

Article

Evaluation of the Regions of Attraction of Higher-Dimensional Hyperbolic Systems Using Extended Dynamic Mode Decomposition

Camilo Garcia-Tenorio ^{1,2}, Duvan Tellez-Castro ², Eduardo Mojica-Nava ² and Alain Vande Wouwer ^{1,*}¹ Systems, Estimation, Control and Optimization (SECO), Université de Mons, 7000 Mons, Belgium² Departamento de Ingeniería Eléctrica y Electrónica—Mecánica y Mecatrónica, Universidad Nacional de Colombia, Carrera 30 No. 45-03, Bogotá 111321, Colombia

* Correspondence: alain.vandewouwer@umons.ac.be

Abstract: This paper provides the theoretical foundation for the approximation of the regions of attraction in hyperbolic and polynomial systems based on the eigenfunctions deduced from the data-driven approximation of the Koopman operator. In addition, it shows that the same method is suitable for analyzing higher-dimensional systems in which the state space dimension is greater than three. The approximation of the Koopman operator is based on extended dynamic mode decomposition, and the method relies solely on this approximation to find and analyze the system's fixed points. In other words, knowledge of the model differential equations or their linearization is not necessary for this analysis. The reliability of this approach is demonstrated through two examples of dynamical systems, e.g., a population model in which the theoretical boundary is known, and a higher-dimensional chemical reaction system constituting an original result.

Keywords: regions of attraction; extended dynamic mode decomposition; Koopman operator; nonlinear systems



Citation: Garcia-Tenorio, C.; Tellez-Castro, D.; Mojica-Nava, E.; Vande Wouwer, A. Evaluation of the Regions of Attraction of Higher-Dimensional Hyperbolic Systems Using Extended Dynamic Mode Decomposition. *Automation* **2023**, *4*, 57–77. <https://doi.org/10.3390/automation4010005>

Academic Editor: Eyad H. Abed

Received: 1 November 2022

Revised: 19 January 2023

Accepted: 27 January 2023

Published: 4 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The analysis of nonlinear systems often focuses on the stability of the fixed point or equilibrium point of the system, especially the global stability of this unique point. When the fixed point is not unique, the concept of the region of attraction (ROA) is as important as the stability of the point. In general terms, the ROA represents the extent to which a disturbance can drive the system away from a stable equilibrium point so that it can still return to it. In other words, the ROA indicates from which initial conditions the system converges to a stable point. This analysis is useful in biological or ecological systems, where there are several equilibrium points in which one or all species vanish, and other points where the species coexist [1,2].

There are traditional techniques for approximating the ROA of asymptotically stable equilibrium points, such as the computation of the level sets of a local Lyapunov function [3], the backwards integration of systems from saddle equilibrium points to approximate the boundary of the ROA [4], or the computation of the level sets of a local energy function, similar to the Lyapunov approach. As local energy functions are constant along the trajectories of the system, they can provide the stable manifold of the saddle points in the boundary. Consequently, these stable manifolds form the whole boundary of the ROA [4]. Among these techniques, integrating the system back from the saddle points in the boundary does not require knowledge of the local Lyapunov (or energy) function of the system, and provides a less conservative approximation.

All the above methods rely upon the explicit knowledge of a mathematical model, i.e., a set of differential equations that result from mathematical modeling and parameter

identification methods [5–7], and backward and forward integration require that the system be backwards-integrable.

In this paper, we propose a data-driven approach based solely on information gathered either from the collection of experimental data from a physical system or the numerical integration of an existing numerical simulator of arbitrary form and complexity. If experimental data are available, there is no need to derive a mathematical model or identify the model parameters. Such data, along with numerical methods, allow the approximation of the Koopman operator, which has a set of eigenfunctions that can be used to approximate the ROA [8,9]. Rather than describing the time evolution of the system states, the Koopman operator describes the evolution of the eigenfunctions [10]. Analyzing these eigenfunctions allows identification of invariant subspaces that acquire specific characteristics of a dynamical system. For example, if an eigenfunction has an associated eigenvalue equal to one, the value of this function is invariant along the trajectories of the dynamical system. As a consequence, an eigenfunction with an associated eigenvalue equal to one defines an invariant subspace useful for capturing specific characteristics of the dynamical system.

A number of recent studies [11–14] have pointed out that eigenfunctions that have an associated eigenvalue equal to one are useful for nonlinear system analysis. However, they do not formalize the characteristics of these functions theoretically in order to perform the analysis. In [11], the authors use time averages of observables, i.e., the average of arbitrary functions of a trajectory of the state from a particular initial condition. These functions converge to the unitary (or near unitary) eigenfunction, and as such their level sets provide visual information on the state space partition. This procedure provides insight into the partition of the state space in two-dimensional systems or slices of three-dimensional ones. However, this analysis does not provide a criterion for obtaining the partition of the state space, as it only provides visual information, and this is the reason it is limited to low-dimensional systems. In addition, for calculation of the time averages, it is necessary to have the solution (or numerical integration) of the difference/differential equations from every point of the state space under consideration in order to obtain the graph depicting the level sets. Another notable approach is that from [13], which uses the isostables of a system. An isostable of a stable equilibrium point is a set of initial conditions with trajectories that converge synchronously to the attractor, that is, their trajectories simultaneously intersect subsequent isostables along a trajectory that converges to a stable point. Here, the definition of isostability is taken from the magnitude of the slowest Koopman operator eigenfunction, the level sets of which provide the isostables. Another important contribution in [14] is a global stability analysis and the approximation of the region of attraction based on spectral analysis of the Koopman eigenfunctions. The global stability analysis comes from the zero-level sets of the Koopman eigenfunctions related to the Koopman eigenvalues, for which the real part is less than zero. The approximation of the region of attraction is taken from the traditional analysis of Lyapunov functions, which are built upon the same Koopman eigenfunctions and eigenvalues as in the global stability analysis. To obtain these eigenfunctions and eigenvalues, the authors used a numerical method based on Taylor and Bernstein polynomials, where it is necessary to know the analytic vector field of the system. Similar to the Lyapunov and energy function-based methods, these methods require calculation of the level sets of a particular function, which yields the same problems with regarding the classification of an arbitrary initial condition in higher dimensional spaces.

Probably the most notable contribution that comes close to developing a purely data-driven technique is from Williams et al., with their development of the extended dynamic mode decomposition algorithm [9]. The authors provide more insight into the determination of the ROA for the particular case of a Duffing oscillator with two basins of attraction, and analyze the leading (unitary) eigenfunction to determine which basin of attraction a point belongs to. The criterion for this classification is the use of the mean value of the unitary eigenfunction as a threshold. Subsequently, the authors parameterize the ROA by recalculating an approximation of the Koopman operator on each basin. Even though the

development in [9] shows an accurate procedure for obtaining the ROA in the particular case of the Duffing oscillator, there are issues that are not covered in this development. For instance, there is no guarantee of the existence of unitary eigenfunctions when using the approximation methods of the Koopman operator. When a unitary eigenfunction is present, it is often trivial, i.e., the function value is constant in the whole state space, and as a consequence it does not provide information about the system. Furthermore, ROA analysis methods based on the fixed points of the system rely on the model equations and their linearization to determine the location and stability of the fixed point. Moreover, while the mean value of the unitary eigenfunction may work in the particular case of the Duffing oscillator, it is not a generalized criterion. Hence, it may not work for other types of systems, and guidelines on how to generalize it to higher-dimensional systems are lacking.

In this study, the above-mentioned problems are addressed, resulting in the following contributions:

- The development of an algebraic condition, rather than a complex geometric analysis, to determine the region of attraction based on a set of unitary eigenfunctions. This result is supported by theorem (i.e., Theorem 2 in the following and its proof);
- The proposal of an approach which is purely data-driven, i.e., all the necessary information comes from the approximation of the Koopman operator, including the location and local stability of the fixed points;
- The proposal of an approach which is suitable for analyzing higher-dimensional dynamical systems (i.e., those with dimensionality greater than three).

The rest of this paper is organized as follows. The next section introduces the main concepts, namely, the region of attraction, Koopman operator theory, and extended dynamic mode decomposition algorithm. Section 3 presents and discusses the main contribution of this study, i.e., a numerical procedure for evaluating the regions of attractions based on the EDMD and the calculation of eigenfunctions associated with unitary eigenvalues. Next, The methodology is applied to two examples in Section 4, namely, a population model and a higher-dimensional chemical reaction system. In both cases, the data are provided by numerical simulation of a model; however, we stress here that the procedure is applicable to experimental data as well, provided that any such data are available in sufficient numbers to secure a sufficiently accurate approximation. The last section is devoted to our concluding remarks and future research prospects.

The notations A^\top and A^+ are the transpose and pseudoinverse of a matrix, respectively, where $A \in \mathbb{R}^{n \times n}$.

2. Basic Concepts and Methods

2.1. Regions of Attraction

Consider a nonlinear dynamical system $(\mathcal{M}; T(x); k)$ in discrete time 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 $T: \mathcal{M} \rightarrow \mathcal{M}$ is the differentiable vector-valued evolution map, i.e.,

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

The solution to (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. Suppose $x^* \in \mathcal{M}$ is a fixed point of (1), i.e.,

$$T^k(x^*) = x^*. \quad (2)$$

The linearization principle defines the local stability of hyperbolic fixed points, i.e., points that satisfy the Hartman–Grobman theorem [3,15]. This principle states that a fixed point x_s^* is asymptotically stable if the modulus of all the eigenvalues of the Jacobian matrix evaluated at the fixed point is less than one, and unstable otherwise. Additionally, the type- k of a hyperbolic fixed point is defined as the number k of eigenvalues with a modulus

greater than one. If only one eigenvalue has a modulus greater than one, the fixed point x^* is a type-1 point. When the index k of unstable fixed points is equal to or greater than one and less than n the fixed point is called a saddle, denoted by \hat{x}^* . The type-1 saddle points play an important role in the approximation of the ROA.

Eigenvalues λ have a corresponding eigenvector E_λ that spans the eigenspace associated with the eigenvalue. For eigenvalues with a modulus smaller than one, the direct sum of their eigenspaces is the generalized stable eigenspace of a fixed point, i.e., $E^s = \oplus E_{|\lambda|<1}$. Conversely, for eigenvalues with a modulus greater than one, the direct sum of their eigenspaces is the generalized unstable eigenspace of a fixed point, i.e., $E^u = \oplus E_{|\lambda|>1}$. The type of 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 state space \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, it is locally the case that the stable and unstable eigenspaces are tangent to the unstable and stable manifolds of the hyperbolic fixed point. The definitions of these manifolds are

$$W^u(x^*) = \{x \in \mathbb{R}^n : \lim_{k \rightarrow -\infty} T^k(x) = x^*\}, \quad (3)$$

$$W^s(x^*) = \{x \in \mathbb{R}^n : \lim_{k \rightarrow \infty} T^k(x) = x^*\}, \quad (4)$$

for the unstable and stable manifold of a fixed point, respectively, assuming that an inverse for the backward flow exists for T^k .

Prior to statement of the theorem characterizing the ROA of an asymptotically stable point, we denote the ROA of the fixed point as $R_A(x_s^*)$ and the stability boundary as $\partial R_A(x_s^*)$. From the definitions of the unstable and stable manifolds (3) and (4), any system under analysis must satisfy the following three assumptions:

- A1: All the fixed points on $\partial R_A(x_s^*)$ are type-1.
- A2: The $W^u(x^*)$ and $W^s(x^*)$ of the type-1 points on $\partial R_A(x_s^*)$ satisfy the transversality condition.
- A3: Every trajectory that starts on $\partial R_A(x_s^*)$ converges to one of the type-1 points as $k \rightarrow \infty$.

Remark 1. Manifolds A and B in \mathcal{M} satisfy the transversality condition if the intersection of the tangent spaces of A and B span the tangent space of \mathcal{M} .

For any hyperbolic dynamical system that satisfies assumptions (A1–A3), the region of attraction of an asymptotically stable point is,

Theorem 1 ((Th. 9-(10,11)) in [4]). Consider the dynamical system (1) and assume that it satisfies assumptions (A1–A3). Define $\{\hat{x}_i^*\}_{i=1}^P$ as the P type-1 hyperbolic fixed points on the boundary, that is, the stability region of an asymptotically stable fixed point. Then,

1. $\hat{x}_i^* \in \partial R_A(x_s) \iff W^u(\hat{x}_i^*) \cap R_A(x_s^*) \neq \emptyset$
2. $\partial R_A(x_s^*) = \cup W^s(\hat{x}_i^*)$.

Hence, the stability boundary is the union of the type-1 hyperbolic fixed points' stable manifolds on the stability boundary.

Remark 2. Assumption (A1) is a generic property of differentiable dynamical systems, while assumptions (A2) and (A3) must be verified.

2.2. Basics of Koopman Operator Theory

Consider a set of arbitrary functions of state, the so-called “observables” $f(x) : \mathcal{M} \rightarrow \mathbb{C}$ for system (1), that belong to some function space, i.e., $f(x) \in \mathcal{F}$. There is a discrete-time

linear operator U^k , the Koopman operator, which defines the time evolution of these observables, i.e.,

$$[U^k f](x) = f(T^k(x)). \quad (5)$$

The left-hand side of (5) is the time evolution of the observables, while the right-hand side is the time evolution of the state subsequently evaluated by the observables. The Koopman operator is linear but infinite-dimensional, and some form of truncation to a finite-dimensional approximation is required in practice, thereby introducing a trade-off between accuracy and dimensionality.

The linear operator has a spectral decomposition of tuples $\{(\mu_i, \phi_i(x), v_i)\}_{i=1}^{\infty}$ of eigenvalues, eigenfunctions, and modes. The eigenvalues and eigenfunctions satisfy the condition that the corresponding eigenvalue determines the dynamics or time evolution of a specific eigenfunction

$$[U^k \phi_i](x) = \mu_i^k \phi_i(x), \quad (6)$$

and the modes map the linear evolution of eigenfunctions (6) into the original observables by weighting the eigenfunctions

$$f(x) = \sum_{i=1}^{\infty} v_i \phi_i(x). \quad (7)$$

The importance and advantages of the Koopman operator and the diagonalization provided by the spectral decomposition are highlighted by Equations (6) and (7). Hence, the evolution of observables with respect to the spectral decomposition of the Koopman operator is

$$f(T^k(x)) = [U^k f](x) = \sum_{i=1}^{\infty} v_i \mu_i^k \phi_i(x). \quad (8)$$

2.3. Extended Dynamic Mode Decomposition Algorithm

The objective of the EDMD algorithm [9] is to obtain a finite and discrete-time approximation of the Koopman operator based on sampled data of the underlying dynamical system [16,17].

The EDMD algorithm that approximates the discrete-time Koopman operator of (1) requires N pairs of snapshot data, either from the numerical solution of an existing mathematical model or from experimental measurements at a specific sampling time Δt . The sets of snapshot pairs $\{(x_i, y_i)\}_{i=1}^N$ satisfy the relationship $y_i = x_{i+1} = T(x_i)$ and their definition in matrix form is

$$X = [x_1 \quad x_2 \quad \cdots \quad x_N], \quad Y = [y_1 \quad y_2 \quad \cdots \quad y_N]. \quad (9)$$

The “extended” part of the EDMD algorithm consists in the approximation of the Koopman operator on a “lifted” space of the state variables, rather than approximating the state space as in the dynamic mode decomposition algorithm [18]. The “lifting” procedure consists in evaluating the state of the system with the set of observables $\Psi: \mathcal{M} \rightarrow \mathbb{C}^{d \times 1}$; $\Psi(x) = [\psi_1(x), \dots, \psi_d(x)]^\top$. However, the choice of observables for a particular system is an open question; common choices include orthogonal polynomials [19], radial basis functions [9], or an arbitrarily constructed set with polynomial elements, trigonometric functions, logarithmic functions, or any combination of them [20]. Our choice here is a low-rank orthogonal polynomial basis [21–24], where every element of $\Psi(x)$ is the tensor product of n univariate orthogonal polynomials from a set $\{\pi_{\alpha_j}(x_j)\}_{\alpha_j=0}^p$, where α_j is the degree of the polynomial on the j th component of the x vector, and p is the maximum degree of the polynomial. Every component of $\Psi(x)$ is provided by

$$\psi_l(x) = \prod_{j=1}^n \pi_{\alpha_j}(x_j). \quad (10)$$

Additionally, the low-rank orthogonal polynomial basis comes from the truncation scheme for the non-empty finite set of indices α , where the choice of indices is based on q -quasi-norms (the quantity $\|\cdot\|_q$ is not a norm, as it does not satisfy the triangle inequality), as follows: $\alpha = \{\alpha \in \mathbb{N}^n : \|\alpha\|_q \leq p\}$, with

$$\|\alpha\|_q = \left(\sum_{i=1}^n \alpha_i^q \right)^{\frac{1}{q}}. \quad (11)$$

This choice of truncation scheme has three advantages. First, it reduces the number of elements in the set of observables $\Psi(x)$, and therefore handles the curse of dimensionality problem when the dimension of the state space grows and available computational resources are limited. The second advantage is empirical, and refers to the numerical stability of the least squares solution for the approximation of the Koopman operator. Having a reduced orthogonal basis improves the condition number of the matrices involved in the computation. The third advantage is the reduction of the amount of data required to compute the approximation. In comparison with other types of observable functions, the decomposition provides accurate approximations with significantly fewer data points.

To illustrate the advantages of the methodology in comparison with the previous literature on the subject, the authors of [9] used 1000 radial basis functions to obtain the approximation of a benchmark problem, e.g., the Duffing oscillator with two basins of attraction. About 10^4 data-points are required to compute the approximation of the Koopman operator and to subsequently find the level sets of the stable manifold and divide the two regions of attraction. In comparison, our previous work [25] allows the same results to be achieved with a polynomial basis of 25 elements and a dataset of about 10^3 data points.

For a detailed description of the use of p - q quasi-norms for the EDMD algorithm, we refer to previous works by the authors [23,24].

Furthermore, the approximation of the discrete-time d -dimensional Koopman operator U_d satisfies the following condition [9]:

$$\Psi(Y) = U_d \Psi(X) + r(X), \quad (12)$$

where $r(X) \in \mathcal{F}$ is the residual term to be minimized in order to find U_d . This minimization accepts a closed-form solution within the least mean squares problem, where the objective function has the form

$$\|r(X)\|^2 = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|\Psi(y_i) - U_d \Psi(x_i)\|_2^2, \quad (13)$$

and the solution is

$$U_d \triangleq AG^+, \quad (14)$$

where $G, A \in \mathbb{C}^{d \times d}$ are square matrices provided by

$$G = \frac{1}{N} \sum_{i=1}^N \Psi(x_i) \Psi(x_i)^\top, \quad (15)$$

$$A = \frac{1}{N} \sum_{i=1}^N \Psi(y_i) \Psi(x_i)^\top. \quad (16)$$

The dimension of the basis of observable functions and the parameters of each individual function have an impact on the stability of the numerical algorithm. In general, poor selection of the observable functions leads to a poor condition number on the regression matrix G .

Hence, it is often necessary to use a Moore–Penrose pseudo-inverse function. Even though this method provides a solution, it leads to propagation of the error in the approximation of the Koopman operator, especially when calculating the spectral decomposition of the Koopman matrix \mathcal{U}_d to obtain the eigenfunctions and modes. The use of orthogonal polynomials and p - q quasi-norm reduction yields a better conditioned regression matrix G .

The finite-dimensional and discrete-time approximation of the Koopman operator from the EDMD algorithm has eigenvalues $M = \text{diag}(\mu_1, \dots, \mu_d)$, right eigenvectors $\Xi = [\zeta_1, \dots, \zeta_d]$, and left eigenvectors, $\Xi^{-1} = W^*$. Furthermore, the approximation of the eigenfunctions $\Phi = [\phi_1, \dots, \phi_d]^\top$ comes from weighting the set of observables with the matrix of left eigenvectors [9,16]:

$$\Phi^\top(x) = \Psi(x)^\top W^*. \quad (17)$$

Using (7), the recovery of the original observables $\Psi(x)$ from the set of eigenfunctions is provided by

$$\Psi(x) = \Xi\Phi(x). \quad (18)$$

Using (8), the time evolution of observables according to the spectrum of the Koopman operator is provided by

$$\Psi(T^k(x)) = \Xi M^k \Phi(x). \quad (19)$$

The common practice to recover the state is to include the functions that capture each of the states in the set of observables, i.e., $\psi_i(x) = x_i$. The value of the time evolution of the states is the value of these particular elements of $\Psi(x)$. The matrix $B \in \mathbb{R}^{d \times n}$ recovers these values from (19), where B is a matrix of unit vectors \hat{e}_i indexing the n observables that captures every one of the states, i.e., $f_i(x) = x_i$. As for a polynomial basis, these elements are not always present, and B is a matrix of unit vectors indexing injective observables. To clarify the importance of using the inverse of injective observables against other methods to recover the state, we refer to previous works by the authors [23,24].

The evolution of observables (19) is used to define the state evolution map according to the approximation of the Koopman decomposition $\{(\mu_i, \phi_i(x), v_i)\}_{i=1}^d$, i.e., evolving (1) k times from an initial condition x_0

$$x(k) = T^k(x_0) = \Psi_B^{-1} \left(B^\top \Xi M^k \Phi(x_0) \right), \quad (20)$$

where $\Psi_B^{-1}(x)$ denotes the inverse of the injective observables selected by B .

Note that in addition to the forward evolution of the states, the discrete-time approximation provides their evolution in reverse time:

$$x(-k) = \Psi_B^{-1} \left(B^\top \Xi M^{-k} \Phi(x_0) \right). \quad (21)$$

In conclusion, the EDMD allows for a linear discrete-time representation on an extended space of a nonlinear system, along with the advantages that the induced spectrum and evolution maps provide. These advantages are the core of this analysis of the ROA.

For completeness, and in order to state both the advantages and the shortcomings of this approximation, note that the EDMD algorithm assumes knowledge of the full state of the system, and does not take noisy signals into account. Therefore, the application of the algorithm in real-world scenarios requires complementing its development with an observer able to handle these issues.

3. Evaluation of the ROA Using EDMD

In order to obtain the ROA of asymptotically stable fixed points in a data-driven fashion using the EDMD approximation of the Koopman operator, it is necessary to accomplish numerous tasks. First, it is necessary to find the location of the fixed points and determine their stability, especially for the asymptotically stable and type-1 points. Subsequently,

the spectrum of the Koopman operator approximation can be analyzed in keeping with the theoretical concepts from Section 2.1 in order to approximate the ROA of the asymptotically stable fixed points under multistability phenomena.

In order to accomplish this objective, several assumptions must hold in order for the algorithm to be applied. First, the EDMD algorithm must have knowledge of the full state and enough trajectories from the dynamical system in order to have an accurate approximation of the operator. This condition can be checked using the empirical error by comparing the orbits of the system from (1) with the discrete-time Koopman state evolution map (20). Second, assumptions (A1–A3) introduced in Section 2.1 must hold. In general, however, it is necessary to know the differential equation model in order to check assumptions (A2) and (A3). Due to this difficulty, we limit our analysis to mass action systems for illustration of the results in Section 4. This type of system inherently satisfies the assumptions [1].

3.1. Fixed Points Approximation

The first step of the analysis consists in locating the fixed points of the hyperbolic system to assess their stability. A fixed point of the discrete-time nonlinear system (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, as in (2).

To approximate the fixed points, consider the forward and backward state evolution maps of a dynamical system in discrete time provided by the Koopman operator approximation in (20) and (21). If the evolution of these mappings is invariant, i.e., maps to themselves, then the system is at a fixed point. To accurately approximate these points, we propose to formulate a minimization problem where the objective function is the Euclidean norm of the comparison between the state and its approximate discrete-time evolution (20) [25].

Lemma 1 (Fixed points). *Let $(\mathcal{M}; T(x); k)$ be a dynamical system that admits a Koopman operator approximation $(\mathcal{F}_d; \mathcal{U}_d; k)$, and let x^* be a fixed point of $T(x)$. Then,*

$$x^* = \operatorname{argmin}_x \left\| \Psi_B^{-1} \left(B^\top \Xi M^k \Phi(x_0) \right) - x \right\|_2^2. \quad (22)$$

Proof. Let $T^k(x) = \Psi_B^{-1} (B^\top \Xi M^k \Phi(x_0))$ be the flow map of $(\mathcal{M}; T(x); k)$, assume $k = K$ is finite, and define $b(x) = \Psi_B^{-1} (B^\top \Xi M \Phi(x))$ using Definition 2 and $b(x)$; then, the least squares problem

$$J(x) = \frac{1}{2} \|b(x) - x\|^2 \quad (23)$$

with solution

$$x^* = \operatorname{argmin}_x (J(x)), \quad (24)$$

provides the location of the fixed points. \square

Remark 3. *Note that this procedure is possible in the portion of the state space where the data snapshots lie, and corresponds to the fixed points of the nonlinear underlying hyperbolic dynamical system. If the system is not hyperbolic, this condition does not hold.*

Remark 4. *This lemma is only accountable for the location of the fixed points in the state space, and does not provide information about their stability.*

Remark 5. *Equation (20) is a nonlinear discrete-time evolution mapping, and under the definition of fixed point, i.e., $\bar{T}^k(x^*) = x^*$, it is possible to obtain the approximation of the hyperbolic fixed points by solving*

$$\bar{x}^* \subseteq \{x \in \mathbb{R}^n : \bar{T}^k(x) - x = 0\}. \quad (25)$$

In practice, this procedure has two issues. First, as the nonlinear mapping (20) comes from a set of observables, the dimension and complexity of these functions affect the possibility of finding a solution in polynomial time. As a consequence, in higher-order polynomial expansions, even if it is 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. Solving a nonlinear set of equations provides several solutions in the complex plane, some of which correspond to the real-valued solutions 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; therefore using this definition to approximate the fixed points is not feasible.

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

3.2. Stability of Fixed Points

The traditional way of establishing the local stability of a hyperbolic fixed point is through analysis of the Jacobian matrix of the system (1) as evaluated at the fixed point. Our proposed approach is to analyze the state evolution map from the discrete approximation of the Koopman operator in the same way. In other words, the stability comes from the linearization of (20), which is a nonlinear mapping of the state.

Definition 1 (Stability). Let $(\mathcal{M}; T(x); k)$ be a dynamical system that admits a Koopman operator approximation $(\mathcal{F}_d; U_d; k)$. Define $b(x) = \Psi_B^{-1}(B^\top \Xi M^k \Phi(x))$ from the state evolution map (20); then, the local linearization of $b(x)$ at a given point x^* is

$$\begin{aligned} x(k+1) &= \left[\frac{\partial T(x)}{\partial x_1} \dots \frac{\partial T(x)}{\partial x_n} \right] \Big|_{x^*} x(k) \\ &= \left[\frac{\partial b(x)}{\partial x_1} \dots \frac{\partial b(x)}{\partial x_n} \right] \Big|_{x^*} x(k) \\ &= Hx(k). \end{aligned} \quad (26)$$

The local stability of a fixed point x^* according to the eigenvalues $\{\mu_i\}_{i=1}^n$ from the spectral decomposition of the matrix H is provided as follows:

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

Remark 6. Note that the inequalities are strictly greater-than or less-than, which is in accordance with the assumption of hyperbolicity.

With information about the location and stability of the fixed points obtained, the main result of this paper can be formulated, i.e., the approximation of the boundary of the ROA via eigenfunctions of the discrete-time Koopman operator.

3.3. Approximation of the ROA Boundary

This section describes the approximation of the ROA of asymptotically stable fixed points in a data-driven approach using the discrete-time approximation of the Koopman operator. The claim of the main result of this study is that the level sets of a unitary eigenfunction provide the boundary of the ROA of the asymptotically stable fixed points. This holds from the following set of premises:

1. A system that admits a Koopman operator transformation has an infinite set of eigenfunctions.
2. There exist nontrivial eigenfunctions that have an associated eigenvalue equal to one (unitary eigenfunctions).
3. A unitary eigenfunction is invariant along the trajectories of the system (per Equation (6)).

4. The trajectories of the system are level sets of a unitary eigenfunction.
5. The stable manifold of a type-1 saddle point in an n -dimensional system is an $(n - 1)$ -dimensional hypersurface composed by the union of the trajectories that converge to the point (from Equation (4)).
6. The level set of a unitary eigenfunction at a type-1 fixed point is the stable manifold of that point.
7. The boundary of the ROA of an asymptotically stable fixed point is the union of the stable manifolds of the type-1 fixed points in the boundary (from theorem 1).

In other words, if it is possible to capture the $(n - 1)$ -dimensional hypersurfaces that converge to the type-1 saddle points in the boundary of the ROA via the Koopman operator, then it is possible to find the boundary of the ROA. If there are eigenfunctions that are invariant along the trajectories of the system, then evaluating these eigenfunctions on the type-1 saddle points of the system provides the constant values for which the unitary eigenfunctions capture the stable manifold of the saddle points, i.e., the level sets of the unitary eigenfunction. A level set of an arbitrary function $h(x): \mathbb{R}^n \rightarrow \mathbb{C}$ for any constant $c \in \mathbb{C}$ is $\Gamma_c(h(x)) = \{x \in \mathbb{R}^n : h(x) = c\}$.

To guarantee the existence of nontrivial unitary eigenfunctions, we consider the following property of the eigenfunctions of the Koopman operator (see Property 1 in [14] for its analogues in continuous time).

Property 1. Suppose that $\phi_1, \phi_2 \in \mathcal{F}_d$ are Koopman eigenfunctions of an arbitrary system with associated eigenvalues μ_1 and μ_2 , respectively. If $\bar{\phi}(x) = \phi_1^{k_1}(x)\phi_2^{k_2}(x) \in \mathcal{F}_d$ for constants $k_1, k_2 \in \mathbb{C}$, then $\bar{\phi}(x)$ is an eigenfunction constructed upon the product of two arbitrary eigenfunctions with the associated eigenvalue $\bar{\mu} = \mu_1^{k_1}\mu_2^{k_2}$.

Proof. It follows from condition (6) that

$$[U_d \bar{\phi}](x) = [U_d(\phi_1^{k_1}\phi_2^{k_2})](x) \quad (27)$$

$$= ([U_d \phi_1](x))^{k_1} ([U_d \phi_2](x))^{k_2} \quad (28)$$

$$= \mu_1^{k_1} \phi_1^{k_1}(x) \mu_2^{k_2} \phi_2^{k_2}(x). \quad (29)$$

□

This means that with a finite approximation of the Koopman operator, there are infinite possible dependent eigenfunctions, showing that Premise 1 holds. Moreover, from the previous construction, it is possible to find the complex constants k_1 and k_2 for eigenvalues μ_1 and μ_2 such that $\bar{\mu} = 1$ and $\bar{\phi}(x) = \phi_+(x)$, where ϕ_+ denotes an eigenfunction with a unitary associated eigenvalue. From (6), it follows that $\phi_+(x)$ is an eigenfunction that is invariant along the trajectories of the system, as

$$[U_d^k \phi_+](x) = \phi_+(x), \quad (30)$$

which demonstrates that Premises 2 and 4 hold, and explains Premise 3. As a result of the previous development, we can state our main result:

Theorem 2. If there exists a dynamical system $(\mathcal{M}; T(x); k)$ with multiple stable points that admits a Koopman operator approximation $(\mathcal{F}_d; U_d; k)$, then the approximation of the stable manifold of the type-1 saddle points \hat{x}^* in the boundary of the ROA is the level set of a unitary eigenfunction $\phi_+(x)$ with a constant value taken from evaluation of the unitary eigenfunction at the saddle point at the boundary, i.e.,

$$W^s(\hat{x}^*) = \Gamma_{c_{\phi_+(\hat{x}^*)}} \phi_+(x) \quad (31)$$

$$\subseteq \{x \in \mathbb{R}^n : \phi_+(x) = c_{\phi_+(\hat{x}^*)}\}. \quad (32)$$

Proof. From the definition of a stable manifold of a type-1 saddle point (4) in Section 2.1,

$$W^s(\hat{x}^*) = \{x \in \mathbb{R}^n : \lim_{k \rightarrow \infty} T^k(x) = \hat{x}^*\}. \quad (33)$$

Evaluating this manifold with an arbitrary eigenfunction $\phi(x)$ and using the condition of the Koopman operator in (5) provides

$$\phi(W^s(\hat{x}^*)) \approx \{x \in \mathbb{R}^n : \lim_{k \rightarrow \infty} \phi(T^k(x)) = \phi(\hat{x}^*)\} \quad (34)$$

$$\approx \{x \in \mathbb{R}^n : \lim_{k \rightarrow \infty} [U_d^k \phi](x) = \phi(\hat{x}^*)\}. \quad (35)$$

Hence, for this arbitrary eigenfunction the associated eigenvalue must be unitary in order for the equality to hold. Moreover, from (30) it is clear that the time evolution of an eigenfunction with a unitary associated eigenvalue is invariant along the trajectories of the system; in other words, it is independent of time evolution. These two developments imply that

$$\phi(W^s(\hat{x}^*)) = \{x \in \mathbb{R}^n : \phi_+(x) = c_{\phi_+(\hat{x}^*)}\}. \quad (36)$$

Note that the right-hand part of (36) is the definition of a level set for an eigenfunction with a unitary associated eigenvalue $\phi_+(x)$. As the trivial eigenfunction $\phi_{\mu=1} = 1$ belongs to the set of eigenfunctions with unitary associated eigenvalues, the left-hand side holds for a subset of the unitary eigenfunctions $\phi_+(x)$. Therefore, the stable manifold of a type-1 saddle point at the boundary of the ROA is

$$\Gamma_{c_{\phi_+(\hat{x}^*)}} \phi_+(x) \subseteq \{x \in \mathbb{R}^n : \phi_+(x) = c_{\phi_+(\hat{x}^*)}\}. \quad (37)$$

□

In summary, the spectral decomposition of the discrete-time approximation of the Koopman operator can be used to find the nontrivial unitary eigenfunction, which, when evaluated at the type-1 points, characterizes the ROA of the asymptotically stable points.

3.4. Algorithm

The approach presented in Sections 2.3, 3.1 and 3.3 for obtaining the attraction regions of asymptotically stable fixed points is summarized in Algorithm 1, for which the following assumptions hold:

- B1: The system under consideration has multiple hyperbolic fixed points.
- B2: At least one of the fixed points is asymptotically stable.
- B3: There are sufficient snapshot data (either from measurements of the real system or from numerical simulation) to construct a discrete approximation of the Koopman operator. This condition can be checked using the empirical error, which is a comparison between the data from the numerical integration of the system dynamics and the state evolution map from the approximation of the discrete-time Koopman operator.

Algorithm 1 ROA with data-driven discrete Koopman operator.

Ensure: Assumptions (B1–B3) are satisfied

- 1: **Data:** $\{x_i, y_i\}_{i=1}^N$ **return** $\partial R_A(x_s^*)$
- 2: **procedure** ROA($\{x_i, y_i\}_{i=1}^N, q, p$)
- 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: $E_q \leftarrow \text{size}(x^*)$
- 7: **for** $i \leftarrow 1, E_q$ **do**
- 8: $H_i \leftarrow \left[\frac{\partial b(x)}{\partial x_1} \dots \frac{\partial b(x)}{\partial x_n} \right] \Big|_{x_i^*}$
- 9: $\{\mu_{i,j}\}_{j=1}^n \leftarrow \text{eig}(H_i)$
- 10: **if** $|\mu_{i,j}| < 1$ for all $j = 1, \dots, n$ **then**
- 11: $x_i^* \leftarrow x_s^*$
- 12: **else if** $|\mu_{i,j}| < 1$ for some $j = 1, \dots, n$ **then**
- 13: $x_i^* \leftarrow \hat{x}^*$
- 14: **else**
- 15: $x^* \leftarrow x_u^*$
- 16: **end if**
- 17: **end for**
- 18: $\hat{E}_q \leftarrow \text{size}(\hat{x}^*)$
- 19: $E_s \leftarrow \text{size}(x_s^*)$
- 20: $\phi_+(x) \leftarrow \{\phi_1^{k_1}(x)\phi_2^{k_2}(x) : \mu_1^{k_1}\mu_2^{k_2} = 1\}$
- 21: **for** $i \leftarrow 1, E_s$ **do**
- 22: **for** $j \leftarrow 1, \hat{E}_q$ **do**
- 23: $W_j^s(\hat{x}_j^*) \leftarrow \{x \in \mathbb{R}^n : \phi_+(x) = \phi_+(\hat{x}_j^*)\}$
- 24: **end for**
- 25: $\partial R_A(x_i^*) = \cup W^s(\hat{x}_j^*)$
- 26: **end for**
- 27: **end procedure**

4. Simulation Results

In order to test the reliability of the methodology and algorithm, they were applied to a model of competitive exclusion with two state variables. This model has the advantage of having an analytical description of the ROA under certain parameters, and is suitable for graphically showing the effect of eigenfunctions with a unitary eigenvalue. Next, the reliability of the algorithm in a higher-dimensional system composed of five state variables, i.e., a mass-action kinetics (MAK) model is demonstrated. This latter example shows the advantage of not having to calculate level sets or handle complex geometries of the $(n - 1)$ -dimensional stable manifold hyperplanes. The population model and the mass-action kinetics systems are suitable for this analysis because of their geometric properties. They are non-negative compartmental systems that, depending on the parameterization, have hyperbolic fixed points that satisfy the Hartman–Grobman theorem [1]. The hyperbolicity of the fixed points implies that their unstable and stable manifolds intersect transversely at the saddle points. Therefore, Assumptions (A1–A3) are satisfied in this kind of system under the right parameterization.

The numerical integration of the systems from a random set of initial conditions provides the dataset of trajectories for the algorithm. A subset is used to calculate the Koopman operator's approximation via the EDMD algorithm (training set) and another subset is used for testing the algorithm's accuracy. The algorithm empirical error criterion for a number N_s of test trajectories and a length K_i for every one of those trajectories is

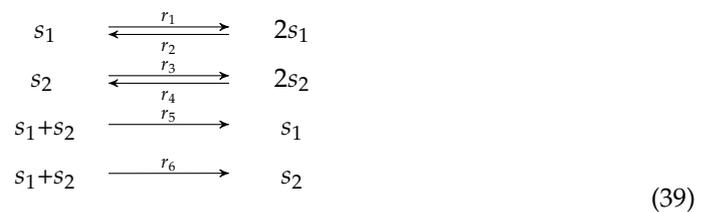
$$e = \frac{1}{\sum_{i=1}^{N_s} K_i} \left(\sum_{i=1}^{N_s} \sum_{k=1}^{K_i} \frac{|T^k(x_i) - U_d^k g(x_i)|}{|T^k(x_i)|} \right), \quad (38)$$

where this error serves to determine the best p - q parameters from a sweep over the different values and to verify whether assumption B3 holds.

From the test set, it is possible to evaluate the convergence from every initial condition. Comparing the real convergence with the classification results of the algorithm provides the accuracy of the method according to the number of misclassified initial conditions over the total.

4.1. A Competition Model

Consider the following network describing the dynamics of a population model where two species compete for the same resource:



where s_1 and s_2 are the competing species. The dynamics of the species depend on six parameters: r_1 and r_3 respectively describe the population growth rates of each species, r_2 and r_4 represent the logistic terms accounting for the competition between members of the same species (resulting in carrying capacities for the environment of r_1/r_2 and r_3/r_4 , respectively), r_5 describes the competitive effect of species s_2 on species s_1 , and r_6 describes the opposite competitive effect between species. The values of these constants determine the potential outcome of the competition. Depending on their values, there can be co-existence or exclusion of one species against the other. The case of interest is the exclusionary 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 at which one species completely takes over the other. The differential equations that describe the population (39) are

$$\dot{x}_1 = r_1 x_1 - r_2 x_1^2 - r_5 x_1 x_2 \quad (40)$$

$$\dot{x}_2 = r_3 x_2 - r_4 x_2^2 - r_6 x_1 x_2, \quad (41)$$

where the parameterization in order 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 points at $(0, 2)$ and $(2, 0)$, defined as x_B^* and x_C^* , respectively. The geometry of this problem provides a simple representation of a type-1 saddle stable manifold, which is the line $x_1 = x_2$, thereby providing a closed formulation for comparison with the stable manifold generated by the construction of the eigenfunction with an associated eigenvalue equal to one.

The numerical integration of the system from 200 uniformly distributed random initial conditions with a $\Delta t = 0.1$ over $x_1, x_2 \in [0 \ 2]$ and stopped upon convergence provides the datasets for approximating and testing the operator via the EDMD algorithm. Each set has 50% of the available trajectories, for a total of 7285 datapoints each.

The application of the algorithm on the training set with a Laguerre polynomial basis and a truncation scheme sweeping over several p - q values provides the best performance when $q = 1.1$ and $p = 3$. This selection produces a polynomial basis of 13 elements of order equal or less than 3 and a Koopman operator of order 13. Figure 1 shows the retained indices after implementing the truncation scheme.

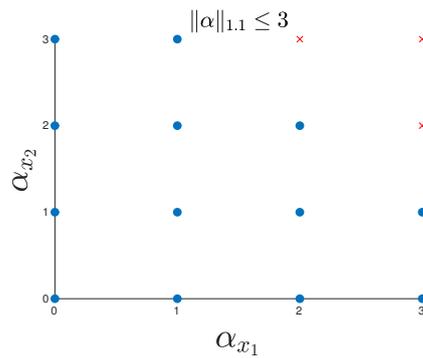


Figure 1. Retained indices for the approximation of the Koopman operator for the competition model when choosing $q = 1.1$ and $p = 3$.

The next step of the method is the location and stability of the fixed points. Their location via Lemma 1 provides an absolute error of 0.15%. Moreover, the method in Section 3.2 accurately provides their stability according to the linearization of the nonlinear state evolution map (20) and the analysis of the eigenvalues λ from the linearization evaluated at the identified fixed points. Table 1 summarizes these results.

Table 1. Competition model fixed points, location, and stability.

	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

Figure 2 depicts the comparison between the theoretical trajectories provided by integration of the differential Equations (40) and (41) and the state evolution map (20) from the same initial conditions. In addition, it depicts the comparison between the theoretical boundary of the attraction regions $x_1 = x_2$ and the boundary established by the level set of the constructed unitary eigenfunction from (32). The percentual error in the classification of the initial conditions, i.e., the number of misclassified initial conditions over the total, is 2%, while the mean absolute error between the boundary of the regions of attraction, i.e., the difference between the $x_1 = x_2$ theoretical boundary and the boundary identified from the level set of the unitary eigenfunction, is 3%.

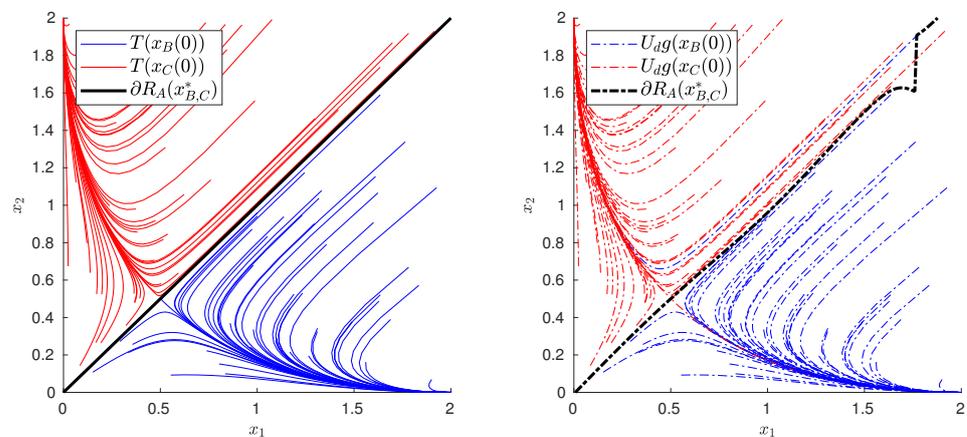


Figure 2. Trajectories and boundary of the asymptotically stable points of (left) the system of differential equations and (right) the Koopman operator and unitary eigenfunction.

Figure 3 shows four eigenfunctions of the Koopman operator, including the trivial eigenfunction with a value that is constant in the whole state space. This eigenfunction is commonly the result of the EDMD algorithm, and does not provide any useful information about the system. Therefore, it is necessary to obtain a nontrivial unitary eigenfunction from (29). To achieve this objective, setting $\mu_1^{k_1} \mu_2^{k_2} = 1$ and computing a solution for k_1 and k_2 results in the desired unitary eigenfunction. However, caution should be exercised with this calculation as there are an infinite number of solutions for the choice of constants. In practice, the method used is to first set one and then calculate the other.

Figure 3 depicts the resulting unitary eigenfunction that captures 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 $\mu_2 = 0.83$. Inspecting the near-unitary eigenfunction with $\mu_1 = 1.07$ shows that it could potentially be used for the classification on its own, and indeed this eigenfunction does provide good results, although not as accurate as the construction of the unitary eigenfunction. At this point, it is not clear how to select the base eigenfunctions properly for the construction of the unitary eigenfunction, and although the algorithm provides accurate results when the base eigenfunctions for the construction of the unitary eigenfunction do not have real-valued eigenvalues, the lack of real-valued eigenfunctions for use in construction degrades the accuracy.

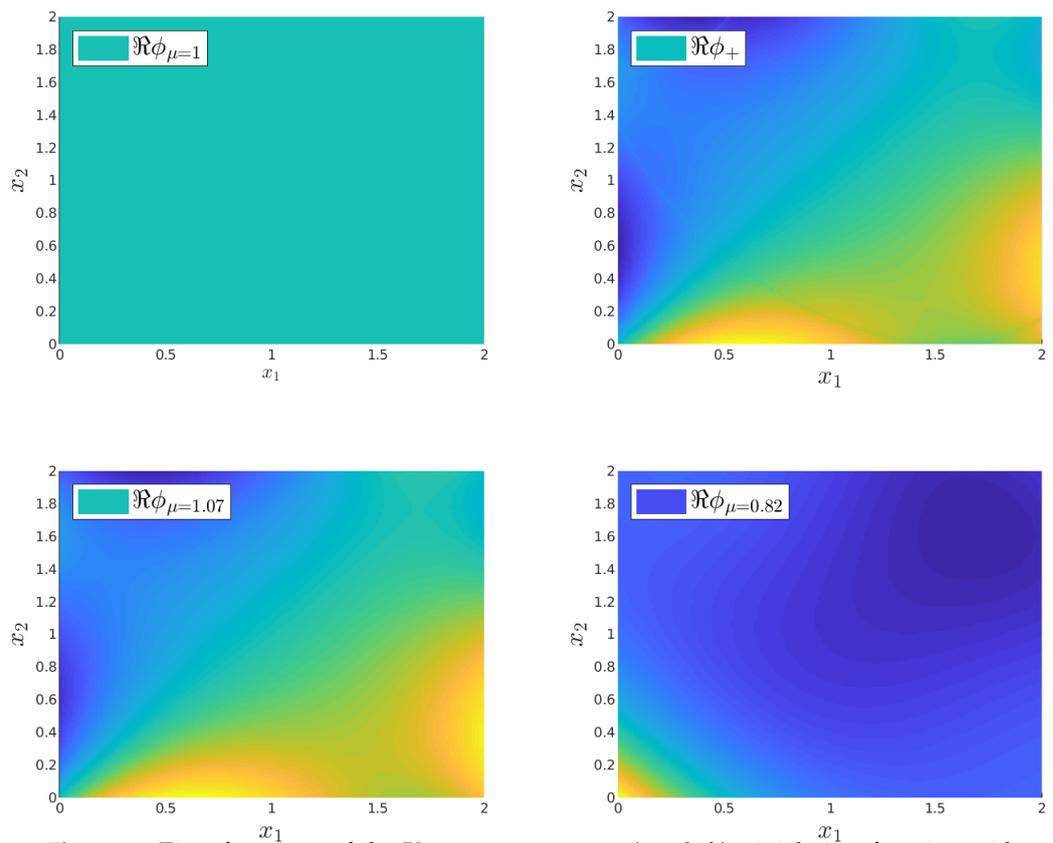
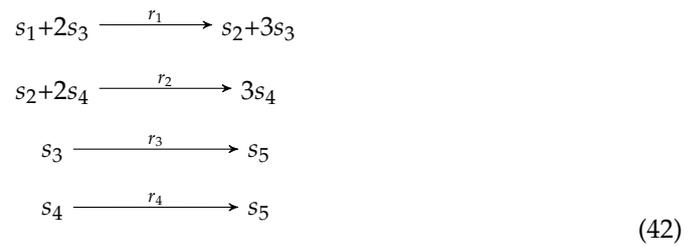


Figure 3. Eigenfunctions of the Koopman operator: (top left) trivial eigenfunction with $\mu = 1$; (top right) constructed eigenfunction ϕ_+ ; (bottom left) first eigenfunction for constructing ϕ_+ ; (bottom right) second eigenfunction for constructing ϕ_+ .

4.2. Mass Action Kinetics

Consider the following network, which describes an auto-catalytic replicator in a continuous-flow stirred tank reactor:



here, there are two species, s_3 and s_4 . Species s_3 consumes substrate s_1 to reproduce and produces the resource s_2 as a byproduct, which is suitable for the reproduction of the second species s_4 . Finally, s_5 represents the dead species from both groups. The constants $r_1 > r_3$ and $r_2 > r_4$ are the paired of replication rate and species death, respectively. Considering d as the dilution rate, that is, the material exchange with the environment, the dynamics of the network (42) are described by the differential equations

$$\dot{x}_1 = -r_1 x_1 x_3^2 + d - dx_1 \tag{43}$$

$$\dot{x}_2 = +r_1 x_1 x_3^2 - r_2 x_2 x_4^2 - dx_2 \tag{44}$$

$$\dot{x}_3 = +r_1 x_1 x_3^2 - r_3 x_3 - dx_3 \tag{45}$$

$$\dot{x}_4 = +r_2 x_2 x_4^2 - r_4 x_4 - dx_4 \tag{46}$$

$$\dot{x}_5 = +r_3 x_3 + r_4 x_4 - dx_5, \tag{47}$$

where the substrate s_1 is the only component of the input flow. The value for the vector of the reaction rates is $k = [7.5 \ 0.3 \ 0.05]^\top$, yielding five fixed points: three asymptotically stable points (the working point, defined as x_A^* , where the two species thrive and coexist, a point where species s_3 thrives and species s_4 washes out, defined as x_C^* , and a wash-out 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 ROA of the working point using the unitary eigenfunction. This example highlights the contributions of this work, as the dimension of the state is 5, instead of being a two-dimensional or three-dimensional system that can be sliced for analysis. The stable manifold of the saddle points is a four-dimensional hypersurface. Therefore, it has a complex geometry, and the criterion for classifying the test set of initial conditions is not trivial. In particular, we show that it is not necessary to deal with complex geometries or higher-dimensional hyperplanes to perform the classification.

The numerical integration of the differential Equations (43)–(47) from 360 randomly distributed initial conditions and $\Delta t = 0.1$ generates the set of orbits for training and testing the algorithm. From this set of orbits, 50% are used to approximate the operator via the EDMD algorithm, while the other 50% to test the accuracy of the state evolution (20) map and the accuracy of the classification, where each set has 79,948 points. Additionally, sweeping over several p - q values provides the best performance for the truncation scheme when $q = 0.8$ and $p = 4$, which leads to an approximation of the Koopman operator of order 163, and thus a set of 163 eigenfunctions, eigenvalues, and modes. From the set of eigenfunctions, there are two eigenfunctions with real eigenvalues closest to one, which are $\mu_1 = 1.00008$ and $\mu_2 = 0.99983$. Considering that the eigenvalues associated with these eigenfunctions are close enough to one, these unitary eigenfunctions are sufficient to perform the analysis.

The identification of the fixed points via Lemma 1 provides an absolute error of 0.15%, and their stability according to the method in Section 3.2 provides an accurate description. Table 2 shows the results from the theoretical and algorithmic location of the fixed points, where the error is the percentual difference between the theoretical and the experimental magnitudes of the vectors that provide the location of the fixed points, i.e.,

$$\text{Error}\%_i = \frac{\| \|x_i^*\| - \|\bar{x}_i^*\| \|}{\|x_i^*\|} \times 100. \tag{48}$$

Table 3 shows the results of taking the norm of the eigenvalues λ from the Hessian matrix of (20) evaluated at each of the previously identified fixed points.

Table 2. Location of 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)	0.00
x_B^*	(0.21, 0.67, 0.30, 0.07, 0.47)	(0.23, 0.62, 0.30, 0.11, 0.49)	1.82
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, 0.30, 0.11, 0.00, 0.17)	2.86
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 3. Stability of fixed points.

	$ \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

The next step in the algorithm is finding a classification scheme that does not depend on the geometry of the four-dimensional hyperplanes that compose the stable manifold of the type-1 saddle points. For this purpose, we use a saddle classifier [24]. This type of classifier divides the set of initial conditions $x_0 \in \mathcal{M}$ of the testing trajectories into a subset $X \subset \mathcal{M}$ and its complement X' , where the evaluation of an arbitrary initial condition with a unitary eigenfunction compared to the evaluation of the type-1 point with the same unitary eigenfunction provides the criterion for the division:

$$X \subseteq \{x_0 \in \mathbb{R}^n : \Re\phi_+(x_0) \geq \Re\phi_+(\hat{x}^*)\}, \quad (49)$$

where \Re denotes the real part of the evaluation. This provides a simple algebraic criterion for the classification of initial conditions into their respective ROAs.

Figure 4 depicts the result of performing the evaluation of the saddle points and the initial conditions of the test set with the unitary eigenfunctions; here, ϕ_{+1} corresponds to the eigenfunction with $\mu_{+1} = 1.00008$ and ϕ_{+2} corresponds to the eigenfunction with $\mu_{+2} = 0.99983$. The test set is carefully chosen such that there are 60 initial conditions that converge to each of the three asymptotically stable fixed points. Furthermore, the division of this set into two subsets results in one set of 90 testing initial conditions when performing the analysis and another set of 90 initial conditions to perform final testing of the classification. Indexing over the analysis set provides the horizontal axis in Figure 4. Moreover, the vertical axis shows the result of evaluating each of the 90 different initial conditions with the trivial unitary eigenfunction ϕ_{1+} , where it is clear that the trivial eigenfunction does not provide any information.

The other two unitary eigenfunctions do provide useful information about the classification scheme. Using ϕ_{+1} , it is clear that evaluating the saddle point \hat{x}_B^* is an accurate criterion for the classification of initial conditions that converge to the asymptotically stable point x_A^* , i.e.,

$$X_A \subseteq \{x_0 \in \mathbb{R}^5 : \phi_{+1}(x_0) \leq \phi_{+1}(\hat{x}_B^*)\}. \quad (50)$$

Furthermore, when performing the same evaluation process with the second unitary eigenfunction ϕ_{+2} it is clear that evaluating the saddle point \hat{x}_D^* is an accurate comparison criterion for classifying the initial conditions that converge to the wash-out point x_E^* , i.e.,

$$X_E \subseteq \left\{ x_0 \in \mathbb{R}^5 : \phi_{+2}(x_0) \leq \phi_{+2}(\hat{x}_D^*) \right\}. \quad (51)$$

Note that the \Re notation is dropped for the two classifiers (50) and (51), as these are real-valued eigenfunctions. The combination of these two conditions specifies the criteria used to classify the initial conditions that converge to the different asymptotically stable fixed points. Figure 5 shows the result of applying the classification criteria over the second set of testing initial conditions. Again, the horizontal axis indexes over the three groups of 30 initial conditions that converge to each of the stable points. Moreover, the vertical axis shows the classification of the initial conditions according to the classifiers, where an incorrect classification has the correct value surrounding it. The evaluation of the points with the second unitary eigenfunction is only for illustration purposes. The final test over this set yields an error of 12% misclassified points.

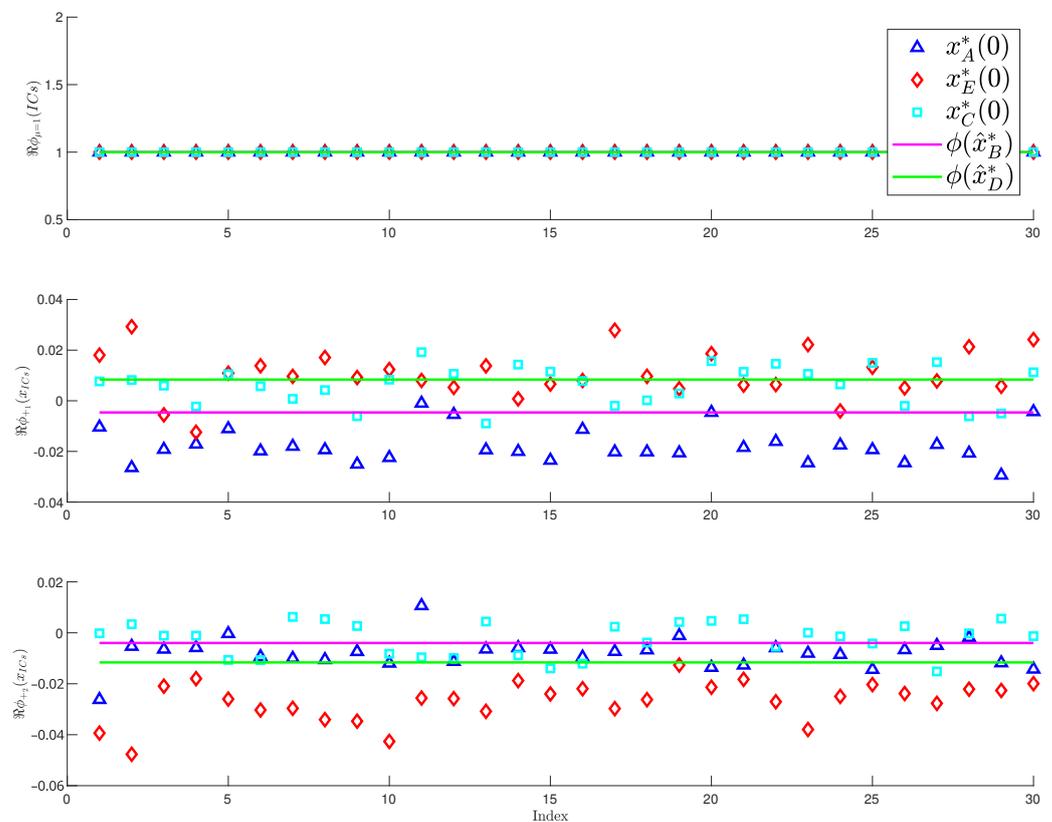


Figure 4. Eigenfunctions with unitary associated eigenvalue: **(top)** trivial; **(center)** $\mu > 1$; **(bottom)** $\mu < 1$.

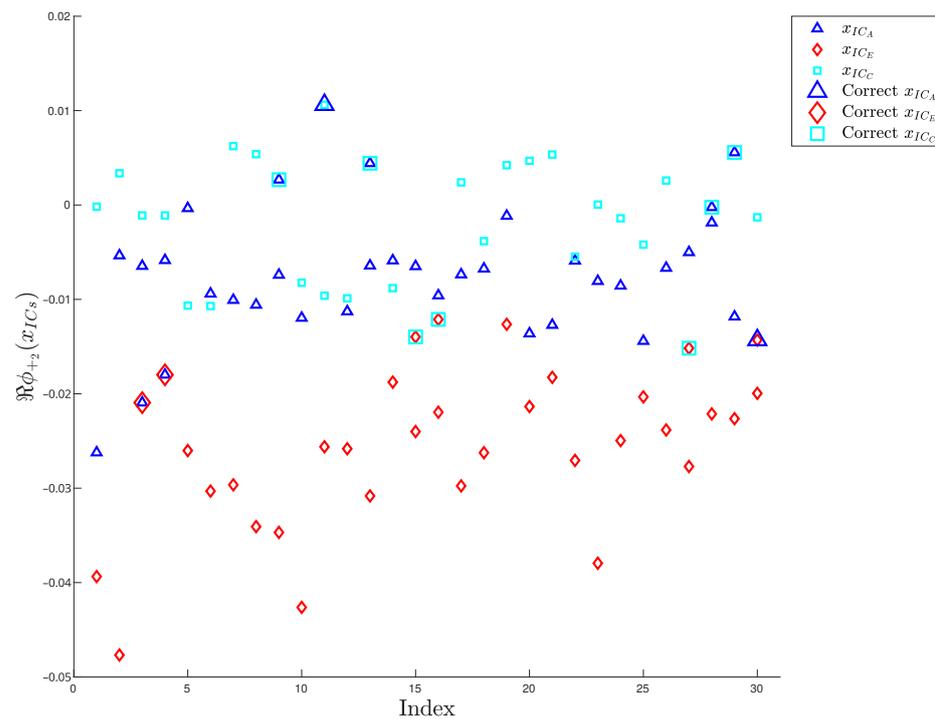


Figure 5. Classification of the initial conditions according to the evaluation of the eigenfunctions on the saddle points.

5. Critical Analysis and Perspectives

Even if we had the tools for extracting all kinds of information regarding the system, there are limitations to 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 and the choice of eigenfunctions to construct the one with the unitary associated eigenvalue are open questions for improving the algorithm and the analysis. The reliability of the algorithm depends on the satisfaction of assumption (B3), i.e., having enough data for the approximation of the operator. This data assumption is enough to perform the analysis, although it has underlying time and space implications. Increasing the dimension of the state necessarily increases the number of elements in the polynomial basis and the computational effort required to calculate the inverse in (14).

Furthermore, there is the problem of the amount and quality of data required to accurately approximate the Koopman operator. For the two case studies, the amount of data is sufficient to produce an accurate nonlinear model of the systems and the integration provides clean data devoid of noise. Moreover, to consider the proposed method over traditional identification techniques, it is necessary to develop experimental design techniques specifically for the calculation of the Koopman operator in order to reduce the necessary data. Additionally, the data from experimental setups are noisy, and often lack the measurements of the whole state. Therefore, it is necessary to complement the algorithm with a noise-filtering observer designed to handle these issues.

Furthermore, the relationship between the eigenfunctions and their associated eigenvalues remains open. There are other dynamical characteristics to be analyzed from this association, as the dynamic behavior of eigenfunctions with associated eigenvalues that match the local spectra of the fixed points is a potential source of information on the underlying dynamical system.

A possible improvement could be in 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 provide more information about the system for control purposes.

6. Conclusions

This paper deals with the analysis of nonlinear dynamical systems using the spectrum of the Koopman operator, more specifically, the construction of invariant sets called unitary eigenfunctions that capture the stable manifold of type-1 saddle points. The union of these stable manifolds is the boundary of the region of attraction of asymptotically stable fixed points, and the evaluation of these unitary eigenfunctions provides a simple algebraic criterion to determine the region of attraction of these stable points. Furthermore, all of this information is obtained through a data-driven approach without the need for the differential equation model. Based solely on data from the system, the algorithm can provide the location and local stability of the fixed points of the system. The extension of these data-driven tools can provide new methods for analyzing complex systems and the ability to reformulate traditional linear controller synthesis methods to work with the discrete-time Koopman operator approximation.

The data-driven representation of the Koopman operator provides the spectral decomposition that can complement traditional techniques for synthesis and control and provide a black-box tool for dealing with systems for which the exact dynamics are challenging to model and identify.

To improve this analysis, it remains necessary to explore the polynomial structure of the system and find relationships between the system and the polynomial bases that provide the approximation in order to optimize their selection. When coupled with more accurate eigenfunctions deduced from optimal truncation schemes, these relationships can allow eigenfunctions to be identified with additional information regarding the nonlinear dynamical system for control purposes, reducing the amount and quality of data necessary for the approximation.

Author Contributions: Conceptualization, C.G.-T., D.T.-C., E.M.-N. and A.V.W.; methodology, C.G.-T., D.T.-C., E.M.-N. and A.V.W.; software, C.G.-T.; writing—original draft preparation, C.G.-T. and A.V.W.; writing—review and editing, C.G.-T., D.T.-C., E.M.-N. and A.V.W.; supervision, E.M.-N. and A.V.W. All authors have read and agreed to the published version of the manuscript.

Funding: C.G.-T. Colciencias-Doctorado Nacional-647/2015 and D.T.-C.: Colciencias-Doctorado Nacional-727/2016.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chellaboina, V.; Bhat, S.; Haddad, W.; Bernstein, D. Modeling and analysis of mass-action kinetics. *IEEE Control Syst. Mag.* **2009**, *29*, 60–78.
2. Bomze, I.M. Lotka-Volterra equation and replicator dynamics: A two-dimensional classification. *Biol. Cybern.* **1983**, *48*, 201–211. [[CrossRef](#)]
3. Khalil, H.K. *Nonlinear Systems*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2002.
4. Chiang, H.D.; Alberto, L.F.C. *Stability Regions of Nonlinear Dynamical Systems: Theory, Estimation, and Applications*; Cambridge University Press: Cambridge, UK, 2015.
5. Garnier, H.; Wang, L. (Eds.) *Identification of Continuous-Time Models from Sampled Data*, 1st ed.; Advances in Industrial Control; Springer: London, UK, 2008; p. 413.
6. Augusiak, J.; Van den Brink, P.J.; Grimm, V. Merging validation and evaluation of ecological models to ‘evaluation’: A review of terminology and a practical approach. *Ecol. Model.* **2014**, *280*, 117–128. [[CrossRef](#)]
7. ElKalaawy, N.; Wassal, A. Methodologies for the modeling and simulation of biochemical networks, illustrated for signal transduction pathways: A primer. *BioSystems* **2015**, *129*, 1–18. [[CrossRef](#)] [[PubMed](#)]
8. Koopman, B.O. Hamiltonian Systems and Transformation in Hilbert Space. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 315–318. [[CrossRef](#)] [[PubMed](#)]
9. Williams, M.O.; Kevrekidis, I.G.; Rowley, C.W. A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. *J. Nonlinear Sci.* **2015**, *25*, 1307–1346. [[CrossRef](#)]

10. Mezić, I. Spectrum of the Koopman Operator, Spectral Expansions in Functional Spaces, and State Space Geometry. *arXiv* **2017**, arXiv:1702.07597.
11. Mezić, I.; Mezić, I. Spectral Properties of Dynamical Systems, Model Reduction and Decompositions. *Nonlinear Dyn.* **2005**, *41*, 309–325. [[CrossRef](#)]
12. Lan, Y.; Mezić, I. Linearization in the large of nonlinear systems and Koopman operator spectrum. *Phys. D Nonlinear Phenom.* **2013**, *242*, 42–53. [[CrossRef](#)]
13. Mauroy, A.; Mezić, I.; Moehlis, J. Isostables, isochrons, and Koopman spectrum for the action–angle representation of stable fixed point dynamics. *Phys. D Nonlinear Phenom.* **2013**, *261*, 19–30. [[CrossRef](#)]
14. Mauroy, A.; Mezić, I. Global Stability Analysis Using the Eigenfunctions of the Koopman Operator. *IEEE Trans. Autom. Control* **2016**, *61*, 3356–3369. [[CrossRef](#)]
15. Coayla-Teran, E.A.; Mohammed, S.E.A.; Ruffino, P.R.C. Hartman-Grobman theorems along hyperbolic stationary trajectories. *Discret. Contin. Dyn. Syst.* **2007**, *17*, 281–292. [[CrossRef](#)]
16. Klus, S.; Koltai, P.; Schütte, C. On the numerical approximation of the Perron-Frobenius and Koopman operator. *J. Comput. Dyn.* **2016**, *3*, 51–79.
17. Korda, M.; Mezić, I. On Convergence of Extended Dynamic Mode Decomposition to the Koopman Operator. *J. Nonlinear Sci.* **2018**, *28*, 687–710. [[CrossRef](#)]
18. Schmid, P.J. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **2010**, *656*, 5–28. [[CrossRef](#)]
19. Koekoek, R.; Lesky, P.A.; Swarttouw, R.F. *Hypergeometric Orthogonal Polynomials and Their q -Analogues*; Springer Monographs in Mathematics; Springer: Berlin/Heidelberg, Germany, 2010.
20. Brunton, S.L.; Proctor, J.L.; Kutz, J.N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 3932–3937. [[CrossRef](#)] [[PubMed](#)]
21. Konakli, K.; Sudret, B. Reliability analysis of high-dimensional models using low-rank tensor approximations. *Probabilistic Eng. Mech.* **2016**, *46*, 18–36. [[CrossRef](#)]
22. Konakli, K.; Sudret, B. Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions. *J. Comput. Phys.* **2016**, *321*, 1144–1169. [[CrossRef](#)]
23. Garcia-Tenorio, C.; Delansnay, G.; Mojica-Nava, E.; Vande Wouwer, A. Trigonometric Embeddings in Polynomial Extended Mode Decomposition—Experimental Application to an Inverted Pendulum. *Mathematics* **2021**, *9*, 1119. [[CrossRef](#)]
24. Garcia-Tenorio, C.; Mojica-Nava, E.; Sbarciog, M.; Wouwer, A.V. Analysis of the ROA of an anaerobic digestion process via data-driven Koopman operator. *Nonlinear Eng.* **2021**, *10*, 109–131. [[CrossRef](#)]
25. Garcia-Tenorio, C.; Tellez-Castro, D.; Mojica-Nava, E.; Wouwer, A.V. Analysis of a Class of Hyperbolic Systems via Data-Driven Koopman Operator. In Proceedings of the 2019 23rd International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 9–11 October 2019; pp. 566–571.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.