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

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

***Abstract*— Koopman operator theory and Willems' fundamental lemma both can provide (approximated) data-driven linear representation for nonlinear systems. However, choosing lifting functions for the Koopman operator is challenging, and the quality of the data-driven model from Willems' fundamental lemma has no guarantee for general nonlinear systems. In this paper, we extend Willems' fundamental lemma for a class of nonlinear systems that admit a *Koopman linear embedding*. We first characterize the relationship between the trajectory space of a nonlinear system and that of its Koopman linear embedding. We then prove that the trajectory space of Koopman linear embedding can be formed by a linear combination of rich-enough trajectories from the nonlinear system. Combining these two results leads to a data-driven representation of the nonlinear system, which bypasses the need for the lifting functions and thus eliminates the associated bias errors. Our results illustrate that both the *width* (more trajectories) and *depth* (longer trajectories) of the trajectory library are important to ensure the accuracy of the data-driven model.**

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

## I. Introduction

Designing controllers for nonlinear systems with approximated linear representations has gained increasing interest. Linear approximations enable the utilization of linear system tools and facilitate computationally efficient model predictive control schemes. Both Koopman operator theory [1] and Willems' fundamental lemma [2] can be applied to construct (approximated) linear representations of nonlinear systems from input and output data, which have shown promising performance in many practical applications [3]–[6].

Koopman operator theory is originally developed for autonomous systems with no input [1]. There are different Koopman operator schemes to handle controlled nonlinear systems, *e.g.*, taking the control sequence as an extended state [7] or considering the control sequence as extra parameters [8]. One key step is to *lift* the state space into a higher-dimensional space, in which the lifted state evolves (approximately) in a linear way. This idea leads to a rigorous framework of Koopman operator theory for autonomous systems [9], and extensions for controlled systems are under extensive development [10]. With (approximated) linear representations using the Koopman operator available, many control techniques have been applied, such as linear optimal control and model predictive control [7], [10]. In all these methods, the accuracy of the Koopman-based approximations depends critically on the lifting functions, and choosing the right set is challenging [11].

Willems' fundamental lemma for linear time-invariant (LTI) systems shows that a rich-enough trajectory library is sufficient to produce a direct data-driven representation of the system evolution [2]. Its wide range of applicability has motivated the search for extensions, including special classes of nonlinear systems, such as Hamerstein and Wiener systems [12], bilinear systems [13], [14], and certain polynomial systems [15]. This data-driven representation can be utilized for linear controller design [16] and also model predictive control [17]. When dealing with non-deterministic or nonlinear systems, it becomes necessary to include suitable regularization terms in predictive control to ensure its performance and increase the size of the trajectory library to construct a good data-driven representation [18]–[20]. Although the benefits of increasing the width of the trajectory library (*i.e.*, collecting more trajectories) are well-recognized in the literature, the importance of enlarging its depth (*i.e.*, extending the trajectory length) is less discussed.

In this paper, we aim to develop an extended Willems' fundamental lemma for nonlinear systems that admit a lifted linear representation under the Koopman operator, which we call *Koopman linear embedding* (see Definition 1). Unlike previous studies [7], [12]–[14], our direct data representation requires no prior knowledge of the lifting functions in the Koopman linear embedding or the nonlinearity of the system. One key idea in our approach is to establish an exact relationship between the trajectory spaces of the nonlinear system and its associated Koopman linear embedding. In general, the latter contains the former. We provide a necessary and sufficient condition for a trajectory of the Koopman linear embedding to be a trajectory of the original nonlinear system (Theorem 1). Motivated by [21, Def. 1], we introduce a new notion of *lifted excitation* for nonlinear systems which accounts for the lifted state in Koopman linear embedding. We show the behavior of Koopman linear embedding can be fully captured by a linear combination of rich enough trajectories from the nonlinear system (Theorem 2). We finally establish a data-driven representation adapted from Willems' fundamental lemma for nonlinear systems with a Koopman linear embedding (Theorem 3). Thus, we can directly utilize the simple-to-build data-driven representation and bypass the need to choose lifting functions.

This work is supported by NSF CMMI 2320697, NSF CAREER 2340713, ONR N00014-23-1-2353, and an Early Career Faculty Development Award from the School of Engineering, UC San Diego. The authors thank Masih Haseli for discussions on Koopman operator theory.

[1]X. Shang and Y. Zheng are with the Department of Electrical and Computer Engineering, UC San Diego; {x3shang, zhengy}@ucsd.edu.

[2]J. Cortés is with the Department of Mechanical and Aerospace Engineering, UC San Diego; cortes@ucsd.edu.

Our data-driven representation can be directly utilized in predictive control. Our approach also illustrates the importance of the *width* and *depth* of the trajectory library, which depends on the "hidden" dimension of the Koopman linear embedding. Both collecting more trajectories (increasing the width) and utilizing longer initial trajectories (increasing the depth) are critical for the data-driven representation of nonlinear systems.

The remainder of this paper is structured as follows. Section II reviews Koopman linear embedding and Willems' fundamental lemma. Section III shows that a Koopman linear embedding of the nonlinear system leads to a direct data-driven representation. Section IV validates our theoretical findings via numerical simulations. We conclude the paper with Section V. Auxiliary proofs/discussions are provided in our report [22].

*Notation:* Given a series of vectors $a_1, \dots, a_n$ and matrices $A_1, \dots, A_n$ with the same column dimension, we denote $\mathrm{col}(a_1, \dots, a_n) := \left[a_1^\mathsf{T}, \dots, a_n^\mathsf{T}\right]^\mathsf{T}$ and $\mathrm{col}(A_1, \dots, A_n) := \left[A_1^\mathsf{T}, \dots, A_n^\mathsf{T}\right]^\mathsf{T}$. We denote the quadratic form $a^\mathsf{T} X a$ as $\|a\|_X^2$ and $\mathrm{diag}(b_1, ..., b_n)$ as a diagonal matrix with $b_1, ..., b_n$ at its diagonal entries. We use $\|A\|_F$ to represent the Frobenius norm of the matrix $A$. Collecting a length-$T$ data sequence $v = \mathrm{col}(v_0, \dots, v_{T-1})$, we represent $v_{p:q} := \mathrm{col}(v_p, \dots, v_q)$ where $p, q \in \mathbb{Z}$ and $T > q \geq p \geq 0$.

## II. Preliminaries and Problem Statement

### A. Koopman linear models for nonlinear systems

Consider a discrete-time nonlinear system

$$x_{k+1} = f(x_k, u_k), \quad y_k = g(x_k, u_k), \tag{1}$$

where $x_k \in \mathbb{R}^n, u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ are the state, input, and output of the system at time $k$, respectively. One key idea of Koopman operator is to lift the state $x_k$ of the original nonlinear system to a higher-dimensional space via a set of lifting functions (often referred to as observables) [9], where the evolution of these observables becomes (approximately) linear.

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

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

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

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

$$z_{k+1} = Az_k + Bu_k, \quad y_k = Cz_k + Du_k, \tag{3}$$

with matrices $A, B, C$ and $D$ of 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. Even when 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 [7] for details.

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) [10]. We organize the measured input-state-output data sequence of (1) as

$$\mathrm{X} = \left[x_0, \dots, x_{n_\mathrm{d}-2}\right], \quad \mathrm{X}^+ = \left[x_1, \dots, x_{n_\mathrm{d}-1}\right],$$
$$\mathrm{U} = \left[u_0, \dots, u_{n_\mathrm{d}-2}\right], \quad Y = \left[y_0, \dots, y_{n_\mathrm{d}-2}\right].$$

With the lifting functions (2), we compute the lifted state as

$$\mathrm{Z} = \left[\Phi(x_0), \dots, \Phi(x_{n_\mathrm{d}-2})\right], \quad \mathrm{Z}^+ = \left[\Phi(x_1), \dots, \Phi(x_{n_\mathrm{d}-1})\right].$$

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

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

It is not necessary to collect the data points in sequence and we can also use data pairs $(x_i, x_i^+, u_i, y_i)$ where $x_i^+ = f(x_i, u_i), y_i = g(x_i, u_i), i = 0, \dots, n_\mathrm{d} - 1$ (see [7] for details).

The choice of observables affects (4) significantly. Even if a Koopman linear embedding exists for (1), we may not know the correct observables (2) for such a Koopman linear embedding. An inexact choice can lead to significant modeling errors [11]. In the literature, common choices for (2) include Gaussian kernel, polyharmonic splines, and thin plate splines [10]. However, none of them can guarantee 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

$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k, \tag{5}$$

where the state, input and output at time $k$ are denoted as $x_k \in \mathbb{R}^{\tilde{n}}$, $u_k \in \mathbb{R}^{\tilde{m}}$ and $y_k \in \mathbb{R}^{\tilde{p}}$, respectively. We consider system (5) from the behavioral (*i.e.*, trajectory) perspective. The key idea is that a linear combination of rich enough offline trajectories of (5) can represent its whole trajectory space.

Let us recall the notion of persistent excitation [2].

*Definition 2 (Persistently exciting):* The length-$T$ sequence $\omega = \mathrm{col}(\omega_0, \dots, \omega_{T-1})$ is persistently exciting (PE) of order $L$ if its Hankel matrix

$$\mathcal{H}_L(\omega) = \begin{bmatrix} \omega_0 & \omega_1 & \cdots & \omega_{T-L} \\ \omega_1 & \omega_2 & \cdots & \omega_{T-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{L-1} & \omega_L & \cdots & \omega_{T-1} \end{bmatrix}$$

has full row rank.

With the pre-collected input-state-output data in sequence, *i.e.*, $u_\mathrm{d} = \mathrm{col}(u_0, \dots, u_{n_\mathrm{d}-1}), x_\mathrm{d} = \mathrm{col}(x_0, \dots, x_{n_\mathrm{d}-1})$ and $y_\mathrm{d} = \mathrm{col}(y_0, \dots, y_{n_\mathrm{d}-1})$, the following Willems' fundamental lemma is adapted from [23, Theorem 1].

*Lemma 1 (Willems' fundamental lemma):* Consider the LTI system (5). Assume that the Hankel matrix formed by its pre-collected trajectory $H_0 := \mathrm{col}(\mathcal{H}_1(x_{\mathrm{d},0:n_\mathrm{d}-L}), \mathcal{H}_L(u_\mathrm{d}))$ has full row rank. Then, a length-$L$ input-output data sequence $\mathrm{col}(u, y) \in \mathbb{R}^{(\tilde{m}+\tilde{p})L}$ is a valid trajectory of (5) if and only if there exists $g \in \mathbb{R}^{n_\mathrm{d}-L+1}$ 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 persistent excitation of order $L + \tilde{n}$ for the input sequence $u_\mathrm{d}$ is sufficient to

guarantee the full rank of $H_0$ [2]. Utilizing Lemma 1, we can build a data-driven representation for system (5). We use $u_{\text{ini}} = \text{col}(u_{k-T_{\text{ini}}}, \dots, u_{k-1})$ and $u_{\text{F}} = \text{col}(u_k, \dots, u_{k+N-1})$ to represent the most recent past input trajectory of length-$T_{\text{ini}}$ and the future input trajectory of length-$N$ and $L = T_{\text{ini}} + N$ (similarly for $y_{\text{ini}}, y_{\text{F}}$). Let us partition the Hankel matrix by its first $T_{\text{ini}}$ rows (*i.e.*, $U_{\text{P}}, Y_{\text{P}}$) and the last $N$ rows (*i.e.*, $U_{\text{F}}, Y_{\text{F}}$) as

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

From Lemma 1, $\text{col}(u_{\text{ini}}, y_{\text{ini}}, u_{\text{F}}, y_{\text{F}})$ is a valid trajectory of (5) if and only if there exists $g \in \mathbb{R}^{n_{\text{d}}-T_{\text{ini}}-N+1}$ 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 $T_{\text{ini}}$ is no smaller than its observability index, then $y_{\text{F}}$ in (6) is unique given an initial trajectory $(u_{\text{ini}}, y_{\text{ini}})$ and any future input $u_{\text{F}}$ [17]. 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 [17] with many successful applications [4]–[6].

### 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. 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 (3) 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 (3) and that of the original nonlinear system (1). A key insight of this characterization is to guarantee the uniqueness of the subsequent output given the leading input-output trajectory and the subsequent input, and observability is not needed for ensuring the uniqueness as long as the length of the initial trajectory is large enough, *i.e.*, the Hankel matrix has a sufficient *depth*. 2) Motivated by lack of controllability, we introduce a new notion of lifted excitation for the offline data collection, which has a similar flavor to [21, Def. 1] that focuses on a special case of affine systems. With these two technical tools, we 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 [7], without the need of identifying the linear model (3). This bypasses the challenging problem of selecting the lifting functions, with the added remarkable benefit of eliminating the associated bias errors.

## III. From Koopman Linear Embeddings to Data-driven Representations

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 also discuss a special case of affine systems considered in [21].

### A. Two behavior spaces

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

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

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

Note that $\mathcal{B}_1|_L$ is a nonlinear set while $\mathcal{B}_2|_L$ is a linear subspace in $\mathbb{R}^{(m+p)L}$. Intuitively, 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) and assume 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 $L > n_z$. Then, $\text{col}(u, y) \in \mathcal{B}_1|_L$ if and only if its leading sequence of length $n_z$ (*i.e.*, $\text{col}(u_{0:n_z-1}, y_{0:n_z-1})$) 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 $\text{col}(u_{0:n_z-1}, y_{0:n_z-1})$. We need a technical lemma to prove the second statement in Theorem 1.

*Lemma 2:* Consider an LTI system (5). Fix an initial trajectory $\text{col}(u_{0:L_1-1}, y_{0:L_1-1}) \in \mathbb{R}^{(\tilde{m}+\tilde{p})L_1}$ of length-$L_1$, where $L > L_1 \geq \tilde{n}$. Given any subsequent input $u_{L_1:L-1} \in \mathbb{R}^{\tilde{m}(L-L_1)}$ (future input), the subsequent output $y_{L_1:L-1} \in \mathbb{R}^{\tilde{p}(L-L_1)}$ (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 its proof is provided in our technical report [22].

**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 $\text{col}(u_{0:n_z-1}, y_{0:n_z-1})$ is part of $\text{col}(u, y)$. We here prove "if" part. Suppose $\text{col}(u_{0:n_z-1}, y_{0:n_z-1}) \in \mathcal{B}_1|_{n_z}$ and we let $\tilde{y}_{n_z:L-1}$ be the corresponding outputs from the nonlinear system (1) for the rest of inputs $u_{n_z:L-1}$, *i.e.*,

$$\text{col}(u, y_{0:n_z-1}, \tilde{y}_{n_z:L-1}) \in \mathcal{B}_1|_L\ .$$

Then, it is clear that $\text{col}(u, y_{0:n_z-1}, \tilde{y}_{n_z:L-1}) \in \mathcal{B}_2|_L$ utilizing the first statement. From Lemma 2, the outputs of the linear system (3) are uniquely determined by $u_{n_z:L-1}$ when $\text{col}(u_{0:n_z-1}, y_{0:n_z-1}) \in \mathcal{B}_2|_{n_z}$ are given. Thus, we must have

$$\tilde{y}_{n_z:L-1} = y_{n_z:L-1},$$

indicating the whole trajectory satisfies $\text{col}(u, y) \in \mathcal{B}_1|_L$. ∎

We might be tempted to estimate an initial state $x_0$ or $z_0$ from $\text{col}(u_{0:n_z-1}, y_{0:n_z-1}) \in \mathcal{B}_1|_{n_z}$. However, since we do not assume observability of the Koopman linear embedding (3), the initial state $z_0$ cannot be uniquely determined from $\text{col}(u_{0:n_z-1}, y_{0:n_z-1})$. As confirmed in Lemma 2, the unobservable part of the initial state does not affect the uniqueness of the input-output trajectory. If the Koopman linear embedding is observable with observability index $l_{\text{o}}$, the required length of the leading sequence can be decreased to $l_{\text{o}}$. In this

case, the initial state $z_0$ can be estimated, and the uniqueness of the input-output trajectory is also guaranteed. The initial trajectory can thus be shorter as $T_{\text{ini}} \geq l_{\text{o}}$ in Theorem 3.

### 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.*, $\text{col}(u_{\text{d}}^i, y_{\text{d}}^i) \in \mathcal{B}_1|_L, i = 1, \ldots, l$) that are rich enough to represent $\mathcal{B}_2|_L$. For this, we propose the following definition of lifted excitation.

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

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

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

This notion of lifted excitation generalizes [21, Def. 1], that focuses only on affine systems. Our notion is suitable for any nonlinear system with a Koopman linear embedding. If (3) is controllable, we can design the inputs to ensure the lifted excitation. Given a single trajectory $u_{\text{d}}, y_{\text{d}}$ with length-$n_{\text{d}}$, we obtain $l$ trajectories with length-$L$ using column vectors of $\text{col}(\mathcal{H}_L(u_{\text{d}}), \mathcal{H}_L(y_{\text{d}}))$ where $l = n_{\text{d}} - L + 1$. If $u_{\text{d}}$ is PE of order $L+n_{\text{z}}$, the $l$ trajectories above satisfy lifted excitation. If we collect multiple trajectories $u_{\text{m}}^i, y_{\text{m}}^i, i = 1, \ldots, q$ (see [23]), we require $u_{\text{m}}^1, \ldots, u_{\text{m}}^q$ to be collectively PE of order $L + n_{\text{z}}$, *i.e.*, $\text{rank}\big(\begin{bmatrix}\mathcal{H}_{L+n_{\text{z}}}(u_{\text{m}}^1), \ldots, \mathcal{H}_{L+n_{\text{z}}}(u_{\text{m}}^q)\end{bmatrix}\big) = m(L+n_{\text{z}})$. The column vectors of $\text{col}(\mathcal{H}_L(u_{\text{m}}^i), \mathcal{H}_L(y_{\text{m}}^i))$, $i = 1, \ldots, q$ satisfies the lifted excitation in Definition 3.

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

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

Due to the page limit, we put the proof to our report [22]. Theorem 2 allows us to use the trajectories from the nonlinear system (1) that provide lifted excitation of order $L$ (always exist; see [22, App. B.1]) to represent any length-$L$ trajectory of the Koopman linear embedding (3). Given trajectories of the nonlinear system, it is still an open problem to determine whether a Koopman linear embedding exists, which is possible when infinite rich enough data is provided (see details in [22, App. B.2]). Combining Theorems 1 and 2 leads to a direct data-driven representation for a nonlinear system (1) with Koopman linear embedding (3), as we describe in our next result. Given a trajectory library $H_{\text{d}} = \text{col}(U_{\text{d}}, Y_{\text{d}})$ in (9), where each column is a trajectory of length $L = T_{\text{ini}} + N$ from the nonlinear system (1), we partition matrices $U_{\text{d}}$ and $Y_{\text{d}}$ as

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

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

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

$$u_{\text{ini}} = \text{col}(u_{k-T_{\text{ini}}}, \ldots, u_{k-1}), \quad y_{\text{ini}} = \text{col}(y_{k-T_{\text{ini}}}, \ldots, y_{k-1}).$$

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

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

This result is a combination of Theorems 1 and 2. Due to the page limit, we provide some details in our report [22]. 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 lifted excitation for nonlinear systems in Definition 3 that enables Theorem 2, and 2) a sufficiently long initial trajectory $\text{col}(u_{\text{ini}}, y_{\text{ini}})$ from the nonlinear system that ensures Theorem 1. We remark that if $n_{\text{z}}$ is big, a large amount of data is needed to satisfy the conditions in Theorem 3 (which is expected since the linear behavior becomes richer).

We here remark that Theorem 3 illustrates the importance of increasing the *width* and *depth* of the trajectory library (9) and (10). While the benefits of increasing width are well-recognized in the literature, the importance of enlarging the depth has been overlooked. Collecting more trajectories to increase the width of (9) contributes to the lifted excitation condition (see Theorem 2). On the other hand, fixing the prediction horizon $N$, a sufficient *depth* ensures the initial trajectory is long enough in (10), which guarantees that the trajectory in the space of the Koopman linear embedding is also a valid trajectory for the nonlinear system (see Theorem 1). Furthermore, Theorem 3 shows the required width and depth of the trajectory library depend on the "hidden" dimension of the Koopman linear embedding of the nonlinear system. Note that estimating the Koopman order $n_z$ is non-trivial, even if the Koopman linear embedding exists (see discussions in [22, App. B.2 and B.3]).

According to Theorem 3, the data-driven representation (11) is equivalent to the Koopman linear embedding. This can be directly integrated with predictive control at each time $k$ as

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

where $R \succ 0, Q \succeq 0$, $y_{\text{r}}$ denotes the reference output trajectory, and $u_{\text{F}} \in \mathcal{U}$ is the input constraint. For nonlinear systems with Koopman linear embedding, there is no need to use any lifting functions, which are instead required by most existing Koopman-based model predictive control approaches [7], [10].

*Remark 1:* The literature on extending Willems' fundamental lemma to nonlinear cases [12]–[14] requires system dynamics, with additional constraints for their nonlinear structures. When the nonlinearities exist in input or output (*e.g.*, Hamerstein systems, Wiener systems), a change of variables

is needed [12]. Our data-driven representation (11) has no additional constraints and requires no knowledge of lifting functions. We only need a lifted excitation condition and a sufficiently long initial trajectory. The work [24] integrates Willems' fundamental lemma with the learning of lifting functions and [25] forms the trajectory library with lifted states, which also requires learning lifting functions. In contrast, we show that learning these lifting functions is redundant for nonlinear systems with Koopman linear embedding.

### C. A special case: Affine systems

Consider an affine system of the form

$$x_{k+1} = A_2 x_k + B_2 u_k + e, \quad y_k = C_2 x_k + D_2 u_k + r, \quad (13)$$

where $x_k \in \mathbb{R}^{\bar{n}}$, $u_k \in \mathbb{R}^{\bar{m}}, y_k \in \mathbb{R}^{\bar{p}}$ are the state, input and output at time $k$, respectively and $e \in \mathbb{R}^{\bar{n}}, r \in \mathbb{R}^{\bar{p}}$ are two constant vectors. The result in [21, Theorem 1] presents a data representation for (13), which we reproduce below.

*Theorem 4 ( [21, Theorem 1]):* Given the pre-collected trajectories $u_{\rm d}, x_{\rm d}, y_{\rm d}$ of (13), suppose

$$H_A := \text{col}(\mathcal{H}_L(u_{\rm d}), \mathcal{H}_1(x_{{\rm d},0:n_{\rm d}-L}), \mathbb{1}) \quad (14)$$

has full row rank. Then, a length-$L$ input-output data sequence $\text{col}(u_{\rm ini}, y_{\rm ini}, u_{\rm F}, y_{\rm F}) \in \mathbb{R}^{(\bar{m}+\bar{p})L}$ is a valid trajectory of (13) if and only if there exists $g \in \mathbb{R}^{n_{\rm d}-L+1}$ such that

$$\text{col}(U_{\rm P}, Y_{\rm P}, U_{\rm F}, Y_{\rm F})g = \text{col}(u_{\rm ini}, y_{\rm ini}, u_{\rm F}, y_{\rm F}), \ \sum g = 1. \quad (15)$$

A unique feature in Theorem 4 is that the coefficient $g$ should have an affine constraint, since (13) is affine. Here, we show that any affine system (13) has an exact Koopman linear embedding, and thus our main Theorem 3 naturally applies. Choose a vector of lifting functions $z(x) = [\phi_1(x), \dots, \phi_n(x), 1]^\mathsf{T}$ where $\phi_i(x) : \mathbb{R}^{\bar{n}} \to \mathbb{R}$ is the $i$-th element of the state, *i.e.*, $\phi_i(x) = x_i$. Then, we have a Koopman linear embedding for the affine system (13) as

$$z_{k+1} = \begin{bmatrix} A_2 & e \\ \mathbb{0} & 1 \end{bmatrix} z_k + \begin{bmatrix} B_2 \\ \mathbb{0} \end{bmatrix} u_k, \ \ y_k = \begin{bmatrix} C_2 & r \end{bmatrix} z_k + D_2 u_k. \quad (16)$$

Consequently, with the lifted excitation and an initial trajectory of length $\bar{n}+1$ in Theorem 3, the data-representation (11) is necessary and sufficient for the behavior of (13). In this case, we need a slightly longer initial trajectory, but no affine constraint on $g$ is needed. We note that the data matrices $H_A$ in (14) and $H_K$ in (8) become the same, thus the lifted excitation and the data requirement in Theorems 3 and 4 are identical. Our data-representation in Theorem 3 works for any nonlinear systems with Koopman linear embedding while it is unclear how to design constraints on $g$ to extend Theorem 4.

*Remark 2:* General nonlinear systems may not admit a Koopman linear embedding, even in infinite dimensions [26]. However, Koopman linear models are widely used to approximate nonlinear dynamics and have demonstrated good performance in real applications when the approximation error is small [3]–[5]. In [21], the affine data-driven representation (15) is continuously updated using the most recent trajectory, approximating the local linearization of the nonlinear system. It is an interesting direction to integrate Theorem 3 with the approach in [21] for a broader class of nonlinear systems.

## IV. Numerical Experiments

In this section, we present numerical experiments to compare the prediction and control performance of four linear representations: 1) our proposed Data-Driven Koopman representation (DD-K) (11), 2) the Data-Driven Affine representation (DD-A) (15), 3) the standard Koopman linear approximation (3) from EDMD (4) (EDMD-K) and 4) the Deep Neural Network Koopman representation (DNN-K) in [27].

### A. Experiment Setup

We consider the following nonlinear system

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} 0.99 x_{1,k} \\ 0.9 x_{2,k} + x_{1,k}^2 + x_{1,k}^3 + x_{1,k}^4 + u_k \end{bmatrix},$$

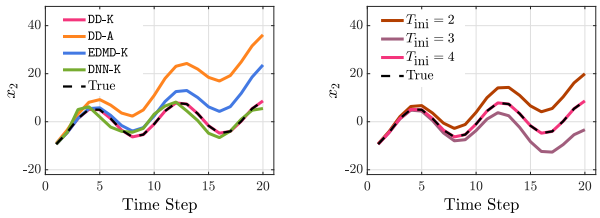
with output $y_k = x_k$, and state $x = \text{col}(x_1, x_2) \in \mathbb{R}^2$ and input $u \in \mathcal{U} := [-5, 5]$. We choose the lifted state as $z := \text{col}(x_1, x_2, x_1^2, x_1^3, x_1^4)$, and its Koopman linear embedding is

$$z_{k+1} = \begin{bmatrix} 0.99 & 0 & 0 & 0 & 0 \\ 0 & 0.9 & 1 & 1 & 1 \\ 0 & 0 & 0.99^2 & 0 & 0 \\ 0 & 0 & 0 & 0.99^3 & 0 \\ 0 & 0 & 0 & 0 & 0.99^4 \end{bmatrix} z_k + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_k,$$

and $y_k = z_k(1:2)$. This linear embedding is not controllable but has an observability index 4. In our experiments, the prediction horizon is $N = 20$, and the lengths of the initial trajectory $\text{col}(u_{\rm ini}, y_{\rm ini})$ are 4 and 2 for DD-K and DD-A (the entire trajectory length $L$ is 24 for DD-K and 22 for DD-A), respectively. Then, we collect a single trajectory of length-52 and obtain 29 length-24 trajectories for DD-K, which is the minimum necessary data length to make $H_K$ in (8) a square matrix. For the EDMD method, we simulate 200 trajectories with 200 time steps. The lifting functions are chosen to be the state of (1) and 300 thin plate spline radial basis functions whose center $x_0$ is randomly selected with uniform distribution from $[-1, 1]^2$ and has the form $\phi(x) = \|x - x_0\|_2^2 \log(\|x - x_0\|_2)$. To learn the lifting functions with DNN, we collect $10^5$ trajectories with 100 steps, and we used 2 hidden layers with 64 units to learn 32 lifting functions (*i.e.*, the DNN's output is 32). The parameters in (12) are set as $R = I_N$ and $Q = I_N \otimes \text{diag}(0, 100)$, and the prediction model is replaced by (3) and (15) for EDMD-K and DD-A, respectively. A regularization term $\text{reg} = \lambda_g \|g\|_2 + \lambda_y \|\sigma_y\|_2$ is added for the DD-A with variables $g, \sigma_y$ and $\lambda_g = 400, \lambda_y = 2 \times 10^5$ to ensure feasibility and numerical stability.

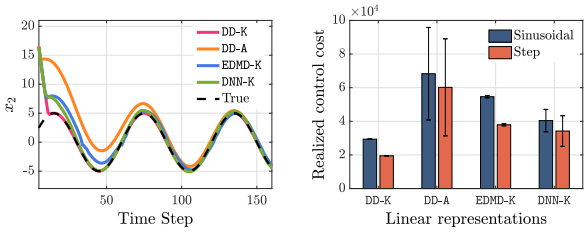
### B. Prediction and Control Performance

We first compare the prediction performance for the four methods, and we also illustrate our DD-K with different initial trajectory lengths. Given the future input $u_{\rm F}(k) = 5\sin(\pi k/4)$, Figure 1(a) displays results for the four methods. The predicted output trajectories of DD-K with different initial trajectory lengths are shown in Figure 1(b). As expected from Theorem 3, the predicted trajectory from DD-K with $T_{\rm ini} = 4$ is the same as the true trajectory (see red and black dashed curves in Figure 1). However, trajectories from DD-A and EDMD-K (see orange and blue curves in Figure 1(a)) and the DD-K with initial trajectory length 2 and 3 (see purple and brown curves in Figure 1(b)) deviate from the true trajectory. The trajectory from DNN-K is relatively close to the true trajectory (see the green curve in Figure 1(a)) but its performance varies significantly for different pre-collected data sets. For DD-A, the affine constraint is inaccurate for this non-affine system. For EDMD-K and DNN-K, their lifting functions are




(a) Comparison of linear models (b) Comparison of different $T_{\text{ini}}$

Fig. 1. Prediction of $\boldsymbol{x_2}$ with a given input $\boldsymbol{u}_{\text{F}}$. In (a) and (b), the red curve is the predicted trajectory of DD−K with $\boldsymbol{T}_{\text{ini}} = \mathbf{4}$. The orange, blue and green curves in (a) are DD−A, EDMD−K and DNN−K, and the brown and purple curves in (b) are DD−K with initial trajectory of $\boldsymbol{T}_{\text{ini}} = \mathbf{2}$ and $\mathbf{3}$.




(a) Tracking performance (b) Realized control cost

Fig. 2. Control performance using different linear representations. (a) Tracking Performance. (b) Realized control costs.

not guaranteed to form an invariant space (cf. [11]), and both EDMD and DNN have approximation errors.

We next compare control performance of predictive controllers utilizing different linear representations. Our goal is to make $x_2$ track two types of reference trajectories: 1) Sinusoidal wave $y_{\text{r},k} = \text{col}(0, 5\sin(\pi k/30))$ and 2) Step signal $y_{\text{r}} = \text{col}(0, 5)$. We consider the realized control cost that is computed as $\|u^*\|_R^2 + \|y^* - y_r\|_Q^2$ where $u^*$ is the computed control input and $y^*$ is the actual trajectory after applying $u^*$. The results are shown in Figure 2, and the realized control cost is averaged over 100 data sets since the performance of these models is related to the pre-collected data. From Figure 2(a), we can observe the controller with DD−K can track the reference trajectory perfectly (see red and black dashed curves). EDMD−K and DNN−K can also track the reference trajectory closely and DNN−K converges faster to the reference trajectory (see green and blue curves) while applying DD−A has a longer transition phase. The realized control cost in Figure 2(b) further demonstrates DD−A $>$ EDMD−K $>$ DNN−K $>$ DD−K for both sinusoidal and step signals. We remark that models obtained from DNN have large variances and we eliminated the problematic models whose realized control cost is larger than $10^5$. The inaccurate prediction of EDMD−K, DNN−K and DD−A (see Figure 1(a)) leads to the tracking error. Finally, we note that both DNN−K and EDMD−K require an order of magnitude more data while failing to achieve the same performance as our method DD−K, and that the performance of DNN−K is highly dependent on the pre-collected data set.

## V. Conclusions

We have developed an extended Willems' fundamental lemma for nonlinear systems that admit a Koopman linear embedding. Our results eliminate the non-trivial process of selecting lifting functions and illustrate that the required size of the trajectory library is related to the dimension of the Koopman linear embedding. Future directions include developing data-driven models for nonlinear systems with approximated Koopman linear embeddings, analyzing the effect of adaptively updating the trajectory library as in [21] and considering coupling nonlinearities between state and input [8], [28].

## References


[1] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *PNAS*, vol. 17, no. 5, pp. 315–318, 1931.

[2] J. C. Willems *et al.*, "A note on persistency of excitation," *Syst. Control Lett.*, vol. 54, no. 4, pp. 325–329, 2005.

[3] D. A. Haggerty *et al.*, "Control of soft robots with inertial dynamics," *Sci. Rob.*, vol. 8, no. 81, p. eadd6864, 2023.

[4] E. Elokda *et al.*, "Data-enabled predictive control for quadcopters," *Int. J. Robust Nonlinear Control*, vol. 31, no. 18, pp. 8916–8936, 2021.

[5] J. Wang *et al.*, "DeeP-LCC: Data-enabled predictive leading cruise control in mixed traffic flow," *IEEE Trans. Control Syst. Technol.*, vol. 31, no. 6, pp. 2760–2776, 2023.

[6] X. Shang *et al.*, "Decentralized robust data-driven predictive control for smoothing mixed traffic flow," *arXiv preprint arXiv:2401.15826*, 2024.

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

[8] M. Haseli and J. Cortés, "Modeling nonlinear control systems via Koopman control family: Universal forms and subspace invariance proximity," *arXiv preprint arXiv:2307.15368*, 2023.

[9] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dyn.*, vol. 41, pp. 309–325, 2005.

[10] A. Mauroy *et al.*, *Koopman Operator in Systems and Control*, 2020.

[11] M. Haseli and J. Cortés, "Learning Koopman eigenfunctions and invariant subspaces from data: Symmetric subspace decomposition," *IEEE Trans. Autom. Control*, vol. 67, no. 7, pp. 3442–3457, 2021.

[12] 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.

[13] Z. Yuan and J. Cortés, "Data-driven optimal control of bilinear systems," *IEEE Control Syst. Lett.*, vol. 6, pp. 2479–2484, 2022.

[14] I. Markovsky, "Data-driven simulation of generalized bilinear systems via linear time-invariant embedding," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 1101–1106, 2022.

[15] M. Guo *et al.*, "Data-driven stabilization of nonlinear polynomial systems with noisy data," *IEEE Trans. Autom. Control*, vol. 67, no. 8, pp. 4210–4217, 2021.

[16] J. Berberich *et al.*, "Robust data-driven state-feedback design," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1532–1538.

[17] J. Coulson *et al.*, "Data-enabled predictive control: In the shallows of the deepc," in *18th European Control Conference*, 2019, pp. 307–312.

[18] F. Dörfler *et al.*, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 883–897, 2022.

[19] V. Breschi *et al.*, "Data-driven predictive control in a stochastic setting: A unified framework," *Automatica*, vol. 152, p. 110961, 2023.

[20] X. Shang and Y. Zheng, "Convex approximations for a bi-level formulation of data-enabled predictive control," in *6th Annual Learning for Dynamics & Control Conference*. PMLR, 2024, pp. 1071–1082.

[21] J. Berberich *et al.*, "Linear tracking MPC for nonlinear systems—part ii: The data-driven case," *IEEE Trans. Autom. Control*, vol. 67, no. 9, pp. 4406–4421, 2022.

[22] X. Shang *et al.*, "Willems' fundamental lemma for nonlinear systems with koopman linear embedding," *arXiv:2409.16389 (report)*, 2024.

[23] H. J. Van Waarde *et al.*, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," *IEEE Control Syst. Lett.*, vol. 4, no. 3, pp. 602–607, 2020.

[24] Y. Lian *et al.*, "Koopman based data-driven predictive control," *arXiv preprint arXiv:2102.05122*, 2021.

[25] M. Lazar, "Basis-functions nonlinear data-enabled predictive control: Consistent and computationally efficient formulations," in *2024 European Control Conference (ECC)*. IEEE, 2024, pp. 888–893.

[26] L. C. Iacob *et al.*, "Koopman form of nonlinear systems with inputs," *Automatica*, vol. 162, p. 111525, 2024.

[27] H. Shi and M. Q.-H. Meng, "Deep Koopman operator with control for nonlinear systems," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7700–7707, 2022.

[28] M. Alsalti *et al.*, "Data-based control of feedback linearizable systems," *IEEE Trans. Autom. Control*, vol. 68, no. 11, pp. 7014–7021, 2023.