View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Open Repository and Bibliography - Luxembourg

Linear identification of nonlinear systems: A lifting technique based on the Koopman operator

Alexandre Mauroy and Jorge Goncalves

***Abstract*— We exploit the key idea that nonlinear system identification is equivalent to linear identification of the so-called Koopman operator. Instead of considering nonlinear system identification in the state space, we obtain a novel *linear* identification technique by recasting the problem in the infinite-dimensional space of observables. This technique can be described in two main steps. In the first step, similar to a component of the Extended Dynamic Mode Decomposition algorithm, the data are lifted to the infinite-dimensional space and used for linear identification of the Koopman operator. In the second step, the obtained Koopman operator is "projected back" to the finite-dimensional state space, and identified to the nonlinear vector field through a linear least squares problem. The proposed technique is efficient to recover (polynomial) vector fields of different classes of systems, including unstable, chaotic, and open systems. In addition, it is robust to noise, well-suited to model low sampling rate datasets, and able to infer network topology and dynamics.**

I. Introduction

Operator-theoretic techniques rely on the idea of lifting nonlinear dynamical systems to an infinite dimensional space, where those systems have a linear representation. For instance, the so-called Koopman operator is a linear operator that describes the evolution of observable-functions along the trajectories of the system. Its spectral properties have been investigated over the past years and related to geometric properties of the system (see e.g. [5]). On the one hand, this body of work yielded novel methods for the analysis of nonlinear systems described by a known vector field (e.g. global stability analysis [4], global linearization [2], analysis of monotone systems [10]). On the other hand, the Koopman operator approach was shown to be conducive to data analysis [8] and related to numerical methods such as Dynamic Mode Decomposition (DMD) [9], yielding new techniques for the analysis of nonlinear systems described by trajectory data points (snapshots). In this context, the present paper aims at bridging these two directions of analysis, connecting data to vector field.

While Koopman operator techniques have been previously considered in the framework of system analysis, we present a first attempt –to the authors knowledge– to develop them for system identification. Exploiting the lifting obtained through the Koopman operator, our key idea is to *identify a linear operator in the space of observables*, instead of identifying a nonlinear system in the state space. Since Koopman operator and state space representations are both equivalent, identifying the operator is indeed sufficient to recover the nonlinear dynamical system. Because the Koopman operator is defined in an infinite-dimensional space, we consider its projection onto a finite basis of functions. The proposed method can somehow be interpreted, in the space of observables, as an analog of methods developed in [1], [6], which use dictionary functions in the state space. Our method proceeds in two main steps: from state space to space of observables, and back to state space. In the first step, a finite-dimensional projection of the Koopman operator is identified from data through classical linear identification, a method that is related to a component of the so-called Extended Dynamic Mode Decomposition (EDMD) algorithm [12]. In the second step, the projected operator is connected to the vector field of the nonlinear system through an analytic derivation of the infinitesimal generator (following similar steps as in [4]). The expansion of the vector field in the finite basis is then obtained by solving a linear least squares problem.

Thanks to the lifting technique, the proposed method has the advantage of relying only on linear techniques (i.e. least squares regression), in spite of the fact that the dynamical system is nonlinear. In contrast to recent (direct) methods (e.g. [1]), it is a (indirect) method that does not rely on the estimation of time derivatives, which cannot be performed when the sampling rate is too low. Therefore, the method is well-suited to low sampling, and also robust to both measurement noise and process noise. In addition, it works efficiently with different kinds of behaviors, including unstable and chaotic systems (although limited so far to polynomial vector fields) and can be applied to several small datasets obtained from multiple trajectories. This flexibility is well-suited to identification problems such as those encountered in the context of system biology.

The rest of the paper is organized as follows. In Section II, we introduce the system identification problem and present the lifting technique obtained through the Koopman operator. Our numerical method for nonlinear system identification based on the lifting technique is described in Section III. Extensions of the method to open systems (i.e. with outputs or process noise) are discussed in Section IV. In Section V, the method is illustrated with several numerical examples, and applied to network reconstruction. Concluding remarks and perspectives are given in Section VI.

A. Mauroy and J. Goncalves are with the Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 4367 Belvaux, Luxembourg alexandre.mauroy@uni.lu, jorge.goncalves@uni.lu

II. Koopman operator technique for system identification

A. Problem statement

Consider a (unknown) dynamical system

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n \tag{1}$$

with a vector field $\mathbf{F}(\mathbf{x})$ that is assumed to be polynomial. We denote by $\varphi : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ the flow induced by the system, i.e. $t \mapsto \varphi^t(\mathbf{x_0}) = \varphi(t, \mathbf{x_0})$ is a solution of (1) associated with the initial condition $\mathbf{x_0}$.

We address the problem of identifying the vector field $\mathbf{F}$ given K snapshot pairs $(\mathbf{x}_k, \mathbf{y}_k) \in \mathbb{R}^{n\times 2}$. The data points are such that $\mathbf{y}_k = \varphi^{T_s}(\mathbf{x}_k)(\mathbf{1} + \mathbf{v}_k)$, where each component of the measurement noise $\mathbf{v}_k$ is a Gaussian random variable with zero mean and standard deviation σ_{meas}. The sampling period T_s is assumed to be identical for all pairs $(\mathbf{x}_k, \mathbf{y}_k)$. Note that the data points $\mathbf{x}_k$ and $\mathbf{y}_k$ can belong to different trajectories. Stochastic systems with process noise and systems with inputs will also be considered in Section IV.

B. Lifting technique using the Koopman operator

The system (1) is described by its flow φ^t in the state space $\mathbb{R}^n$. Alternatively, the system can be lifted to an infinite-dimensional space $\mathcal{F}$ of observable-functions $f : \mathbb{R}^n \to \mathbb{R}$. In this space, it is described by the semigroup of Koopman operators $U^t : \mathcal{F} \to \mathcal{F}$, $t \geq 0$, which governs the evolution of the observables along the trajectories:

$$U^t f = f \circ \varphi^t \quad \forall f \in \mathcal{F}\,. \tag{2}$$

This lifting technique is of interest since it leads to a linear (but infinite-dimensional) representation of the system. Even if the system (1) is nonlinear, the operator is always linear since we have

$$U^t(a_1 f_1 + a_2 f_2) = a_1 f_1 \circ \varphi^t + a_2 f_2 \circ \varphi^t = a_1 U^t f_1 + a_2 U^t f_2\,.$$

In addition, there is a bijective relationship between the Koopman operator and its associated system. We denote by $L = \lim_{t\to 0}(U^t - I)/t$ the infinitesimal generator of the Koopman semigroup, which also satisfies

$$U^t = e^{Lt} = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k\,. \tag{3}$$

Provided that the vector field $\mathbf{F}$ is continuously differentiable, it can be shown that

$$L = \mathbf{F} \cdot \nabla\,, \tag{4}$$

where ∇ denotes the gradient (see e.g. [3]). This implies that there is a one-to-one correspondence between the Koopman operator and the vector field.

This paper exploits the equivalence between the two descriptions of the system, proposing to identify the linear Koopman operator in the space of observables. Figure 1 illustrates the two main steps of the method. In the first step, data are lifted to the space of observables, providing the evolution of observables (at some points in the state space). Then classical linear identification is used to obtain an approximation of the Koopman operator (in a finite-dimensional subspace of $\mathcal{F}$). This corresponds to a component of the so-called Extended Dynamic Mode Decomposition (EDMD) algorithm [12]. In the second step, the vector field $\mathbf{F}$ is obtained through the equalities (3) and (4).

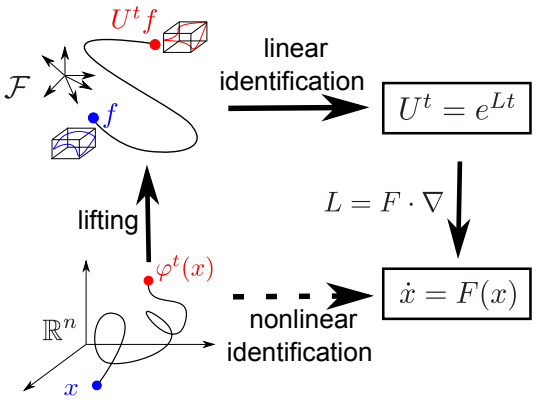

Fig. 1. Classical nonlinear system identification is performed directly in the state space. In contrast, the proposed Koopman operator identification technique consists of two steps: (1) lifting of the data and linear identification of the Koopman operator; (2) identification of the vector field using (3) and (4).

III. Numerical method

This section describes the numerical scheme derived from the Koopman operator identification method. Lifting of the data and identification of the operator (step 1) are explained in the first subsection. Identification of the vector field (step 2) is described in the second subsection.

Conceptually, the method works in the infinite-dimensional space $\mathcal{F}$ of observables. But to be numerically tractable, it is developed on a finite-dimensional subspace $\mathcal{F}_N$ which is spanned by a basis of N linearly independent functions $\{p_k\}_{k=1}^N$. In our case, we will choose a basis of monomials with total degree less or equal to m:

$$p_k(\mathbf{x}) \in \{x_1^{s_1} \cdots x_n^{s_n} | (s_1, \ldots, s_n) \in \mathbb{N}^n, s_1 + \cdots + s_n \leq m\}\,.$$

This choice is motivated by the fact that the vector field is assumed to be polynomial. Other bases of functions (e.g. Fourier basis, radial basis functions) could be considered to generalize the method. For practical purposes, the sequence of monomials should be characterized by some order (e.g. lexicographic order, weight order). The number of functions in the basis is equal to $N = (n+m)!/(n!\,m!)$. We denote by $\mathbf{p}(\mathbf{x}) = (p_1(\mathbf{x}), \ldots, p_N(\mathbf{x}))^T$ the vector of basis monomials.

The Koopman operator U^t can be projected onto the subspace $\mathcal{F}_N$ and represented by a matrix $\overline{\mathbf{U}}$. Considering

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{p}(\mathbf{x}) \in \mathcal{F}_N \tag{5}$$

and

$$\mathbb{P}U^t f(\mathbf{x}) = \mathbf{b}^T \mathbf{p}(\mathbf{x}) \in \mathcal{F}_N \tag{6}$$

where $\mathbb{P} : \mathcal{F} \to \mathcal{F}_N$ is a projection operator, we have

$$\overline{\mathbf{U}}\,\mathbf{a} = \mathbf{b}\,, \qquad \overline{\mathbf{U}} \in \mathbb{R}^{N\times N}\,. \tag{7}$$

The matrix $\overline{\mathbf{U}}$ provides a (approximate) finite-dimensional linear representation of the nonlinear system. This representation is not obtained through local linearization techniques and is valid globally in the state space. In the numerical scheme described below, we compute $\overline{\mathbf{U}}$ from data (step 1), and then connect the obtained matrix to the vector field of the nonlinear system (step 2).

A. Step 1: Lifting of the data and Koopman operator identification

We now describe the method used to compute the matrix $\overline{\mathbf{U}}$ given snapshot pairs $\{(\mathbf{x}_k, \mathbf{y}_k)\}_{k=1}^K$, which corresponds to a component of the EDMD algorithm (see [12] for more details).

For each snapshot pair $(\mathbf{x}_k, \mathbf{y}_k) \in \mathbb{R}^{n\times 2}$, we construct the new pair $(\mathbf{p}(\mathbf{x_k}), \mathbf{p}(\mathbf{y_k})) \in \mathbb{R}^{N\times 2}$. The data are somehow lifted to (a subspace of) the space of observables and provide snapshots of trajectories of particular observables (i.e. the basis functions). Provided that the measurement noise $\mathbf{v}_k$ is small, we have $\mathbf{y}_k \approx \varphi^{T_s}(\mathbf{x}_k)$. It follows from (2) and (5) that

$$\mathbb{P}U^{T_s} f(\mathbf{x}_k) = \mathbf{a}^T \mathbb{P}U^{T_s} \mathbf{p}(\mathbf{x_k}) \approx \mathbf{a}^T \mathbb{P}\mathbf{p}(\mathbf{y_k})$$

and from (6) and (7) that

$$\mathbb{P}U^{T_s} f(\mathbf{x}_k) = \mathbf{a}^T \overline{\mathbf{U}}^T \mathbf{p}(\mathbf{x}_k)\,,$$

so that

$$\mathbf{p}(\mathbf{x}_k)^T\, \overline{\mathbf{U}} \approx \mathbb{P}\mathbf{p}(\mathbf{y}_k)^T \tag{8}$$

where $\overline{\mathbf{U}}$ denotes here the matrix representation of $\mathbb{P}U^{T_s}$. The last equality implies that the jth column $\mathbf{c}_j$ of $\overline{\mathbf{U}}$ is so that $\mathbf{c}_j^T \mathbf{p} \approx \mathbb{P}U^{T_s} p_j$. If the projection $\mathbb{P}$ is chosen so that it yields the least squares fit at the points $\mathbf{x}_k$, $k = 1, \ldots, K$ (i.e. finite-dimensional Galerkin projection), the matrix $\overline{\mathbf{U}}$ is given by the least squares solution

$$\overline{\mathbf{U}} = (\mathbf{P_x})^\dagger\, \mathbf{P_y} \tag{9}$$

with the $K \times N$ matrices

$$\mathbf{P_x} = \begin{pmatrix} \mathbf{p}(\mathbf{x}_1)^T \\ \vdots \\ \mathbf{p}(\mathbf{x}_K)^T \end{pmatrix} \qquad \mathbf{P_y} = \begin{pmatrix} \mathbf{p}(\mathbf{y}_1)^T \\ \vdots \\ \mathbf{p}(\mathbf{y}_K)^T \end{pmatrix} \tag{10}$$

and where $\mathbf{P}^\dagger$ denotes the pseudoinverse of $\mathbf{P}$.

B. Step 2: Identification of the vector field

We assume that the polynomial vector field $\mathbf{F} = (F_1, \ldots, F_n)^T$ is of total degree less or equal to m_F, so that we can write

$$F_j(\mathbf{x}) = \sum_{k=1}^{N_F} w_k^j p_k(\mathbf{x}) \qquad j = 1, \ldots, n \tag{11}$$

where $N_F = (m_F + n)!/(m_F! n!)$ is the number of monomials p_k of total degree less or equal to m_F. The identification problem is therefore reduced to identifying nN_F coefficients w_k^j.

1) Derivation of the Koopman infinitesimal generator: Our goal is to derive the matrix $\overline{\mathbf{L}}$ which represents the projection of the infinitesimal generator (4) on the basis of monomials p_k, and show that this is a function of the unknown coefficients w_k^j of the vector field. Consider the linear operators

$$L_k^j = p_k\, \frac{\partial}{\partial x_j}\,, \quad j = 1, \ldots, n\,, \quad k = 1, \ldots, N_F\,.$$

It is clear that the operators L_k^j map polynomials of total degree less or equal to m_1 to polynomials of total degree less or equal to $m_2 = m_1 + m_F - 1$. Denote by $\mathbf{p}^m(\mathbf{x})$ the vector of monomials of total degree less or equal to m. The projection of L_k^j on the basis of monomials is represented by the matrix $\overline{\mathbf{L}}_k^j \in \mathbb{R}^{N_2\times N_1}$, with $N_1 = (m_1 + n)!/(m_1! n!) \leq N_2 = (m_2 + n)!/(m_2! n!)$, which is so that

$$L_k^j f(\mathbf{x}) = (\overline{\mathbf{L}}_k^j \mathbf{a})^T \mathbf{p}^{m_2}(\mathbf{x}) \tag{12}$$

for all $f(\mathbf{x}) = \mathbf{a}^T \mathbf{p}^{m_1}(\mathbf{x})$. We have also $L_k^j \mathbf{p}^{m_1}(\mathbf{x}) = (\overline{\mathbf{L}}_k^j)^T \mathbf{p}^{m_2}(\mathbf{x})$, which implies that the lth column of $\overline{\mathbf{L}}_k^j$ corresponds to the expansion of $L_k^j p_l$ in the basis of monomials $\mathbf{p}^{m_2}$. Defining an index function $\Psi = (\psi_1(k), \ldots, \psi_n(k))$ that encodes the order of the monomials in the vector $\mathbf{p}$, i.e. $p_k(\mathbf{x}) = x_1^{\psi_1(k)} \cdots x_n^{\psi_n(k)}$, we have

$$L_k^j p_l = \psi_j(l)\, p_{\Psi^{-1}(\Psi(k)+\Psi(l)-\mathbf{e}_j)}$$

where $\mathbf{e}_j \in \mathbb{R}^n$ is the jth unit vector. It follows that the entries of $\overline{\mathbf{L}}_k^j$ are given by

$$\left[\overline{\mathbf{L}}_k^j\right]_{il} = \begin{cases} \psi_j(l) & \text{if } \Psi(i) = \Psi(k) + \Psi(l) - \mathbf{e}_j\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{13}$$

Note that the matrices $\overline{\mathbf{L}}_k^j$ can also be obtained by multiplying a multiplication matrix and a differentiation matrix (see [4] for more details). Finally it follows from (4) and (11) that

$$L = \sum_{j=1}^{n} \sum_{k=1}^{N_F} w_k^j\, L_k^j \tag{14}$$

and the matrix $\overline{\mathbf{L}} \in \mathbb{R}^{N_2\times N_1}$ given by

$$\overline{\mathbf{L}} = \sum_{j=1}^{n} \sum_{k=1}^{N_F} w_k^j\, \overline{\mathbf{L}}_k^j \tag{15}$$

represents the projection of the Koopman infinitesimal generator on the basis of monomials p_k.

2) Computation of the vector field: The coefficients w_k^j of the vector field are finally computed by using the relationship between $\overline{\mathbf{U}}$ (obtained from data in (9)) and $\overline{\mathbf{L}}$ (which depends on w_k^j in (15)).

It can be shown easily (using (3)) that, if $\overline{\mathbf{U}}$ and $\overline{\mathbf{L}}$ represent the projection of U^t and L, respectively, then $\overline{\mathbf{U}} = e^{\overline{\mathbf{L}}t}$. This property and (9) imply that we can obtain an estimation of $\overline{\mathbf{L}}$ from data:

$$\overline{\mathbf{L}}_{data} = \frac{1}{T_s} \log((\mathbf{P_x})^\dagger\, \mathbf{P_y}) \in \mathbb{R}^{N\times N}\,. \tag{16}$$

The function $\log$ denotes the (principal) matrix logarithm.

Remark 1 (Matrix logarithm): The real eigenvalues of $\overline{\mathbf{U}} = (\mathbf{P_x})^\dagger \mathbf{P_y}$ shall not be negative, so that the principal logarithm exists and is real. However, numerical simulations show that $\overline{\mathbf{U}}$ can have negative real eigenvalues when the number of data points is too small. In this case, there exists no real matrix $\log(\overline{\mathbf{U}})$ and the method might fail.

When the Koopman infinitesimal generator admits complex eigenvalues, the sampling period T_s must be sufficiently small so that the eigenvalues of $T_s\overline{\mathbf{L}}$ lie in the strip $\{z \in \mathbb{C} : |\Im\{z\}| < \pi\}$. If this condition is not satisfied, $\log(\exp(T_s\overline{\mathbf{L}})) \neq T_s\overline{\mathbf{L}}$ and the accuracy of the method is altered. This issue is discussed with more details in [13]. $\diamond$

Using (15), we obtain the equality

$$\overline{\mathbf{L}} = \sum_{j=1}^{n}\sum_{k=1}^{N_F} w_k^j \, \overline{\mathbf{L}}_k^j = \hat{\overline{\mathbf{L}}}_{data} \tag{17}$$

where $\hat{\overline{\mathbf{L}}}_{data}$ is the $N_2 \times N_1$ matrix constructed with the N_1 columns and N_2 rows of $\overline{\mathbf{L}}_{data}$ associated with monomials of total degree less or equal to m_1 and m_2, respectively (note that we will typically choose $N = N_2 \geq N_1$ for the size of $\overline{\mathbf{U}}$ and $\overline{\mathbf{L}}_{data}$). By disregarding the $N_2 - N_1$ remaining columns of $\overline{\mathbf{L}}_{data}$, we only consider monomials that are mapped by L onto the span of basis monomials (of total degree less or equal to m_2), for which the finite Galerkin projection (i.e. the identity in this case) is exact.

The matrix equality (17) corresponds to a linear set of equations (with nN_F unknowns w_k^j) that is overdetermined. Its least squares solution yields an approximation of the coefficients w_k^j:

$$\begin{pmatrix} w_1^1 \\ \vdots \\ w_{N_F}^n \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{vec}(\overline{\mathbf{L}}_1^1) & \dots & \mathbf{vec}(\overline{\mathbf{L}}_{N_F}^n) \\ | & & | \end{pmatrix}^\dagger \mathbf{vec}(\hat{\overline{\mathbf{L}}}_{data}) \tag{18}$$

where $\mathbf{vec}$ stands for the vectorization of the matrix. Note that more advanced techniques could be used to solve (17), for instance promoting sparsity of the vector of unknowns w_k^j by adding a L^1-norm penalty term to (17), see e.g. the LASSO method [11]. We leave this for future research.

Remark 2: Each entry of the matrix $\overline{\mathbf{L}}$ yields an equality in (18). However, if a given entry of $\overline{\mathbf{L}}_k^j$ is zero for all j and k, then the corresponding entry of $\overline{\mathbf{L}}$ does not depend on the coefficients w_k^j and the related equality in (18) can be disregarded. In particular, when $m_1 = 1$, the number of effective equalities is equal to nN_F and the problem is not overdetermined. Moreover, (17) can be solved directly in this case since each coefficient w_k^j is equal to an entry of $\overline{\mathbf{L}}$. $\diamond$

Remark 3: The identification problem could also be performed at the level of the Koopman semigroup. However solving the equality $\overline{\mathbf{U}} = e^{\overline{\mathbf{L}}T_s}$ (with a square matrix $\overline{\mathbf{L}}$) amounts to solving a (nonconvex) nonlinear least squares problem. Numerical simulations suggest that better results are obtained by solving (17). $\diamond$

IV. Extensions to open systems

In this section, we show that the proposed method is well-suited to identify open systems that are driven by a known input or by a white noise (i.e. process noise).

A. Systems with inputs

Consider an open dynamical system of the form

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u}(t)) \tag{19}$$

with $\mathbf{x} \in \mathbb{R}^n$ and with the input $u \in \mathcal{U} : \mathbb{R}^+ \to \mathbb{R}^p$. We define the associated flow $\varphi : \mathbb{R}^+ \times \mathbb{R}^n \times \mathcal{U}$ so that $t \mapsto \varphi(t, \mathbf{x}, \mathbf{u}(\cdot))$ is a solution of (19) with the initial condition $\mathbf{x}$ and the input $\mathbf{u}(\cdot)$. Following the generalization proposed in [7] for discrete-time dynamical systems, we consider observables $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ and define the semigroup of Koopman operators

$$U^t f(\mathbf{x}, \mathbf{u}) = f(\varphi^t(\mathbf{x}, \mathbf{u}(\cdot) = \mathbf{u}), \mathbf{u})$$

where $\mathbf{u}(\cdot) = \mathbf{u}$ is a constant input. In this case, $\mathbf{u}$ can be considered as additional state variables and the above operator is the classical Koopman operator for the augmented system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \mathbf{u})$, $\dot{\mathbf{u}} = \mathbf{0}$. In particular, the infinitesimal generator is still given by (4).

It follows that the method proposed in Sections III-A and III-B can be used if the input can be considered as constant between two snapshots (zero-order hold assumption), or equivalently if the sampling rate is high enough. The matrix $\overline{\mathbf{U}}$ is obtained with snapshot pairs $([\mathbf{x}_k, \mathbf{u}_k], [\mathbf{y}_k, \mathbf{u}_k]) \in \mathbb{R}^{(n+p)\times 2}$ and the rest of the procedure follows on similar lines with the augmented state space $\mathbb{R}^{n+p}$. The efficiency of the identification method in the case of systems with inputs is shown in Section V-B.

B. Process noise

We have considered so far only measurement noise. We will show that the proposed method is also robust to process noise. Consider a system described by the stochastic differential equation

$$\dot{x}_k = F_k(\mathbf{x}) + \eta_k(t) \tag{20}$$

where $\eta_k(t)$ is a white noise that satisfies $\mathbb{E}[\eta_k(s)\eta_j(t)] = \sigma_{proc}^2 \delta_{kj} \delta(t - s)$ (where $\mathbb{E}$ denotes the expectation). We define the flow $\varphi : \mathbb{R}^+ \times \mathbb{R}^n \times \Omega$, where Ω is the probability space, such that $t \mapsto \varphi(t, \mathbf{x}, \omega)$ is a solution of (20). In this case, the semigroup of Koopman operators is defined by (see e.g. [5])

$$U^t f(\mathbf{x}) = \mathbb{E}[f(\varphi(t, \mathbf{x}, \omega))]$$

and its infinitesimal generator is given by

$$Lf = \mathbf{F} \cdot \nabla f + \frac{\sigma_{proc}^2}{2} \Delta f$$

where $\Delta = \sum_k \partial^2 / \partial x_k^2$ denotes the Laplacian operator that accounts for diffusion. Note that the infinitesimal generator is related to the so-called Kolmogorov backward equation.

Now we can show that the numerical scheme of the proposed identification method does not need to be adapted

to take account of process noise. As explained in [12], the first step of the method (Section III-A) is still valid for identifying the matrix $\overline{\mathbf{U}}$. In the second step (Section III-B), the procedure is the same, except that one has to consider the Laplacian operator whose projection on the basis of monomials is represented by a matrix $\overline{\mathbf{D}} \in \mathbb{R}^{N_2 \times N_1}$. The equality (17) is then replaced by

$$\sum_{j=1}^{n} \sum_{k=1}^{N_F} w_k^j \, \overline{\mathbf{L}}_k^j + \frac{\sigma^2}{2} \overline{\mathbf{D}} = \hat{\overline{\mathbf{L}}}_{data}$$

where σ is an additional unknown. While the operators L_k^j map monomials of total degree m to monomials of total degree greater or equal to $m-1$, the Laplacian operator maps monomials of total degree m to monomials of total degree $m-2$. Therefore all nonzero entries of $\overline{\mathbf{D}}$ correspond to zero entries of $\overline{\mathbf{L}}_k^j$, so that the addition of the diffusion term only modifies entries of $\overline{\mathbf{L}}$ which are zero—and do not depend on w_k^j—when there is no process noise. In other words, the diffusion term does not affect the equalities on w_k^j, whose solution is still given by (18). In Section V-B, an example illustrates the robustness of the method against process noise.

V. Numerical examples

In this section, we illustrate our numerical method with several examples. We show that the method is efficient to recover the vector field of different kinds of systems, including unstable, chaotic, and open systems. The method is also applied to network reconstruction.

For all the simulations, we consider measurement noise with standard deviation $\sigma_{meas} = 0.01$. We also choose the parameters $m_1 = 1$ and $m_F = 3$, so that $m_2 = m_1 + m_F - 1 = 3$.

A. Periodic, unstable, and chaotic systems

We first consider three systems that exhibit different kinds of behavior.

a) Van der Pol oscillator: The dynamics are given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= (1 - x_1^2) x_2 - x_2 \end{aligned}$$

and possess a stable limit cycle. The data are obtained by taking 3 snapshots at $t \in \{0, 0.5, 1\}$ (sampling period $T_s = 0.5$) for 10 trajectories with initial conditions on $[-1, 1]^2$.

b) Unstable equilibrium: The dynamics are given by

$$\begin{aligned} \dot{x}_1 &= 3\,x_1 + 0.5\,x_2 - x_1 x_2 + x_2^2 + 2\,x_1^3 \\ \dot{x}_2 &= 0.5\,x_1 + 4\,x_2 \end{aligned}$$

and are characterized by an unstable equilibrium at the origin. The data are obtained by taking 2 snapshots at $t \in \{0, 0.1\}$ (sampling period $T_s = 0.1$) for 20 trajectories with initial conditions on $[-1, 1]^2$.

c) Chaotic Lorenz system: The dynamics are given by

$$\begin{aligned} \dot{x}_1 &= 10(x_2 - x_1) \\ \dot{x}_2 &= x_1(28 - x_3) - x_2 \\ \dot{x}_3 &= x_1 x_2 - 8/3\,x_3 \end{aligned}$$

and exhibit a chaotic behavior. The data are obtained by taking 31 snapshots over $[0, 1]$ (sampling period $T_s = 0.033$) for 10 trajectories with initial conditions on $[-20, 20]^3$.

For each model, we identify the coefficient w_k^j of the vector field and compute the root mean square error

$$\text{RMSE} = \sqrt{\frac{1}{n N_F} \sum_{j=1}^{n} \sum_{k=1}^{N_F} \left((w_k^j)_{estim} - (w_k^j)_{exact} \right)^2}.$$

where $(w_k^j)_{estim}$ and $(w_k^j)_{exact}$ are the estimated and exact values of the coefficients w_k^j. In order to compare the results obtained with different systems, we also compute the normalized RMSE: NRMSE = RMSE/$\overline{w}$ where $\overline{w}$ is the average value of the nonzero coefficients $|(w_k^j)_{exact}|$. The RMSE and NRMSE averaged over 10 experiments are given in Table I. The results show that the method performs well, even in the case of unstable and chaotic systems (for which the NRMSE is slightly larger). Note that the accuracy increases with the number of data points (which is low here) but a comprehensive study of this effect is beyond the scope of this paper.

	RMSE	NRMSE
a) Van der Pol	0.035	0.035
b) Unstable equilibrium	0.157	0.092
c) Chaotic Lorenz	0.569	0.074

TABLE I

(Normalized) root mean squared error over 10 simulations.

B. Input and process noise

We consider the forced Duffing system

$$\dot{x}_1 = x_2 \tag{21}$$

$$\dot{x}_2 = x_1 - x_1^3 - 0.2\,x_2 + 0.2\,x_1^2\,u \tag{22}$$

where u is the input. With $u(t) = \cos(t)$, we obtain 51 snapshots over $[0, 10]$ (sampling period $T_s = 0.2$) for 5 trajectories with initial conditions on $[-1, 1]^2$. Our identification method provides a good estimation of the vector field (including the forcing term $0.2\,x_1^2\,u$). The RMSE computed over all coefficients (including those related to the forcing term) and averaged over 10 simulations is equal to 0.025.

We now consider the effect of process noise and replace the forcing term in (21) by a strong white noise with $\sigma_{proc} = 1$ (note that we still add measurement noise with $\sigma_{meas} = 0.01$). Data points are obtained as previously, but for 10 trajectories computed with the Euler-Maruyama scheme. The method is robust against process noise, and recovers the vector field with a RMSE (averaged over 10 simulations) equal to 0.078.

C. Network reconstruction

We consider a 12-dimensional system of the form

$$\dot{x}_j = -\xi_j x_j + \sum_{k=1}^{3} \zeta_{j,k}\, x_{\nu^k_{j,1}}^{\sigma^k_{j,1}} x_{\nu^k_{j,2}}^{\sigma^k_{j,2}} \qquad j = 1, \ldots, 12 \quad (23)$$

where the coefficients ξ_j are chosen according to a uniform distribution on $[0, 1]$ and $\zeta_{j,k}$ are distributed according to a Gaussian distribution of zero mean and standard deviation equal to one. The subscripts $\nu^k_{j,1}$ and $\nu^k_{j,2}$ are integers randomly chosen on $\{1, \ldots, 12\}$ and the exponents $\sigma^k_{j,1}$ and $\sigma^k_{j,2}$ are also integers randomly chosen on $\{0, 1, 2, 3\}$ so that $\sigma^k_{j,1} + \sigma^k_{j,2} \in \{2, 3\}$. The first term in (23) is a linear term that ensures local stability of the origin. The other terms are quadratic and cubic nonlinearities that define an interaction network between the different states. We say that there is a directed link from x_l to x_j if $\nu^k_{j,r} = l$ and $\sigma^k_{j,r} \neq 0$ for some $k \in \{1, 2, 3\}$ and $r \in \{1, 2\}$.

For 500 different trajectories with initial conditions on $[-1, 1]^{12}$, we consider three snapshots obtained at time $t \in \{0, 0.5, 1\}$ (sampling period $T_s = 0.5$). Figure 2 shows that, with this dataset, the method provides an accurate estimation of all the coefficients of the polynomial vector field (RMSE $= 0.021$). Note that we consider a total of 1500 data points to identify $12N_F = 5460$ coefficients w^j_k. As we decrease the number of data points, results are less accurate but still acceptable (RMSE $= 0.106$ with 400 trajectories and RMSE $= 0.151$ with 300 trajectories). More advanced methods promoting sparsity (e.g. LASSO method [11]) could further improve the accuracy of the results.

Network reconstruction aims at predicting links in the network. Very small coefficients are mainly due to measurement noise and have an exact value equal to zero. Also we consider that there is a link from x_l to x_j if at least one coefficient w^j_k related to a monomial containing x_l is above a given threshold value. With a threshold value equal to 0.1, we obtain a true positive rate (i.e. number of correctly identified links divided by the actual number of links) equal to 0.875 and a false positive rate (i.e. number of incorrectly identified links divided by the actual number of missing links) equal to 0.023. It is noticeable that, in addition to this good performance, the method also provides the nature of the links (e.g. quadratic, cubic) and their weight (i.e. values w^j_k).

VI. Conclusion

We have proposed a novel method for the identification of nonlinear systems. This method is based on a lifting technique which recasts nonlinear system identification as the identification of the linear Koopman operator in an infinite-dimensional space of observables. A key advantage of the method is that it relies only on linear techniques and is an indirect method that does not require the estimation of time derivatives. It follows that the method is well-suited to model low sampling rate data sets. It is also robust to noise and efficient to recover the vector field of different classes of systems.

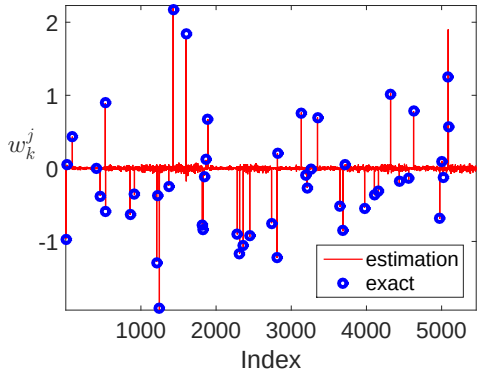

Fig. 2. The method is efficient to recover all the coefficients of the polynomial vector field of a 12-dimensional system. The exact nonzero values w^j_k are depicted in blue. The estimated values are in red.

This paper presents preliminary results that open the door to further developments of Koopman operator lifting techniques for nonlinear system identification (e.g. use other bases to identify non-polynomial vector fields). Since the method is based on the (finite-dimensional) truncation of the (infinite-dimensional) Koopman operator, a theoretical study of its convergence should also be provided in future work.

VII. Acknowledgments

The authors acknowledge support from the Luxembourg National Research Fund (FNR 14/BM/8231540).

References

[1] S. L. Brunton, L. P. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.

[2] Y. Lan and I. Mezić. Linearization in the large of nonlinear systems and Koopman operator spectrum. *Physica D*, 242:42–53, 2013.

[3] A. Lasota and M. C. Mackey. *Chaos, Fractals, and Noise: stochastic aspects of dynamics*. Springer-Verlag, 1994.

[4] A. Mauroy and I. Mezić. Global stability analysis using the eigenfunctions of the Koopman operator. *IEEE Transactions On Automatic Control*, 2016. In press.

[5] I. Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1-3):309–325, 2005.

[6] W. Pan, Y. Yuan, J. Goncalves, and G.-B. Stan. A sparse Bayesian approach to the identification of nonlinear state-space systems. *IEEE Transactions On Automatic Control*, 61(1):182–187, 2016.

[7] L. P. Proctor, S. L. Brunton, and J. N. Kutz. Generalizing Koopman operator theory to allow for inputs and control. http://arxiv.org/abs/1602.07647.

[8] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 641:115–127, 2009.

[9] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.

[10] A. Sootla and A. Mauroy. Operator-theoretic characterization of eventually monotone systems. http://arxiv.org/abs/1510.01149.

[11] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[12] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley. A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. *Journal of Nonlinear Science*, pages 1–40, June 2015.

[13] Z. Yue, J. Thunberg, L. Ljung, and J. Goncalves. Identification of Sparse Continuous-Time Linear Systems with Low Sampling Rate: Exploring Matrix Logarithms. https://arxiv.org/abs/1605.08590.