Willems' fundamental lemma for nonlinear systems with Koopman linear embedding

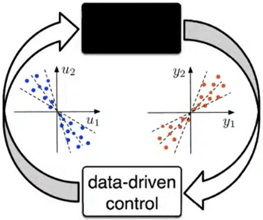
[1]**Xu Shang**, [2]Jorge Cortés, [1]Yang Zheng

[1]Department of Electrical and Computer Engineering, University of California San Diego
[2] Department of Mechanical and Aerospace Engineering, University of California San Diego

ACC 2025

Control of Unknown Systems in a Data-rich World

Key insight of data-driven control: Data $\rightarrow$ Model $+$ Error $\rightarrow$ Control.

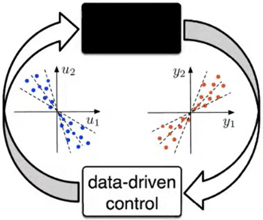


(Markovsky, Huang, and Dörfler 2023)

Control of Unknown Systems in a Data-rich World

Key insight of data-driven control: Data $\rightarrow$ Model $+$ Error $\rightarrow$ Control.

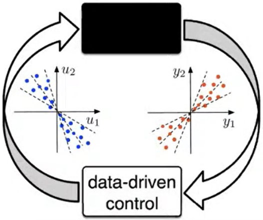
Linear representation:



(Markovsky, Huang, and Dörfler 2023)

Control of Unknown Systems in a Data-rich World

Key insight of data-driven control: Data $\rightarrow$ Model $+$ Error $\rightarrow$ Control.



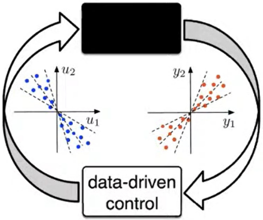
(Markovsky, Huang, and Dörfler 2023)

Linear representation:

- Parametric model: Koopman operator theorem $\rightarrow$ Koopman model predictive control (Koopman 1931; Korda and Mezić 2018)

Control of Unknown Systems in a Data-rich World

Key insight of data-driven control: Data → Model + Error → Control.



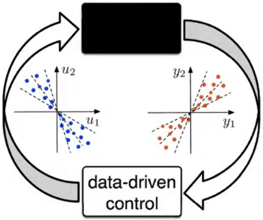
(Markovsky, Huang, and Dörfler 2023)

Linear representation:

- Parametric model: Koopman operator theorem → Koopman model predictive control (Koopman 1931; Korda and Mezić 2018)
- Data-driven model: Willems' fundamental lemma → Data-EnablEd predictive control (Willems *et al.* 2005; Coulson *et al.* 2019)

Control of Unknown Systems in a Data-rich World

Key insight of data-driven control: Data $\rightarrow$ Model $+$ Error $\rightarrow$ Control.

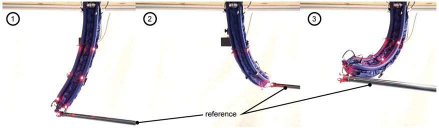


(Markovsky, Huang, and Dörfler 2023)

Linear representation:

- Parametric model: Koopman operator theorem $\rightarrow$ Koopman model predictive control (Koopman 1931; Korda and Mezić 2018)
- Data-driven model: Willems' fundamental lemma $\rightarrow$ Data-EnablEd predictive control (Willems *et al.* 2005; Coulson *et al.* 2019)

(a) Connected vehicles (Wang et al. 2023)



(b) Soft robots (Haggerty *et al.* 2023)

Problem Statement

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 $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ are the state, input, and output of the system.

Problem Statement

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 $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ are the state, input, and output of the system.

Key problem: The linear representation is generally inaccurate for nonlinear systems.

Problem Statement

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 $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ are the state, input, and output of the system.

Key problem: The linear representation is generally inaccurate for nonlinear systems.

Can we construct an accurate linear model for a specific class of nonlinear systems?

Problem Statement

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 $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^p$ are the state, input, and output of the system.

Key problem: The linear representation is generally inaccurate for nonlinear systems.

Can we construct an accurate linear model for a specific class of nonlinear systems?

Goal: Construct an accurate linear representation for the unknown nonlinear system that admits an accurate Koopman linear model from data.

Data-driven Representation via Koopman Operator Theorem

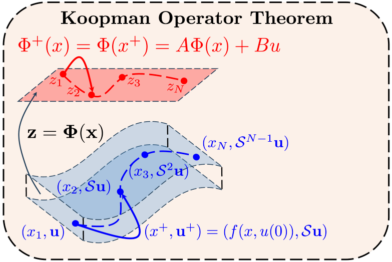
Key idea: Lift state x_k of nonlinear system to a high-dimensional space

Definition 1 (Koopman linear embedding)

There exist a set of lifting functions (observables)

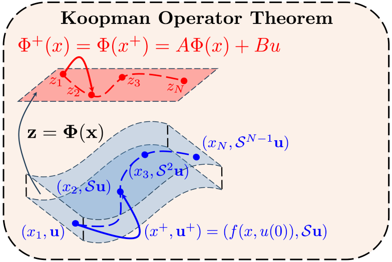
$$\Phi(x_k) := \mathrm{col}(\phi_1(x_k), \ldots, \phi_{n_z}(x_k)) \in \mathbb{R}^{n_z},$$

which propagates linearly along trajectories of the nonlinear system (1).



Data-driven Representation via Koopman Operator Theorem

Key idea: Lift state x_k of nonlinear system to a high-dimensional space



Definition 1 (Koopman linear embedding)

There exist a set of lifting functions (observables)

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

which propagates linearly along trajectories of the nonlinear system (1).

- Let $z_k = \Phi(x_k) \in \mathbb{R}^{n_z}$, the Koopman (parametric) linear model is

$$z_{k+1} = Az_k + Bu_k, \quad y_k = Cz_k + Du_k. \tag{2}$$

- Given x_{ini}, we can obtain $z_{\text{ini}} = \Phi(x_{\text{ini}})$.

Data-driven Representation via Koopman Operator Theorem

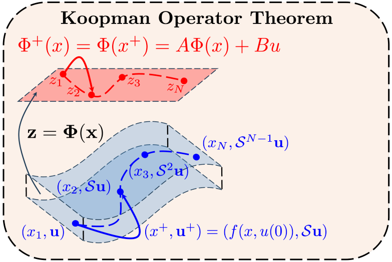
Key idea: Lift state x_k of nonlinear system to a high-dimensional space

Definition 1 (Koopman linear embedding)

There exist a set of lifting functions (observables)

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

which propagates linearly along trajectories of the nonlinear system (1).



- Let $z_k = \Phi(x_k) \in \mathbb{R}^{n_z}$, the Koopman (parametric) linear model is

$$z_{k+1} = Az_k + Bu_k, \quad y_k = Cz_k + Du_k. \tag{2}$$

- Given x_{ini}, we can obtain $z_{\text{ini}} = \Phi(x_{\text{ini}})$.

There does not exist a systematic way to select lifting functions.

Willems' Fundamental Lemma

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

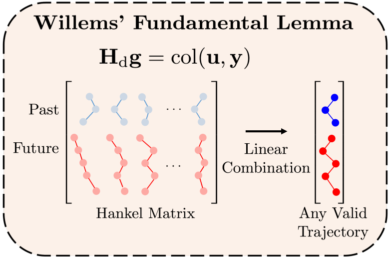
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{3}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (3) *if and only if there exists* g *such that*

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

for rich enough H_{d}.



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

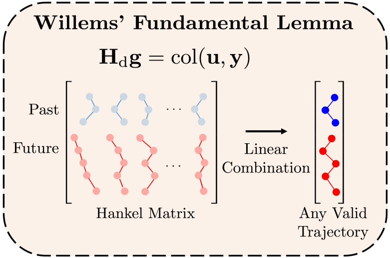
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{3}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (3) *if and only if there exists* g *such that*

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

for rich enough H_d.



- Suppose we have $(u_\mathrm{ini}, y_\mathrm{ini})$ with length T_ini

$$u_{1:T_\mathrm{ini}} = u_\mathrm{ini}, \; y_{1:T_\mathrm{ini}} = y_\mathrm{ini} \quad \Rightarrow \quad H_\mathrm{d} g = \mathrm{col}(u_\mathrm{ini}, y_\mathrm{ini}, u_\mathrm{F}, y_\mathrm{F}).$$

Willems' Fundamental Lemma

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

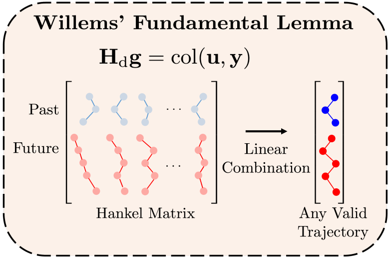
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{3}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (3) *if and only if there exists* g *such that*

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

for rich enough H_d.



- Suppose we have $(u_\mathrm{ini}, y_\mathrm{ini})$ with length T_ini

$$u_{1:T_\mathrm{ini}} = u_\mathrm{ini}, \; y_{1:T_\mathrm{ini}} = y_\mathrm{ini} \quad \Rightarrow \quad H_\mathrm{d} g = \mathrm{col}(u_\mathrm{ini}, y_\mathrm{ini}, u_\mathrm{F}, y_\mathrm{F}).$$

The Willems' fundamental lemma can only be directly applied for LTI systems.

Motivation

Observations:

- **Koopman operator theorem:**
  Nonlinear system $\rightarrow$ Linear system
  High-dimensional state space, same input and output.
- **Willems' fundamental lemma:**
  Unknown linear system $\rightarrow$ Data-driven model
  Directly using input-output trajectories.

Motivation

Observations:

- **Koopman operator theorem:**
  Nonlinear system $\rightarrow$ Linear system
  High-dimensional state space, same input and output.
- **Willems' fundamental lemma:**
  Unknown linear system $\rightarrow$ Data-driven model
  Directly using input-output trajectories.

Key insight:
Unknown nonlinear system $\Rightarrow$ Unknown linear system $\Rightarrow$ Data-driven model

Motivation

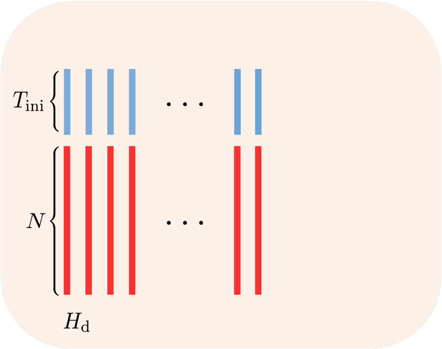
Observations:

- **Koopman operator theorem:**
  Nonlinear system $\rightarrow$ Linear system
  High-dimensional state space, same input and output.
- **Willems' fundamental lemma:**
  Unknown linear system $\rightarrow$ Data-driven model
  Directly using input-output trajectories.

Key insight:
Unknown nonlinear system $\Rightarrow$ Unknown linear system $\Rightarrow$ Data-driven model

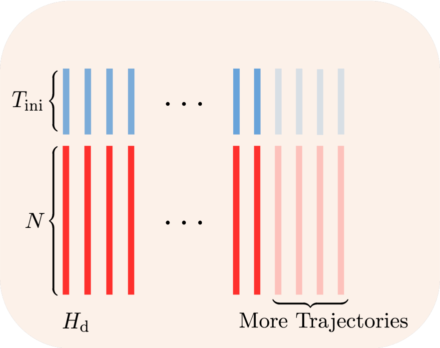
No need to choose lifting functions and the data-driven linear model is accurate.

Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d}.
Predict: length-N future trajectory.

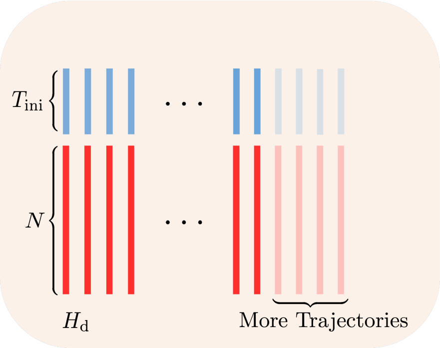
Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d}.
Predict: length-N future trajectory.

- Benefits of collecting more trajectories are well-recognized.

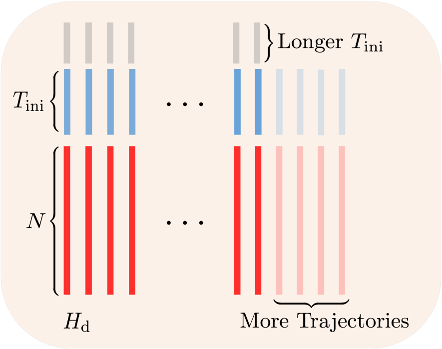
Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d}.
Predict: length-N future trajectory.

- Benefits of collecting more trajectories are well-recognized.

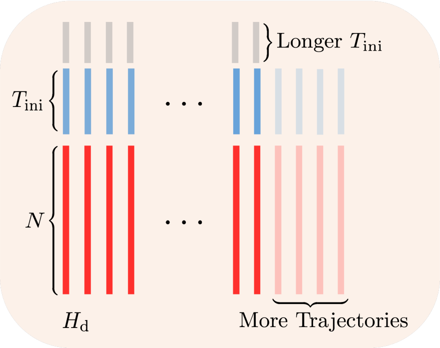
Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d}.
Predict: length-N future trajectory.

- Benefits of collecting more trajectories are well-recognized.
- Importance of extending initial trajectory is overlooked.

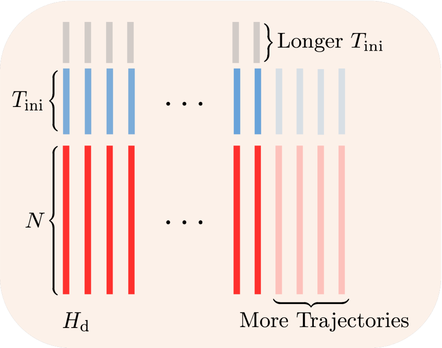
Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d} .
Predict: length-N future trajectory.

- Benefits of collecting more trajectories are well-recognized.
- Importance of extending initial trajectory is overlooked.

Main Result: *Extended Willems' fundamental lemma*



Given: initial trajectory (T_{ini}), trajectory library H_{d}.
Predict: length-N future trajectory.

- Benefits of collecting more trajectories are well-recognized.
- Importance of extending initial trajectory is overlooked.

Main result

Suppose trajectory library H_{d} is rich enough and $T_{\text{ini}} \geq n_{\text{z}}$. The $\text{col}(u_{\text{ini}}, y_{\text{ini}}, u_{\text{F}}, y_{\text{F}})$ is a valid trajectory of nonlinear system if and only if it is in the range space of H_{d}.

Trajectory Spaces of nonlinear system and Koopman linear model:

$$\mathcal{B}_{\mathrm{NL}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, x(0) = x_0 \in \mathbb{R}^n,\ \text{nonlinear system dynamics holds} \right\},$$

$$\mathcal{B}_{\mathrm{K}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, z(0) = z_0 \in \mathbb{R}^{n_z},\ \text{Koopman linear embedding holds} \right\}.$$

Key insight:

1) Willems' fundamental lemma is valid for Koopman linear model, $\mathrm{range}(H_{\mathrm{d}}) = \mathcal{B}_{\mathrm{K}}$.

2) Behavior space of the Koopman model is larger than nonlinear system, $z_0 = \Phi(x_0)$.

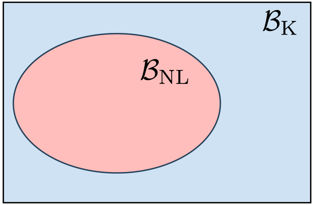
Trajectory Spaces of nonlinear system and Koopman linear model:

$$\mathcal{B}_{\mathrm{NL}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, x(0) = x_0 \in \mathbb{R}^n, \text{ nonlinear system dynamics holds} \right\},$$

$$\mathcal{B}_{\mathrm{K}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, z(0) = z_0 \in \mathbb{R}^{n_z}, \text{ Koopman linear embedding holds} \right\}.$$

Key insight:

1) Willems' fundamental lemma is valid for Koopman linear model, $\mathrm{range}(H_{\mathrm{d}}) = \mathcal{B}_{\mathrm{K}}$.
2) Behavior space of the Koopman model is larger than nonlinear system, $z_0 = \Phi(x_0)$.



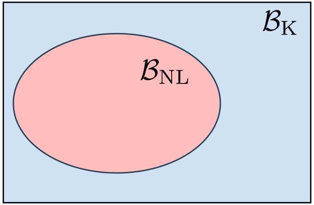
Trajectory Spaces of nonlinear system and Koopman linear model:

$$\mathcal{B}_{\mathsf{NL}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, x(0) = x_0 \in \mathbb{R}^n, \text{ nonlinear system dynamics holds} \right\},$$

$$\mathcal{B}_{\mathsf{K}} = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \mid \exists\, z(0) = z_0 \in \mathbb{R}^{n_z}, \text{ Koopman linear embedding holds} \right\}.$$

Key insight:

1) Willems' fundamental lemma is valid for Koopman linear model, $\mathrm{range}(H_{\mathsf{d}}) = \mathcal{B}_{\mathsf{K}}$.
2) Behavior space of the Koopman model is larger than nonlinear system, $z_0 = \Phi(x_0)$.



- Only trajectories from nonlinear system are available, *i.e.*, columns of H_{d} from $\mathcal{B}_{\mathsf{NL}}$.
- The $\mathrm{col}(u, y) \in \mathcal{B}_{\mathsf{K}}$ is a necessary condition for $\mathrm{col}(u, y) \in \mathcal{B}_{\mathsf{NL}}$.

Behavior Spaces Characterization & Construction

Theorem 1

Suppose $H_{\mathrm{d}} := \begin{bmatrix} u_{\mathrm{d}}^1 & \dots & u_{\mathrm{d}}^l \\ y_{\mathrm{d}}^1 & \dots & y_{\mathrm{d}}^l \end{bmatrix}$ *formed by* l *trajectories from* $\mathcal{B}_{\mathrm{NL}}$ *satisfies lifted excitation, i.e.,* $H_K := \begin{bmatrix} u_{\mathrm{d}}^1 & \dots & u_{\mathrm{d}}^l \\ \Phi(x_0^1) & \dots & \Phi(x_0^l) \end{bmatrix}$ *has full row rank. Then, we have* $\mathrm{range}(H_{\mathrm{d}}) = \mathcal{B}_{\mathrm{K}}$.

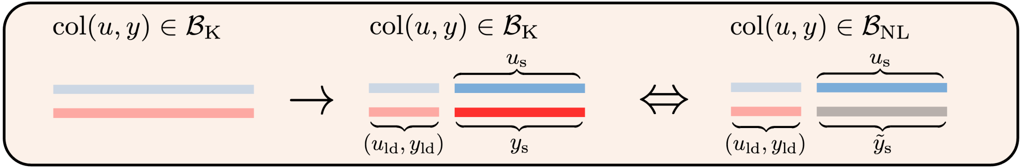
Behavior Spaces Characterization & Construction

Theorem 1

Suppose $H_{\mathrm{d}} := \begin{bmatrix} u_{\mathrm{d}}^1 & \cdots & u_{\mathrm{d}}^l \\ y_{\mathrm{d}}^1 & \cdots & y_{\mathrm{d}}^l \end{bmatrix}$ *formed by* l *trajectories from* $\mathcal{B}_{\mathrm{NL}}$ *satisfies lifted excitation, i.e.,* $H_K := \begin{bmatrix} u_{\mathrm{d}}^1 & \cdots & u_{\mathrm{d}}^l \\ \Phi(x_0^1) & \cdots & \Phi(x_0^l) \end{bmatrix}$ *has full row rank. Then, we have* $\mathrm{range}(H_{\mathrm{d}}) = \mathcal{B}_{\mathrm{K}}$.

Theorem 2

Let $\mathrm{col}(u, y) \in \mathcal{B}_{\mathrm{K}}$. *Then,* $\mathrm{col}(u, y) \in \mathcal{B}_{\mathrm{NL}}$ *if and only if its leading sequence of length-*n_z *(i.e.,* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}})$*) is a valid trajectory of the nonlinear system.*



Data-driven Representation of Nonlinear Systems

Theorem 3

Suppose trajectory library H_{d} *satisfies lifted excitation and* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}) \in \mathcal{B}_{\mathrm{NL}}, (T_{\mathrm{ld}} \geq n_{\mathrm{z}})$. *The sequence* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}})$ *is a valid trajectory of nonlinear system if and only if there exists* g *such that*

$$H_{\mathrm{d}} g = \mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}}).$$

Theorem 3

Suppose trajectory library H_{d} satisfies lifted excitation and $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}) \in \mathcal{B}_{\mathrm{NL}}, (T_{\mathrm{ld}} \geq n_{\mathrm{z}})$*. The sequence* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}})$ *is a valid trajectory of nonlinear system if and only if there exists g such that*

$$H_{\mathrm{d}} g = \mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}}).$$

Remark 1

We require no additional information as we can let

$$u_{\mathrm{ld}} = u_{\mathrm{ini}}, \;\; y_{\mathrm{ld}} = y_{\mathrm{ini}}, \;\; u_{\mathrm{s}} = u_{\mathrm{F}}, \;\; y_{\mathrm{s}} = y_{\mathrm{F}}, \tag{5}$$

in predicting the future output trajectory of the nonlinear system.

Theorem 3

Suppose trajectory library H_{d} *satisfies lifted excitation and* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}) \in \mathcal{B}_{\mathrm{NL}}, (T_{\mathrm{ld}} \geq n_z)$*. The sequence* $\mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}})$ *is a valid trajectory of nonlinear system if and only if there exists* g *such that*

$$H_{\mathrm{d}} g = \mathrm{col}(u_{\mathrm{ld}}, y_{\mathrm{ld}}, u_{\mathrm{s}}, y_{\mathrm{s}}).$$

Remark 1

We require no additional information as we can let

$$u_{\mathrm{ld}} = u_{\mathrm{ini}}, \; y_{\mathrm{ld}} = y_{\mathrm{ini}}, \; u_{\mathrm{s}} = u_{\mathrm{F}}, \; y_{\mathrm{s}} = y_{\mathrm{F}}, \tag{5}$$

in predicting the future output trajectory of the nonlinear system.

Theorem 3 provides an accurate data-driven linear representation.

- Requires no prior knowledge of the nonlinearity or suitable observables;
- Illustrates the importance of increasing the width and depth of the trajectory library.

Numerical Simulations

We consider the following nonlinear system

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

Choose the lifted state as $z := \mathrm{col}(x_1, x_2, x_1^2, x_1^3, x_1^4)$, leading to:

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

Numerical Simulations

We consider the following nonlinear system

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

Choose the lifted state as $z := \mathrm{col}(x_1, x_2, x_1^2, x_1^3, x_1^4)$, leading to:

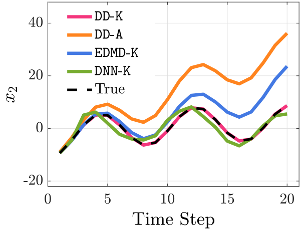
$$z_{k+1} = \begin{bmatrix} 0.99 & 0 & 0 & 0 & 0 \\ 0 & 0.9 & 1 & 1 & 1 \\ 0 & 0 & 0.99^2 & 0 & 0 \\ 0 & 0 & 0 & 0.99^3 & 0 \\ 0 & 0 & 0 & 0 & 0.99^4 \end{bmatrix} z_k + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_k,$$

We compare the prediction and control performance of four linear representations:

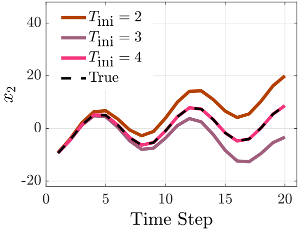
1. our proposed Data-Driven Koopman representation (`DD-K`) (5).
2. Data-Driven Affine representation (`DD-A`) (Berberich *et al.* 2022).
3. Koopman linear approximation (2) from EDMD (`EDMD-K`) (Mauroy *et al.* 2020)
4. Deep Neural Network Koopman representation (`DNN-K`) (Shi and Meng 2022).

Prediction performance:

- DD-K (red) aligns with the true trajectory (black).



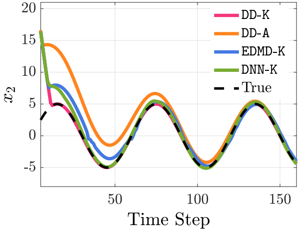
(a) Comparison of linear models



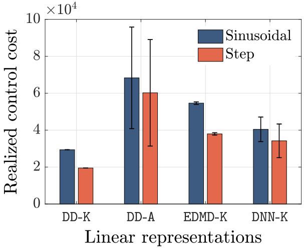
(b) Comparison of different T_{ini}

Control performance:

- Realized control cost: DD-A > EDMD-K > DNN-K > DD-K.



(a) Tracking performance



(b) Realized control cost

Conclusion and Future work

Conclusion:

For nonlinear systems that admit an accurate Koopman linear embedding,

- The Willems' fundamental lemma can be directly applied with
  1. a rich enough trajectory library;
  2. a sufficiently long initial trajectory.
- The selection process of lifting functions is not necessary and can be eliminated.

Future work:

- Linear representation for nonlinear systems with approximated Koopman linear model. (Shang, Li, and Zheng 2025)
- Analyze the effect of adaptively updating the trajectory library.
- Consider coupling nonlinearities between state and input.

Thank you for your attention!

Q & A

References I

- Berberich *et al.*, Julian (2022). "Linear Tracking MPC for Nonlinear Systems—Part II: The Data-Driven Case". In: *IEEE Trans. Autom. Control* 67.9, pp. 4406–4421. DOI: 10.1109/TAC.2022.3166851.
- Coulson *et al.*, Jeremy (2019). "Data-enabled predictive control: In the shallows of the DeePC". In: *18th European Control Conference*, pp. 307–312.
- Haggerty *et al.*, David A (2023). "Control of soft robots with inertial dynamics". In: *Sci. Rob.* 8.81, eadd6864.
- Koopman, Bernard O (1931). "Hamiltonian systems and transformation in Hilbert space". In: *PNAS* 17.5, pp. 315–318.
- Korda, Milan and Igor Mezić (2018). "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control". In: *Automatica* 93, pp. 149–160.

References II

- Markovsky, Ivan, Linbin Huang, and Florian Dörfler (2023). "Data-driven control based on the behavioral approach: From theory to applications in power systems". In: *IEEE Control Systems Magazine* 43.5, pp. 28–68.
- Mauroy *et al.*, Alexandre (2020). *Koopman Operator in Systems and Control*.
- Shang, Xu, Zhaojian Li, and Yang Zheng (2025). "Dictionary-free Koopman Predictive Control for Autonomous Vehicles in Mixed Traffic". In: *arXiv preprint arXiv:2504.06240*.
- Shi, Haojie and Max Q-H Meng (2022). "Deep Koopman operator with control for nonlinear systems". In: *IEEE Robotics and Automation Letters* 7.3, pp. 7700–7707.
- Wang, Jiawei et al. (2023). "DeeP-LCC: Data-enabled predictive leading cruise control in mixed traffic flow". In: *IEEE Transactions on Control Systems Technology* 31.6, pp. 2760–2776.
- Willems *et al.*, Jan C (2005). "A note on persistency of excitation". In: *Syst. Control Lett.* 54.4, pp. 325–329.

Extended Dynamic Decomposition

Matrix parameters A, B, C and D computation:

1. Organize the measured input-state-output data sequence
$$\mathrm{X} = \left[x_0, \ldots, x_{n_\mathrm{d}-2}\right], \quad \mathrm{X}^+ = \left[x_1, \ldots, x_{n_\mathrm{d}-1}\right],$$
$$\mathrm{U} = \left[u_0, \ldots, u_{n_\mathrm{d}-2}\right], \quad Y = \left[y_0, \ldots, y_{n_\mathrm{d}-2}\right].$$
2. Choose observables and compute lifted state
$$\mathrm{Z} = \left[\Phi(x_0), \ldots, \Phi(x_{n_\mathrm{d}-2})\right], \quad \mathrm{Z}^+ = \left[\Phi(x_1), \ldots, \Phi(x_{n_\mathrm{d}-1})\right].$$
3. Least-squares approximations
$$\begin{aligned} (A, B) &\in \underset{A,B}{\operatorname{argmin}} \ \|\mathrm{Z}^+ - A\mathrm{Z} - B\mathrm{U}\|_F^2, \\ (C, D) &\in \underset{C,D}{\operatorname{argmin}} \ \|\mathrm{Y} - C\mathrm{Z} - DU\|_F^2. \end{aligned} \tag{6}$$

Remark 2

The choice of observables affects (6) significantly. Even if a Koopman linear embedding exists for (1), we may not know the correct observables **??** *for such a Koopman linear embedding.*