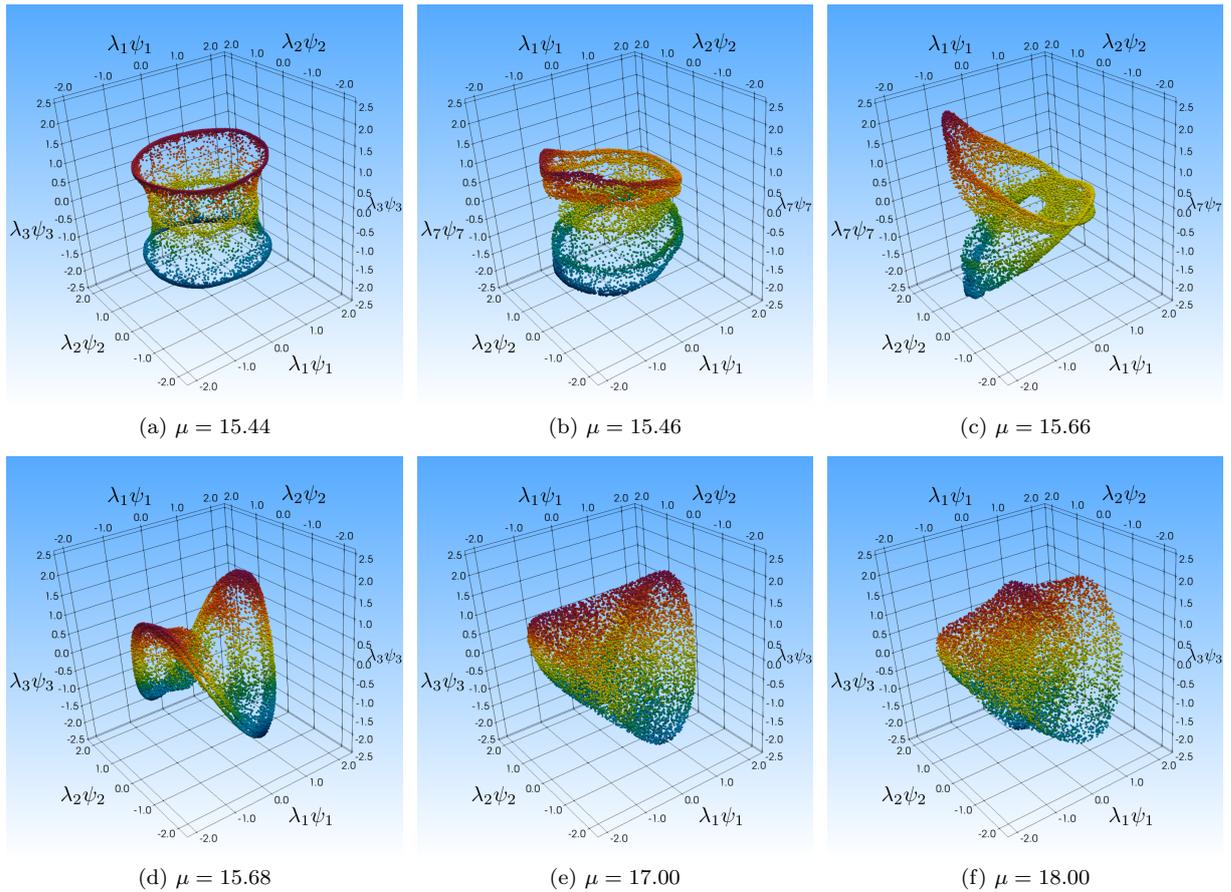


Figure 5.13: Intrinsic geometry of the unstable manifold of the Kuramoto-Sivashinsky equation (5.4) for several parameter values μ . The coloring is according to the chosen third diffusion coordinate. (a)-(b) The cylindrical shape revealed for $\mu = 15$ in Figure 5.9 is still visible. (b) Note that for $\mu = 15.46$ we have to consider the seventh diffusion coordinate already. (c) The clear separation of the two limit cycles is fading. (d) For $\mu = 15.68$ the cylindrical shape is lost, but another structure starts to appear. Note that we consider the third diffusion coordinate again. (e-f) This structure remains dominant up to $\mu = 18$.

Finally, we note that in [ZDG19] the parameter value $\mu = 32$ is also considered. For this parameter the global attractor consists of the union of a two-dimensional unstable manifold of the origin, two one-dimensional unstable manifolds of so-called bi-tri states (which form part of the boundary of the two-dimensional unstable manifold of the origin), a stable bimodal state and an unstable bimodal state whose two-dimensional unstable manifold accumulates in a limit cycle (cf. [JJK01, Figure 1 (c)] and [ZDG19, Figure 10 (b)]).

5.4 Plane Poiseuille Flow

Lastly, we compute the unstable manifold of the *edge state* of a plane Poiseuille flow (PPF) for a Reynolds number of $Re = 650$ [ZE15]. More precisely, we consider the Navier-Stokes equations (NSE) with (infinite) channel geometry where x points in the streamwise, z in the spanwise and y in the wall-normal direction. In addition to that, we employ periodic boundary conditions in x and z and no-slip (Dirichlet) boundary conditions at the wall. We apply a base profile that is parabolic in y , i.e. the laminar profile in dimensionless units is $u_l(t) = [U(y), 0, 0]$ with $U(y) = U_0(1 - y^2)$, where we choose $U_0 = 2/3$. For the calculations we consider a computational domain of length 2π , width π and height 2 and restrict the computations to the subspace that is symmetric to reflections at the midplane and to spanwise reflections at the plane defined by $z = 0$, i.e.,

$$\begin{aligned} s_y[u, v, w](x, y, z) &= [u, -v, w](x, -y, z), \\ s_z[u, v, w](x, y, z) &= [u, v, -w](x, y, -z) \end{aligned}$$

(cf. [ZE15]). In this setting the edge state (cf. Figure 5.14) is a periodic solution with one unstable direction.

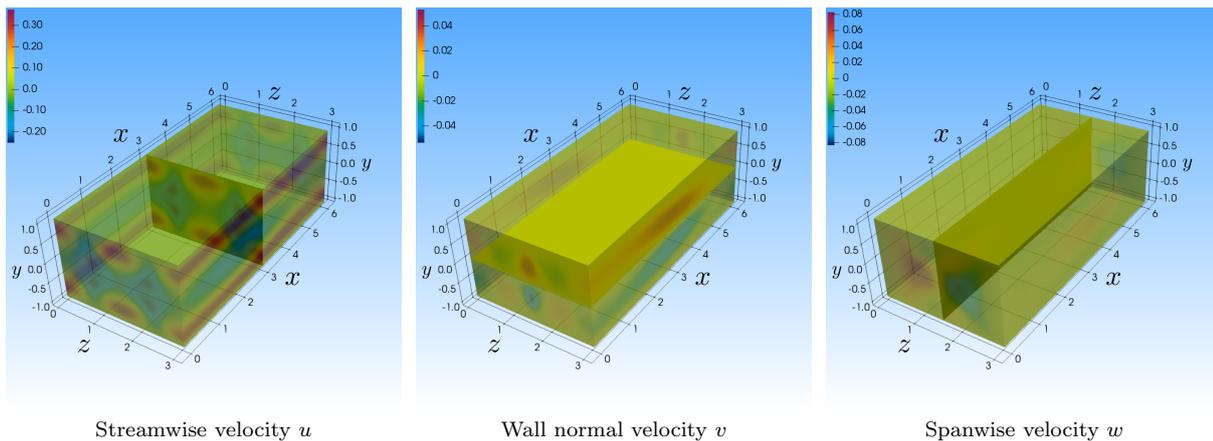


Figure 5.14: Visualization of the edge state in PPF. The coloring is according to the velocity components, respectively. Note that $v = 0$ in the midplane.

This so-called chaotic saddle separates initial conditions that decay directly to the laminar profile and those that swing up to turbulent dynamics [SYE06, SEY07]. Typically, such a state is found by the method of edge tracking described in [SYE06, TI03]. Also note that edge states have practical relevance, for instance in drag reduction [Gra14], and they can also be stabilized [LKZE20].

For the approximation of the embedded unstable manifold of the edge state we first compute a long-term simulation for $T = 700$ time units using a step size of $dT = 1$ which yields 700 snapshots for the computation of the POD basis (see Figure 5.15 for a visualization of its first three modes). Here we choose the parameter S such that $\varepsilon(S) > 0.99$ (cf. (4.18)). Since the edge state has only one unstable direction, an embedding dimension of $k = 5$ should suffice. In fact, the edge state is a periodic solution and, hence, with that choice we also obtain a one-to-one image of the two dimensional manifold of the whole periodic orbit.

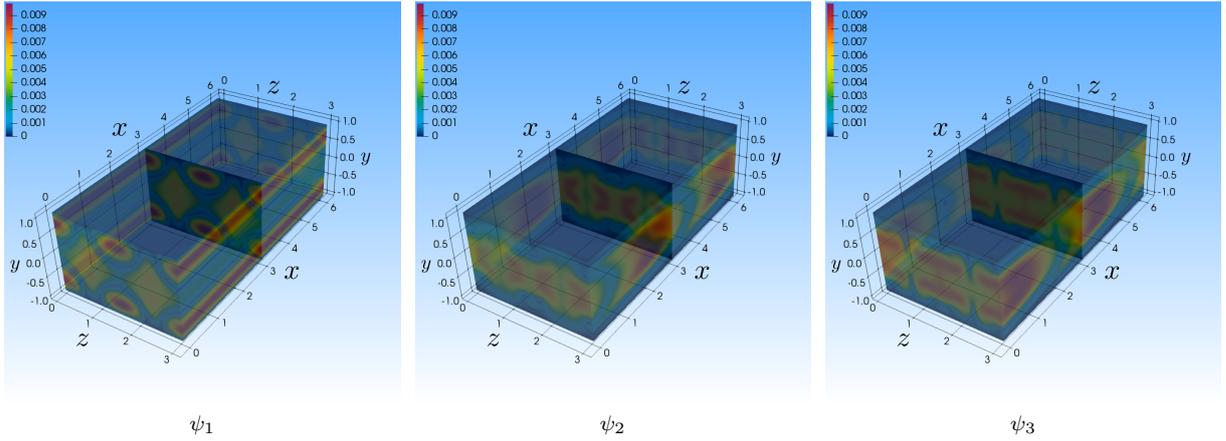


Figure 5.15: Visualization of the first three POD modes obtained by the methods of snapshots. The coloring is according to the magnitude of the flow field.

According to the embedded long-term simulation (see Figure 5.16 (a)) the box

$$Q = [-11, 61] \times [-36, 36] \times [-36, 36] \times [-34, 38] \times [-36, 36] \subseteq \mathbb{R}^5$$

should be large enough for the calculations and we initialize a fine partition \mathcal{P}_s of Q for $s = 45$. As the dynamical system we consider the time- T -map of the flow with $T = 7$ which is approximately the period of the edge state. We also utilize Remark 3.23 (c) and additionally add all the boxes in between, i.e., which are visited during the integration at the time steps $t_i = ih$, $i = 1, \dots, 350$ where $h = 0.02$. This allows us to approximate the manifold of the whole periodic orbit. In Figure 5.16 we show projections of the obtained covering. As expected the embedded manifold connects the edge state with the laminar profile given by the origin in POD coordinates. In the other direction it appears as if the embedded unstable manifold accumulates in another periodic solution.

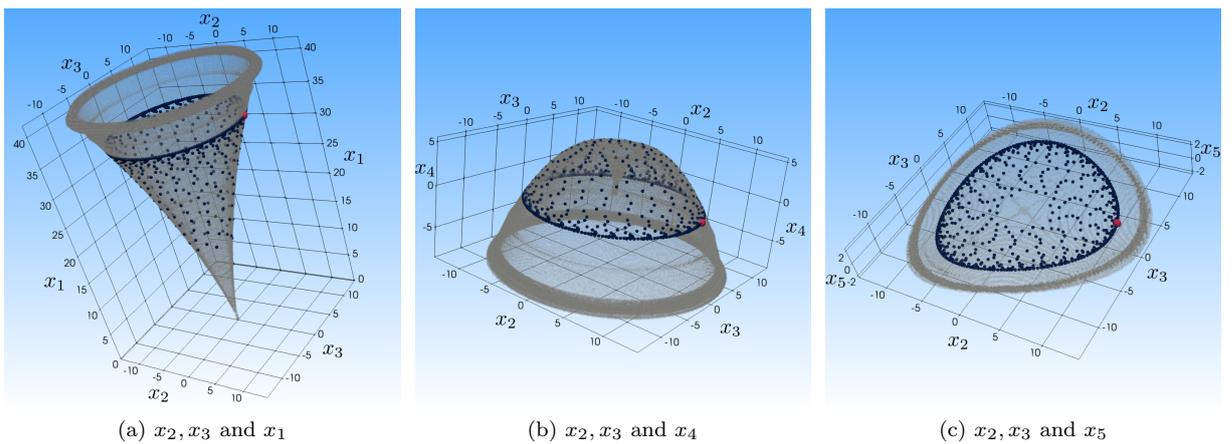


Figure 5.16: Three-dimensional projections of the embedded unstable manifold of the edge state in PPF. Here we show the edge state in red as one point on its periodic orbit, whereas the embedded long-term simulation is shown in blue.

For the sake of completeness we embed the unstable manifold into diffusion coordinates as well. According to Algorithm 9 we sample $m = 10^4$ landmarks out of $M \approx 3.3 \cdot 10^4$ box midpoints after 50 iterations. The dimension estimator developed in Section 4.2.1 yields an optimal bandwidth of $\varepsilon \approx 0.32$ and an intrinsic dimension of $d_{\text{int}} \approx 2.08$ (see Figure 5.17 (a)), whereas the box-counting dimension is estimated by $d_{\text{box}} \approx 2.04$ (see Figure 5.17 (b)). Observe that we almost find the correct dimension of two in both cases.

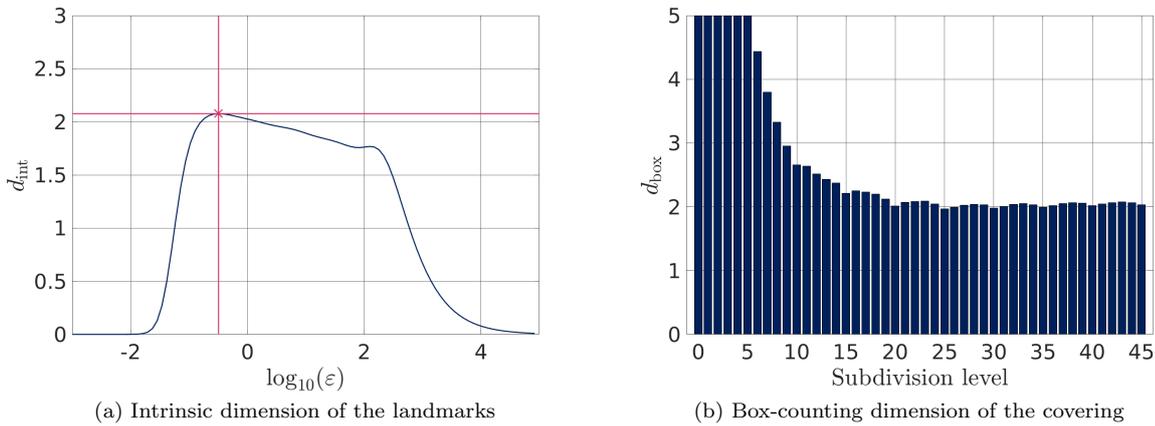


Figure 5.17: Estimation of the dimension of embedded unstable manifold of the edge state. (a) Intrinsic dimension. (b) Box-counting dimension. Observe that the exact dimension of two is only slightly overestimated in both cases.

In Figure 5.18 we show the corresponding diffusion maps embedding scaled by the constant function ψ_0 . It turns out that the coordinates shown in (a) can be compared to the first three POD coefficients (see Figure 5.16 (a)). Analogously, the second, third and fourth diffusion coordinates look similar to the projection shown in Figure 5.16 (b).

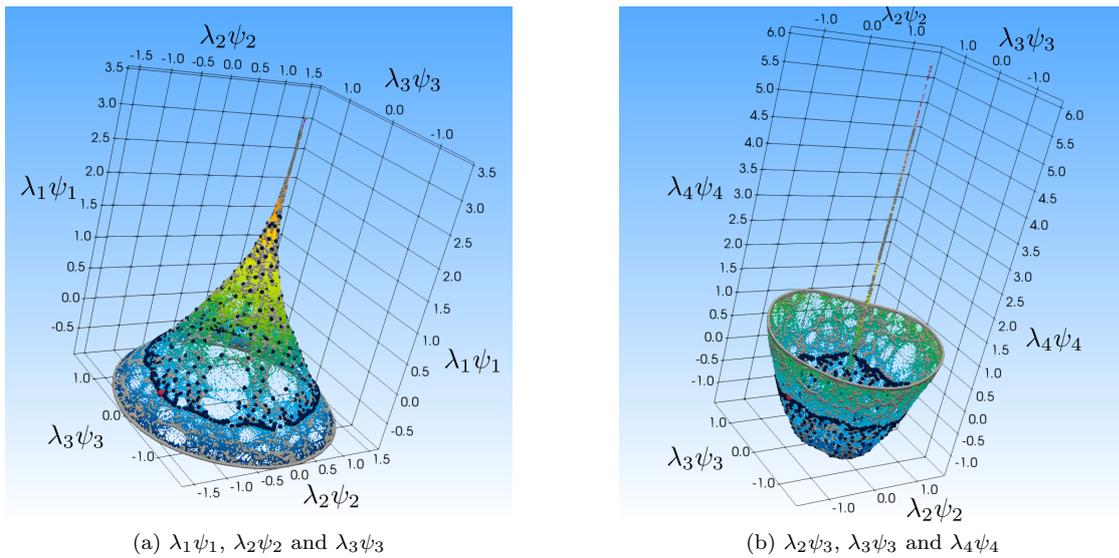


Figure 5.18: Diffusion coordinates of the unstable manifold of the edge state. In gray we show the selected landmarks. The edge state is illustrated in red, whereas a long-term simulation is shown in dark blue. We also embed all box midpoints using Algorithm 8 which we color according to $\lambda_1\psi_1$ in (a) and $\lambda_4\psi_4$ in (b).

6 Conclusion and Outlook

The central goal in this thesis is the numerical analysis of (parameter-dependent) infinite-dimensional dynamical systems. As for finite-dimensional systems one typically aims to characterize the long-term dynamical behavior of its trajectories. To this end, invariant sets such as the global attractor or unstable manifolds crucially influence the complexity of the dynamics. In this thesis we extended the classical set-oriented numerical schemes for the approximation of invariant sets by subdivision and/or continuation techniques to parameter-dependent infinite-dimensional dynamical systems. We utilized embedding techniques for the construction of the dynamically equivalent (*parameter-dependent*) *core dynamical system* defined on a finite-dimensional observation space. This system is then used for the approximation of corresponding embedded invariant sets. Moreover, we extended the subdivision scheme for the approximation of parameter-dependent attractors. This path following method allowed us to efficiently track the attractor with respect to the parameter without restarting the entire procedure anew. In this context we numerically realized a set-valued linearization that served as prediction step followed by a corrector step given by a modified selection step. These one-to-one images of the original invariant sets contained in the infinite-dimensional state space then allow a further analysis regarding their intrinsic geometry. Manifold learning techniques such as diffusion maps are tools that precisely reveal the underlying geometry. In this thesis we adapted the diffusion map scheme such that it can be efficiently applied on the generated coverings. We developed a set-oriented landmark selection scheme that computes points which sample the covering sufficiently well. Furthermore, we presented an intrinsic dimension estimator for point clouds approximating finite-dimensional sets.

6.1 Embedding Techniques

The basis for our analysis of infinite-dimensional systems are the embedding [Theorems 2.18](#) and [2.21](#). Given a finite-dimensional set \mathcal{A} those theorems state that a prevalent set of Lipschitz maps is one-to-one on \mathcal{A} and its corresponding inverse is Hölder continuous, provided the embedding dimension $k \in \mathbb{N}$ is large enough. Hereby, the maximal possible Hölder exponent is controlled by the thickness exponent τ of \mathcal{A} and the chosen embedding dimension (see [\(2.20\)](#) and [\(2.26\)](#)). In the case where the ambient space is a Hilbert space the sharp upper limit for $k \rightarrow \infty$ is $\frac{2}{2+\tau}$. Hence, for zero thickness the bound is one, i.e., \mathcal{A} admits an embedding for any positive $\theta < 1$. This is also true in the context of Banach spaces. However, in [Theorem 2.21](#) it is assumed that the thickness exponent is less than one and it is an interesting open problem whether there is a result that extends [Theorem 2.21](#) for thickness $\tau \geq 1$.

As noted the thickness exponent plays a crucial role for the smoothness of the embedding and hence it is desirable to compute or at least bound the thickness exponent. Even though there are some results that relate the thickness exponent to other quantities such as the Lipschitz deviation [[OR10](#)], the dual thickness exponent [[MR19](#), Section 4] or the box-counting dimension (see [Proposition 2.17](#)), there is no general framework for estimating the thickness exponent. In particular, there is also no numerical scheme for its approximation.

6.1.1 The Core Dynamical System (CDS)

By utilizing the embedding techniques discussed in [Chapter 2](#) we first defined the core dynamical system on the one-to-one image of the finite-dimensional invariant set of interest. Afterwards we employed [Theorem 2.27](#) on the inverse of the observation map in order to extend the CDS on the entire \mathbb{R}^k . The obtained system is a finite-dimensional system that exactly reproduces the original dynamics on the invariant set of the infinite-dimensional dynamical system. However, this approach has two main drawbacks when used for the numerical analysis. On the one hand for each evaluation of the CDS the underlying infinite-dimensional system has to be evaluated as well as opposed to other surrogate models. In general this is computationally expensive since for instance a complex PDE such as the Navier-Stokes equation has to be numerically solved. This problem has been addressed in [[Zie18](#)]. There the numerical effort has been reduced by modifying both the selection step (3.4) in [Algorithm 1](#) and continuation step (3.17) in [Algorithm 6](#). In addition to that, the author combined Koopman Operator based models [[Mez13](#), [WKR15](#), [KKS16](#)] with the CDS for a further numerical improvement. On the other hand it is not really clear how to choose the observation map R and we will discuss this issue in the next section in more detail.

6.2 The Choice of the Observation Map

Even though in principle (almost) every observation map can be used for the construction of the CDS several factors have to be taken into account. First of all, the map R should not be too complex in the sense that its evaluation is computationally less challenging than numerically solving the infinite-dimensional system. In addition, there should be an appropriate realization of the inverse map E such that the conditions in (2.31) are at least approximately satisfied. This is why in this thesis linear observation maps were used which particularly allow for a natural (initial) choice for designing the inverse map E (see [Remark 4.2](#)) that can be improved by using additional information (see [Sections 4.1.1](#) and [4.1.2](#)). For DDEs we observed the state at equidistant points whereas for PDEs we projected the function onto a POD basis and observed the corresponding coefficients with respect to that basis. For simple systems the eigenvalues σ_i containing the amount of information of the POD modes quickly decay such that the basis can be truncated. However, for complex dynamics such as turbulent flows the POD basis may be not suitable anymore. For instance, it is well known that POD neglects small scale structures by construction, regardless of whether these may be important for the dynamics. Hence, other basis such as Fourier modes [[DBS09](#)] or DMD modes [[Sch10](#)] and variations of thereof [[CTR12](#), [WPGG13](#)] may be more appropriate.

6.2.1 Diffusion Maps

In this thesis we first approximated the embedded object of interest using observations such as delay coordinates and POD coefficients. Afterwards, we applied diffusion maps on the generated data to learn the intrinsic geometry. In future research one can aim to approximate the invariant set in diffusion coordinates right away such that we construct the core dynamical system with diffusion maps as the observation map R . To this end, one has to solve several problems.

For nonlinear maps, such as the diffusion maps embedding (see [Algorithm 7](#) and [Algorithm 8](#)), it is in general difficult – or even impossible – to construct an appropriate inverse E , i.e. we have to reverse the diffusion maps embedding. In particular, for given $x \in \mathbb{R}^k$ a point $u \in \mathbb{R}^N$, which is a discretization of the state in the infinite-dimensional space \mathcal{X} , has to be computed such that $R(u) = x$ holds at least approximately, where R is the diffusion map. A first approach to this problem is to compute the nearest images $x_i \in \mathbb{R}^k$ of the landmarks $u_i \in \mathbb{R}^N$ to x and then use the corresponding landmarks $u_i \in \mathbb{R}^N$ to obtain $u \in \mathbb{R}^N$, e.g., by computing their mean. We note that for an exact computation of the unknown point u it is sufficient to know the pairwise distances of u and its neighboring landmarks u_i . In this case one can combine multidimensional scaling (MDS) and procrustes analysis [[Seb09](#)]. First, one computes $\tilde{u} \in \mathbb{R}^N$ and $\tilde{u}_i \in \mathbb{R}^N$ such that their pairwise distances match those of u and u_i using MDS. Then one determines a linear transformation T of the points \tilde{u}_i to best conform them to the points u_i . The desired point u is then given by $u = T(\tilde{u})$.

Moreover, we have to specify an initial set of landmarks in the infinite-dimensional state space (see [Algorithm 7](#)). Note that our proposed landmark selection scheme [9](#) is not applicable for this kind of problem since there is no existing underlying covering (yet). For chaotic dynamical systems a long-term simulation of the system may be used, but for higher dimensions of the invariant set the “uniformity” of samples of the set will play a role. A promising idea though is to update the initial landmarks during the set-oriented schemes. To this end, the kernel, respectively the diffusion coordinates, and the box covering has to be updated whenever the set of landmarks changes, e.g., additional points are added.

The Set-Oriented Landmark Selection Scheme

In [Section 4.2.2](#) we developed a landmark selection scheme that suits for generating samples of a covering that are in some sense equally distributed. The illustrative example in [Figure 4.4](#) showed that the point cloud spreads apart and forms a grid like structure, where the type of the grid depends on the chosen metric. For instance, the euclidean norm in $2d$ generates equilateral triangles as observed in the example, while choosing the Chebychev distance (the maximum metric) instead yields squares in the two dimensional plane. In higher dimensions these structures are translated to equilateral regular hyper tetrahedra and hyper cubes, respectively.

In further research it would be interesting to explore if [Algorithm 9](#) actually converges to such a grid like structure (far from the boundary of the set). In particular, we suggest that this tool may also be helpful for investigating the dispersion problem of a swarm of agents inside a static shape in the plane (see, e.g., [[CP06](#)]).

Bibliography

- [AG93] E. L. Allgower and K. Georg. Continuation and path following. *Acta numerica*, 2:1–64, 1993.
- [AG03] E. L. Allgower and K. Georg. *Introduction to Numerical Continuation Methods*. Society for Industrial and Applied Mathematics, 2003.
- [AHD07] O. Arino, M. L. Hbid, and E. A. Dads. *Delay Differential Equations and Applications: Proceedings of the NATO Advanced Study Institute held in Marrakech, Morocco, 9-21 September 2002*, volume 205. Springer Science & Business Media, 2007.
- [Arn12] V. I. Arnold. *Geometrical methods in the theory of ordinary differential equations*, volume 250. Springer Science & Business Media, 2012.
- [BCC12] F. Brauer and C. Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 2. Springer, 2012.
- [BH16] T. Berry and J. Harlim. Variable bandwidth diffusion kernels. *Applied and Computational Harmonic Analysis*, 40(1):68–96, 2016.
- [BHL93] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25(1):539–575, 1993.
- [BJ70] H. Banks and M. Q. Jacobs. A differential calculus for multifunctions. *Journal of Mathematical Analysis and Applications*, 29(2):246–272, 1970.
- [BK86] D. Broomhead and G. P. King. Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2):217–236, 1986.
- [BM01] C. Bose and R. Murray. The exact rate of approximation in Ulam’s method. *Discrete & Continuous Dynamical Systems*, 7(1):219–235, 2001.
- [BN03] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.
- [BPV⁺04] Y. Bengio, J.-f. Paiement, P. Vincent, O. Delalleau, N. L. Roux, and M. Ouimet. Out-of-sample extensions for LLE, Isomap, MDS, Eigenmaps, and spectral clustering. In *Advances in neural information processing systems*, pages 177–184, 2004.
- [BS87] A. Brandstätter and H. L. Swinney. Strange attractors in weakly turbulent couette-taylor flow. *Physical Review A*, 35:2207–2220, 1987.
- [BSS⁺83] A. Brandstätter, J. Swift, H. L. Swinney, A. Wolf, J. D. Farmer, E. Jen, and P. J. Crutchfield. Low-dimensional chaos in a hydrodynamic system. *Physical Review Letters*, 51:1442–1445, 1983.
- [BZ13] A. Bellen and M. Zennaro. *Numerical methods for delay differential equations*. Oxford university press, 2013.
- [Cas89] M. Casdagli. Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, 35(3):335–356, 1989.
- [CCRFJG11] Y. Chalco-Cano, H. Román-Flores, and M. Jiménez-Gamero. Generalized derivative and π -derivative for set-valued functions. *Information Sciences*, 181(11):2177–2188, 2011.

- [CEES93] P. Collet, J. P. Eckmann, H. Epstein, and J. Stubbe. A global attracting set for the Kuramoto-Sivashinsky equation. *Communications in Mathematical Physics*, 152(1):203–214, 1993.
- [CFNT89] P. Constantin, C. Foias, B. Nicolaenko, and R. Temam. *Integral manifolds and inertial manifolds for dissipative partial differential equations*, volume 70. Springer New York, 1989.
- [CFT85] P. Constantin, C. Foias, and R. Temam. Attractors representing turbulent flows. *Memoirs of the American Mathematical Society*, 53(314), 1985.
- [CG97] F. R. Chung and F. C. Graham. *Spectral graph theory*. American Mathematical Society, 1997.
- [Cha00] A. Chatterjee. An introduction to the proper orthogonal decomposition. *Current Science*, 78(7):808–817, 2000.
- [Chi03] C. Chicone. Inertial and slow manifolds for delay equations with small delays. *Journal of Differential Equations*, 190(2):364–406, 2003.
- [CL06a] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.
- [CL06b] R. R. Coifman and S. Lafon. Geometric harmonics: a novel tool for multi-scale out-of-sample extension of empirical functions. *Applied and Computational Harmonic Analysis*, 21(1):31–52, 2006.
- [CLL⁺05] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the national academy of sciences*, 102(21):7426–7431, 2005.
- [CMRV05] T. Caraballo, P. Marìn-Rubio, and J. Valero. Autonomous and non-autonomous attractors for differential equations with delays. *Journal of Differential Equations*, 208(1):9–41, 2005. in honor of George Sell.
- [Con13] R. Condorelli. Applied nonlinear dynamical system in social science. A nonlinear model for social control system: an application to Italian coercion system. *Quality & Quantity*, 47(2):1173–1198, 2013.
- [CP06] R. Cohen and D. Peleg. Local algorithms for autonomous robot systems. In P. Flocchini and L. Gasieniec, editors, *Structural Information and Communication Complexity*, pages 29–43. Springer Berlin Heidelberg, 2006.
- [CSSS08] R. R. Coifman, Y. Shkolnisky, F. J. Sigworth, and A. Singer. Graph laplacian tomography from unknown random projections. *IEEE Transactions on Image Processing*, 17(10):1891–1899, 2008.
- [CTR12] K. K. Chen, J. H. Tu, and C. W. Rowley. Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analyses. *Journal of Nonlinear Science*, 22(6):887–915, 2012.
- [DBS09] A. Duggeby, K. S. Ball, and M. Schwaenen. Structure and dynamics of low reynolds number turbulent pipe flow. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1888):473–488, 2009.

-
- [DDJS99] P. Deuffhard, M. Dellnitz, O. Junge, and C. Schütte. Computation of essential molecular dynamics by subdivision techniques. In P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel, editors, *Computational Molecular Dynamics: Challenges, Methods, Ideas*, pages 98–115. Springer Berlin Heidelberg, 1999.
- [Del20] L. Dell’Anna. Solvable delay model for epidemic spreading: the case of Covid-19 in Italy. *Scientific reports*, 10(1):1–10, 2020.
- [DFH⁺09] M. Dellnitz, G. Froyland, C. Horenkamp, K. Padberg-Gehle, and A. Sen Gupta. Seasonal variability of the subpolar gyres in the southern ocean: a numerical investigation based on transfer operators. *Nonlinear Processes in Geophysics*, 16(6):655–663, 2009.
- [DFHP09] M. Dellnitz, G. Froyland, C. Horenkamp, and K. Padberg. On the approximation of transport phenomena – a dynamical systems approach. *GAMM-Mitteilungen*, 32(1):47–60, 2009.
- [DFJ01] M. Dellnitz, G. Froyland, and O. Junge. The algorithms behind GAIO — Set oriented numerical methods for dynamical systems. In *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, pages 145–174. Springer-Verlag, 2001.
- [DFS00] M. Dellnitz, G. Froyland, and S. Sertl. On the isolated spectrum of the Perron-Frobenius operator. *Nonlinearity*, 13(4):1171, 2000.
- [DG03] D. L. Donoho and C. Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100(10):5591–5596, 2003.
- [DGHN88] C. R. Doering, J. D. Gibbon, D. D. Holm, and B. Nicolaenko. Low-dimensional behaviour in the complex Ginzburg-Landau equation. *Nonlinearity*, 1(2):279, 1988.
- [DGM⁺05] M. Dellnitz, K. A. Grubits, J. E. Marsden, K. Padberg, and B. Thiere. Set oriented computation of transport rates in 3-degree of freedom systems: the rydberg atom in crossed fields. *Regular and chaotic dynamics*, 10(2):173–192, 2005.
- [DH96] M. Dellnitz and A. Hohmann. The computation of unstable manifolds using subdivision and continuation. In H. W. Broer, S. A. van Gils, I. Hoveijn, and F. Takens, editors, *Nonlinear Dynamical Systems and Chaos*, pages 449–459. Birkhäuser Basel, 1996.
- [DH97] M. Dellnitz and A. Hohmann. A subdivision algorithm for the computation of unstable manifolds and global attractors. *Numerische Mathematik*, 75(3):293–317, 1997.
- [DH12] M. Dellnitz and C. Horenkamp. The efficient approximation of coherent pairs in non-autonomous dynamical systems. *Discrete & Continuous Dynamical Systems*, 32(9):3029–3042, 2012.
- [DHZ16] M. Dellnitz, M. Hessel-von Molo, and A. Ziessler. On the computation of attractors for delay differential equations. *Journal of Computational Dynamics*, 3(1):93–112, 2016.

- [DJ99] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM Journal on Numerical Analysis*, 36(2):491–515, 1999.
- [DJ02] M. Dellnitz and O. Junge. Set oriented numerical methods for dynamical systems. *Handbook of dynamical systems*, 2:221–264, 2002.
- [DJ06] M. Dellnitz and O. Junge. Set oriented numerical methods in space mission design. *Modern Astrodynamics*, 1:127–153, 2006.
- [DJK⁺05] M. Dellnitz, O. Junge, W. S. Koon, F. Lekien, M. W. Lo, J. E. Marsden, K. Padberg, R. Preis, S. D. Ross, and B. Thiere. Transport in dynamical astronomy and multibody problem. *International Journal of Bifurcation and Chaos*, 15(03):699–727, 2005.
- [DJL⁺05] M. Dellnitz, O. Junge, M. W. Lo, J. E. Marsden, K. Padberg, R. Preis, S. D. Ross, and B. Thiere. Transport of Mars-crossing asteroids from the Quasi-Hilda region. *Physical Review Letters*, 94:231102, 2005.
- [DKZ17] M. Dellnitz, S. Klus, and A. Ziessler. A set-oriented numerical approach for dynamical systems with parameter uncertainty. *SIAM Journal on Applied Dynamical Systems*, 16(1):120–138, 2017.
- [Dri68] R. D. Driver. On Ryabov’s asymptotic characterization of the solutions of quasi-linear differential equations with small delays. *SIAM Review*, 10(3):329–341, 1968.
- [DS88] N. Dunford and J. T. Schwartz. *Linear Operators. Part I: General Theory*. Wiley-Interscience, 1988.
- [Dug51] J. Dugundji. An extension of Tietze’s theorem. *Pacific J. Math.*, 1(3):353–367, 1951.
- [EE18] D. E. Edmunds and W. D. Evans. *Spectral theory and differential operators*. Oxford University Press, 2018.
- [Fah00] M. Fahl. *Trust-region methods for flow control based on reduced order modelling*. PhD thesis, University of Trier, 2000.
- [Fal14] K. Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2014.
- [Far82] J. D. Farmer. Chaotic attractors of an infinite-dimensional dynamical system. *Physica D*, 4:366–393, 1982.
- [FBF77] J. H. Friedman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 3(3):209–226, 1977.
- [FD03] G. Froyland and M. Dellnitz. Detecting and locating near-optimal almost-invariant sets and cycles. *SIAM Journal on Scientific Computing*, 24(6):1839–1863, 2003.
- [FHR⁺12] G. Froyland, C. Horenkamp, V. Rossi, N. Santitissadeekorn, and A. S. Gupta. Three-dimensional characterization and tracking of an agulhas ring. *Ocean Modelling*, 52-53:69–75, 2012.
- [FJK⁺88] C. Foias, M. S. Jolly, I. G. Kevrekidis, G. R. Sell, and E. S. Titi. On the computation of inertial manifolds. *Physics Letters A*, 131(7):433–436, 1988.

-
- [FLS10] G. Froyland, S. Lloyd, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. *Physica D: Nonlinear Phenomena*, 239(16):1527–1541, 2010.
- [FOBK12] K. Flaßkamp, S. Ober-Blöbaum, and M. Kobilarov. Solving optimal control problems by exploiting inherent dynamical systems structures. *Journal of Nonlinear Science*, 22(4):599–629, 2012.
- [For13] O. Forster. *Analysis 2*. Springer Fachmedien Wiesbaden, 2013.
- [FP09] G. Froyland and K. Padberg. Almost-invariant sets and invariant manifolds — connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D: Nonlinear Phenomena*, 238(16):1507–1523, 2009.
- [FPET07] G. Froyland, K. Padberg, M. H. England, and A. M. Treguier. Detection of coherent oceanic structures via transfer operators. *Physical Review Letters*, 98:224503, 2007.
- [FR99] P. K. Friz and J. C. Robinson. Smooth attractors have zero „thickness“. *Journal of mathematical analysis and applications*, 240(1):37–46, 1999.
- [Fre80] H. I. Freedman. *Deterministic mathematical models in population ecology*, volume 57. Marcel Dekker Incorporated, 1980.
- [Fro05] G. Froyland. Statistically optimal almost-invariant sets. *Physica D: Nonlinear Phenomena*, 200(3):205–219, 2005.
- [FS87] J. D. Farmer and J. J. Sidorowich. Predicting chaotic time series. *Physical Review Letters*, 59:845–848, 1987.
- [FST88] C. Foias, G. R. Sell, and R. Temam. Inertial manifolds for nonlinear evolutionary equations. *Journal of Differential Equations*, 73(2):309–353, 1988.
- [Gan71] G. Gandolfo. *Economic dynamics: methods and models*, volume 16. Elsevier, 1971.
- [GH13] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 2013.
- [GKD19] R. Gerlach, P. Koltai, and M. Dellnitz. Revealing the intrinsic geometry of finite dimensional invariant sets of infinite dimensional dynamical systems. *preprint: arXiv:1902.08824*, 2019.
- [Gra14] M. D. Graham. Drag reduction and the dynamics of turbulence in simple and complex fluids. *Physics of Fluids*, 26(10):101301, 2014.
- [GS02] M. Golubitsky and I. Stewart. *The Symmetry Perspective*. Birkhäuser Basel, 2002.
- [GZ20] R. Gerlach and A. Ziessler. The approximation of invariant sets in infinite dimensional dynamical systems. In O. Junge, O. Schütze, G. Froyland, S. Ober-Blöbaum, and K. Padberg-Gehle, editors, *Advances in Dynamics, Optimization and Computation*, pages 66–85. Springer International Publishing, 2020.
- [GZED20] R. Gerlach, A. Ziessler, B. Eckhardt, and M. Dellnitz. A set-oriented path following method for the approximation of parameter dependent attractors. *SIAM Journal on Applied Dynamical Systems*, 19(1):705–723, 2020.

- [Hal64] J. H. Halton. Algorithm 247: Radical-inverse quasi-random point sequence. *Communications of the ACM*, 7(12):701–702, 1964.
- [Hal80] J. K. Hale. Some recent results on dissipative processes. In *Functional differential equations and bifurcation*, pages 152–172. Springer, 1980.
- [Hal10] J. K. Hale. *Asymptotic behavior of dissipative systems*. Number 25 in American Mathematical Society, 2010.
- [Hir12] M. W. Hirsch. *Differential topology*, volume 33. Springer Science & Business Media, 2012.
- [HK99] B. R. Hunt and V. Y. Kaloshin. Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces. *Nonlinearity*, 12(5):1263–1275, 1999.
- [HL93] J. K. Hale and S. M. V. Lunel. *Introduction to Functional Differential Equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, 1993.
- [HLBR12] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, 2012.
- [HNZ86] J. M. Hyman, B. Nicolaenko, and S. Zaleski. Order and complexity in the Kuramoto-Sivashinsky model of weakly turbulent interfaces. *Physica D: Nonlinear Phenomena*, 23(1-3):265–292, 1986.
- [HR89] J. K. Hale and G. Raugel. Lower semicontinuity of attractors of gradient systems and applications. *Annali di Matematica pura ed applicata*, 154(1):281–326, 1989.
- [Hsu87] C. S. Hsu. *Cell-to-Cell Mapping*. Springer New York, 1987.
- [HSY92] B. R. Hunt, T. Sauer, and J. A. Yorke. Prevalence: A translation-invariant „almost every“ on infinite-dimensional spaces. *Bulletin of the American Mathematical Society*, 27(2):217–238, 1992.
- [Huk06] J. Huke. Embedding nonlinear dynamical systems: A guide to Takens’ theorem. *MIMS Preprint*, 2006.
- [JJK01] M. E. Johnson, M. S. Jolly, and I. G. Kevrekidis. The Oseberg transition: Visualization of global bifurcations for the Kuramoto–Sivashinsky equation. *International Journal of Bifurcation and Chaos*, 11(01):1–18, 2001.
- [JK17] O. Junge and I. G. Kevrekidis. On the sighting of unicorns: A variational approach to computing invariant sets in dynamical systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(6):063102, 2017.
- [JKT90] M. Jolly, I. Kevrekidis, and E. Titi. Approximate inertial manifolds for the Kuramoto-Sivashinsky equation: Analysis and computations. *Physica D: Nonlinear Phenomena*, 44(1):38–60, 1990.
- [Jol89] M. S. Jolly. Explicit construction of an inertial manifold for a reaction diffusion equation. *Journal of Differential Equations*, 78(2):220–261, 1989.
- [JR15] T. Jackson and A. Radunskaya. *Applications of dynamical systems in biology and medicine*, volume 158. Springer, 2015.
- [Jun00] O. Junge. Rigorous discretization of subdivision techniques. In *Equadiff 99*, pages 916–918. World Scientific, 2000.

-
- [Jä05] K. Jänich. *Topologie*. Springer Berlin Heidelberg, 2005.
- [Kak40] S. Kakutani. Some characterizations of euclidean space. *Japanese journal of mathematics*, 16(0):93–97, 1940.
- [KH95] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Number 54 in Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1995.
- [KKS16] S. Klus, P. Koltai, and C. Schütte. On the numerical approximation of the Perron-Frobenius and Koopman operator. *Journal of Computational Dynamics*, 3(1):51–77, 2016.
- [Kle13] A. Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013.
- [Klo06] P. E. Kloeden. Upper semi continuity of attractors of delay differential equations in the delay. *Bulletin of the Australian Mathematical Society*, 73(2):299–306, 2006.
- [KM44] S. Kakutani and G. W. Mackey. Two characterizations of real Hilbert space. *Annals of Mathematics*, 45(1):50–58, 1944.
- [KM13] V. Kolmanovskii and A. Myshkis. *Introduction to the theory and applications of functional differential equations*, volume 463. Springer Science & Business Media, 2013.
- [KNS90] I. G. Kevrekidis, B. Nicolaenko, and J. C. Scovel. Back in the saddle again: A computer assisted study of the Kuramoto-Sivashinsky equation. *SIAM Journal on Applied Mathematics*, 50(3):760–790, 1990.
- [KO99] H. Keller and G. Ochs. *Numerical Approximation of Random Attractors*, pages 93–115. Springer New York, 1999.
- [KO03] B. Krauskopf and H. M. Osinga. Computing geodesic level sets on global (un)stable manifolds of vector fields. *SIAM Journal on Applied Dynamical Systems*, 2(4):546–569, 2003.
- [KO07] B. Krauskopf and H. M. Osinga. *Computing Invariant Manifolds via the Continuation of Orbit Segments*, pages 117–154. Springer Netherlands, 2007.
- [KOD⁺05] B. Krauskopf, H. M. Osinga, E. J. Doedel, M. E. Henderson, J. Guckenheimer, A. Vladimirovsky, M. Dellnitz, and O. Junge. A survey of methods for computing (un)stable manifolds of vector fields. *International Journal of Bifurcation and Chaos*, page 763–791, 2005.
- [KT76] Y. Kuramoto and T. Tsuzuki. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Progress of Theoretical Physics*, 55(2):356–369, 1976.
- [KT05] A. K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM Journal on Scientific Computing*, 26(4):1214–1233, 2005.
- [KTZ15] A. A. Khan, C. Tammer, and C. Zălinescu. *Derivatives and Epiderivatives of Set-Valued Maps*, pages 399–508. Springer Berlin Heidelberg, 2015.
- [Kua93] Y. Kuang. *Delay differential equations: with applications in population dynamics*. Academic press, 1993.

- [KV99] K. Kunisch and S. Volkwein. Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition. *Journal of optimization theory and applications*, 102(2):345–371, 1999.
- [KW20] P. Koltai and S. Weiss. Diffusion maps embedding and transition matrix analysis of the large-scale flow structure in turbulent rayleigh–bénard convection. *Nonlinearity*, 33(4):1723–1756, feb 2020.
- [Lad85] O. A. Ladyzhenskaya. Finite-dimensionality of bounded invariant sets for Navier-Stokes systems and other dissipative systems. *Journal of Soviet Mathematics*, 28(5):714–726, 1985.
- [Lee13] J. M. Lee. Smooth manifolds. In *Introduction to Smooth Manifolds*, pages 1–31. Springer, 2013.
- [LF17] A. W. Long and A. L. Ferguson. Landmark diffusion maps (L-dMaps): Accelerated manifold learning out-of-sample extension. *Applied and Computational Harmonic Analysis*, 2017.
- [LKZE20] M. Linkmann, F. Knierim, S. Zammert, and B. Eckhardt. Linear feedback control of invariant solutions in channel flow. *Journal of Fluid Mechanics*, 900:A10, 2020.
- [LLL⁺02] Y. C. Liang, H. P. Lee, S. P. Lim, W. Z. Lin, K. H. Lee, and C. G. Wu. Proper orthogonal decomposition and its applications – part I: Theory. *Journal of Sound and Vibration*, 252(3):527–544, 2002.
- [Lor63] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.
- [LR99] J. A. Langa and J. C. Robinson. Determining Asymptotic Behavior from the Dynamics on Attracting Sets. *Journal of Dynamics and Differential Equations*, 11(2):319–331, 1999.
- [MB20] K. S. Morupisi and C. Budd. An analysis of the periodically forced PP04 climate model, using the theory of non-smooth dynamical systems. *IMA Journal of Applied Mathematics*, 86(1):76–120, 2020.
- [Mez13] I. Mezić. Analysis of fluid flows via spectral properties of the Koopman operator. *Annual Review of Fluid Mechanics*, 45(1):357–378, 2013.
- [MFE04] J. Moehlis, H. Faisst, and B. Eckhardt. A low-dimensional model for turbulent shear flows. *New Journal of Physics*, 6(1):56, 2004.
- [MFE05] J. Moehlis, H. Faisst, and B. Eckhardt. Periodic orbits and chaotic sets in a low-dimensional model for shear flows. *SIAM Journal on Applied Dynamical Systems*, 4(2):352–376, 2005.
- [MG77] M. C. Mackey and L. Glass. Oscillation and chaos in physiological control systems. *Science*, 197(4300):287–289, 1977.
- [MN96] J. Málek and J. Nečas. A finite-dimensional attractor for three-dimensional flow of incompressible fluids. *Journal of Differential Equations*, 127(2):498–518, 1996.
- [MR19] A. Margaritis and J. C. Robinson. Embedding properties of sets with finite box-counting dimension. *Nonlinearity*, 32(10):3523–3547, 2019.

-
- [MS03] K. Mittmann and I. Steinwart. On the existence of continuous modifications of vector-valued random fields. *Georgian Mathematical Journal*, 10(2):311–317, 2003.
- [MW96] A. Miranville and X. Wang. Upper bound on the dimension of the attractor for nonhomogeneous Navier-Stokes equations. *Discrete & Continuous Dynamical Systems*, 2(1):95–110, 1996.
- [NP02] P. W. Nelson and A. S. Perelson. Mathematical analysis of delay differential equation models of HIV-1 infection. *Mathematical biosciences*, 179(1):73–94, 2002.
- [OHK06] W. Ott, B. Hunt, and V. Kaloshin. The effect of projections on fractal sets and measures in Banach spaces. *Ergodic Theory and Dynamical Systems*, 26(3):869–891, 2006.
- [OR10] E. J. Olson and J. C. Robinson. Almost bi-lipschitz embeddings and almost homogeneous sets. *Transactions of the American Mathematical Society*, 362(01):145–168, 2010.
- [OY05] W. Ott and J. Yorke. Prevalence. *Bulletin of the American Mathematical Society*, 42(3):263–290, 2005.
- [Ped99] M. Pedersen. *Functional analysis in applied mathematics and engineering*, volume 31. CRC press, 1999.
- [Pei17] S. Peitz. *Exploiting structure in multiobjective optimization and optimal control*. PhD thesis, Paderborn University, 2017.
- [PGK02] M. Pauly, M. Gross, and L. Kobbelt. Efficient simplification of point-sampled surfaces. In *IEEE Visualization, 2002. VIS 2002.*, pages 163–170, 2002.
- [Pop00] S. B. Pope. *Turbulent Flows*. Cambridge University Press, 2000.
- [PWS⁺11] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé. Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics*, 134(17):174105, 2011.
- [Rob94] J. C. Robinson. Inertial manifolds for the Kuramoto-Sivashinsky equation. *Physics Letters A*, 184(2):190–193, 1994.
- [Rob99] J. C. Robinson. Global attractors: topology and finite-dimensional dynamics. *Journal of Dynamics and Differential Equations*, 11(3):557–581, 1999.
- [Rob05] J. C. Robinson. A topological delay embedding theorem for infinite-dimensional dynamical systems. *Nonlinearity*, 18(5):2135–2143, 2005.
- [Rob08] J. C. Robinson. A topological time-delay embedding theorem for infinite-dimensional cocycle dynamical systems. *Discrete & Continuous Dynamical Systems-B*, 9(3&4, May):731, 2008.
- [Rob09] J. C. Robinson. Linear embeddings of finite-dimensional subsets of Banach spaces into Euclidean spaces. *Nonlinearity*, 22(4):711–728, 2009.
- [Rob10] J. C. Robinson. *Dimensions, Embeddings, and Attractors*. Cambridge Tracts in Mathematics. Cambridge University Press, 2010.
- [Rob13] J. C. Robinson. Attractors and finite-dimensional behaviour in the 2d Navier-Stokes equations. *ISRN Mathematical Analysis*, 2013:291823, 2013.

- [Row05] C. W. Rowley. Model reduction for fluids, using balanced proper orthogonal decomposition. *International Journal of Bifurcation and Chaos*, 15(03):997–1013, 2005.
- [RR06] M. Renardy and R. C. Rogers. *An introduction to partial differential equations*, volume 13. Springer Science & Business Media, 2006.
- [RS00] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [SBDH03] J. Stark, D. S. Broomhead, M. E. Davies, and J. Huke. Delay embeddings for forced systems. II. Stochastic forcing. *Journal of Nonlinear Science*, 13(6):519–577, 2003.
- [Sch10] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.
- [Seb09] G. A. F. Seber. *Multivariate observations*, volume 252. John Wiley & Sons, 2009.
- [SEY07] T. M. Schneider, B. Eckhardt, and J. A. Yorke. Turbulence transition and the edge of chaos in pipe flow. *Physical Review Letters*, 99:034502, 2007.
- [SFM10] N. Santitissadeekorn, G. Froyland, and A. Monahan. Optimally coherent sets in geophysical flows: A transfer-operator approach to delimiting the stratospheric polar vortex. *Phys. Rev. E*, 82:056311, 2010.
- [SHD01] C. Schütte, W. Huisinga, and P. Deuffhard. Transfer operator approach to conformational dynamics in biomolecular systems. In B. Fiedler, editor, *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, pages 191–223. Springer Berlin Heidelberg, 2001.
- [Sir87] L. Sirovich. Turbulence and the dynamics of coherent structures part I: Coherent structures. *Quarterly of applied mathematics*, 45(3):561–571, 1987.
- [Siv77] G. Sivashinsky. Nonlinear analysis of hydrodynamic instability in laminar flames—I. Derivation of basic equations. *Acta Astronautica*, 4(11):1177–1206, 1977.
- [SK20] C. Schlosser and M. Korda. Converging outer approximations to global attractors using semidefinite programming. *preprint: arXiv:2005.03346*, 2020.
- [ST06] S. Siegmund and P. Taraba. Approximation of box dimension of attractors using the subdivision algorithm. *Dynamical Systems*, 21(1):1–24, 2006.
- [Sta99] J. Stark. Delay embeddings for forced systems. I. Deterministic forcing. *Journal of Nonlinear Science*, 9(3):255–332, 1999.
- [Ste70] E. M. Stein. *Singular integrals and differentiability properties of functions*, volume 2. Princeton university press, 1970.
- [Ste00] I. Stewart. The Lorenz attractor exists. *Nature*, 406(6799):948–949, 2000.
- [SV10] E. W. Sachs and S. Volkwein. POD-Galerkin approximations in PDE-constrained optimization. *GAMM-Mitteilungen*, 33(2):194–208, 2010.
- [SYC91] T. Sauer, J. A. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65(3):579–616, 1991.
- [SYE06] J. D. Skufca, J. A. Yorke, and B. Eckhardt. Edge of chaos in a parallel shear flow. *Physical Review Letters*, 96:174101, 2006.

-
- [Tak81] F. Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer, 1981.
- [TDSL00] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.
- [Tem97] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, volume 68. Springer New York, 1997.
- [Tes12] G. Teschl. *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012.
- [TI03] S. Toh and T. Itano. A periodic-like solution in channel flow. *Journal of Fluid Mechanics*, 481:67–76, 2003.
- [Tuc99] W. Tucker. The Lorenz attractor exists. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 328(12):1197–1202, 1999.
- [Tur91] G. Turk. Generating textures on arbitrary surfaces using reaction-diffusion. *SIGGRAPH Computer Graphics*, 25(4):289–298, 1991.
- [Tur92] G. Turk. Re-tiling polygonal surfaces. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 55–64, 1992.
- [Tur08] P. Turchin. Arise 'cliodynamics'. *Nature*, 454(7200):34–35, 2008.
- [TvBP13] T. Tervonen, G. van Valkenhoef, N. Bağtürk, and D. Postmus. Hit-and-run enables efficient weight generation for simulation-based multiple criteria decision analysis. *European Journal of Operational Research*, 224(3):552–559, 2013.
- [TW94] R. Temam and X. M. Wang. Estimates on the lowest dimension of inertial manifolds for the Kuramoto-Sivashinsky equation in the general case. *Differential Integral Equations*, 7(3-4):1095–1108, 1994.
- [Ula60] S. M. Ulam. *A collection of mathematical problems*. Interscience Publishers, New York, 1960.
- [Vol11] S. Volkwein. Model reduction using proper orthogonal decomposition. *Lecture Notes*, pages 1–43, 2011.
- [Wal95a] F. Waleffe. Hydrodynamic stability and turbulence: Beyond transients to a self-sustaining process. *Studies in applied mathematics*, 95(3):319–343, 1995.
- [Wal95b] F. Waleffe. Transition in shear flows. nonlinear normality versus non-normal linearity. *Physics of Fluids*, 7(12):3060–3066, 1995.
- [Wer11] D. Werner. *Funktionalanalysis*. Springer Berlin Heidelberg, 2011.
- [Whi36] H. Whitney. Differentiable manifolds. *Annals of Mathematics*, pages 645–680, 1936.
- [Whi44] H. Whitney. The self-intersections of a smooth n -manifold in $2n$ -space. *Annals of Mathematics*, 45(2):220–246, 1944.
- [Wig88] S. Wiggins. *Global Bifurcations and Chaos*. Springer New York, 1988.

- [WKR15] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley. A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, Dec 2015.
- [WPGG13] A. Wynn, D. S. Pearson, B. Ganapathisubramani, and P. J. Goulart. Optimal mode decomposition for unsteady flows. *Journal of Fluid Mechanics*, 733:473–503, 2013.
- [ZDG19] A. Ziessler, M. Dellnitz, and R. Gerlach. The numerical computation of unstable manifolds for infinite dimensional dynamical systems by embedding techniques. *SIAM Journal on Applied Dynamical Systems*, 18(3):1265–1292, 2019.
- [ZE15] S. Zammert and B. Eckhardt. Crisis bifurcations in plane Poiseuille flow. *Physical Review E*, 91:041003, 2015.
- [ZGD20] A. Ziessler, R. Gerlach, and M. Dellnitz. A set-oriented predictor-corrector method for parameter dependent delay differential equations. *in preparation*, 2020.
- [Zie18] A. Ziessler. *Analysis of infinite dimensional dynamical systems by set-oriented numerics*. PhD thesis, Paderborn University, 2018.
- [ZZ04] Z. Zhang and H. Zha. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM journal on scientific computing*, 26(1):313–338, 2004.