When Linear Models Aren't Enough: Forcing as Closure in Chaotic Time Series

Yuer Tang, Justin Baker
yuertang17@g.ucla.edu

Abstract

Koopman theory lifts nonlinear dynamics into a linear but infinite-dimensional space of observables. Any data-driven implementation requires finite-dimensional truncation, where linear closure can fail. HAVOK models chaotic systems from a single scalar time series using delay embeddings, producing a linear system with a sparse forcing term. While effective in practice, the mathematical origin of this forcing has remained unclear.

Key result. We show that the HAVOK forcing term arises from the failure of Koopman invariance under finite-dimensional truncation, and acts as a closure correction for unresolved dynamics.

Koopman: why chaos can look linear

Let M denote the state space and $\Phi^t : M \to M$ the flow map of a dynamical system. Instead of tracking states, Koopman theory studies observables

$$\mathcal{G} = \{g : M \to \mathbb{C}\},$$

equipped with a Hilbert space structure.
The Koopman operator $K^t : \mathcal{G} \to \mathcal{G}$ acts by composition,

$$(K^t g)(x) = g(\Phi^t(x)).$$

Although the state evolution Φ^t may be nonlinear or chaotic, K^t is linear in g and satisfies the semigroup property

$$K^{t+s} = K^t K^s.$$

Key idea. Nonlinear dynamics in state space correspond to linear dynamics in an infinite-dimensional space of observables.

Finite Observations and the Closure Problem In practice, only finitely many observables can be measured or learned from data. Let

$$\mathcal{V}_r = \operatorname{span}\{v_1, \dots, v_r\} \subset \mathcal{G}$$

denote a finite-dimensional approximation.

Linear closure requires Koopman invariance:

$$K\mathcal{V}_r \subset \mathcal{V}_r, \qquad Kv_i = \sum_{j=1}^{r} a_{ij} v_j.$$

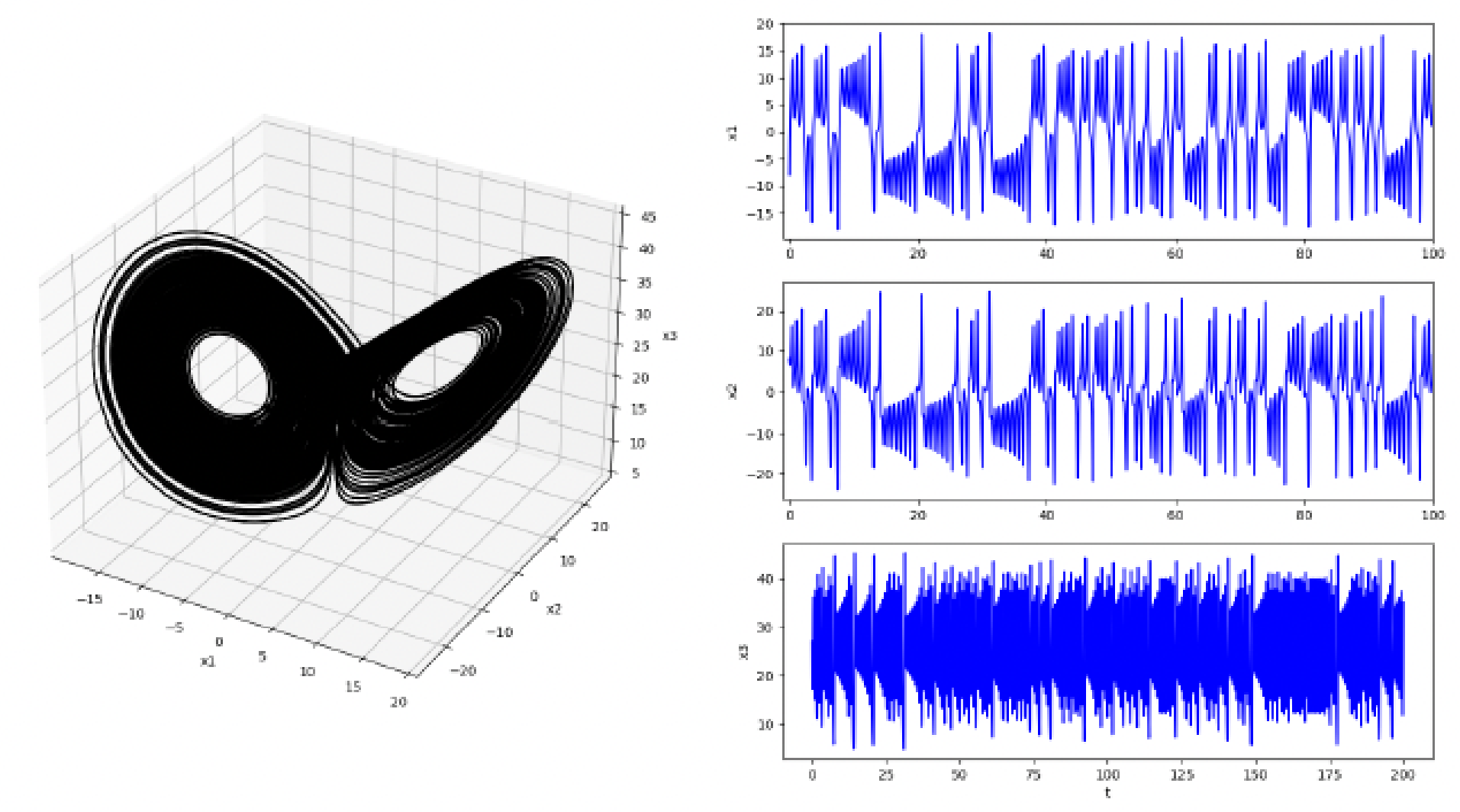
However, because the Koopman operator acts on an infinite-dimensional space, finite truncations are generically *not* invariant:

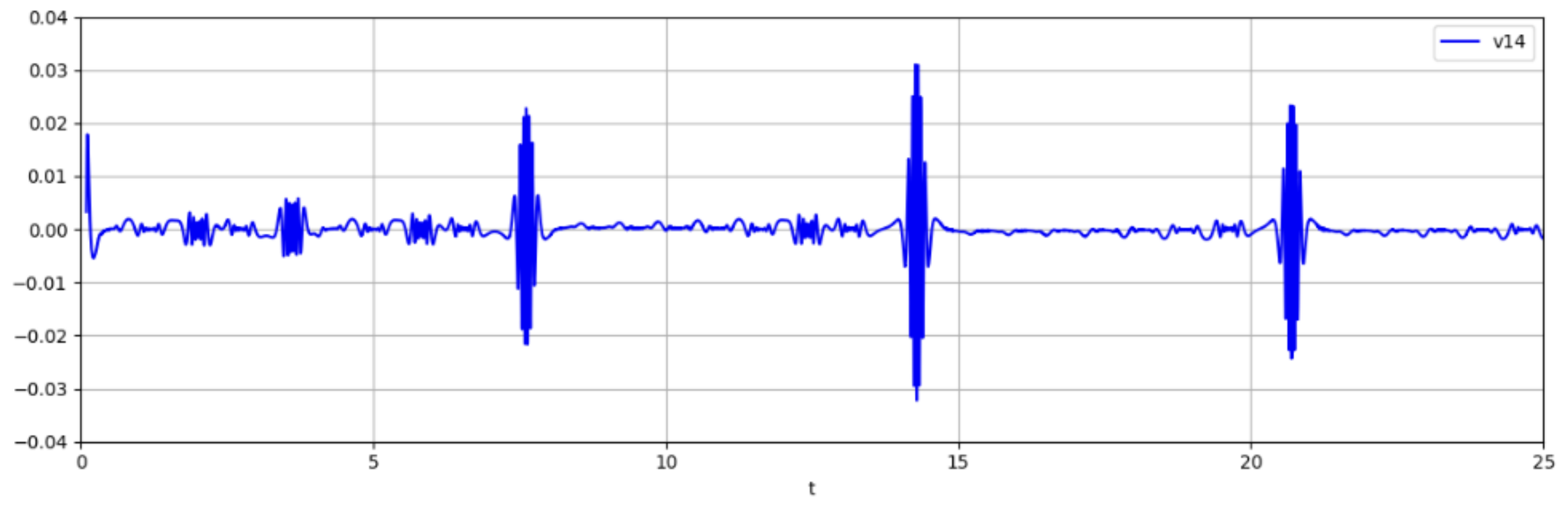
$$K\mathcal{V}_r \not\subset \mathcal{V}_r.$$

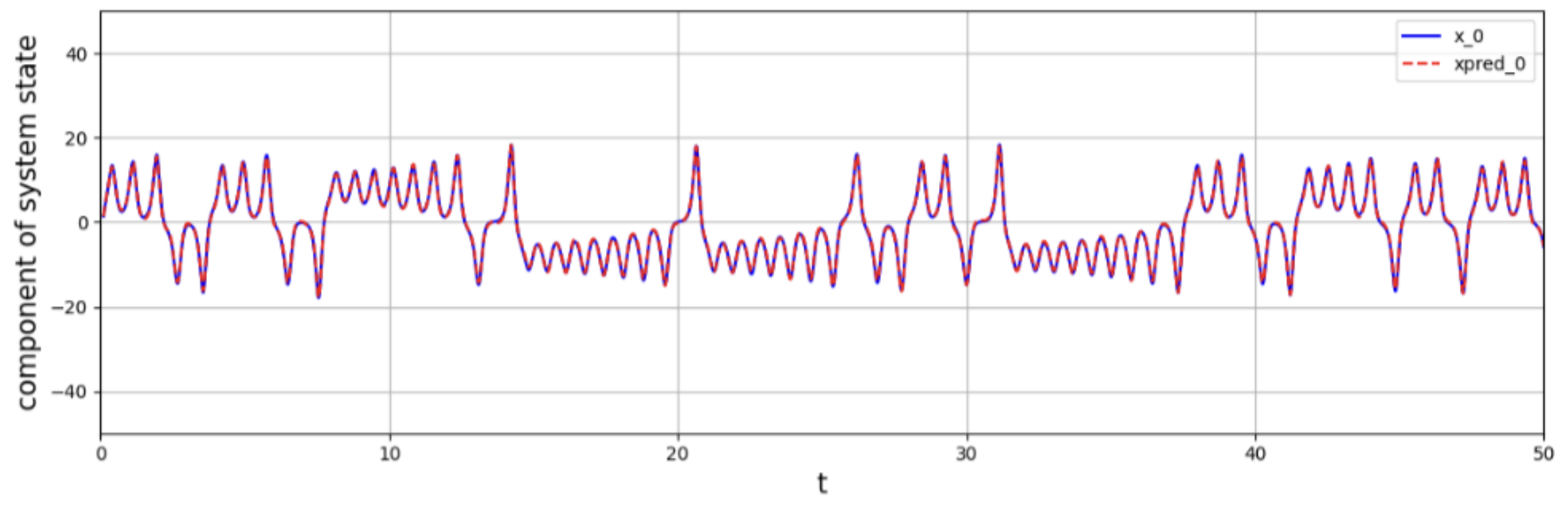
As a result, purely linear reduced models fail to capture the full dynamics.

Toy Examples: Lorenz and Rössler Systems

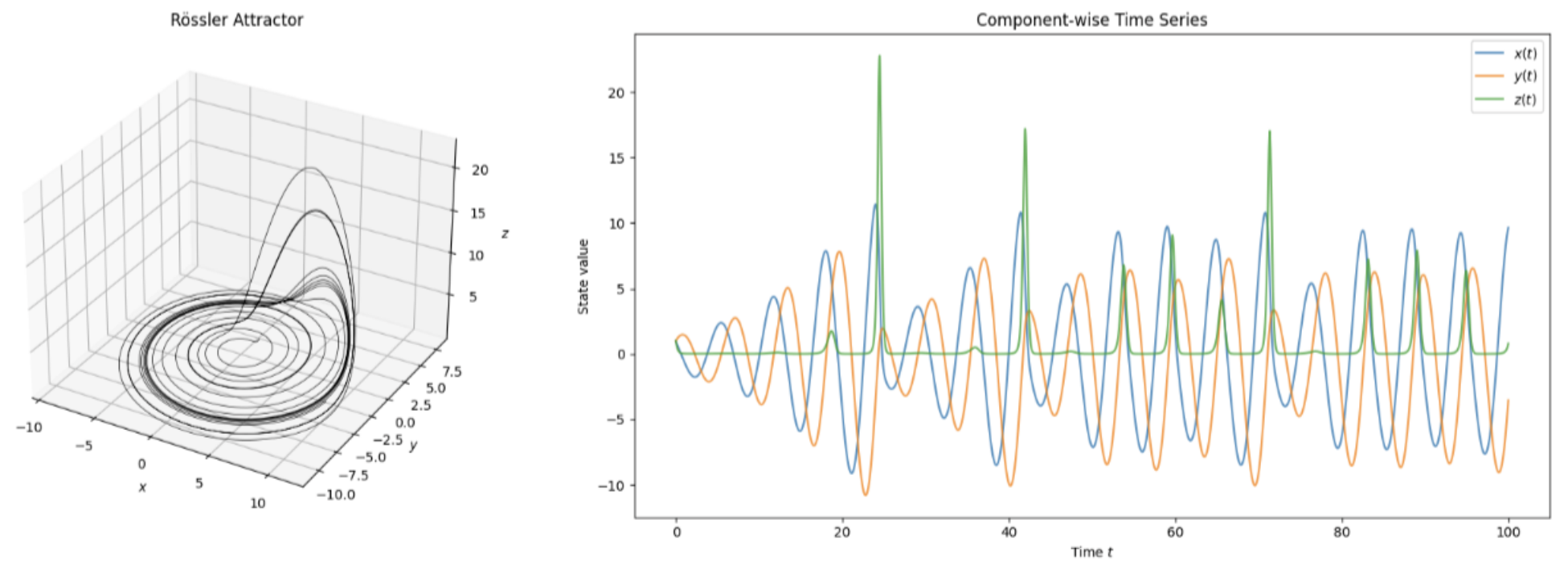
Lorenz system.

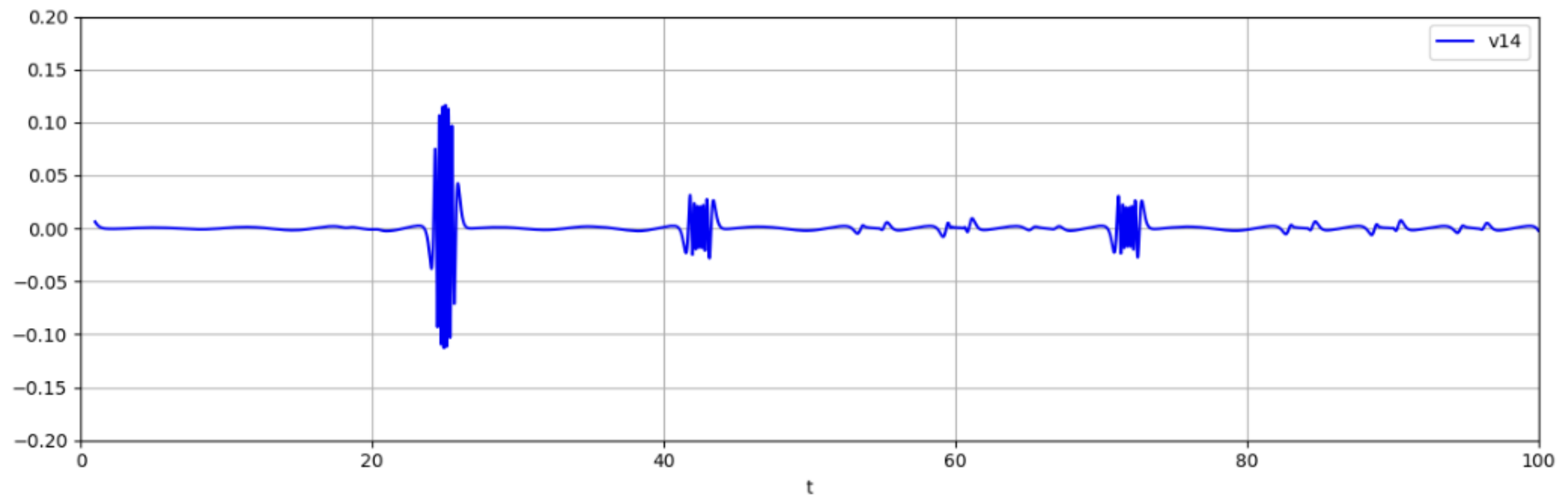


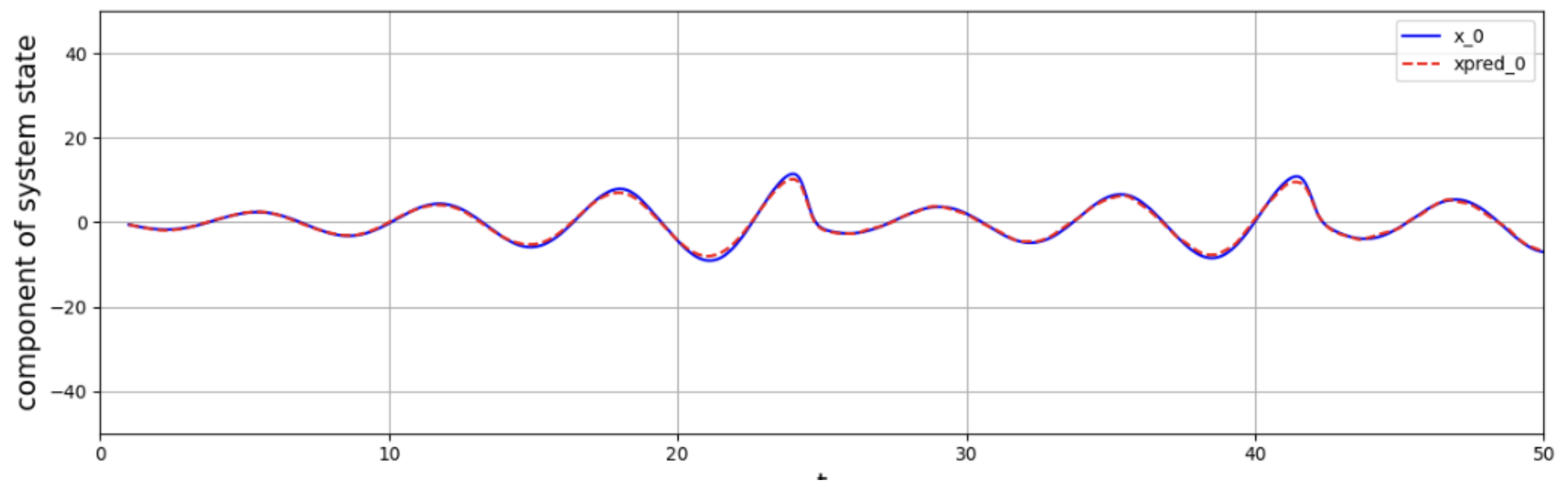




Rössler system.







HAVOK: Linear Modeling from a Single Time Series

HAVOK starts from a single scalar measurement

$$x(t) = g(\mathbf{x}(t)),$$

and constructs delay-coordinate vectors

$$\mathbf{x}(t) = (x(t),\, x(t-\tau),\, \dots,\, x(t-(m-1)\tau)).$$

Stacking these vectors produces a Hankel matrix

$$H = \begin{bmatrix} x(t_1) & x(t_2) & \cdots & x(t_p) \\ x(t_2) & x(t_3) & \cdots & x(t_{p+1}) \\ \vdots & \vdots & \ddots & \vdots \\ x(t_q) & x(t_{q+1}) & \cdots & x(t_m) \end{bmatrix}.$$

Applying SVD,

$$H = U\Sigma V^*,$$

and truncating to rank r yields dominant, data-driven coordinates that capture the leading temporal structure.

$$H \approx U_r \Sigma_r V_r^*,$$

Hankel Matrix as Approximate Koopman Evolution

$$H = \begin{bmatrix} x(t_1) & Kx(t_1) & \cdots & K^{p-1}x(t_1) \\ Kx(t_1) & K^2x(t_1) & \cdots & K^p x(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ K^{q-1}x(t_1) & K^q x(t_1) & \cdots & K^{m-1}x(t_1) \end{bmatrix}.$$

Because this subspace is incomplete, it is generally not invariant under the action of $\mathcal{K}$. Specifically, for dominant modes $v_i \in \mathcal{V}_r$, one typically has

$$\mathcal{K}v_i \notin \operatorname{span}\{v_1, \dots, v_r\},$$

HAVOK Forcing as Koopman Closure

Let $v(t)$ denote the resolved HAVOK coordinates obtained from truncated SVD. Projecting Koopman evolution onto this subspace yields

$$\dot{v}(t) = Av(t) + Bw(t),$$

where $w(t)$ captures the component generated by Koopman evolution outside the resolved subspace.