Dictionary-free Koopman Predictive Control for Autonomous Vehicles in Mixed Traffic

[1]**Xu Shang**, [2]Zhaojian Li, [1]Yang Zheng

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

[2] Department of Mechanical Engineering, Michigan State University

CCTA 2025

Section

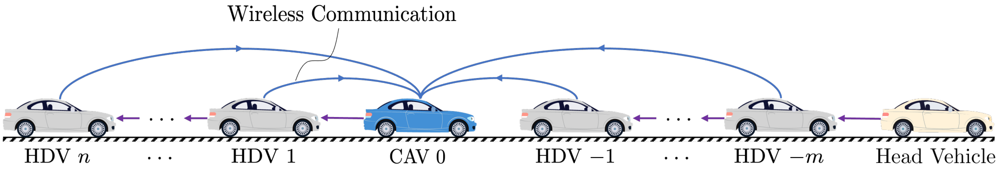
Mixed Traffic System

Background: Small perturbations of vehicle motion may propagate into large periodic speed fluctuations, which lower traffic efficiency and reduce driving safety. Connected and autonomous vehicles (CAVs) have great potential to mitigate traffic jams.

Mixed Traffic System

Background: Small perturbations of vehicle motion may propagate into large periodic speed fluctuations, which lower traffic efficiency and reduce driving safety. Connected and autonomous vehicles (CAVs) have great potential to mitigate traffic jams.

Single-lane Car Following Scenario



- **Mixed traffic** with one CAV multiple HDVs.
- **Mitigate the traffic wave** or **Follow the trajectory** of head vehicle.

Problem Statement

Traditional Method:

- Car-following model: $a_i(t) = F_i(s_i(t), \dot{s}_i(t), v_i(t))$.
- Linearized model-based dynamics around equilibrium state:

$$\textbf{NL:}\quad \begin{aligned} x(k+1) &= f(x(k), u(k)), \\ y(k) &= x(k), \end{aligned} \quad\Longrightarrow\quad \textbf{LL:}\quad \begin{aligned} \tilde{x}(k+1) &= A\tilde{x}(k) + B\tilde{u}(k), \\ \tilde{y}(k) &= C\tilde{x}(k). \end{aligned}$$

Problem Statement

Traditional Method:

- Car-following model: $a_i(t) = F_i(s_i(t), \dot{s}_i(t), v_i(t))$.
- Linearized model-based dynamics around equilibrium state:

$$\textbf{NL:}\ \begin{aligned} x(k+1) &= f(x(k), u(k)), \\ y(k) &= x(k), \end{aligned} \quad \Longrightarrow \quad \textbf{LL:}\ \begin{aligned} \tilde{x}(k+1) &= A\tilde{x}(k) + B\tilde{u}(k), \\ \tilde{y}(k) &= C\tilde{x}(k). \end{aligned}$$

Key Challenge:

- Unknown, nonlinear and nondeterministic behavior of HDVs
  $\rightarrow$ The parametric model (**LL**) is non-trivial to accurately obtain.
- Change of the equilibrium state of the mixed traffic system
  $\rightarrow$ The parametric model (**LL**) needs to be re-identified when the equilibria changes.

Traditional Method:

- Car-following model: $a_i(t) = F_i(s_i(t), \dot{s}_i(t), v_i(t))$.
- Linearized model-based dynamics around equilibrium state:

$$\textbf{NL:}\ \begin{aligned} x(k+1) &= f(x(k), u(k)), \\ y(k) &= x(k), \end{aligned} \quad \Longrightarrow \quad \textbf{LL:}\ \begin{aligned} \tilde{x}(k+1) &= A\tilde{x}(k) + B\tilde{u}(k), \\ \tilde{y}(k) &= C\tilde{x}(k). \end{aligned}$$

Key Challenge:

- Unknown, nonlinear and nondeterministic behavior of HDVs
  $\rightarrow$ The parametric model (**LL**) is non-trivial to accurately obtain.
- Change of the equilibrium state of the mixed traffic system
  $\rightarrow$ The parametric model (**LL**) needs to be re-identified when the equilibria changes.

Goal: Design CAV control input for **unknown system (NL)** using a linear representation over **the operating region** to mitigate the traffic wave.

Data-driven Linear Representation

Available Data:

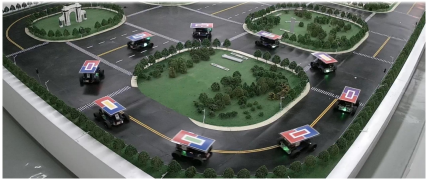
- Input/output trajectory of length-T, *i.e.*, $u_{\mathrm{d}}, y_{\mathrm{d}}$ (Offline data).
- Recent past input/output sequence of length-T_{ini}, *i.e.*, $u_{\mathrm{ini}}, y_{\mathrm{ini}}$ (Online data).

Available Data:

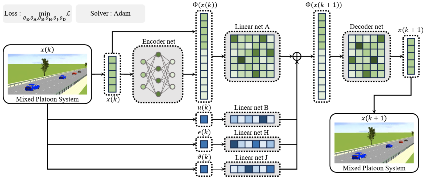
- Input/output trajectory of length-T, *i.e.*, $u_{\mathrm{d}}, y_{\mathrm{d}}$ (Offline data).
- Recent past input/output sequence of length-T_{ini}, *i.e.*, $u_{\mathrm{ini}}, y_{\mathrm{ini}}$ (Online data).

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) Data-driven model (Wang et al. 2023)



(b) Koopman parametric model (Li et al. 2025)

Data-driven Representation via Koopman Operator Theorem

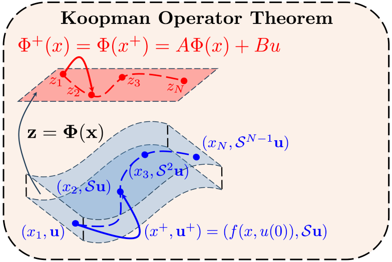
Key idea: Lift state x_k of nonlinear system to a high-dimensional space

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 (**NL**).



Data-driven Representation via Koopman Operator Theorem

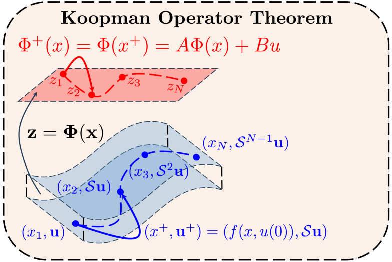
Key idea: Lift state x_k of nonlinear system to a high-dimensional space

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 (**NL**).



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

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

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

Data-driven Representation via Koopman Operator Theorem

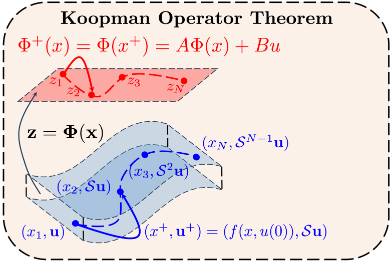
Key idea: Lift state x_k of nonlinear system to a high-dimensional space

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 (**NL**).



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

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

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

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

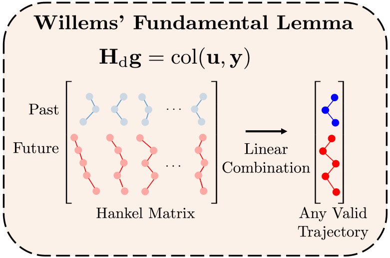
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{2}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (2) *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

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

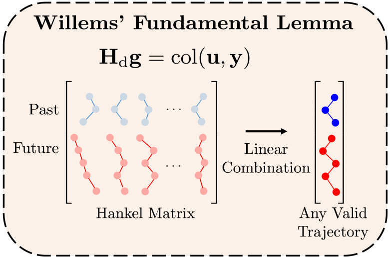
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{2}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (2) *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

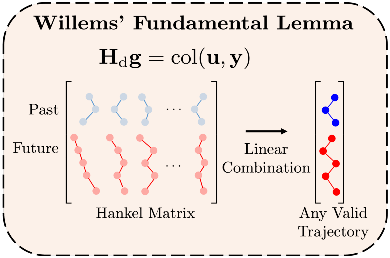
$$x_{k+1} = A_1 x_k + B_1 u_k, \quad y_k = C_1 x_k + D_1 u_k. \tag{2}$$

Lemma 1 (Willems' fundamental lemma)

A length-L input-output data sequence $\mathrm{col}(u, y)$ *is a valid trajectory of* (2) *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 which requires local linearization.

Section

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 Koopman 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 Koopman model

No need to choose lifting functions and the (approximated) data-driven linear model is for the nonlinear system.

Mix Traffic System with Accurate Koopman Model

Definition 2

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

Theorem 1

Suppose trajectory library H_{d} *satisfies lifted excitation. The sequence* $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u, y)$ *with* $T_{\mathrm{ini}} \geq n_{\mathrm{z}}$ *is a valid trajectory of nonlinear system if and only if there exists* g *such that*

$$H_{\mathrm{d}} g = \mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u, y).$$

Mix Traffic System with Accurate Koopman Model

Definition 2

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

Theorem 1

Suppose trajectory library H_{d} *satisfies lifted excitation. The sequence* $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u, y)$ *with* $T_{\mathrm{ini}} \geq n_{\mathrm{z}}$ *is a valid trajectory of nonlinear system if and only if there exists* g *such that*

$$H_{\mathrm{d}} g = \mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, u, y).$$

Theorem 1 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.

Mixed Traffic System with Inaccurate Koopman Model

Search the Koopman linear model minimizing the distance to the collected data $u_\mathrm{d}, y_\mathrm{d}$.

$$\begin{aligned} \min_{A,B,C,D,\bar{y},\Phi\subset\mathcal{F}} \quad & \|\bar{y} - y_\mathrm{d}\|_2^2 \\ \text{subject to} \quad & \mathrm{col}(u_\mathrm{d}, \bar{y}) \in \mathcal{B}_\mathrm{K}(A, B, C, D, \Phi), \end{aligned} \tag{3}$$

Mixed Traffic System with Inaccurate Koopman Model

Search the Koopman linear model minimizing the distance to the collected data $u_\mathrm{d}, y_\mathrm{d}$.

$$\begin{aligned} \min_{A,B,C,D,\bar{y},\Phi\subset\mathcal{F}} \quad & \|\bar{y} - y_\mathrm{d}\|_2^2 \\ \text{subject to} \quad & \mathrm{col}(u_\mathrm{d}, \bar{y}) \in \mathcal{B}_\mathrm{K}(A, B, C, D, \Phi), \end{aligned} \tag{3}$$

Relax (3) via a set of constraints developed from linear system identification techniques.

$$\begin{aligned} \min_{\bar{H}_y,K} \quad & \|H_y - \bar{H}_y\|_F^2 \\ \text{subject to} \quad & \mathrm{rank}(\bar{H}) = mL + n_\mathrm{z}, && \text{(4a)} \\ & \bar{Y}_\mathrm{F} = K \ \mathrm{col}(U_\mathrm{P}, \bar{Y}_\mathrm{P}, U_\mathrm{F}), && \text{(4b)} \\ & K = \begin{bmatrix} K_p & K_f \end{bmatrix}, \ K_f \in \mathcal{L}, && \text{(4c)} \\ & \bar{H}_y \in \mathcal{H}, && \text{(4d)} \end{aligned}$$

where $H_y, \bar{H}_y$ and $\bar{H}$ represent $\mathrm{col}(Y_\mathrm{P}, Y_\mathrm{F})$, $\mathrm{col}(\bar{Y}_\mathrm{P}, \bar{Y}_\mathrm{F})$ and $\mathrm{col}(U_\mathrm{P}, \bar{Y}_\mathrm{P}, U_\mathrm{F}, \bar{Y}_\mathrm{F})$.

Sequential Optimization

We utilize an iterative algorithm to solve (4).

Step 1 (Low-rank approximation): We first only consider constraint (4a).

$$\begin{aligned} \mathbf{\Pi_L(H_y)}: \min_{\bar{H}_y} \quad & \|H_y - \bar{H}_y\|_F \\ \text{subject to} \quad & \text{rank}(\bar{H}) = mL + n_{\mathbf{z}}, \end{aligned}$$

Sequential Optimization

We utilize an iterative algorithm to solve (4).

Step 1 (Low-rank approximation): We first only consider constraint (4a).

$$\begin{aligned} \boldsymbol{\Pi}_{\mathbf{L}}(\mathbf{H_y}) : \min_{\bar{H}_y} \quad & \|H_y - \bar{H}_y\|_F \\ \text{subject to} \quad & \mathrm{rank}(\bar{H}) = mL + n_z, \end{aligned}$$

Step 2 (Causality projection): We then tackle constraints (4b) and (4c)

$$\begin{aligned} \boldsymbol{\Pi}_{\mathbf{C}}(\mathbf{H_{y_1}}) : \min_{\bar{H}_y, K} \quad & \|H_{y_1} - \bar{H}_y\|_F \\ \text{subject to} \quad & \bar{Y}_{\mathrm{F}} = K \ \mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}_1}, U_{\mathrm{F}}), \\ & K = \begin{bmatrix} K_p & K_f \end{bmatrix}, \ K_f \in \mathcal{L}, \end{aligned}$$

Sequential Optimization

We utilize an iterative algorithm to solve (4).

Step 1 (Low-rank approximation): We first only consider constraint (4a).

$$\begin{aligned} \mathbf{\Pi_L(H_y)} : \min_{\bar{H}_y} \quad & \|H_y - \bar{H}_y\|_F \\ \text{subject to} \quad & \operatorname{rank}(\bar{H}) = mL + n_z, \end{aligned}$$

Step 2 (Causality projection): We then tackle constraints (4b) and (4c)

$$\begin{aligned} \mathbf{\Pi_C(H_{y_1})} : \min_{\bar{H}_y, K} \quad & \|H_{y_1} - \bar{H}_y\|_F \\ \text{subject to} \quad & \bar{Y}_\mathrm{F} = K \ \mathrm{col}(U_\mathrm{P}, Y_{\mathrm{P}_1}, U_\mathrm{F}), \\ & K = \begin{bmatrix} K_p & K_f \end{bmatrix}, \ K_f \in \mathcal{L}, \end{aligned}$$

Step 3 (Hankel structure projection): We finally project H_{y_2} to a Hankel matrix set via averaging skew-diagonal elements and denote it as $H_{y_3} := \mathbf{\Pi_H(H_{y_2})}$.

Algorithm Iterative SLRA

Input: U_{P}, U_{F}, Y_{P}, Y_{F}, n_{z}, ϵ

1: $H_y \leftarrow \mathrm{col}(Y_{\mathrm{P}}, Y_{\mathrm{F}})$, $H_{y_3} \leftarrow H_y$;

2: **repeat**

3: $\quad H_{y_1} \leftarrow \Pi_{\mathrm{L}}(H_{y_3})$ $\quad$ (`Low-rank approx`);

4: $\quad H_{y_2} \leftarrow \Pi_{\mathrm{C}}(H_{y_1})$ $\quad$ (`Causality proj`);

5: $\quad H_{y_3} \leftarrow \Pi_{\mathrm{H}}(H_{y_2})$ $\quad$ (`Hankel proj`);

6: **until** $\|H_{y_1} - H_{y_3}\|_F \leq \epsilon \|H_{y_1}\|_F$

Output: $H_y^* = H_{y_1}$

Sequential Optimization

Algorithm Iterative SLRA

Input: U_{P}, U_{F}, Y_{P}, Y_{F}, n_{z}, ϵ

$H_y \leftarrow \mathrm{col}(Y_{\mathrm{P}}, Y_{\mathrm{F}})$, $H_{y_3} \leftarrow H_y$;
repeat
  $H_{y_1} \leftarrow \Pi_{\mathrm{L}}(H_{y_3})$ `(Low-rank approx);`
  $H_{y_2} \leftarrow \Pi_{\mathrm{C}}(H_{y_1})$ `(Causality proj);`
  $H_{y_3} \leftarrow \Pi_{\mathrm{H}}(H_{y_2})$ `(Hankel proj);`
until $\|H_{y_1} - H_{y_3}\|_F \leq \epsilon \|H_{y_1}\|_F$

Output: $H_y^* = H_{y_1}$

- All three optimization problems have analytical solutions.
- This algorithm converges practically.
- The output data-driven Koopman model is guaranteed to satisfy the causality constraint.

Data-driven Koopman linear model (approximated):

$\bar{H}^* = \mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}^*, U_{\mathrm{F}}, Y_{\mathrm{F}}^*)$ where $\mathrm{col}(Y_{\mathrm{P}}^*, Y_{\mathrm{F}}^*) := H_y^*$.

Dictionary-free Koopman model predictive control

We propose the dictionary-free Koopman model predictive control (`DF-KMPC`)

$$\begin{aligned} \min_{g,u\in\mathcal{U},y\in\mathcal{Y}} \quad & \|y-y_{\mathrm{r}}\|_Q + \|u\|_R \\ \text{subject to} \quad & \bar{H}^* g = \mathrm{col}(\Pi_{\mathrm{ini}}(\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})), u, y), \end{aligned} \tag{5}$$

where Π_{ini} projects the intial trajectory to the range space of $\mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}^*)$.

Dictionary-free Koopman model predictive control

We propose the dictionary-free Koopman model predictive control (`DF-KMPC`)

$$\begin{aligned} \min_{g, u \in \mathcal{U}, y \in \mathcal{Y}} \quad & \|y - y_{\mathrm{r}}\|_Q + \|u\|_R \\ \text{subject to} \quad & \bar{H}^* g = \mathrm{col}(\Pi_{\mathrm{ini}}(\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})), u, y), \end{aligned} \tag{5}$$

where Π_{ini} projects the intial trajectory to the range space of $\mathrm{col}(U_{\mathrm{P}}, Y_{\mathrm{P}}^*)$.

Advantages of `DF-KMPC`:

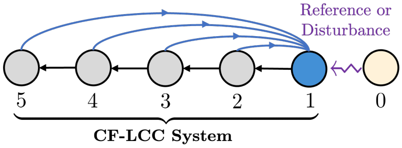
- Do not require extra regularizers to do implicit system identification.
- Do not require relaxation on the initial condition via slack variables.
- Bypass the selection of lifting functions.
- Avoid re-identification when the equilibrium point changes

Section

Numerical Simulations

System Setup:

1 CAV, 4 HDVs behind.



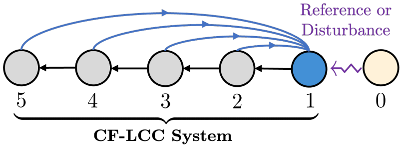
Scenarios:

- Ring road scenario
  → traffic wave mitigation performance.
- Real-data highway scenario
  → velocity tracking performance.

Numerical Simulations

System Setup:

1 CAV, 4 HDVs behind.



Scenarios:

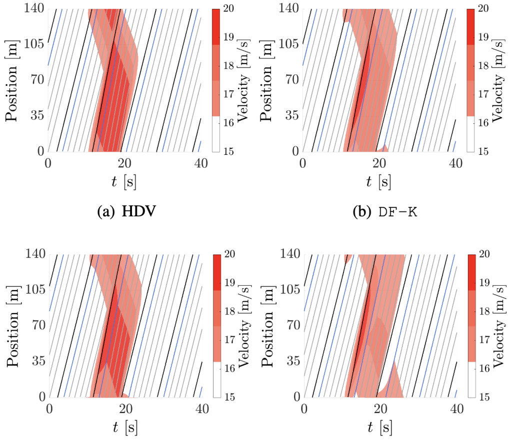
- Ring road scenario
  → traffic wave mitigation performance.
- Real-data highway scenario
  → velocity tracking performance.

We compare the control performance of three different approaches:

1. our proposed Dictionary-free Koopman MPC (`DF-KMPC`) (5).
2. Koopman linear approximation (1) from EDMD (`EDMD-K`) (Mauroy *et al.* 2020).
3. Deep Neural Network Koopman representation (`DNN-K`) (Shi and Meng 2022).

Ring Road Scenario

- When all vehicles are HDV, the traffic wave propagates along the vehicle chain without vanishing.
- Both DF-K and DNN-K can effectively prevent and dampen the propagation of the traffic wave.
- The EDMD-K can mitigate traffic wave but not it is not significant.

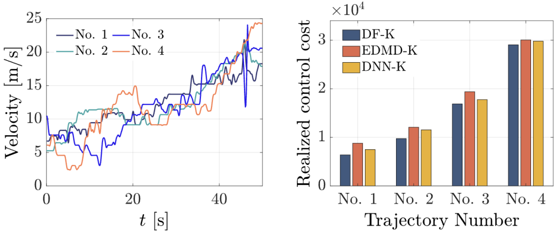


(a) HDV (b) DF-K (c) EDMD-K (d) DNN-K

Real-data Highway Scenario

Realization cost:

- The realization cost illustrates EDMD-K>DNN-K>DF-K.
- DNN-K provides comparable performance in some trajectories but requires much more data.



(a) Velocity trajectorties

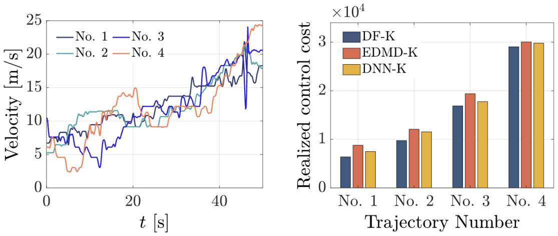
(b) Realization cost

Realization cost:

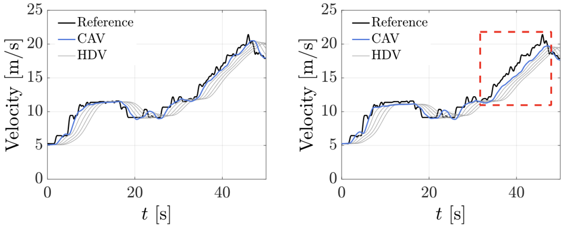
- The realization cost illustrates EDMD-K>DNN-K>DF-K.
- DNN-K provides comparable performance in some trajectories but requires much more data.

Control performance:

- Improper set of lifting functions in EDMD-K can lead to relatively large deviation in velocity tracking.



(a) Velocity trajectorties (b) Realization cost



(a) DF-K (b) EDMD-K

Section

1. Introduction
   - Problem formulation
   - Data-driven linear representation
2. Dictionary-free Koopman MPC
   - Accurate Linear model
   - Approximated Linear model
3. Numerical Simulation
4. Conclusion

Conclusion and Future work

Conclusion:

- We provide a systematic procedure to obtain a (approximated) data-driven representation for the Koopman linear model.
- We develop the `DF-KMPC` for CAV control in mixed traffic.

Our approach 1) bypasses the selection of lifting functions and 2) does not require updating when the equilibrium state changes.

Future work:

- Consider the modeling error and develop a robust `DF-KMPC`.
- Construct a switching algorithm with Koopman linear models computed for different operating regions.
- Test the `DF-KMPC` in larger-scale traffic simulation and real platform.

Thank you for your attention!

Q & A

References I

- Coulson *et al.*, Jeremy (2019). "Data-enabled predictive control: In the shallows of the DeePC". In: *18th European Control Conference*, pp. 307–312.
- 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.
- Li, Shuai et al. (2025). "Robust Nonlinear Data-Driven Predictive Control for Mixed Vehicle Platoons via Koopman Operator and Reachability Analysis". In: *arXiv preprint arXiv:2502.16466*.
- Mauroy *et al.*, Alexandre (2020). *Koopman Operator in Systems and Control*.
- 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.

References II

- 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
$$\begin{aligned} (A, B) &\in \operatorname*{argmin}_{A,B} \|\mathrm{Z}^+ - A\mathrm{Z} - B\mathrm{U}\|_F^2, \\ (C, D) &\in \operatorname*{argmin}_{C,D} \|\mathrm{Y} - C\mathrm{Z} - DU\|_F^2. \end{aligned} \tag{6}$$

Remark 1

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