A Willems' Fundamental Lemma for Nonlinear Systems with Koopman Linear Embedding

Xu Shang[1], Jorge Cortes[2], *Fellow, IEEE*, and Yang Zheng[1], *Member, IEEE*

***Abstract*— In this letter, we prove the data-driven representation adapted from Willems' fundamental lemma is accurate for nonlinear systems that admit a Koopman linear embedding. Precisely, we first propose a necessary and sufficient condition under which the trajectory of the Koopman linear embedding is guaranteed to be valid for the nonlinear system. We then illustrate a rich enough trajectory library of the nonlinear system can fully capture the behavior of the Koopman linear embedding. The existence of the simple-to-build data-driven representation implies we can bypass the non-trivial lifting function selection process. Moreover, our derivation demonstrates the size of the trajectory library of the data-driven representation relates to the dimension of the "hidden" Koopman linear embedding of the nonlinear system. Numerical experiments validate our theoretical findings.**

***Index Terms*— Data-driven control; Willems' Fundamental Lemma; Nonlinear systems; Koopman Operator**

I. Introduction

Designing controllers for nonlinear systems with linear approximations has gained increasing interest. An approximated linear representation enables the utilization of techniques from the linear system tools and facilitates computationally efficient model predictive control schemes [1]. Both the Koopman operator theorem [2] and Willems' fundamental lemma [3] can be applied for the construction of the linear representation and have shown promising prediction performance in real applications [4], [5]. However, ensuring the accuracy of the linear representation remains a critical challenge for both approaches.

The Koopman operator theorem is originally developed for analyzing autonomous systems and then utilized for modeling the controlled system by taking the control sequence as an extended state. The propagation of the nonlinear system is considered as the evolution of lifting functions in a high-dimensional function space. By considering the lifting functions as a new "state", the Koopman linear embedding of the nonlinear system in the high-dimensional space can be obtained. Based on this linear model, many control techniques have been developed such as linear feedback controllers for the linear quadratic regulator [6] and the Koopman model predictive control [1]. However, the prediction accuracy of the Koopman linear embedding highly depends on the choice of the lifting functions, and there is no systematic approach to obtain the exact lifting functions.

Willems' fundamental lemma is first proposed for the deterministic linear time-invariant (LTI) system, which utilizes a rich enough trajectory library to formulate a data-driven representation for the system. It is further extended for nonlinear systems such as Hamerstrin systems and Wiener systems [7] and applied to the class of polynomial systems [8]. Then, the data-driven representation can be utilized for feedback controller design [9] and as the predictive model in model predictive control [10] for the nonlinear system. For the nonlinear or non-deterministic system, it has been found collecting more trajectories with a longer initial trajectory can improve the accuracy of the data-driven representation [11], [12] but its mechanism has not been discussed. Furthermore, the existence of an accurate data-driven representation and the requirement for the trajectory library beyond the deterministic LTI system are not fully explored.

In this letter, we extend Willems' fundamental lemma for nonlinear systems that admit an accurate Koopman linear embedding. Our key idea is to constrain the behavior of the Koopman linear embedding, which has an accurate data-driven representation, to the corresponding trajectory set of the nonlinear system. In particular, our main contributions include: 1) We illustrate the relationship between trajectory spaces of the nonlinear system and its associated Koopman linear embedding. We provide a necessary and sufficient condition for identifying trajectories in the intersection of these spaces. 2) We introduce a new definition of persist excitation for nonlinear systems which relates to lifted states of the Koopman linear embedding. Our derivation shows the behavior of Koopman linear embedding can be fully captured by the linear combination of rich enough nonlinear system's trajectories. 3) We finally prove the data-driven representation adapted from Willems' fundamental lemma is accurate for the nonlinear system with an accurate Koopman linear embedding. Thus, we can directly utilize the simple-to-build data-driven representation as the linear model and bypass the process of choosing the lifting function, which is non-trivial and can lead to bias errors. Moreover, our derivation illustrates the required width and depth of the trajectory library depends on the dimension of the associated Koopman linear embedding, which can be much larger than the original system. Thus,

This work is supported by NSF CMMI-2320697, NSF CAREER 2340713 and an Early Career Faculty Development Award from the Jacobs School of Engineering at UC San Diego.

[1]S. Xu and Y. Zheng are with the Department of Electrical and Computer Engineering, University of California San Diego; {x3shang, zhengy}@ucsd.edu

[2]J. Cotes is with the Department of Mechanical and Aerospace Engineering, University of California San Diego, {cortes}@ucsd.edu

both collecting more trajectories (increasing the width) and enlarging the length of the initial trajectory (increasing the depth) are important, while the latter one is usually ignored in the literature.

The remainder of this letter is structured as follows. Section II briefly reviews basic knowledge of Koopman linear embedding and Willems' fundamental lemma and then provides the problem statement. The proof of an accurate Koopman linear embedding of the nonlinear system leads to an accurate data-driven representation is presented in Section III. Section IV validates our theoretical findings via numerical simulations. We conclude the letter in Section V.

II. Preliminaries and Problem Statement

In this section, we review preliminaries on Koopman linear embedding and the standard Willem's fundamental lemma.

A. Koopman linear models for nonlinear systems

Consider a discrete-time nonlinear system

$$\begin{aligned} x(k+1) &= f(x(k), u(k)), \\ y(k) &= g(x(k), u(k)), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ are the state, input, and output of the system, respectively. One key idea of Koopman operator theory is to lift the state x of the origin nonlinear system to a higher dimensional space via a set of lifting functions (often referred to as observables), where the evolution of these observables is assumed to be linear (or approximately linear). Koopman operator theory for autonomous systems with no control is well developed [2], and there are different ways of generalizing the Koopman operator to controlled systems

In this paper, we consider an important case of Koopman linear embedding for nonlinear systems.

Definition 1 (Koopman Linear Embedding): The nonlinear system (1) is called to admit a Koopman linear embedding if there exists a set of linearly independent lifting functions $\phi_1(\cdot), \phi_2(\cdot), \ldots, \phi_{n_z}(\cdot) : \mathbb{R}^n \to \mathbb{R}$ such that the lifted state

$$\Phi(x(k)) := \mathrm{col}(\phi_1(x(k)), \ldots, \phi_{n_z}(x(k))) \in \mathbb{R}^{n_z}, \tag{2}$$

propagates linearly along all trajectories of (1) and the output $y(k)$ is a linear map of $\Phi(x(k))$ and $u(k)$.

For a nonlinear system admitting a Koopman linear embedding (2), the new lifted state $z(k) = \Phi(x(k)) \in \mathbb{R}^{n_z}$ satisfies

$$\begin{aligned} z(k+1) &= Az(k) + Bu(k), \\ y(k) &= Cz(k) + Du(k), \end{aligned} \tag{3}$$

with matrices A, B, C, D having appropriated dimensions. Note that we normally have $n_z \gg n$ and the matrix pair (A, B) and (C, A) in (3) may not be controllable or observable.

In the case that an exact Koopman linear embedding does not exist, many existing studies (especially in predictive control) often use the linear model (3) to approximate the dynamics of the observables (2); see [1].

After choosing the observables (2), we can compute the matrices A, B, C and D for the linear model (3) using extended dynamic model decomposition (EDMD) [13]. We organize the measured input-state-output data points of (1) as

$$\mathrm{X} = \begin{bmatrix} x_1, \ldots, x_{n_\mathrm{d}} \end{bmatrix}, \quad \mathrm{X}^+ = \begin{bmatrix} x_1^+, \ldots, x_{n_\mathrm{d}}^+ \end{bmatrix},$$
$$\mathrm{U} = \begin{bmatrix} u_1, \ldots, u_{n_\mathrm{d}} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1, \ldots, y_{n_\mathrm{d}} \end{bmatrix},$$

where $x_i^+ = f(x_i, u_i), y_i = g(x_i, u_i), i = 1, \ldots, n_\mathrm{d}$. These data points do not necessarily come from a single trajectory. With the lifting functions (2), we compute the lifted state as

$$\mathrm{Z} = \begin{bmatrix} \Psi(x_1), \ldots \Psi(x_{n_\mathrm{d}}) \end{bmatrix}, \quad \mathrm{Z}^+ = \begin{bmatrix} \Psi(x_1^+), \ldots \Psi(x_{n_\mathrm{d}}^+) \end{bmatrix}.$$

Then, we obtain the matrices A, B, C and D via two least-squares approximations:

$$\begin{aligned} (A, B) &\in \underset{A,B}{\mathrm{argmin}} \ \|\mathrm{Z}^+ - A\mathrm{Z} - B\mathrm{U}\|_F^2, \\ (C, D) &\in \underset{C,D}{\mathrm{argmin}} \ \|\mathrm{Y} - C\mathrm{Z} - DU\|_F^2. \end{aligned} \tag{4}$$

It is clear that the set of observables affects (4) significantly. Even if an exact Koopman linear embedding exists for (1), we may not know the observables (2) for such a Koopman linear embedding. An inexact choice can lead to significant modeling errors (see [14], [15]). In the literature, common choices for (2) are Gaussian kernel, polyharmonic splines, thin plate splines [16], or the RBF neural network [17]. However, none of them can guarantee to obtain an exact linear model even when a Koopman linear embedding exists.

B. Willems' Fundamental Lemma

Willems' fundamental lemma is established for linear time-invariant (LTI) system of the form

$$\begin{aligned} x(k+1) &= A_1 x(k) + B_1 u(k), \\ y(k) &= C_1 x(k) + D_1 u(k), \end{aligned} \tag{5}$$

where the state, input and output of the system are denoted as $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$, respectively. We consider the LTI system (5) from the behavior (*i.e.*, trajectory) perspective. The key idea is that a linear combination of rich enough offline trajectories of (5) is equivalent to its whole trajectory space.

Let us recall the notion of persistent excitation [3].

Definition 2 (Persistently exciting): The length-T sequence $\omega = \mathrm{col}(\omega(1), \omega(2), \ldots, \omega(T))$ is persistently exciting of order L if its Hankel matrix has full row rank:

$$\mathcal{H}_L(\omega) = \begin{bmatrix} \omega(1) & \omega(2) & \cdots & \omega(T-L+1) \\ \omega(2) & \omega(3) & \cdots & \omega(T-L+2) \\ \vdots & \vdots & \ddots & \vdots \\ \omega(L) & \omega(L+1) & \cdots & \omega(T) \end{bmatrix}.$$

With the measured input-state-output data in sequence, *i.e.*, $u_\mathrm{d} = \mathrm{col}(u(0), \ldots, u(n_\mathrm{d}-1)), x_\mathrm{d} = \mathrm{col}(x(0), \ldots, x(n_\mathrm{d}-1)), y_\mathrm{d} = \mathrm{col}(y(0), \ldots, y(n_\mathrm{d}-1)$, the following Willems' fundamental lemma is adapted from [18, Theorem 1].

Lemma 1 (Willems' fundamental lemma): Consider the LTI system (5). Assume the Hankel matrix formed by its pre-collected trajectory $H_0 := \mathrm{col}(\mathcal{H}_1(x_{\mathrm{d},1:n_\mathrm{d}-L+1}), \mathcal{H}_L(u_\mathrm{d}))$ has full row rank. Then, a length L input-output sequence $\mathrm{col}(u, y) \in \mathbb{R}^{(m+p)L}$ is a valid trajectory of (5) if and only if there exists a $g \in \mathbb{R}^{n_\mathrm{d}-L+1}$ such that

$$\begin{bmatrix} \mathcal{H}_L(u_\mathrm{d}) \\ \mathcal{H}_L(y_\mathrm{d}) \end{bmatrix} g = \begin{bmatrix} u \\ y \end{bmatrix}.$$

Lemma 1 does not require the controllability of (5) since it directly imposes a full-rank condition on H_0 that involves the state sequence. If (5) is controllable, then the persistently excitation of order $L+n$ for the input sequence u_d is sufficient to guarantee the full rank of H_0 [18].

Utilizing Lemma 1, we can build a data-driven representation for system (5). We use $u_{\text{ini}} = \text{col}(u(k-T_{\text{ini}}), u(k-T_{\text{ini}}+1), \ldots, u(k-1))$ and $u_{\text{F}} = \text{col}(u(k), u(k+1), \ldots, u(k+N-1))$ to represent the most recent past input trajectory of length T_{ini} and the future input trajectory of length N, respectively (similarly for $y_{\text{ini}}, y_{\text{F}}$). Accoridingly, let us partition the Hankel matrix from its pre-collected data as

$$\begin{bmatrix} U_{\text{P}} \\ U_{\text{F}} \end{bmatrix} := \mathcal{H}_L(u_{\text{d}}), \quad \begin{bmatrix} Y_{\text{P}} \\ U_{\text{F}} \end{bmatrix} := \mathcal{H}_L(y_{\text{d}}).$$

From Lemma 1, the sequence $\text{col}(u_{\text{ini}}, y_{\text{ini}}, u_{\text{F}}, y_{\text{F}})$ is a valid trajectory of (5) if and only if there exists a $g \in \mathbb{R}^{T-T_{\text{ini}}-N+1}$ such that

$$\text{col}(U_{\text{P}}, Y_{\text{P}}, U_{\text{F}}, Y_{\text{F}})g = \text{col}(u_{\text{ini}}, y_{\text{ini}}, u_{\text{F}}, y_{\text{F}}). \tag{6}$$

Furthermore, if (5) is observable and T_{ini} is no smaller than its observability index, then y_{F} in (6) is unique given an initial trajectory $(u_{\text{ini}}, y_{\text{ini}})$ and any future input u_{F} [19]. Intuitively, if (5) is observable, the initial trajectory $(u_{\text{ini}}, y_{\text{ini}})$ allows us to uniquely determine the corresponding initial state. This data-driven representation (6) has been widely used in predictive control [10] with many successful applications [20], [21].

C. Problem Statement

In this paper, we aim to extend the data-driven representation (6) from LTI systems to nonlinear systems (1) that admit a Koopman linear embedding (see Definition 1). At this stage, one may be tempted to directly apply Willems' fundamental lemma to the Koopman linear model (3) and get a similar data-driven representation as (6). However, there are two unsolved challenges for this process:

1) the Koopman linear model (3) may be neither controllable nor observable;
2) the behavior space of the Koopman linear model (6) is much larger than the behavior space of the original nonlinear system (1).

We propose two innovations to resolve the challenges above. 1) We first characterize the relationship between the behavior space of the Koopman linear model (6) and that of the original nonlinear system (1). A key insight of this characterization is that observability is not needed for the data-driven representation (6) as long as the length of the initial trajectory is large enough, *i.e.*, the Hankel matrix has a sufficient depth. 2) We introduce a new notion of persistent excitation for the offline data collection, which has the same spirit as [22, Def. 1] that focuses on a special case of affine systems. With these two technical tools, we will establish a direct data-driven representation in the form of (6) for nonlinear systems that admit a Koopman linear embedding. This representation requires no knowledge of the lifting functions (2) (as long as they exist). Our representation can be directly utilized in Koopman model predictive control [1], but with no need of identifying the linear model (3). This is very beneficial as it eliminates the bias errors in choosing the lifting functions (or observables) which is non-trivial.

III. From Koopman Linear Embedding to Data-driven Representation

In this section, we develop the main technical result that directly represents the nonlinear system with Koopman linear embedding using its input and output data. We refer to it as an extended Willems' fundamental lemma. We also discuss a special case of affine systems [22].

A. Two behavior spaces

Consider the space of length-L trajectories for the nonlinear system (1) and the Koopman linear embedding (3):

$$\mathcal{B}_1|_L = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \mathbb{R}^{(m+p)L} \mid \exists\, x(0) = x_0 \in \mathbb{R}^n, \text{(1) holds} \right\}, \tag{7a}$$

$$\mathcal{B}_2|_L = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \mathbb{R}^{(m+p)L} \mid \exists\, z(0) = z_0 \in \mathbb{R}^{n_z}, \text{(3) holds} \right\}. \tag{7b}$$

It is clear that $\mathcal{B}_1|_L$ is a nonlinear set while $\mathcal{B}_2|_L$ is a linear subspace in $\mathbb{R}^{(m+p)L}$. We have $\mathcal{B}_1|_L \subseteq \mathcal{B}_2|_L$, *i.e.*, the behavior space of the Koopman linear embedding is larger than that of the original nonlinear system. Our first result characterizes the relationship between these two behavior spaces.

Theorem 1: Consider the nonlinear system (1). Suppose it admits a Koopman linear embedding (3).

1) We have $\mathcal{B}_1|_L \subset \mathcal{B}_2|_L, \forall L \geq 1$, *i.e.*, all trajectories of system (1) are also trajectories of (3);
2) Let $\text{col}(u, y) \in \mathcal{B}_2|_L$, where $L \geq n_z$. Then, $\text{col}(u, y) \in \mathcal{B}_1|_L$ if and only if its leading sequence of length n_z (denoted as $\text{col}(u_{1:n_z}, y_{1:n_z})$) is a valid trajectory of (1).

Theorem 1 reveals that while the space $\mathcal{B}_2|_L$ is larger, we can characterize its subset corresponding to $\mathcal{B}_1|_L$ using the initial leading sub-sequence $\text{col}(u_{1:n_z}, y_{1:n_z})$. We need a technical lemma to prove the second statement in Theorem 1.

Lemma 2: Consider an LTI system (5) that may be uncontrollable or unobservable. Fix an initial trajectory $\text{col}(u_{1:L_1}, y_{1:L_1}) \in \mathbb{R}^{(m+p)L_1}$ of length $L_1 \geq n$. Given any subsequent input $u_{L_1+1:L} \in \mathbb{R}^{m(L-L_1)}$ (future input), the subsequent output $y_{L_1+1:L} \in \mathbb{R}^{p(L-L_1)}$ (future output) is unique.

Note that Lemma 2 works for any LTI systems and requires no observability or controllability. This result is not difficult to establish and we put its proof details in the appendix.

Proof of Theorem 1: The first statement is obvious from Definition 1. Let $\text{col}(u, y) \in \mathcal{B}_1|_L$ be arbitrary. By definition, we can find $x_0 \in \mathbb{R}^n$ such that $\text{col}(u, y)$ satisfies the evolution in (1). Then, with the lifted initial state $z_0 = \Phi(x_0) \in \mathbb{R}^{n_z}$, $\text{col}(u, y)$ satisfies the evolution in (3). Thus, $\text{col}(u, y) \in \mathcal{B}_2|_L$.

For the second statement, the "only if" part is trivial as $\text{col}(u_{1:n_z}, y_{1:n_z})$ is part of $\text{col}(u, y)$. We here prove "if" part. Suppose $\text{col}(u_{1:n_z}, y_{1:n_z}) \in \mathcal{B}_1|_{n_z}$ and we let $\tilde{y}_{n_z+1:L}$ be the corresponding outputs from the nonlinear system (1) for the rest of inputs $u_{n_z+1:L}$, *i.e.*,

$$\text{col}(u, y_{1:n_z}, \tilde{y}_{n_z+1:L}) \in \mathcal{B}_1|_L \,.$$

Then, it is clear that $\text{col}(u, y_{1:n+1}, \tilde{y}_{n+2:L}) \in \mathcal{B}_2|_L$ utilizing the first statement. From Lemma 2, the outputs of the linear system (3) are uniquely determined by $u_{n_z+1:L}$ when $\text{col}(u_{1:n_z}, y_{1:n_z}) \in \mathcal{B}_2|_{n_z}$ are given. Thus, we must have $\tilde{y}_{n_z+1:L} = y_{n_z+1:L}$. This indicates that the whole trajectory satisfies $\text{col}(u, y) \in \mathcal{B}_1|_L$. This completes the proof. □

We might be tempted to estimate an initial state $x(0)$ or $z(0)$ from $\mathrm{col}(u_{1:n_z}, y_{1:n_z}) \in \mathcal{B}_1 |_{n_z}$. However, since we do not assume observability of the Koopman linear embedding (3), the initial state $z(0)$ cannot be uniquely determined from $\mathrm{col}(u_{1:n_z}, y_{1:n_z})$. As confirmed in Lemma 2, the unobservable part of the initial state does not affect the uniqueness of input-output trajectory.

B. Data-driven representation of nonlinear systems

While the trajectory space of the Koopman linear embedding (3) is larger than that of the nonlinear system (1), we can use the trajectories from (1) (*i.e.*, $\mathrm{col}(u_{\mathrm{d}}^i, y_{\mathrm{d}}^i) \in \mathcal{B}_1 |_L, i = 1, \dots, l$) that are rich-enough to represent $\mathcal{B}_2|_L$. For this, we propose the following definition of persistence of excitation.

Definition 3: Consider a nonlinear system (1) with a Koopman linear embedding (3). We say l trajectories of length-L from (1), $\mathrm{col}(u_{\mathrm{d}}^i, y_{\mathrm{d}}^i) \in \mathcal{B}_1|_L, i = 1, \dots, l$ with $l \geq mL + n_{\mathrm{z}}$ are persistently exciting of order L, if the following matrix

$$H_K := \begin{bmatrix} u_{\mathrm{d}}^1 & u_{\mathrm{d}}^2 & \dots & u_{\mathrm{d}}^l \\ \Phi(x_0^1) & \Phi(x_0^2) & \dots & \Phi(x_0^l) \end{bmatrix} \in \mathbb{R}^{(mL+n_z)\times l} \tag{8}$$

has full row rank, where $x_0^i \in \mathbb{R}^n$ is the initial state for each trajectory $\mathrm{col}(u_{\mathrm{d}}^i, y_{\mathrm{d}}^i), i = 1, \dots, l$.

This notion of persistent excitation is similar to [22, Def. 1] that focuses on only affine systems. Our notion is suitable for nonlinear systems with a Koopman linear embedding. If the Koopman liner embedding (3) is controllable, then a persistently exciting of order $L + n_{\mathrm{z}}$ for the input sequence is sufficient to gurantee that (8) has full row rank. If we collect multiple trajectories (see [18]), we require the multiple input sequences $u_{\mathrm{m}}^1, \dots, u_{\mathrm{m}}^q$ to be collectively persistently exciting of order $L + n_{\mathrm{z}}$, that is

$$\mathrm{rank}\big(\left[\mathcal{H}_{L+n_z}(u_{\mathrm{m}}^1), \dots, \mathcal{H}_{L+n_z}(u_{\mathrm{m}}^q) \right] \big) = m(L + n_z).$$

Theorem 2: Consider a nonlinear system (1) with a Koopman linear embedding (3). Suppose that l trajectories of length-L from (1), $\mathrm{col}(u_{\mathrm{d}}^i, y_{\mathrm{d}}^i) \in \mathcal{B}_1|_L, i = 1, \dots, l$, are persistently exciting of order L. Then, a length-L sequence $\mathrm{col}(u, y)$ is a valid trajectory of the Koopman linear embedding (3) if and only if there exists $g \in \mathbb{R}^l$ such that

$$H_{\mathrm{d}} g = \mathrm{col}(u, y)$$

where

$$H_{\mathrm{d}} := \begin{bmatrix} u_{\mathrm{d}}^1 & u_{\mathrm{d}}^2 & \dots & u_{\mathrm{d}}^l \\ y_{\mathrm{d}}^1 & y_{\mathrm{d}}^2 & \dots & y_{\mathrm{d}}^l \end{bmatrix} \in \mathbb{R}^{(m+p)L\times l}. \tag{9}$$

The proof of this result is postponed to the appendix. Theorem 2 allows us to use the trajectories from the nonlinear system (1) that are persistently exciting of order L to represent any length-L trajectory of the Koopman linear embedding (3).

Combining Theorem 2 and Theorem 1 leads to a direct data-driven representation for the nonlinear system (1) that admits a Koopman linear embedding (3). We collect a trajectory library $H_{\mathrm{d}} = \mathrm{col}(U_{\mathrm{d}}, Y_{\mathrm{d}})$ in (9), where each column is a trajectory of the nonlinear system (1) with length $L = T_{\mathrm{ini}} + N$. We partition matrices U_{d} and Y_{d} as

$$\begin{bmatrix} U_{\mathrm{P}} \\ U_{\mathrm{F}} \end{bmatrix} := U_{\mathrm{d}}, \quad \begin{bmatrix} Y_{\mathrm{P}} \\ Y_{\mathrm{F}} \end{bmatrix} := Y_{\mathrm{d}}, \tag{10}$$

where U_{P} and U_{F} consist of the first T_{ini} rows and the last N rows of U_{d}, respectively (similarly for Y_{P} and Y_{F}).

Theorem 3: Consider a nonlinear system (1) with a Koopman linear embedding (3). We collect a data library H_{d} in (9) with $l \geq mL + n_{\mathrm{z}}$ trajectories of length $L = T_{\mathrm{ini}} + N$ where $T_{\mathrm{ini}} \geq n_{\mathrm{z}}$. Suppose these l trajectories are persistently exciting of order L. At time k, denote the most recent input-output sequence $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})$ with length T_{ini} from (1) as

$$u_{\mathrm{ini}} = \mathrm{col}(u(k - T_{\mathrm{ini}}), u(k - T_{\mathrm{ini}} + 1), \dots, u(k - 1)),$$
$$y_{\mathrm{ini}} = \mathrm{col}(y(k - T_{\mathrm{ini}}), y(k - T_{\mathrm{ini}} + 1), \dots, u(k - 1)),$$

For any future input $u_{\mathrm{F}} = \mathrm{col}(u(k), u(k + 1), \dots, u(k + N - 1))$, the sequence $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u_{\mathrm{F}}, y_{\mathrm{F}})$ is a valid length-L trajectory of (1) if and only if there exists $g \in \mathbb{R}^l$ such that

$$\mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}, U_{\mathrm{F}}, Y_{\mathrm{F}}) g = \mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u_{\mathrm{F}}, y_{\mathrm{F}}). \tag{11}$$

Proof: This result is a combination of Theorems 1 and 2. Since the pre-collected data is persistently exciting of order L, Theorem 2 confirms that $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u_{\mathrm{F}}, y_{\mathrm{F}})$ is a valid trajectory with length $L = T_{\mathrm{ini}} + N$ of the Koopman linear embedding (3) if and only if there exists a vector $g \in \mathbb{R}^l$ such that (11) holds. In addition, Theorem 1 guarantees that the length-L trajectory $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u_{\mathrm{F}}, y_{\mathrm{F}})$ of the Koopman linear embedding (3) is a valid trajectory of the nonlinear system (1) if and only if $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})$ is a trajectory of the nonlinear system (1). This is obvious true. ■

Theorem 3 gives a direct data-driven representation of nonlinear systems with a Koopman linear embedding from its input and output data. This data-driven representation requires no knowledge of the lifting functions (2) as long as they exist. Also, we do not require the Koopman linear embedding (3) to be controllable or observable. Two key enablers for Theorem 3 are 1) our notion of persistent excitation for nonlinear systems in Definition 3 that enables Theorem 2, and 2) a sufficiently long initial trajectory $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})$ from the nonlinear system that ensures Theorem 1.

We remark that Theorem 3 illustrates the benefits and importance of increasing the width and depth of the trajectory library. While the benefits of increasing width is well-recognized in the literature as a fact, its mechanism remains unclear and the importance of enlarging the depth is often overlooked in the literature. Precisely, collecting more trajectories to increase the width ensures the satisfaction of the persistent excitation for nonlinear systems (corresponds to Theorem 2). On the other hand, fixing the prediction horizon N, large enough depth ensures the initial trajectory is sufficiently long, under which the trajectory in the space of the Koopman linear embedding is also a valid trajectory for the nonlinear system (corresponds to Theorem 1). Our derivation shows the required width and depth of the trajectory library depend on the dimension of the "hidden" Koopman linear embedding of the nonlinear system which can be much larger than the dimension of the original system. Thus, more columns are needed to fully capture the behavior of the Koopman linear embedding and more rows are required to ensure the trajectory is valid for the nonlinear system. Moreover, Theorem 3 shows the data-driven representation is equivalent to the Koopman linear embedding which can be directly integrated with pre-

dictive control as

$$\begin{aligned} \min_{u\in\mathcal{U},y} \quad & \|u\|_R^2 + \|y - y_\mathrm{r}\|_Q^2 \\ \text{subject to} \quad & (11) \end{aligned} \tag{12}$$

where $R, Q \succeq 0$ and y_r denotes the reference output trajectory. Thus, there is no need to choose the lifting function which is required by most existing Koopman model predictive control literature and can lead to bias errors.

Remark 1: The existing literature on extending Willems' fundamental lemma to nonlinear systems [7], [23] requires prior knowledge of the system, with additional constraints imposed to account for the system's nonlinear structure in the final representation. Typically, the nonlinearities exist in the input or output (*e.g.*, Hamerstrin systems and Wiener systems). We forcus on a different class of nonlinear systems where nonlinearity contains in system dynamics and the structure is unknown. We also do not introduce additional constraints in the final representation. Regarding the connection between the data-driven representation and the Koopman linear embedding, the work [24] mainly integrates Willems' fundamental lemma with the learning of lifting functions while we show learning these lifting functions is redundant under certain conditions.

C. A special case: Affine systems

We here demonstrate the affine system can be considered as a special case in our framework and compare it with the existing result in [22]. The discrete time-invariant affine system has the following form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + e, \\ y(k) &= Cx(k) + Du(k) + r, \end{aligned} \tag{13}$$

where $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^m, y(k) \in \mathbb{R}^p$ are the state, input, output at time k respectively and $e \in \mathbb{R}^n, r \in \mathbb{R}^p$ are two constant vectors. We here present the extension of Willems' fundamental lemma for affine system proposed in [22, Theorem 1] for the completeness of our work.

Theorem 4: Suppose the pre-collected trajectory $u_\mathrm{d}, x_\mathrm{d}, y_\mathrm{d}$ of (13) satisfies $H_A := \mathrm{col}(\mathcal{H}_L(u_\mathrm{d}), \mathcal{H}_1(x_{\mathrm{d},1:n_\mathrm{d}-L+1}), \mathbb{1})$ has full row rank where $\mathbb{1}$ represents matrix with captible size and all its elements are 1. The length-L input-output sequence $\mathrm{col}(u, y) \in \mathbb{R}^{(m+p)L}$ is a valid trajectory of (13) if and only if there exists a $g \in \mathbb{R}^{n_\mathrm{d}-L+1}$ such that

$$\mathrm{col}(U_\mathrm{P}, Y_\mathrm{P}, U_\mathrm{F}, Y_\mathrm{F})g = \mathrm{col}(u_\mathrm{ini}, y_\mathrm{ini}, u_\mathrm{F}, y_\mathrm{F}), \ \sum g = 1. \tag{14}$$

In our framework, since the affine system has a corresponding Koopman linear embedding, the affine constraint in (13) can be eliminated if the initial trajectory is sufficiently long. Precisely, we can choose a vector of lifting functions that is $z(x) = [\phi_1(x), \dots, \phi_n(x), 1]^\mathsf{T}$ where $\phi_i(x) : \mathbb{R}^n \to \mathbb{R}$ is the i-th element of the state, *i.e.*, $\phi_i(x) = x_i$. Then, the Koopman linear embedding for the affine system (13) is as follows:

$$\begin{aligned} z(k+1) &= \begin{bmatrix} A & e \\ \mathbb{0} & 1 \end{bmatrix} z(k) + \begin{bmatrix} B \\ \mathbb{0} \end{bmatrix} u(k), \\ y(k) &= \begin{bmatrix} C & r \end{bmatrix} z(k) + Du(k). \end{aligned} \tag{15}$$

Under the condition that the data requirement in Theorem 3 is satisfied, (6) is guaranteed to be an accurate data-driven representation for the affine system (13).

We then compare the data requirement for utilizing Theorem 3 and Theorem 4 to construct the accurate data-driven representation of the affine system. With the specific choice of lifting functions, H_A and H_K are equivalent in this case. Thus, for controllable matrix pair (A, B), the data requirements of them with length-L trajectory are the same which is to make u_d to be persistently exciting of order $L + n + 1$ [25]. The minimum data length satisfies the requirement is to make $\mathcal{H}_{L+n+1}(u_\mathrm{d})$ a square matrix that is $(m+1)(L+n)+m$. For fixed prediction horizon N, the minimum trajectory length for Theorem 3 and Theorem 4 is $N + n + 1$ and $N + n$, respectively. Our framework needs to have a longer initial trajectory to account for the higher dimensional lifted states, which increases the length of the pre-collected trajectory by $m + 1$. However, the requirement of Theorem 3 does not rely on the specific structure of the affine system is general for all nonlinear systems whose Koopman linear embedding is controllable and has lifted state dimension $n + 1$.

IV. Numerical Experiments

In this section, we validate Theorem 3 by conducting experiments on a nonlinear system that admits an accurate Koopman linear embedding[1]. We compare the prediction and control performance of three linear representations for the nonlinear system which are the accurate data-driven representation (`DD-R`) (11), the approximated affine data-driven representation (`DD-R-A`) (14) and the Koopman linear embedding (`Koopman-R`) (3) given by the EDMD (4).

A. Experiment Setup

We consider the following nonlinear system

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} 0.99x_1(k) \\ 0.9x_2(k) + x_1(k)^2 + x_1(k)^3 + x_1(k)^4 + u(k) \end{bmatrix}, \\ y(k) &= \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}, \end{aligned}$$

where $x = \mathrm{col}(x_1, x_2) \in \mathbb{R}^2$, $y \in \mathbb{R}^2$ and $u \in \mathcal{U} := [-5, 5]$. By choosing the lifted state as $z := \mathrm{col}(x_1, x_2, x_1^2, x_1^3, x_1^4)$, its Koopman linear embedding can be written as

$$\begin{aligned} z(k+1) &= \begin{bmatrix} 0.99 & 0 & 0 & 0 & 0 \\ 0 & 0.9 & 1 & 1 & 1 \\ 0 & 0 & 0.99^2 & 0 & 0 \\ 0 & 0 & 0 & 0.99^3 & 0 \\ 0 & 0 & 0 & 0 & 0.99^4 \end{bmatrix} z(k) + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u(k), \\ y(k) &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} z(k), \end{aligned}$$

which is observable with lag 4.

We here introduce parameters and settings for constructing the data-driven model, the Koopman linear embedding, and the predictive controller. The future trajectory length is chosen as $N = 20$, and the lengths of the initial trajectory are 4 and 2 for the `DD-R` and `DD-R-A`, respectively. Then, we collect a single trajectory with length 52 which is the minimum necessary data length to make H_K a square matrix. For the Koopman linear embedding, we simulate 50 trajectories with 200 time steps. The lifting functions are chosen to be the state of the nonlinear system (1) (*i.e.*, $\phi_1(x) = x_1, \phi_2(x) = x_2$) and 150

[1]All our implementation and experimental scripts can be found at .

thin plate spline radial basis functions whose center x_0 is randomly selected with uniform distribution from $[-1,1]^2$ and has the form $\phi(x) = \|x - x_0\|_2^2 \log(\|x - x_0\|_2)$. The parameters of the controller (12) are set as $R = I_N$ and $Q = I_N \otimes \mathrm{diag}(0,1)$ and the predictive model is replaced by (3) and (14) for `Koopman-R` and `DD-R-A`, respectively. The regularization term $\mathrm{reg} = \lambda_g \|g\|_2 + \lambda_y \|\sigma_y\|_2$ is added to the objective function for the `DD-R-A` with extra slack variables g, σ_y and $\lambda_g = 400, \lambda_y = 2 \times 10^6$ to ensure the feasibility and numerical stability. We refer the interested readers to [10], [22] for detailed construction.

B. Performance Comparison

We first validate `DD-R` is an accurate linear representation for the nonlinear system when the initial trajectory is sufficiently long and compare the prediction performance for different linear models. Given the future input $u_{\mathrm{F}}(k) = 5\sin(\pi k/4)$, the predicted future output from various linear models are shown in Figure 1. It is obvious that the predicted trajectory from `DD-R` is aligned with the actual system trajectory (see red and black dashed curves) while trajectories from `DD-R-A` and the `Koopman-R` deviate from the true trajectory (see orange and blue curves). For the `DD-R-A`, the prediction error comes from two folds: (1) the system is not affine so that the affine constraint $\sum g = 1$ in combining the pre-collected trajectories is inaccurate; (2) the initial trajectory is not long enough which leads to the non-uniqueness of the initial state. For the `Koopman-R`, the selected lifting functions are not guaranteed to form an invariant space under the dynamic of the nonlinear system (*i.e.*, The element of $\Phi(f(x,u))$ is not in the $\mathrm{span}(\phi_1(x), \ldots, \phi_n(x), u)$). Thus, the model obtained from the EDMD contains approximation error which comes from the projection of $\Phi(f(x,u))$ to $\mathrm{span}(\phi_1(x), \ldots, \phi_n(x), u)$.

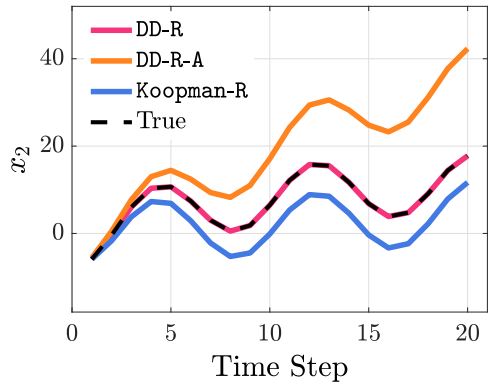

Fig. 1. Prediction performance of different linear models. The black dashed curve denotes the true trajectory of $\boldsymbol{x_2}$ for the given future input. The red, orange and blue curves represent the predicted trajectories of models `DD-R`, `DD-R-A` and `Koopman-R`, respectively.

We then demonstrate the necessity of utilizing a sufficient long initial trajectory. For the same input sequence $u_{\mathrm{F}}(k) = 5\sin(\pi k/4)$, the predicted output trajectory utilizing the same setting of `DD-R` with different lengths of initial trajectories are shown in Figure 2. We note that, for $T_{\mathrm{ini}} < 4$ (*i.e.*, smaller than the lag of the system), there exists infinite many predicted trajectories that satisfy (**??**) but not aligned with the true trajectory (see purple and brown curves). However, increasing the length of the initial trajectory limits the freedom of the deviation. Precisely, in this case, potential initial states of the system form an affine space, and the longer initial trajectory can decrease its dimension. On the other perspective, the prediction error is caused by the rank difference between matrices $\mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}, U_{\mathrm{F}}, Y_{\mathrm{F}})$ and $\mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}, U_{\mathrm{F}})$ which will decrease as the increasing of the initial trajectory length due to the property of the LTI system. This finding also emphasizes that increasing the length of the initial trajectory (*i.e.*, the depth of the trajectory library) is important for ensuring the model accuracy, which is not fully explored in the literature.

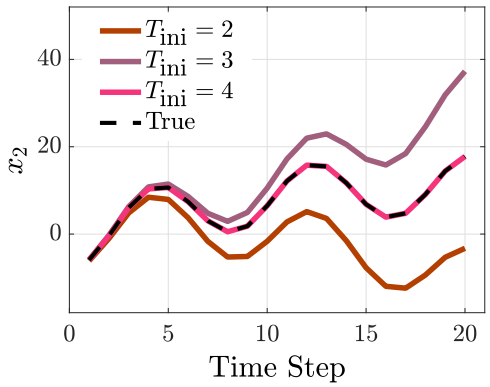

Fig. 2. Prediction performance of model `DD-R` with different initial trajectory lengths. The black dashed curve represents the true trajectory of $\boldsymbol{x_2}$ with fixed future input. The brown, purple and red curves correspond to choose the initial trajectory length as $\mathbf{2}$, $\mathbf{3}$ and $\mathbf{4}$, repsectively.

We finally present the control performance of the predictive controller utilizing different linear representations. We require the state x_2 follow a sinusoidal wave (*i.e.*, $y_{\mathrm{r}}(k) = \mathrm{col}(0, 0.5\sin(\pi k/30))$) and the controller is activated at time step 5. The tracking performances for using different prediction models are shown in Figure 3. We can observe the controller with `DD-R` can track the reference trajectory perfectly (see red and black dashed curves). Using `Koopman-R` can also track the reference trajectory closely (see blue curve) after serval periods while applying `DD-R-A` has a relatively large offset (see orange curve). Although the optimal control input is recomputed at each time step, the inaccurate prediction of `Koopman-R` and `DD-R-A` (see Figure 1) leads to the tracking error. We also note that the Koopman linear embedding needs much more data to achieve good control performance compared with `DeePC` (around 200 times).

V. Conclusion

In this letter, we consider nonlinear systems that admit an accurate Koopman linear embedding and present it leads to an equivalent accurate data-driven representation. We extend Willems' fundamental lemma for general LTI systems and further show the adapted data-driven representation from it is accurate for the nonlinear system. We analyze the data requirement for constructing the accurate data-driven representation based on the Koopman linear embedding of the nonlinear

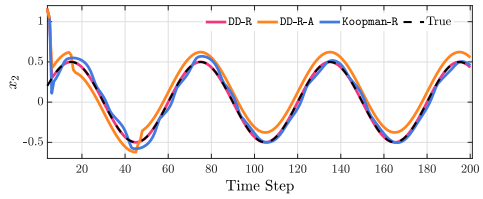

Fig. 3. Control performance of predictive controllers using different linear models. The black dashed curve denotes the reference trajectory of $\boldsymbol{x_2}$. The red, orange and blue curves denote track performances for controllers using models `DD-R`, `DD-R-A` and `Koopman-R`, respectively.

system. Numerical experiments have validated our theoretical findings, which show that the data-driven representation is accurate and has superior control performance. Interesting future directions include exploring the relationship between the data-driven representation and the approximated Koopman linear embedding and analyzing the effect of updating the trajectory library for the data-driven representation as a local Koopman linear embedding.

APPENDIX

A. Proof 1

Proof: We first derive the general form for all potential initial states x_0 for the given leading L_1 data sequence $\mathrm{col}(u_{1:L_1}, y_{1:L_1})$. We then present output trajectories which starts from different potential initial states x_0 with fixed input u are the same.

For the given leading L_1 data sequence, the output trajectory can be decomposed as follows:

$$y_{1:L_1} = \mathcal{T}_{L_1} u_{1:L_1} + \mathcal{O}_{L_1} x_0,$$

where we have

$$\mathcal{T}_{L_1} = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{L_1-2}B & CA^{L_1-3}B & \cdots & D \end{bmatrix}, \mathcal{O}_{L_1} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{L_1-1} \end{bmatrix}.$$

Let $v = y_{1:L_1} - \mathcal{T}_{L_1} u_{1:L_1}$, without loss of generality, the initial state can be represented as

$$x_0 = x_p + \hat{x}$$

where $x_p = \mathcal{O}_{L_1}^{\dagger} v$ and $\hat{x} \in \mathrm{null}(\mathcal{O}_{L_1})$.

For the fixed input u, the trajectory starts from the potential initial state can be derived as

$$y_{1:L} = \mathcal{T}_L u_{1:L} + \mathcal{O}_L x_0 = \mathcal{T}_L u_{1;L} + \mathcal{O}_L x_p + \mathcal{O}_L \hat{x}.$$

Since $\hat{x} \in \mathrm{null}(\mathcal{O}_{L_1})$, from the Cayley–Hamilton theorem, we have $\hat{x} \in \mathrm{null}(\mathcal{O}_L)$. That leads to

$$y_{\mathrm{d},1:L} = \mathcal{T}_L u_{\mathrm{d},1:L} + \mathcal{O}_L x_p,$$

which completes the proof. ■

B. Proof 2

Proof: Since all input-output trajectories of (1) are trajectories of (3), it is standard to derive

$$\begin{bmatrix} u_{\mathrm{d}}^i \\ y_{\mathrm{d}}^i \end{bmatrix} = \begin{bmatrix} I_{mL} & \mathbb{0}_{mL\times n_z} \\ \mathcal{T}_L & \mathcal{O}_L \end{bmatrix} \begin{bmatrix} u_{\mathrm{d}}^i \\ z_0^i \end{bmatrix}, \qquad i = 1, \ldots, l \tag{16}$$

where $z_0^i = \Phi(x_0^i)$ and

$$\mathcal{T}_L = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{L-2}B & CA^{L-3}B & \cdots & D \end{bmatrix}, \mathcal{O}_L = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{bmatrix}.$$

Thus, the trajectory library can be represented as

$$H_{\mathrm{d}} = \begin{bmatrix} I_{mL} & \mathbb{0}_{mL\times n_z} \\ \mathcal{T}_L & \mathcal{O}_L \end{bmatrix} H_K = \begin{bmatrix} I_{mL} & \mathbb{0}_{mL\times n_z} \\ \mathcal{T}_L & \mathcal{O}_L \end{bmatrix} \begin{bmatrix} U_{\mathrm{d}} \\ Z_0 \end{bmatrix}, \tag{17}$$

where we have $U_{\mathrm{d}} = \begin{bmatrix} u_{\mathrm{d}}^1 & u_{\mathrm{d}}^2 & \ldots & u_{\mathrm{d}}^l \end{bmatrix}$ and $Z_0 = \begin{bmatrix} \Phi(x_0^1), \ldots, \Phi(x_0^l) \end{bmatrix}$.

First, it is clear that for any $g \in \mathbb{R}^l$, we have $H_{\mathrm{d}} g \in \mathcal{B}_2|_L$ since each column of H_{d} belongs to $\mathcal{B}_2|_L$ and $\mathcal{B}_2|_L$ is a linear subspace in $\mathbb{R}^{(m+p)L}$.

Second, let $\mathrm{col}(u, y) \in \mathcal{B}_2|_L$ be an arbitrary length-L trajectory of (3). We can write it as

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} I_{Lm} & \mathbf{0}_{Lm\times n_z} \\ \mathcal{T}_L & \mathcal{O}_L \end{bmatrix} \begin{bmatrix} u \\ z_0 \end{bmatrix}, \tag{18}$$

with some initial state $z_0 \in \mathbb{R}^{n_z}$. Since $\mathrm{col}(U_{\mathrm{d}}, Z_0)$ has full row rank of $mL + n_z$ by assumption, its column rank is also $mL + n_z$ and thus there exists a $g \in \mathbb{R}^l$ such that $\mathrm{col}(U_{\mathrm{d}}, Z_0) g = \mathrm{col}(u, z_0)$. Substituting this into (17) and (18), we derive $H_{\mathrm{d}} g = \mathrm{col}(u, y)$. This completes the proof. ■

REFERENCES

[1] M. Korda and I. Mezić, "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control," *Automatica*, vol. 93, pp. 149–160, 2018.

[2] B. O. Koopman, "Hamiltonian systems and transformation in hilbert space," *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315–318, 1931.

[3] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[4] D. A. Haggerty, M. J. Banks, E. Kamenar, A. B. Cao, P. C. Curtis, I. Mezić, and E. W. Hawkes, "Control of soft robots with inertial dynamics," *Science robotics*, vol. 8, no. 81, p. eadd6864, 2023.

[5] E. Elokda, J. Coulson, P. N. Beuchat, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for quadcopters," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 18, pp. 8916–8936, 2021.

[6] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz, "Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control," *PloS one*, vol. 11, no. 2, p. e0150171, 2016.

[7] J. Berberich and F. Allgöwer, "A trajectory-based framework for data-driven system analysis and control," in *2020 European Control Conference (ECC)*. IEEE, 2020, pp. 1365–1370.

[8] M. Guo, C. De Persis, and P. Tesi, "Data-driven stabilization of nonlinear polynomial systems with noisy data," *IEEE Transactions on Automatic Control*, vol. 67, no. 8, pp. 4210–4217, 2021.

[9] J. Berberich, A. Koch, C. W. Scherer, and F. Allgöwer, "Robust data-driven state-feedback design," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1532–1538.

[10] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the deepc," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 307–312.

[11] N. Lawrence, P. Loewen, S. Wang, M. Forbes, and B. Gopaluni, "Deep hankel matrices with random elements," in *6th Annual Learning for Dynamics & Control Conference*. PMLR, 2024, pp. 1579–1591.

[12] K. Zhang, Y. Zheng, C. Shang, and Z. Li, "Dimension reduction for efficient data-enabled predictive control," *IEEE Control Systems Letters*, 2023.

[13] 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*, vol. 25, pp. 1307–1346, 2015.

[14] M. Haseli and J. Cortés, "Learning koopman eigenfunctions and invariant subspaces from data: Symmetric subspace decomposition," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3442–3457, 2021.
[15] ——, "Generalizing dynamic mode decomposition: Balancing accuracy and expressiveness in koopman approximations," *Automatica*, vol. 153, p. 111001, 2023.
[16] X. Zhang, W. Pan, R. Scattolini, S. Yu, and X. Xu, "Robust tube-based model predictive control with koopman operators–extended version," *arXiv preprint arXiv:2108.13011*, 2021.
[17] M. Lazar, "Basis functions nonlinear data-enabled predictive control: Consistent and computationally efficient formulations," *arXiv preprint arXiv:2311.05360*, 2023.
[18] H. J. Van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 602–607, 2020.
[19] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.
[20] X. Shang, J. Wang, and Y. Zheng, "Decentralized robust data-driven predictive control for smoothing mixed traffic flow," *arXiv preprint arXiv:2401.15826*, 2024.
[21] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler, "Decentralized data-enabled predictive control for power system oscillation damping," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 3, pp. 1065–1077, 2021.
[22] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Linear tracking mpc for nonlinear systems—part ii: The data-driven case," *IEEE Transactions on Automatic Control*, vol. 67, no. 9, pp. 4406–4421, 2022.
[23] I. Markovsky, "Data-driven simulation of generalized bilinear systems via linear time-invariant embedding," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 1101–1106, 2022.
[24] Y. Lian, R. Wang, and C. N. Jones, "Koopman based data-driven predictive control," *arXiv preprint arXiv:2102.05122*, 2021.
[25] A. Martinelli, M. Gargiani, M. Draskovic, and J. Lygeros, "Data-driven optimal control of affine systems: A linear programming perspective," *IEEE Control Systems Letters*, vol. 6, pp. 3092–3097, 2022.