

# A multichannel generalization of the HAVOK method for the analysis of nonlinear dynamical systems

Carlos Colchero,<sup>\*</sup> Jorge Perez,<sup>†</sup> Alvaro Herrera,<sup>‡</sup> and Oliver Probst<sup>§</sup>

*School of Engineering and Sciences, Tecnológico de Monterrey,  
Av. Eugenio Garza Sada 2051 Sur, Monterrey CP 64700, Mexico*

(Dated: September 18, 2025)

By extending Takens' embedding theorem (1981), Deyle and Sugihara (2011) provided a theoretical justification for using parallel measurement time series to reconstruct a system's attractor. Building on Takens' framework, Brunton et al. (2017) introduced the Hankel alternative view of Koopman (HAVOK) algorithm, a data-driven approach capable of linearizing chaotic systems through delay embeddings. In this work, a modified version of the original algorithm is presented (*m*HAVOK), a practical realization of Deyle and Sugihara's generalized embedding theory. *m*HAVOK extends the original algorithm from one to multiple input time series and introduces a systematic approach to separating linear and nonlinear terms. An  $R^2$ -informed quality score is introduced and shown to be a reliable guide for the selection of the reduced rank. The algorithm is tested on the familiar Lorenz system, as well as the more sophisticated Sprott system, which features different behaviors depending on the initial conditions. The quality of the reconstructions is assessed with the Chamfer distance, validating how *m*HAVOK allows for a more accurate reconstruction of the system dynamics. The new methodology generalizes HAVOK by allowing the analysis of multivariate time series, fundamental in real life data-driven applications.

## GLOSSARY

HAVOK	Hankel Alternative view of Koopman	$\mathbf{a}_j$	$j$ -th column of $\mathbf{A}$
<i>m</i> HAVOK	multiple Hankel Alternative view of Koopman	$\bar{\mathbf{v}}_i$	temporal mean of $\dot{\mathbf{v}}_i$
SVD	Singular Value Decomposition	$\hat{\mathbf{A}}_r$	system matrix
$\{\mathbf{x}_n\}_{n=0}^{G-1}$	$G$ state variables	$\mathbf{B}_r$	input matrix
$M$	embedding dimension	$\mathbf{u}(t)$	$\mathbf{V}_f$
$f(\mathbf{x}_n)$	observable	$\mathbf{C}$	output matrix
$\Psi[n]$	discrete delay vector	$\mathbf{D}$	zero matrix
$\Psi_{GL}$	multichannel delay vector	$\mathbf{x}(t)$	simulated state vector
$\tau_s$	user-defined delay	$\mathbf{y}(t)$	simulated output
$\mathbf{h}_k$	block of observables	$\mathbf{x}(0)$	first row of $\mathbf{V}_c$
$\mathbf{H}$	block Hankel matrix	$N_s$	simulated time steps
$N$	size of the time series	$\hat{\mathbf{H}}(t)$	simulated Hankel matrix
$Q$	number of channels	$\mathbf{U}_c$	core left singular vectors
$l$	$\min(QM, N - M - 1)$	$\Sigma_c$	diagonal matrix of core singular values
$\mathbf{U}_l$	left singular vectors	$.(tr)$	Matrix employed on the training set
$\mathbf{S}_l$	diagonal matrix of singular values	$.(ev)$	Matrix employed on the evaluation set
$\mathcal{V}_l^T$	eigen-time-delay coordinates	$\sigma_k$	$k$ -th singular value
$r$	cutoff rank of $\mathcal{V}_l^T$	$q$	percentile cutoff
$\mathbf{V}$	matrix with first $r$ columns of $\mathcal{V}_l^T$	$\mathcal{Q}(r)$	quality score
$\tau$	user-defined threshold	$C$	subset of rank-values
$\dot{\mathbf{V}}$	derivative of the dynamical modes	$\kappa(r)$	condition number of $\mathbf{B}_r$
$\dot{\mathbf{v}}_j$	$j$ -th column of $\dot{\mathbf{V}}$	$R_q^2$	$R^2$ between the $q$ -th channel of the original and the reconstructed system
$I_c$	index set of core dimensions	$\overline{R_{rec}^2}(r)$	mean of $R_q^2$
$I_f$	index set of forcing dimensions	$\mathbb{V}$	explained variance
$\mathbf{V}_c$	$\mathbf{V}$ matrix with the core dimensions	$T$	simulated time
$\mathbf{V}_f$	$\mathbf{V}$ matrix with the forcing dimensions	$\Delta t$	time step
$\mathbf{A}_r$	Matrix describing the evolution of core dimensions	$\Omega$	Set of subsampled points
		$\zeta$	Set of constructed points
		$d_c(\Omega, \zeta)$	Chamfer distance between $\Omega$ and $\zeta$

\* A00835489@tec.mx

† A00835537@tec.mx

‡ A00836274@tec.mx

§ oprobst@tec.mx

## 1. Introduction

The linearization of chaotic dynamical systems has attracted considerable attention in recent years [1–5]. A great deal of this work can be traced back to a seminal paper by Bernard Koopman in 1931 [6], who proposed a Hilbert space description of classical mechanics, enabling the analysis of nonlinear systems through a linear operator. This operator, later known as the Koopman operator, acts on functions of the state space, also known as observables, by evolving them linearly in time, even when the system’s evolution is nonlinear [7]. With the rise of modern data analysis, particularly for data-driven analysis of dynamical systems, Koopman’s theory has seen a new surge in popularity [8–11].

A number of works [12–18] have been sparked by pioneering work conducted by Igor Mezic and co-workers on the spectral analysis of the Koopman operator. Mezic demonstrated that, under ergodic conditions, the Koopman operator is unitary, allowing for the decomposition of its spectrum into a discrete and a continuous part [19]. The discrete part is associated with an almost-periodic and recurrent behavior, such as rotations on a torus, periodic orbits or coherent structures [20]. On the other hand, the continuous spectrum captures dispersive and mixing behavior exhibited by chaotic attractors or ergodic systems that lack quasi-periodic components [21]. In general, such dynamical systems exhibit a mixture of both discrete and continuous spectral components.

Building on this perspective, Brunton et al. [22] formulated the Hankel alternative view of Koopman (HAVOK) algorithm, a data-driven approach capable of linearizing chaotic systems. While not explicitly based on Koopman spectral theory, HAVOK reflects a similar approach by decomposing the dynamics into recurrent, structured are associated with the discrete spectrum, and irregular components that resemble continuous or mixed spectral behavior. In this algorithm, a delay embedding is constructed from a single time series and arranged into a Hankel matrix. Theoretically, this construction generates a Krylov subspace of observables [23], which approximates the action of the Koopman operator. Then, a Singular Value Decomposition (SVD) is performed, thereby identifying dynamical modes within the embedding. Only the leading dynamical modes are selected for further processing, requiring the definition of a cutoff rank  $r$ . The most suitable value of  $r$  was determined empirically. The last remaining mode is interpreted as a forcing component driving the chaotic behavior of the system. Subsequently, a linear regression between the time derivatives of the first  $r-1$  dynamical modes and the modes themselves is performed, leading to the construction of a dynamical matrix. Addition-

ally, a second regression is carried out between the same derivatives and the last mode, allowing for the identification of a forcing matrix. Finally, both the dynamical and forcing matrices are combined into a linear system, from which time series can be simulated.

Although the algorithm was shown to successfully reconstruct several chaotic attractors, a number of limitations have become apparent [24–26]. First and foremost, HAVOK is limited to single-input single-output systems, precluding the use of multiple measurement time series. The algorithm also assumes that only a single nonlinear term provides the forcing of the system, which is taken as the last column of the reduced rank embedding, but no objective criterion for this separation approach was offered. Given the selection of the cutoff rank also remained arbitrary, it was unclear if HAVOK can be used for systems with unknown dynamics. Furthermore, another somewhat more subtle issue with HAVOK is the lack of a clear instruction for inverting the output time series from the eigen-time-delay-coordinate frame back to the original coordinates.

The use of a single time series for the reconstruction of the system dynamics in HAVOK is based on Takens’ well-known embedding theorem [27]. Building on earlier work by Packard et al. [28] and extended to fractal sets by Sauer et al. [29], this theorem guarantees a diffeomorphic reconstruction of an attractor from a single observable given enough data points. Nevertheless, there may be practical limitations to this approach. Deyle and Sugihara [30] proved generalized embedding theorems, showing that state-space reconstruction can be improved by including multiple measurement time series, thereby generalizing Takens’ original result. Importantly, they proved that the reconstructed system remains diffeomorphic to the original attractor. This approach reduces the number of delays required from any time series, making the reconstruction less reliant on a single source of information. Such improvement is crucial, since Letellier and Aguirre [31] showed that if an observable remains invariant under a symmetry transformation, it will fail to unfold the system’s topology when an embedding is performed on it, leading to a symmetry-blind reconstruction. Hence, single-input methods like HAVOK will fail to reconstruct the system if the observable is not properly selected.

In this work, the limitations of the HAVOK method pointed out above are addressed. Firstly, the method is extended to allow for multiple measurement time series to be processed together in block-Hankel matrices, providing a practical implementation of Deyle and Sugihara’s work [30]. This extension will be referred to as  $m$ HAVOK. The convenience of using multichannel information, albeit in the context of Dynamical Mode Decomposition (DMD), was already demonstrated by

Arbabi and Mezic [23], and the use of block-Hankel matrices was proposed as early as 1985 by Juang and Pappa [32] for modal parameter estimation and model reduction. However, to the best knowledge of the authors, no multichannel version of the HAVOK algorithm has been presented so far. It will be shown that *mHAVOK* allows for an improved reconstruction of the system dynamics, preventing symmetry-blindness reconstructions. Furthermore, a natural way of tracing the system trajectories in the original coordinate space is included, as well as an objective criterion for selecting the cutoff rank of the leading dynamical modes.

Regarding the second limitation of the original HAVOK method, i.e., the selection of nonlinear components, a systematic approach to solving this problem is also included in *mHAVOK*. Instead of assuming the last dynamical mode from the reduced rank SVD must be classified as nonlinear, *mHAVOK* performs a regression-based technique in order to allow for the identification of nonlinear terms. It will be shown that, in general, several nonlinear terms emerge and their inclusion is critical for an optimal reconstruction of the system dynamics. Finally, the quality of reconstruction is assessed quantitatively using similarity metrics as opposed to previous work [22], where only qualitative comparisons were provided.

It should be noted that *mHAVOK*, very much like HAVOK, has been designed as a technique capable of reconstructing the dynamics of (generally nonlinear) systems from observed data alone, i.e., as a data-driven technique. However, in order to validate the accuracy of the method, reference cases are required. In this work, time series of observables obtained from two test systems, the Lorenz and the Sprott system, determined from numerical solutions of the constitutive equations, were used as the input to *mHAVOK*. The Lorenz system, a widely used reference case, is used for the illustration of some basic insights and for comparison with the original (1D) HAVOK method. The more challenging Sprott system, on the other hand, allows the new method proposed in this work to demonstrate its full potential, which is why this study case is explored in a somewhat greater detail.

The paper is organized as follows: First, Sections 2A–2C explain the extension of the HAVOK method to multiple inputs, present a novel component-classification scheme for the dynamical modes, and describe the procedure for reconstructing the system in the original coordinate system. Next, Sections 2D–2F provide an objective criterion for selecting the cutoff rank  $r$ , explain how the accuracy of the system's reconstruction is quantitatively assessed, and show the simulation setup for the two tested systems: Lorenz and Sprott. Then, Sections 3 and 4 present the simulation

results, showing how the aforementioned modifications remarkably improve state-space reconstruction, particularly in the challenging Sprott system. Finally, Section 5 summarizes the work by presenting the main contributions, limitations found, and future research lines.

## 2. Methods and simulation setup

### A. Generalized Embedding

Consider a discrete trajectory of unobserved state variables  $\{\mathbf{x}_n\}_{n=0}^{G-1}$  evolving on a compact state space  $\mathcal{M}$ . Given a smooth function of the state variables  $f : \mathcal{M} \rightarrow \mathbb{R}$ , the discrete delay vector of length  $M$  and delay  $\tau_s \in \mathbb{N}$  is defined by

$$\Psi[n] = (f(\mathbf{x}_n), f(\mathbf{x}_{n+\tau_s}), \dots, f(\mathbf{x}_{n+[M-1]\tau_s})). \quad (1)$$

In the literature [18–21],  $f(\mathbf{x}_n)$  is referred to as an observable. Under Takens' hypotheses, the delay map in Eq. (1) is an embedding for generic observables, so that  $\Psi(\mathcal{M})$  is diffeomorphic to the original system. In HAVOK [22], columns of such delay vectors are stacked to form a Hankel matrix with uniform delays.

An observable  $f$  is defined as symmetry-blind with respect to a system symmetry  $\mathcal{G}$  if  $f(\mathbf{x}_n) = f(\mathcal{G} \cdot \mathbf{x}_n) \forall \mathbf{x}_n \in \mathcal{M}$ . If an observable satisfies this condition, it will fail to unfold the system's topology when an embedding is performed on it [31].

The previously discussed single-channel limitations can be avoided by extending the delay map to multiple observables, an extension provided by Deyle and Sugihara's generalized embedding theorems [30]. Consider  $Q$  observables denoted by  $\{f_q\}_{q=0}^{Q-1}$ . In each of them, a multichannel delay vector of  $M$  delays is collected as follows:

$$\Psi_{GL}[n] = (f_0(\mathbf{x}_{n+k\tau_s}), \dots, f_{Q-1}(\mathbf{x}_{n+k\tau_s}))_{k=0}^{M-1}, \quad (2)$$

where  $\Psi_{GL}$  has the length  $MQ$ . For a fixed embedding dimension  $D$ , one may trade depth and channels so that  $D = MQ$ , an improvement from the single source delay, where  $D = M$ . With this considered, a block Hankel matrix is formed considering  $\tau_s = 1$ , similar to the original formulation by Brunton et al. [22]. For each delay  $k = \{0, \dots, N-1\}$ , the stacked observable block is given by:

$$\mathbf{h}_k := \begin{bmatrix} f_0(\mathbf{x}_k) \\ f_1(\mathbf{x}_k) \\ \vdots \\ f_{Q-1}(\mathbf{x}_k) \end{bmatrix} \in \mathbb{R}^Q. \quad (3)$$

Arranging these blocks into the Hankel matrix yields:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_{N-M} \\ \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_{N-M+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{M-1} & \mathbf{h}_M & \cdots & \mathbf{h}_{N-1} \end{bmatrix}, \quad (4)$$

where  $N$  is the number of time measurements, i.e., the size of the time series. Each column within the block Hankel matrix constitutes a time-delay vector  $\Psi[n]$ , such as the one presented in Eq. (1).

The array in Eq. (4) extends the delay embedding to several input dimensions, with the matrix size being  $(QM, N - M + 1)$ . This embedding allows for several spatiotemporally correlated channels to coexist, potentially enriching the system's reconstruction by reducing noise correlations occurring in one-dimensional embeddings [30].

Having constructed the block Hankel matrix, a thin Singular Value Decomposition (SVD) is performed [33],

$$\mathbf{H} = \mathbf{U}_l \mathbf{S}_l \mathcal{V}_l^T, \quad (5)$$

where  $l = \min(QM, N - M + 1)$ . Similar to the original HAVOK method [22], this decomposition identifies an orthonormal basis for the dynamical modes of the system. In practical implementations, singular values are usually ordered from largest to smallest, as well as their associated vectors in  $\mathcal{V}_l$  and  $\mathbf{U}_l$ . This ordering affects the indexing of the SVD modes, but it does not permute the rows or columns of the original matrix  $\mathbf{H}$ . Therefore, channel ordering and block positions in the original matrix are preserved after an SVD, regardless of the sorting procedure.

$\mathbf{U}_l$  is an orthonormal matrix sized  $(QM, QM)$ , whose columns form a basis for the column space of  $\mathbf{H}$ . These columns are often referred to as spatial or structural modes, since they capture coherent patterns across data rows. In the context of Hankel matrices, each row is a shifted time window, so the spatial modes capture coherent structures within different time windows. This motivates the use of multiple channels in the embedding process, revealing coherent patterns across different measurement channels. Accordingly, using a block

array enhances the reconstruction quality and better captures the global dynamics of the system.

On the other hand,  $\mathcal{V}_l^T$  is a rectangular matrix sized  $(QM, N - M + 1)$ , whose rows contain the time coefficients of the structural modes in  $\mathbf{U}_l$ . For that reason, the rows of  $\mathbf{S}_l \mathcal{V}_l^T$  are often referred to as the eigen-time-delayed coordinates, since they serve as a transformed time series corresponding to the primary structural modes in the Hankel matrix.

## B. Linear regression in embedded space

Having constructed the generalized embedding, the eigen-time-delay coordinates are truncated by retaining the first  $r$  columns of  $\mathcal{V}_l$ . Let  $\mathbf{V}$  represent the reduced rank matrix.  $r$  plays a fundamental role in the original HAVOK model, since the  $r$ -th column of  $\mathbf{V}$  is assumed to be the forcing term of the linearized model [22]. Given that the appropriate value of  $r$  is not known a priori, a systematic approach is required.

To address these questions, each of the columns in  $\mathbf{V}$  are classified based on a goodness-of-fit criterion, allowing for the identification of nonlinear components. As will be shown below, the number of nonlinear terms is generally not limited to one, and the inclusion of all nonlinear components meeting an appropriate selection criterion is critical to the correct reconstruction of the system dynamics.

A first regression problem is solved to separate linear and nonlinear components at a fixed rank  $r$ . The first step involves the calculation of the derivative of the eigen-time-delay coordinates  $\dot{\mathbf{V}}$  using a high-order central difference, which enhances numerical accuracy and reduces sensitivity to noise. Then, a linear regression is performed in an attempt to find  $\mathbf{A}_r$ , i.e., the matrix that best describes the changes in time of the eigen-time-delay coordinates:

$$\min_{\mathbf{A}_r} \|\dot{\mathbf{V}} - \mathbf{V} \mathbf{A}_r^T\|_F, \quad (6)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. Therefore, each column of the time derivative matrix  $\dot{\mathbf{V}}$  is approximated as

$$\dot{\mathbf{v}}_j \approx \mathbf{V} \mathbf{a}_j, \quad (7)$$

where  $\dot{\mathbf{v}}_j$  and  $\mathbf{a}_j$  are the  $j$ -th columns of  $\dot{\mathbf{V}}$  and  $\mathbf{A}_r$ , respectively. Next, the coefficient of determination  $R^2$  is calculated for each of the regressed columns,

$$R_i^2 = 1 - \frac{\|\dot{\mathbf{v}}_i - \mathbf{V}\mathbf{a}_i\|^2}{\|\dot{\mathbf{v}}_i - \bar{v}_i \cdot \mathbf{1}\|^2}, \quad (8)$$

where  $\|\cdot\|$  denotes the Euclidean norm in  $L^2$ ,  $\bar{v}_i$  is the temporal mean of  $\dot{\mathbf{v}}_i$  and  $\mathbf{1}$  is a vector of ones. Linear and nonlinear components are then classified by their goodness of fit:

$$\mathcal{I}_c = \{i : R_i^2 \geq \tau\}, \quad (9)$$

$$\mathcal{I}_f = \{1, \dots, r\} \setminus \mathcal{I}_c, \quad (10)$$

where  $\tau$  is a user-defined threshold,  $\mathcal{I}_c$  is the index set of core dimensions obtained with the given threshold and  $\mathcal{I}_f$  is the index set for the identified nonlinear dimensions. Evidently,

$$r_c + r_f := |\mathcal{I}_c| + |\mathcal{I}_f| = r, \quad (11)$$

where  $|\cdot|$  denotes the cardinality of the indexed sets. Since the two sets are disjoint, the cardinality of their union equals  $r$ . At this point,  $r$  continues to be an arbitrary-fixed value.

Next, the temporal coordinate matrix  $\mathbf{V}$  is redefined to only contain the core dimensions identified in the first regression stage:

$$\mathbf{V}_c := [\mathbf{v}_i]_{i \in \mathcal{I}_c}, \quad (12)$$

where  $\mathbf{v}_i$  is the  $i$ -th column of  $\mathbf{V}$ . Therefore,  $\mathbf{V}_c$  only retains the columns that display a linear evolution. Similarly, it is possible to define a time series matrix that corresponds to nonlinear components:

$$\mathbf{V}_f := [\mathbf{v}_i]_{i \in \mathcal{I}_f}. \quad (13)$$

To describe the time evolution of core dimensions (which show linear behavior), a second regression problem is defined. The objective is to obtain two matrices, one for the evolution of nonlinear terms,  $\mathbf{B}_r$ , and the other one for the evolution of core dimensions,  $\hat{\mathbf{A}}_r$ :

$$\min_{\hat{\mathbf{A}}_r, \mathbf{B}_r} \|\dot{\mathbf{V}}_c - \mathbf{V}_c \hat{\mathbf{A}}_r^T - \mathbf{V}_f \mathbf{B}_r^T\|_F. \quad (14)$$

Subsequently, to simulate embedding dynamics, i.e., the dynamics of the eigen-time-delay coordinates, the

retrieved system matrices are provided as input to the `lsim` function from Python's SciPy library, which simulates the following linear system in the familiar state-space form:

$$\dot{\mathbf{x}} = \hat{\mathbf{A}}_r \mathbf{x} + \mathbf{B}_r \mathbf{u}, \quad (15)$$

$$\mathbf{y} = \mathbf{C} \mathbf{x} + \mathbf{D} \mathbf{u}, \quad (16)$$

where  $\mathbf{u} := \mathbf{V}_f$ ,  $\mathbf{C}$  is an  $r_c \times r_c$  identity matrix,  $\mathbf{D}$  is a zero matrix,  $\mathbf{x}$  is the simulated state vector and  $\mathbf{y}$  is the simulated output. The initial condition  $\mathbf{x}(0)$  is a column vector given by the first row of  $\mathbf{V}_c$ , i.e.,

$$\mathbf{x}(0) = \mathbf{V}_c(0, :). \quad (17)$$

With this procedure, the core eigen-time-delay coordinates are simulated by integrating the linear dynamical system defined by the matrices  $\hat{\mathbf{A}}_r$  and  $\mathbf{B}_r$ . The simulated matrix  $\mathbf{y}$  will have dimensions  $(N_s, r_c)$ , where  $N_s \leq N$  denotes the number of simulated time steps.

### C. Reconstructed Attractor

Having reconstructed the time series of the eigen-time-delay coordinates  $\mathbf{y}$ , the simulated block Hankel matrix  $\hat{\mathbf{H}}$  is retrieved by reapplying the SVD factorization:

$$\hat{\mathbf{H}} = \mathbf{U}_c \Sigma_c \mathbf{y}^T \in \mathbb{R}^{QM \times (N-M+1)}, \quad (18)$$

where  $\mathbf{U}_c$  and  $\Sigma_c$  represent the SVD matrices of the core components. As previously noted, the SVD factorization preserves channel ordering and block positions in the reconstructed block-Hankel matrix.

Given  $Q \geq 3$  observables, a three-dimensional representation of the embedded attractor can be obtained by plotting the columns from the first block of the simulated Hankel matrix. If  $Q < 3$ , an embedded attractor in  $M$ -dimensional space is recovered. If the observable from which the embedding was obtained is smooth and generic, then it is diffeomorphic to the original attractor, according to Takens' theorem [27]. Such attractor can be visualized in three-dimensional space (assuming the original attractor is three-dimensional). To do so, the columns from three different row indices were chosen in a way that there is little correlation between them.

In cases where a single observable was provided, the selected row indices were  $k = 0$ ,  $k = M/2$  and  $k =$

$M$ . This selection allowed for the visualization of the embedded attractor in three-dimensional space. If  $Q = 2$ , only a single delay is required; in this case, a delay from any of the two inputted time series was chosen so that  $k = M$ .

In order to determine the goodness of the reconstruction of the system dynamics, the coefficient of determination between the original data set (assumed to have  $Q$  input channels) and the reconstructed attractor (which features  $Q$  reconstructed channels), has been calculated. Let  $X_{q,n}^{\text{rec}} := \hat{\mathbf{H}}(q, n)$ , where  $q \in \{0, \dots, Q-1\}$ ,  $q$  is the channel index (rows of a Hankel block) and  $n$  is the time index for the reconstructed data series (columns of the Hankel matrix). When compared to the data series from the original Hankel matrix  $X_{q,n}^{\text{og}} = \mathbf{H}(q, n)$ , the goodness of fit is given by:

$$R_q^2 = 1 - \frac{\sum_{\forall n} (X_{q,n}^{\text{og}} - X_{q,n}^{\text{rec}})^2}{\sum_{\forall n} (X_{q,n}^{\text{og}} - \bar{X}_q^{\text{rec}})^2}, \quad (19)$$

where  $\bar{X}_q^{\text{rec}}$  is the mean of  $X_{q,n}^{\text{rec}}$  over the columns (time indices). Finally, a simple average is performed:

$$\overline{R_{\text{rec}}^2} := \frac{1}{Q} \sum_{q=0}^{Q-1} R_q^2. \quad (20)$$

#### D. Automated rank selection algorithm

In Section 2B we explained how linear and nonlinear components in eigen-delay space can be distinguished by a certain regression procedure, once a value of the cutoff rank  $r$  had been selected. However, nothing was said about how to identify such rank  $r$ . In Ref. [22], the authors found that, at least in the case of the Lorenz System, the precise value of  $r$  was not critical for a good reconstruction of the system dynamics. For mHAVOK, particularly in systems other than the Lorenz system, reconstruction quality changes drastically for different rank choices. Therefore, an appropriate scheme for the optimal selection of  $r$  had to be devised. In the following, two schemes will be described: a variance-informed  $r$  selection, used as a reference case, and an  $R_{\text{rec}}^2$ -informed quality score, one of the innovations of this work.

##### 1. Variance-informed $r$ selection

Given that both HAVOK and mHAVOK are based on SVD techniques, it seems plausible to use a criterion involving the explained variance for the selection of the cutoff rank [34, 35]. This approach has been implemented by several authors [36, 37]. It should be noted that Ref. [35] has a more sophisticated approach to cutoff rank selection, focusing on signal-to-noise ratio, something that has been stressed by Brunton et al. [22].

The reference method implemented in this work is based on the following: given the hierarchically sorted SVD singular values from Eq. (5), the explained variance for a given number of modes is calculated. Taking this into account, a cutoff rank is defined such that a given threshold of explained variance  $\mathbb{V}$  is achieved,

$$\mathbb{V} = \frac{\sum_{k=1}^r \sigma_k^2}{\sum_{k=1}^{MQ} \sigma_k^2} (100\%), \quad (21)$$

where  $\sigma_k$  is the  $k$ -th singular value. As previously mentioned, this method is only employed as a reference and does not constitute an original contribution of this work.

##### 2. $R^2$ -informed quality score

The novel cutoff rank selection method proposed in this work is based on the dynamical matrix  $\mathbf{B}_r$ . This matrix is retrieved for a range of selected ranks  $r \in [r_{\min}, r_a]$ , where  $r_{\min}$  is the smallest rank for which mHAVOK successfully identifies at least one linear mode and  $r_a$  is an arbitrary upper bound. For each  $\mathbf{B}_r$ , a singular value decomposition is performed,

$$\mathbf{B}_r = U \Sigma V^T. \quad (22)$$

Then, the quotient between the maximum and minimum singular values is stored for each  $r \in [r_{\min}, r_a]$ . Mathematically,

$$\kappa(r) := \sigma_{\max}(\mathbf{B}_r)/\sigma_{\min}(\mathbf{B}_r). \quad (23)$$

The rationale behind this criterion stems from the condition number of a matrix [38], which is defined as the quotient between the maximum and minimum singular values and measures how sensible a matrix is to perturbations in its input. In practice,  $r$  values associated with a high quotient enable the system's linearization in embedded space.

With this considered, a subset of high scoring rank values is selected by introducing a percentile cutoff.

The upper  $q\%$  of retrieved scores are stored in a subset  $C \subseteq [r_{\min}, r_a]$ , while the remaining  $(100 - q)\%$  are removed. For each  $r \in C$ , a 70/30 training-testing split is performed on the (possibly multiple) measurement time series. Working on the training set, the routine determines the matrices  $\hat{\mathbf{A}}_r$  and  $\mathbf{B}_r$  from the linear system, as well as the index sets  $\mathcal{I}_c$  and  $\mathcal{I}_f$ . The nonlinear terms are obtained from a projection of the training set onto the evaluation set [22]:

$$\mathcal{V}_l^{(\text{ev})} = \mathbf{S}_l^{-1(\text{tr})} \mathbf{U}_l^{T(\text{tr})} \mathbf{H}^{(\text{ev})}, \quad (24)$$

where  $\cdot^{(\text{tr})}$  and  $\cdot^{(\text{ev})}$  are matrices related to the training and evaluation sets, respectively. Having obtained the eigen-time-delay coordinates for the evaluation data, the nonlinear terms are given by:

$$\mathbf{V}_f^{(\text{ev})} = [\mathbf{v}_i^{(\text{ev})}]_{i \in \mathcal{I}_f^{(\text{tr})}} \quad (25)$$

where  $\mathbf{v}_i^{(\text{ev})}$  are columns of  $\mathcal{V}_k^{(\text{ev})}$ . Using the projected nonlinear terms and the dynamical matrices of the training period, the reconstructed time series for the evaluation set are simulated. Then, the mean  $\overline{R}_{\text{rec}}^2$  defined in Eq. (20) is calculated between the original and the reconstructed evaluation time series after reapplying the SVD factorization.

Each  $r$ -candidate in  $C$  is scored on its ability to reconstruct the evaluation data set, which is provided by the following quality score:

$$\mathcal{Q}(r) := 1 - \overline{R}_{\text{rec}}^2(r), \quad (26)$$

The optimal rank value  $r_{\text{opt}}$  corresponds to the one with a minimal quality score,

$$r_{\text{opt}} = \arg \min_{r \in C} \mathcal{Q}(r). \quad (27)$$

Equivalently,  $\mathcal{Q}(r_{\text{opt}}) = \min[\mathcal{Q}(r)] \forall r \in C$ . We posit that this methodology provides an optimal criterion for rank selection. In order to test the validity of the rank selection algorithm, an independent metric for the fidelity of the system's reconstruction is required.

A summary of this rank selection process, as well as the 70/30 training-testing split on the measurement time series, is presented on Fig. 1.

## E. Quantitative assessment of the goodness of reconstruction

In this work, the Chamfer distance [39] was implemented to measure the accuracy of the system's reconstruction. This metric is calculated in the original coordinate space and provides a measure of the generalized distance between sets of points. Considering prior work on HAVOK only provided a descriptive reconstruction in eigen-delay space [22, 36], the current methodology builds on and goes beyond these works.

In our case, having applied the previously mentioned 70/30 split, we are interested in the distance between the simulated and reconstructed attractors for the evaluation data sets. To ensure computational efficiency, a random subsampling is first performed on the original index set  $I_R = \{0, 1, \dots, N - 1\}$ , defining a new index set  $I_S \subset I_R$  of size  $S < N$ . Given the generated matrices  $X_{n,q}^{\text{og}}$  and  $X_{n,q}^{\text{rec}}$  defined in Section 2C, two corresponding subsampled sets of points are defined: one for the points from the original attractor,

$$\Omega = \{X_{n,:}^{\text{og}} : n \in I_S\}, \quad (28)$$

and one for the reconstructed attractor,

$$\zeta = \{X_{n,:}^{\text{rec}} : n \in I_S\}. \quad (29)$$

Then, both data sets are rescaled by the min-max procedure [40] to ensure repeatability and standardization, i.e., by setting  $x_i \rightarrow (x_i - x_{\min}) / (x_{\max} - x_{\min}) \forall x_i \in \zeta \wedge \Omega$ .  $x_{\max}$  and  $x_{\min}$  denote, respectively, the maximum and minimum values in each of the sets. These normalized subsampled sets serve as the basis for geometric validation.

With this considered, the Chamfer distance  $d_c$  quantifies the average closest-point distance between both sets; the smaller the distance, the higher the geometrical resemblance between the two sets of points. Mathematically, the Chamfer distance is defined as a function  $d_C : \mathbb{R}^{Q \times S} \times \mathbb{R}^{Q \times S} \rightarrow \mathbb{R}$ , where  $Q$  is the number of channels and  $S$  is the subsampled size, such that

$$d_c(\Omega, \zeta) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \min_{z \in \zeta} \|\omega - z\| + \frac{1}{|\zeta|} \sum_{z \in \zeta} \min_{\omega \in \Omega} \|z - \omega\|, \quad (30)$$

where  $|\cdot|$  stands for the cardinality of the sets and  $\|\cdot\|$  is the  $Q$ -dimensional Euclidean distance. In other

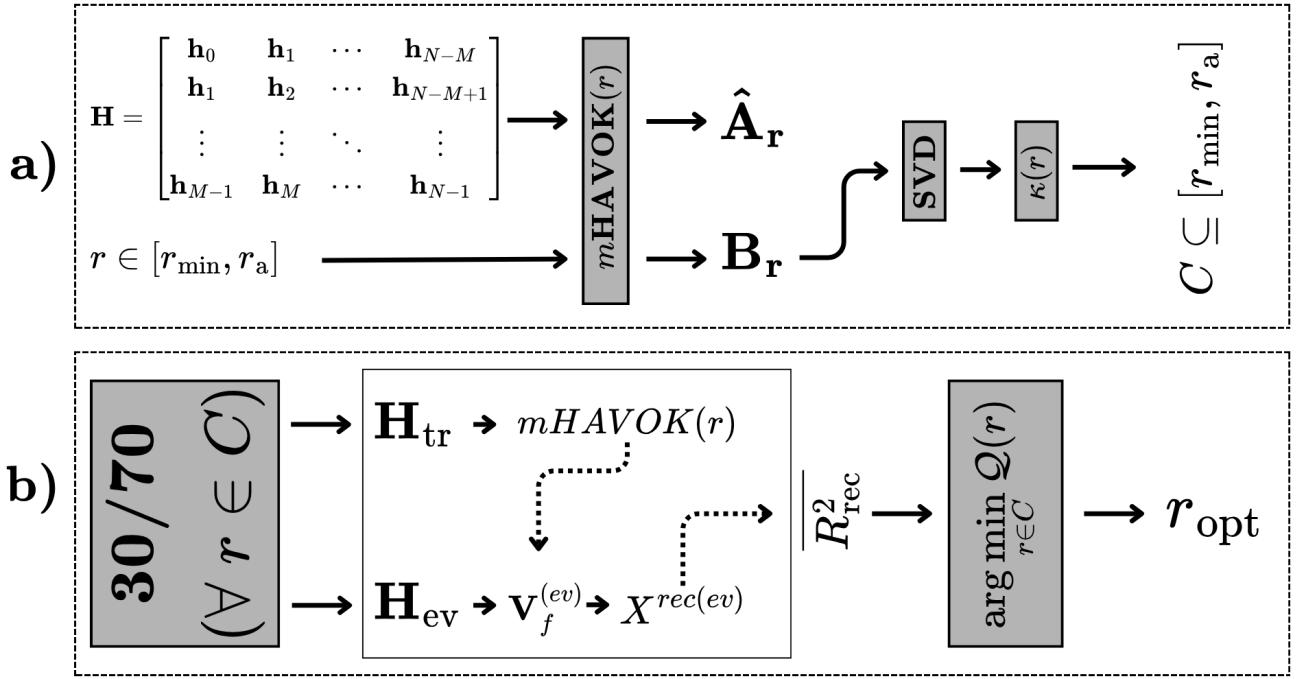


Figure 1: General pipeline for the rank selection procedure developed and utilized throughout this work. The workflow starts in a), where a block Hankel matrix is built and a variety of cutoff ranks  $r \in [r_{\min}, r_a]$  are applied in order to obtain the subset  $C$  of rank-candidates. In b), a condensed illustration of the 70/30 split procedure is presented, where the training set serves as input for  $mHAVOK$ , and the evaluation set provides  $\overline{R}_{\text{rec}}^2$ . Finally, the optimal rank  $r_{\text{opt}}$  is defined as the one that minimizes the quality score  $\mathcal{Q}(r)$ .

words, for each  $\omega \in \Omega$ , the minimum distance with respect to all points in  $\zeta$  is calculated. All minimum distances are summed and then normalized. Similarly, a corresponding procedure is carried out for all points  $z \in \zeta$ . Finally, the Chamfer distance  $d_c$  is calculated as the sum of the two partial distances.

The Hausdorff distance [41] was also employed to quantify the goodness of the reconstructions, but it showed similar results to those from the Chamfer distance; the corresponding results have therefore been omitted for brevity.

## F. Test systems studied

### 1. Lorenz system

The  $mHAVOK$  method introduced in this work was first tested on the Lorenz system, which is given by

$$\dot{x} = \sigma(y - x), \quad (31)$$

$$\dot{y} = x(\rho - z) - y, \quad (32)$$

$$\dot{z} = xy - \beta z, \quad (33)$$

where  $\mathbf{r} = (x, y, z)^T \in R^3$  and the parameter values used throughout this work are

$$\sigma = 10, \quad \rho = 28, \quad \beta = \frac{8}{3}.$$

The time step was chosen to be  $\Delta t = 0.001$ , while the selected simulation time was  $T = 200$ . With this configuration, the original Lorenz attractor shown in Fig. 2 was obtained.

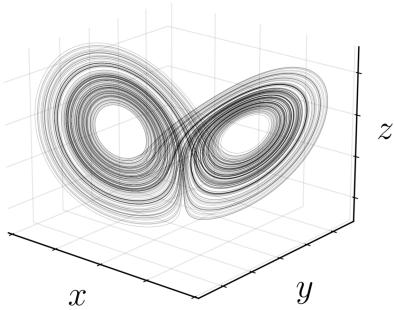


Figure 2: Lorenz attractor generated using Eqs. (31), (32) and (33).

$Q$  channels were introduced into the embedding in Eq. (4) as observables from the state variables  $\mathbf{x} = (x, y, z)$ . The optimal rank value  $r_{\text{opt}}$  was calculated using the methodology in Section 2D2. The embedding dimension was set to  $M = \lceil 100/Q \rceil$  per channel, where  $\lceil \cdot \rceil$  denotes the ceiling function. Although this choice is empirical, it aligns with prior work in HAVOK and delay-coordinate embeddings for chaotic systems [22, 37], where values of  $M \in [100, 150]$  were shown to be sufficient to unfold the attractor's geometry. The Lorenz system was chosen because of its importance as a test case in dynamical systems modeling, as well as its central role in the original HAVOK paper [22]. Its role in this work is mainly to illustrate some of the qualitative improvements of mHAVOK over HAVOK, such as its insensitivity to symmetry blindness and its capability of accurately reconstructing trajectories in original space. It is also used to illustrate the appearance of multiple nonlinear components and the importance of including them in the reconstruction procedure.

## 2. Sprott system

The algorithm's ability to reconstruct complex dynamical systems is tested through the exploration of the Sprott system, which is not globally ergodic [42]. This system is described by the following coupled differential equations:

$$\dot{x} = y + 2xy + xz, \quad (34)$$

$$\dot{y} = 1 - 2x^2 + yz, \quad (35)$$

$$\dot{z} = x - x^2 - y^2. \quad (36)$$

Depending on the initial conditions  $\mathbf{r}_0 = (x_0, y_0, z_0)^T$ , the system either converges to a strange attractor with a noninteger, fractal dimension, or resides in a non-attracting toroidal orbit, which is invariant under the

system's flow [43]. Numerically, the embedding dimension was increased to  $M = \lceil 500/Q \rceil$ . The rest of the simulation parameters remained the same as in the Lorenz system;  $T = 200$  and  $\Delta t = 0.001$ . Fig. 3 shows Sprott system for both types of initial conditions.

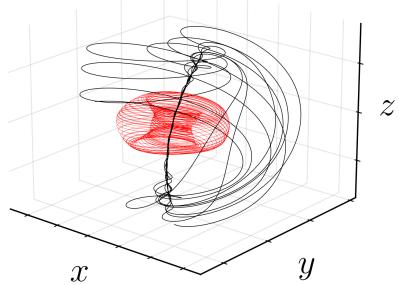


Figure 3: Sprott system generated using Eqs. (34), (35) and (36). The initial conditions were  $\mathbf{r}_0 = (1, 0, 0)^T$  (red) and  $\mathbf{r}'_0 = (2, 0, 0)^T$  (black).

## 3. Results - Lorenz system

In order to demonstrate the current methodology's ability to handle several input time series, the embedding coordinates for the Lorenz system are generated for a varying number of channels, generating multiple time delay vectors (Eq. (2)). Particularly, Fig. 4 shows the three most significant columns of  $\mathbf{V}_c$  defined in Eq. (12), i.e., the three most important linear or core dimensions, demonstrating how providing the model with an increased number of channels allows detecting complex inter-channel relationships. For instance, the pairwise relationship between  $v_3$  and  $v_2$  turns from an uncorrelated pattern in Fig. 4a, to a two-lobe structure reminiscent of the Lorenz attractor, visualized in Fig. 4b.

Moreover, the relationship between  $v_1$  and  $v_2$  changes when introducing two input channels, transitioning from a compressed geometry to the characteristic butterfly shape of the original attractor. Finally, the relationship between  $v_1$  and  $v_3$  appears twisted in the unidimensional embedding, as it potentially lacks vertical information from the original system.

Previous work [30, 31] already pointed out the limitations of using  $f(\mathbf{x}_n) = z_n$  as input channel. This is shown in Fig. 4a, where the reconstructed attractor can be seen to miss the characteristic two-lobe structure of the Lorenz system. In this work, we expand on these findings, showing how each input channel contributes unique geometrical information from the original attractor. Introducing  $f(\mathbf{x}_n) = x_n$  into the  $f(\mathbf{x}_n) = z_n$  embedding allows for the recovery of the two-lobe structure, since it includes the bimodally distributed mode

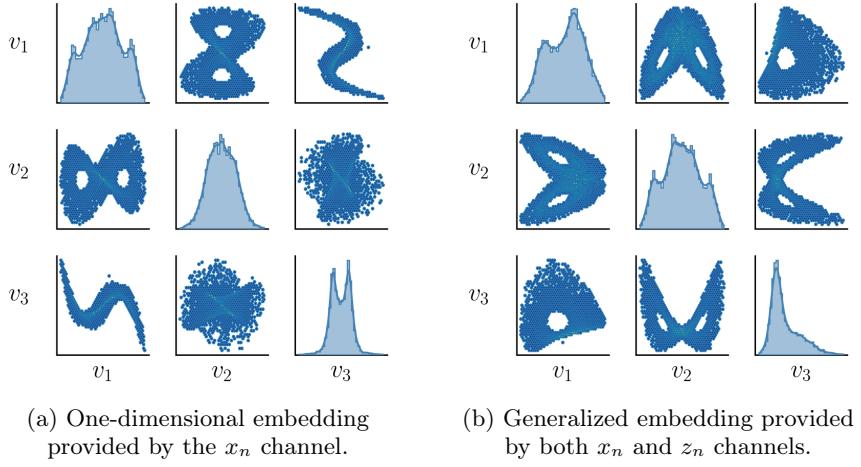


Figure 4: Pairwise relationships between the principal modes in *mHAVOK* embeddings.  $v_i$  are the core dynamic modes.

$v_1$  observed in both instances of Fig. 4. However, introducing  $f(\mathbf{x}_n) = z_n$  alongside the  $f(\mathbf{x}_n) = x_n$  channel has also proven to be beneficial, as it inserts the right-skewed  $v_3$  mode into the embedding, as shown in Fig. 4b. This mode appears to contain radial and vertical information from the original attractor, which is not present in the one-dimensional embedding.

Interestingly, when introducing several measurement channels, multiple independent forcing components are identified, which had to be appropriately classified and separated during *mHAVOK*'s regression. The methodology presented in Section 2B guarantees the identification of multiple nonlinear components, filtering modes that feature  $R^2$  scores lower than  $\tau = 0.95$  during training, as mentioned in the explanations of Eq. (8).

Channels	$r_{\text{opt}}$	$ \mathcal{I}_c $	$ \mathcal{I}_f $
$f(\mathbf{x}_n) = x_n$	7	6	1
$f(\mathbf{x}_n) = y_n$	11	10	1
$f(\mathbf{x}_n) = z_n$	18	11	7
$f(\mathbf{x}_n) = (x_n, y_n)$	9	8	1
$f(\mathbf{x}_n) = (x_n, z_n)$	8	6	1
$f(\mathbf{x}_n) = (y_n, z_n)$	7	5	2
$f(\mathbf{x}_n) = (y_n, z_n^2)$	5	3	2
$f(\mathbf{x}_n) = (x_n, y_n, z_n)$	18	15	3
$f(\mathbf{x}_n) = (x_n, y_n, x_n + z_n)$	9	7	2

Table 1: Component classification of dynamical modes for different input channels using  $\tau = 0.95$ . The optimal rank  $r_{\text{opt}}$  was determined using Eq. (27) with a percentile cutoff  $q = 30\%$ .

Table 1 confirms the presence of several components with a low  $R^2$  score during *mHAVOK*'s component classification. Additionally, it shows how single channel embeddings are susceptible to increased noise correlation, as illustrated by the component classification of  $f(\mathbf{x}_n) = z_n$ , where the number of nonlinear components is significantly higher than for other single channel embeddings. Hence, this particular embedding potentially contains noisy modes. On the other hand, the generalized  $f(\mathbf{x}_n) = (x_n, z_n)$  embedding features only one nonlinear component, demonstrating how an extended embedding may contribute to well-behaved *mHAVOK* reconstructions.

As noticed,  $r_{\text{opt}}$  changes depending on the input channels. For instance, attempting a single channel reconstruction for  $z_n$  using  $r = 11$  results in a disorganized reconstructed attractor, while it guarantees a successful reconstruction for  $y_n$ . Moreover, selecting appropriate measurement time series can drastically reduce the optimal rank value. For example, performing linear combinations of input channels may reduce the number of required dynamical modes, likely because of uncorrelated noise components from each channel canceling out when combined.

In order to assess the success of the algorithm in each of the presented paradigms, each of the reconstructed time series is transformed back to the original coordinate system using the methodology in Section 2C. Fig. 5 displays the reconstructed attractors for some input channel embeddings in the original coordinate space. First off, Fig. 5a showcases how the embedding of  $f(\mathbf{x}_n) = z_n$  fails to represent the original attractor's topology, as reported in References [30, 31]. Then, Fig. 5b shows how *mHAVOK* successfully reproduces the attractor presented in Deyle and Sugihara's

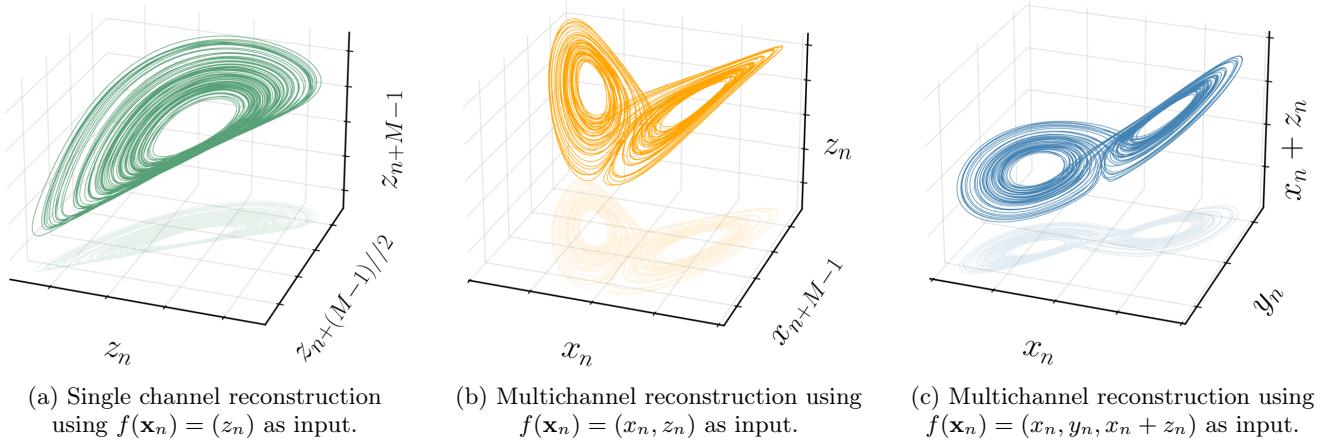


Figure 5: Reconstructed attractors in the original coordinates for multiple observables.

work, where two input channels  $f(\mathbf{x}_n) = (x_n, z_n)$  are included [30]. Complementarily, the algorithm's ability to reproduce the Lorenz attractor is showcased in Fig. 5c, where three input channels  $f(\mathbf{x}_n) = (x_n, y_n, x_n + z_n)$  are provided. Interestingly, the reconstructed attractor is visually similar to the original, even when a linear combination of state variables is included in one of the input channels. Such a result is useful when reconstructing a broader range of real world dynamical systems from experimental data, where multiple sensor readings are available. In these cases, it is common that complicated functions of the system's state variables are measured rather than the state variables themselves.

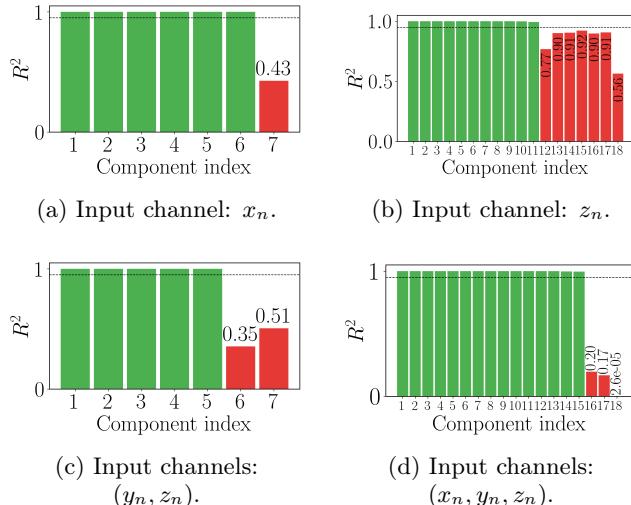


Figure 6: Component classification for different input channels. The dotted line indicates  $\tau = 0.95$ . Green components are linear, red are nonlinear.

Fig. 6 exhibits the  $R^2$  score defined in Eq. (8) for multiple input channels. In consistency with the find-

ings of Brunton et al. [22], the last component in the embedding has a low  $R^2$  score in comparison to those from its neighbors. As shown in Fig. 6, multiple channel embeddings may contain more than one nonlinear component, generalizing Brunton's assertion.

Additionally, Fig. 7 confirms how this generalization is an essential step to reconstruct the system's dynamics from multiple observables. If only the last dynamical mode is used for forcing, the system's reconstruction remarkably degrades, as demonstrated by the marked difference between Fig. 7a, where only the last component has been considered as the forcing term, and Fig. 7b, where all nonlinear components were used for forcing.

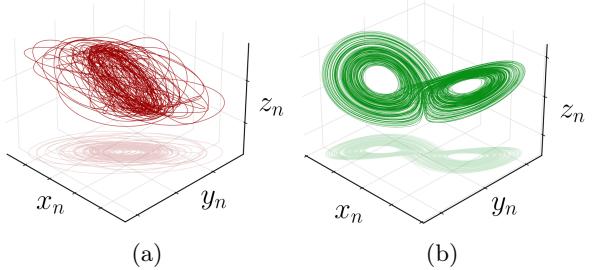


Figure 7: Reconstruction of the Lorenz attractor using (a) the component classification analogous to the original HAVOK method, i.e., considering only one forcing component, and (b) using the component classification scheme proposed by mHAVOK. The input channels were  $(x_n, y_n, z_n)$ .

#### 4. Results - Sprott system

In comparison to the Lorenz attractor, the reconstruction of Sprott system poses a major challenge,

since depending on the initial conditions, the system evolves to a basin of attraction or to an invariant set, as mentioned in Section 2F. For instance, the invariant torus shown in Fig. 3 is neutrally stable for trajectories initialized precisely on it, but does not attract an open neighborhood in phase space [42]. Consequently, marginally distinct initial conditions will send the trajectory outside the torus and towards the strange attractor.

### A. Generalized embeddings

As previously mentioned, the Sprott system features two types of solutions: the conservative torus and the strange attractor, which resembles a cord. Interestingly, this system is invariant under the transformation  $\mathcal{G} : (x, y, z) \rightarrow (x, -y, -z)$  [42]. The torus solutions feature a  $180^\circ$  rotational symmetry about the  $x$ -axis, while the cord is an attractor in forward time and a repellor in reversed time, as well as symmetric under a  $180^\circ$  rotation about the  $x$ -axis. Recalling the argument by Letellier and Aguirre, symmetry-blind observables defined in Section 2A will fail to reconstruct the attractors' topology [31].

Unlike the Lorenz attractor, Sprott system is highly susceptible to the choice of observables and often requires several measurements of the state variables to recover either the strange attractor, the invariant torus, or both. A single time series delay map  $\Psi[n]$  was constructed using information from  $f(\mathbf{x}_n) = x_n$  and  $f(\mathbf{x}_n) = y_n$  with  $M = 1000$ . Then, the transformation  $\mathcal{G}$  was applied to verify invariance of observables. Fig. 8 demonstrates the symmetry blindness of  $\mathcal{G}$ -even measurements from both the torus and the strange attractor. As can be noticed, even observables such as  $x_n$  coincide with their  $\mathcal{G}$  images. In contrast, odd observables are not set-wise coincident. Hence, the attractor features  $\mathcal{G}$ -even symmetry blindness.

The apparent symmetry blindness in both the torus and the strange attractor relates directly to the previously mentioned findings in Ref. [31], where it is shown that observables invariant under a system's symmetry do not provide any information about the lost equivariance, obscuring reconstructions. The authors classified these reconstructions as image systems, where attractors might be dynamically equivalent but stripped of the underlying symmetry properties [31].

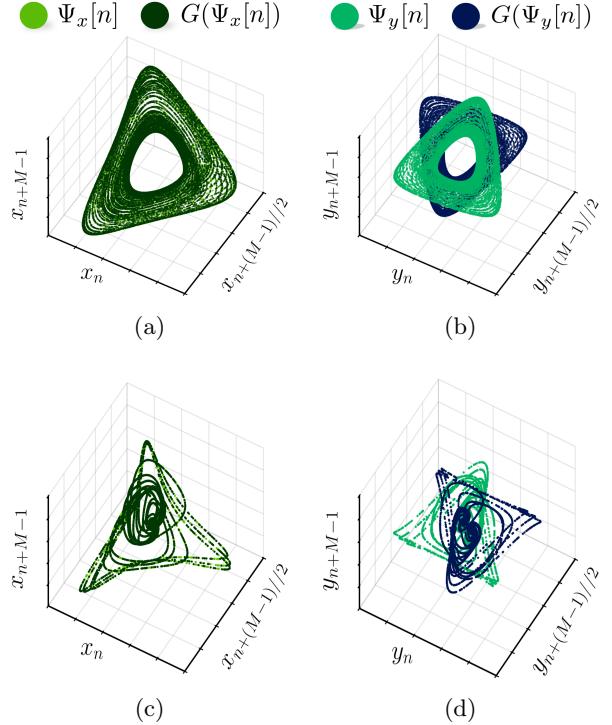


Figure 8: Reconstructions in real space of the Sprott system based on single-channel delay embeddings and their image under  $\mathcal{G}$ . (a, b) Trajectories located on the invariant torus with input channels  $x_n$  and  $y_n$ . (c, d) Trajectories near or on the chord-like attractor with input channels  $x_n$  and  $y_n$ .

Due to the symmetry blindness of the torus' observables, state-space reconstruction using mHAVOK may require multiple observables. As previously mentioned, generalized embeddings provide the necessary machinery by combining several measurement channels, even across different initial conditions. In Fig. 9, mHAVOK was run with different combinations of input channels, plotting the reconstructed attractors employing the methodology in Section 2C. As shown, single even channel reconstructions collapse the torus in Fig. 9a, obscuring the system's symmetry properties. When single odd channel reconstructions were performed, the reconstructions better mirror the original system, as shown in Figs. 9c and 9d. These results confirm the fact that symmetry blind observables negatively impact mHAVOK's ability to reconstruct attractors by obscuring geometrical features; the multichannel generalization provides a safe mechanism when dealing with symmetry blindness, as demonstrated on Fig. 9e, where all three channels were used in the reconstruction.

The dissipation  $\langle y + z \rangle$  defined in [42] was calculated for the reconstructed observables in Fig. 9e. The dissipation for the torus' reconstruction was  $\langle y_1 + z_1 \rangle =$

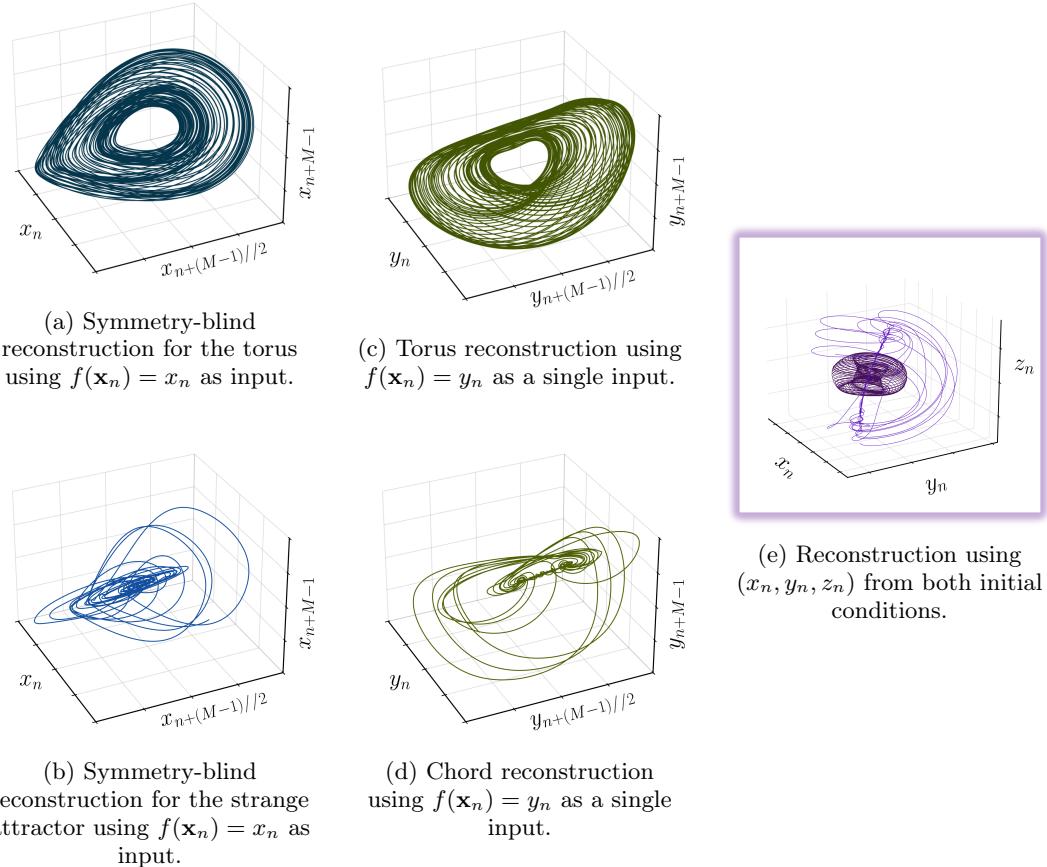


Figure 9: Reconstructed attractors in the original coordinates for different channel inputs.

0.001, in agreement with Sprott’s work [42]. On the other hand, the dissipation for the attractor’s reconstruction was,  $\langle y_2 + z_2 \rangle = -0.044$ , indicating negative dissipation within the basin of attraction, as indicated by Ref. [42]. These results demonstrate that delay maps can faithfully reproduce geometric and dynamical properties given sufficient observable information.

Beyond geometry and dissipation, it is important to evaluate the reconstruction’s ability to classify nonlinear components that drive the dynamics. Similar to the work by Brunton et al. [22], *mHAVOK* was employed as a diagnostic tool to identify nonlinear forcing contributions, as visualized in Fig. 10.

Reportedly, the conservative torus has Lyapunov exponents  $(0, 0, 0)$  [42]. Therefore, *mHAVOK* should in principle return a purely linear model, which is largely confirmed by the time series of Fig. 10a. In practice, the appearance of small nonlinear terms in the torus is attributed to numerical noise and truncation artifacts, highlighting the need for a posterior analysis tool that discerns noise from true forcing. On the other hand, Fig. 10b features nonlinear terms during intermittent

events. Therefore, they are likely to be related to true system forcing, making them essential for a faithful reconstruction.

## B. Automated rank selection algorithm

Throughout this section, *mHAVOK* was run by simultaneously using the channels  $f(\mathbf{x}_n) = (x_n, y_n, z_n)$  from both initial conditions, which are necessary to reconstruct the invariant torus and the chord-like attractor, respectively. The rationale behind this decision will be discussed in the following (Subsection 4C).

### 1. Variance-informed $r$ selection

At first sight, it would seem plausible to use a variance-informed rank selection using Eq. (21), since predominant dynamical modes should amount to high percentages of explained variance. However, Fig. 11 ex-

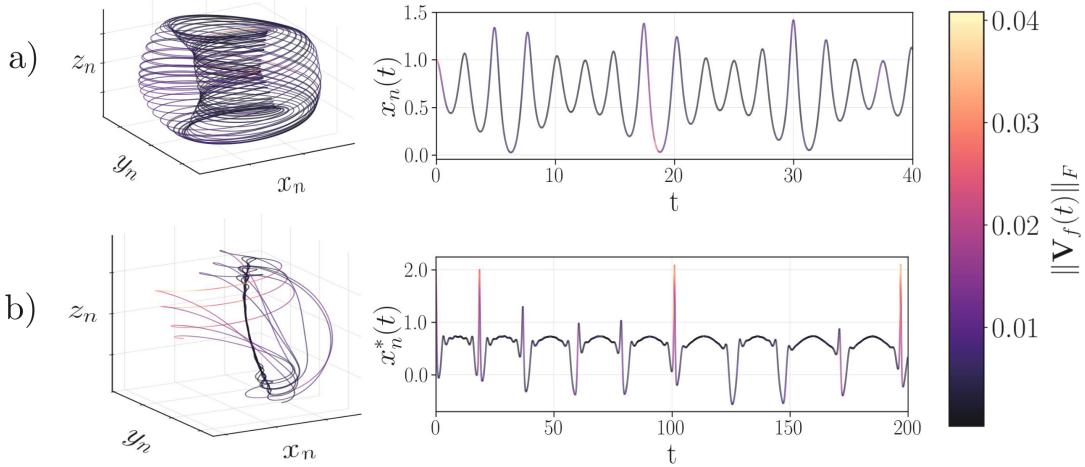


Figure 10: Nonlinear components  $\mathbf{V}_f$  identified by  $m\text{HAVOK}$  for a) the torus and b) the strange attractor.

hibits how  $r = 9$ , the smallest rank for which  $m\text{HAVOK}$  successfully identified one linear mode, already amounts to more than 99% of the total variance, with adjacent rank values showing similar results. However, this particular value, as well as its neighbors, yields deficient reconstructions, as will be shown in the following (Fig. 13a).

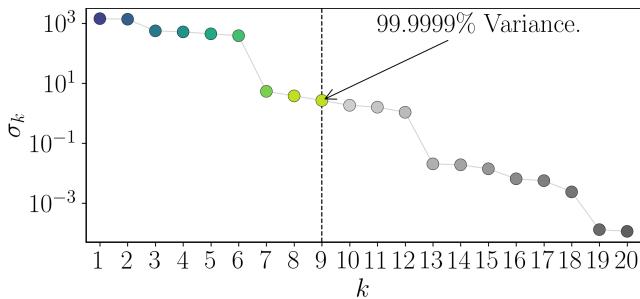


Figure 11: Singular values for the SVD of the block Hankel matrix in Eq. (5).

Therefore, traditional cutoff criteria based on explained variance are not adequate to work as a proper  $r$ -selection criterion, motivating a more objective algorithm capable of determining the optimal rank value.

## 2. $R^2$ -informed quality score

For a range of values  $r \in [9, 25]$ , the condition number  $\kappa(\mathbf{B}_r)$  defined in Subsection 2D2 was calculated, providing a first filter for the possible  $r$ -candidates. In general, remarkable peaks in the condition number are connected to optimal reconstructions in multichannel settings. For instance, Fig. 12 displays a prominent

peak at  $r = 19$ , which corresponds to the rank yielding the best coefficient of determination  $\overline{R_{\text{rec}}^2}$ . While such peaks provide useful guidance, in practice several candidates need to be examined in order to identify the optimal rank. The following figure highlights the 30% best rank candidates for Sprott system.

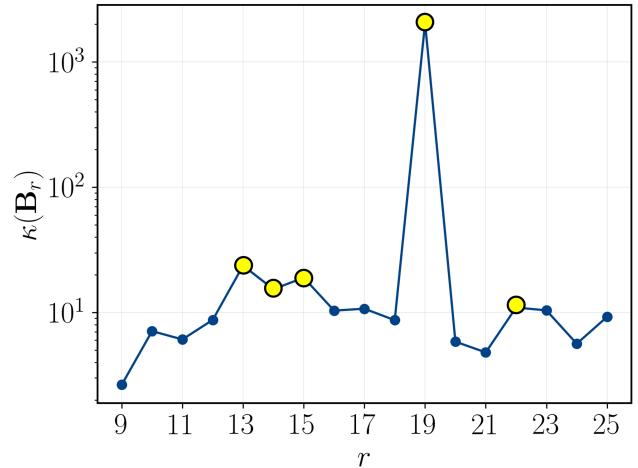


Figure 12: Condition number of  $\mathbf{B}_r$  for multiple rank values. The marked values denote ranks with the 30% highest condition number.

## C. Quantitative evaluation

The Chamfer distance  $d_C$  was calculated for some rank values, providing an objective measure of the goodness of the reconstructions. Fig. 13 demonstrates how  $d_C$  tends to decrease as  $r$  increases, reaching a minimum at  $r_{\text{opt}} = 19$ . As opposed to the rank selection in the

Lorenz system which is not critical, as already pointed out by Brunton et al. [22], the Sprott system is highly susceptible to the cutoff rank. For instance,  $r = 20$ , one step ahead of the optimal rank value  $r_{\text{opt}}$ , yields a distance  $d_C = 204 \times 10^{-3}$ , which is almost 40 times greater than  $d_C(r_{\text{opt}})$ . Therefore, a slight rank variation generates significantly different reconstructions.

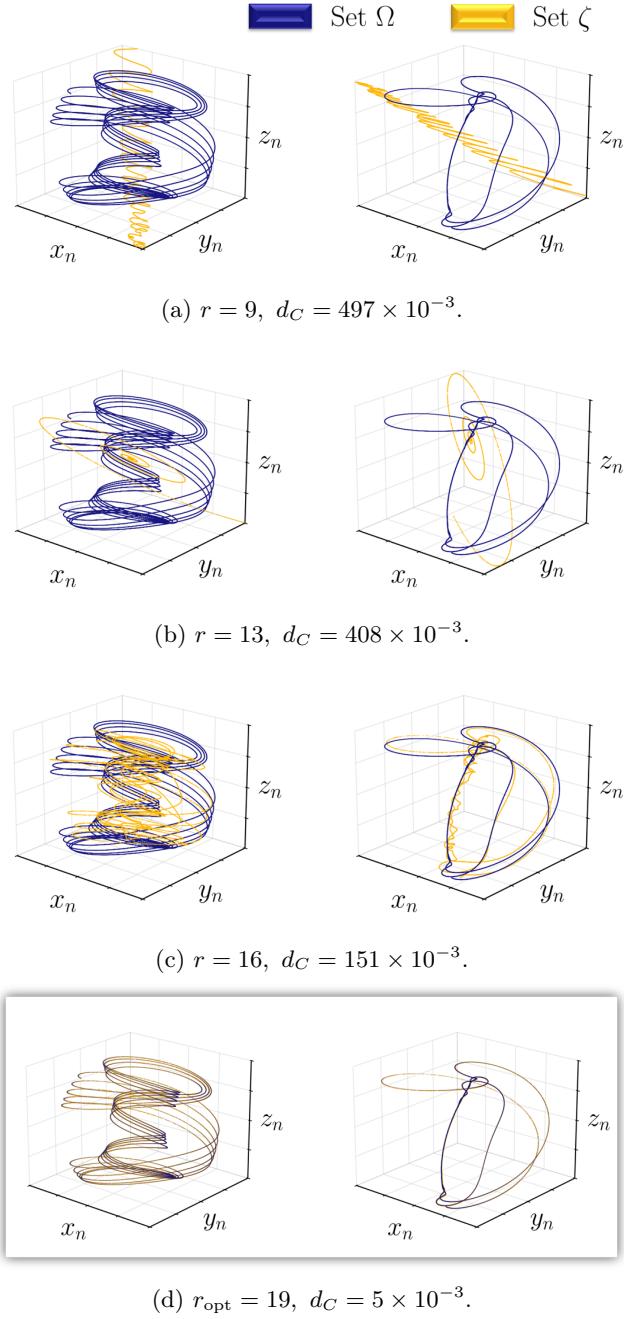


Figure 13: Comparison of Chamfer distances for different rank values. The sets  $\Omega$  and  $\zeta$  are defined in Section 2E.

Interestingly, Fig. 14 shows how simultaneously providing six channels,  $f(\mathbf{x}_n) = (x_n, y_n, z_n)$  for both initial conditions (for the invariant torus and the chord-like attractor, respectively), yields a considerable decrease in  $d_C$  in comparison to reconstructing the torus and the strange attractor as separate inputs. This can be attributed to the fact that noise correlation is decreased by introducing more observables, as discussed by Delye and Sugihara [30].

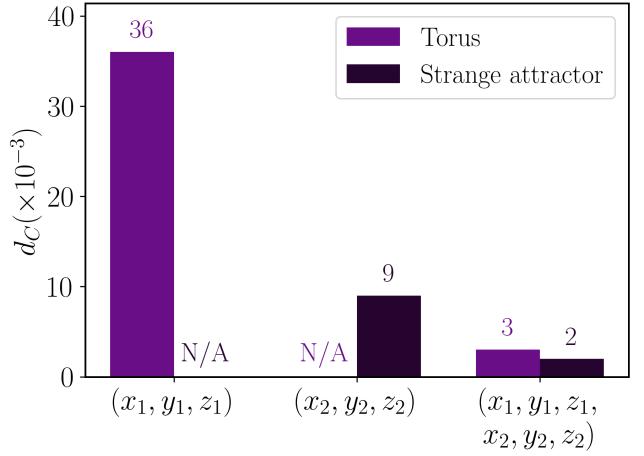


Figure 14: Chamfer distance  $d_C$  with respect to the input channels. The simulations were performed using the optimal rank value  $r_{\text{opt}}$  for each case.  $(x_1, y_1, z_1)$  and  $(x_2, y_2, z_2)$  represent, respectively, the initial conditions from which the torus and strange attractor can be reconstructed.

## 5. Conclusions

A generalized multichannel extension of the HAVOK framework ( $m$ HAVOK) has been developed and validated through simulation. By applying the novel methodology to the Lorenz and Sprott systems, the model's ability to recover the full underlying dynamics is enhanced by incorporating multiple and functionally diverse observables. The algorithm was driven by the key insight that providing the model with more than one informative input signal, alongside a generalized component classification scheme, enriches the embedding and ultimately enables a complete unfolding of the system dynamics. This was particularly evident in the case of the Sprott system, where faithful reconstruction required information from separate initial conditions.

The extension to multichannel embeddings was not only natural but necessary given the practical limitations of real-world measurements, which often consist of indirect or function-based observations. By allowing

the integration of such diverse input sources, *mHAVOK* positions itself as a practical and flexible tool. The proposed model is readily applicable in scenarios involving sensor fusion or multiscale dynamics, where capturing the global structure of the attractor is essential for understanding or control.

This framework opens a clear direction for future work by incorporating predictive models for the forcing components. Since *mHAVOK* relies on known or precomputed forcing signals, extending the methodol-

ogy to forecast those inputs would allow the transition from reconstruction to full forecasting. This would significantly enhance the autonomy and applicability of the model across scientific and engineering domains.

Future work may explore adaptive or data-driven selection of  $M$  based on embedding dimension tests, as well as automated threshold techniques. Additionally, a rigorous and well-structured proof of the connection between Koopman's theory and *mHAVOK* is yet to be done and will be discussed in future articles.

- 
- [1] G. Haller and B. Kaszás, *Nonlinear Dynamics* **112**, 18639 (2024).
  - [2] M. Elin and D. Shoikhet, *Linearization models for complex dynamical systems: Topics in univalent functions, functional equations and semigroup theory*, Vol. 208 (Springer Science & Business Media, 2011).
  - [3] M. Yassen, *Chaos, Solitons & Fractals* **26**, 913 (2005).
  - [4] S. Vaidyanathan and C. Volos, *Advances and applications in chaotic systems*, Vol. 636 (Springer, 2016).
  - [5] M. Budišić, R. Mohr, and I. Mezić, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22**, 047510 (2012).
  - [6] B. O. Koopman, *Proceedings of the National Academy of Sciences* **17**, 315 (1931).
  - [7] Y. Lan and I. Mezić, *Physica D: Nonlinear Phenomena* **242**, 42 (2013).
  - [8] M. Santos Gutiérrez, V. Lucarini, M. D. Chekroun, and M. Ghil, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **31**, 053116 (2021).
  - [9] Q. Li, F. Dietrich, E. M. Boltt, and I. G. Kevrekidis, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**, 103111 (2017).
  - [10] E. Kaiser, J. N. Kutz, and S. L. Brunton, *Machine Learning: Science and Technology* **2**, 035023 (2021).
  - [11] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, *Journal of Nonlinear Science* **25**, 1307 (2015).
  - [12] A. Mauroy and I. Mezić, *IEEE Transactions on Automatic Control* **61**, 3356 (2016).
  - [13] E. M. Boltt, Q. Li, F. Dietrich, and I. Kevrekidis, *SIAM Journal on Applied Dynamical Systems* **17**, 1925 (2018).
  - [14] M. J. Colbrook and A. Townsend, *Communications on Pure and Applied Mathematics* **77**, 221 (2024).
  - [15] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *Journal of Fluid Mechanics* **641**, 115 (2009).
  - [16] M. Korda and I. Mezić, *Automatica* **93**, 149 – 160 (2018).
  - [17] D. Giannakis, *Applied and Computational Harmonic Analysis* **47**, 338 (2019).
  - [18] M. Korda, M. Putinar, and I. Mezić, *Applied and Computational Harmonic Analysis* **48**, 599 (2020).
  - [19] I. Mezić, *Nonlinear Dynamics* **41**, 309 (2005).
  - [20] I. Mezić, in *2015 54th IEEE Conference on Decision and Control (CDC)* (2015) pp. 7034–7041.
  - [21] S. E. Otto and C. W. Rowley, *Annual Review of Control, Robotics, and Autonomous Systems* **4**, 59 (2021).
  - [22] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, *Nature Communications* **8**, 19 (2017).
  - [23] H. Arbabi and I. Mezić, *SIAM Journal on Applied Dynamical Systems* **16**, 2096 (2017).
  - [24] X. Cai and Y. Qian, *Algorithms* **17** (2024), 10.3390/a17090388.
  - [25] G. Yılmaz Bingöl, O. A. Soysal, and E. Günay, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **34**, 083143 (2024).
  - [26] J. Yang, J. Zhao, J. Song, J. Wu, C. Zhao, and H. Leng, *Entropy* **24** (2022), 10.3390/e24030408.
  - [27] F. Takens, in *Dynamical Systems and Turbulence, Warwick 1980*, edited by D. Rand and L.-S. Young (Springer, Berlin, Heidelberg, 1981) pp. 366–381.
  - [28] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, *Phys. Rev. Lett.* **45**, 712 (1980).
  - [29] T. Sauer, J. A. Yorke, and M. Casdagli, *Journal of Statistical Physics* **65**, 579 (1991).
  - [30] E. R. Deyle and G. Sugihara, *PLOS ONE* **6**, 1 (2011).
  - [31] C. Letellier and L. A. Aguirre, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **12**, 549 (2002).
  - [32] J.-N. Juang and R. S. Pappa, *Journal of Guidance, Control, and Dynamics* **8**, 620 (1985).
  - [33] J. Demmel, “6. Singular Value Decomposition,” in *Templates for the Solution of Algebraic Eigenvalue Problems* (Society for Industrial and Applied Mathematics, 2000) pp. 135–147.
  - [34] R. Mahmoudvand and M. Zokaei, *Chil. J. Stat.* **3**, 43 (2011).
  - [35] M. Gavish and D. L. Donoho, *IEEE Transactions on Information Theory* **60**, 5040 (2014).
  - [36] X. Cai and Y. Qian, *Algorithms* **17** (2024), 10.3390/a17090388.
  - [37] R. G. Moulder, E. Martynova, and S. M. B. and, *Multivariate Behavioral Research* **58**, 441 (2023), pMID: 35001769.
  - [38] A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, *SIAM Journal on Numerical Analysis* **16**, 368 (1979).
  - [39] H. Fan, H. Su, and L. Guibas, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society, 2017) pp. 2463–2471.
  - [40] S. G. Patro and D.-K. K. Sahu, *IARJSET* (2015),

- 10.17148/IARJSET.2015.2305.
- [41] W. Rucklidge, ed., “The Hausdorff distance,” in *Efficient Visual Recognition Using the Hausdorff Distance* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996) pp. 27–42.
- [42] J. Sprott, Physics Letters A **378**, 1361 (2014).
- [43] A. Algaba, M. Merino, B.-W. Qin, and A. J. Rodríguez-Luis, Physics Letters A **383**, 1441 (2019).