Koopman operator theory for systems with control

Mircea Lazar

m.lazar@tue.nl

Constrained Control of Complex Systems, EE

Eindhoven University of Technology

European Control Conference 2026

July 8, Reykjavík, Iceland

Outline

1. Classical Koopman Operator Theory
2. Koopman with control: From composition operator to Koopman model
3. Koopman with control: From data to Koopman model
4. Illustrative Examples and Concluding remarks

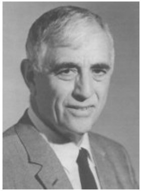
Outline

The Koopman composition operator for autonomous systems

Autonomous system $x_{t+1} = F_0(x_t)$, $F_0 : X \to X$, with observables $\varphi \in \mathcal{H}_x := L^2(X, \mu_x)$.

Koopman (composition) operator, linear in φ:

$$\boxed{(C_{F_0}\varphi)(x) = \varphi \circ F_0(x)}, \qquad C_{F_0} : \mathcal{H}_x \to \mathcal{H}_x.$$



B. O. Koopman, *PNAS, 1931.*

Well-posedness (invariance)

$$\varphi \circ F_0 \in \mathcal{H}_x, \qquad \forall \varphi \in \mathcal{H}_x.$$

Classical setting: F_0 measure-preserving and bijective $\Rightarrow$ C_{F_0} unitary $\Rightarrow$ invariance.

Koopman model: Exact infinite-dimensional lifted dynamics

Lift the state with a basis $\{\psi_{i,x}\}_{i\geq 1}$ of $\mathcal{H}_x$:

$$\Psi_x(x) = \begin{bmatrix} \psi_{1,x}(x) \\ \psi_{2,x}(x) \\ \vdots \end{bmatrix}, \qquad z_t := \Psi_x(x_t).$$

Lifted dynamics: the composition operator applied entrywise

$$\boxed{z_{t+1} = \big(C_{F_0}\Psi_x\big)(x_t) = \Psi_x \circ F_0(x_t) = \mathcal{K} z_t}$$

- An exact representation, linear in the observables, of the nonlinear map F_0;
- Orthonormal basis of $L^2(X, \mu_x)$ and boudedness of C_{F_0} enables convergence of data-driven finite-dimensional Koopman models to $\mathcal{K}$.[1]

[1] M. Korda & I. Mezić, *J. Nonlinear Sci.* 28:687–710, 2018.

Koopman model: The lifting picture

$$
\begin{tikzcd}[column sep=huge, row sep=large]
z_t \arrow[r, "z_{t+1} = \mathcal{K}\, z_t \quad \text{(linear)}"] & z_{t+1} \arrow[d, "\text{decode } D"] \\
x_t \arrow[u, "\text{lift } \Psi_x"] \arrow[r, "x_{t+1} = F_0(x_t) \quad \text{(nonlinear)}"'] & x_{t+1}
\end{tikzcd}
$$

The nonlinear F_0 becomes exactly linear on the lifted state; the picture every controlled extension will try to preserve.

Outline

Overview of the discussed approaches

For $x_{t+1} = F(x_t, u_t)$ the autonomous construction does not apply directly. We analyze existing approaches using the same 3 aspects: 1) the composition operator (domain/codomain); 2) the invariance property; and 3) the Koopman model class.

(i) linear		(ii) bilinear		(iii) KCF		(iv) GeKo
one $\mathcal{K}$ on $X \times \ell(U)$	→	$m+1$ frozen $\mathcal{K}_u$	→	the family $\{\mathcal{K}_u\}_{u \in U}$	→	one $\mathcal{K}$ on $\mathcal{H}_x \otimes \mathcal{H}_u$
$z^+ = Az + Bu$		$z^+ = Az + Bu + \sum_i N_i z u_i$		$z^+ = A(u)\, z$		$z^+ = \mathcal{K}(z \otimes v)$
Korda–Mezić '18		Goswami–Paley '17–'22		Haseli–Cortés '25–'26		Lazar '25–'26

If the technical details get too dense, hang on for "The lifting picture". For full derivations we refer to the cited papers.

(i) Linear Koopman models: lifting with infinite input sequences

1. Composition operator (Korda–Mezić, *Automatica*, 2018)

Extended state $\chi := (x, \boldsymbol{u}) \in X \times \ell(U)$, $\boldsymbol{u} = (u_0, u_1, \dots)$ an *infinite input sequence*; extended dynamics $\chi_{t+1} = S(\chi_t) := \big(F(x_t, \boldsymbol{u}_t(0)),\ \sigma \boldsymbol{u}_t\big)$ with the left shift $(\sigma \boldsymbol{u})(i) := \boldsymbol{u}(i+1)$. Koopman operator on extended observables:

$$\boxed{(\mathcal{K}\phi)(\chi) = \phi \circ S(\chi)}, \qquad \mathcal{K} : \mathcal{H} \to \mathcal{H}, \quad \mathcal{H} = L^2\big(X \times \ell(U), \mu_\chi\big),$$

Key idea: appending the future inputs makes the dynamics *autonomous* again, so the classical construction applies verbatim on the extended space.

(i) Linear Koopman models: structure of $\mathcal{K}$ and the finite model

2. Assumptions and the structure of $\mathcal{K}$

Restrict to the subspace spanned by state-only observables $\psi_{i,x}(x)$ and the first input coordinate $\boldsymbol{u}(0)$. On this subspace, the $L^2(\mu_\chi)$-orthogonal projection of $\mathcal{K}$ (its *compression* $P_V\mathcal{K}|_V$) is block-structured:

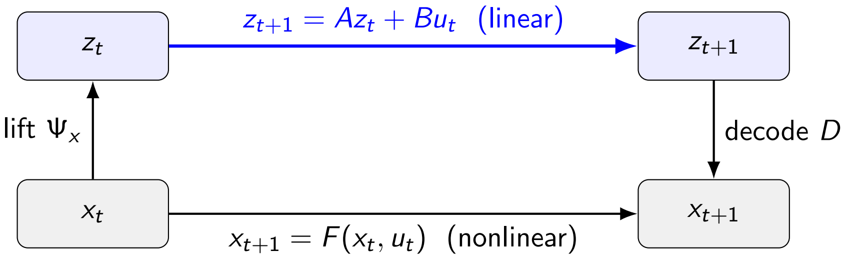
$$\begin{bmatrix} \Psi_x(x_{t+1}) \\ \boldsymbol{u}_{t+1}(0) \end{bmatrix} \approx \begin{bmatrix} A & B \\ * & * \end{bmatrix} \begin{bmatrix} \Psi_x(x_t) \\ \boldsymbol{u}_t(0) \end{bmatrix} \quad \Rightarrow \quad \text{only the state rows } [\,A \quad B\,] \text{ are kept.}$$

The shifted input coordinates are *not modeled*: future inputs are decision variables or known.

3. Finite-dimensional model

$$\boxed{z_{t+1} = Az_t + Bu_t}, \qquad x_t = Dz_t.$$

(i) Linear Koopman models: the lifting picture



One linear map on the lifted state; the input enters additively, unlifted.

Convergence? functions linear in *unlifted* u do not form a basis in general. Convergence to an exact Koopman form is not guaranteed.

(ii) Bilinear Koopman models: constant-input operators

1. Composition operator(s)

Fix (freeze) $u \in U$ and set $F_u := F(\cdot, u)$: each constant-input system $x^+ = F_u(x)$ is autonomous, with Koopman operator

$$\boxed{(\mathcal{K}_u \varphi)(x) = \varphi \circ F_u(x)}, \qquad \mathcal{K}_u : \mathcal{H}_x \to \mathcal{H}_x \quad \text{(one operator per } u\text{).}$$

2. Underlying assumptions

- Control-affine dynamics: $\dot{x} = f_0(x) + \sum_{j=1}^{m} g_j(x) u_j$;
- the Koopman generator of the frozen flow $\mathcal{K}_u^t$, $\mathcal{L}_u \varphi := \lim_{t \searrow 0} \frac{1}{t}(\mathcal{K}_u^t \varphi - \varphi) = \nabla \varphi^\top f_u$ pointwise, is affine in u: $\mathcal{L}_u = \mathcal{L}_0 + \sum_j u_j (\mathcal{L}_{e_j} - \mathcal{L}_0) \Rightarrow$ exact *bilinear* lifted ODE $\dot{z} = Az + Bu + \sum_j N_j z\, u_j$ (Goswami & Paley, *TAC*, 2022);
- discrete time: $\mathcal{K}_u^{T_s} = e^{T_s \mathcal{L}_u}$, and the exponential of an affine map is not affine: bilinearity holds up to $O(T_s^2)$, exact as $T_s \to 0$.

(ii) Bilinear Koopman models: SafEDMD and the finite model

Discrete-time bilinear surrogate (SafEDMD)

Use only $m+1$ members of the constant-input family, $\mathcal{K}_0, \mathcal{K}_{e_1}, \ldots, \mathcal{K}_{e_m}$, learned from data generated with the frozen inputs $u = 0$ and $u = e_j$, and interpolate affinely: $\mathcal{K}_u \approx \mathcal{K}_0 + \sum_{j=1}^{m} u_j \big(\mathcal{K}_{e_j} - \mathcal{K}_0\big)$.

Interpolated Koopman generators/operators: Peitz, Otto & Rowley, *SIAM JADS* 19(3), 2020.

3. Finite-dimensional model

$$\boxed{z_{t+1} = Az_t + Bu_t + \sum_{j=1}^{m} N_j\, z_t\, u_{j,t}}$$

Guarantees: proportional error bounds w.r.t. $\|x\|$, $\|u\|$ enable controller design with closed-loop guarantees; the control-affine assumption is structural; non-affine input maps are not directly captured. Strässer, Schaller, Worthmann, Berberich & Allgöwer, "SafEDMD," *Automatica* 185:112732, 2026.

(ii) Bilinear Koopman models: the lifting picture

$$
\begin{tikzcd}
z_t \arrow[r, "z^+ = Az + Bu + \sum_j N_j\, z\, u_j"] & z_{t+1} \arrow[d, "\text{decode } D"] \\
x_t \arrow[u, "\text{lift } \Psi_x"] \arrow[r, "x_{t+1} = F(x_t, u_t) \quad \text{(nonlinear)}"'] & x_{t+1}
\end{tikzcd}
$$

Convergence? yes for input-affine dynamics: it requires exact bilinearization in u and $T_s \to 0$.

(iii) Koopman control family: all constant-input operators at once

1. Composition operator(s)

The *Koopman control family* (KCF) is the same constant-input family as in (ii), now taken in its entirety:

$$\boxed{\{\mathcal{K}_u\}_{u\in U}, \qquad \mathcal{K}_u\varphi = \varphi \circ F_u}, \qquad \mathcal{K}_u : \mathcal{F} \to \mathcal{F} \;\; \forall u \in U,$$

$\mathcal{F}$: one *common* linear space of *state* observables $\varphi : X \to \mathbb{K}$ (not functions of (x, u)), invariant under *every* frozen-input composition: $\varphi \circ F_u \in \mathcal{F}\ \forall u$; $\mathcal{F} \subseteq L^2(X, \mu_x)$.

2. Underlying assumption: a common invariant subspace

$V \subset \mathcal{F}$ is invariant under the KCF if $\mathcal{K}_u\varphi \in V$ for all $\varphi \in V$, $u \in U$. *Theorem (Haseli–Cortés):* the KCF admits a finite-dimensional common invariant subspace $V = \mathrm{span}(\Psi)$ **iff** the dynamics admit the input-state separable form

$$\Psi(x^+) = A(u)\,\Psi(x), \qquad A : U \to \mathbb{K}^{n\times n}.$$

M. Haseli & J. Cortés, *Automatica* 185:112722, 2026.

(iii) Koopman control family: normal subspaces and the finite model

Constructing $A(u)$ via the augmented operator

Augmented system $(x^+, u^+) = (F(x,u), u)$ with operator $\mathcal{K}^{\mathrm{aug}}$ on $\mathcal{F}^{\mathrm{aug}}$. If a normal subspace $\Upsilon(x,u) = \begin{bmatrix} I \\ G(u) \end{bmatrix} \Psi(x)$ is invariant under $\mathcal{K}^{\mathrm{aug}}$, then $V = \mathrm{span}(\Psi)$ is KCF-invariant.

3. Finite-dimensional model

$$\boxed{\Psi(x^+) = \big(A_{11} + A_{12}\, G(u)\big)\, \Psi(x)}$$

The input enters through a general matrix-valued $G(u)$.

Special cases: linear ($z^+ = Az + Bu$) and bilinear lifted models are input-state separable with particular $G(u)$; exactness hinges on the existence of the common invariant subspace, quantified by *invariance proximity* (later).

(iii) Koopman control family: the lifting picture

$$
\begin{tikzcd}
z_t \arrow[rrr, "z_{t+1} = (A_{11} + A_{12}G(u_t))\, z_t"] & & & z_{t+1} \arrow[d, "\text{decode } D"] \\
x_t \arrow[u, "\text{lift } \Psi_x"] \arrow[rrr, "x_{t+1} = F(x_t, u_t) \quad \text{(nonlinear)}"'] & & & x_{t+1}
\end{tikzcd}
$$

Convergence? adding more observables does not improve invariance in general; one settles for approximate invariant subspaces and Koopman forms.

So far, a discrepancy w.r.t. the classical Koopman operator is apparent: the observables basis and underlying natural convergence mechanism is missing.

(iv) Generalized Koopman (GeKo) operator: the main idea

For a controlled system $x_{t+1} = F(x_t, u_t)$, $F : X \times U \to X$, lift the state in $\mathcal{H}_x$ and the input in $\mathcal{H}_u$ *independently*, and compose with F:

$$\boxed{(C_F\varphi)(x,u) := \varphi \circ F(x,u)}, \qquad C_F : \mathcal{H}_x \to \mathcal{H}, \quad \mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_u \cong L^2(X \times U,\ \mu_x \otimes \mu_u).$$

Generalized invariance condition

$$\varphi \circ F \in \mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_u, \qquad \forall \varphi \in \mathcal{H}_x.$$

Key change: the image $\varphi \circ F$ depends on x *and* u, so the target space is no longer $\mathcal{H}_x$ but $\mathcal{H}$.

Key feature: the input u is not treated as an artificial state or parameter of the dynamics.

Lifted observables

Stack the observables into *lifted vectors*:

$$\Psi_x(x) = \mathrm{col}(\psi_{1,x}(x), \psi_{2,x}(x), \ldots), \qquad \Psi_u(u) = \mathrm{col}(\psi_{1,u}(u), \psi_{2,u}(u), \ldots),$$

and form the product lifting

$$\boxed{\Psi(x,u) := \Psi_x(x) \otimes \Psi_u(u) = \mathrm{col}(\psi_1, \psi_2, \ldots), \quad \psi_l = \psi_{i,x}\, \psi_{j,u}}.$$

Goal: prove that there exists an exact lifted dynamics $\Psi_x(F(x,u)) = \mathcal{K}\, \Psi(x,u)$.

Intermezzo: Riesz bases

A sequence $\{\psi_i\}_{i\geq 1}$ is a *Riesz basis* of a Hilbert space $\mathcal{G}$ if, for constants $0 < \alpha \leq \beta < \infty$,

$$\alpha \sum_i |c_i|^2 \leq \Big\| \sum_i c_i \psi_i \Big\|_{\mathcal{G}}^2 \leq \beta \sum_i |c_i|^2, \quad \forall (c_i) \in \ell^2.$$

F. Riesz

Synthesis operators of the Riesz bases, both bounded and boundedly invertible:

$$W_x : \ell^2 \to \mathcal{H}_x, \;\; W_x c = \sum_i c_i \psi_{i,x}; \qquad W : \ell^2 \to \mathcal{H}, \;\; Wd = \sum_j d_j \psi_j.$$

Explicit form: $\mathcal{K} = Q_x R^{-1}$

Result (existence and form)

If $\{\psi_{i,x}\}$, $\{\psi_{i,u}\}$ are Riesz bases of $\mathcal{H}_x$, $\mathcal{H}_u$ and $\psi_{i,x} \circ F \in \mathcal{H}$ for all i, then R is invertible and $\mathcal{K} = Q_x R^{-1}$.

$$R = \begin{bmatrix} \langle \psi_1, \psi_1 \rangle_{\mathcal{H}} & \langle \psi_1, \psi_2 \rangle_{\mathcal{H}} & \cdots \\ \langle \psi_2, \psi_1 \rangle_{\mathcal{H}} & \langle \psi_2, \psi_2 \rangle_{\mathcal{H}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \qquad Q_x = \begin{bmatrix} \langle \psi_{1,x} \circ F, \psi_1 \rangle_{\mathcal{H}} & \langle \psi_{1,x} \circ F, \psi_2 \rangle_{\mathcal{H}} & \cdots \\ \langle \psi_{2,x} \circ F, \psi_1 \rangle_{\mathcal{H}} & \langle \psi_{2,x} \circ F, \psi_2 \rangle_{\mathcal{H}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

R is the Gram matrix of the product basis; these inner products can be approximated from data.[2]

[2] Builds on the construction for *autonomous* systems by H. H. Asada, "Global, unified representation of heterogeneous robot dynamics using composition operators: a Koopman direct encoding method," *IEEE/ASME Trans. Mechatronics* 28(5):2633–2644, 2023.

Relation between $\mathcal{K}$ ($\mathcal{K}_0$) and C_F (C_{F_0})

$$
\begin{tikzcd}
\ell^2 \arrow[r, "\mathcal{K}_0^\top"] \arrow[d, "W_x"'] & \ell^2 \arrow[d, "W_x"] \\
\mathcal{H}_x \arrow[r, "C_{F_0}"'] & \mathcal{H}_x
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
\ell^2 \arrow[r, "\mathcal{K}^\top"] \arrow[d, "W_x"'] & \ell^2 \arrow[d, "W"] \\
\mathcal{H}_x \arrow[r, "C_F"'] & \mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_u
\end{tikzcd}
$$

Autonomous: $\mathcal{K}_0^\top = W_x^{-1} C_{F_0} W_x$

With control input : $\mathcal{K}^\top = W^{-1} C_F W_x$

Coordinate representation of C_F

From composition operator to matrix $\mathcal{K}$

With the standard ℓ^2 basis $\{e_i\}$, synthesis acts by $W_x e_i = \psi_{i,x}$, $W e_j = \psi_j$. Since $\psi_{i,x} \circ F = \sum_j \mathcal{K}_{ij}\, \psi_j$ (coordinates = the i-th *row* of $\mathcal{K}$),

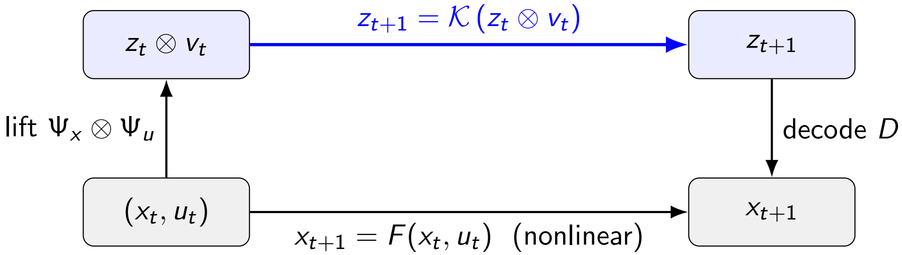
$$C_F\, W_x\, e_i = \sum_j \mathcal{K}_{ij}\, \psi_j = W \underbrace{\textstyle\sum_j \mathcal{K}_{ij}\, e_j}_{=\mathcal{K}^\top e_i} = W\, \mathcal{K}^\top e_i.$$

Since the above relation is valid for all i it follows that:

$$C_F\, W_x = W\, \mathcal{K}^\top \quad \Longrightarrow \quad \boxed{\mathcal{K}^\top = W^{-1} C_F\, W_x}.$$

Above $C_F\, W_x$ denotes operator *composition*, and $W^{-1} C_F\, W_x$, the coordinate representation, is the operator analogue of a change of basis.

(iv) GeKo: the lifting picture



Both states and inputs are lifted ($v_t = \Psi_u(u_t)$); one *exact* linear bounded operator. Convergence?

GeKo: strong operator convergence

Operator convergence ($n_z, n_v \to \infty$)

$\mathcal{K}^{(n_z,n_v)} = \Pi_{n_z}\mathcal{K}\,\Pi_{n_z,n_v}$ with oblique Petrov–Galerkin projections (Riesz, not orthonormal); the projections factor, $\Pi_{n_z,n_v} = \Pi^x_{n_z} \otimes \Pi^u_{n_v}$, and are uniformly bounded by $\sqrt{\beta/\alpha}$, hence

$$\mathcal{K}^{(n_z,n_v)}c - \mathcal{K}c = \Pi_{n_z}\mathcal{K}(\Pi_{n_z,n_v} - I)c + (\Pi_{n_z} - I)\mathcal{K}c \;\to\; 0.$$

$$\boxed{\mathcal{K}^{(n_z,n_v)} \longrightarrow \mathcal{K} \qquad \text{(strong operator convergence)}}$$

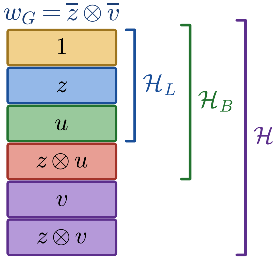
Applicable to non-measure-preserving, general $F(x,u)$ and Riesz dictionaries.

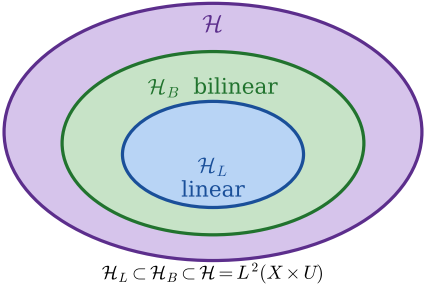
Mirrors the convergence proof for the autonomous case[3]: without u ($\mathcal{H} = \mathcal{H}_x$, $\mathcal{K}_0$, C_{F_0}) the same Galerkin projection argument recovers classical EDMD convergence.

[3] M. Korda & I. Mezić, *J. Nonlinear Sci.* 28:687–710, 2018.

Relevant implication: a nested model-class hierarchy

Linear and bilinear models satisfy $\mathcal{H}_L \subset \mathcal{H}_B \subset \mathcal{H}$. All models share the output projection $\Pi_{n_{\bar{z}}}$; only the *domain* projection differs: $\mathcal{K}_L = \Pi_{n_{\bar{z}}}\mathcal{K}\,\Pi^L$, $\mathcal{K}_B = \Pi_{n_{\bar{z}}}\mathcal{K}\,\Pi^B$, $\mathcal{K} = \Pi_{n_{\bar{z}}}\mathcal{K}\,\Pi_{n_{\bar{z}},n_{\bar{v}}}$.





Monotone floors $\varepsilon_L \geq \varepsilon_B \geq \varepsilon$ enable data-driven *Koopman model-class selection*.

The four routes at a glance

Approach	Operator: domain $\to$ codomain	Key assumption	Finite model
Linear	$\mathcal{K} : \mathcal{H} \to \mathcal{H}$ on $X \times \ell(U)$	exact iff F linear in the lifted coordinates	$z^+ = Az + Bu$
Bilinear	family $\{\mathcal{K}_0, \mathcal{K}_{e_1}, \dots, \mathcal{K}_{e_m}\}$ on $\mathcal{H}_x$	control-affine F; $T_s \to 0$	$z^+ = Az + Bu + \sum_j N_j z u_j$
KCF	family $\{\mathcal{K}_u\}_{u \in U} : \mathcal{F} \to \mathcal{F}$	common finite-dim. invariant subspace	$\Psi(x^+) = A(u)\Psi(x)$
GeKo	single $C_F : \mathcal{H}_x \to \mathcal{H}_x \otimes \mathcal{H}_u$	bounded R–N derivative; Riesz bases	$z^+ = \mathcal{K}(z \otimes v)$

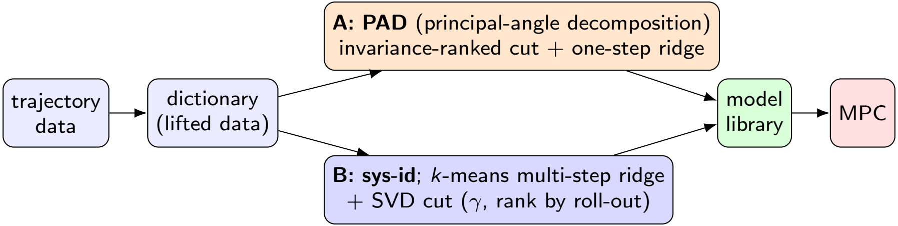
Note: with full coupling (KCF) or sparsity (GeKo) the two finite forms can be rendered *structurally identical*; infinite dim. forms are still fundamentally different as well as their corresp. invariance and convergence properties.

For a detailed comparison and EDMDc for all 4 routes, including KCF: I. Mezić, J. Cortés, K. Worthmann, M. Lazar & A. Lederer, "Koopman operator theory: fundamentals, control, and applications," arXiv:2607.01819, 2026.

Outline

Learning Koopman models from data: two pipelines

One-step least squares $\neq$ small roll-out error: residuals compound through the recursion. Two independent pipelines, each producing linear, bilinear and GeKo models:



Every model in the library is verified on two measures:

- invariance of its subspace (principal-angle certificates);
- true-state roll-out error on test data.

Common data-driven setup

Data snapshots $\{(x_t, u_t, x_{t+1})\}_{t=0}^{T-1}$, no knowledge of F. State dictionary $\psi^x : \mathbb{R}^d \to \mathbb{R}^{n_z}$, input dictionary $\psi^u : \mathbb{R}^m \to \mathbb{R}^{n_v}$ (where needed). Data matrices:

$$Z_X = [\, z_0 \;\; \cdots \;\; z_{T-1} \,], \quad Z_X^+ = [\, z_1 \;\; \cdots \;\; z_T \,], \quad U = [\, u_0 \;\; \cdots \;\; u_{T-1} \,], \quad V_U = [\, v_0 \;\; \cdots \;\; v_{T-1} \,].$$

Khatri–Rao (column-wise Kronecker) product

$$\boxed{P \odot Q := [\, p_0 \otimes q_0 \mid \cdots \mid p_{T-1} \otimes q_{T-1} \,]}$$

The t-th column collects the *product observables* evaluated at snapshot t.

EDMDc: linear and bilinear forms

Linear EDMDc (Korda–Mezić)

$$[\,\hat{A}\quad \hat{B}\,] = \arg\min_{A,B} \left\| Z_X^+ - AZ_X - BU \right\|_{\mathrm{Fr}}^2 = Z_X^+ W^\top (WW^\top)^\dagger, \qquad W := \begin{bmatrix} Z_X \\ U \end{bmatrix}.$$

Bilinear EDMDc (Khatri–Rao regressors in u)

$$\Theta := \begin{bmatrix} Z_X \\ U \\ U \odot Z_X \end{bmatrix}, \qquad [\,\hat{A}\ \hat{B}\ \hat{N}_1 \cdots \hat{N}_m\,] = Z_X^+ \Theta^\top (\Theta\Theta^\top)^\dagger.$$

Both are one-step least squares; Tikhonov regularization $(\cdot)^\dagger_\gamma = (\cdot)^\top [(\cdot)(\cdot)^\top + \gamma I]^{-1}$ used throughout.

EDMDc: GeKo form

GeKo: Khatri–Rao EDMDc

Fit one matrix advancing the tensor-product lift one step:

$$\boxed{\hat{\mathcal{K}} = Z_X^+ \, (Z_X \odot V_U)^\dagger \in \mathbb{R}^{n_z \times n_z n_v}}$$

The product Hilbert-space basis expansion; this structure is what enabled convergence (shown in the previous part).

Decoder (all model forms)

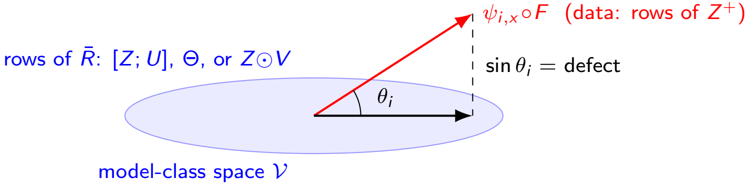
Map the lifted state back to the state, by least squares on the same data:

$$\boxed{\hat{D} = X^+ \, (Z_X^+)^\dagger = X^+ (Z_X^+)^\top (Z_X^+ (Z_X^+)^\top + \gamma I)^{-1} \in \mathbb{R}^{d \times n_z}}, \qquad \hat{x} = \hat{D}\, z.$$

The state is included in the dictionary, so decoding is exact in full coordinates; after PAD/SVD reduction, $\hat{D}$ is refitted in the reduced coordinates, per model.

Pipeline A: PAD; invariance-maximizing reduction

Rank directions by invariance (principal angles):



Like the angle between two vectors, extended to subspaces: with normalized bases $Q_{\bar{R}}$, Q_{Z^+} of the regressor $\bar{R}$ and image Z^+, one SVD of the overlap $Q_{\bar{R}}^\top Q_{Z^+}$ gives all angles, sorted. Keep the smallest-angle directions (rank by the roll-out criterion), then one-step ridge, small fixed γ.

PAD (autonomous): Conradie et al., arXiv:2603.15091, 2026. PAD for control: this tutorial + software.

Pipeline A: invariance proximity for systems with control

Per model class, compare two subspaces of functions of (x, u), each spanned by data rows:

- model-class space $\mathcal{V} = \mathrm{rowspan}(\bar{R})$: $[Z; U]$ (linear), Θ (bilinear), $Z \odot V$ (GeKo);
- image space $\mathcal{I} = \mathrm{rowspan}(Z^+)$: the state observables pushed one step through the dynamics, $\psi_{i,x} \circ F$ on the data.

The overlap SVD gives the full angle spectrum $\theta_1 \le \cdots \le \theta_q$ (the invariance plots); the certificate is the worst angle:

Invariance proximity of a controlled model class

$$\boxed{I_\infty = \sin \theta_{\max}}$$

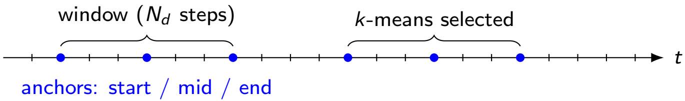
no operator in the class can predict the worst image direction with one-step relative error below I_∞; $I_\infty = 0 \iff$ an exact finite Koopman model of the class exists.

Computed identically for full and reduced (PAD/SVD) models.

Invariance proximity & CCI (autonomous/KCF setting): Haseli & Cortés, *Automatica* 185:112722, 2026; here applied to the controlled model classes.

Pipeline B: system identification; multi-step regression

Tell the regression that prediction is more than one step: fit on sliding windows along trajectories, not isolated one-step pairs.



k-means clusters candidate windows by their lifted anchors (start/mid/end) and keeps one representative per cluster: coverage of the dynamics without redundant data. Stack the shift-k blocks, $\Phi = [\Phi_1 \cdots \Phi_{N_d}]$, $W = [W_1 \cdots W_{N_d}]$; one convex ridge problem:

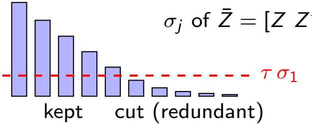
$$\boxed{\hat{\mathcal{K}} = W\Phi^\top(\Phi\Phi^\top + \gamma I)^{-1}}$$

γ tuned on the decoded roll-out (raw or re-lifted: deployment-matched).
Why "SysID" label: estimation and selection by multi-step prediction error on validation data.

Lazar, arXiv:2606.02938, 2026.

Pipeline B: SVD reduction; rank by the roll-out criterion

Kernel/random dictionaries are intentionally oversized: cut by energy, i.e., the data variance σ_j^2 along each SVD direction (state and input lifts separately), then refit:



- Reduced lifts $\tilde{\Psi}_x = U_{z,r_z}^\top \Psi_x$, $\tilde{\Psi}_u = U_{v,r_v}^\top \Psi_u$ preserve the model form;
- rank r: smallest with roll-out error close to the full model, $r_k^x = \|D\hat{Z}_k - DZ_k\|_F / \|DZ_k\|_F$ at $k = N_d$;
- why not energy alone: the roll-out error need not decrease as energy-ranked directions are added; low-energy directions can be dynamically essential, so the rank is validated on roll-out;
- multi-step refit in the reduced coordinates.

Lazar, arXiv:2606.02938, 2026.

Koopman MPC

Prediction with the generic (GeKo) model form; it recovers the bilinear ($\Psi_u = [1, u^\top]^\top$) and linear ($N_j = 0$) cases; prediction horizon H:

Koopman model predictive control problem

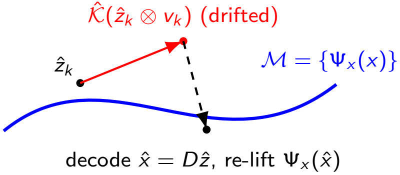
$$\min_{u_0,\ldots,u_{H-1}} \sum_{k=0}^{H-1} \|Dz_k - x_{\mathrm{ref}}\|^2_{Q_x} + \|u_k\|^2_R + \|\Delta u_k\|^2_{R_\Delta} + \|Dz_H - x_{\mathrm{ref}}\|^2_{Q_N}$$

$$\text{s.t.} \quad z_{k+1} = \mathcal{K}\big(z_k \otimes \Psi_u(u_k)\big), \qquad z_0 = \Psi_x(x_t), \qquad \|u_k\|_\infty \le u_{\max}.$$

- solved by SQP with exact (adjoint) gradients, warm-started with the shifted previous solution; linear class $\Rightarrow$ convex QP;
- prediction optionally with re-lift (next slide).

Re-lift: the practical fix for MPC predictions

Without perfect invariance, roll-outs drift off the lifted manifold $\mathcal{M} = \{\Psi_x(x)\}$:



- Restores consistency; complementary to both pipelines (deployment-matched γ: tune on the *re-lifted* roll-out);
- Remark: PAD models minimize drift; hence the re-lift effect should be *less dramatic* by construction.

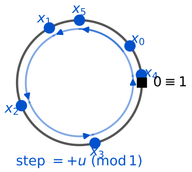
Mauroy & Goncalves, *IEEE TAC*, 2019; van Goor, Mahony, Schaller & Worthmann, *CDC*, 2023; van Goor et al., arXiv:2506.17817, 2025.

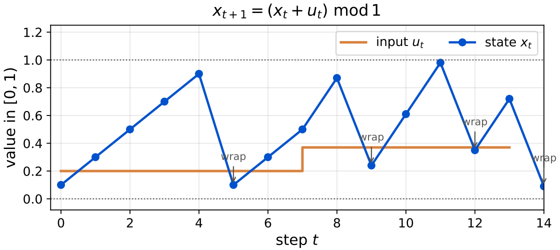
Outline

Example: controlled rotation on the circle

State x on the unit circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (identify $x \sim x+1$); input u rotates it:

$$x_{t+1} = (x_t + u_t) \bmod 1, \qquad x, u \in [0, 1).$$





$\mathbb{T} = \mathbb{R}/\mathbb{Z}$: the circle: points differing by an integer are the *same* point.

$x \bmod 1$: fractional part $x - \lfloor x \rfloor$; passing 1 wraps back to 0.

u: the control: a larger u gives a larger rotation per step.

Classical Koopman example: rotation on the circle; eigenfunctions are the Fourier modes (Budišić, Mohr & Mezić, "Applied Koopmanism", *Chaos* **22**, 047510, 2012, *Example 14*).

Building the Koopman operator

Operator = compose observable with the map

An *observable* φ is a function of the state; evaluate it *after* one step:

$$(C_F\varphi)(x,u) = \varphi \circ F(x,u) = \varphi\big((x+u) \bmod 1\big).$$

The input enters in a nonlinear way

Apply it to a Fourier mode $\varphi(x){=}e^{2\pi imx}$:

$$C_F\, e^{2\pi imx} = e^{2\pi im(x+u)} = {\color{blue}e^{2\pi imx}}\, {\color{red}e^{2\pi imu}}.$$

The mod1 drops out (period 1).

Observables (the lifts)

Same Fourier modes for state and input $(n_z{=}n_v{=}2N{+}1)$:

$$z{=}\Psi_x(x){=}\big(e^{2\pi inx}\big)_{|n|\le N},$$
$$v{=}\Psi_u(u){=}\big(e^{2\pi inu}\big)_{|n|\le N}.$$

Model classes $\mathcal{H}_L \subset \mathcal{H}_B \subset \mathcal{H}$

${\color{blue}\mathcal{K}_L}$	linear	$z^+{=}\mathcal{A}z{+}\mathcal{B}u$
${\color{blue}\mathcal{K}_B}$	bilinear	$z^+{=}\mathcal{A}z{+}\mathcal{B}u{+}\mathcal{N}(z\otimes u)$
${\color{red}\mathcal{K}}$	GeKo	$z^+{=}\mathcal{K}(z\otimes v)$

Ground truth operator $C_F\varphi = \varphi \circ F$

Probe observable and its exact image

One probe with *infinitely many* Fourier modes (Jacobi–Anger), $I_{|m|}(1)$=1.27, 0.57, . . . :

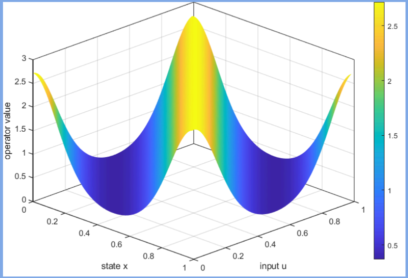
$$\varphi = e^{\cos 2\pi x} = \sum_m I_{|m|}(1)\, e^{2\pi i m x}.$$

C_F acts *mode by mode* (each $\times\, e^{2\pi i m u}$), giving a closed form:

$$C_F\varphi = \sum_m I_{|m|}(1)\, e^{2\pi i m x} e^{2\pi i m u} = e^{\cos 2\pi(x+u)}.$$

Exact ground truth can be computed with *no* truncation.

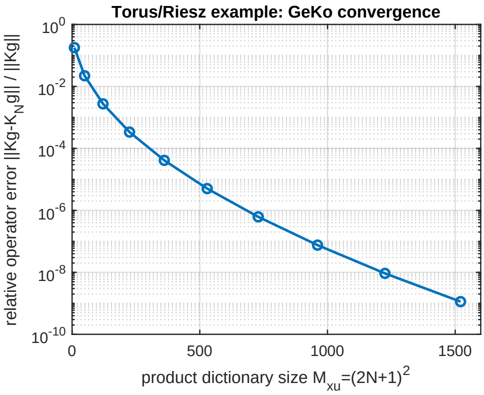
Operator action $C_F\varphi$ on $\mathbb{T}^2 = \mathbb{T}_x \times \mathbb{T}_u$

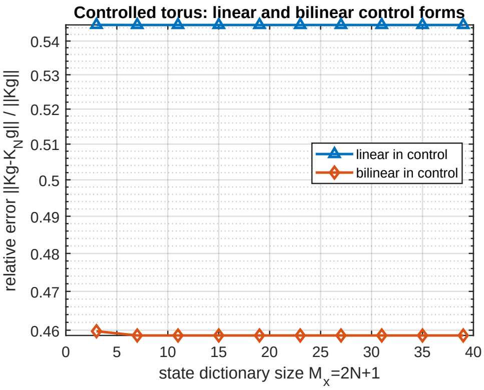


Rotating u slides the profile along x; Koopman models must match the coupled $e^{2\pi i m x} e^{2\pi i m u}$ ridge.

Convergence test: learned $\hat{\mathcal{K}}^{(n_z,n_v)}\varphi$ (modes $|n| \leq N$) vs. exact $C_F\varphi \Rightarrow$ the gap is the truncation tail $\to 0$.

Example: convergence vs. the nested model-class floors

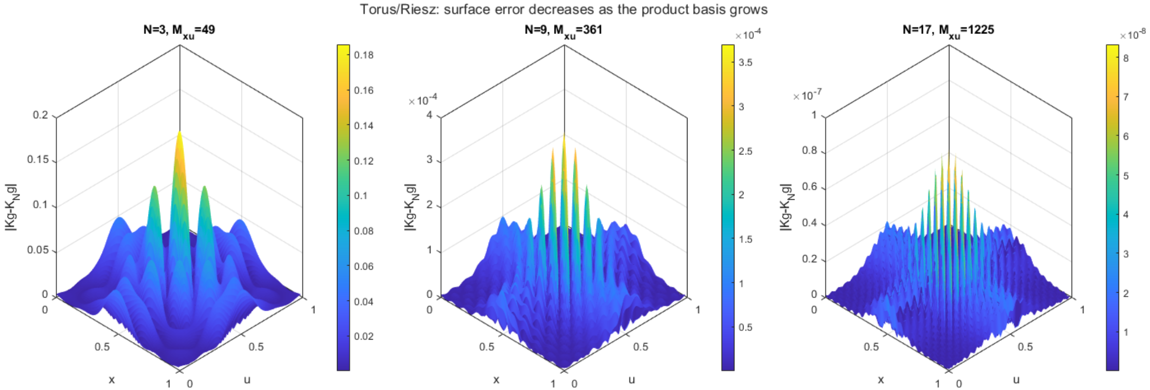




GeKo $\hat{\mathcal{K}}^{(n_z,n_v)} \to \mathcal{K}$: strong (norm) convergence, reaching 10^{-9} at $N=19$. Riesz Gram spectrum stays in $[\alpha, \beta] \approx [0.42, 1.82]$.

Linear/bilinear hit structural floors $\varepsilon_L \approx 0.54 \geq \varepsilon_B \approx 0.46$, *flat in* n_z: a line in u cannot represent the sinusoids $e^{2\pi imu}$.

Example 1: 3D error convergence



Convergence of the error as the number of observables increases in 3D.

Example: forced Duffing oscillator with different input functions

Dynamics

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = (1 - x_1^2)\, x_2 - x_1 - x_1^3 + h(u)$$

Three input maps

Increasing difficulty:

$h(u) = u$ (affine), $\qquad h(u) = u + 0.5u^3$ (cubic, monotone), $\qquad h(u) = \sin u$ (non-monotone!)

Sampling $T_s = 0.05$ s (RK4), $|u| \leq 3$; multisine excitation, 10 training trajectories (incl. small ICs near the origin), one held-out test trajectory.

Setup used throughout: mixed state dictionary (monomials deg $\leq 3 \oplus 30$ RFFs, $n_z = 40$) and monomial input dictionary $\Psi_u(u) = [\,1,\ u,\ u^2,\ u^3\,]^\top$ ($n_v = 4$, incl. the constant).

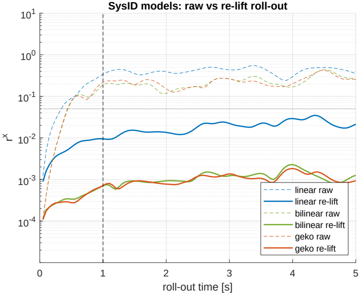
Invariance proximity of the model library

Worst-case one-step defect $I_\infty = \sin\theta_{\max}$, per input map, pipeline and class:

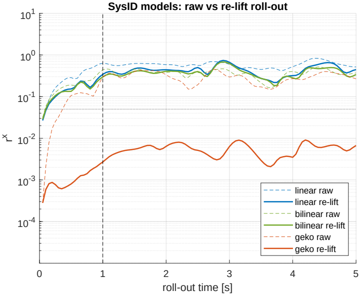
$h(u)$	pipeline	linear	bilinear	GeKo
u	PAD	0.754	0.744	0.406
	SysID	0.551	0.418	0.368
$u + 0.5u^3$	PAD	0.911	0.670	0.417
	SysID	0.756	0.510	0.359
$\sin u$	PAD	0.671	0.580	0.517
	SysID	0.472	0.439	0.400

Note: I_∞ is a conservative index: worst case over *all* image directions, for a fixed, generic dictionary (typical-direction defects: 10^{-3}–10^{-4}); training a neural dictionary *against* invariance can drive $I_\infty \sim 10^{-2}$ (Haseli & Cortés, 2026). Here: comparative use, the ranking, not the absolute value.

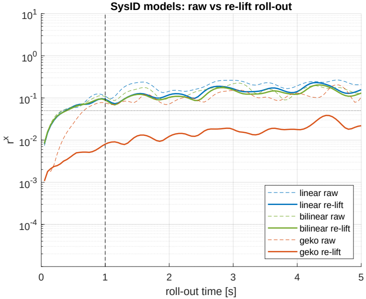
Roll-out without & with re-lift (SysID models)



$h(u) = u$



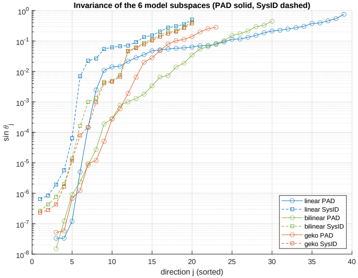
$h(u) = u + 0.5u^3$



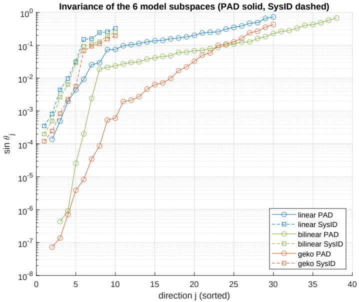
$h(u) = \sin u$

SysID trio, raw (dashed) vs re-lift (solid): re-lift reduces the roll-out error for all input maps (if the Koopman model is structurally valid).

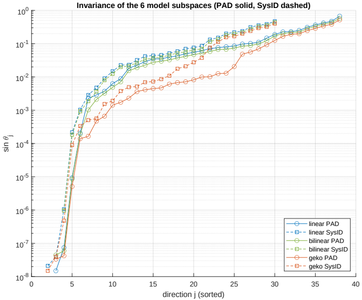
Angle spectra of the model subspaces



$h(u) = u$



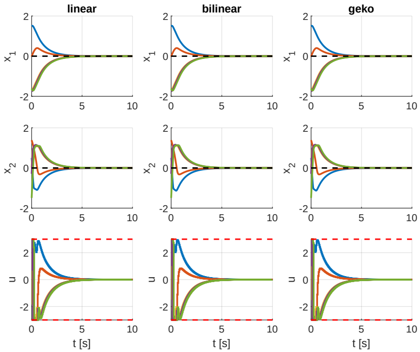
$h(u) = u + 0.5u^3$



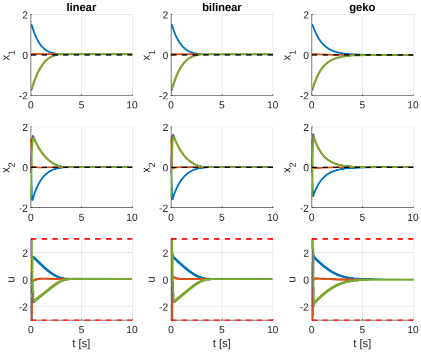
$h(u) = \sin u$

Sorted $\sin \theta_j$ of the 6 model subspaces (PAD solid, SysID dashed): near-invariant in most directions; a few hard directions set I_∞.

Koopman MPC ($H = 15$), SysID models



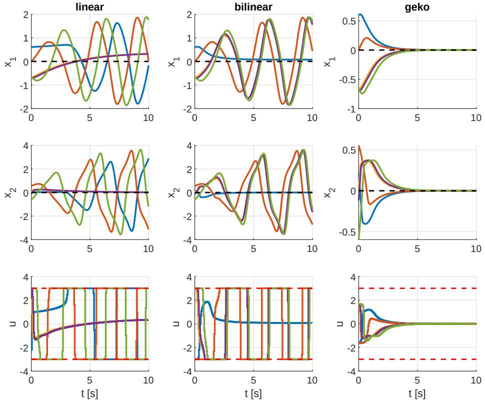
$h(u) = u$



$h(u) = u + 0.5u^3$

Note: affine and cubic (monotone) actuation: all three classes stabilize; an average fit survives.

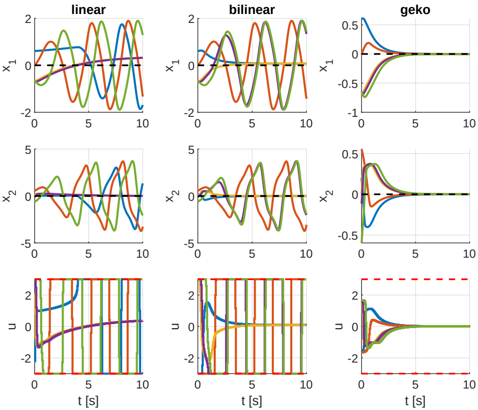
Koopman MPC ($H = 15$), the non-monotone case



$h(u) = \sin u$, SysID models.

Note: non-monotone actuation: an average fit is not sufficient; only GeKo stabilizes all ICs.

Closed-loop model falsification



Closed-loop test: each model's MPC is run with the trusted GeKo model + re-lift as the plant; under sin u, linear and bilinear fail on the surrogate too: the library falsifies its own members, plant-free.

Consistency analysis over different/larger dictionaries

Same experiment, $h(u) = u + 0.5u^3$, three state dictionaries (Chebyshev tensor, deg ≤ 9: Conradie et al., arXiv:2603.15091, 2026):

dictionary (n_z)	pipeline	linear	bilinear	GeKo
mixed (40)	PAD	0.911	0.670	0.417
	SysID	0.756	0.510	0.359
RFF (103)	PAD	0.990	0.584	0.459
	SysID	0.555	0.439	0.363
Chebyshev (100)	PAD	0.986	0.569	0.437
	SysID	0.999	0.793	0.536

Note: the ranking is dictionary-robust: GeKo attains the smallest I_∞ in every entry, across all three dictionaries, input maps and pipelines; richer dictionaries drive the linear class to the ceiling $I_\infty \approx 1$.

Concluding remarks

- Convergence first: the GeKo product-basis construction yields a provably convergent Koopman representation ($\mathcal{K}^\top = W^{-1} C_F W_x$); its predicted advantage is observed: best I_∞ and roll-out error across both pruning pipelines;

- Two complementary measures: invariance proximity I_∞ certifies, *a priori*, the worst-case defect of a model class; the true-state roll-out measures deployment error; use both for model selection;

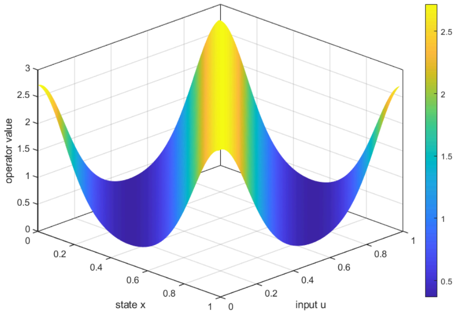
- An empirical link inviting theory: multi-step regression with SVD reduction, tuned only on the roll-out error, implicitly promotes invariance;

- Always re-lift in MPC: restoring lifted-state consistency reduces the roll-out error over the prediction horizons (for a valid Koopman model class);

- Open problems: open-loop scores cannot certify closed-loop viability; validate in closed-loop against the most trusted (convergent) model; less conservative invariance indexes.

Thank you for your attention!

Koopman operator theory for systems with control



The generalized Koopman operator $C_F\varphi = \varphi \circ F(x,u)$.

Supporting papers: generalized Koopman arXiv:2508.07494; multi-step regression arXiv:2606.02938.

Questions?