Data-Driven Analysis of Mass-Action Kinetics

Camilo Garcia-Tenorio, Duvan Tellez-Castro, Eduardo Mojica-Nava, and Jorge Sofronye

Abstract—The physical interconnection of spatial distributed biochemical systems has some advantages when dealing with large-scale problems that require separated agents to be controlled locally but with an overall objective. The analysis and control of such systems becomes a difficult task, because of the nonlinear nature of the dynamics of the individual subsystems, and the added complexity due to the interconnection. Therefore, analysis tools from a data-driven perspective are employed in contrast with the analytical classical way that becomes intractable once the subsystems grow in size or complexity.

Index Terms—Attraction region, Eigenvalue analysis, Koopman operator, Mass action kinetics, Stoichiometric networks.

I. INTRODUCTION

The problem of wastewater treatment is a current issue in the path of sustainable intelligent cities. Modern cities can plan their growth and development centered on these perspectives, but older cities whose development was not sustainable centered cannot accommodate the large areas a centralized wastewater treatment plant occupies. Along this trend, city expansion could be focused on decentralized wastewater treatment in favor of better use of space and the avoidance of large facilities [1]. There are some challenges to be addressed for the decentralized management of water. One of them is in the dimensionality of the problem due to interconnection; the whole system has to guarantee the collaboration of subsystems to get a handle on an adequate effluent. On the other hand, the whole interconnected system in the wastewater treatment allows for its resilience in terms of collaboration between individual subsystems when one or several are operating under nonideal conditions.

Biochemical processes for the wastewater treatment process such as the activated sludge treatment can be modeled by mass action kinetics (MAK). This representation, based on the stoichiometric network of biochemical processes give a mathematically rich polynomial structure based on concentrations [2]. There are plenty of analysis tools to handle on such structure [3], [4]. The problem arises when there is uncertainty in the reaction constants [5], when there is uncertainty, the general behavior of the system cannot be analyzed quantitatively in an exact manner, only qualitative characteristics are present. Also, there are problems when an interconnection component is present on the system, that is a system of systems. This interconnection component induces a rise in the number of variables and can induce complex behavior. When this happens, classical tools have to be complemented with theories handling on large-scale systems and their behavior in order to analyze the overall behavior and performance of the system.

The objective of this paper is to show how classical techniques such as Lyapunov, or passivity (Energy-based) analysis can be coupled with operator theory (Koopman operator) in order to have a measure of the individual and overall behavior and performance. There are many tools for the analysis of polynomial or algebraic systems [3], [6], [7]. The problem with these is that they can not get a handle on complexity; i.e., when the dimensionality or the order of the system grows, specifically in the case of complexity for polynomial systems [8]–[11]. Traditional methods for the non-existence of periodic orbits such as Bendixon criteria or comparison functions need much effort for its implementation in high dimensions [12], [13]. Another problem is that usually with the classic techniques (e.g., Lyapunov, input-output stability) the way to develop a global stability analysis, is checking if there are other points of equilibria or limit cycles [14]. But even the first step of finding all the equilibria points for the analysis of nonlinear systems is a hard problem. Therefore, complex dynamic systems require tools for their analysis that transcend beyond the current knowledge in control theory [15]. The escalating behavior in the number of equilibria and the characteristic of the equilibria do not allow for a simple and straightforward approach. Still, the overall performance has to be guaranteed under interconnection. Therefore, the approach proposed in this paper is to use operator based algorithms coupled with classical analysis tools. This allows us to have a better measure of the performance of the system; the characterization of attraction regions for this particular case when we only have data snapshots of the system and a general or qualitative picture of the process. Although the analysis tools are intended for the interconnection of systems, the proposed approach is exposed for the analysis of a single system.

We show that based on the level sets of the main eigenfunction of the Koopman operator we can extract information concerning the stable manifold of saddle points, and therefore, we can characterize the stability boundary of the attraction region of asymptotically stable equilibrium points.

This paper is organized as follows. In Section II, we present the analysis of a system based on MAK where the problem of complexity is presented by the order of the differential equation and the existence of multiple equilibria. In Section III, we show the current methods for finding the Koopman operator based on extended dynamic mode
decomposition, an useful theorem for the characterization of attraction regions, and the way we use the operator to find these regions. In Section IV we present our results for a particular problem within the MAK modeling strategy. A discussion which describes implications and advantages of the analysis is presented in Section V. Finally, in Section VI we present some conclusions.

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 vector \( x \in \mathbb{R}^n \), \( x^* \) denotes complex conjugate transpose, and \(|x|\) represents the Euclidean norm. The space \( n \times n \) of real matrices is designates by \( \mathbb{R}^{n \times n} \). For a complex number \( \lambda \), \(|\lambda|\) represents its norm. For any set \( A \), \( \text{Card}(A) \) denotes the closure of \( A \). Finally, the operator \( \bigotimes \) is defined as the product term to term. The vector exponentiation \( M^{z_0} \) is defined term by term.

II. LOCAL DYNAMICS

The distributed approach for treating the wastewater and flood control problem provides that the subsystems be represented in a suitable form. Biochemical dynamical systems can be represented by mass action kinetics (MAK) [2], [16]–[19], defined by autonomous nonlinear dynamical systems of the form

\[
\dot{x} = f(x, u), \quad x(0) = x_0, \quad t \in \mathbb{R}_+,
\]

where \( x(t) \in \mathbb{R}^n \) is the state of the system, \( u(t) \in \mathbb{R}^p \) is the input or control signal of the system, and \( f : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n \) is continuous in \( x \) and \( u \). Dynamical systems under the MAK structure produce differential equations in a Euclidean form, as they represent concentrations, they cannot be negative, i.e., they live in the first orthant of the Euclidean space. Starting from a set of species \( S = \{s_1, s_2, \cdots, s_n\} \in \mathbb{R}^n \) with their respective concentrations \( X = [x_1, x_2, \cdots, x_n]^{\top} \in \mathbb{R}^n \) such that \( \sum_{j=1}^{n} x_j = 1 \), a set of reactions \( \kappa = [k_1, k_2, \cdots, k_m]^{\top} \in \mathbb{R}^m \), and a stoichiometric network given by

\[
\sum_{j=1}^{n} \alpha_{i,j} x_j \bigotimes \sum_{j=1}^{n} \beta_{i,j} x_j,
\]

In matrix form, the reactants \( \alpha_{i,j} \) and the products \( \beta_{i,j} \) can be represented as \( A = [\alpha_{i,j}] \) and \( B = [\beta_{i,j}] \), \( A, B \in \mathbb{R}^{m \times n} \) respectively. With this at hand, the differential equation governing the stoichiometric network (2) is

\[
\dot{x} = [B - A]^{\top} (\kappa \otimes x^k),
\]

where the matrix exponentiation gives a vector \( y^A \in \mathbb{R}^m \) such that \( y_j = \prod_{k=1}^{n} x_k^{n^j} \). Through the whole paper we are going to analyze the next example based on a stoichiometric network.

Example 1: Consider the simple network of an autocatalytic replicator [20] on a continuous flow stirred tank reactor (CFSTR) described by the network

\[
\begin{align*}
s_1 + 2s_2 & \quad \xrightarrow{k_1} 3s_2 \\
\quad s_2 & \quad \xrightarrow{k_2} s_3 \\
s_1 & \quad \xrightarrow{g} s_0 \quad \xrightarrow{g} s_2 \\
& \quad \xrightarrow{g} s_3,
\end{align*}
\]

where \( s_1 \) is the resource the species \( s_2 \) consumes to replicate, \( s_3 \) is the death species in the tank, \( s_0 \) is the environment, from which input reactants or species come from and go to at the input and output. \( k_1 > k_2 \) are the replicator rate constant and the species death constant respectively. \( g \) is the in/out-flow of the system, and \( x_i^n \) are the concentration of reactants and product in the input feed. The differential equation for this system arises from:

\[
A = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \kappa = \begin{bmatrix} k_1 \\ k_2 \\ g \\ g \\ g \end{bmatrix},
\]

where the effect of the input concentrations has been neglected because \( x_i^n = [1 0 0]^{\top} \), which as a result gives the set of differential equations

\[
\begin{align*}
\dot{x}_1 &= -k_1 x_1 x_2^2 + g - gx_1 \\
\dot{x}_2 &= +k_1 x_1 x_2^2 - k_2 x_2 - gx_2 \\
\dot{x}_3 &= +k_2 x_2 - gx_3,
\end{align*}
\]

where the line concerning the dynamics of the \( s_0 \) species, which represents the exchange of material with the environment was discarded because it has no dynamics (i.e., \( \dot{x}_0 = 0 \)). Note that any of the equations for the state variables can be discarded as it is a linear combination of the other two, then, the system can be treated as a two state variable problem (e.g., \( x_1 \) and \( x_2 \)).

Definition 1: A vector \( x^* \in \mathbb{R}^m \) satisfying \([B - A]^{\top} (\kappa \otimes x^A) = 0\) is an equilibrium of (3).

Example 1 (Continued):

\[
x^* = \left\{ (x_1, x_2, x_3) \in \mathbb{R}_+^3 : [1 0 0]^{\top}, \begin{bmatrix} 1 - \gamma + \sqrt{\gamma \sqrt{\gamma k_1}} \\ 2 \sqrt{\gamma k_1} \\ 2 \sqrt{\gamma k_1} (g + k_2) \end{bmatrix}, \begin{bmatrix} k_2 (\gamma + \sqrt{\gamma k_1}) \\ 2 \sqrt{\gamma k_1} (g + k_2) \end{bmatrix}, \begin{bmatrix} 1 + \gamma - \sqrt{\gamma \sqrt{\gamma k_1}} \\ 2 \sqrt{\gamma k_1} (g + k_2) \end{bmatrix}, \begin{bmatrix} k_2 (\gamma + \sqrt{\gamma k_1}) \\ 2 \sqrt{\gamma k_1} (g + k_2) \end{bmatrix} \right\}
\]

where \( \gamma \) is defined as

\[
\gamma(k_1, k_2, g) = \sqrt{-4g_2^2 - 8gk_2 + k_1g - 4k_2^2},
\]

and the equilibria is only defined for the interval in which \( \gamma \) is real.
An analysis of the nature of the equilibria for the simple system based on classical techniques would be very difficult, the nature of the equilibria for every equilibrium depends on the values of \( k_1, k_2, \) and \( g \).

**Example 1 (Continued):** For the first point in (6) a simple choice of Lyapunov function \( V(x) \), see [7] for details on Lyapunov based stability,

\[
V(x) = -\sum_{i=1}^{m} x_i^* \ln \left( \frac{x_i}{x_i^*} \right) ,
\]

where \( V(x^*) = 0 \) and due to Jensen’s inequality \( \mathbb{E}f(x) \geq f(\mathbb{E}x) \),

\[
\begin{align*}
V(x) &= -\sum_{i=1}^{m} x_i^* \ln \left( \frac{x_i}{x_i^*} \right) \\
&\geq -\ln \left( \sum_{i=1}^{m} x_i x_i^* \right) \\
&= 0.
\end{align*}
\]

Now, for the derivatives of \( V(x) \) along the trajectories of \( x \), and due to the fact that \( x \in [0, 1] \) we have

\[
\begin{align*}
\dot{V}(x) &= \begin{bmatrix} x_1^2 & x_2^2 & x_3^2 \\ x_1 & x_2 & x_3 \end{bmatrix} \dot{x} \\
&= -k_1 x_2^2 + g - \frac{g}{x_1} \\
&\leq 0.
\end{align*}
\]

Note that \( \dot{V}(x) \neq 0 \) for all \( x \neq x_c \) and therefore the first equilibrium point is asymptotically stable, and, as \( V(x) \) is not radially unbounded; the first equilibrium is not globally asymptotically stable. As \( x_3 \) can be expressed as a linear combination of \( x_1 \) and \( x_2 \), there is no uncertainty induced by the absence of \( x_3 \) in the derivative along the trajectories of the Lyapunov function.

For the second and third equilibria in (6), the dynamic behavior is more complex, as they have a dependency on the parameters, and the second equilibrium produces a Hopf bifurcation [21], [22] that changes the nature of the equilibria. The dynamic behavior of the equilibria with respect to different values of the in/out-flow constant \( g \), and some fixed values of the reaction constants \( k_1 = 10 \) and \( k_2 = 0.1 \) is presented in Fig. 1.

Although this process gives important information about of the local behavior of the different equilibria, a better analysis would consider regions of attraction of these equilibria for control and optimization of the process.

Now, consider that the system belongs to a network of interacting subsystems that are set to collaborate into achieving a common goal while being controlled locally for their individual performance, such a subsystems interconnection would yield a structure such as

\[
\dot{x}_i = f_i(x_i(t), u_i(t)) + g(x(t)) , \{ i \}^N_s,
\]

where \( f_i(x_i(t), u_i(t)) : \mathbb{R}^{n_i} \times \mathbb{R}^{m_i} \rightarrow \mathbb{R}^{n_i} \) represents the local dynamics, and \( g(x) : \mathbb{R}^{\Sigma_r m} \rightarrow \mathbb{R}^{n_i} \) represents the effect due to interconnection, these for each of the \( N_s \) subsystems. When considering interconnection the dimension of the dynamical system grows. Depending on how large is this growth, on the order of the individual systems, the number of individual systems and the interconnection structure, the simple task of finding the equilibria may not be possible analytically, and numerical methods may not give solutions within acceptable computing times [23].

The proposed approach combines two techniques for the solution of the equilibria analysis of the interconnected system via numerical methods. These present two advantages with respect to the classical approach. The first is in terms of dimensionality, and the second is in terms of dependence on the model. The operator based technique, such as the Koopman operator, relies on a data set of snapshot pairs of the system, e.g., the full state observable of the system space with a specific sampling time [24], [25]. This approach allows us to complement algorithmic techniques such as [26], to characterize nonlinear dynamical systems. These techniques depend on the model and the need of backward integration to better characterize the system. In the next section the extended dynamic mode decomposition (EDMD) and the Stability Regions of Nonlinear Dynamical Systems (SR) are going to be explained as well as the way in which these two can be combined and complemented.

### III. The Koopman Operator and Attraction Regions

The approach for the analysis of the dynamical characteristics of the systems, as a single entity and in interconnection is based on EDMD and SR. The following algorithm for EDMD is from [25], and references therein where the theory behind the algorithm is presented.

#### A. Extended Dynamic Mode Decomposition

Consider the discrete time dynamical system \( (\mathcal{M}_d; \eta; F^\eta_d) \), where \( \mathcal{M}_d \subseteq \mathbb{R}^N \) is the state space, \( \eta \in \mathbb{Z} \) is the discrete time, and \( F^\eta_d : \mathcal{M}_d \rightarrow \mathcal{M}_d \) is the evolution operator. The Koopman \( \mathcal{K} \) operator acts on functions of the state space.
\[ \psi \in \mathcal{F}, \text{ where } \psi : \mathcal{M} \to \mathbb{C}, \text{i.e., defines the evolution of the observable } \psi \text{ governed by the linear operator } \mathcal{K}, \text{ defining a new dynamical system } (\mathcal{F}; \mathcal{M}; \mathcal{K}). \] For the case of autonomous continuous time dynamical systems such as (1), there is an associated continuous flow map \( F^t_c(x) : \mathcal{M} \to \mathcal{M} \) (i.e., the solution of the differential equation for \( t \in \mathbb{R}^+ \)) that defines the continuous time dynamical system \((\mathcal{M}; F^t_c)\). For this dynamical system, we can define the Koopman operator \( \mathcal{K}_{\Delta t} \), associated with the flow of the system in a fixed time interval \( \Delta t \), defining again a new dynamical system \((\mathcal{F}; \eta; \mathcal{K}_{\Delta t})\), i.e., a dynamical system defined by its discrete time evolution operator \( \mathcal{K}_{\Delta t} \).

The Koopman operator based algorithm requires \( P \) pairs of data snapshots, either from a real system or a computationally integrated one at a specific sampling \( \Delta t \). The snapshot pairs, \( \{(x_p, y_p)\}_{p=1}^P \) are organized in data sets

\[ X = [x_1 \ x_2 \ \ldots \ x_P], \quad Y = [y_1 \ y_2 \ \ldots \ y_P], \]

where \( x_i, y_i \in \mathcal{M} \), the space state of the system, and a dictionary of observables \( \mathcal{D} = \{\psi_1, \ldots, \psi_{N_k}\} \), where \( \psi_i \in \mathcal{F} \) are the observables that define the vector valued function \( \Psi : \mathcal{M} \to \mathbb{C}^{1 \times N_k} \) where

\[ \Psi(x) = [\psi_1(x) \ \psi_2(x) \ \ldots \ \psi_{N_k}(x)]. \]

The goal now is to generate the finite-dimensional approximation of the infinite linear operator \( \mathcal{K} \), defined as \( K \in \mathbb{C}^{N_k \times N_k} \) and calculated

\[ K \triangleq G_k^+ A_k, \]

where \( + \) denotes pseudoinverse, and matrices \( G_k, A_k \in \mathbb{C}^{N_k \times N_k} \) are defined by

\[ G_k = \frac{1}{P} \sum_{p=1}^P \Psi(x_p)^* \Psi(x_p), \]

\[ A_k = \frac{1}{P} \sum_{p=1}^P \Psi(x_p)^* \Psi(y_p), \]

where \( * \) denotes complex conjugate. With the \( K \) matrix, its eigenvalues \( M_k = \text{diag}(\mu_1, \ldots, \mu_{N_k}) \), and right eigenvectors \( \Xi = [\xi_1, \ldots, \xi_{N_k}] \) can be calculated. Koopman eigenfunctions can be defined as

\[ \Phi = \Psi(x)\Xi, \]

with \( \Xi \) as the matrix of right eigenvectors, its inverse is the matrix of left eigenvectors, \( \Xi^{-1} = W_k^* \). Now, defining the full-state observable, \( g_k(x) = x \) such that \( g_k(x) : \mathcal{M} \to \mathbb{R}^N \), where \( g_k(x) \in \mathcal{F}_D \subseteq \mathcal{F} \), i.e., the full state observable lies in the space spanned by \( \mathcal{D} \), which is a subspace of \( \mathcal{F} \). With this result, the full state observable can be expressed as

\[ g_k(x) = (\Psi(x)B_k)^\top, \]

where \( B_k \in \mathbb{R}^{N \times N_k} \) is an appropriate matrix of weights that relate the leading Koopman eigenfunctions to each of the full state observables values for them to accurately capture the state. The literature lacks a proper way of characterizing these vectors that compose the matrix \( B_k \). With the appropriate weights for each of the leading eigenfunctions, the full state observable can be calculated as

\[ g_k(x) = V_k \Phi(x)^\top, \quad V_k = (W_k^* B_k)^\top, \]

where the \( i \)th column of \( V_k \) is the \( i \)th Koopman mode, leaving us with an appropriate definition of the full state observable in terms of Koopman eigenfunctions and modes of the Koopman operator \( K \). This definition serves to define the evolution of the states in terms of the Koopman characterization \( \{(\mu_k, \phi_k, v_k)\}_{k=1}^N \) consisting of \( N_k \) Koopman eigenvalues, eigenfunctions, and modes. Therefore, the \( \eta \)th application of the state evolution \( F^\eta(x) = x_\eta \) from the initial condition \( x_0 \) is

\[ x(\eta) = F^\eta(x_0) = K^n g_k(x_0) = (M_k^n \circ V_k) \Phi(x_0)^\top. \]

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

\[ x(-\eta) = (M_k^{-\eta} \circ V_k) \Phi(x)^\top. \]

Remark 1: The Koopman eigenfunctions \( \Phi(x) \) contain information about the equilibria of the system in terms of where they are, and the functions reveal the stable and unstable manifold of saddle point equilibria. With the deduction of the Koopman eigenfunctions and the evolution of the system, the stability boundary of asymptotically stable equilibria can be characterized in the following way.

B. Stability Boundary

The theory for stability boundary for systems that contain saddle points that characterize the attraction region of an asymptotically stable equilibrium point was developed by [26]. Equilibria is divided between asymptotically stable equilibria \( x_s \), and saddle/hyperbolic equilibria \( \hat{x} \). For each of them we can define the attraction region of the asymptotically stable equilibria defined as \( A_s(x_s) \triangleq \{x \in \mathbb{R}^n : \lim_{t \to \infty} F^t_c(x) = x_s\} \), the boundary of the attraction region as \( \partial A_s(x_s) = A_s(x_s) \cap (\mathbb{R}^n \setminus A_s(x_s)) \), the stable manifold of the saddle point \( W^s(\hat{x}) = \{x \in \mathbb{R}^n : F^t_c(x) \to \hat{x} \text{ as } t \to \infty\} \), and the unstable manifold of the saddle point \( W^u(\hat{x}) = \{x \in \mathbb{R}^n : F^t_c(x) \to \hat{x} \text{ as } t \to -\infty\} \) (assuming there is an inverse for \( F^t_c \) for the backward flow of the system). With these definitions, we can present the main theorem for the characterization of attraction regions [26, Th. 4-8].

**Theorem 1:** For a nonlinear autonomous dynamical systems (1) which satisfies the following assumptions.

1) All equilibria on the stability boundary are saddle points.
2) The stable and unstable manifolds of the equilibria on the stability boundary satisfy the transversality condition.
3) Every trajectory on the stability boundary approaches one of the saddle point equilibria as \( t \to \infty \).
Let \( \{ \hat{x}_i \}_{i \in \mathbb{N}} \) be the saddle equilibria on the stability boundary \( \partial A \) of an asymptotically stable equilibrium point \( x_s \). Then

(a) \( \hat{x}_i \in \partial A \) if and only if \( W^u(\hat{x}_i) \cap A(x_s) \neq \emptyset \)

(b) \( \partial A(x_s) = \cup W^s(\hat{x}_i) \).

Next, we are going to show how the eigenfunctions from EDMD can be used to characterize the attraction region of an asymptotically stable equilibrium point.

C. Koopman Operator for Attraction Regions

The characterization of the attraction region of an asymptotically stable equilibrium point requires that the stable manifolds of saddle points on the boundary of attraction be characterized. For this purpose, we propose the use of some important results coming from the operator representation that EDMD gives. First, we are going through some description and assumptions of the kind of systems being evaluated.

The systems under evaluation are under the MAK modeling strategy, these have a nice polynomial structure coming from the stoichiometric network of the system. Although we know the polynomial structure arising from MAK, we assume that the reaction constants for the process are not known, but that they are constant nonetheless, i.e., they are independent of the reactor temperature, either because the system is adiabatic or, from the assumption that a proper controller has been setup for this purpose. This gives us some notion of the characteristics of the equilibria for the system, but not the whole picture.

We also assume that there is information concerning several trajectories of the system, i.e., the system has been tested from several initial conditions such that a set of snapshots can be used to characterize the whole dynamic spectra, which is possible due to the fact that MAK lie in the unitary manifold of the first orthonth. These snapshots correspond to a full measurement of the state at a specified sampling time. Now, the procedure is to identify the attraction regions, characterized by the stability boundary, defined by the stable manifold of saddle points in the boundary.

Consider the autonomous nonlinear system (1) whose dynamics come from (3), based on a stoichiometric network such as (4). Recall from the EDMD procedure that the forward and backward flow of the system can be achieved in terms of the eigenvalues, eigenfunctions, and Koopman modes, (7) and (8). Therefore if any of these trajectories is invariant with respect to the flow map or evolution operator of the system \( F^u(x_0) \), then the system is in a fixed point.

**Definition 2 (Fixed Point):** Let \( x^* \) be a fixed point of the dynamical system \( (M; F^t(x)) \), then the flow map \( F^t(x^*) \) maps into itself. i.e.,

\[
F^t(x^*) = x^*.
\]

As it is the case, we propose minimization of the squared norm of the difference between \( x \) and the flow map of the system over the state space in order to find the equilibria.

**Proposition 1 (Fixed points):** Let \( (M; F^t(x)) \) be a dynamical system that accepts an operator representation \( (\mathcal{F}; \eta; \mathcal{K}_{\Delta t}) \) based on EDMD, then its evolution operator is defined by (7), and its equilibria can be found by solving

\[
x^* = \min_x \| (M^t_k \circ V_k)\Phi(x)^T - x \|
\]

where all the local minima represent a fixed point of the system.

**Proof:** This is a consequence of Def 2.

**Remark 2:** Note that this procedure is possible under the MAK paradigm as the system is defined on the unitary manifold of the first orthonth, and therefore a complete exploration of the space state from several initial conditions is possible to find all the equilibria of the system.

**Remark 3:** This proposition is only accountable for the location of the equilibria in the space state, it does not give information about its stability characteristic.

With the location of the equilibria, the known structure of the dynamical system based on the stoichiometric network comes into place in order to know the proper way to analyze the system. There are techniques that consider the reversibility of the stoichiometric network, and the deficiency of the system in order to draw conclusions about the number of equilibria and their behavior [2], [3], [16]–[19], [27], [28]. We are focusing our efforts for the case where a saddle equilibrium can characterize the attraction region of an asymptotically stable equilibrium point via the stable manifold of a saddle equilibrium, therefore, we assume that by linearization, or by using the structure of the system and its properties, the saddle point is identifiable. With this information, the next proposition is the main result of this work, and is the fact that the stable manifold of a saddle equilibrium has a level set on the main eigenfunction describing the system.

**Proposition 2:** Let \( (M; F^t(x)) \) be a dynamical system that accepts an operator representation \( (\mathcal{F}; \eta; \mathcal{K}_{\Delta t}) \) based on EDMD, then its main eigenfunction \( \phi_+(x) \) corresponding to the eigenvalue with the largest norm is constant along the trajectory of the stable manifold of the saddle equilibria. Therefore, the stable manifold is defined by the level set

\[
L_{W^s(x)}(\phi_+(x)) = \{ (x_1, \ldots, x_n) : \phi_+(x_1, \ldots, x_n) = \phi_+(\hat{x}) \}
\]

**Proof:** This is a consequence from the fact that \( K^\dagger \phi_i = \lambda_i \phi_i \), and the linearity of the operator [29], [30].

IV. Application Example

Consider the application example of, the auto-catalytic simple system (4), assume that the structure is known, and that we are working in the region where the in/out-flow \( g \) induces a stable focus on the third equilibrium point.

**Example 1 (Continued):** We run the system 100 times with a length of 21 points per run with a \( \Delta t = 0.125 \), that gives us a total of 2,000 snapshots. With that information, we apply EDMD from Section III-A with a choice of dictionary composed by 25 entries of Hermite Polynomials. The system under study is the reduced dynamics due to the dependence of \( x_3 \) on the other two, that is, (5) without the last equation. The result is an EDMD based Koopman operator that
effectively represents the dynamics of the system with 25 eigenfunctions $\phi(x)$ that characterize the dynamical structure of the system. The first and second eigenfunctions have the associated eigenvalue with the largest norm, and their evaluation on the unitary manifold that contains the dynamics can be seen in Fig. 2, where the level set, obtained by evaluating the main eigenfunction as in Proposition 2 is also shown. The result of this approach is presented in

![Image](image_url)

**Fig. 2.** $\phi_0(x)$ and the level set $L_{W^s(\hat{z})}(\phi_0(x))$ that characterizes the stable manifold of the saddle point.

**Fig. 3.** Evolution of trajectories with initial conditions based on the stable manifold of the saddle criteria.

As can be seen in Fig. 2, where the level set, obtained by evaluating the main eigenfunction as in Proposition 2 is also shown. The result of this approach is presented in

\[
\begin{align*}
\text{IC}_\bullet &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \leq 1\} \cup \\
&\quad \{(x_1, x_2) \in \mathbb{R}^2 : (x_1, x_2) \geq L_{W^s(\hat{z})}(\phi_0(x))\},
\end{align*}
\]

and in solid diamonds the initial conditions that satisfy

\[
\begin{align*}
\text{IC}_\bullet &= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \leq 1\} \cup \\
&\quad \{(x_1, x_2) \in \mathbb{R}^2 : (x_1, x_2) < L_{W^s(\hat{z})}(\phi_0(x))\}.
\end{align*}
\]

Taking trajectories from these distinct initial conditions gives the result as to which of the equilibria they converge, where it can be seen that there is some error due to the numeric nature of the algorithm, despite the error, a measure of the minimal distance between the focus equilibrium and the level set gives a measure on robustness of that equilibria with respect to any possible perturbations, i.e., 0.5 for the case under consideration.

V. Discussion

Our proposed approach relies on the proper analysis of eigenfunctions. This paper only covers the analysis of the level sets of eigenfunctions. This analysis lets us characterize the stable manifold of saddle points to obtain the attraction region of an asymptotically stable equilibrium. This is among the tools to extract information out of the eigenfunctions [30]–[32] to properly analyze a dynamical system. These numeric tools represent a base for the future analysis of complex systems in order to synthesize controllers for polynomial dynamical systems (e.g., MAK). Although there are other ways to analyze systems in terms of their attraction regions, we only need the structure of the system in order to perform the analysis. A qualitative representation is enough.

Even if we had the tools for extracting all kinds of information regarding the system, there are limitations with respect to the deduction of the eigenfunctions. The algorithms are not clear on the appropriate number of samples or the distribution of initial conditions in order to have an accurate representation of the system. There is no clear criteria to select a dictionary and its size. This produces inaccuracies in the eigenfunction, and as can be seen in Fig. 3, some trajectories with initial conditions in the estimated region for the trivial equilibrium converge to the stable focus.

A possible improvement would be with respect to the choice of the main eigenfunction of the operator. All eigenfunctions capture different information and properties (relevant or not, accurate or not). A linear combination of the different level sets that show a similar behavior could yield better results. Unfortunately, the shortcoming stated earlier and the fact that the choice of dictionary affects the information that a particular eigenfunction captures, there is no proper way of selecting the main eigenfunctions.

These data-driven techniques allow us to work over broader spectra of the state space; when dealing with systems where the exact dynamics are unknown, in presence of uncertainties in the system parameters or in the presence of perturbation that can get the system out of its operation point or even destabilize it. Therefore, these data-driven approach can complement classical control synthesis techniques to get a handle on optimality and robustness.

To improve on the analysis, we need to explore the polynomial structure of MAK and find the relationship with the polynomials of the dictionary. This will allow us to construct
dictionaries which will provide eigenfunctions with more information regarding the characterization of the system. This way, reduced order dictionaries could be employed to have a better handle on complexity.

VI. CONCLUSIONS

In conclusion, eigenfunctions coming from operator based algorithms provide information for the system analysis; in this particular case, the equilibria of the system and the attraction regions of asymptotically stable equilibria. The problem relies on the proper construction of eigenfunctions and the extraction of the information they provide.

The accuracy of the estimation of the stability boundary depends on the number of available snapshots, the kind of dictionary used to obtain the Koopman operator matrix, and the size of the dictionary.

We can conclude that the operator based analysis of dynamical systems is far from complete; we need a stronger theoretical framework in order to analyze systems with inputs, interconnection and feedback controllers.

REFERENCES

[1] T. A. Larsen, S. Hoffmann, C. Luthi, B. Truffer, and M. Maurer, “Emerging solutions to the water challenges of an urbanizing world,” Science, vol. 352, pp. 928–933, May 2016.
[2] V. Chellaboina, S. Bhat, W. Haddad, and D. Bernstein, “Modeling and analysis of mass-action kinetics,” IEEE Control Systems Magazine, vol. 29, pp. 60–78, Aug 2009.
[3] Q. H. Wassim M. Haddad Vijay Sekhar Chellaboina, Nonnegative and Compartmental Dynamical Systems. Princeton University Press, 2010.
[4] D. Bernstein and S. Bhat, “Nonnegativity, reducibility, and semistability of mass action kinetics,” in Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No.99CH36304), vol. 3, pp. 2206–2211, IEEE, 1999.
[5] R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petsche, and J. H. Schabily, “Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory,” The Journal of Chemical Physics, vol. 59, pp. 3873–3878, Oct 1973.
[6] W. M. Haddad and V. Chellaboina, Nonlinear dynamical systems and control : a Lyapunov-based approach. Princeton University Press, 2008.
[7] H. K. Khalil, Nonlinear systems. Prentice Hall, 3rd ed., 2002.
[8] A. A. Ahmadi and P. A. Parrilo, “Converse results on existence of sum of squares Lyapunov functions,” in IEEE Conference on Decision and Control and European Control Conference, (Orlando FL, USA), pp. 6516–6521, IEEE, Dec 2011.
[9] A. Majumdar, A. A. Ahmadi, and R. Tedrake, “Control and verification of high-dimensional systems with DSOS and SDSOS programming,” in Proceedings of the IEEE Conference on Decision and Control, no. February, pp. 394–401, 2014.
[10] J. Anderson and A. Papachristodoulou, “Dynamical system decomposition for efficient, sparse analysis,” in 49th IEEE Conference on Decision and Control (CDC), (Atlanta GA, USA), pp. 6565–6570, IEEE, Dec 2010.
[11] P. Giesl and S. Hafstein, “Review on computational methods for Lyapunov functions,” Discrete and Continuous Dynamical Systems-Series B, vol. 20, no. 8, pp. 2291–2331, 2015.
[12] M. Budišić, R. Mohr, and I. Mezić, “Applied koopmanism a,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 22, no. 4, p. 47510, 2012.
[13] S. Skar, R. Miller, and A. Michel, “On nonexistence of limit cycles in interconnected systems,” IEEE Transactions on Automatic Control, vol. 26, no. 3, pp. 669–676, 1981.
[14] F. Cuesta, F. Gordillo, J. Aracil, and A. Ollero, “Stability analysis of nonlinear multivariable Takagi-Sugeno fuzzy control systems,” IEEE Transactions on Fuzzy Systems, vol. 7, no. 5, pp. 508–520, 1999.
[15] G. M. Dimirovski, Complex Systems: Relationships between Control, Communications and Computing, vol. 55. Springer, 2016.
[16] M. Feinberg, “Lectures on chemical reaction networks,” 1979.