### Universidad Nacional de Colombia Université de Mons

Contribution to the computation of regions of attraction of nonlinear systems based on the extended dynamic mode decomposition -Application to the anaerobic digestion

Author Camilo García Tenorio Supervisors Eduardo Mojica Nava Alain Vande Wouwer

March 29, 2021



### Abstract

The main topic of the thesis is the data-driven identification of the region of attraction (ROA) of asymptotically stable equilibrium points. Although this is the main computational contribution, satisfying the underlying conditions to make this possible constitutes most of the work of the thesis. To achieve an accurate data-driven approximation of the ROA in systems with multiple fixed or equilibrium points it is necessary to properly complete a series of steps parting from some trajectories of the system, i.e., assuming there is no access to the differential or difference model equation. The main condition is an accurate approximation of the Koopman operator because it provides a set of eigenfunctions where a particular composition of them gives another non-trivial eigenfunction with an associated eigenvalue that is unitary. The main property of this eigenfunction is that it gives the stable manifold of saddle points in the boundary of the ROA, where this stable manifold is in fact, the actual boundary of the ROA. Therefore, for this whole procedure to work, it it also necessary to have an approximation of the location and stability of the fixed points of the system, recalling that the only input to the algorithm is a set of trajectories of the system. Consequently, the algorithm must be an appropriate approximation of the dynamics of the system and be able to provide a difference equation able to give the location and stability of fixed points upon further traditional non-linear system analysis. The algorithm that has the potential to achieve these requisites is the extended dynamics mode decomposition (EDMD) algorithm, where most of the work of this thesis focuses in transforming the potential into actual. For the most part, the development focus is on the numerical stability of the algorithm, reducing the computational effort and necessary steps to perform the approximation. Techniques such as the p-q-quasi norm

reduction of orthogonal polynomials and polynomial element elimination according to its error, ensures that smaller bases perform the approximations while guaranteeing the existence of solutions because of the orthogonality property. Improvements such as the recovery of the state via the inverse of univariate order-one polynomials reduce the number of necessary matrix inversions. Finally, a priori expansions of the state with arbitrary trigonometric functions or any other kind of elemental functions, expand the possible types of systems that the algorithm can handle. As a consequence of these improvements, the thesis achieves the original objectives of analyzing systems and controlling sets of interconnected systems in a data-driven context. Finally, the main application of the thesis is the analysis of the ROA to the anaerobic digestion process, where the analysis of multi-stability phenomena that guarantees the proper operation of the reactor is of paramount importance.

## Contents

1	Inti	roducti	ion	1
<b>2</b>	Lite	erature	e Review	13
	2.1	Region	ns of Attraction	13
	2.2	Koopr	nan Operator & EDMD	17
	2.3	Interc	onnected Systems	18
	2.4	Summ	ary	21
3	Dyı	namica	l Systems, ROA & EDMD	<b>23</b>
	3.1	Discre	ete-Time Non-linear Systems	24
	3.2	Region	ns of Attraction	28
	3.3	The K	Coopman Operator	31
		3.3.1	Koopman Operator Example	33
	3.4	Exten	ded Dynamic Mode Decomposition	35
		3.4.1	Observables	39
		3.4.2	Reduction by p-q-quasi Norms	41
		3.4.3	Reduction by Polynomial Accuracy	43
		3.4.4	EDMD Example	44
		3.4.5	Approximating the Koopman Operato	50
	3.5	EDMI	D for Control	52

<ul> <li>3.6</li> <li>3.7</li> <li>3.8</li> <li><b>App</b></li> <li>4.1</li> </ul>	3.5.2 Interce 3.6.1 3.6.2 3.6.3 Bioche Summ Fixed	Inverted Pendulum: Experimental Results	<ul> <li>61</li> <li>64</li> <li>65</li> <li>67</li> <li>75</li> <li>77</li> <li>81</li> <li>83</li> </ul>
<ul> <li>3.6</li> <li>3.7</li> <li>3.8</li> <li><b>App</b></li> <li>4.1</li> </ul>	Interce 3.6.1 3.6.2 3.6.3 Bioche Summ Fixed	onnected Dynamical Systems	<ul> <li>64</li> <li>65</li> <li>67</li> <li>75</li> <li>77</li> <li>81</li> <li>83</li> </ul>
3.7 3.8 <b>App</b> 4.1	3.6.1 3.6.2 3.6.3 Bioche Summ Fixed	A Case Study: Two Duffing Oscillators Development of the EDMD Algorithm Controlling the Interconnection	65 67 75 77 81 <b>83</b>
3.7 3.8 <b>App</b> 4.1	3.6.2 3.6.3 Bioche Summ <b>Droxim</b> Fixed	Development of the EDMD Algorithm Controlling the Interconnection	67 75 77 81 <b>83</b>
<ul><li>3.7</li><li>3.8</li><li>App</li><li>4.1</li></ul>	3.6.3 Bioche Summ <b>proxim</b> Fixed	Controlling the Interconnection	75 77 81 <b>83</b>
3.7 3.8 <b>App</b> 4.1	Bioche Summ <b>proxim</b> Fixed	ary	77 81 <b>83</b>
3.8 <b>App</b> 4.1	Summ proxim Fixed	ary	81 <b>83</b>
<b>Ap</b> 4.1	oroxim Fixed	ating the ROA via Koopman	83
4.1	Fixed		
		Points	83
	4.1.1	Fixed Points Example	85
4.2	Stabili	ity of Fixed Points	86
	4.2.1	Stability Example	87
4.3	Appro	eximation of the ROA boundary	88
	4.3.1	ROA Example	93
4.4	Algori	thm $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	95
4.5	Summ	nary	97
4.6	Additi	ional Examples	97
	4.6.1	Lotka-Volterra Model	98
	4.6.2	Mass Action Kinetics	101
4.7	Discus	ssion	105
Ana	aerobic	e Digestion Process	109
5.1	Proble	em Statement	109
5.2	Anaer	obic Digestion Analysis	113
	5.2.1	Training and Testing Data	114
	5.2.2	Approximating the Koopman operator	117
	<ul> <li>4.2</li> <li>4.3</li> <li>4.4</li> <li>4.5</li> <li>4.6</li> <li>4.7</li> <li><b>Ana</b></li> <li>5.1</li> <li>5.2</li> </ul>	<ul> <li>4.1 Fixed</li> <li>4.1 4.2 Stabil</li> <li>4.2.1</li> <li>4.3 Approx</li> <li>4.3 Approx</li> <li>4.3.1</li> <li>4.4 Algori</li> <li>4.5 Summ</li> <li>4.6 Addit:</li> <li>4.6.1</li> <li>4.6.2</li> <li>4.7 Discuss</li> <li>Anaerobic</li> <li>5.1 Proble</li> <li>5.2 Anaer</li> <li>5.2.1</li> <li>5.2.2</li> </ul>	<ul> <li>4.1 Fixed Points</li></ul>

6	Conclusio	ns	135
	5.3 Discus	ssion $\ldots$	132
	5.2.4	Approximating the ROA	126
	5.2.3	Discrete-Time Approximation Result	123
	523	Discrete-Time Approximation Result	

# List of Figures

1.1	Ball between two valleys where there are three equilibrium or fixed points: the left one is the desired stable point, in the center there is an unstable point, and the right is the undesired stable point.	2
2.1	Lyapunov vs. Energy function ROA	15
2.2	Backward evolution for the ROA	16
3.1	Discrete-time system with an AS point at the origin, un- stable points at the corners and unstable-saddles at the edges.	25
3.2	Region of attraction as the union of the stable manifolds (solid black lines) of the saddle fixed points	31
3.3	Transversality condition violation by the stable and un- stable manifold of saddle points	32
3.4	Retained basis terms for the low-rank polynomial basis	42
3.5	Training and testing set for the approximation of the discrete- time Koopman operator via the EDMD algorithm with reduced basis	47
3.6	Test orbits from the EDMD algorithm with different p-q truncation schemes for a Laguerre Polynomial basis. Solid lines are the theoretical trajectories and dashed lines are the approximation by the Koopman operator	48

3.7	Test orbits from the EDMD algorithm with polynomial ac- curacy truncation scheme for a Laguerre Polynomial basis. Solid lines are the theoretical trajectories and dashed lines are the approximation by the Koopman operator	50
3.8	Eigenfunction of the discrete-time approximation of the Koopman operator for the Duffing equation with a reduced Laguerre polynomial basis.	51
3.9	Pendulum on a cart	53
3.10	Pendulum trajectories with sinusoidal input	55
3.11	Testing trajectories of the traditional computation of the EDMD algorithm with a Jacobi orthogonal polynomial basis up to order four. Solid lines are the orbits from the numerical integration of the ODE and dashed lines the approximation by the EDMD algorithm.	57
3.12	Testing trajectories of the traditional EDMD computation against the p-q-EDMD of polynomial elements up to order two. Solid lines are the orbits from the numerical integra- tion of the ODE, dash-dotted lines are the approximation of the EDMD algorithm, and dashes lines are the approx- imation by the p-q-EDMD algorithm	58
3.13	Testing Trajectories of the traditional EDMD against the p-q-EDMD with trigonometric embeddings. Solid lines are the orbits from the numerical integration of the ODE, dash-dotted lines are the approximation of the EDMD algorithm, and dashes lines are the approximation by the p-q-Trigonometric EDMD algorithm.	61
3.14	Feedback Digital Pendulum 33-005-PCI	62
3.15	Data filtering of the test set for the p-q-Trigonometric EDMD. Solid lines are the orbits from the experimental set-up and dashed lines are the orbits from the Kalman	
	filter	64

3.16	Testing trajectories of the experimental set-up with the p-q-Trigonometric EDMD. Solid lines are the orbits from the filtered experimental set-up, dash-dotted lines are the approximation of the EDMD algorithm, and dashes lines are the approximation by the p-q-Trigonometric EDMD algorithm.	66
3.17	Subsystem feedback interconnection with the network. $% {\displaystyle \sum} {\displaystyle $	67
3.18	Subsystem local controller with measurement of the inter- connection state.	68
3.19	Two interconnected Duffing oscillators: 4 trajectories con- stitute the training set and 2 trajectories are kept for test- ing purposes.	69
3.20	Test orbits from the non-affine forced EDMD algorithm. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm	72
3.21	Test orbits from the input-affine forced EDMD algorithm. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm	73
3.22	Test orbits from the input-affine forced EDMD algorithm, where the interconnection input is unknown. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm	75
3.23	MPC synthesis for the interconnected Duffing equation.	77
4.1	Initial conditions and location of the fixed points for the Duffing equation	86
4.2	Level plot of the unitary eigenfunction of the discrete-time approximation of the Koopman operator for the Duffing equation problem.	94
4.3	Stable manifold approximation and initial condition clas- sification based on the evaluation of the unitary eigenfunc- tion in the saddle point.	95

4.4	Retained indices for the approximation of the Koopman operator for the Lotka-Volterra model with a choice of $q = 1.1$ and $p = 3$	100
4.5	Trajectories and boundary of the asymptotically stable points of: a) The system differential equation, and b) The Koopman operator and the eigenfunction with unitary as- sociated eigenvalue	101
4.6	Eigenfunctions of the Koopman operator. a) Trivial eigen- function with $\mu = 1$ , b) Constructed eigenfunction $\phi_+$ , c) First eigenfunction for constructing $\phi_+$ , and d) Second eigenfunction for constructing $\phi_+$	102
4.7	Eigenfunctions with unitary associated eigenvalue a) Trivial b) $\mu > 1$ c) $\mu < 1 $	105
4.8	Classification of the initial conditions with respect to the evaluation of the eigenfunctions on the saddle points	106
5.1	Phase plane three different realizations of the anaerobic digestion process	113
5.2	Training and testing sets of snapshots for the approximation of the discrete-time Koopman operator. Initial conditions are uniformly distributed in the boundary of $S_x$ .	116
5.3	Full basis vs. truncated basis	119
5.4	Empirical error for every test of the 3D grid of parameter- izations.	122
5.5	Location of fixed points in the state space from the so- lution of the minimization problem (4.1) from different initial conditions.	125
5.6	Training orbits, testing orbits, and fixed points with their respective approximations based on Koopman linear pre- dictor and analysis algorithms	126
5.7	Theoretical, saddle based classification of initial condi- tions, and eigenfunction with the corresponding saddle	107
	point and stable manifold	127

5.8	Theoretical, saddle based classification of initial condi- tions, and eigenfunction with the corresponding saddle point and stable manifold	128
5.9	Theoretical, saddle based classification of initial condi- tions, and eigenfunction with the corresponding saddle point and stable manifold	129
5.10	Intersection of the stable manifold different realization s with different $u$ , and constant concentrations at the input.	130
5.11	Intersection of the stable manifold different realization s with different $u$ , and constant concentrations at the input.	131
5.12	Intersection of the stable manifold for different realizations with different $x_{in_2}$ , and constant dulition rate, and $x_{in_1}$ concentration at the input.	132

## List of Tables

3.1	Fixed points, stability and $type$ - $k$ for the arbitrary non- linear system (3.2)	28
3.2	Basis for the first state variable $x_1$ . Solution to the Laguerre ODE: $x_1\ddot{y} + (1 - x_1)\dot{y} + \alpha_1 y = 0$ where $\dot{y} = \frac{dy}{dx_1}$ .	40
3.3	Parameters of the inverted pendulum on a cart	54
3.4	Empirical error for each polynomial basis used in the p-q Trigonometric EDMD.	65
3.5	p, q, dimension, and empirical error for the non-affine approximation of the interconnected duffing equation	71
3.6	p, q, dimension, and empirical error for the input-affine approximation of the interconnected duffing equation	74
4.1	Duffing equation fixed points location.	86
4.2	Fixed point stability for the Duffing equation	88
4.3	Lotka-Volterra equation fixed points, location and stability.	100
4.4	Location of fixed points	104
4.5	Stability of fixed points.	104
5.1	Model parameters	112
5.2	Algorithm Parameters.	115
5.3	Basis for the first state variable $x_1$ . Solution to the Jacobi ODE with parameters $\eta = 0.5$ and $\nu = 1. \ldots \ldots$	117

- 5.4 Grid optimization results for different types of polynomials.122

# Chapter 1 Introduction

The concept of stability is ubiquitous in the world, just by riding a bicycle, a person experiences the act of maintaining stability. It is a concept that pervades any type of systems, physical, social, economical, ecological, or biochemical among others. A possible definition of stability is the property of a system that causes it to return to its original condition when it experiences a disturbance from the equilibrium state. Being these equilibrium states the main concept of system analysis and the synthesis of forces or moments that drive them to a desired equilibrium, understanding equilibrium as a condition where a system remains unchanged in the absence of external influences. Within the analysis of these equilibrium states, it is important to determine if they are unstable or stable, and if they are the latter, the analysis turns toward its robustness; the ability of the system to return to its equilibrium state after small disturbances. As a consequence of not having just one stable point, i.e., the multi-stability phenomena, there is a limit to the disturbance such that the system does not converges to another (possibly undesired) stable point. How big is this disturbance limit, gives a measure of the stability of the point, and this measure is related to the concept of region of attraction (ROA). A simple example of the concept of stability and equilibrium points is a ball that can move between two wells, depicted in Figure 1.1. The system has three equilibrium points, the bottom of the two wells and the top of the hill where the wells are stable points and the hill is unstable. Assuming that the objective is to keep the ball in the well

on the left, any disturbance is bounded such that the ball does not pass through the unstable point at the hill. Otherwise it will converge to the undesired well on the right. Meaning that the ROA of the desired state is the area to the left of the unstable point at the hill. These unstable points and the analysis of certain invariant subspaces that they posses are the foundation for the identification of the ROA of stable points.





The subject matter of *control systems* deals with the analysis and synthesis algorithms for these equilibrium states, called equilibrium points in the continuous-time analysis, and fixed points when the system is in discrete-time. Given that the results of this thesis are in discrete-time, the remainder of this thesis will refer to these equilibrium states as fixed points.

When dealing with the analysis of fixed points, systems with unique equilibrium states are rare, the norm are systems that deal with the multistability phenomena. Conversely, the analysis of non-linear systems often focuses on the stability and especially, the global stability of systems with unique fixed points. Whereas the concept of the ROA is as important as its stability. Questions such as: how far a disturbance can drive the system away from an equilibrium point without leaving the region where it can converge back to it? Or, from which initial conditions does the system converge to the desired point? Are as important as determining the local stability of fixed points.

Traditional model-based methods such as Lyapunov, dissipativity, passivity (Energy-based) or input-output based techniques for non-linear systems [Haddad and Chellaboina, 2008, Khalil, 2002, Garcia-Tenorio et al., 2016] are suitable for low dimensional dynamical systems or low-order systems, but they cannot get a handle on complexity [Ahmadi and Parrilo, 2011, Majumdar et al., 2014, Anderson and Papachristodoulou, 2010, Giesl and Hafstein, 2015, Budišić et al., 2012, Skar et al., 1981, Cuesta et al., 1999], not only for the difficulty of the analysis but also because modeling complex systems is a difficult task. Therefore, complex dynamic systems require tools for their analysis that transcend beyond the current knowledge in control theory [Dimirovski, 2016]. The escalating behavior in the number of equilibria, and the modeling difficulties do not allow for a simple and straightforward approach. Still, the overall performance has to be guaranteed.

In order to solve this issue, instead of approaching it with the traditional model based techniques, data-driven frameworks are a suitable alternative to get a solution. Within these data-driven strategies, that include artificial intelligence, machine learning and the whole plethora of algorithms encapsulated by these techniques, the extended dynamic mode decomposition (EDMD) that under certain conditions gives an approximation of the Koopman operator is a promising alternative that makes possible the use of a modern approach where traditional analysis methods still apply.

The Koopman operator framework provides a way of having eigenfunctions  $\phi$  of a linear operator U be observables or measures of a particular dynamical system, i.e., the evolution of eigenfunctions relate to the evolution of the states of the system [Koopman, 1931]. For discrete-time systems, the application of U on  $\phi$  gives the same function scaled by its corresponding eigenvalue  $\mu$  as in

$$U\phi = \mu\phi. \tag{1.1}$$

This approach is suitable for an analysis of the evolution of eigenfunctions rather than the analysis of state evolution [Budišić et al., 2012, Mezić, 2017]. The trade-off with this approach concerns linearity and dimensionality, finite-dimensional nonlinear systems described by an infinitedimensional linear system.

Given that this operator is linear, it has a spectral decomposition that contains information about the underlying non-linear dynamical system. Furthermore, the analysis of eigenfunctions allows the identification of invariant subspaces that capture specific characteristics of the dynamical system. If an eigenfunction has an eigenvalue equal to one, the value of this function is invariant along the trajectories of the dynamical system. Coupled with the fact that finding the ROA of a stable fixed point relies on the ability to approximate some invariant subspaces of the system state space, for a specific type of system, these unitary eigenfunctions give the solution to the multi-stability phenomena.

An advantage of using the Koopman operator is that there are several methods to have a finite-dimensional approximation of this infinitedimensional operator via snapshot data of the system with the application of, among others, the EDMD algorithm [Williams et al., 2015]. A solution that is closely related to modern data-driven and machine learning techniques, where the analysis comes solely from information gathered either from the numerical integration of a nonlinear differential equation, or from the measurements of a real system. In contrast with traditional techniques that rely upon the explicit knowledge of the differential equation that comes from modeling and identification methods [Garnier and Wang, 2008, Augusiak et al., 2014, ElKalaawy and Wassal, 2015], where the identification also relies on the available data of the system, the EDMD avoids the modeling and identification tasks. Moreover, it also gives an approximation of the system dynamics with the additional benefit of having an expanded linear approximation of some measurements or observables of a system, opening up the possibility of adapting traditional linear system analysis techniques to these approximations.

Juxtaposing the analysis of eigenfunctions and the linear evolution of observables provides a set of tools for accurately handling nonlinear dynamical systems. Giving the possibility to find the fixed points, their stability, and finally, analyze the spectral decomposition to determine the ROA of asymptotically stable points. Additionally, the linear evolution of observables allows for the application of linear control techniques such as model predictive control (MPC) for the synthesis of controllers that drive the system into the desired equilibrium [Korda and Mezić, 2018b].

In order to take advantage of these benefits; eigenfunctions and linear predictors, it is necessary to have an accurate approximation of the Koopman operator from the available data. Although, there are several methods to compute the approximation [Williams et al., 2015, 2016, Kaiser et al., 2018, Li et al., 2017], all of those have some inherent difficulties that hinder the fulfillment of the objectives of this thesis. Accordingly most of the development of this thesis focuses in the refinement of the EDMD algorithm using orthogonal polynomial basis as a means to overcome those difficulties. The main contribution in this regard is in the amount of necessary data to calculate the approximation of the discretetime Koopman operator. In contrast to the available methods, the algorithm developed in this thesis achieves an accurate representation with one order of magnitude reduction in the amount of data. Another reason for the improvement of the algorithm is with regards to the dimension of the basis, for the regular formulations the dimension of the observables grows exponentially with the addition of state variables, due to the curse of dimensionality problem, increasing the number of state variables or the maximum order of the polynomials hinders the possibility of calculating the discrete time approximation of the operator. Therefore, the use of p-q-quasi norm reduction methods, based on the proposed ideas for reliability analysis in polynomial chaos expansion [Konakli and Sudret, 2016a,b], gives a truncation scheme on the polynomial basis that serve as the observables for the discrete-time approximation of the EDMD. Moreover, the accuracy of the linear predictors and eigenfunctions is directly related to the contribution of each of the observables on the approximation. Hence, an error criterion on the contribution of these individual elements, in conjunction with an error threshold based on the error of the individual elements of the basis that must remain, serves to eliminate those elements that contribute to the inaccuracies in the approximation.

Indeed, there are some polynomial elements that must remain in the basis since the use of orthogonal polynomials is sufficient to easily recover the state. Some of the available methods include elements in the set of observables that capture directly each of the states, with the risk of breaking the orthogonality of the basis, and making the problem impossible to solve or numerically unstable. Whereas machine learning techniques, such as radially basis functions and Kernel methods need to solve a second minimization problem to recover the state as a linear combination of the observables, at the risk of having inaccurate solutions, or no solution at all.

The solution to this dilemma, of including or not elements that directly capture the state is the selection of univariate and order one injective polynomial elements for the recovery of the state [Garcia-Tenorio et al., 2020]. There is no need to break the orthogonality of the set while still being able to recover the state as a linear function of the observables, without a second optimization solution. In addition to the recovery of the state, the exponential growth of the maximum order and dimension of the set of observables based on orthogonal polynomial has a solution via p-q-quasi norms, and polynomial accuracy methods.

In summary, the methods and algorithms achieve the following analysis objectives:

- 1. Approximate the discrete-time Koopman operator with the least amount of data, and with a reduced set of observables, both in dimension and maximum order of the polynomials.
- 2. Approximate the location of the system fixed points, which is made possible by the linear predictors of the observables that can easily recover the state.
- 3. Accurately determine their stability (asymptotically stable, unstable or saddle points) because the approximation also provides a nonlinear evolution map of the state. This mapping can be linearized and evaluated at the fixed points to asses stability.
- 4. Identify the saddle points in the boundary of the ROA.
- 5. Generate an eigenfunction with unitary associated eigenvalue or in the ideal case, have this eigenfunction present in the original approximation.

6. Find an algebraic criterion that accurately classifies an arbitrary initial condition in the state space to its corresponding attractor, with the benefit that the criterion is a comparison between real numbers, therefore, suitable to handle higher dimensional systems.

It is important to emphasize that the analysis is solely data-driven, and although this thesis makes use of the differential model equations, it is only to produce the data for the analysis. This data-driven presupposition also restricts the analysis to hyperbolic systems, this as that we are not able to characterize a limit cycle or chaotic behavior without the differential equation. Without this restriction, the results of this thesis still apply, with the caveat that there are better alternatives for two dimensional systems, or slices of three dimensional ones using time averages [Mezić, 2005], or for a higher dimensional system, finding the isostables of a dynamical system by the Laplace averages of the differential equation forward integration [Mauroy et al., 2013].

#### Production

1. Journal: Garcia-Tenorio Camilo, Gilles Delansney, Mojica-Nava Eduardo, Vande Wouwer Alain, "Trigonometric Embeddings in Polynomial EDMD and Linear Representations of Interconnected Oscillators" in "Submitted".

Journal presenting the results of applying trigonometric, or in general any functional embeddings into the set of state variables of the system. These embeddings serve to improve the accuracy of the EDMD algorithm by incorporating a priori knowledge of the differential equation in to the set of observables while preserving the benefits of the p-q and polynomial accuracy reduction methods.

2. Journal: Garcia-Tenorio Camilo, Tellez-Castro Duvan, Mojica-Nava Eduardo, Vande Wouwer Alain, "Hyperbolic Systems Region of Attraction via EDMD Koopman Operator" in "Submitted".

Journal paper covering an in depth analysis of the data driven algorithm for obtaining the ROA of asymptotically stable fixed points. The method covers the key aspects in dynamical system analysis for the estimation of the ROA: Finding the fixed points of the system, giving their stability, and estimating the boundary of the ROA for the stable points.

3. Journal: Garcia-Tenorio Camilo, Sbarciog Mihaela, Mojica-Nava Eduardo, Vande Wouwer Alain, "Analysis of the ROA for the Anaerobic Digestion Process via Data-Driven Koopman Operator" in "Accepted in Nonlinear Engineering"

Journal paper covering an in depth analysis of the multi-stability problem for the Anaerobic digestion process. This work also presents in detail the reduction methods for the observables based on orthogonal polynomials.

 Colloquium/Abstract: Garcia Tenorio Camilo, Mojica-Nava Eduardo, Vande Wouwer Alain, "Linear Predictors for Interconnected Systems: a Koopman Operator Approach" in "39th Benelux Meeting on Systems and Control", 128, Elspeet, The Netherlands, 2020.

Abstract that shows the general idea for the linear predictors when dealing with external forcing signals and interconnection inputs.

5. Conference: Garcia-Tenorio Camilo, Tellez-Castro Duvan, Mojica-Nava Eduardo, Vande Wouwer Alain, "Analysis of a Class of Hyperbolic Systems via Data-Driven Koopman Operator" in "23rd International Conference on System Theory, Control and Computing (ICSTCC)", Sinaia, Romania, 2019.

Conference paper that describes some of the methods that achieve the approximation of the ROA for the multi-stability problem. Specializes in the study of polynomial systems under the MAK modeling paradigm for biochemical reaction networks. Precedes the ROA Journal paper submission.

 Colloquium/Abstract: Garcia-Tenorio Camilo, Vande Wouwer Alain, Mojica-Nava Eduardo, "Biochemical Reaction Networks, Dynamics and Regions of Attraction: A Koopman Operator Based Approach" in "38th Benelux meeting on Systems and Control", Lommel, Belgique, 2019.

Abstract that presents the preliminary concepts for the formulation of the solutions based on the Koopman operator for the ROA problem.  Colloquium/Abstract: Garcia-Tenorio Camilo, Vande Wouwer Alain, Mojica-Nava Eduardo, "Order reduction for EDMD based on observable error: Application to the anaerobic digestion model" in "Cape Forum", Liege, Belgique, 2019.

Abstract that contains the first approach for the order and dimensionality reduction of orthogonal polynomials as an observable basis for the approximation of the Koopman operator. Exemplified via the two-dimensional representation of the AD process. Precedes the Journal submission with the complete analysis of the AD process and observables reduction methods.

 Colloquium/Abstract: Garcia-Tenorio Camilo, Sbarciog Mihaela, Mojica-Nava Eduardo, Vande Wouwer Alain, "Anaerobic Digestion, Dynamics and Regions of Attraction: A Koopman Operator Based Approach" in "9th IWA Specialized Conference on Sustainable Viticulture, Winery Wastes &; Agri-industrial Wastewater Management", Mons, Belgium, 2019.

Abstract preceding the Journal paper on the analysis of the ROA for the AD process. The abstract shows the feasibility of using the operator based techniques to analyze the process.

 Conference: Garcia-Tenorio Camilo, Quijano Nicanor, Mojica-Nava Eduardo, Sofrony Jorge, "Bond-Graph model-based for IDA-PBC" in "2016 IEEE Conference on Control Applications (CCA)", Buenos Aires, Argentina, 2016.

Conference paper with preliminary research concerning the use of Bond-Graph models to systematically solve the interconnection and damping assignment problem under passivity based control techniques.

This thesis is organized as follows: Chapter 2 is a literature review that covers some history and important references for the development of the thesis. Chapter 3 contains the main theoretical concepts necessary for the development of the thesis. Section 3.1 contains the theoretical aspects from discrete-time nonlinear dynamical systems. Under this discrete-time framework, section 3.2 shows how to perform the analysis under the multi-stability phenomena to determine the ROA of stable points. Next, Section 3.3 presents the Koopman operator, introduces its spectrum, and gives an example of a particular type of discrete-time dynamical systems that accepts a closed analytical form. Given that this operator is more often approximated than calculated, Section 3.4 shows how to perform the approximation of the discrete part of the operator based on the EDMD algorithm. In addition to the EDMD algorithm, the aforementioned section shows the developed improvements to the algorithm that allow the calculation of more accurate approximation with a reduction of the necessary data. These improvements are the reduction by p-q-quasi norms, polynomial accuracy and trigonometric embeddings. For the latter, Section 3.5 shows a detailed description of how to perform the embeddings along with the algorithmic concepts to handle systems with forcing inputs. In the context of the EDMD algorithm, for input signals and interconnected systems, Section 3.6 shows how to handle the linear predictors coming from the EDMD to handle such combination of signals to synthesize controllers for the interconnected systems. For closing this Chapter, Section 3.7 presents the method of mass action systems for dealing with reaction and their corresponding differential equations. Next, Chapter 4 shows the main results of the thesis regarding the datadriven approximation of the ROA using the tools from the previous chapter, describing how to handle the fixed points in Section 4.1, and their stability in Section 4.2 and the method to analyze the fixed points of the state space via the approximation of the Koopman operator to get the ROA of the stable points. The remaining of the Chapter shows several examples to demonstrate the effectiveness of the algorithm. Chapter 5 shows a detailed analysis of the anaerobic digestion process using the EDMD-Koopman based approach for the estimation of the ROA, giving a detailed description on the different methods proposed by the thesis. Finally, Chapter 6 gives some additional discussion and conclusions regarding the subject matter of the thesis and its development.

**Notation**  $\mathbb{C}$  denotes the field of complex numbers.  $\mathbb{R}$  and  $\mathbb{R}_+$  denote the field of real and nonnegative real numbers, respectively. For any matrix  $A \in \mathbb{R}^{n \times n}$ ,  $A^{\top}$  denotes transpose,  $A^+$  denotes its pseudoinverse, and ||x|| represents the Euclidean norm. For a complex number  $\lambda$ ,  $|\lambda|$  represents its absolute value. For any set A,  $\overline{A}$  denotes its closure. For a set of eigenspaces  $E^i$ ,  $\oplus$  represents their direct sum. The vector exponentiation  $M^{\pm \eta}$  is defined term by term. A level set of an arbitrary function h(x) for any constant c is  $\Gamma(h(x)) = \{x \in \mathbb{R}^n \colon h(x) = c\}.$ 

### Chapter 2

### Literature Review

The main focus of the thesis concerns the approximation of the ROA, more specifically, its boundary via the analysis of eigenfunctions from the Koopman operator. Additionally, some of the most important contributions in this subject matter are in the refinement of the EDMD algorithm that is a suitable data-driven technique for the approximation of the discrete part of the Koopman operator. Finally, the thesis contributes some techniques for handling control synthesis algorithms for interconnected systems, again by improving on the current algorithmic techniques of the EDMD algorithm. Therefore, this literature review will cover the relevant topics in order of importance, that is, regions of attraction, the Koopman operator and EDMD, and finally, interconnected systems.

#### 2.1 Regions of Attraction

The analysis of dynamical systems with multiple equilibrium points cover their location, the local stability characteristics and for Asymptotically Stable (AS) equilibrium points, the ROA that describes the set of initial conditions that converge to a specific AS point. The concept of ROA is crucial for the analysis of a dynamical system, as it is often necessary to maintain the system in one of the AS equilibrium points, and guarantee that any perturbation does not make the system converge into an undesired equilibrium or diverge. Traditionally, the ROA in the multi-stability problem comes from the maximal Lyapunov set. This is a domain that contains the equilibrium point where the local Lyapunov function is positive definite and its derivative along the trajectories is negative definite. Meaning, that the estimate of the ROA is the largest level set of the Lyapunov function where the aforementioned conditions hold [Khalil, 2002], where the constant value for this largest level set is often called the critical level value.

Even though there are several optimization methods to find local Lyapunov functions from linear programing or linear matrix inequalities [Peter Giesl, 2015], this is not an easy task, especially when the dimensionality of the system grows. Additionally, the level sets of the Lyapunov function are a conservative approach, given that the Lyapunov function is not unique, and it is a great challenge to find a function whose level sets maximize the estimate of the ROA.

In addition to the Lyapunov methods, energy function methods give a less conservative approximation of the ROA. These methods are restricted to dynamical systems that accept a global or a local energy function, and the process for either case is similar, the difference between them is in the method to find the critical level values. If the system accepts a global function, the theory is equivalent to the Lyapunov function, that is, to find the critical level value of the energy function where the energy function is non-negative, and its derivative along the trajectories is decreasing. The critical level value comes from evaluating this global energy function in the saddle points of the system and selecting the point with the lowest value after evaluation. If the system accepts a local energy function, the critical level value comes from the evaluation of the closest saddle point with the local energy function. The difference between energy and Lyapunov functions, is that for the latter, the largest level set does not necessarily intercept with a saddle point in the boundary as it is the case with either form of energy functions. Assume there is an arbitrary system where the origin is stable and around these point are several unstable and saddle points. Additionally, assume that there is an arbitrary Lyapunov and energy function. Figure 2.1 shows such a system with the potential outcome of the approximation of the ROA via either method. The derivatives of Lyapunov (left) function along the trajectories of the system are no longer decreasing when outside the square,

while the energy function (right) is less conservative because its critical value corresponds to the function evaluated at the nearest saddles (that are equidistant to the origin).



Figure 2.1: Lyapunov vs. Energy function ROA.

The feasibility of getting a solution with either of these methods is questionable. Starting from the difficulties in getting a function that satisfies the conditions, coupled with the difficulties of calculating level sets, those level sets are points of a n-1 dimensional hyperplane of the state space, where the classification of an arbitrary initial condition on higher than three dimensional systems is not trivial. This classification depends on the geometry of the stable manifold, and the ability to interpolate between the identified points to perform the comparison.

Along with the traditional analysis of the differential equation via Lyapunov or energy functions, there are some methods that rely on the numerical integration of the differential equation in reverse time. This process is summarized as follows: First, locate all the equilibrium points of the system and determine their local stability by the Jacobian matrix. Then, find the stable eigenvectors of the saddle points, next, find the intersection of these eigenvectors with an  $\epsilon$ -ball of the saddle point, i.e., a small *n*-dimensional hyper-sphere around the saddle point, and integrate numerically to check which of the asymptotically stable points ROA is under analysis. Finally, perform the backward integration, or a trajectory reversing algorithm from several initial conditions close enough to the intersection point between the stable eigenvectors and the  $\epsilon$ -ball to get the boundary of the ROA. In conclusion, the method gives a set of trajectories that converge to the saddle point in the boundary. Although this method can yield an accurate approximation of the boundary of the ROA, it is computationally intensive, and has the same geometry and interpolation problems of the Lyapunov and energy function methods. Figure 2.2 shows a summary of this process: from the fixed points (green: stable dot, unstable crosses, saddle asterisks) there is an  $\epsilon$ -ball from the saddles. In the intersection of the stable manifold (red lines) and the  $\epsilon$ -ball (yellow dots) the forward integration of the system gives the stable point under analysis. Finally, The backward integration from initial conditions near the intersection (green dotted lines) gives an approximation of the boundary of the ROA.



Figure 2.2: Backward evolution for the ROA.

In summary, the method gives a set of trajectories that converge to

the saddle point in the boundary. Although this method can yield an accurate approximation of the boundary, it is computationally intensive, and has the same geometry and interpolation problems of the Lyapunov and energy function methods. For a detailed approximation of the ROA in the problem at hand using backward integration we refer to [Sbarciog et al., 2010b].

#### 2.2 Koopman Operator & EDMD

The Approximation of the Koopman operator (1.1) for control has its origins in the work of Koopman [1931], where he extends the linear transformations of the Hilbert space given in harmonic equations and ergodic theory to Hamiltonian systems. Up until the 2000s, the theory had applications in ergodic, chaos, and measure theory analysis via the adjoint Perron-Frobenius operator [Lasota and Yorke, 1982]. In the mid-2000s Mezić [2005] established the relation that the Koopman operator has with linear decompositions such as Karhunen-Loeve, singular value, proper orthogonal, and dynamic mode decomposition (DMD). From this relation, the Koopman mode decomposition (KMD) was able to exploit the spectral characteristics of the operator to analyze nonlinear systems, where KMD is a means to reduce complex flow models and analyze the properties of the system near the attractors.

The possibility to get the spectrum of nonlinear systems for analysis gave rise to the extended dynamic mode decomposition (EDMD) [Williams et al., 2015]. The objective then shifted, from the reduced analysis near the attractor to a higher dimensional representation. This representation results in the exchange of dimensionality with complexity, where low-dimensional nonlinear systems have a high-dimensional representation with the advantage of the spectral analysis.

There are several variants of the EDMD algorithm, from norm-based expansions, radial-basis functions, kernel-based [Williams et al., 2016], orthogonal polynomials, and their variations [Kaiser et al., 2018, Li et al., 2017]. These representations give tools for analyzing nonlinear systems via the spectral decomposition and are the fundamentals for developing synthesis algorithms such as the EDMD for control [Proctor et al., 2016].

Within the theoretical framework under consideration, that is, the use of an approximation of the Koopman operator to analyze the system, there are some recent studies [Mezić, 2005, Lan and Mezić, 2013, Mauroy et al., 2013, Mauroy and Mezic, 2016] that lead to our theoretical framework and algorithm. Most notably, in [Mezić, 2005] the authors propose a visual method where the time average of an observable (i.e., an arbitrary function of the state space) approximates the eigenfunction of the Koopman operator whose eigenvalue is approximately unitary. Although this method provides visual information in two-dimensional systems or slices of three dimensional ones, the method does not give a criterion to determine the convergence of an arbitrary point, and is not feasible to analyze higher dimensional systems. Another notable approach is [Mauroy et al., 2013], where the authors calculate isostables of the system. An isostable of a stable equilibrium point is a set of points, or initial conditions that converge synchronously to the attractor, that is, they simultaneously intersect subsequent isostables along the trajectory to the stable point. The definition of an isostable comes from the magnitude of the Koopman operator slowest eigenfunction, whose level sets give the isostables. Similar to the Lyapunov and energy function based methods, this approach requires the calculation of the level sets of a particular function, which yields the same problems related to those methods regarding the classification of an arbitrary initial condition in higher dimensional systems.

#### 2.3 Interconnected Systems

The concept of interconnected dynamical systems has been a topic of research since the early '70s [Siljak, 1972, Grujic and Siljak, 1973, Sezer and Hüseyin, 1978], where the main focus was on large-scale and multilevel nonlinear dynamic systems interconnected by a function of their output variables. The analysis consists of the way of attaining the system stability conditions based on the local stability conditions and the analysis of the interconnection functions of the system. The basis for the analysis of the system is Lyapunov vector functions (LVF) [Bellman, 1962], in which a vector of distinct Lyapunov functions is defined based on the functions of the different subsystems. A simple algebraic condition on the interconnection matrix gives the interconnection's overall stability. The local dynamics of the local state variables, the interconnection functions, and a comparison vector function give the interconnection matrix. These methods rely on the ability to find suitable Lyapunov and comparison functions, which can be a difficult task. Furthermore, they rely on having the subsystem's dynamics identified.

Also, in the '70s, and with a different approach, Willems [Willems, 1972b] set the bases for the energetic analysis of dynamical systems and the energetic characteristics of the subsystems interconnection. Willem's stability analysis consists of the identification of the overall storage function of the system, all of these subject to the condition of the system having a "Neutral" interconnection constraint (i.e., the interconnecting system must be lossless). Another approach for linear systems from Shih-Ho Wang and Davison [1973] consists of the decentralization of large scale system's controllers, with the system considered as a whole, and each subsystem stabilized via local output feedback with segregated controllers. Controllable and observable modes introduce the notion of "fixed modes" that prove stability when they are in the left-hand part of the complex plane.

In the '80s, two approaches emerged, Ioannou [Ioannou, 1986] considers the subsystems as linear with a low relative degree of interconnection to other systems. For the synthesis of controllers, he designs local controllers assuming there is no noise in the output of the system and sets the interconnection input to zero. He handles the interconnection with a stability analysis that assumes that the noise in the system and the interconnection inputs are bounded signals. The second approach is from Linnemann [1984] that considers the stabilizability and reachability conditions of the linear subsystems and the stability conditions of the dynamics of the linear interconnections. He concludes that a system is stabilizable by decentralized output feedback if the subsystems are stabilizable by centralized output feedback, and the interaction system is stable.

The prevalent ideas in the '90s are by Siljak [1991], where large-scale systems are interconnections of physical entities or pure mathematical artifices. These descriptions are partitioned suitably in order to take advantage of the structural features that reduce the dimensionality of the analysis and synthesis methods via graph-theoretical methods. The approach considers the overall system as a collection of individual subsystems whose individual Lyapunov functions are combined to perform the analysis.

Another interesting approach comes from Maschke and van der Schaft [1996], where the notion of large-scale is not present; instead, they develop the paradigm of network interconnection, where the different interconnection structures appearing in the network models are analyzed presenting interconnections as energy exchanging components in some geometric space in correspondence with Dirac structures. This analysis is an extension of Willems [1972a,b] energetic analysis, where the Dirac structures represent the energetic exchange in the interconnection of Lagrangian and Hamiltonian dynamical formalisms.

More recently, an interesting approach to interconnected systems is in terms of decentralized control as population games, and evolutionary dynamics [Quijano et al., 2017]. The main idea of this approach is to use a multi-agent system method where the interactions between agents comes from a game-theoretical approach. The local controller problem is a game where the available control inputs are associated with strategies and the control objective as a pay-off function.

For the analysis of interconnected systems via the Koopman operator, there are two recent developments. Mauroy and Hendrickx [2017a] propose a method in which the application of the EDMD algorithm on some sparse measurements of some nodes of the network, and possibly from measurements of just one node of the network gives a spectral decomposition. From the graph-theoretical properties of the EDMD spectral decomposition, they deduce the topology of the network, i.e., the arrangement of the different edges and nodes. Meaning that the algorithm is a graph identification tool when the topology of the network is unknown. More recently Mauroy and Hendrickx [2017b] extended their previous theory for forced linear systems in the edges of the network.
## 2.4 Summary

The current literature covers several aspects that give the starting point for the development of the thesis. A key aspect in all of the previous developments is the availability of a differential equation model to perform all the calculations. Given that it is not always possible to satisfy this assumption and when possible, several identification techniques rely on numerical and optimization methods to parametrize the models, the use of algorithmic techniques such as the EDMD is a promising approach that relies on the same data, yet bypasses the nonlinear models to give linear predictors of the state in an extended space of functions. Therefore, improving the algorithm such that it gives more accurate approximations with less amount of data is a valuable line of research.

On the subject of linear approximations in an extended space of functions, there is the Koopman operator, an infinite transformation of the state space into a function space where the evolution of a set of so-called eigenfunctions, relate to the evolution of the original state space. The availability of Koopman operator and the EDMD algorithm, and the fact that under some assumptions on the amount of data and the size of the extended space of functions, the EDMD converges to the point spectrum of the Koopman operator, continuing with the line of research of the possibilities that the combination of these two paradigms is also a promising approach. Specifically concerning the Koopman operator, a deeper characterization of the properties of the available eigenfunctions and their relation to the analysis of nonlinear systems is an interesting line of research.

An interesting property of the eigenfunctions of the discrete-time Koopman operator is that their eigenvalues determine their evolution (1.1). From this property a nontrivial eigenfunction that has a unitary eigenvalue must be invariant, its value must not change along the trajectories of the states of the system. Matching this phenomena with the fact that the analysis of attraction regions, or domains of attractions is based on the ability to find invariant sets of these kind, shows that it is worth pursuing the line of research that combines these two concepts for the analysis of non-linear systems.

For the dual aspect of analysis in the subject of *control systems* the-

ory, i.e., synthesis algorithms, the EDMD and the Koopman operator are also an alluring approach. The world of linear systems control has a plethora of techniques for driving systems into a desired state. These techniques are available for driving non-linear systems in a vicinity of the equilibrium points where the linearization of the system is valid. Then again, it is clear the advantageous use of linearization techniques such as the EDMD that provide linear approximations that are valid in the portion of the state space from which the samples come from. Coupled with the availability of the point spectrum of the Koopman operator, it is reasonable that there is an opportunity to bring and adapt linear systems techniques into the expanded linearization from the EDMD approximation of the Koopman operator.

Exploiting the linear predictor from the EDMD algorithm not only comes from the possibility to adapt linear systems techniques, these approximations are versatile in the use of different forms of observables for the state variables, and the inputs. Notably, in the context of interconnected systems, some inputs are not driven by a controller although their action affects the dynamics of the system, for these inputs, there it not necessary to have linear observables, instead, polynomial observables have the possibility to provide more accurate approximations according to this inputs, while keeping the linear effect on the driven inputs. These paradigm is specially useful for synthesizing model predictive control algorithms that require linear driven inputs. The combination of all these methods has the potential of providing a tool for decentralized control of interconnected systems.

# Chapter 3

# Dynamical Systems, ROA & EDMD

In order to present the methods based on the discrete time approximation of the Koopman operator via the EDMD algorithm, this section will cover preliminary concepts of dynamical systems theory that are necessary to understand the subsequent development. As any traditional control systems analysis, this section starts with the definition of fixed points and presents the concept of stability based on the classical Lyapunov analysis. Along with these definitions, this chapter presents the Hartman-Grobman theorem that defines what is a hyperbolic fixed point. Finally, this chapter presents the multi-stability problem along with the theoretical aspects that define the region of attraction.

Given that all of the theoretical preamble is developed for nonlinear system with explicit knowledge of the differential model equations, the next development of this theoretical chapter is the basic theory of the Koopman operator and the EDMD algorithm. Besides the theoretical foundation for the subsequent developments, this chapter also includes the algorithmic improvements to the calculation of the EDMD. Those improvements are the reduction methods by p-q-quasi norms and polynomial accuracy, the expansion of the state of the system by trigonometric embeddings and the production of bi-linear predictors with mixed affine and non-affine inputs for handling interconnected systems. Furthermore, this chapter presents the synthesis of controllers for these interconnected systems making use of these linear predictors.

To conclude the chapter, and in line with the general topic of this work, the last section shows how to get polynomial models from reaction networks that serve to illustrate the different developments of the thesis due to their geometrical properties.

## 3.1 Discrete Nonlinear Systems

Consider the autonomous nonlinear discrete-time system  $(\mathcal{M}; T(x); k)$ , with state variables  $x \in \mathcal{M}$  where  $\mathcal{M} \subseteq \mathbb{R}^n$  is the nonempty compact state space,  $k \in \mathbb{Z}_0^+$  is the discrete time, and  $T: \mathcal{M} \to \mathcal{M}$  is the differentiable vector-valued evolution map, i.e.,

$$x(k+1) = T(x(k)), \quad x_0 = x(0).$$
 (3.1)

The solution to (3.1) is the successive application of T from an initial condition  $x_0 \in \mathcal{M}$  at k = 0, i.e.,  $x_k = T^k(x_0) \in \mathcal{M}$ , which is an infinite sequence called a trajectory of the system.

As an example, consider an arbitrary non-linear system in discretetime with two states [Chiang and Alberto, 2015], depicted in Figure 3.1 with evolution map

$$x_1(k+1) = x_1(k)^3 + \frac{3}{4}x_1(k)$$
(3.2)

$$x_2(k+1) = x_2(k)^3 + \frac{3}{4}x_2(k)$$
(3.3)

Suppose  $x^* \in \mathcal{M}$  is a fixed point of (3.1); i.e.,

**Definition 3.1.1** (Fixed Point). Let  $x^*$  be a fixed point of the dynamical system  $(\mathcal{M}; T(x); k)$ , then the flow map  $T^k(x^*)$  maps into itself, i.e.,

$$T^k(x^*) = x^*. (3.4)$$

To obtain the fixed points in the Example (3.2-3.3), it is necessary to solve

$$x = x^3 + \frac{3}{4}x, (3.5)$$



Figure 3.1: Discrete-time system with an AS point at the origin, unstable points at the corners and unstable-saddles at the edges.

whose solution is x = 0 or  $x = \pm 0.5$ , i.e., there are nine fixed points, one at the origin, two at  $(0, \pm 0.5)$ , two at  $(\pm 0.5, 0)$  and four at  $(\pm 0.5, \pm 0.5)$ .

**Definition 3.1.2** (Stability). The fixed point  $x^*$  of x(k+1) = T(x(k)) is

a. Stable if, for each  $\epsilon(x^*) > 0$ , there is a  $\delta = \delta(\epsilon) > 0$  such that

$$\|x_0 - x^*\| < \delta \Rightarrow \|T^k(x) - x^*\| < \epsilon.$$
(3.6)

- b. Unstable if it is not stable.
- c. Asymptotically stable if,  $\delta$  can be chosen such that

$$||x_0 - x^*|| < \delta \Rightarrow \lim_{k \to \infty} T^k(x) = x^*$$
(3.7)

d. Globally asymptotically stable if,  $x^*$  is unique and  $\delta$  can be chosen arbitrarily large.

Returning to the example at hand, the origin satisfies condition a and condition c of Definition 3.1.2 making it an asymptotically stable (AS) point while other fixed points are unstable.

In order to determine the stability of an equilibrium point it is very useful to understand the qualitative behavior near the equilibrium point. The linearization principle around an equilibrium point states that the analysis of eigenvalues of the linear state transition matrix gives the stability of the point. The analysis of eigenvalues and eigenvectors under consideration comes from systems that satisfy the Hartman-Grobman theorem [Khalil, 2002, Coayla-Teran et al., 2007, Giesl and Hafstein, 2015].

**Theorem 3.1.1** (Hartman-Grobman). Consider system (3.1) with a fixed point  $x^*$  and consider the linear system x(k + 1) = Ax(k) where A is the Jacobian  $DT(x^*)|_{x^*}$ . If the modulus of all the eigenvalues of A are different from one (i.e., Hyperbolic), then there exists a homomorphism H, defined on a neighborhood of  $x^*$  such that, for all initial points  $x_0$  along the flows of the system  $T^k(x)$ 

$$H \circ T^k(x_0) = A^k H(x_0)$$
 (3.8)

is satisfied. This means that the homeomorphism preserves the sense of the flows of the nonlinear system and is chosen to preserve parametrization by time.

A hyperbolic fixed point of a non-linear discrete time system is a point where all the eigenvalues of the linearization at the point have a modulus different from one. According to the Hartman-Grobman theorem in discrete-time a hyperbolic fixed point accepts a local linearization that preserves the behavior of the non-linear system. Evaluating this condition on the current example gives,

Ì

$$DT(x) = \begin{bmatrix} 3x_1^2 + \frac{3}{4} & 0\\ 0 & 3x_2^2 + \frac{3}{4} \end{bmatrix}$$
$$DT(x)|_{(0,0)} = \begin{bmatrix} \frac{3}{4} & 0\\ 0 & \frac{3}{4} \end{bmatrix}$$
(3.9)

$$DT(x)|_{(0,\pm0.5)} = \begin{bmatrix} \frac{3}{4} & 0\\ 0 & \frac{3}{2} \end{bmatrix}$$
 (3.10)

$$DT(x)|_{(\pm 0.5,0)} = \begin{bmatrix} \frac{3}{2} & 0\\ 0 & \frac{3}{4} \end{bmatrix}$$
 (3.11)

$$DT(x)|_{(\pm 0.5, \pm 0.5)} = \begin{bmatrix} \frac{3}{2} & 0\\ 0 & \frac{3}{2} \end{bmatrix},$$
 (3.12)

given that all the eigenvalues of the different fixed points are real and different from one, they are all hyperbolic. Furthermore, the modulus of the eigenvalues of a particular hyperbolic fixed point give its corresponding local stability.

**Theorem 3.1.2** (Local stability). The local stability of an equilibrium point  $x^*$  of a linear discrete time system with respect to its eigenvalues  $\mu_i$  are:

- a. The equilibrium point  $x^*$  is asymptotically stable if  $|\mu_i| < 1$  for all eigenvalues of its state transition matrix.
- b. The equilibrium point  $x^*$  is unstable if  $|\mu_i| > 1$  for one or more eigenvalues of its state transition matrix.
- c. The equilibrium point  $x^*$  is a saddle point if some but not all  $|\mu_i| < 1$ , which means that the equilibrium has some modal components that converge to it.

The linearization principle, along with the local stability of fixed points states that a fixed point  $x_s^*$  is AS if the modulus of all the eigenvalues of the Jacobian matrix evaluated at the fixed point are less than one, and unstable otherwise with an index k equal to the number of eigenvalues with modulus greater than one. A hyperbolic fixed point is of type-k if k eigenvalues of the Jacobian matrix have modulus greater than one. If only one eigenvalue has modulus greater than one  $x^*$  is a *type-one* fixed point. When the index k of unstable fixed points is equal or greater than one, and less than n, the fixed point is called a saddle; denoted by  $\hat{x}^*$ . The *type-one* saddle points play an important role in the approximation of the ROA.

Returning to the discrete-time Example (3.2-3.3), and analyzing the nine fixed points of the system based on Theorem 3.1.2 gives as a result that the origin is asymptotically stable, the edges of the "square" are saddle points of *type-one* and the corners are unstable. Table 3.1 summarizes these results.

Table 3.1: Fixed points, stability and type-k for the arbitrary non-linear system (3.2)

Equilibrium point	Stability	type- $k$
(0, 0)	AS	type-0
$(0, \pm 0.5)$	Saddle	type- $one$
$(\pm 0.5, 0)$	Saddle	type- $one$
$(\pm 0.5, \pm 0.5)$	Unstable	type-2

When the AS fixed point is not unique, the concept of attraction regions of the AS fixed points defines the subset of the state space where an arbitrary initial condition converges to any of them.

# **3.2** Regions of Attraction

The central question is: how far can an initial condition be from the AS equilibrium point in order that the discrete-time nonlinear evolution map (3.1) converges to it? Or, how big can a disturbance from an AS be such that state converges to another AS point or diverges? These questions are especially relevant for the case when the equilibrium point is not unique, and there are several AS points. For that case the ROA is the open set from which every initial condition converges to a particular

AS point  $x_s^*$ , defined as

$$R_A(x_s^*) \equiv \left\{ x \in \mathbb{R}^n : \lim_{k \to \infty} T^k(x) = x_s^* \right\}.$$
 (3.13)

The least conservative approach is to characterize the stability boundary of the attraction region, defined as the closure of the set (3.13) without its interior

$$\partial R_A(x_s^*) = \overline{R_A(x_s^*)} \setminus \operatorname{int}(R_A(x_s^*)).$$
(3.14)

For the particular case of dynamical systems, this boundary comes from the analysis of saddle fixed points, which have modes that converge to them, and modes that diverge from them. The method consists of finding the stable manifolds of the saddle points, where the union of the stable manifolds of saddle points in the stability boundary of an AS gives the entire boundary. The objective is to approximate these stable manifolds.

From the linearization principle it is known that eigenvalues  $\lambda$  with modulus less than one are stable eigenvalues with stable eigenvectors and generalized eigenspace  $E_{\lambda}$  that spans the stable eigenspace  $E^s$  of the fixed point; i.e.,  $E^s = \oplus E_{|\lambda|<1}$ . Conversely, eigenvalues greater than one are unstable eigenvalues with unstable eigenvectors and generalized eigenspace that spans the unstable eigenspace  $E^u$  of the fixed point; i.e.,  $E^u = \oplus E_{|\lambda|>1}$ . The type of the hyperbolic fixed point defines the dimension of the corresponding eigenspaces,  $E^u \in \mathbb{R}^k$  and  $E^s \in \mathbb{R}^{n-k}$ . The space state  $\mathbb{R}^n$  is the direct sum of the two invariant stable and unstable eigenspaces  $\mathbb{R}^n = E^s \oplus E^u$ .

As the Hartman-Grobman theorem establishes a one-to-one correspondence between the nonlinear system and its linearization, locally, the stable and unstable eigenspaces are tangent to the stable and unstable manifolds of the hyperbolic fixed point. The definitions of these manifolds are

$$W^{s}(x^{*}) = \{ x \in \mathbb{R}^{n} \colon \lim_{k \to \infty} T^{k}(x) = x^{*} \},$$
(3.15)

$$W^{u}(x^{*}) = \{x \in \mathbb{R}^{n} \colon \lim_{k \to -\infty} T^{k}(x) = x^{*}\},$$
 (3.16)

for the stable and unstable manifold respectively, assuming there is an inverse for  $T^{k}(x)$ .

With these definitions, the theorem characterizing the ROA for asymptotically stable fixed points, with stability boundary  $\partial R_A(x_s^*)$ , is taken from [Chiang and Alberto, 2015, Th. 9-(10,11)], where the system satisfies the following assumptions:

**Assumption 1.** All the fixed points on the stability boundary are saddle points.

**Assumption 2.** The stable and unstable manifolds of the fixed points on the stability boundary satisfy the transversality condition<sup>1</sup>.

**Assumption 3.** Every trajectory on the stability boundary approaches one of the saddle points as  $k \to \infty$ .

**Theorem 3.2.1.** Consider an autonomous nonlinear discrete-time dynamical system (3.1), satisfying Assumptions 1-3. Let  $\{\hat{x}_i^*\}_{i=1}^P$  be the *P* saddle points on the stability boundary of an asymptotically stable fixed point. Then,

a. 
$$\hat{x}_i^* \in \partial R_A(x_s)$$
 if and only if  $W^u(\hat{x}_i^*) \cap R_A(x_s^*) \neq \emptyset$   
 $h \quad \partial R_A(x_s^*) = \cup W^s(\hat{x}^*)$ 

$$0. \ OR_A(x_s) = \bigcirc W \ (x_i).$$

Hence, the stability boundary is the union of the stable manifolds of the hyperbolic fixed points on the stability boundary.

To illustrate the role of the stable manifolds of saddle points in the boundary, recall the non-linear example from Section 3.1 equation (3.2). This system has four *type-one* saddle points at  $(0, \pm 0.5)$  and  $(\pm 0.5, 0)$ , where their corresponding stable and unstable eigenspaces are

$$E^s(0,0.5) = \{ x \in \mathbb{R}^2 \colon x_2 = 0.5 \}$$
(3.17)

$$E^{u}(0,0.5) = \{x \in \mathbb{R}^{2} \colon x_{1} = 0\}$$
(3.18)

$$E^{s}(0, -0.5) = \{ x \in \mathbb{R}^{2} \colon x_{2} = -0.5 \}$$
(3.19)

$$E^{u}(0, -0.5) = \{ x \in \mathbb{R}^{2} \colon x_{1} = 0 \}$$
(3.20)

$$E^s(0.5,0) = \{ x \in \mathbb{R}^2 \colon x_1 = 0.5 \}$$
(3.21)

 $E^{u}(0.5,0) = \{x \in \mathbb{R}^{2} \colon x_{2} = 0\}$ (3.22)

$$E^{s}(-0.5,0) = \{x \in \mathbb{R}^{2} \colon x_{1} = -0.5\}$$
(3.23)

 $E^{u}(-0.5,0) = \{ x \in \mathbb{R}^{2} \colon x_{2} = 0 \}.$ (3.24)

<sup>&</sup>lt;sup>1</sup>For manifolds A and B in  $\mathcal{M}$ , the transversality condition is satisfied if the tangent spaces of the intersection between A and B span the tangent space of  $\mathcal{M}$ .

Figure 3.2 depicts the stable eigenspaces of these *type-one* saddle points with solid black lines, and highlights the ROA of the AS point in the origin as the union of these sets.



Figure 3.2: Region of attraction as the union of the stable manifolds (solid black lines) of the saddle fixed points.

**Remark 1.** Assumption 1 is a generic property of differentiable dynamical systems, while Assumptions 2 and 3 must be verified. Figure 3.3 shows the importance of Assumption 2, where a part of the stable manifold of  $\hat{x}_A^*$ does not intercept with the stability boundary. For this particular case, the transversality condition is not satisfied because part of the tangent space of the unstable manifold of  $\hat{x}_A^*$  is equal to a part of the tangent space of the stable manifold of  $\hat{x}_B^*$ .

# 3.3 The Koopman Operator

This section contains the basic theory of the Koopman operator for systems that accept a discrete-time point spectrum. The main quality of this operator, related to the underlying dynamical system is the linear evolution of a set of observables (later defined as eigenfunctions) that capture the nonlinear evolution of the system.





Figure 3.3: Transversality condition violation by the stable and unstable manifold of saddle points.

Consider the autonomous nonlinear discrete-time dynamical system (3.1), and a set of observation functions

$$f(x)\colon \mathcal{M}\to\mathbb{C}$$

that belong to some function space, i.e.,  $f(x) \in \mathcal{F}$ . For these functions there is a discrete operator  $U^k$ , where the action of this operator on observables defines their evolution. This is the Koopman operator, and its relationship with the observables (i.e., arbitrary functions of the states) and the states of the system is

$$\left[U^{k}f\right](x) = f\left(T^{k}(x)\right).$$
(3.25)

In other words, the time-evolution of observations is the observations of the time-evolution of states. The trade-off with this approach concerns linearity and dimensionality, a finite-dimensional nonlinear system being described by an infinite-dimensional linear one. The linearity of the operator allows it to have a spectral decomposition of tuples  $\{(\mu_i, \phi_i(x), v_i)\}_{i=1}^{\infty}$  of eigenvalues, eigenfunctions and modes that contain information on the underlying dynamical system. The eigenvalues and eigenfunctions satisfy the condition that the corresponding eigenvalue determines the dynamics associated with a specific eigenfunction as

$$[U^k \phi_i](x) = \mu_i^k \phi_i(x), \qquad (3.26)$$

and the Koopman modes serve to recover the observation functions f(x), i.e., they map the linear evolution of eigenfunctions (3.26) into the original observables by weighting the eigenfunctions as

$$f(x) = \sum_{i=1}^{\infty} v_i \phi_i(x).$$
 (3.27)

The importance and advantages of the Koopman operator and the diagonalization provided by the spectral decomposition are highlighted by (3.26) and (3.27), which show that the eigenvalue associated to an eigenfunction defines its time evolution and that the modes recover the value of the defined set of observables f(x). Therefore, the evolution of observables with respect to the spectral decomposition of the Koopman operator is

$$f\left(T^{k}(x)\right) = \left[U^{k}f\right](x) = \sum_{i=1}^{\infty} v_{i}\mu_{i}^{k}\phi_{i}(x).$$
(3.28)

#### 3.3.1 Koopman Operator Example

The exact solution to the Koopman operator of a dynamical system is available for a family of polynomial difference equations [Brunton et al., 2016a]. This closed form solution is only possible for systems with a unique fixed point. Therefore, there is no exact solution for the problem of approximating the ROA. Consider the following autonomous nonlinear discrete-time dynamical system

$$x_1(k+1) = \gamma x_1(k)$$
  

$$x_2(k+1) = \delta x_2(k) + (\gamma^2 - \delta) x_1^2(k).$$
(3.29)

The solution of the system under the Koopman operator requires the definition of a set of observable functions whose evolution is defined by the operator as in (3.25). For this particular case (see [Brunton et al., 2016a] for the deduction of the set of observables), the set of observables and the Koopman system are

$$f(x) = \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \end{bmatrix} \Rightarrow f(x(k+1)) = \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \delta & (\gamma^2 - \delta) \\ 0 & 0 & \gamma^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1^2 \end{bmatrix}.$$
 (3.30)

The spectral decomposition of the operator gives the set of eigenvalues  $\mu_1 = \gamma$ ,  $\mu_2 = \gamma^2$  and  $\mu_3 = \delta$ , and the sets of left and right eigenvectors, defined as  $W^*$  and  $\Xi$  respectively and given by

$$W^{\star} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \qquad \qquad \Xi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \qquad (3.31)$$

From the left eigenvectors and the set of observables f(x), the definition of the set of eigenfunctions  $\Phi(x)$  of the Koopman operator is the observables weighted by the left eigenvector, i.e.,  $\Phi(x) = (f(x)^{\top} W^{\star})^{\top}$ , which gives

$$\Phi(x) = \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{bmatrix} = \left( \begin{bmatrix} x_1 & x_2 & x_1^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \right)^{\top} = \begin{bmatrix} x_1 \\ x_1^2 \\ x_2 - x_1^2 \end{bmatrix}.$$
(3.32)

With the complete spectral decomposition of the Koopman operator for system 3.29, with right eigenvectors  $\Xi$ , eigenvalues  $M = \text{diag}(\gamma, \gamma^2, \delta)$ , and using (3.28), the evolution of observables in matrix form is

$$f(x(k+1)) = \Xi M \Phi(x(k))$$
  
=  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \gamma^2 & 0 \\ 0 & 0 & \delta \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_1^2(k) \\ x_2(k) - x_1^2(k) \end{bmatrix},$  (3.33)

and the evolution of states  $[x_1, x_2]^{\top} = B^{\top} f(x)$  is

$$\begin{aligned} x(k+1) &= B^{\top} \Xi M \Phi(x(k)) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}^{\top} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \gamma & 0 & 0 \\ 0 & \gamma^2 & 0 \\ 0 & 0 & \delta \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_1^2(k) \\ x_2(k) - x_1^2(k) \end{bmatrix} \\ &= \begin{bmatrix} \gamma x_1 \\ \delta x_2 + (\gamma^2 - \delta) x_1^2 \end{bmatrix}, \end{aligned}$$
(3.34)

where  $B \in \mathbb{R}^{3\times 2}$  is a selection matrix with a unitary value per column in the index of the observables that capture every one of the states. In this particular case, the set of observables contains the functions that capture every one of the states, i.e.,  $f_1(x) = x_1$  and  $f_2(x) = x_2$ . Note that (3.34) circles back to the original dynamical system (3.29).

It is often difficult to obtain a closed analytical form for the Koopman operator, and due to the fact that for the multiple stability problem at hand, there is no known closed analytical form of the discrete-time Koopman operator it is necessary to use approximation techniques to circumvent this difficulty. Algorithms such as the EDMD approximate a truncated version of the Koopman operator with data from the dynamical system [Korda and Mezić, 2018a]. Furthermore, complementing the algorithm with order and dimension reduction methods, coupled with the retention of order one polynomial elements in the basis, provides a way to have accurate approximations of polynomial systems that provide all the necessary components to perform the analysis.

# 3.4 Extended Dynamic Mode Decomposition

The objective of the EDMD algorithm [Williams et al., 2015] is to get a truncated approximation of the discrete-time Koopman operator based on sampled data of the underlying system [Klus et al., 2016, Korda and Mezić, 2018a]. Consider the autonomous nonlinear discrete dynamical system (3.1), the EDMD algorithm to approximate the Koopman operator in discrete-time requires N pairs of data snapshots, either from a real system or a numerical simulation at a specific sampling  $\Delta t$ . The snapshot pairs,  $\{(x_i, y_i)\}_{i=1}^N$ , where  $y_i = T(x_i)$ , are organized in data sets

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_N \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 & y_2 & \dots & y_N \end{bmatrix}.$$
 (3.35)

The "extended" part of the EDMD algorithm consists in the approximation of the Koopman operator of a "lifted" space of the state variables, rather than approximating the state space as in the DMD algorithm [Schmid, 2010]. The "lifting" procedure consists in evaluating the state space of the system on a vector-valued function of observables  $\Psi = [\psi_1 \cdots \psi_d]^T : \mathcal{M} \to \mathbb{C}^{1 \times d}$  [Williams et al., 2015]. The approximation of the Koopman operator will be of dimension d. This operator has to satisfy the condition

$$\Psi(y) = U_d \Psi(x) + r(x), \qquad (3.36)$$

where  $r(x) \in \mathcal{F}$  is the residual term to minimize in order to find  $U_d$ . This problem is closely related to the least mean squares problem, polynomial chaos expansion [Sudret, 2008], statistical estimation [Walter and Pronzato, 1997], and machine learning [Neumaier, 2003] problems. Considering that the best approximation is the solution that minimizes the residual term, a solution of the objective function

$$\|r(x)\|^{2} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \|\Psi(y_{i}) - U_{d}\Psi(x_{i})\|_{2}^{2}, \qquad (3.37)$$

gives an accurate approximation of  $U_d$ . The solution corresponds with the ordinary least-squares [Williams et al., 2015], which has a closed analytical form given by,

$$U_d \triangleq AG^+, \tag{3.38}$$

with matrices  $G, A \in \mathbb{C}^{d \times d}$  defined by

$$G = \frac{1}{N} \sum_{i=1}^{N} \Psi(x_i) \Psi(x_i)^{\top}, \qquad (3.39)$$

$$A = \frac{1}{N} \sum_{i=1}^{N} \Psi(x_i) \Psi(y_i)^{\top}.$$
 (3.40)

**Remark 2.** A critical factor for the effectiveness of the solutions to (3.37) is the choice of observables  $\Psi$ . Although there are several proposed procedures, the choice of the so-called dictionary elements is still an open problem.

For a formal deduction of solution (3.38), consider the problem (3.37) in matrix form.

$$\min_{U_d} J = \min_{U_d} \frac{1}{2} \|\Psi(Y) - U_d \Psi(X)\|_2^2.$$
(3.41)

The solution to the minimization problem is

$$J = \frac{1}{2} (\Psi(Y) - U_d \Psi(X)) (\Psi(Y) - U_d \Psi(X))^{\top}$$
(3.42)

$$= \frac{1}{2} (\Psi(Y) - U_d \Psi(X)) (\Psi(Y)^{\top} - \Psi(X)^{\top} U_d^{\top})$$
(3.43)

$$= \frac{1}{2} (U_d \Psi(X) \Psi(X)^\top U_d^\top - \Psi(Y) \Psi(X)^\top U_d^\top + \Psi(Y) \Psi(Y)^\top)$$
(3.44)

$$\frac{\mathrm{d}}{\mathrm{d}U_d}J = \frac{1}{2}(2U_d\Psi(X)\Psi(X)^\top - 2\Psi(Y)\Psi(X)^\top)$$
(3.45)

$$0 = U_d \Psi(X) \Psi(X)^\top - \Psi(Y) \Psi(X)^\top$$
(3.46)

$$U_d = (\Psi(Y)\Psi(X)^{\top})(\Psi(X)\Psi(X)^{\top})^+$$
(3.47)

$$= AG^+. (3.48)$$

Solution (3.38) is suitable when matrix G is not ill-conditioned. Otherwise, solutions based on the Moore-Penrose pseudoinverse or regularized least-squares give a more accurate result. The former being the most accurate, and the latter being computationally efficient and able to provide sparse approximations of  $U_d$ . For the regularized leastsquares case, the objective is to find a set of d linear fitting functions  $\{h_j(x) = U_{d_j}\Psi(x)^T\}_{j=1}^d$  of the 'lifted" space, where j indexes the rows of the operator that minimize the error function:

$$E(U_{d_j}) = \frac{1}{N} \sum_{i=1}^{N} L(\psi_j(y_i), h_j(x_i)) + \alpha R(U_{d_j}), \qquad (3.49)$$

where L is the loss function that measures the model fit, R is the regularization term, and  $\alpha \in \mathbb{R}_+$  is a nonnegative parameter. The choice of L gives different types of classifiers [Rosasco et al., 2004], while different regularization terms R provide various solutions to the over fitting problem that arises from ill-conditioned problems.

The common choices for the loss function are the Ridge regression, absolute value loss, and support vector machines (epsilon-intensive) functions [Rosasco et al., 2004]. For the regularization term, the common choices are  $l_1$ ,  $l_2$ , and a linear combination of the two (elastic net) norms. With the solution to (3.36), and the spectral decomposition  $M, \Xi, W^*$ , where  $M = \text{diag}(\mu_1, \ldots, \mu_d)$  is the diagonal matrix of eigenvalues,  $\Xi = [\xi_1, \ldots, \xi_d]$  is the aforementioned matrix of right eigenvectors, and  $W^* = \Xi^{-1}$  is the matrix of left eigenvectors. The definition of the eigenfunctions of the Koopman operator according to EDMD algorithm (according to Williams et al. [2015])  $\Phi = [\phi_1, \ldots, \phi_d]^{\top}$  are

$$\Phi^{\top}(x) = \Psi(x)^{\top} W^{\star}, \qquad (3.50)$$

from (3.26) that describes the evolution of eigenfunctions according to their eigenvalue, the evolution of eigenfunctions from the EDMD algorithm is

$$\Phi(x(k)) = M^k \Phi(x_0), \qquad (3.51)$$

with the evolution of eigenfunctions, the right eigenvectors or modes serve to recover the original set of observables

$$\Psi(x) = \Xi \Phi(x), \tag{3.52}$$

and the time evolution of observables according to the spectrum of the Koopman operator [Williams et al., 2015] is

$$\Psi(T^k(x)) = \Xi M^k \Phi(x). \tag{3.53}$$

To recover the state, we define a matrix  $B \in \mathbb{R}^{(d \times n)}$  with a unitary value per column in the index of an injective observable-function and zeros otherwise. This matrix selects the functions whose inverse gives the value of the state. If we define this set of functions selected by B as  $\Psi_B(x)$ , then, with (3.53), the approximation of the state evolution map  $T^k(x) = x_k$  from the initial condition  $x_0$  in terms of the approximation of the Koopman decomposition  $\{(\mu_i, \phi_i(x), v_i)\}_{i=1}^d$  is

$$\bar{x}(k) = \bar{T}^k(x_0) = \Psi_B^{-1} \left( B^\top \Xi M^k \Phi(x_0) \right).$$
(3.54)

Note that the Koopman representation of the system also allows for a clean representation of the backward evolution of the states given by

$$\bar{x}(-k) = \Psi_B^{-1} \left( B^\top \Xi M^{-k} \Phi(x_0) \right).$$
(3.55)

Given that the approximation of the discrete-time Koopman operator comes from a set of orbits of the system, the comparison between these theoretical orbits against the evolution map of the Koopman operator gives a measure for the algorithms accuracy. The empirical error for a number of  $N_s$  trajectories, where each has a number of  $N_k$  data points is

$$e = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{|T^k(x_i) - \bar{T}^k(x_i)|}{|T^k(x_i)|}.$$
(3.56)

A low error over a set of test trajectories of the Koopman state evolution map, gives a Koopman operator approximation whose eigenfunctions are accurate enough for the subsequent calculations related to them.

#### 3.4.1 Observables

A critical part of the approximation via the EDMD algorithm is the choice of observables. These functions of the state  $\psi_l(x)$  usually consists of an orthogonal basis of polynomials [Koekoek et al., 2010], radial basis functions, or an arbitrarily constructed set with polynomial elements, and trigonometric functions, among others [Brunton et al., 2016b].

The choice of orthogonal polynomials for the approximation of the Koopman operator has the advantage that the rank and dimension of the vector-valued function of observables  $\Psi$  can be reduced [Konakli and Sudret, 2016a,b], and the state can be recovered by the inverse of injective observable-functions. A sequence of orthogonal polynomials  $\{\pi_{\alpha}(x)\}_{\alpha=0}^{p}$  where  $\pi_{\alpha}(x)$  is a univariate (i.e.,  $x \in \mathbb{R}$ ) polynomial of degree  $\alpha \in \mathbb{N}_{+}$  up to order p. This sequence is defined over a range [a, b] where some inner product between distinct elements is zero, i.e.,  $\langle \pi_i(x), \pi_j(x) \rangle = 0$  for  $i \neq j$ , and satisfies a particular ordinary differential equation. For example, the set of Laguerre polynomials, which is defined over the range  $[0, \infty]$  satisfies the ordinary differential equation

$$x\frac{\mathrm{d}^2 y(x)}{\mathrm{d}x^2} + (1-x_1)\frac{\mathrm{d}y(x)}{\mathrm{d}} + \alpha y(x) = 0, \qquad \alpha = 1, 2, \dots$$
(3.57)

Consider an autonomous nonlinear discrete-time dynamical system such as (3.29), where the number of state variables n = 2, and the selected polynomial basis consists of Laguerre polynomials. Table 3.2 shows the set of univariate polynomials on the first component of the state space vector (i.e., j = 1) for a choice of maximum polynomial degree p = 3, and a set of indices  $\alpha_1 = \{0, 1, 2, 3\}$ .

Table 3.2: Basis for the first state variable  $x_1$ . Solution to the Laguerre ODE:  $x_1\ddot{y} + (1 - x_1)\dot{y} + \alpha_1 y = 0$  where  $\dot{y} = \frac{dy}{dx_1}$ .

$J(\alpha_1)$	ODE	$\pi_{\alpha_1}(x_1) = y(x_1)$
$J(0_1)$	$x_1 \ddot{y} + (1 - x_1) \dot{y} = 0$	1
$J(1_{1})$	$x_1 \ddot{y} + (1 - x_1) \dot{y} + y = 0$	$-x_1 + 1$
$J(2_{1})$	$x_1 \ddot{y} + (1 - x_1)\dot{y} + 2y = 0$	$\frac{1}{2}x_1^2 - 2x_1 + 1$
$J(3_1)$	$x_1\ddot{y} + (1 - x_1)\dot{y} + 3y = 0$	$-\frac{1}{6}x_1^3 + \frac{3}{2}x_1^2 - 3x_1 + 1$

With these univariate polynomials, every element of the vector-valued function of observables is the tensor product of n univariate polynomials, that is,

$$\psi_l(x) = \prod_{j=1}^n \pi_{\alpha_j}(x_j), \qquad l = 1, \dots, d.$$
(3.58)

If we consider a full degree p and full dimension vector-valued function of observables, the order and dimension grows exponentially with the number of state variables n. The maximum degree of the polynomials in  $\Psi(x)$  is  $p^n$  and the dimension of the whole set is

$$d = \dim \Psi(x) = \binom{n+p+1}{p+1} = \frac{(n+p+1)!}{(p+1)!n!}.$$
 (3.59)

The exponential growth of the order produces an over-fitting problem in the discrete approximation of the Koopman operator via the EDMD algorithm, and the growth in the dimension has the course of dimensionality problem, especially in the definition of the discrete Koopman matrix from (3.38) that needs the inverse of matrix G. To avoid these problems it is necessary to restrict the maximum order of the polynomial basis and its dimension.

#### 3.4.2 Reduction by p-q-quasi Norms

The proposed reduction of the order and dimension of the polynomial basis is via a p-q-quasi norm<sup>2</sup> reduction first introduced in the solution of polynomial chaos problems [Konakli and Sudret, 2016a]. Define the q-quasi norm as,

$$\|\alpha\|_q = \left(\sum_{i=1}^n \alpha_i^q\right)^{\frac{1}{q}},\tag{3.60}$$

for a polynomial element  $\psi_l(x)$  consider a maximum degree  $p \in \mathbb{N}$ , a quasi norm  $q \in \mathbb{R}_+$  and define the set of indexes  $\alpha_l$  as

$$\alpha_l = \{ \alpha \in \mathbb{N}^n \colon \|\alpha\|_q \le p \}. \tag{3.61}$$

Given that each element of the polynomial basis is the tensor product of the univariate polynomials (3.58) in the retained indexes given by (3.61), this approach reduces the maximum order, and the overall dimension of the vector-valued function of observables.

Figure 3.4 shows the effect of the truncation scheme for the previous example on Laguerre polynomials (3.57) for the case where the maximum polynomial order is p = 3, and the number of state variables is n = 2. Certainly, reducing the value of the quasi norm q, reduces the maximum order and dimension of the vector valued function of observables.

Every element of the polynomial basis comes from the tensor product in (3.58), and a full basis, i.e., with  $q = \infty$ , has 16 elements resulting from all the available combinations of indices in two state variables with a univariate order less or equal to p. The available combinations for this example are

This selection gives the following set of orthogonal polynomials as the vector-valued function of observables for the approximation of the

<sup>&</sup>lt;sup>2</sup>The quantity  $\|\cdot\|_q$  is not a norm because it does not satisfy the triangle inequality.



Figure 3.4: Retained basis terms for the low-rank polynomial basis.

discrete-time Koopman operator.

$$\Psi(x) = \begin{bmatrix} J(0_1)J(0_2) \\ J(1_1)J(0_2) \\ J(2_1)J(0_2) \\ \vdots \\ J(0_1)J(1_2) \\ J(1_1)J(1_2) \\ \vdots \\ J(3_1)J(3_2) \end{bmatrix} = \begin{bmatrix} 1 \\ -x_1 + 1 \\ \frac{1}{2}x_1^2 - 2x_1 + 1 \\ \vdots \\ -x_2 + 1 \\ (-x_1 + 1)(-x_2 + 1) \\ \vdots \\ (-\frac{1}{6}x_1^3 + \frac{3}{2}x_1^2 - 3x_1 + 1)(-\frac{1}{6}x_2^3 + \frac{3}{2}x_2^2 - 3x_2 + 1) \end{bmatrix}$$
(3.63)

If the selection of truncation scheme is a value q = 0.6, the set of retained indices reduces the polynomial basis dimension from 16 elements to 8, and the maximum order from 9 to 3. That is, the remaining indices to perform the tensor product on univariate polynomials are,

$$\alpha = \begin{cases} 0 & 1 & 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 & 3 \end{cases},$$
(3.64)

This new set of indices redefines the vector valued function of observables

$$\Psi(x) = \begin{bmatrix} J(0_1)J(0_2)\\ J(1_1)J(0_2)\\ J(2_1)J(0_2)\\ J(3_1)J(0_2)\\ J(0_1)J(1_2)\\ J(0_1)J(1_2)\\ J(0_1)J(2_2)\\ J(0_1)J(2_2)\\ J(0_1)J(3_2) \end{bmatrix} = \begin{bmatrix} 1\\ -x_1+1\\ \frac{1}{2}x_1^2 - 2x_1 + 1\\ -\frac{1}{6}x_1^3 + \frac{3}{2}x_1^2 - 3x_1 + 1\\ -x_2 + 1\\ (-x_1 + 1)(-x_2 + 1)\\ \frac{1}{2}x_2^2 - 2x_2 + 1\\ -\frac{1}{6}x_2^3 + \frac{3}{2}x_2^2 - 3x_2 + 1 \end{bmatrix}$$
(3.65)

Even though this reduction method eliminates most of the higher order polynomial elements in the vector valued function of observables, it does not guarantee that the remaining elements give an accurate approximation of the Koopman operator. Some polynomial elements in the basis rather than reducing the error of the approximation, will increase it. Therefore, a second reduction method based on the individual error of each element of the basis reduces the dimension even further, and gives a more accurate approximation.

#### 3.4.3 Reduction by Polynomial Accuracy

The idea of the reduction is to calculate the error of the individual multivariate polynomials and eliminate the elements whose error is greater than a specified threshold. The approximation of the Koopman operator from the EDMD algorithm comes from a set of the system orbits. This set of orbits comes from one of two distinct sets, one for solving the least-squares problem (9) and one for testing the accuracy of the solution. Evaluating the posterior time event Y in every element of the vector-valued function, and comparing it with the effect of the operator on the evaluation of the anterior time event X, gives a metric of the contribution of the individual multivariate elements to the accuracy of the solution. Every element of this error criterion,  $\epsilon = (\epsilon_1, \dots, \epsilon_d)$  is

$$\epsilon_l = \frac{1}{N} \sum_{i=1}^{N} \frac{|\psi_l(y_i) - U_{d_l}\psi_l(x_i)|}{|\psi_l(y_i)|}$$
(3.66)

where  $U_{d_l}$  is the *l*-th row of the Koopman operator matrix.

Recall that for the case of orthogonal polynomials, the way to recover the state is by the inverse of n injective multivariate elements, if these elements are the ones with index one in each of the state variables, the recovery is a linear vector-valued function of the observed values. Therefore, the threshold  $\bar{\epsilon}$  for the elimination of elements is the maximum of the errors with an index equal to one,

$$\bar{\epsilon} = \max(\epsilon \colon \|\alpha_l\|_1 = 1) \tag{3.67}$$

and the multivariate polynomial elements that stay as a part of the reduced vector-valued function of observables  $\Psi_R(x)$  are

$$\Psi_R(x) = \{\Psi(x) \colon \epsilon \le \bar{\epsilon}\}.$$
(3.68)

The application of the p-q-quasi norms and polynomial error reductions allow for the approximation of the Koopman operator of the underlying systems with significant fewer basis elements than the current methods and increases the accuracy of the operator. This increase in accuracy allows for the training of the operator with less training trajectories.

#### 3.4.4 EDMD Example

The proposed method for obtaining the low-order polynomials based on the p-q-quasi norms and the error of the multivariate elements is by applying a greedy approach for the calculation of a suboptimal p-q parametrization, and a second calculation of the operator based on the elimination of the multivariate elements.

The greedy approach to find the suboptimal approximation of the Koopman operator relies on the error from the comparison on the test set of orbits, and the predicted orbits of the operator. Consider (3.36), where the k-th application of the operator  $U_d$  gives the evolution of the observables  $\Psi(x)$  up to that time. From this evolved state, consider a matrix  $B \in \mathbb{R}^{d \times n}$  with a unitary value per column in the position where an element of the vector-valued function has an index equal to one, i.e., every column of matrix B is defined as

$$B_l = \{e_l \colon \|\alpha_l\|_1 = 1\}.$$
(3.69)

Consider the set of indices from the reduced p-q-quasi norm reduction of Laguerre polynomials with p = 3 and q = 0.6 in (3.64). The indices that satisfy the condition  $\|\alpha\| = 1$  are the second and fifth polynomial elements. Hence, the *B* matrix for that particular case is

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^{\top},$$
(3.70)

These two polynomial elements are the vector valued function whose inverse recovers the state from the values of  $\Psi(x)$ . This selection gives a state-observable-function  $\Psi_B(x)$  whose inverse recovers the state. For this particular case of Hermite polynomials with the selected truncation scheme

$$\Psi_B^{-1}(x) = \begin{bmatrix} -\psi_2 + 1 & -\psi_5 + 1 \end{bmatrix}^\top.$$
 (3.71)

This serves to define the state evolution map in terms of the discrete approximation of the Koopman decomposition  $\{(\mu_i, \phi_i(x), v_i)\}_{i=1}^d$  (for the aforementioned example, d = 8). The k-th application of the state evolution map  $T^k(x) = x_k$  from the initial condition  $x_0$  is

$$\hat{x}(k) = \Psi^{-1} \left( B^{\top} U_d^k \Psi(x(0)) \right).$$
(3.72)

**Remark 3.** Note that this approximation of the evolution of states is different from (3.54), that describes the evolution as a function of eigenfunctions and their corresponding eigenvalues, while for this case, the approximation is a function of observables evolution weighted by the Koopman matrix Ud. This is an important distinction because the evolution of observables is less prone to error because it only depends on the inverse of G. Conversely, the evolution of eigenfunctions not only depends on the inverse of G, it also depends on the spectral decomposition of  $U_d$ , where additional numerical calculation increments the empirical error (3.56).

The first step of the process is to select a set of training and testing trajectories, and to generate the N pairs of data snapshots. All throughout this thesis, these sets come from the integration of ordinary differential equations from different initial conditions. For an accurate approximation of the discrete-time Koopman operator, there are two important aspects to consider. First, if an orbit reaches steady state, the integration must stop in order to avoid redundant data at steady state that causes the condition number of matrix G to increase making the matrix ill-conditioned and very sensitive to approximation errors. This is the reason why the solution (3.38) is in terms of the Moore-Penrose pseudo inverse. If G has a condition number sufficiently small, the accuracy of the solution increases with the use of the inverse, i.e.,

$$U_d = AG^{-1}.$$
 (3.73)

The second aspect to consider is that in the construction of the N pairs of data snapshots, the last value of each trajectory must be discarded. Given that the integration must stop at steady state, usually, the last value of the trajectory is less than one  $\Delta t$  away from the penultimate, and this small difference decreases the accuracy of the approximation.

For an illustrative example, consider the Duffing equation, a benchmark case in the literature for analyzing the accuracy of Koopman-based methods. The system has different types of behavior depending on the parametrization. This example will consider the case with damping and two basins of attraction.

$$\dot{x}_1 = x_2 \tag{3.74}$$

$$\dot{x}_2 = -\delta x_2 - x_1(\beta + \alpha x_1^2). \tag{3.75}$$

The parametrization for this particular case is  $\delta = 0.5$ ,  $\beta = -1$ ,  $\alpha = 1$ . This parametrization describes a system with two stable focus points at  $(\pm 1, 0)$  and a saddle point at (0, 0). The samples for the system come from 4 uniformly distributed initial conditions over  $x_1, x_2 \in [-2 \ 2]$  and are obtained by numerical integration with a  $\Delta t = 0.1$  until they reach one of the steady states. This means that the number of points per trajectory is not the same. Along with the samples, the training and test sets are 50% of the total samples each. Figure 3.5 shows the testing and training trajectories. Note that for each set, there is one trajectory that converges to each of the asymptotically stable equilibrium points. For the case depicted, the total amount of points in the training set is 404.

From the trajectories of the training set, the next step is to generate the X and Y snapshots sets. Recall that excluding the last element of the orbits increases the accuracy. Furthermore, given that an (x, y)snapshot pair is the value at a time instant, and the value at the next



Figure 3.5: Training and testing set for the approximation of the discretetime Koopman operator via the EDMD algorithm with reduced basis.

one, it is possible to concatenate all the available pairs in two sets for the evaluation of the A and G matrices. Consider a set of  $Tn_r$  training trajectories, each with an  $N_k$  number of points, i.e.,  $Tn_r \in \mathbb{R}^{N_k \times n}$ . The X and Y set are:

$$X = \begin{bmatrix} Tn_1(1:N_{k_1}-2)^\top & Tn_2(1:N_{k_2}-2)^\top & \cdots & Tn_r(1:N_{k_r}-2)^\top \end{bmatrix}^\top$$
(3.76)  
$$Y = \begin{bmatrix} Tn_1(2:N_{k_1}-1)^\top & Tn_2(2:N_{k_2}-1)^\top & \cdots & Tn_r(2:N_{k_r}-1)^\top \end{bmatrix}^\top$$
(3.77)

The third step of the process is to select a set of q norms and a set of maximum multivariate order p. For each p, sweep along the parameter q calculating the error (3.56), then select the suboptimal parameters from the approximation that gives the least one. For the Duffing equation under consideration, the selection of orthogonal polynomials is the same Laguerre basis as for the illustration of the p-q-quasi norm reduction methods with a parameter sweep of p = 3 and  $q = \begin{bmatrix} 0.4 & 0.5 & 0.7 & \infty \end{bmatrix}$ , that includes the infinity norm for illustration purposes of the effect of

the reduction. Figure 3.6 shows the polynomial elements for each of the q values in the sweep, where it is clear that for the case where there is no truncation, the error increases compared to q = 0.7. Additionally, it is important to remark that the proposed algorithm makes an accurate approximation of the system with just one trajectory that converges to each of the asymptotically stable equilibrium points in the training set.



Figure 3.6: Test orbits from the EDMD algorithm with different p-q truncation schemes for a Laguerre Polynomial basis. Solid lines are the theoretical trajectories and dashed lines are the approximation by the Koopman operator

With the first approximation of the discrete-time Koopman operator, the next step is to compute the error (3.66) for the individual multivariate elements and select the elements that will stay on the polynomial basis according to (3.68). Figure 3.7 shows how the second reduction has an effect on the polynomial basis for the different choices of q. The "best" value q = 0.7 does not lead to a polynomial reduction, and to illustrate the procedure, we will focus on q = 0.5 (which is the "worst" value). In this case,

$$\epsilon = \begin{bmatrix} 0.0000 & 0.0017 & 0.0019 & 0.0132 & 0.0165 & 1.1708 & 0.0192 & 0.0637 \end{bmatrix},$$
(3.78)

\_

where the error threshold is the maximum between the second and fifth elements because these correspond to the order one single variable observables from (3.65). As the last three elements in the error vector are larger than the fifth, they must be discarded from the basis where the reduced polynomial basis has indices

$$\alpha = \begin{cases} 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{cases}, \tag{3.79}$$

as a consequence, this polynomial basis is of reduced dimension with an error less than the previous approximation that only had a p-q-quasi norm reduction.

$$\Psi(x) = \begin{bmatrix} J(0_1)J(0_2)\\ J(1_1)J(0_2)\\ J(2_1)J(0_2)\\ J(3_1)J(0_2)\\ J(0_1)J(1_2) \end{bmatrix} = \begin{bmatrix} 1\\ -x_1+1\\ \frac{1}{2}x_1^2 - 2x_1 + 1\\ -\frac{1}{6}x_1^3 + \frac{3}{2}x_1^2 - 3x_1 + 1\\ -x_2 + 1 \end{bmatrix}$$
(3.80)

Note that the polynomial reduction improves the error  $\epsilon$  from 11 to 5.34, which is better but still not as good as the basis corresponding to q = 0.7 with  $\epsilon = 0.634$ . However, the second reduction does give an accurate approximation of five elements that can be enough for control purposes reducing the computational complexity of any synthesis algorithm that makes use of it. Along with this line of thought, and continuing with the example, the approximation of the dynamical system evolution, from (3.38) and (3.72) is,

$$\Psi(k+1) = \begin{bmatrix} 1.0000 & 0.0000 & -0.0000 & 0.0000 & 0.0000 \\ -0.0723 & 0.9171 & 0.0880 & -0.0294 & 0.0967 \\ -0.0448 & -0.2259 & 1.2794 & -0.1041 & 0.1129 \\ 0.0401 & -0.4680 & 0.6016 & 0.7709 & 0.1009 \\ 0.5446 & -1.6488 & 1.7501 & -0.5846 & 0.9407 \end{bmatrix} \Psi(k)$$

$$\hat{x} = \begin{bmatrix} -\psi_2 + 1 \\ -\psi_5 + 1 \end{bmatrix}, \qquad (3.81)$$

giving a linear system of the observables  $\Psi$  and a nonlinear system according to the state x.

\_



Figure 3.7: Test orbits from the EDMD algorithm with polynomial accuracy truncation scheme for a Laguerre Polynomial basis. Solid lines are the theoretical trajectories and dashed lines are the approximation by the Koopman operator.

#### 3.4.5 Approximating The Koopman Operator

Recall that from the solution of either (3.38) or (3.73) depending on the condition number of G, the spectral decomposition of eigenvalues Mand the left eigenvectors  $W^*$  give the set of eigenfunctions of the discretetime approximation of the Koopman operator according to (3.50). As the left eigenvectors come from the inverse of the right ones, this extra computational step hinder the accuracy of eigenfunctions for subsequent analysis. Therefore, in practice, the algorithm calculates the transpose of  $U_d$  as

$$U_d^+ = G \backslash A, \tag{3.83}$$

and proceeds to calculate the spectral decomposition  $[W^*, M] = \text{eig}(U_d^{\top})$ , where both functions are presented in the Matlab® notation that gives the best results. Applying this functions to the example of the Duffing equation gives a spectral decomposition,

```
\Gamma 1.0 + 0.0i
                    0.0000 + 0.0000i
                                       0.0000 + 0.0000i
                                                         0.0000 + 0.0000i
                                                                           0.0000 + 0.0000i
         0.0 + 0.0i
                    0.9565 + 0.1450i
                                       0.0000 + 0.0000i
                                                         0.0000 + 0.0000i
                                                                           0.0000 + 0.0000i
  M =
         0.0 + 0.0i
                     0.0000 + 0.0000i
                                       0.9565 - 0.1450i
                                                         0.0000 + 0.0000i
                                                                           0.0000 + 0.0000i
         0.0 + 0.0i
                     0.0000 + 0.0000i
                                       0.0000 + 0.0000i
                                                         0.9976 + 0.0087i
                                                                           0.0000 + 0.0000i
        0.0+0.0i
                     0.0000 + 0.0000i
                                      0.0000 + 0.0000i
                                                         0.0000 + 0.0000i
                                                                           0.9976 - 0.0087i
                                                                                          (3.84)
        -1.0 + 0.0i
                     0.2066 - 0.0519i
                                         0.2066 + 0.0519i
                                                             -0.1591 + 0.0241i
                                                                                -0.1591 - 0.0241i
        0.0 + 0.0i
                     -0.6508 - 0.0209i
                                        -0.6508 + 0.0209i
                                                            0.5732 + 0.0119i
                                                                                 0.5732 - 0.0119i
W^{\star} =
       -0.0 + 0.0i
                     0.6889 + 0.0000i
                                         0.6889 + 0.0000i
                                                             -0.7505 + 0.0000i
                                                                                -0.7505 + 0.0000i
        -0.0 + 0.0i
                    -0.2294 + 0.0064i
                                        -0.2294 - 0.0064i
                                                             0.2865 - 0.0120i
                                                                                 0.2865 + 0.0120i
        -0.0 + 0.0i
                     -0.0155 + 0.0559i
                                         -0.0155 - 0.0559i
                                                             -0.0064 - 0.0002i
                                                                                -0.0064 + 0.0002i
                                                                                          (3.85)
```

The application of (3.50) gives a set of five eigenfunctions of the truncated approximation of the Koopman operator. Figure 3.8 shows the respective real and imaginary parts of the evaluation of the eigenfunction in the set from which the initial conditions come, i.e.,  $x_1, x_2 \in [-2 \ 2]$ . The surface colors show the different values of the evaluated function, where the yellow (lighter) regions are the bigger values and the blue (darker) regions the smaller. As previously stated, this representation is not suitable for analysis, only for control, because the eigenvalues of the eigenfunctions are near unitary and the surface should reflect an approximation to the two basins of the original system.



Figure 3.8: Eigenfunction of the discrete-time approximation of the Koopman operator for the Duffing equation with a reduced Laguerre polynomial basis.

# 3.5 EDMD for Control

This section presents a description of the traditional algorithm for dealing with forced systems, along with its improvements via embedding trigonometric functions into the state of the system. These embeddings are an a priori expansion before the application of the EDMD. It also shows via simulations and real experiments, the advantages of modeling a pendulum on a cart over the current alternative methods. So far the EDMD algorithm deals with unforced or autonomous systems, and given that for synthesizing controllers there is a manipulation of the different forces that inject energy into the system, the EDMD algorithm needs to be modified such that it can capture the dynamics according to the inputs. The relation of the inputs and the algorithm is established by Korda and Mezić [2018b], where they conclude for the synthesis of controllers, especially in the model predictive control framework, the driven inputs of the system must be affine. Although these approximations for control come from modifying the EDMD algorithm, the relation that they have with the Koopman operator is have not been established yet. Therefore, the remaining of this sections will refer to these as linear predictor, and not the approximation of the Koopman operator.

Consider the non-autonomous nonlinear system  $(\mathcal{M}, \mathcal{U}, T(x, u), k)$  in discrete time, with state variables  $x \in \mathcal{M}$  where  $\mathcal{M} \subseteq \mathbb{R}^n$  is the nonempty compact state space, forcing signals  $u \in \mathcal{U}$  where  $\mathcal{U} \subseteq \mathbb{R}^r$  is the nonempty compact input space,  $k \in \mathbb{Z}_0^+$  is the discrete time, and  $T: \mathcal{M} \times \mathcal{U} \to \mathcal{M}$ is the differentiable vector-valued evolution map, i.e.,

$$x(k+1) = T(x(k), u(k)),$$
(3.86)

where a trajectory, or an orbit of the system is the sequence of states  $(x_i)_{i=0}^k$  that come from the solution of (3.86), which is the successive application of the non-linear mapping T from an initial condition  $x_0 \in \mathcal{M}$  at k = 0 and a specific sequence of forcing signals  $u \triangleq (u_i)_{i=0}^{k-1}$ .

For example, consider a pendulum and a moving cart attached by a swivel that allows the pendulum to rotate freely. The cart wheels rotate on a rail and a DC motor drives the whole system. The available information from two encoders are the displacement of the cart and the angular rotation of the pendulum. Figure 3.9 depicts the experimental set-up where x is the horizontal displacement of the cart and  $\theta$  is the angle of the pendulum with respect to the vertical axis. Mass and energy balances give a set of ordinary differential equations that describe the dynamics of the system.



Figure 3.9: Pendulum on a cart

The system equations, which depend on the masses of the cart (M) and pendulum (m), the length of the pendulum rod (l), the linear damping  $(\mu_x)$  of the cart wheels with the rails and the gravitational constant (g), are given by

$$\dot{x} = v$$
  

$$\dot{v} = \frac{Gu - \mu_x v - ml\theta_v^2 \sin(\theta)}{M + m \sin^2(\theta)}$$
  

$$\dot{\theta} = \theta_v$$
  

$$\dot{\theta}_v = \frac{Gu - \mu_x v - ml\theta_v^2 \sin(\theta)) \cos(\theta) + (M + m)g\sin(\theta)}{l - ml\cos^2(\theta)}, \qquad (3.87)$$

where the states are the cart displacement x, the cart velocity v, the rod angle  $\theta$  and the angular velocity  $\theta_v$ . The model also considers a gain G between the voltage of the motor u and the resulting force on the cart. The parameter values come either from the available data of the manufacturer (Feedback Instruments Ltd) or from an identification of the parameters from a set of data collected in preliminary experiments with the system. Table 3.3 lists the value of these parameters.

	Description	Value	Units
M	Cart Mass	1.12	[kg]
m	Pendulum Mass	0.0905	[kg]
g	Gravity	9.81	$[m \cdot s^{-2}]$
l	Pendulum length	0.365	[m]
$\mu_x$	Cart Friction	6.65	$[\cdot]$
G	Tension-Force Gain	7.5	$[\cdot]$

Table 3.3: Parameters of the inverted pendulum on a cart.

Let us now consider a set of seven orbits or trajectories obtained by numerical integration of (3.87) with the assumption that all the state variables of the system are available. For the accuracy of the algorithm, it is necessary that the trajectory samples are collected at a constant rate, which is chosen equal to 0.01 seconds, both in the numerical simulation and the experimental study. This set of orbits sampled at a constant  $\Delta t$ correspond to the solution of the non-linear discrete-time mapping (3.86).

Figure 3.10 depicts the discrete-time evolution of some of these trajectories with their respective forcing signals. These trajectories serve as the available data to approximate the non-linear dynamics via the EDMD algorithm.

As EDMD is a data-driven approach, the set of trajectories is divided into training and testing sets for the approximation and validation of the algorithm. In our case study, the selection is five orbits for training the linear predictor, and two for testing (Figure 3.10). Finally, the snapshot data is defined as a set of tuples  $\{(x_i, y_i, u_i)\}$ , where  $y_i = T(x_i, u_i)$ . From these tuples, the snapshot matrices are given by:

$$X = \begin{bmatrix} x_1 & \dots & x_N \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 & \dots & y_N \end{bmatrix}, \quad U = \begin{bmatrix} u_1 & \dots & u_N \end{bmatrix}, \quad (3.88)$$

according to the traditional EDMD algorithm [Williams et al., 2015] and the extension for including the inputs of the system [Korda and Mezić, 2018b]. The rationale behind the approach is to get linear predictors of the state evaluated on a vector-valued function of observables  $\Psi(x) =$ 



Figure 3.10: Pendulum trajectories with sinusoidal input.

 $[\psi_1(x), \cdots, \psi_d]^\top : \mathcal{M} \to \mathbb{C}^{d \times 1}$  where d is the dimension of the set of observables that must satisfy the condition

$$\Psi(y) = U_{d,x}\Psi(x) + U_{d,u}u + r(x,u)$$
(3.89)

 $r(x, u) \in \mathcal{F}$  is the residual term that has to be minimized in order to find matrices  $U_{d,x}$  and  $U_{d,u}$ . This leads to the least squares problem

$$\|r(x,u)\|^{2} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \|\Psi(y_{i}) - U_{d,x}\Psi(x_{i}) - U_{d,u}u\|_{2}^{2}, \qquad (3.90)$$

which as a closed-form solution

$$\begin{bmatrix} U_{d,x} & U_{d,u} \end{bmatrix} = \Psi(Y) \begin{bmatrix} \Psi(X) \\ U \end{bmatrix} \begin{pmatrix} \Psi(X) \\ U \end{bmatrix} \begin{bmatrix} \Psi(X) \\ U \end{bmatrix}^{\top} \end{pmatrix}^{-1}.$$
 (3.91)

As the purpose of these predictors is the synthesis of controllers, it is necessary to recover the state from the observable functions. Thus, an additional least squares problem for the best projection matrix of x onto the span of  $\Psi$  has to be solved. The projection must satisfy the condition

$$\bar{x} = U_{d,c}\Psi(x) + r_c(x),$$
 (3.92)

where  $r_c(x)$  is the residual term to be minimized to find  $U_{d,c}$ , that accepts a pseudo-inverse based least squares solution

$$U_{d,c} = X\Psi(X)^+.$$
 (3.93)

Instead of solving (3.93) to recover the state, a common practice is to include the state in the set of observables, so that matrix  $U_{d,c} = [I_n, 0_{n \times (d-n)}]$  after reordering the observables such that the state vector form the first *n* elements.

The EDMD formulation (3.91) and (3.93) can be used with a basis of Jacobi type monomials for the approximation of the pendulum dynamics (3.87), i.e., a sequence of orthogonal polynomials  $\{\pi_{\alpha}(x)\}_{\alpha=0}^{p}$  where  $\pi_{\alpha}(x)$  is a univariate polynomial (i.e., a polynomial in only one of the state variables  $x_i$ , i = 1, ..., n) of degree  $\alpha \in \mathbb{N}_+$  up to order p according to Equation (3.58).

For the approximation of the pendulum dynamics, the order p = 1can be chosen, giving a set of observables of dimension d = 17 with maximum order 4, i.e., the product of all the univariate monomials  $\psi_{17} = \prod_{i=1}^{4} 5x_i + 1$  for parameters a = 5 and b = 3 of a Jacobi type monomial J(1, a, b, x) = a/2 - b/2 + x(a/2 + b/2 + 1). The five trajectories of the training set sum 1110 data points, and the two trajectories of the testing set sum 490 data points. Figure 3.11 shows the trajectories produced by the EDMD, where the empirical error  $\epsilon = 1/N \sum_{i=1}^{N} |x_i - \bar{x}_i|$  is 1.31. While it is possible to perform the approximation with higher p values, the results do not necesserally improve. For p = 2 as the maximum order of the univariate polynomials, the basis has dimension d = 81with a maximum order of eight, which shows that these two numbers grow exponentially with the maximum univariate order p. Moreover, the empirical error increases to 1.77 in this particular case, which is also the sign of numerical issues.

Indeed, there are some inherent numerical instabilities with the method. First, the EDMD algorithm is a linear map on the function space that the set of observables span, and the accuracy of the solution depends on the characteristics of this set. Choosing an orthogonal basis with an observable that corresponds to a constant generally improves the performance of the approximation. In contrast, adding the state to the set of observables is prone to break the orthogonality of the observables, depending on their actual choice. As a consequence, the square matrix  $[\Psi(X), \ U]^{\top}[\Psi(X), \ U]$  in (3.91) becomes singular or close to singular, and while replacing the inverse by a Moore-Penrose pseudo inverse can partly alleviate this issue, the result can still be inaccurate. Second, pre-


Figure 3.11: Testing trajectories of the traditional computation of the EDMD algorithm with a Jacobi orthogonal polynomial basis up to order four. Solid lines are the orbits from the numerical integration of the ODE and dashed lines the approximation by the EDMD algorithm.

serving an orthogonal basis, without explicitly including the states as observables poses a similar matrix inversion problem, when there is no solution to the projection of the state space onto the set of observables.

The way out of this dilemna is the selection of order-one, univariate, injective, polynomial elements for the recovery of the state, which implies that there is no need of breaking the orthogonality of the set while still being able to recover the state as a linear function of the observables, completely avoiding the burden of a matrix inversion as described in Section 3.4.1. Besides, the exponential growth of the maximum order and dimension of the set of observables based on orthogonal polynomial has a solution via p-q-quasi norms and polynomial accuracy methods as described in Section 3.4.3.

In our case study of the inverted pendulum, we can consider the same orthogonal basis of Jacobi polynomials with a sweep of p-q parameters corresponding to: p = [2, 3, 5,] and q = [0.3, 0.5, 0.7, 0.9, 1.1, 1.3]. Although there are 18 possible combinations, some p-q parametrizations produce equal basis. There are only 12 distinct sets of observables, ranging from 6 elements of maximum order 2 to 173 elements of maximum order 5. The result of the reduction gives a sub-optimal basis of 33 elements of maximum order 3 that achieves an empirical error of 0.62, as compared to 1.31 for the traditional EDMD. Note that the reduction provides the possibility to test higher-order polynomial basis than in the traditional form, where a basis of maximum order 5 would count 625 elements. Figure 3.12 shows the performance of the sub-optimal basis of the p-q-EDMD in comparison to the traditional EDMD.



Figure 3.12: Testing trajectories of the traditional EDMD computation against the p-q-EDMD of polynomial elements up to order two. Solid lines are the orbits from the numerical integration of the ODE, dashdotted lines are the approximation of the EDMD algorithm, and dashes lines are the approximation by the p-q-EDMD algorithm.

Although the use of reduced orthogonal polynomials for the p-q-EDMD provides a method to improve the accuracy of the algorithm while avoiding computationally heavy high-order and dimensional solutions, the accuracy of the algorithm for systems that have trigonometric components or an arbitrary behavior like exponentials or logarithms is not enough. Therefore, the next section introduces the concept of trigonometric embeddings.

### 3.5.1 Trigonometric Embeddings

In this section, we consider the problem of representing dynamic systems with oscillatory behavior by polynomial expansions and the possibility to increase the parsimony of the approximation by the inclusion of trigonometric functions as elementary units of the polynomial expansion. Similar embeddings could be used for any particular behavior, using associated functions. Similar to the idea that took the dynamic mode decomposition algorithm to the extended version, where instead of performing a regression on the states, the extended method considers a set of functions of the state (the so-called observables), the trigonometric embeddings, or more generally function embeddings, provide specific functions of the state conveying particular information.

Consider the discrete-time dynamical system (3.86) and assume that a subset of the state  $x^{tg} \subseteq x$  has trigonometric components in the difference equation T(x, u). For each of these state variables an extension of the state space is defined including a pair of sine and cosine:

$$x_{e} = \begin{bmatrix} x_{e_{1}} \\ \vdots \\ x_{e_{2m+1}} \end{bmatrix} = \begin{bmatrix} x_{1} \\ \vdots \\ x_{n} \\ \sin x_{1}^{tg} \\ \cos x_{1}^{tg} \\ \vdots \\ \sin x_{m}^{tg} \\ \cos x_{m}^{tg} \end{bmatrix}, \qquad (3.94)$$

where  $m \leq n$ . Using this extended set of variables, the approximation of the dynamics can be achieved via the p-q-EDMD.

Consider for example an arbitrary discrete-time dynamical system  $(\mathcal{M}, \mathcal{U}, T(x, u), k)$ , where n = 2, r = 2 and  $x^{tg} = x_2$ , i.e., the system has two state variables where the second one has a trigonometric component and two inputs within the non-linear mapping T(x, u). Moreover a Hermite basis of orthogonal polynomials of univariate elements up to order 2, i.e.,  $\pi_{\alpha=[0, 1, 2]}(x) = [1, 2x, 4x^2 - 2]$ , is used. For illustration purposes, assume an arbitrary p-q parametrization p = 3 and q = 0.7. The resulting basis of polynomials for the approximation of the dynamics

is

$$\Psi(x) = \begin{bmatrix} 1\\ 2x_1\\ 4x_1^2 - 2\\ 2x_2\\ 4x_1x_2\\ 4x_2^2 - 2\\ 2\cos(x_2)\\ 4x_1\cos(x_2)\\ 4x_2\cos(x_2)\\ 4\cos^2(x_2) - 2\\ 2\sin(x_2)\\ 4x_2\sin(x_2)\\ 4x_2\sin(x_2)\\ 4x_2\sin(x_2)\\ 4\cos^2(x_2) - 2 \end{bmatrix}.$$
(3.95)

Note that the embeddings are not restricted to trigonometric functions, but could for instance include logarithmic, exponential or hyperbolic functions, if the non-linear mapping T(x, u) has state variables with such a behavior. However, the functional embeddings and particularly the trigonometric embeddings, that add two more variables for each trigonometric state, increases exponentially the dimension of the set of observables and it is necessary to resort to the previous p-q-quasi norm reduction.

F

The application of trigonometric embeddings to the orbits of the pendulum problem, considering that only the angle  $\theta$  has trigonometric components, reduces the empirical error of the approximation, from the aforementioned value of 0.62 (provided by p-q-EDMD) to 0.17. To achieve these results, we consider a p-q sweep where p = [2, 3, 4, 5], q = [0.3, 0.5, 0.7, 0.9, 1.1, 1.3] and the available orthogonal polynomials in Matlab. The sub-optimal solution is a Laguerre polynomial basis with parameters p = 4 and q = 0.7. Although the approximation error in the test set is reduced, the inclusion of the two extra trigonometric variables, and the increased p value produces a basis of 65 elements. Figure 3.13 depicts the result of the algorithm in comparison with the benchmark EDMD.



Figure 3.13: Testing Trajectories of the traditional EDMD against the pq-EDMD with trigonometric embeddings. Solid lines are the orbits from the numerical integration of the ODE, dash-dotted lines are the approximation of the EDMD algorithm, and dashes lines are the approximation by the p-q-Trigonometric EDMD algorithm.

## 3.5.2 Inverted Pendulum: Experimental Results

For illustration purposes, and to compare numerically various expansions, we consider the real-life application provided by a Feedback Digital Pendulum 33-005-PCI (Figure 3.14).

This pendulum is the same as described in the beginning of this Section. A DC motor drives the cart along the rail to which a pendulum is attached. The available experimental set-up provides, through two encoders, the noisy measurements of the cart position and the pendulum angle every 0.01 [s]. Therefore, the output equation is given by:

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ v \\ \theta \\ \theta_v \end{bmatrix} + w_n,$$
(3.96)

where  $w_n \sim \mathcal{N}(0, \sigma^2)$ .

However, the knowledge of the cart velocity and the angular velocity



Figure 3.14: Feedback Digital Pendulum 33-005-PCI

are necessary to compute an approximation through the EDMD algorithm. A simple differentiation of the displacements gives an amplification of the noise, impeding the possibility of getting accurate approximations of the dynamics. This can be alleviated by the design of two Kalman filters based on simple kinematic expressions. The Kalman filter [Paul Zarchan, 2009] is a model-based technique which allows recovering on-line estimations by blending the prediction of a mathematical model with the available on-line measurements. Here, the idea is to avoid using a full differential equation model such as (3.87), as this would be contradictory with the objective of developing a data-driven EDMD model. Hence, the method relies on basic kinematic relations

$$x_{k+1} = x_k + v_k \cdot \Delta t + a_k \cdot 0.5 \cdot \Delta t^2$$
  

$$v_{k+1} = v_k + a_k * \Delta t$$
  

$$a_{k+1} = a_k,$$
(3.97)

where the acceleration is assumed constant. These expressions are used

as predictor while the state update equation of the Kalman filter is

$$\begin{bmatrix} \hat{x}_k \\ \hat{v}_k \\ \hat{a}_k \end{bmatrix} = \begin{bmatrix} \hat{x}_{k-1} \\ \hat{v}_{k-1} \\ \hat{a}_{k-1} \end{bmatrix} + K_k \begin{bmatrix} x_k - \hat{x}_{k-1} \\ \frac{x_k - \hat{x}_{k-1}}{\Delta t} \\ \frac{x_k - \hat{x}_{k-1}}{0.5\Delta t^2} \end{bmatrix},$$
(3.98)

where  $\hat{\cdot}$  is the estimate of the corresponding variable,  $x_k$  is the measurement of the cart position or pendulum angle and  $K_k$  is the Kalman gain that gets updated in every time-step of the algorithm according to the estimation covariance matrix (solution to a Ricatti equation), and the measurement uncertainty. Furthermore, each of the position-velocity pairs has an independent filter with their corresponding parametrization.

For generating the experimental data, the pendulum starts at the stable point  $\theta(0) = \pi$  and is exited with a sinusoidal signal at various frequencies and amplitudes, i.e.,

$$\begin{bmatrix} x_0\\v_0\\\theta_0\\\theta_{v0} \end{bmatrix} = \begin{bmatrix} 0\\0\\\pi\\0 \end{bmatrix}, \ u = A \cdot \sin(\omega t + \phi), \tag{3.99}$$

where the ranges of the different parameters of the forcing signal are:  $A \in (0.1, 1), \omega \in (\pi, 3\pi), \phi \in (0, 2\pi)$ . The selection of these parameters ensures that the cart movement does not exceed the track limits. Figure 3.15 depicts the result of gathering and filtering the experimental data.

The application of the p-q trigonometric EDMD on the experimental data is a linear predictor suitable for controller synthesis. To this end, we consider a p-q sweep where p = [2, 3, 4], q = [0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3], a trigonometric embedding over the third state  $\theta$  and an evaluation of all the available orthogonal polynomials in Matlab. Additionally, and for the consistency of the results, the training and testing sets correspond with the ones used in the simulation results. Table 3.4 shows empirical errors of the best sub-optimal solution for each polynomial basis.

The sub-optimal solution is a Chebyshev polynomial of the first type with parameters p = 4, q = 0.9 and 85 observables that gives an empirical error of 0.57. Figure 3.16 shows the results of the approximation in the test set.



Figure 3.15: Data filtering of the test set for the p-q-Trigonometric EDMD. Solid lines are the orbits from the experimental set-up and dashed lines are the orbits from the Kalman filter.

## **3.6** Interconnected Dynamical Systems

An interconnected dynamical system is a collection of dynamical systems, and an interconnection topology that describes the energetic exchange between them, i.e., which output of system A injects energy into which input of system B.

Consider a system  $\Sigma = \left( \{\Sigma_i\}_{i=1}^N, G_{i,j} \right)$  given as a collection of dynamical systems

$$\Sigma_i : \begin{cases} \dot{x}_i = f_i(x_i) + u_i + G_{i,j} \\ y_i = h_i(x_i) \end{cases} , \qquad (3.100)$$

where  $x_i \in \mathbb{R}^{n_i}$  (i = 1, ..., M) is the state of the  $i^{th}$  subsystem, M is the number of subsystems,  $u_i \in \mathbb{R}^{r_i}$  is the input of the  $i^{th}$  subsystem, and  $y_i \in \mathbb{R}^{m_i}$  is the output of the  $i^{th}$  subsystem. The input  $u_i$  can be divided into external inputs  $u_i$ , and internal inputs  $G_{i,j}$ , that come from the neighboring subsystems of the network.

Figure 3.17 depicts this situation, where the first subsystem in the network provides energy to its neighbors and vice-versa. This interaction

Polynomial basis	Empirical error
Hermite	0.8237
Legendre	0.8249
Laguerre	0.8816
ChebyshevT	0.5731
ChebyshevU	0.8245
Gegenbauer	0.8254
Jacobi	0.8231

Table 3.4: Empirical error for each polynomial basis used in the p-q Trigonometric EDMD.

potentially changes the location and stability of the first subsystem fixed points.

From a controller synthesis perspective, where the external inputs must drive the system to a desired state, the controller must account for the effect of the interconnection input. Figure 3.18 shows the topology of the controller assuming that it has information of the interconnection states.

The objective is to generate linear predictors that accurately describe the local dynamics of a system without having the complete information of the driving force given by the neighboring systems. Next section introduces the concept of linear predictors with a combination of affine, and non-affine input signals.

## 3.6.1 A Case Study: Two Duffing Oscillators

For illustration purposes, consider an interconnection of two Duffing oscillators where the parameters yield a dynamic behavior with two basins of attraction (see Figure 3.19). The differential equations that describe this system are



Figure 3.16: Testing trajectories of the experimental set-up with the p-q-Trigonometric EDMD. Solid lines are the orbits from the filtered experimental set-up, dash-dotted lines are the approximation of the EDMD algorithm, and dashes lines are the approximation by the p-q-Trigonometric EDMD algorithm.

$$\dot{x}_{1,1} = x_{1,2}$$

$$\dot{x}_{1,2} = 0.5x_{1,2} + x_{1,1} - x_{1,1}^3 + 2.2x_{2,1} + u_1$$

$$\dot{x}_{2,1} = x_{2,2}$$

$$\dot{x}_{2,2} = 0.5x_{2,2} + x_{2,1} - x_{2,1}^3 + 2.2x_{1,2} + u_2,$$
(3.101)

where the first state variable of each system drives the second one of its counterpart, e.g.,  $G_{i,j} = 2.2x_{i,j}$ . Figure 3.19 shows the behavior of the two systems where the inputs are randomly selected step signals over  $u_1, u_2 \in [-2 \ 2]$ . The numerical integration with a time step  $\Delta t = 0.05$ , and six randomly selected initial conditions  $x_{i,j}(0) \in [-2 \ 2]$ , gives a set of orbits that can be exploited as dataset for the EDMD algorithm. The training set consists of four trajectories, while two are kept for testing. All the trajectories are of different lengths because the accuracy of the algorithm decreases with the presence of redundant data at the asymptotically stable points, and the integration is stopped upon convergence. As a consequence, the sample size for the approximation of the linear predictors is 3.324 where 2.150 are available for training and 1.174 for testing.



Figure 3.17: Subsystem feedback interconnection with the network.

Recall that it is necessary to have an input-affine EDMD approximation to apply a model predictive control algorithm. Although this is the case for the input space, it is not necessary to have that restriction for the interconnection inputs. Therefore, the suitable approximation for the synthesis algorithms is hybrid. In order to get such a behavior, first consider an approximation where none of the input are affine and the state of the system is expanded to include the inputs similarly to the approach by Korda [Korda and Mezić, 2018b].

### 3.6.2 Development of the EDMD Algorithm

Consider first a non-affine approximation where the set of state variables is expanded to include the inputs as states, as proposed in [Korda and Mezić, 2018b]. This means that the inputs are state variables of the system with zero dynamics, i.e., u(k + 1) - u(k) = 0.

With the trajectories of the training and testing sets and the previous rationale for handling the inputs of the system, define the sets of snapshots as



Figure 3.18: Subsystem local controller with measurement of the interconnection state.

$$XU = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ x_{\Sigma,1} & x_{\Sigma,2} & \cdots & x_{\Sigma,N} \\ u_1 & u_2 & \cdots & u_N \end{bmatrix} \quad YU_+ = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ y_{\Sigma,1} & y_{\Sigma,2} & \cdots & y_{\Sigma,N} \\ u_{1+} & u_{2+} & \cdots & u_{N+} \end{bmatrix}$$
(3.102)

where the + subscript denotes the time shift of the input by one  $\Delta t$ and  $x_{\Sigma}$  are the interconnection input variables. Generally, this shifted value of the input is not available for the last time instance in the orbits, however, given that numerical integration stops at convergence, the last sample of the simulation is usually not a complete  $\Delta t$  after the penultimate, and keeping this value decreases the accuracy of the algorithm. Additionally, note that this representation includes any number of inputs as state variables, and in turn, it is not limited to single-input systems.

The lifting process consists in evaluating the new state vector with a vector valued function of observables  $\Psi = [\psi_1 \cdots \psi_d]^T : \mathcal{M} \times \mathcal{U} \to \mathbb{C}^{1 \times d}$ , i.e.,



Figure 3.19: Two interconnected Duffing oscillators: 4 trajectories constitute the training set and 2 trajectories are kept for testing purposes.

$$\Psi(x,u) = \begin{bmatrix} \Psi_x(x) \\ \Psi_{xu}(x,u) \\ \Psi_u(u) \end{bmatrix} = \begin{bmatrix} \psi_1(x) \\ \vdots \\ \psi_{dx}(x) \\ \psi_{dx+1}(x,u) \\ \vdots \\ \psi_{dxu}(x,u) \\ \psi_{dxu+1}(u) \\ \vdots \\ \psi_{du}(u) \end{bmatrix}.$$
(3.103)

69

For the selection of these functions, we will continue with orthogonal polynomials that have the advantage of improving the accuracy of the approximation. Additionally, benefiting from the aforementioned reduction methods for the dimension and maximum order of the basis achieve an accurate approximation while avoiding computationally expensive basis [Garcia-Tenorio et al., 2020].

The conditions that these observables must satisfy to get the nonaffine predictor is

$$\Psi_x(y, u_+) = U_d \Psi(x, u) + r(x, u), \qquad (3.104)$$

where  $r(x, u) \in \mathcal{F}$  is the residual term to minimize in order to find the regression or linear predictor matrix  $U_d$ . Considering that the best approximation minimizes the residual term, the following objective function

$$\|r(x,u)\|^{2} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \|\Psi(y_{i}, u_{i+}) - U_{d}\Psi(x_{i}, u_{i})\|_{2}^{2}, \qquad (3.105)$$

gives an accurate approximation of  $U_d$ , through the ordinary least-squares solution [Williams et al., 2015], that has a closed analytical form given by

$$G = \Psi(XU)\Psi(XU)^{\top} \qquad A = \Psi(YU^{+})\Psi(XU)^{\top}, \qquad (3.106)$$

$$U_d^{\top} = G \backslash A. \tag{3.107}$$

Even though this approximation is suitable for recovering the spectrum, the evolution of the input space is not important for controller synthesis algorithms. Recall from (3.103) that there is a division of the polynomial basis in subsets that depend on just the state, the input or both. Assume that the dimension of these subsets is:  $\dim \Psi_x(x) \in \mathbb{N}^{dx}$ ,  $\dim \Psi_{x,u} \in \mathbb{N}^{dxu}$  and  $\dim \Psi_u \in \mathbb{N}^{du}$ . To get a linear predictor matrix Lfrom the sets of snapshots, define a  $A_L$  matrix as

$$A_L = \Psi_x(Y)\Psi(XU)^\top, \qquad (3.108)$$

i.e., from the original definition of matrix A (3.106) slice out the last dxu + du rows from matrix A to get this new matrix  $A_L$ . Then, the linear predictor that describes the evolution of state observables is

$$L = A_L G^{-1}, (3.109)$$

or from the original definition of the EDMD matrix for the non-affine forced approximator (3.107), this linear predictor matrix comes from slicing the last dxu + du rows of  $U_d$ . Furthermore, slicing the columns of the linear predictor matrix into the matrix bocks that weight each of the observables sets gives the evolution of state observables as

$$\Psi(y) = L_x \Psi_x(x) + L_{xu} \Psi_{xu}(x, u) + L_u \Psi_u(u), \qquad (3.110)$$

where  $L_x = L(1 : dx)$ ,  $L_{x,u} = L(dx + 1 : dx + dxu)$  and  $L_u = L(dx + dxu + 1 : end)$ . The order one observables that capture each of the state variables. Accordingly, the approximation of the state evolution map  $T^k(x, u) = x_k$  for a known sequence of inputs and from the initial condition  $x_0$  in terms of the state observables is

$$\hat{x}(k+1) = \Psi_B^{-1}(B^\top \Psi_x(y(k))).$$
(3.111)

Consider Jacobi orthogonal polynomials and a p-q parameter sweep where  $p = [3 \ 4 \ 5]$  and  $q = [0.4 \ 0.5 \ 0.6]$ . The result of the non-affine algorithm on each of the subsystems with their corresponding training and testing sets is depicted in Figure 3.20, and the summary of the results is in Table 3.5.

Table 3.5: p, q, dimension, and empirical error for the non-affine approximation of the interconnected duffing equation.

	p	q	$\dim$	$\epsilon$
System 1	4	0.4	13	0.74
System 2	4	0.4	13	0.65

Even though this representation is not suitable for control, it illustrates the advantages of the choice of orthogonal polynomials with p-qquasi norm reduction. The linear predictors are able to approximate the



Figure 3.20: Test orbits from the non-affine forced EDMD algorithm. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm

dynamics of two systems, each with four state variables with the least amount of data, and a substantial reduction in the amount of polynomial elements. Additionally, this approximation is suitable for analysis based on the spectral decomposition of the Koopman operator as shown in Chapter 4. Likewise, the eigenfunctions of the Koopman operator should be useful to analyze forced systems, not only in terms of equilibrium points, their stability and regions of attraction, but also in terms of bifurcations and limit cycles. These extensions are still an open subject for research.

Consider next the restriction of the observables on the input space to be affine, i.e., the set of observables for this input signal is the constant function, i.e.,  $\Psi_u = u$ , while the interconnection inputs remain state variables. This restriction modifies the set of observables that becomes,

$$\Psi(x, x_{\Sigma}, u) = \begin{bmatrix} \Psi_x(x) \\ \Psi_{xu}(x, x_{\Sigma}) \\ u \end{bmatrix} = \begin{bmatrix} \psi_1(x) \\ \vdots \\ \psi_{dx}(x) \\ \psi_{dx+1}(x, x_{\Sigma}) \\ \vdots \\ \psi_{dxu}(x, x_{\Sigma}) \\ u_1 \\ \vdots \\ u_d \end{bmatrix}.$$
 (3.112)

The same polynomial basis and p-q parameter sweep is used as in the previous developments, and the result of the input-affine approximation of each of the subsystems with their corresponding training and testing sets is depicted in Figure 3.21, and the summary of the results is in Table 3.6.



Figure 3.21: Test orbits from the input-affine forced EDMD algorithm. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm

Notice that the suboptimal q value for the input-affine approximation is larger than in the non-affine case, and, as a consequence, the dimension of the polynomial basis should be larger, which is not the case. Indeed, all the polynomials on the input space are eliminated leaving only the

	p	q	$\dim$	$\epsilon$
System 1	4	0.5	10	1.14
System 2	4	0.5	10	1.73

Table 3.6: p, q, dimension, and empirical error for the input-affine approximation of the interconnected duffing equation.

constant terms. As a consequence, the empirical error increases, but the approximation is suitable for controller synthesis, and depending on the case, the accuracy could be enough for control.

In addition, this approximation requires the value of the interconnection input, which is a drawback of the algorithm, which is not suitable for decentralized controllers where the available information is the local state. The solution to this issue goes along the line of the statistical, geometric and observability properties of the interconnection inputs.

For illustrative purpose, consider the input-affine approximation without the knowledge of the interconnection signal, i.e., this variable enters the linear predictor as another state variable of the system where the only information provided to the predictor is the initial condition. Figure 3.22 depicts how the linear predictors give an approximation of their own state, and the interconnection input. For this particular case, all the states, including the interconnection, converge to the vicinity of their corresponding asymptotically stable equilibrium point, and are accurate for the first time steps of the trajectory, meaning that an MPC controller synthesis with short control horizon should be suitable for driving the system.

Although this contribution is a step further into the possibility of controlling such systems, the algorithm still relies on the partial knowledge of the energy injected by neighboring subsystems. Consequently, the method is not suitable yet for decentralized control, where there is only information of the local state.



Figure 3.22: Test orbits from the input-affine forced EDMD algorithm, where the interconnection input is unknown. Solid lines are the theoretical trajectories and dashed lines are the approximation by the algorithm.

### 3.6.3 Controlling the Interconnection

Coupled with the extensions to the formulation of the EDMD based linear predictors for interconnected systems, the synthesis technique to drive them will be the traditional multiple-input multiple-output model predictive control (MPC) with integral action [Di Ruscio, 2013].

Consider a non-autonomous linear discrete time system

$$\begin{aligned}
x(k+1) &= A_L x(k) + B_L u(k) \\
y(k) &= C_L x(k),
\end{aligned} (3.113)$$

with state variables  $x \in \mathcal{M} \subseteq \mathbb{R}^n$ , input space  $u(k) \in \mathcal{U} \subseteq \mathbb{R}^l$ , output

space  $y \in \mathcal{M}_y \subseteq \mathbb{R}^m$ , discrete-time  $k \in \mathbb{N}_0^+$ , and system matrices  $A_L$ ,  $B_L$  and  $C_L$ , that describe the evolution of the system from an arbitrary initial condition  $x_0 = x(0) \in \mathcal{M}$  for a sequence of inputs  $u \triangleq (u_i)_{i=0}^k$ .

The classical formulation of the MPC with integral action solves for the rate of change of the input signal  $\Delta u_k = u_k - u_{k-1}$  for each time instance according to the evolution of the system on a prediction and control horizon  $N_{pc}$ . The control sequence that drives the system to the desired output state comes from the solution of the cost function

$$J_{k} = \left(y_{k+1|N_{pc}} - r_{k+1|N_{pc}}\right)^{\top} Q \left(y_{k+1|N_{pc}} - r_{k+1|N_{pc}}\right) + \Delta u_{k|N_{pc}}^{\top} P \Delta u_{k|N_{pc}}^{\top}, \qquad (3.114)$$

where  $r_k \triangleq (r_i)_{i=k}^{k+N_{pc}}$  is the reference signal on the prediction horizon, the matrix  $Q \in \mathbb{R}^{N_{pc}m \times N_{pc}m}$  is block diagonal with  $N_{pc}$  symmetric positive semi-definite matrices and  $P \in \mathbb{R}^{N_{pc}r \times N_{pc}r}$  also a block diagonal matrix of  $N_{pc}$  symmetric positive semi-definite matrices.

The solution is equivalent to a quadratic problem with objective function (3.114) subject to (3.113) and accepts constraints on the maximum and minimum values of the inputs or their rate of change.

With some matrix manipulations, and considering a smaller control horizon than the prediction horizon for computational efficiency, the problem in standard quadratic programming form is

$$\min_{u} u^{\top} H u + h^{\top} u$$
  
s.t.  $L_c u \le b,$  (3.115)

where H, h, are the constant matrix and vector of the problem, and  $L_c$  and b are the matrix and vector for the linear inequality constraints.

The application of the controller to the input-affine and interconnection non-affine predictor for the coupled Duffing equation problem, where the prediction horizon is ten steps and the control horizon four drives both systems to the origin. Figure 3.23 shows the result of applying the algorithm to all the initial conditions from the training and test sets.

This result shows the potential of the algorithm to handle on bigger networks, where the injection of energy comes from several neighboring systems.



Figure 3.23: MPC synthesis for the interconnected Duffing equation.

A potential improvement to the methodology is the achievement of a decentralized controller. Given the fact that it is possible to approximate the interconnection free response of each system, the difference between these two models is a potential solution to the approximation of the interconnection effect.

## 3.7 Biochemical System Models

Reaction networks describe the dynamical behavior in chemical and biochemical systems, and the most commonly used method to model their behavior are polynomial ordinary differential equations. These sets of equations are mathematically well-established, and they have a broad range of numerical solution techniques and solvers. Additionally, they model the systems in a deterministic way, according to the average behavior of the state variables [Feinberg, 1987]. The advantages of using these systems is their non-negativity and compartmental characteristics, coupled with the fact that depending on the parametrization, the fixed points of the system are hyperbolic [Chellaboina et al., 2009]. Meaning that they satisfy the assumptions to perform the analysis present in this thesis.

Each biochemical reaction network gives rise to a dynamical system, which describes the evolution of the species concentrations in the network over time. Each biochemical reaction has three sets associated to it, the set of chemical species S, the set of complexes of the network C, and the set of reactions in the network  $\mathcal{R}$ . Reactions between complexes, composed by species. The following set of reactions constitute a network of 6 reactions between 5 complexes composed by 5 chemical species.

$$s_{1} \xrightarrow{r_{1}} 2s_{2}$$

$$s_{1}+s_{3} \xrightarrow{r_{3}} s_{4}$$

$$r_{5} \xrightarrow{s_{2}+s_{5}} r_{6}$$

$$(3.116)$$

For a network such as (3.116), and for a number of m chemical species, a number of n complexes in the network, and a number of r reactions in the network. We can write the set  $S = \{s_1, s_2, s_3, s_4, s_5\}$ , which has an associated composition vector x of nonnegative molar concentrations per species,  $x \in \mathbb{R}^n_+$ , the set of complexes,  $C = \{s_1, s_2, s_1 + s_3, s_4, s_2 + s_5\}$ , the set,

$$\mathcal{R} = \begin{cases} s_1 \to 2s_2 \\ 2s_2 \to s_1 \\ s_1 + s_3 \to s_4 \\ s_4 \to s_1 + s_3 \\ s_2 + s_5 \to s_1 + s_3 \\ s_4 \to s_2 + s_5 \end{cases}, \qquad (3.117)$$

with the associated reactions rate vector function  $\rho$  :  $\mathcal{R} \to \mathbb{R}^r_+$  that

associates a positive rate constant to each reaction. Having these sets, the definition of a reaction kinetics network is:

**Definition 3.7.1** (Reaction Network). A biochemical reaction network consists in a quadruple  $\{S, C, \mathcal{R}, \rho\}$  such that:

- i. the finite set S is the set of species,
- ii. the finite set C is the finite set of complexes,  $C \subset \mathbb{R}^{S}$ , where each element of the set is the sum of species

$$\sum_{s \in \mathcal{S}} \alpha_s s_s \tag{3.118}$$

where  $\alpha_s$  is the stoichiometric coefficient of the s<sup>th</sup> species.

- iii. The finite set  $\mathcal{R}$  is the set of reactions,  $\mathcal{R} \subset \mathcal{C} \times \mathcal{C}$  such that,
  - a.  $y \to y \notin \mathcal{R}, \quad \forall y \in \mathcal{C}.$
  - b. For each  $y \in C$  there exists  $y' \in C$  such that  $y' \to y \in \mathcal{R}$  or such that  $y \to y' \in \mathcal{R}$ .
- iv. And the vector function  $\rho$  that associates a positive rate constant to each reaction.

Given an initial concentration  $x(0) \in \mathbb{R}^n_+$  of the interacting species, kinetic rates, and the topology of the biochemical network, such as (3.116), These ODEs measure how the concentration x(t) of each species evolves with time. They describe the velocity of the composition changes, the velocity of a given reaction, its derivative with respect to time.

**Definition 3.7.2** (Species ODE). For a reaction network  $\{S, C, \mathcal{R}, \rho\}$ the formation rate function  $f_s(x) : \mathbb{R}^n_+ \to \mathbb{R}$  is

$$f_s(x) = \sum_{y \to y' \in \mathcal{R}} (\alpha'_s - \alpha_s) \rho_{y \to y'} \prod_{s \in \mathcal{S}} x_s^{\alpha_s}.$$
 (3.119)

This formation rate function is the time derivative of species concentration, i.e.,  $\dot{x} = f(x)$ . Consider for example the first species of network (3.116), according to the species differential equation ODE, its formation rate is

$$f_{1}(x) = (0-1)r_{1}(x_{1}^{1}x_{2}^{0}x_{3}^{0}x_{4}^{0}x_{5}^{0}) + (1-0)r_{2}(x_{1}^{0}x_{2}^{2}x_{3}^{0}x_{4}^{0}x_{5}^{0}) + (0-1)r_{3}(x_{1}^{1}x_{2}^{0}x_{3}^{1}x_{4}^{0}x_{5}^{0}) + (1-0)r_{4}(x_{1}^{0}x_{2}^{0}x_{3}^{0}x_{4}^{1}x_{5}^{0}) + (1-0)r_{5}(x_{1}^{0}x_{2}^{1}x_{3}^{0}x_{4}^{0}x_{5}^{1}) = -r_{1}x_{1} + r_{2}x_{2}^{2} - r_{3}x_{1}x_{3} + r_{4}x_{4} + r_{5}x_{2}x_{5}$$
(3.120)

To extend the per-species definition of the ODE to a matrix notation for the complete differential equation. First, define the matrix exponentiation of  $x \in \mathbb{R}^n$  by matrix  $\alpha = [\alpha_{i,j}] \in \mathbb{R}^{r \times n}$  as  $x^{\alpha} \in \mathbb{R}^r$  such that its *i*-th component is the product  $x_1^{\alpha_{i,1}} x_2^{\alpha_{i,2}} \cdots x_n^{\alpha_{i,n}}$ . Then, define  $\circ$  as the element wise Hadamard product, and finally, the MAK ODE is

$$f(x) = (\alpha' - \alpha)^{\top} (x^{\alpha} \circ \rho)$$
(3.121)

According to the previous definition, the ordinary differential equation of reaction network (3.116) comes from the matrices

$$\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \alpha' = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \qquad (3.122)$$

that are the stoichiometric coefficients for reactants and products respectively for each reaction in the network, the vector of reaction rates, and the species in the network. Coupled with (3.121), the ODE that represents the dynamics of the average species concentration is

$$\dot{x} = \begin{bmatrix} -1 & 1 & -1 & 1 & 1 & 0 \\ 2 & -2 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{pmatrix} \begin{bmatrix} x_1 \\ x_2^2 \\ x_1x_3 \\ x_4 \\ x_2x_5 \\ x_4 \end{bmatrix} \circ \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix} \end{pmatrix}$$
(3.123)  
$$-r_1x_1 + r_2x_2^2 - r_3x_1x_3 + r_4x_4 + r_5x_2x_5$$
$$2r_1x_1 - 2r_2x_2^2 - r_5x_2x_5 + r_6x_4$$
$$= -r_3x_1x_3 + r_4x_4 + r_5x_2x_5$$
$$r_3x_1x_3 - r_4x_4 - r_6x_4$$
$$-r_5x_2x_5 + 5r_6x_4$$

## 3.8 Summary

This section presents the traditional theory regarding regions of attraction, the Koopman operator and the approximation of its discretetime truncation via the EDMD algorithm, where the contributions regarding the algorithm are: A basis reduction method based on p-q-quasi norms that reduce the dimension and order of the approximation. As a consequence, increasing the accuracy of the algorithm. On the subject of dimension, a basis reduction method based on polynomial accuracy, that combined with the p-q-quasi norm reduction gives lower dimensional approximation suitable for subsequent analysis and synthesis algorithms that benefit computationally from the use of smaller approximations.

Additionally, and contrary to traditional algorithms, using the inverse of order-one polynomial elements to recover the state does not break the orthogonality of the basis when the single state variable observables are included in the vector valued function of observables. Additionally, there is no need to include additional optimization algorithms in the process to recover the state.

Further expanding the EDMD algorithm for the incorporation of a priori knowledge such that the set of observables includes embeddings of elementary functions that increase the accuracy of the approximation. This method is a step further into the inclusion of information regarding the differential equation for the approximation of the linear predictors. Further research will probably show how to include the knowledge from a mass and energy balance modeling process or empirical knowledge regarding the dynamics to further improve the approximation.

Finally, this section presents linear and expanded approximations of interconnected dynamical systems for control. From these approximations, it is possible to set-up linear MPC algorithms that rely on the knowledge of the state and the partial knowledge of the interconnection input.

## Chapter 4

# Approximating the ROA via Koopman

This section describes our main results for the characterization of the asymptotically stable points region of attraction in the multi-stability scenario with the Koopman operator. In order to achieve this goal in a data-driven approach, we show how our method can find the fixed points of the system, give their stability and finally, analyze the spectrum of the operator to get the ROA. There are three assumptions that must hold for the approximation of the Koopman operator and the application of the algorithm. First, the EDMD algorithm has enough trajectories from the dynamical system to have a proper estimate of the operator. Second, Assumptions 1-3 from Section 3.2 hold. Third, the system is ergodic [Eisner et al., 2015] or in terms of polynomial chaos, it has finite variance [Marelli and Sudret, 2018]. For our particular case, this means that every initial condition in the state space converges to a specific asymptotically stable fixed point.

## 4.1 Fixed Points

The first step of the analysis consists of locating the fixed points to asses their stability. Recall from Definition 3.1.1 in Section 3.1 that a fixed point of the discrete time nonlinear system (3.1) is an invariant

subspace that has the property that whenever the state starts in it, it will remain in it for all future time.

Consider the forward and backward state evolution map of the system given by the Koopman operator in (3.54) and (3.55). If the evolution of these mappings is invariant then the system is at a hyperbolic fixed point. The solution of the minimization of the squared norm of the difference between the state and its evolution map from different initial conditions in the state space gives the location of the fixed points [Garcia-Tenorio et al., 2019].

**Lemma 4.1.1** (Fixed points). Let  $(\mathcal{M}; T(x); k)$  be a dynamical system that accepts an operator representation  $(\mathcal{F}_d; U_d; k)$  based on EDMD, and  $x^*$  is a fixed point of T(x). Then

$$x^{*} = \underset{x}{\operatorname{argmin}} \left\| \Psi_{B}^{-1} \left( B^{\top} \Xi M^{k} \Phi(x_{0}) \right) - x \right\|_{2}^{2}.$$
(4.1)

*Proof.* Let  $T^k(x) = \Psi_B^{-1}(B^{\top} \Xi M^k \Phi(x_0))$  the flow map of  $(\mathcal{M}; T(x); k)$ , assume k = K finite, and define  $b(x) = \Psi_B^{-1}(B^{\top} \Xi M \Phi(x))$ , using Definition 3.1.1 and b(x), the least squares problem

$$J(x) = \frac{1}{2} \|b(x) - x\|^2, \qquad (4.2)$$

with solution

$$x^* = \underset{x}{\operatorname{argmin}} \left( J(x) \right), \tag{4.3}$$

gives the location of the fixed points.

**Remark 5.** This lemma is only accountable for the location of the fixed points in the state space; it does not give information about their stability.

**Remark 6.** Equation (3.54) is a nonlinear, discrete-time evolution mapping, and under the definition of fixed point, i.e.,  $\overline{T}^k(x^*) = x^*$  it is possible to get the approximation of the hyperbolic fixed points by solving

$$\bar{x}^* \subseteq \{x \in \mathbb{R}^n \colon \bar{T}^k(x) - x = 0\}.$$

$$(4.4)$$

In practice, this procedure poses two problems, first, as the nonlinear mapping (3.54) comes from a set of observables, the dimension and complexity of these functions affect the possibility of getting a solution in polynomial type. In high order polynomial expansions, even if it possible to find a solution, the computational time is high. Second, when there is a solution, it is not clear which elements of the subset represent an actual fixed point of the system, i.e., not all the points of the subset are fixed points of the system, while the converse is true. Considering the data driven case, the fixed points of the system are not known a priori and therefore using this definition to approximate the fixed points is not feasible.

With the location of the fixed points, it is necessary to establish their stability.

#### 4.1.1 Fixed Points Example

For illustration purposes, consider the ongoing example of the Duffing equation with damping and two basins of attraction. Recall that this system has a saddle at the origin and two asymptotically stable equilibrium points.

For the approximation of the equilibrium points in Lemma 4.1.1 the initial conditions from the test set are not enough to identify all of the fixed points from a multi-start algorithm. Consequently, a new set of points in the original portion of the state space is generated to start the optimization. Additionally, recall that the evolution of the system based on eigenfunctions according to (3.54) has an additional computational step than the system evolution based on observables in (3.72), which is the spectral decomposition of the Koopman matrix. Therefore, the use of the evolution based on observables improves the accuracy of the location of the fixed points. Table 4.1 shows the numerical result and fig. 4.1 depicts the location of the initial conditions and the result of applying the method.

	Theoretical	Algor	ithmic
$x_A^*$	(-1,0)	(-1.002	2,0.001)
$x_B^*$	$(0,\!0)$	(0.003)	,0.001)
$x_C^*$	(1,0)	(1.00)	03,0)
2 [			
1.5	•	•	$\begin{array}{c} \text{ICs} \\ \text{Sys} \ x^* \end{array}$
1 -		•	Koop $x^*$
0.5		•	•
0 -	•	•	•
-0.5			
-1 -			•
-1.5			
-2			
-2	-1	0	1 2

Table 4.1: Duffing equation fixed points location.

Figure 4.1: Initial conditions and location of the fixed points for the Duffing equation.

## 4.2 Stability of Fixed Points

The traditional way of establishing the local stability of a hyperbolic fixed point is through the analysis of the Jacobian matrix of the system (3.1) evaluated at the fixed point. Our proposed approach is to analyze the state evolution map from the approximation of the Koopman operator in the same way.

**Definition 4.2.1** (Stability Properties). Consider the discrete-time nonlinear system (3.1) that has a representation via the state evolution map as stated in (3.26) from the EDMD algorithm. Define  $b(x) = \Psi_B^{-1}(B^{\top} \Xi M^k \Phi(x))$ as the state evolution map of the system, then the system accepts a local linearization such as

$$\begin{aligned} x(k+1) &= \left[ \frac{\partial T(x)}{\partial x_1} \cdots \frac{\partial T(x)}{\partial x_n} \right] \Big|_{x^*} x(k) \\ &= \left[ \frac{\partial b(x)}{\partial x_1} \cdots \frac{\partial b(x)}{\partial x_n} \right] \Big|_{x^*} x(k) \\ &= Hx(k). \end{aligned}$$
(4.5)

The local stability of a fixed point  $x^*$  with respect to the eigenvalues  $\{\mu_i\}_{i=1}^n$  from the spectral decomposition of H are given by,

- if  $|\mu_i| < 1$  for all i = 1, ..., n then  $x^*$  is asymptotically stable.
- if  $|\mu_i| > 1$  for all i = 1, ..., n then  $x^*$  is unstable.
- if  $|\mu_i| < 1$  for some  $i = 1, ..., n_s$  and  $|\mu_i| > 1$  for some  $i = n_s + 1, ..., n$ , then  $x^*$  is also unstable but has modal components that converge to it, making it a saddle point.

**Remark 7.** Note that the inequalities are strictly greater-than or lessthan, this is in accordance with the hyperbolicity assumption.

With the information about the location and stability of fixed points, the main result of this paper is formulated, i.e., the approximation of the boundary of the ROA via eigenfunctions of the Koopman operator.

### 4.2.1 Stability Example

Continuing the Duffing equation example, recall that the fixed pints of the system are: asymptotically stable for  $x_1^*4$ , unstable saddle for the origin  $x_2^*$ , and asymptotically stable for  $x_3^*$ . The first step to get the stability of fixed points is to calculate the Jacobian matrix of (3.54). However, given that (3.72) is more accurate, the Jacobian matrix according to the evolution of observables is

$$J = \begin{bmatrix} -0.0147x_1^2 + 1.32e - 4x_1 + 1.0 & 0.0967\\ -0.292x_1^2 + 3.81e - 3x_1 + 0.0976 & 0.9407 \end{bmatrix}$$
(4.6)

Notice that this matrix is only in function of  $x_1$  because all of the fixed points are in  $x_2 = 0$ . Next, the evaluation of the Jacobian matrix for the three different values of  $x_1$  and their corresponding eigenvalues are:

$$J|_{x_A^*} = \begin{bmatrix} 0.9901 & 0.0967\\ -0.1986 & 0.9407 \end{bmatrix} \qquad \lambda_A = \begin{cases} 0.9654 + 0.1364i\\ 0.9654 - 0.1364i \end{cases}$$
(4.7)

$$J|_{x_B^*} = \begin{bmatrix} 1.0049 & 0.0967\\ 0.0976 & 0.9407 \end{bmatrix} \qquad \lambda_B = \begin{cases} 0.8705 + 0.0000i\\ 1.0751 + 0.0000i \end{cases}$$
(4.8)

$$J|_{x_C^*} = \begin{bmatrix} 0.9901 & 0.0967\\ -0.1986 & 0.9407 \end{bmatrix} \qquad \lambda_C = \begin{cases} 0.9655 + 0.1336i\\ 0.9655 - 0.1336i \end{cases}$$
(4.9)

Table 4.2 shows the norm of the respective eigenvalues for each case, and the corresponding stability according to the classical linear theory analysis.

Table 4.2: Fixed point stability for the Duffing equation.

	$ \lambda_1 $	$ \lambda_2 $	Stability
$x_A^*$	0.97	0.97	AS
$x_B^*$	0.88	1.08	Saddle
$x_C^*$	0.97	0.97	AS

## 4.3 Approximation of the ROA boundary

This section describes the results for the approximation of the ROA for asymptotically stable fixed points in a data-driven approach using the Koopman operator. The boundary of the ROA is formed by the stable manifold of the saddle fixed points in it (see Figure 3.2). A set of eigenfunctions built from the approximation of the Koopman operator captures the stable manifold of saddle fixed points. Therefore, a set of eigenfunctions captures the ROA of asymptotically stable fixed points.

The set of eigenfunctions that captures the ROA has the characteristic that their associated eigenvalue is unitary. An eigenfunction with an associated eigenvalue equal to one is invariant along the trajectories of the system (if  $\mu_i = 1$  in (3.26),  $[U^k \phi_i](x) = \phi_i(x)$ ). Therefore, the set of eigenfunctions that approximate the ROA are a set of eigenfunctions that are invariant along the trajectories of the system. With this information and the fact that a level set is the curve of a function that is invariant, it follows that the level sets of eigenfunctions with an associated eigenvalue equal to one characterize the ROA of asymptotically stable fixed points. The following shows how these premises hold.

**Theorem 4.3.1.** Let  $(\mathcal{M}; T(x); k)$  be a dynamical system that accepts an operator representation  $(\mathcal{F}_d; U_d; k)$  based on EDMD. If there exists an eigenfunction with associated eigenvalue equal to one defined as  $\phi_+(x)$ , then the stable manifold of a saddle point  $W^s(\hat{x}^*)$  is the level set

$$\Gamma(\phi_{+}(x)) \subseteq \{ x \in \mathbb{R}^{n} \colon \phi_{+}(x) = \phi_{+}(\hat{x}^{*}) \}.$$
(4.10)

Proof. Let  $\phi(x) \in \mathcal{F}_d$  be an eigenfunction that satisfies the condition  $[U_d\phi](x) = \phi(T(x))$ . Let  $\bar{\phi}(x) = \phi_1^{k_1}(x)\phi_2^{k_2}(x)$  be a function constructed upon the product of two arbitrary eigenfunctions with associated eigenvalues  $\mu_1$  and  $\mu_2$  respectively, for constants  $k_1, k_2 \in \mathbb{C}$ . According to (3.25), the relationship between these two eigenfunctions is:

$$\left[U_{d}\bar{\phi}\right](x) = \left[U_{d}(\phi_{1}^{k_{1}}\phi_{2}^{k_{2}})\right](x)$$
(4.11)

$$= ([U_d \phi_1](x))^{k_1} ([U_d \phi_2](x))^{k_2}$$
(4.12)

$$=\mu_1^{k_1}\phi_1^{k_1}(x)\mu_2^{k_2}\phi_2^{k_2}(x).$$
(4.13)

This implies that  $\bar{\phi}(x)$  is an eigenfunction of the Koopman operator U with associated eigenvalue  $\bar{\mu} = \mu_1^{k_1} \mu_2^{k_2}$ . It is possible to find the complex constants  $k_1$  and  $k_2$  for eigenvalues  $\mu_1$  and  $\mu_2$  such that  $\bar{\mu} = 1$ . Meaning that  $\bar{\phi}(x) = \phi_+(x)$ . Evaluating the eigenfunction with associated eigenvalue equal to one on the saddle point  $\phi_+(\hat{x}^*)$  gives the constant value along the trajectories that converge to it. These trajectories are the stable manifold of this saddle point. From (3.15) this manifold is

$$W^{s}(\hat{x}^{*}) = \{ x \in \mathbb{R}^{n} \colon \lim_{k \to \infty} T^{k}(x) = \hat{x}^{*} \},$$
(4.14)

evaluating this manifold with an arbitrary eigenfunction  $\phi(x)$ , and using the condition of the Koopman operator in (3.25) gives

$$\phi(W^s(\hat{x}^*)) = \{ x \in \mathbb{R}^n \colon \lim_{k \to \infty} [U_d^k \phi](x) = \phi(\hat{x}^*) \}.$$
(4.15)

Recall from (3.26) that the eigenvalue determines the dynamics of the associated eigenfunction, meaning that for this arbitrary eigenfunction, the associated eigenvalue must be unitary for the equality to hold, giving an eigenfunction that is invariant under  $U_d$ , this means that the eigenfunction is constant along the trajectories of the system, therefore,

$$\phi(W^s(\hat{x}^*)) = \{ x \in \mathbb{R}^n \colon \phi(x) = \phi(\hat{x}^*) \}.$$
(4.16)

Therefore, by the definition of level set we obtain,

$$\Gamma \phi_+(x) = \{ x \in \mathbb{R}^n \colon \phi_+(x) = \phi_+(\hat{x}^*) \}.$$
(4.17)

**Remark 8.** The level set  $\Gamma$  in (4.10) is a subset and does not hold with equality because the trivial eigenfunction, i.e.,  $\phi_{\mu=1} = 1$  is an eigenfunction with unitary associated eigenvalue where (4.17) is the whole state space  $\mathbb{R}^n$ .

The level set (4.10) gives the ROA boundary, and therefore, it is possible to find a classification rule of the different initial conditions in the state space that converge to any of the asymptotically stable equilibrium points. A geometrical approach to the classification problem is identifying the hyperplane that represents the boundary. This is a common approach when using the backwards integration of the dynamical system to find the set of points that compose the hyperplane. Although this is a possible approach, classifying an initial condition based on the geometry of the manifold is often difficult, especially when the hyperplane describing the boundary is not a function, or the system is of higher than three dimensions.

The approximation of the discrete-time Koopman operator from the EDMD algorithm using orthogonal polynomials as observables gives a set of eigenfunctions whose eigenvalue is unitary or close to unitary. Considering the difficulties of classifying the initial conditions based on the geometry of the stable manifold of saddles points, we propose two different methods based on this set of eigenfunctions to determine the convergence of an arbitrary initial condition.

**Definition 4.3.1** (Saddle classifier). Let  $\{\phi_+^i : \mu_i = 1\}_{i=1}^{\infty}$  be a set nontrivial eigenfunctions of a dynamical system  $(\mathcal{M}; T(x); k)$  that accepts an approximation of the discrete-time Koopman operator  $(\mathcal{F}_d; U_d; k)$  based on EDMD, and let  $\hat{x}^*$  be a type-one saddle point in the boundary of the ROA separating the asymptotically stable points  $x_A^*$  and  $x_B^*$ . The convergence of an arbitrary initial condition  $x_0$  to either asymptotically stable equilibrium points  $x_A^*$  or  $x_B^*$  is

$$x_A \subseteq \{x_0 \in \mathbb{R}^n \colon \Re[\phi_+(x_0) \ge \Re[\phi_+(\hat{x}^*)]\}$$
  

$$x_B \subseteq \{x_0 \in \mathbb{R}^n \colon \Re[\phi_+(x_0) < \Re[\phi_+(\hat{x}^*)]\}$$
(4.18)

**Remark 9.** The points that converge to either asymptotically stable fixed point is a subset and not an equality because not all eigenfunctions with associated eigenvalue equal to one capture this phenomena. The particular characteristics that an eigenfunction must have to satisfy this criteria is still open.

This criterion is closely related to the geometric one, especially because there is the need for a clear separatrix given by the real part eigenfunction evaluation in the boundary saddle point. This implies that for every pair of asymptotically stable fixed points to be classified there must be a an eigenfunction that separates the two. When the dimension of the problem starts to grow, especially in the amount of asymptotically stable fixed points, the rule given by Definition 4.3.1 becomes a combination of the evaluations of the initial condition, compared to the evaluation of the different saddle points in the boundaries.

Because of the invariant nature of the eigenfunction with associated eigenvalue equal to one along the trajectories of the system, we propose a second criterion to classify the initial conditions that is independent on the aforementioned boundary of the ROA given by the level set. Theorem 3.2.1 states that the stable manifold of *type-one* saddle points can be captured with an eigenfunction with associated eigenvalue equal to one. Extrapolating this concept to the asymptotically stable fixed points, it is possible to find a unitary eigenfunction that is invariant along its ROA.

**Definition 4.3.2** (AS classifier). Let  $\{\phi_+^i : \mu_i = 1\}_{i=1}^{\infty}$  be a set nontrivial eigenfunctions of a dynamical system  $(\mathcal{M}; T(x); k)$  that accepts an approximation of the discrete-time Koopman operator  $(\mathcal{F}_d; U_d; k)$  based on EDMD, let  $\{x_I^*\}_{I=1}^L$  be a set of L asymptotically stable fixed points in the space state, and define  $C_l(\phi): \mathbb{C} \to \mathbb{C}$  as an arbitrary function of the

observation. The set of points  $x_I$  that converge from an arbitrary initial condition  $x_0$  to the fixed points  $x_I^*$  is

$$x_{I} \subseteq \{x \in \mathbb{R}^{n} \colon C_{l}(\phi_{+}(x_{0})) = C_{l}(\phi_{+}(x_{I}^{*}))\}$$

$$(4.19)$$

**Remark 10.** The ideal case is to have a real valued constant eigenfunction where  $C_l(\phi(x)) = \phi(x)$ . Although this is theoretically possible, i.e., to find an eigenfunction invariant along the ROA of the asymptotically stable fixed point. The necessary conditions or the characteristics for this to happen is still an open question. For  $C_l(\phi(x)) = \angle(\psi(x))$ ,  $C_l(\phi(x)) = |\phi(x)|$  and  $C_l(\phi(x)) = R(\phi(x))$  there is a chance of getting an accurate classification based on the natural eigenfunction from the EDMD approximation of by constructing the eigenfunction. Nevertheless, the current method to find an eigenfunction that accurately classifies according to either of these functions of  $\phi$  is by checking the necessary conditions for the eigenfunction to be accurate under each rule. As well as the constant case, the conditions that guarantee an accurate behavior is still an open question.

**Remark 11.** Although this eigenfunction is suitable for the classification of all the approximated state space into its respective point of convergence, note that the equality is replaced by a subset because not all eigenfunctions with an associated eigenvalue equal to one satisfy this criterion. To the best of the author knowledge, the criteria to make it an equality are still unknown.

Section 5.2 exemplifies the different results of performing the analysis with the different classification schemes. In general, the available eigenfunctions with a unitary eigenvalue that come directly from the EDMD algorithm or a constructed one from (4.13) give an accurate result by using Definition 4.3.1. For the classification based on the asymptotically stable equilibrium points from Definition 4.3.2, finding an eigenfunction that satisfies this criterion is not trivial.

In summary, the approximation of the Koopman operator serves to find the unitary eigenfunction that characterizes the ROA of asymptotically stable ones.
### 4.3.1 ROA Example

For illustration purposes, consider the ongoing example of the Duffing equation with damping and two asymptotically stable equilibrium points. This system satisfies Assumptions 1-3 from section 3.2 because there is only one fixed point in the stability boundary of the two asymptotically stable ones, which is the saddle point at the origin. Given that it is unique, it cannot violate the transversality condition.

Theorem 3.2.1 states that the stable manifold of the saddle point at the origin is the boundary os the region of attraction of the two asymptotically stable points. Furthermore, theorem 4.3.1 states that a nontrivial eigenfunction with a unitary associated eigenvalue captures this stable manifold. Consider the level plot of fig. 4.2 that shows a unitary eigenfunction of the discrete-time approximation of the Koopman operator. Contrary to the previous cases, this eigenfunction does not come from the approximation using only one trajectory for each of the asymptotically stable points. Instead, the set of eigenfunctions comes from a full basis, i.e.,  $q = \infty$  for a value of p = 4 that gives a set of 25 eigenfunctions. The training and testing sets come from 400 uniformly distributed initial conditions over  $x_1, x_2 \in [-2 \ 2]$  with a  $\Delta t = 0.1$  and contrary to the previous cases, there are only 11 points per trajectory, not a full numeric integration until convergence. Again, the division between the training and testing set is 50 - 50.

Notice that the amount of points for the accurate approximation of eigenfunctions is five times greater than the necessary points for the approximation of the dynamics. Furthermore, the dimension and order of the basis is greater, and given that for this case there are more initial conditions and less points per trajectory, as a result, the available points for the approximation cover uniformly the portion of the space state under evaluation. The reason for this different approach is that the method for building accurate unitary eigenfunctions is still an open question.

Evaluating the unitary eigenfunction at the identified saddle point near the origin gives the constant value of the level set that describes the stable manifold. Identifying this line and classifying according to it is not the best numerical approach. As a consequence, notice that the ROA of each of the points corresponds to a subset of the space state



Figure 4.2: Level plot of the unitary eigenfunction of the discrete-time approximation of the Koopman operator for the Duffing equation problem.

that is either above or below the stable manifold when evaluated by the unitary eigenfunction. Therefore, the classification rule does not rely on the interpolation of points in the manifold to check if a point in the state space is above or below the line. The classification rule is according to the evaluation of an individual point with the unitary eigenfunction and the comparison of the real part of this value to the constant value of the level set as in definition 4.3.1.

Figure 4.3 depicts the approximation of the stable manifold of the saddle point, and the correct classification of the test set of initial conditions, where the application of the classification criterion gives an error of 2.5%. In this figure it is clear that the stable manifold cannot be represented as a function and therefore, its geometry in not a convenient classification rule. If the system under study has more than two state variables, the problem of classifying according to the geometry of the stable hyperplane is worse.



Figure 4.3: Stable manifold approximation and initial condition classification based on the evaluation of the unitary eigenfunction in the saddle point.

# 4.4 Algorithm

The approach presented in section 3.4 and ??-4.3 for obtaining the attraction regions of asymptotically stable fixed points is summarized in Algorithm 1, for which the following assumptions hold.

**Assumption 4.** The system under consideration has multiple hyperbolic fixed points.

Assumption 5. At least one of the fixed points is asymptotically stable.

**Assumption 6.** There is enough snapshot data either from measurements of the real system or a numerical simulation for constructing an approximation of the Koopman operator. Algorithm 1: Approximation of the ROA boundary via the eigenfunctions of the Koopman operator. A data-driven calculation that approximates the fixed points and their stability and the eigenfunction of the Koopman operator, generates the unitary eigenfunction, and analyzes the saddle points in the ROA to get its boundary.

1 For a dynamical system that satisfies Assumptions 4-6; **Data:**  $\{x_i, y_i\}_{i=1}^N$ **Result:**  $\partial R_A(x_s^*)$ **2** initialization; **3**  $[\Phi(x), M] \leftarrow \text{EDMD}(\{x_i, y_i\}_{i=1}^N, q, p);$ 4  $b(x) \leftarrow \Psi_B^{-1}(B^\top \Xi M^k \Phi(x));$ 5  $x^* \leftarrow \operatorname{argmin}_x \|b(x) - x\|_2^2$ ; 6  $E_q \leftarrow \operatorname{size}(x^*);$ 7 for  $i \leftarrow 1$  to  $E_q$  do 8  $\left| H_i \leftarrow \left[ \frac{\partial b(x)}{\partial x_1} \cdots \frac{\partial b(x)}{\partial x_n} \right] \right|_{x_i^*};$  $\{\mu_{i,j}\}_{j=1}^n \longleftarrow \operatorname{eig}(H_i);$ 9 if  $|\mu_{i,j}| < 1$  for all  $j = 1, \ldots, n$  then  $\mathbf{10}$  $x_i^* \leftarrow x_s^*$ 11 else if  $|\mu_{i,j}| < 1$  for some  $j = 1, \ldots, n$  then 12 $x_i^* \leftarrow \hat{x}^*$ 13 else  $\mathbf{14}$  $x^* \leftarrow x^*_u$ 15end 16 17 end 18  $\hat{E}_q \longleftarrow \operatorname{size}(\hat{x}^*);$ 19  $E_s \leftarrow \operatorname{size}(x_s^*);$ **20**  $\phi_+(x) \longleftarrow \{\phi_1^{k_1}(x)\phi_2^{k_2}(x) \colon \mu_1^{k_1}\mu_2^{k_2} = 1\};$ 21 for  $i \leftarrow 1$  to  $E_s$  do for  $j \leftarrow 1$  to  $\hat{E}_q$  do  $\mathbf{22}$  $W_j^s(\hat{x}_j^*) \longleftarrow \{x \in \mathbb{R}^n \colon \phi_+(x) = \phi_+(\hat{x}_j^*)\};$ 23 end  $\mathbf{24}$  $\partial R_A(x_i^*) = \cup W^s(\hat{x}_i^*)$  $\mathbf{25}$ 26 end

# 4.5 Summary

This Section presents the theoretical and algorithmic contributions that allow the approximation of regions of attraction for asymptotically stable fixed points. Regarding the analysis, it is restricted to hyperbolic systems because of the impossibility to determine the location of limit cycles based only on snapshot data of the system. The main component of the method is an accurate calculation of the systems' EDMD, from this decomposition it is possible to locate the hyperbolic fixed points of the system and give their stability along with their *type*. Additionally, an accurate decomposition with a suitable set of observables gives an accurate approximation of the Koopman operator. Which in turn, provides the basis for producing a non-trivial eigenfunction with eigenvalue equal to one. By combining these two results from the EDMD, it is possible to analyze the *type-one* fixed points in order to get an approximation of the ROA for the asymptotically stable fixed points.

## 4.6 Additional Examples

This thesis has proposed a method to find the ROA of asymptotically stable fixed points based on the analysis of the eigenfunctions of the approximation of the Koopman operator. To test the reliability of the results, we first apply the method to a model of competitive exclusion with two state variables to graphically show the effect of the eigenfunction with unitary eigenvalue. Then, a five-state variable mass action kinetics (MAK) model that allows displaying the effectiveness of the algorithm for higher dimensional systems. The Lotka-Volterra and the MAK systems are suitable for the analysis because of their geometrical properties. They are nonnegative compartmental systems that, depending on the parameterization, have hyperbolic fixed points that satisfy the Hartman-Grobman theorem [Chellaboina et al., 2009]. The hyperbolicity of the fixed points implies that their stable and unstable manifolds intersect transversely at the saddle points. Therefore, Assumptions 1-3 are satisfied in this kind of systems under the right parameterization.

The numerical integration of the systems from a random set of initial

conditions gives the dataset of trajectories for the algorithm. A subset is used to calculate the Koopman operator's approximation via the EDMD algorithm, and another subset is used for testing the algorithm's accuracy according to the empirical error (3.56), where this error serves to determine the best p - q parameters from a sweep over the different values.

### 4.6.1 Lotka-Volterra Model

Consider a Lotka-Volterra model where two species compete for the same resource described by the following network:

where  $s_1$  and  $s_2$  are the competing species,  $r_1$  and  $r_3$  describe the growth rates of the species respectively,  $r_2$  and  $r_4$  are the constants that describe competition between members of the same species, and the constants  $r_5$ and  $r_6$  describe the competition between species. The values of these constants predicts the potential outcome of the competition, depending on the parameterization there can be a co-existence or an exclusion of one of the species against the other. The case of interest is the exclusion one, and the objective is to find the set of initial conditions within the state space that lead the model to one of the stable points where one species completely takes over the other.

The ordinary differential equation that describes the dynamics of the model comes from the matrices

$$\alpha = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \\ 0 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad \qquad \alpha' = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad (4.21)$$

The differential equation that describes the network (4.20) according to the matrices (4.21) and (3.121) is

$$\dot{x}_1 = r_1 x_1 - r_2 x_1^2 - r_5 x_1 x_2 \tag{4.22}$$

$$\dot{x}_2 = r_3 x_2 - r_4 x_2^2 - r_6 x_1 x_2, \tag{4.23}$$

where the parameterization for the species to have two particular asymptotically stable fixed points is  $r = [2 \ 1 \ 2 \ 1 \ 3 \ 3]^{\top}$ . With this choice, the system has four fixed points; an unstable point at the origin defined as  $x_A^*$ , a saddle point at (0.5, 0.5) defined as  $x_D^*$  and two asymptotically stable at (0, 2) and (2, 0) defined as  $x_B^*$  and  $x_C^*$  respectively. The geometry of this problem provides a simple representation of the stable manifold of the saddle point which is the line  $x_1 = x_2$ , therefore, giving a closed formulation for the comparison with the stable manifold provided by the operator by the construction of the eigenfunction with an associated eigenvalue equal to one.

The integration of the system from 200 uniformly distributed random initial conditions over  $x_1, x_2 \in [0 \ 2]$  give the datasets for approximating the operator via the EDMD algorithm and the dataset to test the accuracy of the algorithm; both sets, the training and test ones are 50% of the trajectories from the original set. We apply the EDMD algorithm for the training set with a Laguerre polynomial basis where a sweep over the p - q values gives the best performance when q = 1.1 and p = 3 for the truncation scheme. This selection produces a polynomial basis with 13 elements of an order less than 3 and a Koopman operator of order 13. Figure 4.4 shows the retained indices after implementing the truncation scheme.

The identification of the fixed points via Lemma (4.1.1) gives an absolute error of 0.15%, and the linearization of the state evolution map from (3.54) evaluated at the fixed points accurately provides their stability according to the norm of the eigenvalues  $\lambda$ ; these are the eigenvalues of the linearization of the state evolution map (where  $\Psi_B^{-1}(x) =$  $[1-x_1, 1-x_2]^{\top}$  for this particular case) evaluated at the identified fixed points. Table 4.3 summarizes these results.

Figure 4.5 depicts the comparison between the theoretical trajectories given by the integration of the differential equations and the state evolution map (3.54) from the same initial conditions. It also depicts the



Figure 4.4: Retained indices for the approximation of the Koopman operator for the Lotka-Volterra model with a choice of q = 1.1 and p = 3.

			-	,	
	Theoretical	Algorithmic	$ \lambda_1 $	$ \lambda_2 $	Stability
$x_A^*$	(0,0)	(-0.006, -0.006)	1.21	1.22	Unstable
$x_B^*$	(0,2)	(0,2)	0.66	0.82	AS
$x_C^*$	(2,0)	(2,0)	0.82	0.66	AS
$x_D^*$	(0.5, 0.5)	(0.5, 0.5)	0.81	1.10	Saddle

Table 4.3: Lotka-Volterra equation fixed points, location and stability.

comparison between the theoretical boundary of the attraction regions  $x_1 = x_2$  and the boundary given by the level set of the constructed eigenfunction with unitary associated eigenvalue from (4.10). The error in the classification of the initial conditions is of 2%, while the mean absolute error between the boundaries of the regions of attraction is 3%.

Figure 4.6 depicts four eigenfunctions of the Koopman operator, where the problem of the determination of the invariant unitary eigenfunction is apparent. The approximation methods such as the EDMD can yield  $\phi_{\mu=1} = 1$ , which is a trivial eigenfunction that does not provide information about the system. Therefore, it is necessary to get a nontrivial invariant eigenfunction. Recall from (4.13) the relationship that the set of eigenfunctions satisfy, where  $\phi_+(x) = \phi_1^{k_1}(x)\phi_2^{k_2}(x)$  is an eigenfunction of the Koopman operator with associated eigenvalue  $\mu_+ = \mu_1^{k_1}\mu_2^{k_2}$  for a set of complex constants  $k_1$  and  $k_2$ . Therefore, setting  $\mu_1^{k_1}\mu_2^{k_2} = 1$  and computing a solution for  $k_1$  and  $k_2$  gives the desired invariant eigenfunction.

Figure 4.6 also depicts the resulting invariant eigenfunction that cap-



Figure 4.5: Trajectories and boundary of the asymptotically stable points of: a) The system differential equation, and b) The Koopman operator and the eigenfunction with unitary associated eigenvalue.

tures the stable manifold of the saddle point, and the two eigenfunctions with real-valued eigenvalues close to one used for the construction. The eigenvalues of these eigenfunctions are  $\mu_1 = 1.07$  and  $\mu_2 = 0.83$ . It is still not clear how to properly select these eigenfunctions, and although the algorithm gives accurate results when the selected eigenfunctions for the construction of the invariant eigenfunction do not have real-valued eigenvalues, i.e., the complex part of the eigenvalue is different from zero, not having real-valued eigenfunctions for the construction does hinder the accuracy.

### 4.6.2 Mass Action Kinetics

Consider the simple network of an auto-catalytic replicator on a continuous flow stirred tank reactor described by the network

$$s_{1}+2s_{3} \xrightarrow{r_{1}} s_{2}+3s_{3}$$

$$s_{2}+2s_{4} \xrightarrow{r_{2}} 3s_{4}$$

$$s_{3} \xrightarrow{r_{3}} s_{5}$$

$$s_{4} \xrightarrow{r_{4}} s_{5}$$

$$(4.24)$$



Figure 4.6: Eigenfunctions of the Koopman operator. a) Trivial eigenfunction with  $\mu = 1$ , b) Constructed eigenfunction  $\phi_+$ , c) First eigenfunction for constructing  $\phi_+$ , and d) Second eigenfunction for constructing  $\phi_+$ .

where  $s_1$  is the resource the species  $s_3$  consumes to replicate and produce substrate  $s_2$ , which is consumed by species  $s_4$  to replicate, and  $s_5$  is the dead species from both groups.  $r_1 > r_3$ , and  $r_2 > r_4$  are the pairs of replication rate constants and the species death constants. The dynamics of the network (4.24) come from matrices

$$\alpha = \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \alpha' = \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (4.25)$$

The differential equation that describes the network (4.24) according to the matrices (4.25) and (3.121), and considering the material exchange

with the environment is

$$\dot{x}_{1} = -r_{1}x_{1}x_{3}^{2} + d - dx_{1}$$

$$\dot{x}_{2} = +r_{1}x_{1}x_{3}^{2} - r_{2}x_{2}x_{4}^{2} - dx_{2}$$

$$\dot{x}_{3} = +r_{1}x_{1}x_{3}^{2} - r_{3}x_{3} - dx_{3}$$

$$\dot{x}_{4} = +r_{2}x_{2}x_{4}^{2} - r_{4}x_{4} - dx_{4}$$

$$\dot{x}_{5} = +r_{3}x_{3} + r_{4}x_{4} - dx_{5},$$
(4.26)

where d is the in/out-flow (dilution rate) of the system, and the substrate  $s_1$  is the only component in the input flow. The value for the reaction rates vector  $k = [7 5 0.3 0.05]^{\top}$ , yielding five fixed points: three asymptotically stable, i.e., the working point defined as  $x_A^*$ , a point where species  $s_3$  thrives and species  $s_4$  washes-out defined as  $x_C^*$ , and a washout point where the concentration of both species is zero defined as  $x_E^*$ , and two saddle points defined as  $x_B^*$  and  $x_D^*$ . The objective is to find the attraction region of the working point using the eigenfunctions with unitary associated eigenvalue. For this particular case, as the system is not a two-dimensional one and the geometry of the stable manifold is not evident. Therefore, the criteria for classifying the test set of initial conditions is not trivial.

A set of 360 random initial conditions is generated to integrate the system with a  $\Delta t = 0.1$  and get the snapshot data for the approximation of the Koopman operator via the EDMD algorithm. From this set of snapshots, 50% is used to approximate the operator, and the other 50% to test the accuracy of the state evolution map, the eigenfunctions, and the results of the classification scheme based on the unitary eigenfunction. A sweep over the p-q values gives the best performance for the truncation scheme when q = 0.8 and p = 4 which leads to an order of 163 by 163 elements of the approximation of the Koopman operator, and thus, a set of 163 eigenfunctions, eigenvalues and modes. From the set of eigenfunctions, the two eigenfunctions with real eigenvalue closest to one are  $\mu_1 = 1.00008$ , and  $\mu_2 = 0.99983$  are selected to analyze the system. The eigenvalue associated with these eigenfunctions is close enough to one to provide an accurate analysis.

The identification of the fixed points via Lemma 4.1.1 gives an absolute error of 0.15%, and the linearization of the state evolution map from (3.54) evaluated at the fixed points provides an accurate description

of their stability according to the norm of the eigenvalues  $\lambda$ . Table 4.4 shows the results from the theoretical and algorithmic location of the fixed points, Table 4.5 shows the results of taking the norm of the eigenvalues  $\lambda$  of the linearization of the state evolution map (3.54), evaluated at each of the fixed points.

	Theoretical	Algorithmic	Error $\%$
$x_A^*$	(0.23, 0.09, 0.30, 0.54, 0.59)	(0.23, 0.09, 0.30, 0.54, 0.59)	0.08
$x_B^*$	(0.21, 0.67, 0.30, 0.07, 0.47)	(0.23, 0.62, 0.30, 0.11, 0.49)	0.00
$x_C^*$	(0.23, 0.76, 0.30, 0.00, 0.46)	(0.23, 0.76, 0.30, 0.00, 0.46)	0.00
$x_D^*$	(0.76, 0.23, 0.09, 0.00, 0.14)	(0.70, 0.3, 0.11, 0.00, 0.17)	0.54
$x_E^*$	(1.00, 0.00, 0.00, 0.00, 0.00)	(1.00, 0.00, 0.00, 0.00, 0.00)	0.00

Table 4.4: Location of fixed points.

Table 4.5: Stability of fixed points.

					., .	I III	
	$ \lambda_1 $	$ \lambda_2 $	$ \lambda_3 $	$ \lambda_4 $	$ \lambda_5 $	Stability	
$x_A^*$	0.89	0.98	0.98	0.97	0.98	AS	Working Point
$x_B^*$	1.04	0.98	0.98	0.98	0.98	Saddle	
$x_C^*$	0.98	0.98	0.93	0.98	0.98	AS	$x_4$ Wash-out
$x_D^*$	1.05	0.96	0.98	0.98	0.97	Saddle	
$x_E^{*}$	0.90	0.98	0.98	0.97	0.97	AS	Wash-out

Figure 4.7 depicts the results of evaluating the initial conditions of the test set on the eigenfunctions with an associated eigenvalue equal to one. The evaluation of the selected eigenfunction with  $\mu = 1.00008$  (denoted  $\phi_{+1}$ ) at the initial conditions of the test set and the saddle fixed points  $x_B$  and  $x_D$  shows that the evaluation at the former gives an accurate classification of the initial conditions that converge to the asymptotically stable fixed point  $x_A$ . The evaluation of the selected eigenfunction with  $\mu = 0.99983$  (denoted  $\phi_{+2}$ ) at the initial conditions of the test set and the saddle fixed points  $x_B$  and  $x_D$  shows that the evaluation of the selected eigenfunction with  $\mu = 0.99983$  (denoted  $\phi_{+2}$ ) at the initial conditions of the test set and the saddle fixed points  $x_B$  and  $x_D$  shows that the evaluation at the latter gives an accurate classification of the initial conditions that converge to the asymptotically stable fixed point  $x_E$ . The combination of these two conditions specifies the criteria to classify the initial conditions that converge to the different asymptotically stable fixed points. Figure 4.8

shows the result of the classification criteria, where the correct classification surrounds the miss-classified initial conditions, with an error of 12%.



Figure 4.7: Eigenfunctions with unitary associated eigenvalue a) Trivial b)  $\mu > 1$  c)  $\mu < 1$ .

# 4.7 Discussion

Even if we had the tools for extracting all kinds of information regarding the system, there are limitations for the deduction of the eigenfunctions. The right choice of the polynomial basis based on the known structure of the differential equation, the optimal truncation scheme, the choice of eigenfunctions to construct the one with unitary associated eigenvalue are open questions for improving the algorithm and the analysis.



Figure 4.8: Classification of the initial conditions with respect to the evaluation of the eigenfunctions on the saddle points.

Also, there is the problem of the required data to train the Koopman operator properly. For the two exemplified cases, the amount of data is sufficient to produce an accurate nonlinear model of the systems. To consider the proposed method over traditional identification techniques would be necessary to develop experimental design techniques specifically for the calculation of the Koopman operator to reduce the necessary data.

Furthermore, the relation between the eigenfunctions and their associated eigenvalues is still open. There are other dynamical characteristics to be analyzed from this association, as the dynamic behavior of eigenfunctions whose associated eigenvalue matches the local spectra of the fixed points is a potential source of information of the underlying dynamical system.

A possible improvement would be for the choice or construction of

eigenfunctions. All eigenfunctions capture different information and properties (relevant or not, accurate or not), and the analysis is limited to the invariant sets of these eigenfunctions. An analysis based on the spectral and geometric characteristics of real-valued and complex-valued eigenfunctions could give more information about the system for control purposes.

# Chapter 5

# Anaerobic Digestion Process

The biological process through which multiple organisms break down organic matter in the absence of oxygen is called Anaerobic Digestion (AD). The benefits of the AD process include both the production of energy and the production of low sludge. What motivates the use of the AD process on an industrial scale is its capacity of degrading strong and resilient substrates, the low sludge production, and the possibility of making a profit out of the production of methane gas [Mailleret et al., 2003].

# 5.1 Problem Statement

The model of the anaerobic digestion system under study considers two reactions: acidogenesis and methanogenesis. The following reaction network describes the biological transformations:

$$a\xi_1 \xrightarrow{r_1(\xi)} c\xi_2 + \xi_3,$$
  
$$d\xi_2 \xrightarrow{r_2(\xi)} CH_4 + \xi_4.$$
(5.1)

In the first reaction, the acidogenic bacteria  $\xi_3$  consumes the organic substrate  $\xi_1$  for growth and produces volatile fatty acids  $\xi_2$ . In the second

reaction, the methanogenic bacteria  $\xi_4$  uses the volatile fatty acids as a substrate for growth and produces methane. It is necessary to maintain the balance between the acidogenesis and the methanogenesis states in the operation of the AD process to avoid acidification, which is the state of accumulation of volatile fatty acids in the reactor that produces the washout of methanogenic bacteria.

For an ideal continuously stirred tank reactor, the following differential equations describe the system dynamics of the reaction network (5.1):

$$\dot{\xi}_{1} = u(\xi_{\text{in}_{1}} - \xi_{1}) - ar_{1}(\xi), 
\dot{\xi}_{2} = u(\xi_{\text{in}_{2}} - \xi_{2}) + cr_{1}(\xi) - dr_{2}(\xi), 
\dot{\xi}_{3} = -u\xi_{3} + r_{1}(\xi), 
\dot{\xi}_{4} = -u\xi_{4} + r_{2}(\xi),$$
(5.2)

with a methane outflow rate

$$Q(\xi) = qr_2(\xi), \tag{5.3}$$

where  $\xi = [\xi_1 \ \xi_2 \ \xi_3 \ \xi_4]^T \in \mathbb{R}^4_+$  is the state vector,  $u \in \mathbb{R}_+$  is the dilution rate,  $\xi_{in_1}, \xi_{in_2} \in \mathbb{R}_+$  are the concentrations of organic substrate and volatile fatty acids in the influent,  $a, b, c \in \mathbb{R}_+$  are the stoichiometric coefficients, q > 0 is the yield for the methane production, and  $r_1(\xi), r_2(\xi)$  are the reaction rates defined as

$$r_1(\xi) = f_1(\xi_1)\xi_3,\tag{5.4}$$

$$r_2(\xi) = f_2(\xi_2)\xi_4,$$
(5.5)

where the growth functions,  $f_1(\xi_1)$  and  $f_2(\xi_2)$  are based on Monod and Haldane kinetics respectively as

$$f_1(\xi_1) = \mu_{m_1} \frac{\xi_1}{K_{s_1} + \xi_1},$$
  

$$f_2(\xi_2) = \mu_{m_2} \frac{\xi_2}{K_{s_2} + \xi_2 + \frac{\xi_2^2}{K_{i_2}}}.$$
(5.6)

A canonical state space transformation of the AD system can be obtained by considering the partition  $\xi = [\xi_a \ \xi_b]^T$ , where  $\xi_a = [\xi_3 \ \xi_4]^T$  and  $\xi_b = [\xi_1 \ \xi_2]^T$ , and a linear transformation of the states  $x_a = \xi_a$  and  $x_b = \xi_b - C_b C_a^{-1} \xi_a$  [Bastin and Dochain, 1990, Bastin and Van Impe, 1995] as

$$\dot{x}_a = u(w_a - x_a) + C_a \rho(x),$$
  
 $\dot{x}_b = u(w_b - x_b),$  (5.7)

where  $x_a \triangleq [x_3 \ x_4]^{\top}$  are the state variables that represent the acidogenic and methanogenic bacteria respectively,  $x_b \triangleq [x_1 \ x_2]^{\top}$ , are the estate variables that represent a linear combination of the four state variables,  $C_a = I_2, \ \rho(x) \triangleq [\rho_1(x) \ \rho_2(x)]$  are the reaction rates,  $w_a \triangleq [w_3 \ w_4]^{\top} = [0 \ 0]^{\top}$  are the concentration of species in the input flow,  $w_b \triangleq [w_1 \ w_2]^{\top} = [\xi_{in_1} \ \xi_{in_2}]$  are the concentration of substrates in the input flow,

$$C_b = \begin{bmatrix} -a & 0\\ c & -d \end{bmatrix}.$$

is the stoichiometric matrix, and the reaction rates in the canonical state space are

$$\rho_i(x) = r_i(\xi)|_{\xi_a = x_a; \ \xi_b = x_b + C_b C_a^{-1} x_a}, \ i = 1, 2.$$

Unlike the four-state representation (5.2), the canonical state representation is a two-state variable nonlinear differential equation system in the bacterial populations  $x_a$  and a two-state linear differential equation where the state is a linear combination of species and substrates. This linear part is assumed to be fast in comparison with the nonlinear part and therefore its dynamics can be neglected for the analysis of the behavior of the system in the species concentration. As a consequence the problem reduces to a two-state system for which the analysis is easier and can be illustrated in 2D figures of the bacterial concentrations.

In general, the species measurements are not available. They often come from computer sensors from models such as [Bernard et al., 2001] whose identification comes from the available data such as gas flow-rate (sum of all gas compounds), the content of volatile suspended solids, volatile fatty acids, total organic carbon or chemical oxygen demand among others.

Table 5.1. Model parameters						
a	42.14		$K_{s_1}$	7.1	g/l	
c	116.15	$\mathrm{mmol/g}$	$K_{s_2}$	9.28	$\mathrm{mmol/g}$	
d	268	$\mathrm{mmol/g}$	$K_{i_2}$	256	$\mathrm{mmol/g}$	
$\mu_{m_1}$	1.2	day <sup>-1</sup>	$\xi_{in_1}$	[60 - 90]	g/l	
$\mu_{m_2}$	0.74	day <sup>-1</sup>	$\xi_{in_2}$	[160 - 245]	$\mathrm{mmol/l}$	
q	453		u	[0.42 - 0.5]	$day^{-1}$	

Table 5.1: Model parameters

Depending on the magnitude of the dilution rate and substrate concentration at the inflow, the system may possess up to six equilibrium points. In [Sbarciog et al., 2010a] and [Shen et al., 2007], the authors present the regions in the input space for which various numbers of equilibria occur, and their stability. This analysis example will consider the multi-stability case, where there are two asymptotically stable equilibrium points, one representing the desired working point of the system, where the acidogenic and methanogenic bacteria coexist, and the other, the acidification point.

Consider the benchmark presented in [Sbarciog et al., 2010b] with system parameters as in Table 5.1, where the authors present the stability boundaries for the problem using backward integration. For the analysis, the value of the dilution rate is fixed at u = 0.45 [day<sup>-1</sup>], and the concentration of substrates at the input are set in a grid, where  $40 \le w_1 \le 70$  [g/l] and  $170 \le w_2 \le 225$  [mmol/l]. This paper will not consider a unique dilution rate, instead, it considers a three dimensional grid of dilution rate, and input concentrations to perform the analysis.

Figure 5.1 depicts the phase plane orbits of bacteria population for different values of the inlet for the dilution rate and substrate concentration. The system has six equilibrium points; two asymptotically stable, the working point  $x_A^*$  in which the species coexist and the acidification point  $x_B^*$  of methanogenic wash-off. Additionally, the unstable saddle points are: the point that separates the two asymptotically stable  $x_C^*$ , the two acidogenic wash-off points  $x_D^*$ ,  $x_E^*$ , and the total wash-out of the system  $x_F^*$ . Each of these points if of *type-k*, where k is the number of unstable modes of the equilibrium point.

The objective is to get the separation of the state space for the dif-



Figure 5.1: Phase plane three different realizations of the anaerobic digestion process.

ferent parameterizations that specifies if an initial condition converges to either one of the asymptotically stable equilibrium points. Depending on the magnitude of the dilution rate and substrate concentration at the inflow, the system may possess up to six equilibrium points. In [Sbarciog et al., 2010a], the authors present the regions in the input space for which various numbers of equilibria occur and their stability.

# 5.2 Anaerobic Digestion Analysis

The previous sections present the necessary tools to perform the analysis, a methodology to get an approximation of the discrete-time Koopman operator based on orthogonal polynomials, and a couple of degree and dimension reduction methods. This representation gives the possibility to approximate the location of the fixed points of the system, provides their stability, and use the set of eigenfunctions with an associated eigenvalue equal to one in conjunction with the *type-one* saddle points in the boundary of the ROA to determine the convergence of an arbitrary point in the state space. In summary, the required steps to perform the analysis are:

1 From a set of system trajectories, and with the EDMD algorithm of Section 3.4, calculate the discrete-time approximation of the Koopman operator.

- 2 From the discrete-time representation, use (4.1) to find the fixed points of the system.
- 3 Calculate the Jacobian of (3.54), and evaluate it on the fixed points to get their local stability according to the spectral decomposition of the matrix. It is important to remark that these eigenvalues come from the classical stability analysis of linear systems, and are not related to the eigenvalues from the Koopman operator.
- 4 Analyze the ROA of the system, this includes finding an eigenfunction with associated eigenvalue equal to one for the case where the EDMD algorithm fails to provide an accurate one. This eigenfunction comes from (4.13).

Given that the anaerobic digestion system in its reduced form (5.2) is bistable under some operating conditions, where the separatrix between the stability regions is the stable manifold of a unique *type-one* saddle point, we can show the accuracy of the method analyzing this system.

Consider 100 different operating conditions of system (5.2) in a 3D grid of four different dilution rates u, five different values for the organic substrate  $\xi_{in_1}$  and five different values for the volatile fatty acids  $\xi_{in_2}$  with values given in Table 5.2.

### 5.2.1 Training and Testing Data

The EDMD algorithm captures the dynamics of the portion of the state space where the data for the calculation lies. For the case of the eigenfunctions that approximate the ROA of the process, it is especially important to have an even distribution of data on the portion of the state space subject to evaluation.

There are some geometrical properties that hold for the reduced Anaerobic Digestion System [Sbarciog et al., 2010a]. Specifically, it is important to note that the reduced system (5.2) is nonnegative by definition; concentrations cannot be negative, and the canonical representation

Table 5.2: Algorithm Parameters.					
u	[0.42, 0.44, 0.44, 0.5]				
$\xi_{in_1}$	$[60\ 67.5\ 75\ 82.5\ 90]$				
$\xi_{in_2}$	$\begin{bmatrix} 160.0 & 181.25 & 202.5 & 223.75 & 245.0 \end{bmatrix}$				
		Hermite			
	]	Legendre			
	]	Laguerre			
Polynomials	Chebyshev first				
	Chebyshev second				
	Jacobi $(1, 1)$				
	Jac	(0.5, 1)			
p	[5	8 9 10 11]			
q	$[0.6 \ 0.7 \ 0.8 \ 0.9 \ 1.0 \ 1.1]$				
	Min Max	x Average			
Train points	2,943 $5,237$	7 3,811			
Test points	3,253 $5,877$	7 4,199			
$\Delta t$ 0.1 [s]					

- 0 . 1

is also compartmental, where all the trajectories lie inside an invariant set  $S_x$  given by

$$S_x = \left\{ \begin{aligned} \xi_1 &= x_1 - ax_3 \ge 0\\ \xi_2 &= x_2 + cx_3 \ge 0\\ \xi_3 &= x_3 \ge 0\\ \xi_4 &= x_4 \ge 0 \end{aligned} \right\}.$$
(5.8)

From the boundary given by this subspace of  $\mathbb{R}^4$  we define a set of initial conditions for every one of the cases. Figure 5.2 depicts the selection of such initial conditions and their respective orbits for an arbitrary realization, specifically, for u = 0.44,  $\xi_{in_1} = 90$ , and  $\xi_{in_2} = 245$ . The uniform distribution of these initial conditions on the boundary gives enough snapshot data for the approximation and testing of the discretetime Koopman operator.

Given that these trajectories come from the numerical integration of (5.7), an important consideration is the final integration time. The



Figure 5.2: Training and testing sets of snapshots for the approximation of the discrete-time Koopman operator. Initial conditions are uniformly distributed in the boundary of  $S_x$ 

integration algorithm<sup>1</sup> stops when a specific orbit reaches steady state because repeated data in the asymptotically stable fixed points affects the condition number of G in (3.39), and as a consequence it affects the accuracy of the approximation of the Koopman operator from (3.38). When the integration stops, the final  $\Delta t$  of the integration is inconsistent with the fixed time-step of the integration, therefore, affecting also the accuracy of the approximation. The solution to this problem is to discard the last value of every trajectory.

 $<sup>^{1}</sup>MATLAB(R)$  ode23s

### 5.2.2 Approximating the Koopman operator

The first step to approximate the discrete-time Koopman operator is to define a set of orthogonal polynomials that act as observables for the system. Consider for example a set of Jacobi polynomials  $\pi_{\alpha}^{(\eta,\nu)}(x)$  of degree  $\alpha$  with parameters  $\eta$  and  $\nu$  that satisfy the differential equation

$$0 = (1 - x^{2}) \frac{d^{2}}{dx^{2}} y(x) + (\eta - \nu - (\eta + \nu + 2)x) \frac{d}{dx} y(x) + \alpha (\alpha + \eta + \nu + 1) y(x).$$
(5.9)

The solution of this differential equation for a set of nonnegative integers  $\alpha_1 = \{0, 1, 2, 3\}$ , and parameters  $\eta = 0.5$  and  $\nu = 1$  is summarized in Table 5.3, where the subscript on  $\alpha$  indicates that this are the univariate solutions for the first component of the state vector.

Table 5.3: Basis for the first state variable  $x_1$ . Solution to the Jacobi ODE with parameters  $\eta = 0.5$  and  $\nu = 1$ .

$J(\alpha_1)$	$\pi^{0.5,1}_{\alpha_1}(x_1) = y(x_1)$
$J(0_1)$	1
$J(1_1)$	$7x_1/4 - 1/4$
$J(2_1)$	$99x_1^2/32 - 9x_1/16 - 21/32$
$J(3_1)$	$715x_1^3/128 - 143x_1^2/128 - 319x_1/128 + 27/128$

Every element of the polynomial basis comes from (3.58), which gives a total of 16 polynomial elements for the full basis without truncation that correspond to all the available combinations of the products between two univariate polynomials. i.e., the whole basis is

$$\Psi(x) = \begin{bmatrix} J(0_1)J(0_2) \\ J(1_1)J(0_2) \\ J(2_1)J(0_2) \\ \vdots \\ J(0_1)J(1_2) \\ J(1_1)J(1_2) \\ \vdots \\ J(3_1)J(3_2) \end{bmatrix} = \begin{bmatrix} 1 \\ 7x_1/4 - 1/4 \\ 99x_1^2/32 - 9x_1/16 - 21/32 \\ \vdots \\ 7x_2/4 - 1/4 \\ (7x_1/4 - 1/4)(7x_2/4 - 1/4) \\ \vdots \\ (715x_1^3/128 + \dots + 27/128)(715x_2^3/128 + \dots + 27/128) \end{bmatrix}$$
(5.10)

The first reduction method is to apply a truncation scheme on this basis. From the definition of the q-quasi norm in (3.60), and the p-q-quasi norm truncation scheme in (3.61), take for example the values of p = 3 and q = 0.5 to reduce the degree and dimension of the basis. Figure 5.3 depicts the result of applying the truncation scheme, where the resulting number of polynomial elements is 8, effectively reducing the maximum degree from 6 to 3 and the dimension by half. The redefined basis for the observables is

$$\Psi(x) = \begin{bmatrix} J(0_1)J(0_2)\\ J(1_1)J(0_2)\\ J(2_1)J(0_2)\\ J(2_1)J(0_2)\\ J(0_1)J(1_2)\\ J(0_1)J(1_2)\\ J(0_1)J(2_2)\\ J(0_1)J(2_2)\\ J(0_1)J(3_2) \end{bmatrix} = \begin{bmatrix} 1\\ 7x_1/4 - 1/4\\ 99x_1^2/32 - 9x_1/16 - 21/32\\ 715x_1^3/128 - 143x_1^2/128 - 319x_1/128 + 27/128\\ 7x_2/4 - 1/4\\ (7x_1/4 - 1/4)(7x_2/4 - 1/4)\\ 99x_2^2/32 - 9x_2/16 - 21/32\\ 715x_1^3/128 - 143x_2^2/128 - 319x_2/128 + 27/128\\ (5.11) \end{bmatrix}$$

With the definition of the train, test and observables set, the discretetime approximation of the Koopman operator comes from the calculation of matrices A from (3.40), G from (3.39), and finally  $U_d$  from (3.38).

Recall that for the state evolution map of the approximation of the Koopman operator (3.54), and the error threshold (3.67) for the second reduction method, an elemental matrix B with l number of rows, and n number of columns where there is a single unitary element per column



Figure 5.3: Full basis vs. truncated basis

in the index of an injective observable needs to be defined to recover the state and set the error threshold for the elimination of polynomial elements. A simple solution to this problem is to select the univariate polynomial elements of degree one as stated in (3.69) from the basis. This type of selection also gives the advantage that the inverse of the selected function by B is linear. By applying this criterion, for this particular case,

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^{\top},$$
(5.12)

that selects the second and fifth element of the polynomial basis which are the univariate, order-one, and injective linear functions of the set of observables that define the set  $\Psi_B(x) = [\psi_2(x), \psi_5(x)]$ , and whose inverse gives the linear function of observations that recover the state as

$$\Psi_B^{-1} = \begin{bmatrix} \frac{1}{7}(4x_1+1) & \frac{1}{7}(4x_2+1) \end{bmatrix}.$$
 (5.13)

With the definition of B, the approximation is further reduced via the second method, where the threshold for the elimination of polynomial

elements from the basis comes from (3.66), and (3.67) as

$$\bar{\epsilon} = \max\left(\sum_{i=1}^{N} \|\psi_2(y_i) - U_{d_2}\Psi(x_i)\|_2^2,$$
(5.14)

$$\sum_{i=1}^{N} \|\psi_5(y_i) - U_{d_5}\Psi(x_i)\|_2^2 \right), \qquad (5.15)$$

giving a lower degree, and reduced order approximation of the discretetime Koopman operator that accurately satisfies condition (3.36) for the set of observables.

With the reduced discrete-time Koopman operator matrix  $U_d$ , the set of observables  $\Psi(x)$ , the set  $\Psi_B(x)$ , and its corresponding inverse, the linear predictor that approximates the dynamics of the system from an arbitrary initial condition comes from (3.54). Although it is useful to have this linear approximation of the evolution of the states according to the spectral decomposition of the discrete-time approximation of the Koopman operator, in practice, it is enough to use a linear predictor based on the original set of observables as in (3.36). Considering again that  $y = T^k(x)$ ,

$$\Psi_B(x_k) = (U_d^k \Psi(x_0))^\top B \tag{5.16}$$

$$y = \Psi_B^{-1}(\Psi_B(x_k)). \tag{5.17}$$

Equation (5.16) is linear, and describes the linear evolution of observations of the state space. Along with matrix B, it only describes the linear evolution of a set of injective functions whose inverse recovers the state. Given that B selects the order-one observables, (5.17) is also a linear function. The benefit of having the evolution according to the polynomial basis of observables instead of the eigenfunctions of the discrete-time approximation of the Koopman operator is in terms of accuracy. Equation (3.54) needs the left eigenvectors  $W^*$  of matrix  $U_d$ to get the eigenfunctions as in (3.50), and the right eigenvectors  $\Xi$  to recover the value of the evolved observations. To have this full representation it is necessary to calculate the inverse of  $\Xi$ , adding additional error because of the numerical approximation of the higher order inverse. As a consequence, the set of eigenfunctions also suffer accuracy issues. For eliminating this induced numerical error, the solution is to get matrix  $W^*$  as the right eigenvectors of  $U_d^{\top}$ , and apply (3.50) to get the eigenfunctions.

Using the linear evolution of observables (5.16) and (5.17) instead of eigenfunctions, the linear predictor for the case of Jacobi polynomials (5.11) is

$$\Psi_B(x_k) = (U_d^k \Psi(x_0))^\top B = \begin{bmatrix} U_{d_2} \Psi(x_0) & U_{d_5} \Psi(x_0) \end{bmatrix}$$
(5.18)

$$y = \begin{bmatrix} \frac{1}{7} (4U_{d_2}\Psi(x_0) + 1) & \frac{1}{7} (4U_{d_5}\Psi(x_0) + 1) \end{bmatrix}.$$
 (5.19)

In conclusion, these changes in the computation of the approximation give a linear predictor and a set of eigenfunctions that do not have an additional inverse numerical error.

There are still some open problems to solve in the approximation of the discrete-time Koopman operator. For our particular case, in which we select a type of orthogonal polynomial and then apply a p-q-quasi norm truncation scheme, there is no current solution to determine the type of polynomial, and the values of the p-q parameters the give the best performance. Therefore, we perform the approximation of the discrete-time Koopman operator for different types of polynomials with different values for the p-q parameters and evaluate their performance using the empirical error (3.56) to select the suboptimal polynomial type and parameter values. Table 5.2 shows the different type of polynomials, the values for the p-q sweep, and the minimum, maximum, and average number of training and testing points.

Table 5.4 shows the result of averaging the empirical error given by every one of the polynomials, and the total time required to perform the calculation<sup>2</sup>. Despite the fact that a Legendre polynomial has the lowest average empirical error within the set, there are some cases where the accuracy of the Legendre polynomial basis does not hold well enough to determine an accurate classifier.

Figure 5.4 shows the empirical error for each of the parameterizations in the grid, where the Legendre polynomial error in case 44 is 1.65 for this case, u = 0.44,  $\xi_{in_1} = 82.5$ , and  $\xi_{in_2} = 223.75$  that is high enough

 $<sup>^2 \</sup>mathrm{Intel}(\mathrm{R})$ Core<br/>(TM) i<br/>7-7567U CPU @ 3.50GHz, 16GB DDR4 RAM @ 2400MHz

Polynomial	Av. error	Time [m]
Hermite	0.89	31
Legendre	0.64	31
Laguerre	2.75	44
Chebyshev first	0.77	29
Chebyshev second	0.76	30
Jacobi $(1, 1)$	0.71	31
Jacobi (0.5, 1)	0.67	42

Table 5.4: Grid optimization results for different types of polynomials.



Figure 5.4: Empirical error for every test of the 3D grid of parameterizations.

for the EDMD algorithm not to give an accurate approximation of the discrete-time Koopman operator. Therefore, the selection of the type of basis for the whole analysis is Jacobi with parameters  $\eta = 0.5$ , and  $\nu = 1$ . Note that some of the cases have a greater empirical error in which the approximation is not ideal. The reason to have these inconsistencies in the results will be covered in the next section.

### 5.2.3 Discrete-Time Approximation Result

Recall from Theorem 3.2.1, and Definitions 4.3.1 and 4.3.2 that it is necessary to have an accurate location and stability of the fixed points of the system. Either to determine the stable manifold of the *type-one* saddle points in the boundary of the ROA for the classification criterion 4.3.1, or to apply the classification criterion 4.3.2 on the asymptotically stable ones.

For either of the classification criteria to perform correctly it is only necessary to get the location and stability of the working point  $x_A^*$ , the acidification point  $x_B^*$ , and the point that separates the these two  $x_C^*$ . Recall from ?? that the fixed points are given by the solution of (4.1), and in practice, the solution to this equation from different start points gives the whole set of fixed points according to the discrete-time approximation of the Koopman operator. The start points for the identification of the fixed points comes from the trajectories of the test set, where the initial condition of every trajectory is a starting point for the solution. In addition to these points, and for this particular case it was necessary to select a random point along these trajectories because counting only with the initial conditions, there was no available starting point that gave an identification of the acidification wash-out point.

Table 5.5: Anaerobic digestion fixed points for u = 0.47,  $x_{in_1} = 75$ , and  $x_{in_2} = 160$ .

	Theoretical	Algorithmic	
$x_A^*$	(1.67, 1.25)	(1.67, 1.25)	Working point
$x_B^*$	(1.67,  0.00)	(1.67,  0.00)	Acidification
$x_C^*$	(1.67, 0.86)	(1.67, 0.86)	Separation saddle
$x_D^*$	(0.00, 0.13)	(0.00, 0.19)	Acidogenic wash-off
$x_E^*$	(0.00,  0.53)	(0.00,  0.53)	Acidogenic wash-off
$x_F^{\overline{*}}$	(0.00,  0.00)	(0.00,  0.02)	Wash-off

For the stability of the fixed points, consider the linear evolution of observations (5.19) and definition 4.2.1. The application of (4.5) on (5.19), and with  $U_d = (U_{i,j})$ , the Jacobian matrix  $J = (J_{i,j})$  for the nonlinear evolution operator on the states is

$$J_{1,1} = U_{2,2} + \frac{1}{4}U_{2,6}(7x_2 - 1) + \frac{4}{7}\left(U_{2,3}\left(\frac{99x_1}{16} - \frac{9}{16}\right) - U_{2,4}\left(-\frac{2145x_1^2}{128} + \frac{143x_1}{64} + \frac{319}{128}\right)\right)$$
(5.20)  

$$J_{1,2} = U_{2,5} + \frac{1}{4}U_{2,6}(7x_1 - 1) + \frac{4}{7}\left(U_{2,7}\left(\frac{99x_2}{16} - \frac{9}{16}\right) - U_{2,8}\left(-\frac{2145x_2^2}{128} + \frac{143x_2}{64} + \frac{319}{128}\right)\right)$$
(5.21)  

$$J_{2,1} = U_{5,2} + \frac{1}{4}U_{5,6}(7x_1 - 1) + \frac{4}{7}\left(U_{5,3}\left(\frac{99x_2}{16} - \frac{9}{16}\right) - U_{5,4}\left(-\frac{2145x_2^2}{128} + \frac{143x_2}{64} + \frac{319}{128}\right)\right)$$
(5.22)  

$$J_{2,2} = U_{5,5} + \frac{1}{4}U_{5,6}(7x_1 - 1) + \frac{4}{7}\left(U_{5,7}\left(\frac{99x_2}{16} - \frac{9}{16}\right) - U_{5,8}\left(-\frac{2145x_2^2}{128} + \frac{143x_2}{64} + \frac{319}{128}\right)\right) .$$
(5.23)

Evaluating the identified fixed points in the Jacobian matrix and getting the respective eigenvalues give their local stability. Table 5.6 shows the result of this process for the same arbitrary case of the anaerobic digestion system. These result comes from a Jacobi polynomial basis with  $\eta = 0.5$ ,  $\nu = 1$ , p = 9, and q = 0.7.

Table 5.6: Anaerobic digestion fixed points for u = 0.47,  $x_{in_1} = 75$ , and  $x_{in_2} = 160$ .

	$ \lambda_1 $	$ \lambda_2 $	Stability
$x_A^*$	0.6448	0.8210	AS
$x_B^*$	0.6444	0.9834	AS
$x_C^*$	0.6443	1.0212	Saddle
$x_D^*$	1.0160	1.0596	Unstable
$x_E^*$	0.9318	1.0597	Saddle
$x_F^*$	0.9554	1.0589	Saddle

Figure 5.5 shows the theoretical, and approximated fixed points of an arbitrary realization of the anaerobic digestion system from the solution of (5.5) from 24 initial conditions. These come from the orbits of the train set, one for every starting point of the train set, and one from an arbitrary point along those trajectories. The error of the working point, the separating saddle and the acidification point is low enough for every realization of the system for an accurate analysis of the ROA.

The results of approximating the discrete-time Koopman operator are summarized in Figure 5.6 that shows in the first row the most accurate



Figure 5.5: Location of fixed points in the state space from the solution of the minimization problem (4.1) from different initial conditions.

realizations of the anaerobic digestion system, and in the second row, the realizations with the highest empirical error. Although the trajectories for the training set of each of the cases come from the same parameterization for the process with a different value for the dilution rate and input concentrations, not all of the orbits for the training set give the same amount of information. The reduced, nonlinear canonical model for the anaerobic digestion system (5.7) has a movement and bifurcation phenomena for the fixed points depending on the input parameters [Sbarciog et al., 2010a. For the bi-stability case, the displacement of the separating saddle  $x_C^*$  is directly proportional to the values of input concentrations, specially for the volatile fatty acids concentration  $\xi_{in_2}$  (see Figure 5.1), as a consequence, the size of the ROA for the asymptotically stable fixed points changes. The fact that the selection of initial conditions for all the realizations is uniform along the boundary of the state space, makes the available data for the acidification ROA scarce, and reduces the accuracy of the algorithm. A solution to this problem is to generate the data with an experimental design that gives enough information on both regions. The reason to avoid that approach is because for our particular case, it defeats the purpose of analyzing the system without prior knowledge of



the system, otherwise we would have to know the ROA to find the ROA.

Figure 5.6: Training orbits, testing orbits, and fixed points with their respective approximations based on Koopman linear predictor and analysis algorithms.

### 5.2.4 Approximating the ROA

The spectral decomposition of the discrete-time approximation of the Koopman operator gives a set of eigenvalues with their respective eigenfunctions. In the approximation from a Jacobi type of polynomial basis, and under the selected p-q sweep each of the system gets a subset of the eigenfunctions with a unitary or near unitary eigenvalue. For the reduced order anaerobic digestion system, and under the selected set of orbits for the approximation of the Koopman operator the set of eigenfunctions is rich enough so that it contains at least one unitary eigenfunction that characterizes the stable manifold of the separating saddle point for the classification of an arbitrary initial condition. If that was not the case, the eigenfunction with unitary associated eigenvalue comes from (4.13). Several test over the construction of unitary eigenfunctions suggest that selecting the eigenfunctions with a real valued eigenvalue closest to one

improves the performance of the algorithm. Also note that (4.13) has infinite solutions, and as a consequence not all eigenfunctions constructed this way give an accurate result. The selection of the base eigenfunctions and the solutions to the construction to guarantee an accurate approximation for either of the selection criteria is still an open question.

Also, it is important to note that the constant eigenfunction is also part of the unitary eigenfunctions set, i.e.,  $\phi_+(x) = 1$ . In particular, because of the presence of the constant observable when approximating the discrete-time Koopman operator, this trivial eigenfunction is often present in the approximation. As a result, the eigenfunction that gives an accurate classification is a near unitary, or a constructed one.

In this case, and for the 100 different parameterizations, the approximation of the discrete-time Koopman operator gives a set of eigenfunctions whose unitary or near unitary eigenvalue give an accurate classification by the saddle selection rule (4.3.1). This result is depicted in Figure 5.7, where the associated eigenvalue is real, exactly one, and whose associated eigenfunction is nontrivial. The accuracy of this eigenfunction and the application of the saddle classification rule is the comparison of an additional test set of 900 initial conditions of the state space against the algorithm. for the realization under consideration, the error is 3 misclassified initial conditions out of the 900.



Figure 5.7: Theoretical, saddle based classification of initial conditions, and eigenfunction with the corresponding saddle point and stable manifold.

Consider next a realization where the approximation of the discretetime Koopman operator gives a real valued nontrivial eigenfunction whose associate eigenvalue is near unitary, i.e.,  $\mu = 1.0078$ . Although the accuracy decreases, where the number of misclassified initial conditions is seven out of 900, it is till a good approximation of the stable manifold and the convergence of an arbitrary initial condition in the state space.



Figure 5.8: Theoretical, saddle based classification of initial conditions, and eigenfunction with the corresponding saddle point and stable manifold.

For illustration purposes, consider now a case where the associated eigenvalue is not real valued or unitary, where  $\mu = 1.003 + 0.0101i$ . The eigenfunction along with the saddle classification algorithm gives an error of 35 misclassified initial conditions out of 900 where the result is given in Figure 5.9, it shows that the eigenfunction is not monotonic with respect to  $x_4$ , and although the approximation of the stable manifold is accurate, and the non-monotonic portion is completely below the threshold given by the real part of the eigenfunction evaluated at the saddle point. If that was not the case the eigenfunction would give a classification error in the ROA of  $x_A^*$ .

Recall from Section 5.1 that the driving forces in the reduced anaerobic digestion process are the dilution rate, which is the controlled input of the process, and the uncontrolled fluctuations of inlet organic matter and volatile fatty acids concentration. These driving forces produce a displacement of the fixed points, and consequently a displacement of the stable manifold of the separating saddle point in the boundary of the ROA [Shen et al., 2007]. Additionally, consider that the canonical system representation is present because it allows for graphical representations of the bacterial dynamics in a 2D plane. As a consequence the


Figure 5.9: Theoretical, saddle based classification of initial conditions, and eigenfunction with the corresponding saddle point and stable manifold.

movement of the separating saddle stability boundary can be illustrated in  $x_3 - x_4 - x_i$  3D plots where  $x_i = \{u, \xi_{in_1}, \xi_{in_2}\}$  that sweep over one of the parameters while the other two are constant.

Figure 5.10 illustrates two different cases for the  $x_3 - x_4 - u$  case, varying the dilution rate between  $0.42 \le u \le 0.54$  as shown in Table 5.2 and keeping constant values of  $\xi_{in_1} = \{90, 82.5\}$ , and  $\xi_{in_2} = \{160, 224\}$ . The plot shows the displacement of the fixed points, specially how the stable manifold and the working point  $x_A^*$  move toward each other in the  $x_4$  axis, decreasing the methane production rate according to (5.3), and reducing the portion of the state space that converges to the working point.

Consider now the  $x_3 - x_4 - \xi_{in_1}$  that describes an increase in the organic matter input concentration as  $60 \leq \xi_{in_1} \leq 90$  for the values given in Table 5.2 for two different constant values of dilution rate and volatile fatty acid concentration of  $u = \{0.447, 0.473\}$ , and  $\xi_{in_2} = \{160, 224\}$ , using the same fixed values for the input concentration as the previous case. Figure 5.11 depicts the approximation of the stable manifold of the dividing saddle, and shows that for this case, the relation between the organic matter and the dividing saddle is directly proportional in both axis. As a consequence, an increase in the organic matter concentration with the drawback of having a reduction of the working point ROA.

The final case considers the same constant values for the organic



Figure 5.10: Intersection of the stable manifold different realization s with different u, and constant concentrations at the input.

matter concentration at the input feed of  $\xi_{in_1} = \{90, 82.5\}$  as in the firs case, and evaluates for two different values of the dilution rate of  $u = \{0.447, 0.5\}$  while varying for the different values of volatile fatty acids concentration in the input feed given by Table 5.2, i.e.,  $160 \leq \xi_{in_2} \leq 245$ . Figure 5.12 shows the approximation of the dividing saddle stable manifold that gives the boundary of the ROA for the asymptotically stable fixed points. In tis case, the movement of the stable manifold and the working point are on the  $x_4$  axis is directly proportional to the



Figure 5.11: Intersection of the stable manifold different realization s with different u, and constant concentrations at the input.

volatile fatty acids concentration at the input. As a consequence, there is an increase in the methane production and a reduction of the effective state space that converges to the working point.



Figure 5.12: Intersection of the stable manifold for different realizations with different  $x_{in_2}$ , and constant dulition rate, and  $x_{in_1}$  concentration at the input.

## 5.3 Discussion

In general, not only for the case of the anaerobic digestion system, the analysis of the ROA is static, that is, it comes from a fixed value of the input. For the process at hand these are the input concentrations of substrates and the dilution rate. The rationale behind this approach is the movement and bifurcation phenomena present in the equilibrium points of nonlinear systems, for example, consider the movement of the fixed points with respect to the input concentrations and dilution rate as depicted in Figures 5.10 to 5.12. The working point ROA size is inversely proportional to all of the variations in the input, while the converse applies for the movement of the working point. Meaning that a transient and steady state optimization of biogas production benefits from a dynamical analysis instead of a static one, which can reduce the computational burden for this type of optimizations, and can improve the current heuristic approach [Sbarciog et al., 2011].

Under the Koopman operator approach and the discrete-time approximation given by the EDMD algorithm, it is possible to formulate an expansion of the state space defined as the product of the original space and the input space, which is equivalent to setting the inputs as state variables without specific dynamics. Using a polynomial expansion on the input space rather than considering the approximation given by Korda [Korda and Mezić, 2018b] as an input affine system gives the possibility to handle it as a closed one, and as a consequence, analyze the movement of the equilibrium points and the stability boundary in a dynamic way.

The approximation of the discrete-time Koopman operator, either in static or dynamic ways via the EDMD algorithm still has some important problems to be resolved. The first one is the amount of necessary data to have an accurate representation. For the anaerobic digestion process at hand, the type of analysis is possible for an identified system that can be numerically integrated from a set of initial conditions. Even though the reduction methods and the possibility to test with different polynomials gives an accurate representation with a small amount of data, the necessary amount at the moment, makes the analysis unfeasible if the solely available trajectories come from a real digester.

Additionally, the type of data available also poses a problem for the analysis given that the species concentration is not commonly available for the process, while the biogas outflow is. Meaning that the algorithm needs to be extended so that the eigenfunctions from an approximation with partial information are still able to capture the dynamic behavior of the system to approximate all the necessary information: fixed points, stability, unitary eigenfunctions, and classification criteria. Furthermore, for the dynamical analysis, it is necessary to get a classification criteria that not only depends on the current state but also depends on the input concentrations and dilution rate of the system.

The analysis provides a clear criterion that does not depend on the specific geometry of a manifold or the ability to calculate level sets and interpolate between data-points for the classification of an arbitrary point in the state space. As a result, the analysis can be extended to models with more than two reactions and state variables while preserving a clear and simple criterion for the classification.

Given that it is difficult and costly to acquire data from a reactor, this approach also opens up the possibility of developing an analysis based on the commonly available measurements such as the flow of biogas at the output. To make that possible, it is necessary that the approximation and reduction techniques presented here are extended so that they achieve an accurate approximation of the discrete-time Koopman operator for a reduced set of forced system trajectories.

## Chapter 6

## Conclusions

The development of the thesis began with an exploratory process of the available literature for analyzing and controlling interconnected dynamical systems. Part of this inquiry where the traditional modelbased techniques or input-output techniques such as dissipativity, passivity, population dynamics and Lyapunov based techniques for the former. Regarding the latter, techniques from the solution of the Lur'e problem, from the circle criterion up to integral quadratic constraints. In addition to this model based techniques, and in the search for a data-driven alternative able to provide analysis tools for the non-linear dynamical systems, this exploration led to the finding of the Koopman operator and its approximation via the EDMD algorithm.

Upon further research, and considering the potential that the Koopman operator has to analyze the associated spectrum of the discrete-time operator, its eigenvalues and eigenfunctions, the thesis took upon this path as the selected method to achieve the objectives.

The theory regarding the Koopman operator is a powerful tool that is still in its infancy with a lot of potential to tackle on the plethora of problems regarding the subject matter of control systems analysis and synthesis. However, the closed and analytical Koopman operator representation is only available for a small family of polynomial systems, making it a must, the development of tools to accurately approximate it.

Currently, the best tool for the task of approximating the Koopman

operator is the EDMD algorithm, that certainly is at a similar stage of development as the operator, i.e., in its early stages. There are still several problems to solve within the method, where the main, is the selection of observables. Also within these problems is the numerical stability of the algorithm, where more often than not, the problem is in the inverse of matrix G, arguably, the foundation of the algorithm. As a consequence, the inversion of this matrix has to rely on a Moore-Penrose pseudo inverse, instead of an inverse. The reason for this phenomena is the poor condition number that the matrix has after the selection of the observables and the evaluation of the available data for training the algorithm.

Continuing with the assessment of the EDMD algorithm and its relation to the Koopman operator, having an accurate approximation of the dynamics of the system via the former does not guarantee an accurate approximation of the eigenfunctions of the latter. As a consequence of this discrepancy, and given the fact that in the dimensional limit of the set of observables there is a better approximation of the eigenfunctions, there is a compromise between these two concepts.

Given all of the aforementioned drawbacks, the focus of the thesis turned into the improvement of the EDMD algorithm to circumvent them. Giving rise to the use of p-q-quasi norms, the reduction by polynomial accuracy, the inverse of order one univariate polynomial elements to recover the state, and the trigonometric embeddings. With these improvements the accuracy of the predictors and the approximation of the Koopman operator had a significant improvement, not only in the error metrics that determine their effectiveness, they also reduce the necessary amount of trajectories, and therefore the amount of data-points for the approximations. Without the improvements proposed by the thesis, the method for approximating the boundary of the ROA in the problem of the multi-stability analysis would be just a theoretical possibility, instead of a thoroughly tested algorithm. Among the most important achievements of the process for approximating the ROA is the fact that it can handle complex systems with more than a couple of state variables and with several asymptotically stable fixed points. Additionally, it shows potential for the solution of real problems as with the static analysis of the anaerobic digestion process.

In conclusion, there is still a long way to consider the EDMD algorithm and the theoretical concepts of the Koopman operator mature tools. Although the recent contributions for the analysis of dynamical systems are growing and becoming widespread, the forecast of these tools, that are expected to bring some of the numerous methods of linear systems into these linear transformations is still not fulfilled. Moreover, synthesis tools are limited to MPC, while the availability of the point spectrum of the Koopman operator transformation is not yet exploited, in contrast with the linear systems tools that heavily rely on the spectral decomposition of the state transition matrix. As a consequence, the research for closing this gap offers endless possibilities. In that regard, and hopefully, the main contribution of this thesis is grounding the algorithmic methods such that they become more widespread in the control systems community and get more attention for their development.

## Bibliography

- Amir Ali Ahmadi and Pablo A. Parrilo. Converse results on existence of sum of squares Lyapunov functions. In *IEEE Conference on Deci*sion and Control and European Control Conference, pages 6516–6521, Orlando FL, USA, dec 2011. IEEE. ISBN 978-1-61284-801-3. doi: 10.1109/CDC.2011.6161493.
- James Anderson and Antonis Papachristodoulou. Dynamical system decomposition for efficient, sparse analysis. In 49th IEEE Conference on Decision and Control (CDC), pages 6565–6570, Atlanta GA, USA, dec 2010. IEEE. ISBN 978-1-4244-7745-6. doi: 10.1109/CDC.2010. 5717269.
- Jacqueline Augusiak, Paul J. Van den Brink, and Volker Grimm. Merging validation and evaluation of ecological models to 'evaludation': A review of terminology and a practical approach. *Ecological Modelling*, 280:117–128, 2014.
- G. Bastin and D. Dochain. On-line Estimation and Adaptive Control of Bioreactors. Elsevier, Amsterdam, 1990.
- G. Bastin and J.F. F. Van Impe. Nonlinear and Adaptive Control in Biotechnology: A Tutorial. *European Journal of Control*, 1(1):37–53, jan 1995.
- Richard Bellman. Vector Lyanpunov Functions. Journal of the Society for Industrial and Applied Mathematics Series A Control, 1(1):32–34, jan 1962. ISSN 0887-4603. doi: 10.1137/0301003. URL http://epubs. siam.org/doi/10.1137/0301003.

- Olivier Bernard, Zakaria Hadj-Sadok, Denis Dochain, Antoine Genovesi, and Jean Philippe Steyer. Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology and Bioengineering*, 75(4):424–438, 2001. ISSN 00063592. doi: 10.1002/bit.10036.
- Steven L. Brunton, Bingni W. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE*, 11(2):1–19, 2016a.
- Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, apr 2016b.
- Marko Budišić, Ryan Mohr, and Igor Mezić. Applied koopmanism. Chaos: An Interdisciplinary Journal of Nonlinear Science, 22(4): 47510, 2012.
- Vijaysekhar Chellaboina, Sanjay Bhat, Wassim Haddad, and Dennis Bernstein. Modeling and analysis of mass-action kinetics. *IEEE Con*trol Systems Magazine, 29(4):60–78, aug 2009.
- H D Chiang and L F C Alberto. Stability Regions of Nonlinear Dynamical Systems: Theory, Estimation, and Applications. Cambridge University Press, 2015. ISBN 9781316374320.
- Edson A. Coayla-Teran, S. E A Mohammed, and P. R C Ruffino. Hartman-Grobman theorems along hyperbolic stationary trajectories. Discrete and Continuous Dynamical Systems, 17(2):281–292, 2007.
- Federico Cuesta, Francisco Gordillo, Javier Aracil, and Anibal Ollero. Stability analysis of nonlinear multivariable Takagi-Sugeno fuzzy control systems. *IEEE Transactions on Fuzzy Systems*, 7(5):508–520, 1999.
- David Di Ruscio. Model Predictive Control with Integral Action: A simple MPC algorithm. *Modeling, Identification and Control*, 34(3): 119–129, 2013.

- Georgi M Dimirovski. Complex Systems: Relationships between Control, Communications and Computing, volume 55. Springer, 2016.
- Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel. Operator Theoretic Aspects of Ergodic Theory, volume 272 of Graduate Texts in Mathematics. Springer International Publishing, Cham, 2015.
- Nesma ElKalaawy and Amr Wassal. Methodologies for the modeling and simulation of biochemical networks, illustrated for signal transduction pathways: A primer. *BioSystems*, 129:1–18, 2015.
- Martin Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors-I. The deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 42(10):2229–2268, 1987. ISSN 00092509. doi: 10.1016/0009-2509(87)80099-4.
- C. Garcia-Tenorio, D. Tellez-Castro, E. Mojica-Nava, and A. Vande Wouwer. Analysis of hyperbolic systems via data-driven koopman operator. *Submitted to Applied Mathematics and Computation*, 2020.
- Camilo Garcia-Tenorio, Nicanor Quijano, Eduardo Mojica-Nava, and Jorge Sofrony. Bond graph model-based for IDA-PBC. In 2016 IEEE Conference on Control Applications (CCA), pages 1098–1103. IEEE, sep 2016. ISBN 978-1-5090-0755-4. doi: 10.1109/cca.2016.7587953.
- Camilo Garcia-Tenorio, Duvan Tellez-Castro, Eduardo Mojica-Nava, and Alain Vande Wouwer. Analysis of a Class of Hyperbolic Systems via Data-Driven Koopman Operator. In *International Conference on Sys*tem Theory, Control and Computing (ICSTCC), pages 566–571, 2019.
- Hugues Garnier and Liuping Wang, editors. *Identification of Continuoustime Models from Sampled Data*. Advances in Industrial Control. Springer-Verlag London, London, 1 edition, 2008.
- Peter Giesl and Sigurdur Hafstein. Review on computational methods for Lyapunov functions. Discrete and Continuous Dynamical Systems-Series B, 20(8):2291–2331, 2015.
- L. Grujic and D. Siljak. Asymptotic stability and instability of largescale systems. *IEEE Transactions on Automatic Control*, 18(6):636–

645, dec 1973. ISSN 0018-9286. doi: 10.1109/TAC.1973.1100422. URL http://ieeexplore.ieee.org/document/1100422/.

- Wassim M Haddad and VijaySekhar Chellaboina. Nonlinear dynamical systems and control: a Lyapunov-based approach. Princeton University Press, 2008. ISBN ISBN-13: 978-0-6911-3329-4.
- P. Ioannou. Decentralized adaptive control of interconnected systems. *IEEE Transactions on Automatic Control*, 31(4):291-298, apr 1986. ISSN 0018-9286. doi: 10.1109/TAC.1986.1104282. URL http:// ieeexplore.ieee.org/document/1104282/.
- E. Kaiser, J. N. Kutz, and S. L. Brunton. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A: Mathematical, Physical and En*gineering Sciences, 474(2219), 2018.
- Hassan K Khalil. *Nonlinear systems*. Prentice Hall, 3rd edition, 2002. ISBN 0130673897,9780130673893,0131227408,9780131227408.
- Stefan Klus, Péter Koltai, and Christof Schütte. On the numerical approximation of the Perron-Frobenius and Koopman operator. Journal of Computational Dynamics, 3(1):51–79, 2016. doi: 10.3934/jcd. 2016003.
- Roelof Koekoek, Peter A. Lesky, and René F. Swarttouw. *Hypergeometric* Orthogonal Polynomials and Their q-Analogues. Springer Monographs in Mathematics. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- Katerina Konakli and Bruno Sudret. Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions. *Journal of Computational Physics*, 321:1144–1169, sep 2016a.
- Katerina Konakli and Bruno Sudret. Reliability analysis of highdimensional models using low-rank tensor approximations. *Probabilis*tic Engineering Mechanics, 46:18–36, 2016b.
- Bernard O. Koopman. Hamiltonian Systems and Transformation in Hilbert Space. Proceedings of the National Academy of Sciences, 17 (5):315–318, may 1931.

- Milan Korda and Igor Mezić. On Convergence of Extended Dynamic Mode Decomposition to the Koopman Operator. Journal of Nonlinear Science, 28(2):687–710, apr 2018a.
- Milan Korda and Igor Mezić. Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. Automatica, 93:149–160, 2018b. URL http://arxiv.org/abs/1611.03537.
- Yueheng Lan and Igor Mezić. Linearization in the large of nonlinear systems and Koopman operator spectrum. *Physica D: Nonlinear Phe*nomena, 242(1):42–53, 2013.
- A Lasota and James A. Yorke. Exact dynamical systems and the frobenius-perron operator. TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY, 273(1):375–384, 1982.
- Qianxiao Li, Felix Dietrich, Erik M. Bollt, and Ioannis G. Kevrekidis. Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the koopman operator. Chaos: An Interdisciplinary Journal of Nonlinear Science, 27 (10):103111, 2017. doi: 10.1063/1.4993854.
- A. Linnemann. Decentralized control of dynamically interconnected systems. *IEEE Transactions on Automatic Control*, 29(11):1052–1054, nov 1984. ISSN 0018-9286. doi: 10.1109/TAC.1984.1103423. URL http://ieeexplore.ieee.org/document/1103423/.
- Ludovic Mailleret, Olivier Bernard, and J. P. Steyer. Robust regulation of anaerobic digestion processes. *Water Science and Technology*, 48(6): 87–94, 2003. ISSN 02731223.
- Anirudha Majumdar, Amir Ali Ahmadi, and Russ Tedrake. Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *Proceedings of the IEEE Conference on Decision and Control*, number February, pages 394–401, 2014. ISBN 978-1-4673-6090-6. doi: 10.1109/CDC.2014.7039413.
- Stefano Marelli and Bruno Sudret. An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis. *Structural Safety*, 75(August 2017):67–74, 2018.

- B. M. Maschke and A.J. J van der Schaft. Interconnection of systems: the network paradigm. In 35th Conference on Decision and Control, volume 1, pages 207–212. IEEE, 1996. ISBN 0780335902. doi: 10. 1109/CDC.1996.574297.
- A. Mauroy and J. Hendrickx. Spectral identification of networks using sparse measurements. SIAM Journal on Applied Dynamical Systems, 16(1):479–513, 2017a. ISSN 15360040. doi: 10.1137/16M105722X.
- A. Mauroy, I. Mezić, and J. Moehlis. Isostables, isochrons, and koopman spectrum for the action-angle representation of stable fixed point dynamics. *Physica D: Nonlinear Phenomena*, 261:19 30, 2013. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2013.06. 004. URL http://www.sciencedirect.com/science/article/pii/S0167278913001620.
- Alexandre Mauroy and Julien M. Hendrickx. Spectral identification of networks with inputs. In 2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017, volume 2018-Janua, pages 469–474, sep 2017b. ISBN 9781509028733. doi: 10.1109/CDC.2017.8263708.
- Alexandre Mauroy and Igor Mezic. Global Stability Analysis Using the Eigenfunctions of the Koopman Operator. *IEEE Transactions on Au*tomatic Control, 61(11):3356–3369, nov 2016.
- Igor Mezić. Spectral Properties of Dynamical Systems, Model Reduction and Decompositions. *Nonlinear Dynamics*, 41(1):309–325, 2005. ISSN 0924-090X.
- Igor Mezić. Koopman operator spectrum and data analysis, feb 2017.
- Arnold Neumaier. Solving Ill-Conditioned and Singular Linear Systems: A Tutorial on Regularization. *SIAM Review*, 40(3):636–666, 2003. ISSN 0036-1445. doi: 10.1137/s0036144597321909.
- Frank K. Lu Paul Zarchan, Howard Musoff. Fundamentals of Kalman Filtering:: A Practical Approach, volume 232 of Progress in Astronautics and Aeronautics (Volume 232). AIAA (American Institute of Aeronautics & Astronautics), 3 edition, 2009.

- Sigurdur Hafstein Peter Giesl. Review on computational methods for lyapunov functions. Discrete & Continuous Dynamical Systems - B, 20:2291, 2015.
- Joshua L. Proctor, Steven L. Brunton, and J. Nathan Kutz. Dynamic mode decomposition with control. SIAM Journal on Applied Dynamical Systems, 15(1):142–161, 2016.
- Nicanor Quijano, Carlos Ocampo-Martinez, Julian Barreiro-Gomez, German Obando, Andres Pantoja, and Eduardo Mojica-Nava. The Role of Population Games and Evolutionary Dynamics in Distributed Control Systems: The Advantages of Evolutionary Game Theory. *IEEE Control Systems*, 37(1):70–97, feb 2017. ISSN 1066-033X. doi: 10.1109/MCS.2016.2621479. URL http://ieeexplore.ieee.org/document/7823106/.
- Lorenzo Rosasco, Ernesto De Vito, Andrea Caponnetto, Michele Piana, and Alessandro Verri. Are Loss Functions All the Same? Neural Computation, 16(5):1063–1076, 2004. ISSN 08997667. doi: 10.1162/089976604773135104.
- M. Sbarciog, M. Loccufier, and E. Noldus. Determination of appropriate operating strategies for anaerobic digestion systems. *Biochemical En*gineering Journal, 51(3):180–188, 2010a. URL http://dx.doi.org/ 10.1016/j.bej.2010.06.016.
- Mihaela Sbarciog, Mia Loccufier, and Erik Noldus. The estimation of stability boundaries for an anaerobic digestion system. *IFAC Proceedings Volumes*, 43(6):359 364, 2010b. ISSN 1474-6670. doi: https://doi.org/10.3182/20100707-3-BE-2012. 0032. URL http://www.sciencedirect.com/science/article/pii/S1474667016303810. 11th IFAC Symposium on Computer Applications in Biotechnology.
- Mihaela Sbarciog, Mia Loccufier, and Alain Vande Wouwer. On the optimization of biogas production in anaerobic digestion systems<sup>\*</sup>. *IFAC Proceedings Volumes*, 44(1):7150 – 7155, 2011. ISSN 1474-6670. 18th IFAC World Congress.

- Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, aug 2010.
- M. Erol Sezer and Özay Hüseyin. Stabilization of Linear Time-Invariant Interconnected Systems Using Local State Feedback. *IEEE Trans*actions on Systems, Man, and Cybernetics, 8(10):751-756, 1978. ISSN 0018-9472. doi: 10.1109/TSMC.1978.4309846. URL http: //ieeexplore.ieee.org/document/4309846/.
- Shuiwen Shen, Giuliano C Premier, Alan Guwy, and Richard Dinsdale. Bifurcation and stability analysis of an anaerobic digestion model. Nonlinear Dynamics, 48(4):391–408, 2007. ISSN 1573-269X. doi: 10.1007/s11071-006-9093-1. URL https://doi.org/10.1007/ s11071-006-9093-1.
- Shih-Ho Wang and E. Davison. On the stabilization of decentralized control systems. *IEEE Transactions on Automatic Control*, 18(5):473– 478, oct 1973. ISSN 0018-9286. doi: 10.1109/TAC.1973.1100362. URL http://ieeexplore.ieee.org/document/1100362/.
- Dragoslav D. Siljak. Stability of Large-Scale Systems under Structural Perturbations. *IEEE Transactions on Systems, Man, and Cybernetics*, 2(5):657-663, 1972. ISSN 0018-9472. doi: 10.1109/TSMC.1972. 4309194. URL http://ieeexplore.ieee.org/document/4309194/.
- Dragoslav D Siljak. *Decentralized Control of Complex Systems*. Mathematics in science and engineering. Academic Press, 1991.
- S. J. Skar, R. K. Miller, and A. N. Michel. On Nonexistence of Limit Cycles in Interconnected Systems. *IEEE Transactions on Automatic Control*, 26(3):669–676, 1981. ISSN 15582523.
- Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93(7):964–979, 2008. ISSN 09518320.
- Eric Walter and Luc Pronzato. Identification of parametric models from experimental data. Springer-Verlag London, London, 1 edition, 1997. ISBN 978-3-540-76119-8.

- Jan C. Willems. Dissipative dynamical systems Part II: Linear systems with quadratic supply rates. Archive for Rational Mechanics and Analysis, 45(5):352–393, 1972a. ISSN 0003-9527. doi: 10.1007/BF00276494. URL http://link.springer.com/10.1007/BF00276494.
- JanC. Willems. Dissipative dynamical systems part I: General theory. Archive for Rational Mechanics and Analysis, 45(5):321–351, 1972b. ISSN 0003-9527. doi: 10.1007/BF00276493. URL http://dx.doi. org/10.1007/BF00276493.
- Matthew O. Williams, Ioannis G Kevrekidis, and Clarence W Rowley. A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. *Journal of Nonlinear Science*, 25(6): 1307–1346, dec 2015.
- Matthew O. Williams, Clarence W. Rowley, and Ioannis G. Kevrekidis. A kernel-based method for data-driven koopman spectral analysis. *Jour*nal of Computational Dynamics, 2(2):247–265, may 2016. ISSN 2158-2491. doi: 10.3934/jcd.2015005.