RESEARCH ARTICLE | DECEMBER 09 2022

The spatiotemporal coupling in delay-coordinates dynamic mode decomposition

Emil Bronstein ✉ 🔲 ; Aviad Wiegner 🔲 ; Doron Shilo 🔲 ; Ronen Talmon 🔲

Check for updates

View Online   Export Citation   CrossMark

AIP Publishing

The spatiotemporal coupling in delay-coordinates dynamic mode decomposition

View Online   Export Citation   CrossMark

Emil Bronstein,[1,a] ID Aviad Wiegner,[2] ID Doron Shilo,[1] ID and Ronen Talmon[2] ID

AFFILIATIONS

[1]Faculty of Mechanical Engineering, Technion – Israel Institute of Technology, Haifa 3200003, Israel
[2]Viterbi Faculty of Electrical and Computer Engineering, Technion – Israel Institute of Technology, Haifa 3200003, Israel

[a]**Author to whom correspondence should be addressed:** emilbr@campus.technion.ac.il

ABSTRACT

Dynamic mode decomposition (DMD) is a leading tool for equation-free analysis of high-dimensional dynamical systems from observations. In this work, we focus on a combination of DMD and delay-coordinates embedding, which is termed *delay-coordinates DMD* and is based on augmenting observations from current and past time steps, accommodating the analysis of a broad family of observations. An important utility of DMD is the compact and reduced-order spectral representation of observations in terms of the DMD eigenvalues and modes, where the temporal information is separated from the spatial information. From a spatiotemporal viewpoint, we show that when DMD is applied to delay-coordinates embedding, temporal information is intertwined with spatial information, inducing a particular spectral structure on the DMD components. We formulate and analyze this structure, which we term the *spatiotemporal coupling in delay-coordinates DMD*. Based on this spatiotemporal coupling, we propose a new method for DMD components selection. When using delay-coordinates DMD that comprises redundant modes, this selection is an essential step for obtaining a compact and reduced-order representation of the observations. We demonstrate our method on noisy simulated signals and various dynamical systems and show superior component selection compared to a commonly used method that relies on the amplitudes of the modes.

Dynamical systems are abundant in many fields of science and engineering. As dynamical systems are often high-dimensional and complex, their characterization from observations is a coveted goal. A key task in accomplishing this goal is finding a reduced-order representation of a system, where dynamic mode decomposition (DMD) is a prominent tool for this purpose. In recent years, the combination of delay-coordinates embedding and DMD (i.e., delay-coordinates DMD) has been shown to be highly useful in the characterization of dynamical systems, even when these systems are highly nonlinear or chaotic. This combination gives rise to a specific spectral structure that does not exist in ordinary DMD and has not been studied so far. In this work, we formulate and analyze this structure, where we show that the spectral components of delay-coordinates DMD exhibit a spatiotemporal coupling. This coupling suggests that the representations obtained by delay-coordinates DMD can be further reduced. By considering this coupling, we propose not only to decouple the spatial information from the temporal information but also to exploit it to construct an improved reduced-order

representation in an unsupervised fashion. We demonstrate our approach on several dynamical systems that include noisy observations of undamped and damped mechanical oscillators.

I. INTRODUCTION

Time-series analysis, modeling, and prediction are ubiquitous tasks in applied sciences. When the time-series stem from ergodic dynamical systems, learning their phase space in a nonparametric fashion from sufficiently long intervals of observations is possible and has become an active field of research in recent years. This so-called data-driven approach seeks to obtain meaningful, physics-related knowledge, in an *equation-free* manner, shifting the focus from equations-based descriptions to observations-based analysis.[1–6] Existing data-driven methods can be largely divided into two categories.[7] The first is based on a state-space representation, and the focus is typically on finding a map that propagates a state at a given time to a state at a future time. Arguably, the classical approach

16 January 2024 08:57:34

is to approximate the nonlinear dynamics as a collection of locally linear systems on tangent spaces near attractors.[8–10] Another more recent class of methods in this category constructs reduced models of the state space in Euclidean spaces and then finds the propagation rules in the reduced spaces.[1,11–13]

The second category is an operator-theoretic approach that considers observables of the system. In this approach, many recent methods are based on the Koopman operator,[14] which is a linear, infinite-dimensional operator that operates on observables in a Hilbert space and propagates them *linearly* in time.[15,16] While the ability to represent nonlinear dynamics in a linear fashion is perhaps the most notable property of using the Koopman operator, it has several other remarkable attributes. In this line of work,[7,15,17–19] it was shown that the Koopman operator has the ability of capturing the dynamics of linear or nonlinear dynamical systems through its spectral components. More concretely, the time evolution of a dynamical system can be decomposed into spatial patterns, which are often referred to as the Koopman modes and derived from the eigenfunctions of the Koopman operator, and temporal patterns, which are derived from the eigenvalues of the Koopman operator. Therefore, the spectral analysis of the Koopman operator is of high importance for obtaining an informative description of dynamical systems.[7,17,18,20–22]

Using the Koopman operator for analysis poses an important trade-off. On the one hand, it facilitates a linear description of the dynamics and a useful decoupling of the spatial and temporal patterns via its spectral representation. On the other hand, the Koopman operator is infinite dimensional. Therefore, finite-dimensional approximations are necessary for practical purposes. Perhaps, the most common technique for such a finite approximation is the dynamic mode decomposition (DMD), introduced by Schmid and Sesterhenn,[23,24] where the primary goal is to approximate the Koopman eigenvalues and modes in finite spaces based on a finite set of observations.[25] The DMD eigenvalues and modes give rise to a *discrete* spatiotemporal representation of the dynamical system, facilitating a *data-driven* and *equation-free* analysis. Following its appearance, several variants for enhancing the capabilities of DMD have been introduced, e.g., the extended DMD (EDMD),[26] where projections on finite-space dictionaries were used to improve the spatiotemporal representation, least-squares,[27,28] and sparsity promoting[29,30] techniques, as well as approaches for the analysis of compressed data.[31,32] Additional extensions of DMD have been introduced, such as multiresolution DMD (mrDMD),[33] which effectively uncovers multiscale structures in the data, and DMD with control (DMDc),[34] which extracts low-order models of high-dimensional systems that require control.

Indeed, in recent years, DMD has been shown to be a powerful tool, demonstrating remarkable capabilities in a broad range of fields[35–40] and, particularly, in fluid dynamics.[25,41–47] Nevertheless, despite its popularity and success, DMD has some notable limitations. For example, a straightforward application of DMD does not allow for the signal reconstruction of a standing wave.[16,48] In addition, ordinary DMD cannot handle cases when the number of linearly independent DMD modes is smaller than the number of the system's oscillation frequencies.[49] Another notable limitation of DMD concerns one-dimensional signals, which are the output of an abundance of sensors, such as thermocouples,[50] force sensors,[51,52]

and magnetic and acoustic emission sensors.[53,54] The application of DMD to such signals results in a degenerate (scalar) representation.

Combining delay-coordinates embedding and DMD mitigates these limitations.[16,48,49] The delay-coordinates embedding (also referred to as time-delay embedding, or, simply, delay-coordinates) is a method for augmenting past observations to the present observation. This approach dates back to 1981 with the formulation of Takens' embedding theorem,[55] according to which an attractor of a system can be reconstructed up to a diffeomorphism using delay embedding.

Various studies that associate delay-coordinates and DMD-based methods have been presented in recent years. For example, Le Clainche and Vega[49] presented the higher order DMD (HODMD), a global linear method that is capable of uncovering a large number of frequencies of periodic and quasiperodic dynamical systems based on limited and noisy input data. Brunton *et al.* introduced the HAVOK analysis in terms of a Hankel matrix, which successfully represents highly nonlinear and chaotic systems using a linear model and intermittent forcing.[35] Pan and Duraisamy provided the minimal required augmentation number for a perfect recovery of dynamical systems based on their Fourier spectrum.[56] We note that combinations of delay-coordinates and Koopman operator-based methods have also been investigated.[19,57,58]

In this paper, we show that despite its prevalence, application of DMD to *augmented* data (i.e., delay-coordinates DMD) gives rise to a particular spectral structure that does not exist in ordinary DMD applications and, to the best of our knowledge, has not been studied in the existing literature. This structure comprises augmented DMD modes that embody *temporal* information, which is entangled with *spatial* information. We formulate and analyze this entanglement, and term it the *spatiotemporal coupling in delay-coordinates DMD*.

Based on this spatiotemporal coupling, we propose a new approach for obtaining compact and reduced-order representations of dynamical systems from observations, where, similar to ordinary DMD, the spatial and temporal patterns are *decoupled*. Our approach includes solving an inherent challenge of delay-coordinates DMD, where the number of *augmented* DMD components is often larger than the number of intrinsic modes of the dynamical system. In such cases, the augmented DMD components can be divided into two subsets: those that describe the dynamical system, which we term *true*, and those that are a mere artifact of the augmentation, unrelated to the system, which we term *spurious*.

By relying on the difference in their spatiotemporal coupling, we distinguish between the true and spurious DMD components and present a method for selecting the true components. In contrast to ordinary DMD, our method is based on delay-coordinates DMD and is, therefore, capable of analyzing and representing a broader family of observations and signals. We demonstrate the spatiotemporal coupling and the effectiveness of our approach on various simulated dynamical systems.

This paper is organized as follows. Section II briefly describes existing approaches for mode selection. Section III presents the problem formulation. In Sec. IV, we introduce the spatiotemporal coupling in delay-coordinates DMD in detail and reveal the specific relations within augmented DMD modes. Then, in Sec. V, we propose a new method for decoupling the spatial and temporal information, leading to the identification of the true augmented

DMD components and to a compact, reduced-order, and informative representation of the observations. For the purpose of illustration, in Sec. VI, we present the application of our method to a two-mode sine signal. Finally, in Sec. VII, our method is demonstrated on various dynamical systems, outperforming the common method that relies on the amplitudes of the modes at low signal-to-noise ratio (SNR) values.

II. RELATED WORK ON MODE SELECTION

The identification of the dominant DMD components required for optimal, reduced-order representations of dynamical systems (which we term *true*) has been studied in wider contexts, beyond delay-coordinates DMD. Criteria and methods for such a purpose are usually referred to as *mode selection*, where we use the broader term *DMD components selection*. For instance, Rowley *et al.* ordered the DMD modes by their norms (amplitudes)[25]—a method we term the *maximal amplitudes method*. The norms of the modes can be further weighted by the magnitudes of the corresponding eigenvalues to account for modes that have large norms yet decay rapidly, as suggested by Tu *et al.*[48] Schmid *et al.* used a projection of the data sequence on the identified modes, whose coefficients indicate on the significance of the modes.[59] Jovanović *et al.* introduced a sparsity-promoting approach using an addition of a penalty term of the DMD amplitudes, which regularizes the least-square deviation between the linear combination of DMD modes and the snapshots matrix.[29] Tissot *et al.* proposed an energetic criterion, where the amplitude of a DMD mode is weighted by its corresponding temporal coefficient.[60] Sayadi *et al.* proposed a parameterized approach, which utilizes the sparsity promoting DMD, and then reconstructs the modes' amplitudes using a time-dependent coefficient, giving a notion of their significance.[30] Another approach was proposed by Kou and Zhang, which considered the initial conditions and temporal evolution of DMD modes, ordering them according to the integrals of their corresponding time coefficients.[61]

Our work differs from the above studies on mode selection, as it provides, for the first time to the best of our knowledge, a mode selection framework in the specific context of delay-coordinates DMD. Seemingly, the problem of mode selection in delay-coordinates DMD is more challenging due to the existence of spurious DMD modes, as well as the higher dimensions of augmented DMD modes compared to the observations. Nevertheless, we show that despite the additional challenges, delay-coordinates DMD also constitutes a remedy. Specifically, we show that the spatiotemporal coupling in delay-coordinates DMD bears information that facilitates a new method for mode selection.

III. PROBLEM FORMULATION

Consider a dynamical system

$$x_{k+1} = f(x_k), \quad k \in \mathbb{Z}, \tag{1}$$

which evolves in discrete time k on a manifold $\mathcal{M} \subset \mathbb{R}^n$, where $x_k \in \mathcal{M}$ is a state vector, and $f: \mathcal{M} \mapsto \mathcal{M}$. The function f is unknown and could be deterministic or stochastic, e.g., due to the presence of noise. Assume that this discrete time formulation

arises from a continuous-time dynamical system $\dot{x} = \mathcal{A}x$. Specifically, suppose that the state is sampled at a fixed sampling rate $\omega_s = 2\pi/\Delta t$, and $m + 1$ discrete samples $x_k = x(k\Delta t)$ of the state are collected, where $k = 0, 1, \ldots, m$. In this work, following common practice,[49,56,58] we assume that $m > n$.

Our goal is to analyze the dynamics of system (1) based on the finite set of observations $\{x_k\}_{k=0}^m$. For this purpose, we use the DMD approach,[7,23,48] which provides a spatiotemporal representation of the observations, and is described next. First, a linear approximation of the discrete-time system in (1) is employed,

$$x_{k+1} = Ax_k, \quad A \in \mathbb{R}^{n \times n}, \quad \text{and } x_k \in \mathbb{R}^n. \tag{2}$$

To find A, the observations $\{x_k\}_{k=0}^m$ are arranged into two observation matrices

$$X = \begin{bmatrix} | & | & & | \\ x_0 & x_1 & \cdots & x_{m-1} \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times m},$$

$$X' = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times m}. \tag{3}$$

Then, following the exact DMD,[16,48] the singular value decomposition (SVD) of X is computed as

$$X = U\Sigma V^*, \tag{4}$$

where $U \in \mathbb{C}^{n \times r}$, $\Sigma \in \mathbb{C}^{r \times r}$, $V \in \mathbb{C}^{m \times r}$, $(\cdot)^*$ is the complex conjugate transpose, and $r = \mathrm{rank}(X)$. Define $\tilde{A} \in \mathbb{C}^{r \times r}$ by

$$\tilde{A} = U^*X'V\Sigma^{-1}. \tag{5}$$

The so-called DMD eigenvalues, λ_i, are the eigenvalues of $\tilde{A}$ that satisfy

$$\tilde{A}v_i = \lambda_i v_i, \quad i = 1, \ldots, r, \tag{6}$$

where v_i are the corresponding eigenvectors. The so-called exact DMD modes, ϕ_i, are given by[16,48]

$$\phi_i = \frac{1}{\lambda_i} X'V\Sigma^{-1}v_i, \quad i = 1, \ldots, r. \tag{7}$$

Tu *et al.* defined the exact DMD modes, ϕ_i, and showed that $\{\lambda_i, \phi_i\}$ are the eigenvalue-eigenvector pairs of A.[48] Moreover, the authors distinguished between the exact modes ϕ_i and the DMD modes obtained by the standard DMD[24] (termed projected DMD modes). It was noted that the exact and projected modes have the tendency to converge when X and X' have the same column spaces.[16]

The observations x_k can be expressed as

$$x_k = \sum_{i=1}^r \lambda_i^k \phi_i \eta_{0,i}, \tag{8}$$

where $\eta_{0,i}$ is the ith entry of $\eta_0 = \Phi^* x_0$ and Φ is the matrix whose columns are the r leading DMD modes ϕ_i, satisfying $\Phi^*\Phi = I$.

Alternatively, (8) can be written in a matrix form as

$$\boldsymbol{x}_k = \boldsymbol{\Phi}\boldsymbol{c}_k,$$

$$\boldsymbol{\Phi} = \begin{bmatrix} | & | & & | \\ \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_r \\ | & | & & | \end{bmatrix} \in \mathbb{C}^{n \times r},$$

$$\boldsymbol{c}_k = \begin{bmatrix} \lambda_1^k \eta_{0,1} \\ \lambda_2^k \eta_{0,2} \\ \vdots \\ \lambda_r^k \eta_{0,r} \end{bmatrix} \in \mathbb{C}^r. \tag{9}$$

The continuous-time counterpart of the decomposition in (8), which corresponds to the dynamical system $\dot{\boldsymbol{x}} = \mathcal{A}\boldsymbol{x}$, is given by[16]

$$\boldsymbol{x}(t) = \sum_{i=1}^{r} \exp(\mu_i t) \boldsymbol{\psi}_i v_{0,i}, \tag{10}$$

where μ_i and $\boldsymbol{\psi}_i$ are the eigenvalues and eigenvectors of the propagating operator $\mathcal{A}$, $v_{0,i} = \langle \boldsymbol{\psi}_i, \boldsymbol{x}(0) \rangle = \boldsymbol{\psi}_i^* \boldsymbol{x}(0)$, and

$$\lambda_i = \exp(\mu_i \Delta t). \tag{11}$$

In case of a damped oscillating system, when the system comprises natural frequencies ω_i and damping ratios ζ_i, according to (11), we have

$$\mu_i = -\zeta_i \omega_i \pm j\omega_i \sqrt{1 - \zeta^2}, \quad j = \sqrt{-1}, \tag{12}$$

from which we obtain

$$\omega_i = \frac{|\log(\lambda_i)|}{\Delta t}, \quad \zeta_i = -\frac{\Re(\log(\lambda_i))}{\omega_i \Delta t}. \tag{13}$$

In case the system is undamped, $\zeta = 0$ can be substituted into (12) and (13). Equation (8) constitutes a decomposition of the system in (1) into its dynamic modes, such that λ_i and $\boldsymbol{\phi}_i$ hold the temporal and spatial information, respectively. Yet, as described in Sec. I, obtaining this decomposition might be hindered for a variety of reasons, such as standing waves, one- or low-dimensional signals,[16,48] and observation noise (e.g., noise in the measurement equipment). Our goal is to obtain a compact and reduced-order representation of the system similar to Eq. (8), where the challenges mentioned above are present. To accomplish this goal, we use *delay-coordinates DMD* as detailed below.

IV. SPATIOTEMPORAL COUPLING IN DELAY-COORDINATES DMD

In this section, we show that when DMD is applied to delay-coordinates embedding (constituting the delay-coordinates DMD), the strict separation of temporal and spatial information in representations obtained by DMD as in (8) is violated. In other words, the separation where the eigenvalues λ_i bear the temporal information about the dynamical system and the modes $\boldsymbol{\phi}_i$ represent the spatial information no longer exists in delay-coordinates DMD. Then, we show that this seemingly limiting spatial and temporal information entanglement could be harnessed toward the selection of the

DMD components required for accurate and reduced-order characterization of the dynamical system. Specifically, we formulate and analyze the induced spatiotemporal structure of the eigenvalues and modes that arise from delay-coordinates DMD, which we term *augmented* DMD components. For simplicity, we divide the exposition into two stages. First, an augmentation of one sample is considered in Subsection IV A, followed by a generalization to augmentations of several samples, which is presented in Subsection IV B.

We begin by formulating the augmentation, i.e., applying delay-coordinates embedding to the observation matrices $\boldsymbol{X}$ and $\boldsymbol{X}'$ in Eq. (3). Let $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{X}}'$ denote the augmented observation matrices, defined by

$$\hat{\boldsymbol{X}} = \begin{bmatrix} | & | & & | \\ \hat{\boldsymbol{x}}_0 & \hat{\boldsymbol{x}}_1 & \cdots & \hat{\boldsymbol{x}}_{m-s-1} \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{x}_0 & \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{m-s-1} \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_{m-s} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{x}_s & \boldsymbol{x}_{s+1} & \cdots & \boldsymbol{x}_{m-1} \end{bmatrix} \in \mathbb{R}^{[n(s+1)] \times [m-s]},$$

$$\hat{\boldsymbol{X}}' = \begin{bmatrix} | & | & & | \\ \hat{\boldsymbol{x}}_1 & \hat{\boldsymbol{x}}_2 & \cdots & \hat{\boldsymbol{x}}_{m-s} \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_{m-s} \\ \boldsymbol{x}_2 & \boldsymbol{x}_3 & \cdots & \boldsymbol{x}_{m-s+1} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{x}_{s+1} & \boldsymbol{x}_{s+2} & \cdots & \boldsymbol{x}_m \end{bmatrix} \in \mathbb{R}^{[n(s+1)] \times [m-s]}, \tag{14}$$

where $s < m$, $s \in \mathbb{N}$ is termed the *augmentation number*. In the remainder of the paper, we use hats to denote either augmented or augmented-related terms. Note that $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{X}}'$ are Hankel matrices, which are typical in delay-coordinates DMD.[16,35,58]

Same as in the exact DMD,[16,48] $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{X}}'$ are related through

$$\hat{\boldsymbol{X}}' = \hat{\boldsymbol{A}}\hat{\boldsymbol{X}}, \quad \hat{\boldsymbol{A}} \in \mathbb{R}^{[n(s+1)] \times [n(s+1)]}, \tag{15}$$

or in vector form,

$$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}_k, \quad \hat{\boldsymbol{x}}_k, \hat{\boldsymbol{x}}_{k+1} \in \mathbb{R}^{n(s+1)},$$
$$k = 0, \ldots, m - s - 1. \tag{16}$$

Application of the exact DMD to $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{X}}'$ results in $\hat{r}$ DMD eigenvalue-mode pairs denoted by $\{\hat{\lambda}_l, \hat{\boldsymbol{\phi}}_l\}_{l=1}^{\hat{r}}$, where $\hat{r} = \text{rank}(\hat{\boldsymbol{X}})$. Then, similar to Eq. (8), the augmented observations $\hat{\boldsymbol{x}}_k$ can be represented as

$$\hat{\boldsymbol{x}}_k = \sum_{l=1}^{\hat{r}} \hat{\lambda}_l^k \hat{\boldsymbol{\phi}}_l \hat{\eta}_{0,l}, \quad k = 0, \ldots, m - s - 1, \tag{17}$$

where $\hat{\eta}_{0,l}$ is the lth entry of $\hat{\boldsymbol{\eta}}_0 = \hat{\boldsymbol{\Phi}}^* \hat{\boldsymbol{x}}_0$ and $\hat{\boldsymbol{\Phi}}$ is the matrix whose columns are the augmented DMD modes $\hat{\boldsymbol{\phi}}_l$. Equation (17) can be

written in a matrix form as

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{c}}_k,$$

$$\hat{\boldsymbol{\Phi}} = \begin{bmatrix} | & | & & | \\ \hat{\boldsymbol{\phi}}_1 & \hat{\boldsymbol{\phi}}_2 & \cdots & \hat{\boldsymbol{\phi}}_{\hat{r}} \\ | & | & & | \end{bmatrix} \in \mathbb{C}^{[n(s+1)]\times\hat{r}}, \qquad (18)$$

$$\hat{\boldsymbol{c}}_k = \begin{bmatrix} \hat{\lambda}_1^k \eta_{0,1} \\ \hat{\lambda}_2^k \eta_{0,2} \\ \vdots \\ \hat{\lambda}_{\hat{r}}^k \eta_{0,\hat{r}} \end{bmatrix} \in \mathbb{C}^{\hat{r}}.$$

In order to obtain a sufficient amount of augmented DMD components for a full description of the dynamical system, $\hat{r}$ must be at least as large as the number of DMD components required for the description of the unaugmented (original) system in (8); i.e., $\hat{r} \geq r$. Recalling that $\hat{r} = \text{rank}(\hat{\boldsymbol{X}}) \leq \min(n(s+1), m-s)$, fulfillment of $\hat{r} \geq r$ depends on the choice of s. Concretely choosing s such that $n(s+1) \leq m-s$ leads to the range

$$0 \leq s \leq \frac{m-n}{n+1}. \qquad (19)$$

Conversely, the choice $n(s+1) \geq m-s$ imposes $r \leq \hat{r} \leq m-s$, which leads to

$$\frac{m-n}{n+1} \leq s \leq m-r. \qquad (20)$$

In practice, as $r = \text{rank}(\boldsymbol{X})$ is unknown prior to the application of DMD, one can choose $s \leq m-n$ as the upper bound in (20).

By (16) and by recalling that $\hat{\boldsymbol{A}}\hat{\boldsymbol{\phi}}_l = \hat{\lambda}_l \hat{\boldsymbol{\phi}}_l$, the expansion of the augmented observations vector $\hat{\boldsymbol{x}}_{k+1}$ is given by

$$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}_k = \sum_{l=1}^{\hat{r}} \hat{\lambda}_l^k \hat{\lambda}_l \hat{\boldsymbol{\phi}}_l \hat{\eta}_{0,l}, \qquad (21)$$

which, using (18), can be recast in the matrix form as

$$\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{c}}_k,$$
$$\hat{\boldsymbol{\Lambda}} = \text{diag}[\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_{\hat{r}}] \in \mathbb{C}^{\hat{r}\times\hat{r}}. \qquad (22)$$

Equations (17) and (21) show that the temporal propagation of $\hat{\boldsymbol{x}}_k$ can be expressed using the augmented DMD components. Specifically, multiplication of the augmented modes $\hat{\boldsymbol{\phi}}_l$ by their corresponding augmented eigenvalues $\hat{\lambda}_l$ propagates $\hat{\boldsymbol{x}}_k$ one time step to $\hat{\boldsymbol{x}}_{k+1}$. Similarly, $\hat{\boldsymbol{x}}_k$ can be propagated q steps to $\hat{\boldsymbol{x}}_{k+q}$ by

$$\hat{\boldsymbol{x}}_{k+q} = \hat{\boldsymbol{A}}^q \hat{\boldsymbol{x}}_k = \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Lambda}}^q \hat{\boldsymbol{c}}_k, \quad q = 0, \dots, s. \qquad (23)$$

A. Augmentation with one sample

When $s = 1$, $\hat{\boldsymbol{x}}_k \in \mathbb{R}^{2n}$ can be written as $\hat{\boldsymbol{x}}_k = [\boldsymbol{x}_k, \boldsymbol{x}_{k+1}]^T$, splitting it to the top and bottom n entries. Consequently, (18) can be

rewritten as

$$\begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{x}_{k+1} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Phi}}^{(0)} \\ \hat{\boldsymbol{\Phi}}^{(1)} \end{bmatrix} \hat{\boldsymbol{c}}_k, \qquad (24)$$

where the columns of $\hat{\boldsymbol{\Phi}}^{(0)} \in \mathbb{C}^{n\times\hat{r}}$ and $\hat{\boldsymbol{\Phi}}^{(1)} \in \mathbb{C}^{n\times\hat{r}}$ are the top and bottom n entries of the columns of $\hat{\boldsymbol{\Phi}}$. We term each column of $\hat{\boldsymbol{\Phi}}^{(0)}$ or $\hat{\boldsymbol{\Phi}}^{(1)}$ as a *DMD sub-mode* or, simply, *sub-mode*. Considering only the top n entries in Eq. (24) yields

$$\boldsymbol{x}_k = \hat{\boldsymbol{\Phi}}^{(0)} \hat{\boldsymbol{c}}_k. \qquad (25)$$

On the one hand, Eq. (25) shows that $\boldsymbol{x}_k$ can be expressed by the $\hat{r}$ columns of $\hat{\boldsymbol{\Phi}}^{(0)}$. On the other hand, according to (9), it can be represented using only the $r \leq \hat{r}$ columns of $\boldsymbol{\Phi}$. Therefore, Eq. (25) might not be a compact representation of $\boldsymbol{x}_k$, leading to the conjecture that the application of delay-coordinates DMD results in two types of DMD components—those that describe the dynamical system (true) and those that are not related to it and are an artifact of the augmentation (spurious). By applying similar considerations to the bottom n entries in (24), the same conjecture can be made for $\boldsymbol{x}_{k+1}$ and $\hat{\boldsymbol{\Phi}}^{(1)}$. In the following assumption, we make these conjectures more precise. We note that this assumption is supported by an empirical verification in Sec. VII.

Assumption 1: *Without loss of generality, we write the matrices $\hat{\boldsymbol{\Phi}}^{(j)} \in \mathbb{C}^{n\times\hat{r}}$ as $\hat{\boldsymbol{\Phi}}^{(j)} = \left[\hat{\boldsymbol{\Phi}}_{true}^{(j)}, \hat{\boldsymbol{\Phi}}_{spurious}^{(j)}\right]$, $j = 0, 1$, where the r leftmost columns $\hat{\boldsymbol{\Phi}}_{true}^{(j)} \in \mathbb{C}^{n\times r}$ are termed true and the remaining $\hat{r} - r$ columns $\hat{\boldsymbol{\Phi}}_{spurious}^{(j)} \in \mathbb{C}^{n\times[\hat{r}-r]}$ are termed spurious. Assume that the matrix $\hat{\boldsymbol{\Phi}}_{true}^{(j)}$ has full column rank r and that its column space spans the space of observations $\{\boldsymbol{x}_k\}$.*

In other words, we assume that the r leftmost columns of $\hat{\boldsymbol{\Phi}}^{(0)}$ are the ones required for the compact reduced-ordered representation of $\boldsymbol{x}_k$ (hence, true), while the rest $\hat{r} - r$ columns are redundant (hence, spurious). According to (9) and by Assumption 1, implying that the columns of $\hat{\boldsymbol{\Phi}}_{true}^{(0)}$ are linearly independent, $\boldsymbol{x}_k$ can be represented using only the r true columns of $\hat{\boldsymbol{\Phi}}^{(0)}$ and their respective eigenvalues.

Now, consider $\hat{\boldsymbol{x}}_{k+1} \in \mathbb{R}^{2n}$, which for $s = 1$ is $\hat{\boldsymbol{x}}_{k+1} = [\boldsymbol{x}_{k+1}, \boldsymbol{x}_{k+2}]^T$. Then, from (22) and by splitting $\hat{\boldsymbol{\Phi}}$, $\hat{\boldsymbol{\Lambda}}$ and $\hat{\boldsymbol{c}}_k$ to their true and spurious parts as above, we have

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{x}_{k+2} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Phi}}_{true}^{(0)}, \hat{\boldsymbol{\Phi}}_{spurious}^{(0)} \\ \hat{\boldsymbol{\Phi}}_{true}^{(1)}, \hat{\boldsymbol{\Phi}}_{spurious}^{(1)} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\Lambda}}_{true} & 0 \\ 0 & \hat{\boldsymbol{\Lambda}}_{spurious} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{c}}_k^{true} \\ \hat{\boldsymbol{c}}_k^{spurious} \end{bmatrix}, \qquad (26)$$

where $\hat{\boldsymbol{\Lambda}}_{true} = \text{diag}[\hat{\lambda}_1, \dots, \hat{\lambda}_r]$, $\hat{\boldsymbol{\Lambda}}_{spurious} = \text{diag}[\hat{\lambda}_{r+1}, \dots, \hat{\lambda}_{\hat{r}}]$, and $\hat{\boldsymbol{c}}_k^{true}$, $\hat{\boldsymbol{c}}_k^{spurious}$ denote the r and $\hat{r} - r$ expansion coefficients that correspond to the true and spurious components, respectively.

Under Assumption 1, $\boldsymbol{x}_{k+1}$ can be represented using only the true parts of (24) and (26) as

$$\boldsymbol{x}_{k+1} = \hat{\boldsymbol{\Phi}}_{true}^{(1)} \hat{\boldsymbol{c}}_k^{true}, \quad \boldsymbol{x}_{k+1} = \hat{\boldsymbol{\Phi}}_{true}^{(0)} \hat{\boldsymbol{\Lambda}}_{true} \hat{\boldsymbol{c}}_k^{true}. \qquad (27)$$

Consequently, by equating the right-hand side (RHS) terms of Eq. (27), we have

$$\left[\hat{\mathbf{\Phi}}_{\text{true}}^{(1)} - \hat{\mathbf{\Phi}}_{\text{true}}^{(0)}\hat{\mathbf{\Lambda}}_{\text{true}}\right]\hat{c}_k^{\text{true}} = \mathbf{0}, \quad k = 0, \ldots, m - s - 1. \quad (28)$$

Proposition 2: *For distinct augmented DMD eigenvalues $\hat{\lambda}_l$,*

$$\hat{\mathbf{\Phi}}_{true}^{(1)} = \hat{\mathbf{\Phi}}_{true}^{(0)}\hat{\mathbf{\Lambda}}_{true}. \quad (29)$$

Note that assuming distinct DMD eigenvalues is a common practice.[25,49,62]

Proof. Denote $\hat{\mathbf{C}} \in \mathbb{C}^{r \times [m-s]}$ as a matrix whose columns are $\hat{c}_k^{\text{true}}$, $k = 0, \ldots, m - s - 1$. Based on the definition of $\hat{c}_k$ in Eq. (18), $\hat{\mathbf{C}}$ can be recast as a product of a diagonal matrix $\hat{\boldsymbol{\eta}}_0 \in \mathbb{C}^{r \times r}$ and a $[m-s]$th order Vandermonde matrix $\hat{\mathbf{V}}_{\text{and}}^{m-s}(\hat{\lambda}_1, \ldots, \hat{\lambda}_r) \in \mathbb{C}^{r \times [m-s]}$, both consisting of the r true DMD eigenvalues, as

$$\hat{\mathbf{C}} = \begin{bmatrix} | & | & & | \\ \hat{c}_0^{\text{true}} & \hat{c}_1^{\text{true}} & \cdots & \hat{c}_{m-s-1}^{\text{true}} \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \hat{\eta}_{0,1} & 0 & \cdots & 0 \\ 0 & \hat{\eta}_{0,2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \hat{\eta}_{0,r} \end{bmatrix} \begin{bmatrix} 1 & \hat{\lambda}_1 & (\hat{\lambda}_1)^2 & \cdots & (\hat{\lambda}_1)^{m-s-1} \\ 1 & \hat{\lambda}_2 & (\hat{\lambda}_2)^2 & \cdots & (\hat{\lambda}_2)^{m-s-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \hat{\lambda}_r & (\hat{\lambda}_r)^2 & \cdots & (\hat{\lambda}_r)^{m-s-1} \end{bmatrix}$$

$$= \hat{\boldsymbol{\eta}}_0 \hat{\mathbf{V}}_{\text{and}}^{m-s}(\hat{\lambda}_1, \ldots, \hat{\lambda}_r). \quad (30)$$

For distinct DMD eigenvalues, the rank of $\hat{\mathbf{V}}_{\text{and}}^{m-s}(\hat{\lambda}_1, \ldots, \hat{\lambda}_r)$ is $\min(r, m - s)$.[56] Since $r \leq m - s$ [see Eq. (20)], rank$\left(\hat{\mathbf{V}}_{\text{and}}^{m-s}(\hat{\lambda}_1, \ldots, \hat{\lambda}_r)\right) = r$. For a non-trivial representation of the dynamical system, i.e., $\hat{\eta}_{0,l} \neq 0$, $l = 1, \ldots r$, we have that rank$\left(\hat{\boldsymbol{\eta}}_0\right) = r$. Consequently, rank$(\hat{\mathbf{C}}) = r$.

For convenience, we define $\mathbf{\Delta} = \hat{\mathbf{\Phi}}_{\text{true}}^{(1)} - \hat{\mathbf{\Phi}}_{\text{true}}^{(0)}\hat{\mathbf{\Lambda}}_{\text{true}}$. As the columns of $\hat{\mathbf{C}}$ are in the subspace of the solutions to Eq. (28), then dim$\left(\ker(\mathbf{\Delta})\right) = r$. By the rank-nullity theorem, the rank of the matrix $\mathbf{\Delta}$ equals to the matrix's number of columns r minus the dimension of its null space, i.e., dim$\left(\ker(\mathbf{\Delta})\right) = r$. Therefore, rank$(\mathbf{\Delta}) = 0$, implying that $\mathbf{\Delta} = \mathbf{0}$. Hence, $\hat{\mathbf{\Phi}}_{\text{true}}^{(1)} = \hat{\mathbf{\Phi}}_{\text{true}}^{(0)}\hat{\mathbf{\Lambda}}_{\text{true}}$, yielding Eq. (29). $\square$

Denote the columns of $\hat{\mathbf{\Phi}}_{\text{true}}^{(0)}$ and $\hat{\mathbf{\Phi}}_{\text{true}}^{(1)}$ by $\hat{\boldsymbol{\phi}}_l^{(0)}$ and $\hat{\boldsymbol{\phi}}_l^{(1)}$, respectively. Their consideration along with $\hat{\lambda}_l$ in $\hat{\mathbf{\Lambda}}_{\text{true}}$ in (29) yields

$$\hat{\boldsymbol{\phi}}_l^{(1)} = \hat{\boldsymbol{\phi}}_l^{(0)}\hat{\lambda}_l, \quad l = 1, \ldots, r. \quad (31)$$

Proposition 2 and Eq. (31) show the spatiotemporal coupling in delay-coordinates DMD for the special case of $s = 1$. In this case, the sub-modes $\hat{\boldsymbol{\phi}}_l^{(0)}, \hat{\boldsymbol{\phi}}_l^{(1)} \in \mathbb{C}^n$ that comprise the true augmented mode $\hat{\boldsymbol{\phi}}_l \in \mathbb{C}^{2n}$ are related to one another through the corresponding true eigenvalue $\hat{\lambda}_l$ of $\hat{\mathbf{A}}$. The general relations for augmentation with more than one sample are presented in Subsection IV B.

B. Augmentation with several samples

In case $s > 1$, two observation vectors that are augmented s times, denoted by $\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{x}}_{k+q} \in \mathbb{R}^{n(s+1)}$, are considered. Then, from Eqs. (18) and (23), we have

$$\hat{\boldsymbol{x}}_k = \begin{bmatrix} \boldsymbol{x}_k \\ \vdots \\ \boldsymbol{x}_{k+j} \\ \vdots \\ \boldsymbol{x}_{k+s} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{\Phi}}^{(0)} \\ \vdots \\ \hat{\mathbf{\Phi}}^{(j)} \\ \vdots \\ \hat{\mathbf{\Phi}}^{(s)} \end{bmatrix} \hat{c}_k,$$

$$\hat{\boldsymbol{x}}_{k+q} = \begin{bmatrix} \boldsymbol{x}_{k+q} \\ \vdots \\ \boldsymbol{x}_{k+q+i} \\ \vdots \\ \boldsymbol{x}_{k+q+s} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{\Phi}}^{(0)} \\ \vdots \\ \hat{\mathbf{\Phi}}^{(i)} \\ \vdots \\ \hat{\mathbf{\Phi}}^{(s)} \end{bmatrix} \hat{\mathbf{\Lambda}}^q \hat{c}_k, \quad (32)$$

where the matrix $\hat{\mathbf{\Phi}}$ in Eqs. (18) and (23) is viewed as a column-stack of $s + 1$ "sub-modes" matrices $\hat{\mathbf{\Phi}}^{(0)}, \ldots, \hat{\mathbf{\Phi}}^{(s)}$, each of dimensions $n \times \hat{r}$. Without loss of generality, we assume that $j = q + i$, $i, j = 0, \ldots, s$. Thus, both $\hat{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{x}}_{k+q}$ in (32) contain the same unaugmented observation $\boldsymbol{x}_{k+j} = \boldsymbol{x}_{k+q+i}$. Similarly to the case of $s = 1$, by equating the true components of the RHS terms in (32) that correspond to $\boldsymbol{x}_{k+j}$ and $\boldsymbol{x}_{k+q+i}$ and substituting $j = q + i$, we have

$$\left[\hat{\mathbf{\Phi}}_{\text{true}}^{(j)} - \hat{\mathbf{\Phi}}_{\text{true}}^{(i)}\hat{\mathbf{\Lambda}}_{\text{true}}^{j-i}\right]\hat{c}_k^{\text{true}} = \mathbf{0}, \quad k = 0, \ldots, m - s - 1. \quad (33)$$

Next, we generalize Assumption 1 and Proposition 2 for augmentation with more than one sample, where the spatiotemporal coupling in delay-coordinates DMD is bounded from above by B, which is defined and explained in the sequel in Eq. (40).

Assumption 3: *The matrices $\hat{\mathbf{\Phi}}_{true}^{(j)}$, $j = 0, \ldots, B$, comprising the r leftmost columns of $\hat{\mathbf{\Phi}}^{(j)}$, are assumed to have full column ranks and that the column space of each of them spans the space of observations $\{\boldsymbol{x}_k\}$.*

The division of $\hat{\mathbf{\Phi}}$ into $\hat{\mathbf{\Phi}}_{\text{true}}^{(j)} \in \mathbb{C}^{n \times r}$ and $\hat{\mathbf{\Phi}}_{\text{spurious}}^{(j)} \in \mathbb{C}^{n \times [\hat{r}-r]}$, $j = 0, \ldots, B$, as in Assumptions 1 and 3, is illustrated in Fig. 1(a).

Proposition 4: *For distinct augmented DMD eigenvalues $\hat{\lambda}_l$,*

$$\hat{\mathbf{\Phi}}_{true}^{(j)} = \hat{\mathbf{\Phi}}_{true}^{(i)}\hat{\mathbf{\Lambda}}_{true}^{j-i}, \quad i, j = 0, \ldots, B. \quad (34)$$

We omit the proof because it is similar to the proof of Proposition 2.

Consideration of $\hat{\boldsymbol{\phi}}_l^{(i)}$ and $\hat{\boldsymbol{\phi}}_l^{(j)}$, which are the respective ith and jth sub-modes of the true, augmented modes $\hat{\boldsymbol{\phi}}_l$ [i.e., the columns of $\hat{\mathbf{\Phi}}_{\text{true}}^{(i)}$ and $\hat{\mathbf{\Phi}}_{\text{true}}^{(j)}$ in (34), respectively], as well as $\hat{\lambda}_l$ in $\hat{\mathbf{\Lambda}}_{\text{true}}$, $l = 1, \ldots, r$, leads to

$$\hat{\boldsymbol{\phi}}_l^{(j)} = \hat{\boldsymbol{\phi}}_l^{(i)}\hat{\lambda}_l^{j-i}, \quad i, j = 0, \ldots, B, \quad l = 1, \ldots, r. \quad (35)$$

The spatiotemporal coupling in delay-coordinates DMD, generally presented in Proposition 4 and Eq. (35), reveals that the true,
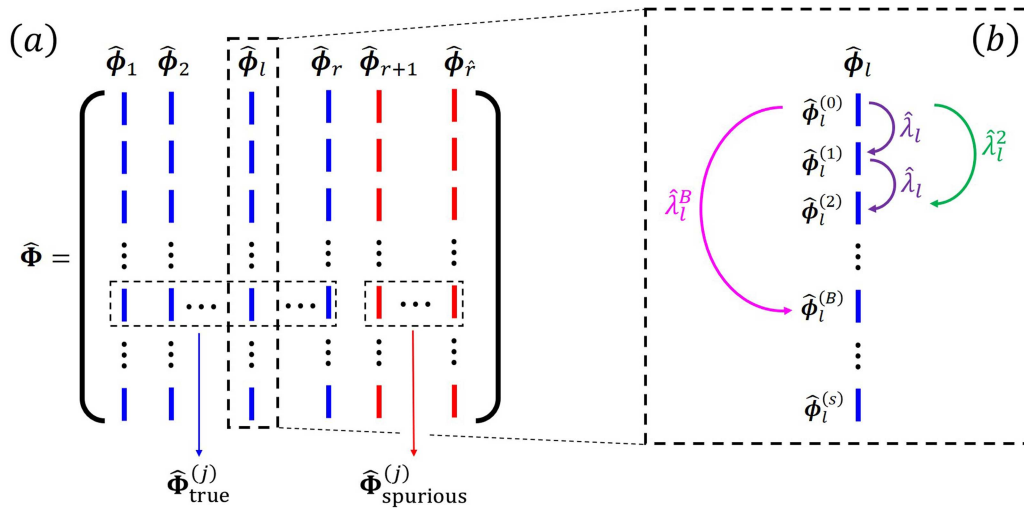
FIG. 1. (a) Division of $\hat{\Phi}$, whose columns are the augmented DMD modes $\hat{\phi}_l$, into $\hat{\Phi}_{\text{true}} \in \mathbb{C}^{[n(s+1)] \times r}$ and $\hat{\Phi}_{\text{spurious}} \in \mathbb{C}^{[n(s+1)] \times [\hat{r}-r]}$. The columns of the former and latter are the r true and $\hat{r} - r$ spurious augmented DMD modes, respectively. (b) Illustration of the spatiotemporal coupling in delay-coordinates DMD, according to Proposition 4 and Eq. (35).

augmented modes $\hat{\phi}_l$ contain not only spatial information, but also temporal information given by $\hat{\lambda}_l$. Importantly, by Proposition 4, this coupling is exhibited only by the true DMD components. In contrast, the considerations leading to this property do not apply to the spurious DMD components, and, therefore, they lack this coupling. This spatiotemporal coupling can be written more explicitly using (35), e.g., as

$$\hat{\phi}_l = \begin{bmatrix} \hat{\phi}_l^{(0)} \\ \hat{\phi}_l^{(1)} \\ \hat{\phi}_l^{(2)} \\ \vdots \\ \hat{\phi}_l^{(j)} \end{bmatrix} = \begin{bmatrix} \hat{\phi}_l^{(0)} \\ \hat{\phi}_l^{(0)} \hat{\lambda}_l \\ \hat{\phi}_l^{(0)} \hat{\lambda}_l^2 \\ \vdots \\ \hat{\phi}_l^{(0)} \hat{\lambda}_l^j \end{bmatrix},$$

$$j = 0, \ldots, B, \qquad (36)$$

$$l = 1, \ldots, r$$

and is illustrated in Fig. 1(b), showing a zoom-in on the single augmented DMD mode $\hat{\phi}_l$. Alternatively, one can obtain the augmented DMD eigenvalue through (35) by

$$\hat{\lambda}_l = \left(\frac{\hat{\phi}_l^{(j)}}{\hat{\phi}_l^{(i)}} \right)^{\frac{1}{j-i}},$$

$$i, j = 0, \ldots, B, \ i \neq j, \qquad (37)$$

$$l = 1, \ldots, r.$$

Thus far, we claimed that the spatiotemporal coupling in Eqs. (34)–(37) holds for augmentation numbers bounded from

above by B. Now, we present this upper bound explicitly and show that it depends on the oscillation frequencies of the underlying continuous-time dynamical system, ω_l. The existence of the upper bound B stems from the Nyquist–Shannon sampling criterion[63,64] provided that the sampling frequency ω_s is sufficiently high, i.e., $\omega_l \leq 0.5\omega_s$ for every ω_l of the system.

By substituting $\Delta t = 2\pi/\omega_s$ into (13), the relationship between the DMD eigenvalue $\hat{\lambda}_l$ and the oscillation frequency ω_l is

$$\hat{\lambda}_l = \exp(\omega_l \Delta t) = \exp\left(\frac{\omega_l}{0.5\omega_s} \pi \right). \qquad (38)$$

Therefore, taking powers $p \in \mathbb{R}$ of $\hat{\lambda}_l$ results in

$$\hat{\lambda}_l^p = \left(\exp\left(\frac{\omega_l}{0.5\omega_s} \pi \right) \right)^p = \exp\left(\frac{p\omega_l}{0.5\omega_s} \pi \right). \qquad (39)$$

Namely, $\hat{\lambda}_l^p$ corresponds, in effect, to a frequency that is p-times faster than the continuous time counterpart of $\hat{\lambda}_l$, i.e., $p\omega_l$. So, the Nyquist–Shannon sampling criterion for this frequency is
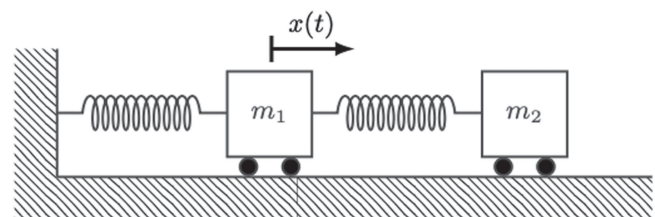

FIG. 2. Two degrees of freedom oscillator, comprised of two masses, m_1 and m_2, connected via two springs.
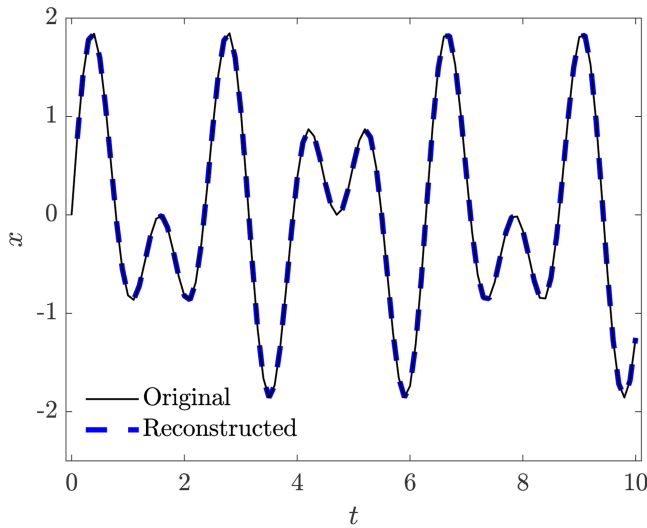
FIG. 3. The two-mode sine signal in Eq. (47), depicted as the black solid line (original signal). The blue dashed line depicts the reconstructed signal according to Algorithm 1 and Eq. (46).
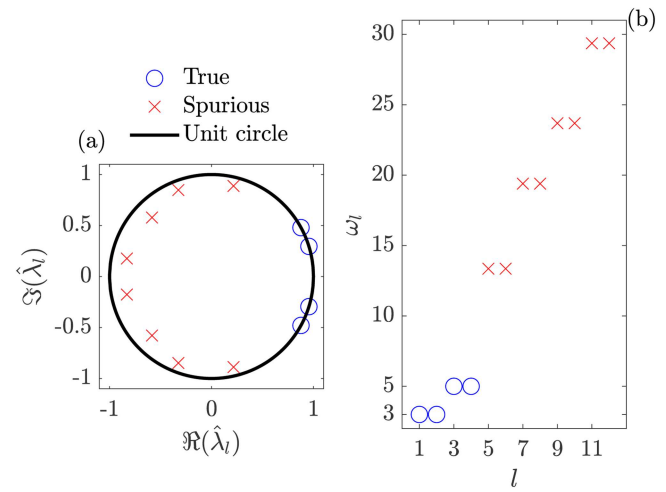

FIG. 4. (a) Polar representation of the augmented DMD eigenvalues, $\hat{\lambda}_l$, that arise from an application of the exact DMD algorithm to the observations $\{\boldsymbol{x}_k\}_{k=0}^{100}$ sampled from signal (47), which were augmented $s = 11$ times. The black line denotes the unit circle. (b) The oscillation frequencies, ω_l, that correspond to $\hat{\lambda}_l$ through Eq. (13). The *true* eigenvalues that are related to the system's harmonics ($\omega_1^{\text{sys}} = 3$ rad/s and $\omega_2^{\text{sys}} = 5$ rad/s) are marked by blue circles, while the red crosses mark the *spurious* eigenvalues that are not related to them.

$p\omega_l \le 0.5\omega_s$, which provides the constraint $p \le 0.5\omega_s/\omega_l$. Accordingly, we define

$$B(\omega_l) = \min\left(s, \left\lfloor \frac{0.5\omega_s}{\omega_l} \right\rfloor\right), \tag{40}$$

where $\lfloor \cdot \rfloor$ is the floor function.

In terms of Eq. (36) and as illustrated in Fig. 1(b), $\hat{\lambda}_l^j$ relates the zeroth sub-mode of $\hat{\boldsymbol{\phi}}_l$ to its jth sub-mode, where $\hat{\boldsymbol{\phi}}_l$ represents the lth DMD mode that corresponds to the oscillation frequency ω_l. Additionally, according to Eq. (39), $\hat{\lambda}_l^j$ represents a j-times faster

frequency than $\hat{\lambda}_l$, and is utilized by the spatiotemporal coupling in delay-coordinates DMD in Proposition 4. Consequently, this spatiotemporal coupling exists only for values of j that satisfy the Nyquist–Shannon sampling criterion (i.e., $j = 0, \ldots, B$). In other words, $\hat{\lambda}_l^j$ for $j > 0.5\omega_s/\omega_l$ represents a frequency that is sampled at a sub-Nyquist rate and, therefore, does not capture the underlying dynamics.
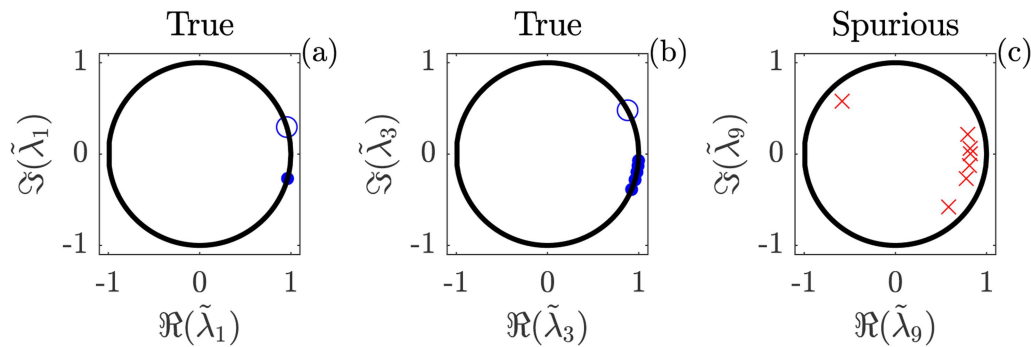

FIG. 5. Polar representation of the computed eigenvalues $\tilde{\lambda}_1^{(j)}, \tilde{\lambda}_3^{(j)}, \tilde{\lambda}_9^{(j)}, j = 1, \ldots, 11$, obtained from Eq. (41). The black lines denote the unit circle. According to the bound $B(\omega_l)$ in Eq. (40), which is valid only for *true* DMD components, the first 10 (a) and 6 (b) values of the *true* $\tilde{\lambda}_1^{(j)}$ and $\tilde{\lambda}_3^{(j)}$, respectively, are the same. These values are denoted by blue circles, which coincide with each other 10 times in (a) and 6 times in (b). Contrarily, the remaining eleventh value $\tilde{\lambda}_1^{(11)}$ in (a), and the five remaining values $\tilde{\lambda}_3^{(j)}$, $j = 7, \ldots, 11$ in (b), which are marked by blue dots, differ from one another and from the true $\hat{\lambda}_1$ or $\hat{\lambda}_3$. (c) Repeated computations of $\tilde{\lambda}_9^{(j)}$ that yield different values, since the relations in Proposition 4 do not hold for *spurious* DMD components.

We showed that the sub-modes within an augmented mode are related to each other through the corresponding eigenvalue [see Proposition 4, Eq. (35) and Fig. 1(b)]. We refer to these relations as the *spatiotemporal coupling in delay-coordinates DMD*. That is, an augmented DMD mode holds temporal information in addition to the spatial information. In addition, we show empirically in Sec. VI that the spurious components do not have this coupling. In Secs. V, VI, and VII, we utilize the spatiotemporal coupling for selection of the true DMD components and further representation and characterization of dynamical systems from observations, even when the observations are corrupted with noise.

V. PROPOSED METHOD

Based on the spatiotemporal coupling in delay-coordinates DMD presented in Proposition 4, we propose a method for obtaining a compact and reduced-order representation of the observations that is analogous to Eq. (8). As discussed in Sec. IV, representation (8) cannot be directly obtained from applications of DMD algorithms to delay-coordinates embedding due to two challenges: (a) the possible existence of both true and spurious augmented DMD components and (b) the existence of redundant sub-modes comprising every augmented DMD mode, which are unnecessary for the representation. Both challenges must be addressed to obtain the desired compact representation, i.e., the selection of the true augmented DMD modes, followed by the selection of the relevant sub-modes within the true modes.

To the best of our knowledge, challenge (a) is not addressed in the current literature in the context of delay-coordinates DMD. In the more general context of the Koopman operator, Pan *et al.* suggested a framework, which is based on multi-task feature learning, for the extraction of the most informative Koopman-invariant components and the pruning of the redundant components by penalizing departure from linear evolution.[21] A common practice to address challenge (b) is to choose the first n entries in each selected augmented DMD mode, ignoring its spatiotemporal coupling. Therefore, we propose a method that considers the spatiotemporal coupling in Proposition 4 and selects the DMD components
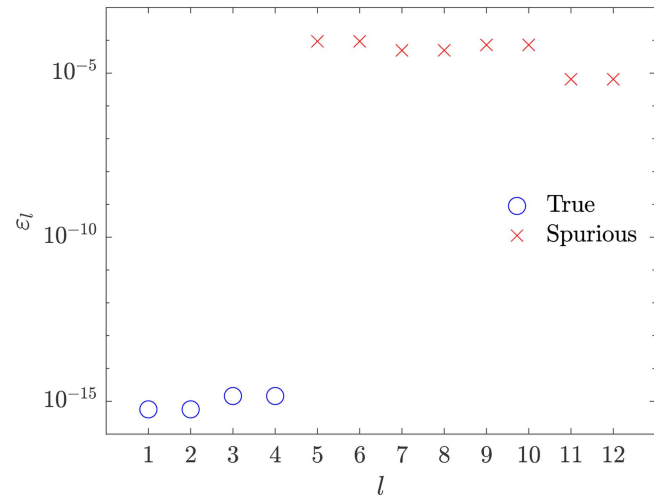

FIG. 6. The absolute errors, ε_l, between the averaged eigenvalues $\langle \tilde{\lambda}_l \rangle$ in Eq. (42) and their corresponding DMD eigenvalues, $\hat{\lambda}_l$ [see Eq. (43)]. The true eigenvalues, $\hat{\lambda}_1, \ldots, \hat{\lambda}_4$ are marked by blue circles, and the spurious eigenvalues, $\hat{\lambda}_5, \ldots, \hat{\lambda}_{12}$, are marked by red crosses. The absolute errors of the spurious $\hat{\lambda}_l$ are larger by at least 10 orders of magnitude than those of the true $\hat{\lambda}_l$, providing a clear distinction between the true and spurious DMD eigenvalues. The markings of the blue circles and red crosses were obtained in a unsupervised manner using Algorithm 1.

required for compact representations of dynamical systems. Our method is detailed in this section and summarized in Algorithm 1. An empirical comparison between our method and the maximal amplitudes method is conducted in Sec. VII, where the superiority of our method is demonstrated in case the observations are corrupted with high levels of noise.

Our method begins by augmenting the observations s times, and applying the exact DMD to them. Generally, the augmentation number s can be chosen either in the range prescribed in (19) or in

---

ALGORITHM 1. Proposed algorithm.

Input: Observations $\{x_k\}_{k=0}^m$ of a dynamical system.

Output: Compact and reduced-order spectral representation of the observations.

1. Augment the data s times according to the range in (20), apply the exact DMD algorithm to the augmented observations, and extract the augmented DMD eigenvalue-mode pairs, $\hat{\lambda}_l$ and $\hat{\boldsymbol{\phi}}_l$, respectively.
2. Calculate the oscillations frequencies ω_l that correspond to each augmented DMD eigenvalue $\hat{\lambda}_l$ using Eq. (13).
3. Compute the bound $B(\omega_l)$ in Eq. (40) for each ω_l.
4. Compute the averaged eigenvalues $\langle \tilde{\lambda}_l \rangle$ by Eq. (42).
5. Calculate the absolute errors, ε_l, between the averaged computed eigenvalues, $\langle \tilde{\lambda}_l \rangle$, and their corresponding (non-averaged) DMD eigenvalues, $\hat{\lambda}_l$, as in Eq. (43).
6. Partition ε_l into two subsets using k-means with $k = 2$, and compute the average values of each subset, as in (44). Select the averaged eigenvalues (and their corresponding modes) related to the subset corresponding to the smaller average value.
7. Use the selected DMD components for building a compact and reduced-order spectral representation of the observations, according to Eq. (46).

---

(20). Yet, the empirical evidence presented in Sec. VII shows that s should be chosen according to (20). Next, the DMD eigenvalue-mode pairs, $\hat{\lambda}_l$ and $\hat{\phi}_l$, respectively, are extracted, followed by a calculation of the oscillation frequencies ω_l by (13), as well as the bounds $B(\omega_l)$ that correspond to each ω_l by (40).

To identify the true components out of all the obtained DMD components, we propose to utilize Eq. (37), where the identification proceeds as follows. The same eigenvalue $\tilde{\lambda}_l$ is repeatedly computed via substitution of different i and j values in Eq. (37). For example, by setting $i = 0$ and $j = 1, \ldots, B(\omega_l)$, the eigenvalue $\tilde{\lambda}_l$ can be computed multiple times by taking the quotients of the sub-modes as

$$\tilde{\lambda}_l^{(j)} = \left(\frac{\hat{\phi}_l^{(j)}}{\hat{\phi}_l^{(0)}} \right)^{\frac{1}{j}}, \quad j = 1, \ldots, B(\omega_l), \qquad (41)$$

where a tilde denotes a computed entity and the superscripted index (j) in $\tilde{\lambda}_l^{(j)}$ denotes the jth computation of this eigenvalue. As stated in Proposition 4, the spatiotemporal coupling in delay-coordinates DMD holds only for true DMD components. Hence, any jth computation of any lth true eigenvalue according to (41) yields the same value (in a noise-free system). Moreover, this value is identical to

the DMD eigenvalue, i.e., $\tilde{\lambda}_l^{(j)} = \hat{\lambda}_l$, for $l = 1, \ldots, r$. Conversely, such computations of spurious DMD eigenvalues, for $l = r + 1, \ldots, \hat{r}$, are expected to yield different computed results upon substitution of varying values of j in (41), as well as results that differ from $\hat{\lambda}_l$. Therefore, (41) contains information that enables the distinction between true and spurious DMD components.

In the presence of observation noise, the identification of true DMD components through (41) can be enhanced by introducing an averaged computed eigenvalue, $\langle \tilde{\lambda}_l \rangle$, e.g., as

$$\langle \tilde{\lambda}_l \rangle = \frac{1}{B(\omega_l)} \sum_{j=1}^{B(\omega_l)} \tilde{\lambda}_l^{(j)}. \qquad (42)$$

In case $\omega_l > 0.5\omega_s$, then $B(\omega_l) = 0$. This scenario can occur, e.g., when s is chosen to be very large, which, in turn, might produce a very large ω_l. Since the sampling rate admits the Nyquist–Shannon sampling criterion for the true oscillation frequencies, such ω_l must be related to a spurious eigenvalue. Therefore, when $B(\omega_l) = 0$, we consider the eigenvalue corresponding to ω_l as spurious and remove it from the computation in (42).
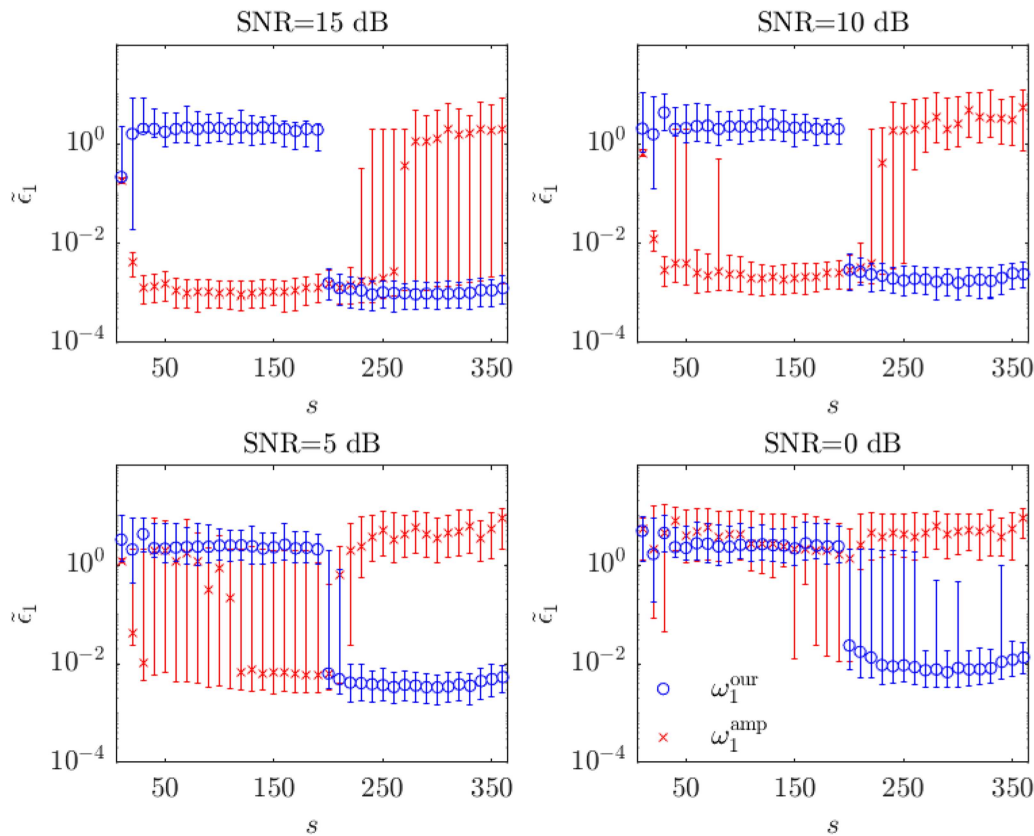

FIG. 7. The medians of the absolute errors, $\tilde{\epsilon}_1$, for augmentation numbers $s = 10, 20, \ldots, 360$ and SNR values of 15, 10, 5, 0 dB. The blue circles and red crosses denote $\tilde{\epsilon}_1$ computed based on our and the maximal amplitudes methods, respectively. The whiskers denote the 25th and 75th percentiles of ϵ_1.

A noteworthy drawback of DMD is its sensitivity to observation noise.[41,65] Different approaches, which are more sophisticated than Eq. (42), tackled this issue. For example, noise-corrected DMD, forward-backward DMD, total least-squares methods,[28,66] and variational formulations,[67] to name but a few. Overall, the sensitivity of DMD to noise-corrupted data still remains an open problem.[68]

Thereafter, the absolute errors, ε_l, are calculated by

$$\varepsilon_l = \left| \langle \tilde{\lambda}_l \rangle - \hat{\lambda}_l \right|, \tag{43}$$

providing an estimation for the difference between the augmented DMD eigenvalue and the average of the eigenvalues computed based on the spatiotemporal coupling in Proposition 4. Alternatively, the absolute errors can be defined as the average of the absolute differences between computed eigenvalues $\tilde{\lambda}_l^{(j)}$ and their corresponding DMD eigenvalue $\hat{\lambda}_l$. We note that this definition yields similar empirical results in the applications studied in this paper.

If the computed eigenvalues $\tilde{\lambda}_l^{(j)}$ for $j = 1, \ldots, B(\omega_l)$ are different from $\hat{\lambda}_l$, then ε_l in Eq. (43) is large, which means that $\hat{\lambda}_l$ is spurious. Contrarily, if $\tilde{\lambda}_l^{(j)}$ are approximately equal to $\hat{\lambda}_l$, then ε_l is small, indicating that $\hat{\lambda}_l$ is true. Therefore, ε_l, $l = 1, \ldots, \hat{r}$, can be

partitioned into two subsets, S_1 and S_2. Without loss of generality, assume that S_1 constitutes the first d values of ε_l, i.e., $\{\varepsilon_1, \ldots, \varepsilon_d\} \in S_1$, where $1 \leq d \leq \hat{r}$, and that S_2 contains the remaining $\hat{r} - d$ values, $\{\varepsilon_{d+1}, \ldots, \varepsilon_{\hat{r}}\} \in S_2$. This partitioning can be carried out by applying standard clustering algorithms to ε_l, e.g., k-means with $k = 2$.

We choose the k-means algorithm because, on the one hand, it is widely used, easy to implement, and does not assume any statistical model as a prior, and one the other hand, it is a centroid-based clustering algorithm. Namely, it is suitable to perform the task of separating two types of data points clustered around two centroids that are related to the true and spurious DMD components.

After clustering, the average absolute errors of subsets S_1 and S_2, $\langle \varepsilon \rangle_{S_1}$ and $\langle \varepsilon \rangle_{S_2}$ are computed as

$$\langle \varepsilon \rangle_{S_1} = \frac{1}{d} \sum_{l=1}^{d} \varepsilon_l, \quad \langle \varepsilon \rangle_{S_2} = \frac{1}{\hat{r} - d} \sum_{l=1}^{\hat{r}-d} \varepsilon_l, \tag{44}$$

and then, the smaller of $\langle \varepsilon \rangle_{S_1}$ and $\langle \varepsilon \rangle_{S_2}$ denotes the subset that contains the ε_l related to the true DMD eigenvalues. The distinction between small and large absolute errors was straightforward in the
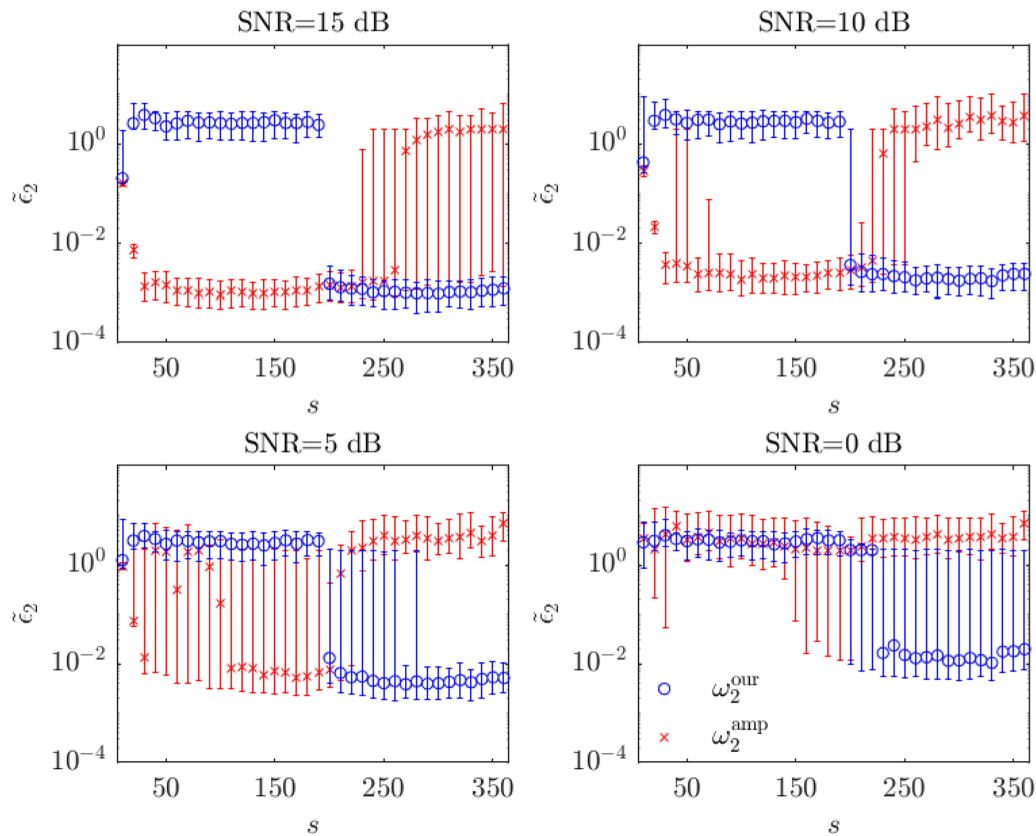

FIG. 8. The medians of the absolute errors, $\tilde{\epsilon}_2$, for augmentation numbers $s = 10, 20, \ldots, 360$ and SNR values of 15, 10, 5, 0 dB. The blue circles and red crosses denote $\tilde{\epsilon}_2$ computed based on our and the maximal amplitudes methods, respectively. The whiskers denote the 25th and 75th percentiles of ϵ_2.

illustrative example (see Sec. VI, Fig. 6). We note that in case the values of ε_l that are related to the two types of components (true and spurious) are small and relatively close to each other, the input of the k-means algorithm can be scaled, e.g., by considering $\log_{10}(\varepsilon_l)$ rather than ε_l; this scaling enhances the difference between the two types and alleviates their distinction via the k-means algorithm.

In some cases, the DMD components of dynamical systems arise in complex conjugate pairs. In such cases, each complex conjugate pair represents the same characteristics of the system (e.g., the same oscillation frequencies ω_l). These pairs embody another degree of redundancy, which can be exploited, for example, as follows. Denote $\hat{\lambda}_q$ as the complex conjugate of $\hat{\lambda}_l$. Then, similarly to (42), an averaged eigenvalue $\langle \tilde{\lambda}_{l,q} \rangle$, which accounts for this redundancy, can be written as

$$\langle \tilde{\lambda}_{l,q} \rangle = \frac{1}{2} \left(\langle \tilde{\lambda}_l \rangle + \overline{\langle \tilde{\lambda}_q \rangle} \right), \qquad (45)$$

where an overline denotes a complex conjugate. As $\langle \tilde{\lambda}_{l,q} \rangle$ is computed over both l and q, the influence of noise on the observations can be further diminished in these systems.

Lastly, once the subset of the true augmented DMD components is identified using Eq. (44), the observations can be represented analogously to (8) as

$$\boldsymbol{x}_k = \sum_{l=1}^{|S_1|} \langle \tilde{\lambda}_l \rangle^k \hat{\boldsymbol{\phi}}_l^{(0)} \hat{\sigma}_{0,l}, \qquad (46)$$

where $\hat{\sigma}_{0,l} = \langle \hat{\boldsymbol{\phi}}_l, \hat{\boldsymbol{x}}_0 \rangle$, $|S_1|$ is the cardinality of S_1, and, without loss of generality, we assume that S_1 is the subset that contains the ε_l values related to the true DMD eigenvalues $\langle \tilde{\lambda}_l \rangle$.

We note that our method has two main shortcomings, which we plan to address in future research. First, compared to mode selection based on maximal amplitudes, our method is computationally heavier. Second, as shown in Sec. VII, our empirical study suggests that the performance is sensitive to the choice of the augmentation number s. We plan to develop a systematic procedure to set the augmentation number, as well as more efficient implementation schemes that mitigate the repeated computation of eigenvalue decomposition. Additional future work may include the demonstration of our method on nonlinear and chaotic dynamical systems, as well as on real-life measurements.
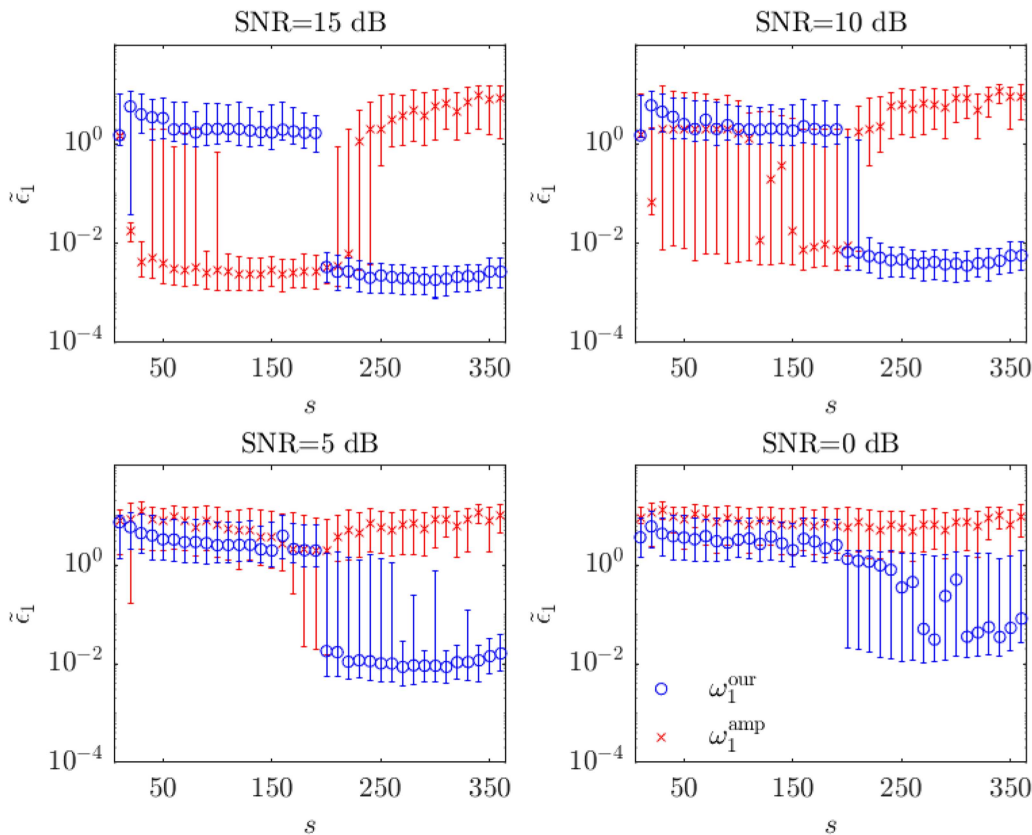

FIG. 9. Same as Fig. 7 but for ω_1 of the quasiperiodic signal.
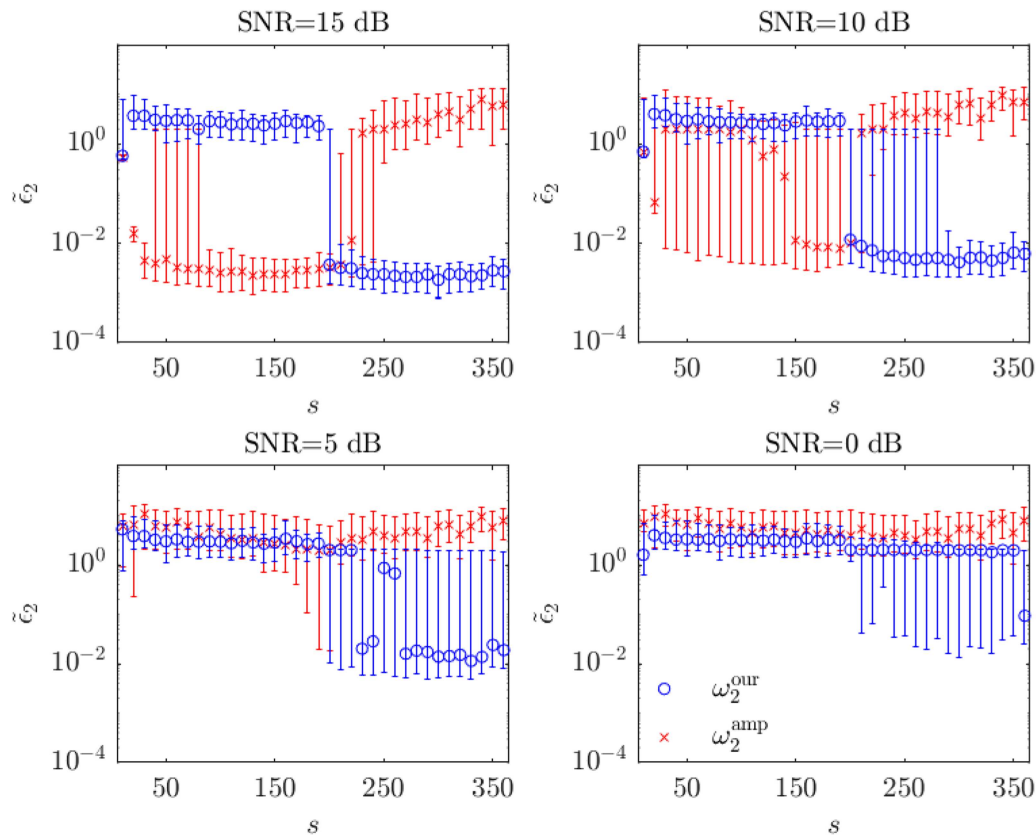
16 January 2024 08:57:34

FIG. 10. Same as Fig. 8 but for ω_2 of the quasiperiodic signal.

VI. ILLUSTRATIVE EXAMPLE

By following the proposed algorithm (Algorithm 1) step by step, we demonstrate the spatiotemporal coupling in delay-coordinates DMD (formulated in Sec. IV) on an illustrative example of a two-mode sine signal. More specifically, we show how Proposition 4 facilitates the identification, in an equation-free manner, of the DMD components required for the representation and characterization of the signal. Evaluation is done by comparing the true signal's oscillation frequencies to the oscillation frequencies that correspond to the DMD eigenvalues selected by our method. Our method can also be visually evaluated through its signal reconstruction compared to the original signal, as shown in Fig. 3.

Consider the following two-mode sine signal:

$$x(t) = \sin(\omega_1^{\text{sys}} t) + \sin(\omega_2^{\text{sys}} t) + \alpha n(t), \qquad (47)$$

where $\omega_1^{\text{sys}} = 3$ rad/s, $\omega_2^{\text{sys}} = 5$ rad/s, $n(t)$ is an additive white standard Gaussian noise, and $\alpha = 10^{-12}$. Figure 2 shows a dynamical system whose coordinate $x(t)$ can represent such a signal. Specifically, signal (47) can be viewed as the position of mass m_1 in a two degrees of freedom (DOF) oscillator, comprised of two masses, m_1 and m_2, connected to each other via springs. We denote the

components that correspond to the system by the superscript sys, i.e., ω_1^{sys} and ω_2^{sys} are the system's harmonics.

The signal is sampled at the sampling rate $\omega_s = 2\pi/\Delta t = 20\pi$ rad/s (corresponding to 10 Hz) and 101 samples (observations) $\{x_k\}_{k=0}^{100}$ are collected, shown by the solid black line in Fig. 3.

We begin the illustration by augmenting the observations with $s = 11$. This choice of s is in accordance with the upper bound on s in (20), yet it does not adhere to the lower bound in the same equation. One of the main purposes of this section is to demonstrate
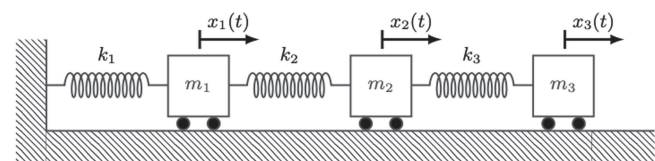

FIG. 11. Three degrees of freedom oscillator, comprised of three masses $m_1 = 1$ kg, $m_2 = 2$ kg, $m_3 = 3$ kg, connected via three springs of constants $k_1 = k_2 = 50$ N/m, $k_3 = 75$ N/m. The system is subjected to the initial conditions $x_1(0) = 1$ m, $x_2(0) = 2$ m, $x_3(0) = 3$ m, $\dot{x}_1(0) = \dot{x}_2(0) = \dot{x}_3(0) = 0$, giving rise to its oscillations.

that Proposition 4 holds for true DMD components and up to the bound $B(\omega_l)$. Consequently, we choose $s = 11$, which is a value that, on the one hand, demonstrates these two statements for system (47), and on the other hand, provides a small enough number of eigenvalues that can be conveniently visualized. That is, this choice is made merely for illustrative purposes.

The exact DMD algorithm is applied to the augmented observations, resulting in $s + 1 = 12$ augmented DMD eigenvalue-mode pairs, $\{\hat{\lambda}_l, \hat{\boldsymbol{\phi}}_l\}_{l=1}^{12}$. Figure 4(a) presents the polar representation of $\hat{\lambda}_l$, and Fig. 4(b) shows the oscillation frequencies, ω_l (corresponding to the eigenvalues by Eq. (13)), sorted in ascending order. It can be observed in Fig. 4(a) that the DMD eigenvalues of this signal arise in complex conjugate pairs. Moreover, these eigenvalues are distinct, satisfying the condition of Propositions 2 and 4.

Since the signal is composed of two oscillation frequencies, ω_1^{sys} and ω_2^{sys}, two complex conjugate DMD eigenvalue pairs $\hat{\lambda}_{1,2}$ and $\hat{\lambda}_{3,4}$ are related to these frequencies—these are the *true* eigenvalues and they are marked by blue circles in Fig. 4. Indeed, Fig. 4(b) corroborates them as true since they correspond to the frequencies 3 and 5 rad/s. Similarly, the remaining eight eigenvalues (and their corresponding modes) are *spurious* and are marked by red crosses in Fig. 4.

Our goal of identifying the true DMD components required for the system description in an equation-free manner can be visually described via Fig. 4. Suppose the observations $\{x_k\}_{k=0}^{100}$ were obtained from an unknown system, i.e., without the knowledge that 3 and 5 rad/s are the system's harmonics. Then, the markings of the blue circles and red crosses would be unknown as well. In this regard, our goal would be the task of identifying the blue circles, which represent the true DMD eigenvalues.

By Eq. (40), the bounds that correspond to the two system's harmonics are computed to be $B(\omega_1^{sys}) = 10$ and $B(\omega_2^{sys}) = 6$. Next, we demonstrate Proposition 4 by showing that the spatiotemporal coupling indeed holds up to these bounds. The choice of $s = 11$ is the minimal value of s that can serve this purpose. As an example, we consider the true eigenvalues $\hat{\lambda}_1$ and $\hat{\lambda}_3$, as well as the spurious eigenvalue $\hat{\lambda}_9$. For each of them, we follow the eigenvalues computation in Eq. (41) eleven times; namely, we substitute $j = 1, \ldots, 11$ for each of the three values $l = 1, 3, 9$ in Eq. (41). Figure 5 presents the polar representations of these computed eigenvalues.

Figure 5(a) shows $\tilde{\lambda}_1^{(1)}, \ldots, \tilde{\lambda}_1^{(10)}$ as blue circles and $\tilde{\lambda}_1^{(11)}$ as a blue dot. As expected and according to $B(\omega_1^{sys}) = 10$, the former 10 computed eigenvalues coincide with each other, and appear as a single circle in Fig. 5(a). Moreover, a comparison between them and
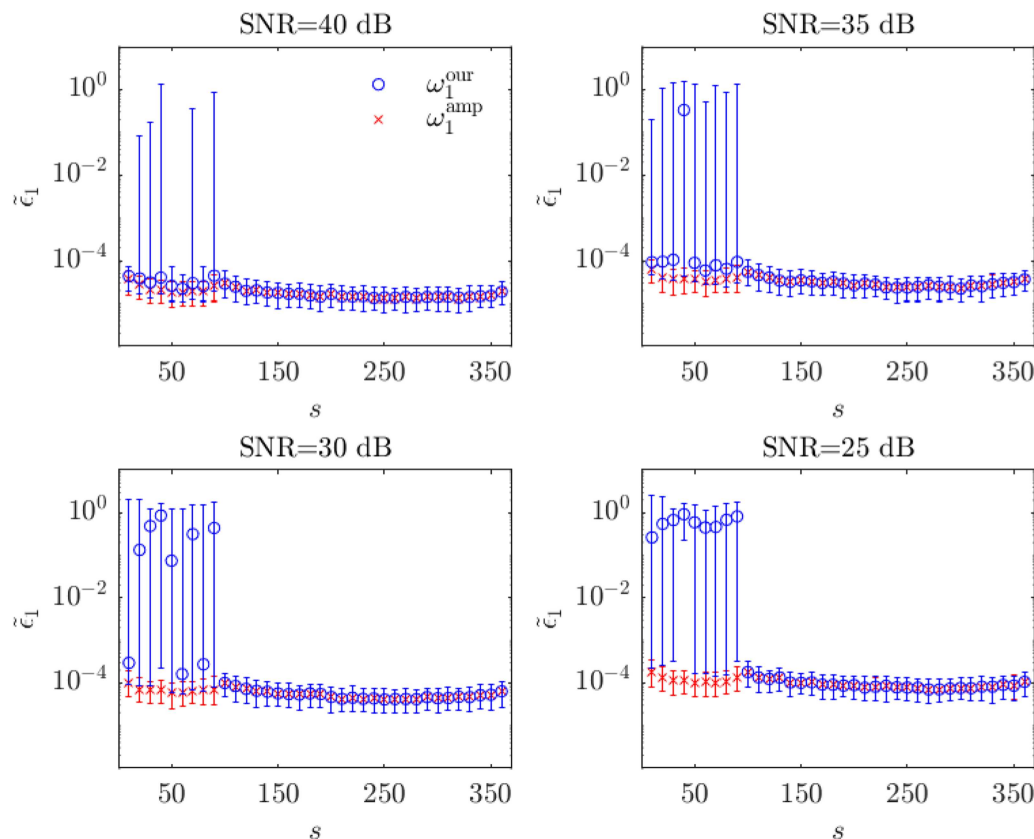

FIG. 12. Same as Fig. 7 but for ω_1 of the 3 DOF oscillator.

their corresponding true DMD eigenvalue, $\hat{\lambda}_1$, reveals that $\tilde{\lambda}_1^{(1)} = \ldots = \tilde{\lambda}_1^{(10)} = \hat{\lambda}_1$. Contrarily, the 11th computed eigenvalue, $\tilde{\lambda}_1^{(11)}$, differs from them, as expected, since its index exceeds the upper bound $B(\omega_1^{sys}) = 10$.

Similarly, Fig. 5(b) shows $\tilde{\lambda}_3^{(1)}, \ldots, \tilde{\lambda}_3^{(6)}$ as blue circles and $\tilde{\lambda}_3^{(7)}, \ldots, \tilde{\lambda}_3^{(11)}$ as blue dots. As $B(\omega_2^{sys}) = 6$, the former 6 computed eigenvalues coincide with each other and appear as a single circle in Fig. 5(b). Further comparison confirms that $\tilde{\lambda}_3^{(1)} = \cdots = \tilde{\lambda}_3^{(6)} = \hat{\lambda}_3$. On the other hand, $\tilde{\lambda}_3^{(7)} \neq \cdots \neq \tilde{\lambda}_3^{(11)} \neq \hat{\lambda}_3$. That is, relations (41) are indeed valid for true DMD components and up to $B(\omega_l)$.

Finally, Fig. 5(c), which relates to the spurious DMD eigenvalue $\hat{\lambda}_9$, depicts the 11 computed eigenvalues $\tilde{\lambda}_9^{(1)}, \ldots, \tilde{\lambda}_9^{(11)}$, which are different from each other. This difference indicates that $\hat{\lambda}_9$ is spurious, as relations (41) are not valid for spurious DMD components.

To conclude the demonstration in Fig. 5, we showed that only *true* DMD components adhere to the spatiotemporal coupling in delay-coordinates DMD in Proposition 4 and only up the bound $B(\omega_l)$ in Eq. (40).

Next, we exploit this spatiotemporal coupling in order to identify the true DMD components required for compact representation of the dynamical system. Accordingly, we compute all the averaged

eigenvalues, $\langle \tilde{\lambda}_l \rangle, l = 1, \ldots, 12$, via Eq. (42), followed by a calculation of their absolute errors, ε_l, by Eq. (43). Figure 6 shows ε_l for each of the 12 eigenvalues in this illustrative example, where $\hat{\lambda}_1, \ldots, \hat{\lambda}_4$ are the true eigenvalues marked by blue circles, and the rest are the spurious eigenvalues marked by red crosses. Evidently, the absolute errors of $\hat{\lambda}_1, \ldots, \hat{\lambda}_4$ are smaller by at least 10 orders of magnitude than the absolute errors of the spurious eigenvalues. Consequently, the true and spurious eigenvalues can be easily distinguished by visual inspection based on Fig. 6.

To implement this identification in an unsupervised fashion, ε_l are clustered into two subsets, S_1 and S_2, using the k-means algorithm with $k = 2$. The resulting clusters are $\{\varepsilon_1, \ldots, \varepsilon_4\} \in S_1$ and $\{\varepsilon_5, \ldots, \varepsilon_{12}\} \in S_2$. Next, the respective average absolute errors of S_1 and S_2 are computed, namely, $\langle \varepsilon \rangle_{S_1}$ and $\langle \varepsilon \rangle_{S_2}$, according to (44). In this example, $\langle \varepsilon \rangle_{S_1} \approx 10^{-15}$ and $\langle \varepsilon \rangle_{S_2} \approx 10^{-3}$. Therefore, subset S_1, corresponding to the smaller value $\langle \varepsilon \rangle_{S_1} < \langle \varepsilon \rangle_{S_2}$ is identified as the subset containing the ε_l that are related to the true DMD eigenvalues.

Based on this identification, the original (sampled) signal is reconstructed using Eq. (46). The accuracy of the reconstruction can be observed in Fig. 3, which shows the original (black solid line) and the reconstructed (blue dashed line) signals. We note that the
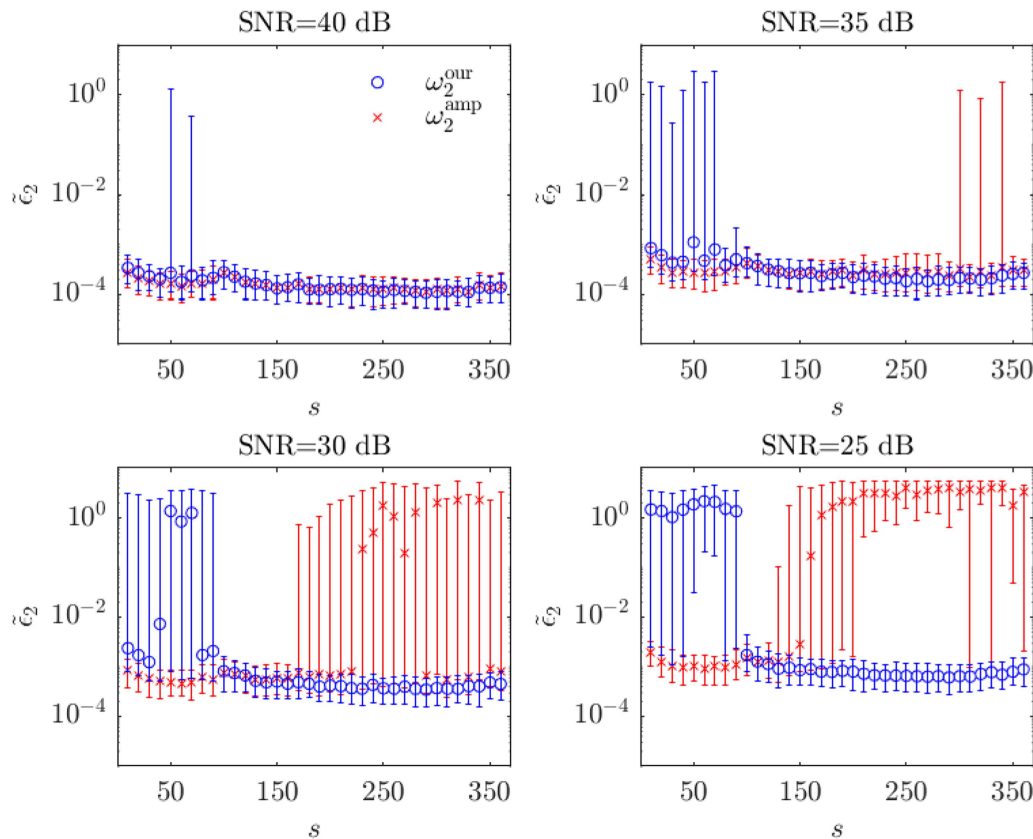

FIG. 13. Same as Fig. 7 but for ω_2 of the 3 DOF oscillator.

markings of the blue circles and red crosses in Fig. 6, as well as the reconstructed signal in Fig. 3, were obtained in a completely unsupervised and automatic manner using Algorithm 1.

To conclude, we emphasize that the input of our algorithm is the samples (observations) of the system, where it operates in an *equation-free* and *unsupervised* fashion to, eventually, extract a reduced-order and optimal spectral representation of the system.

The code that reproduces the results in this section is openly available in the following GitHub link (https://github.com/emilbron stein/spatiotemporal_coupling_in_delay_coordinates_DMD).

VII. SIMULATION RESULTS

A. Two-mode sine signal

Consider the two-mode sine signal in Eq. (47). We examine noise with four different amplitudes α that correspond to the SNRs 15, 10, 5, and 0 dB. We test the performance of the proposed method (detailed in Sec. V and summarized in Algorithm 1) for the augmentation numbers $s = 10, 20, \ldots, 360$ and compare its DMD components selection to the selection method based on maximal amplitudes. For comparison purposes, we focus on the accuracy of the recovery of the oscillation frequencies of the signal obtained

by the two methods. For a fair comparison, we set the number of selected DMD components to 4 (2 oscillation frequencies that appear in 2 complex conjugate pairs correspond to 4 DMD eigenvalues). Note that, in general, our method does not require the number of DMD components as a prior, but infers it from the observations.

The recovery of the oscillation frequencies is evaluated by the absolute errors, ϵ_i, between the system's true oscillation frequencies (ω_1^{sys} and ω_2^{sys}) and the frequencies obtained from the DMD eigenvalues that were selected based on our and the maximal amplitudes methods, denoted by the superscripts *our* and *amp*, respectively. Namely, we calculate

$$\epsilon_i^{\mathrm{our}} = \left| \omega_i^{\mathrm{sys}} - \omega_i^{\mathrm{our}} \right|, \quad \epsilon_i^{\mathrm{amp}} = \left| \omega_i^{\mathrm{sys}} - \omega_i^{\mathrm{amp}} \right|. \quad (48)$$

We repeat the calculations of ϵ_i 500 times, each time with a different generated random noise, and compute the medians of the absolute errors, $\tilde{\epsilon}_i$, as well as their 25th and 75th percentiles. Figure 7 considers the first oscillation frequency, $\omega_1^{\mathrm{sys}} = 3$ rad/s, and presents $\tilde{\epsilon}_1$ for different augmentation numbers s at different SNR values, where the 25th and 75th percentiles of ϵ_1 are denoted by whiskers. Figure 8 shows the same analysis for $\omega_2^{\mathrm{sys}} = 5$ rad/s.

Figures 7 and 8 show that at the larger SNR values (15 and 10 dB), both ω_1^{sys} and ω_2^{sys} can be recovered with small errors based
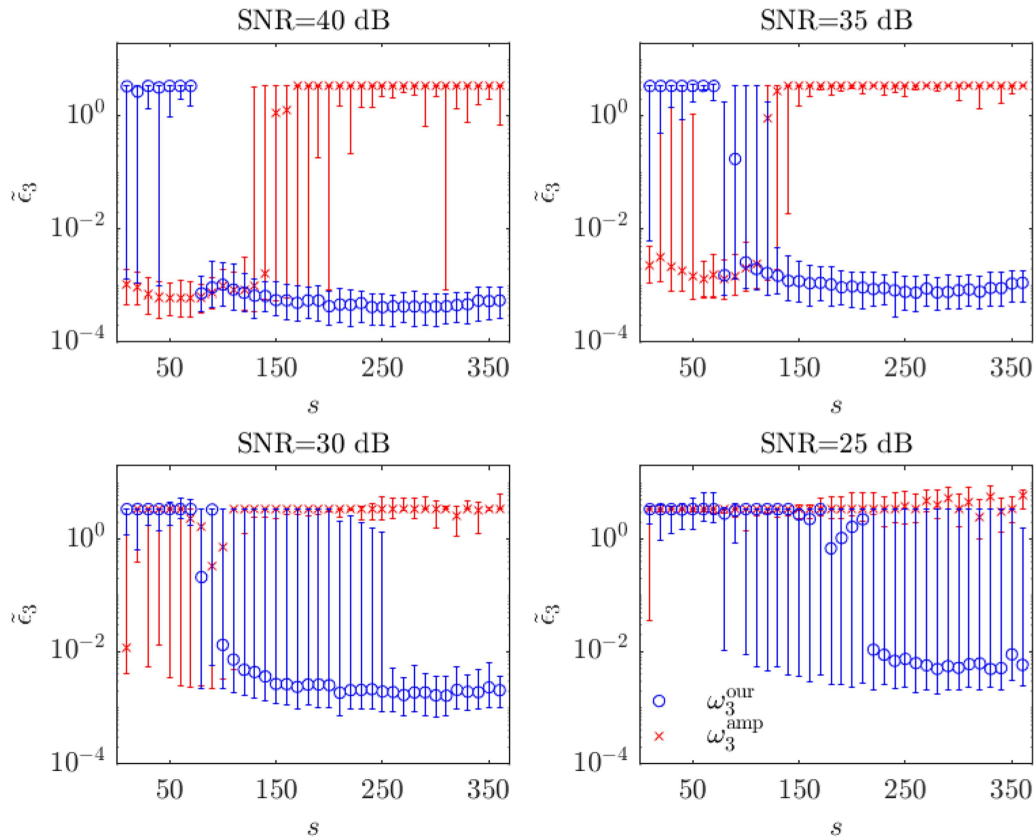
FIG. 14. Same as Fig. 7 but for ω_3 of the 3 DOF oscillator.

ω_1^{sys} and ω_2^{sys} with reasonable errors; namely, the frequencies are obtained with either large error median or large error variance of ϵ_i. Contrarily, our method recovers these frequencies with small errors when the values of s are set in range (20). For SNR value of 0 dB, the maximal amplitudes method leads to large error medians, whereas our method obtains small error medians when the values of s are set in range (20).

B. Quasiperiodic signal

Consider the following noisy quasiperiodic signal

$$x(t) = \sin\left(\sqrt{10}t\right)\sin(t) + \alpha n(t), \qquad (49)$$

which can be recast as

$$x(t) = \frac{1}{2}\cos\left((\sqrt{10}-1)t\right) - \frac{1}{2}\cos\left((\sqrt{10}+1)t\right) + \alpha n(t), \quad (50)$$

where $n(t)$ is a white standard Gaussian noise with varying amplitudes α that correspond to SNR value of 15, 10, 5, and 0 dB.

Similar to the two-mode sine signal simulation (Sec. VII A), we test the capability to uncover the signal's irrational frequencies in (50) using our method and compare it with the maximal amplitudes method. Figures 9 and 10 are the same as Figs. 7 and 8 but
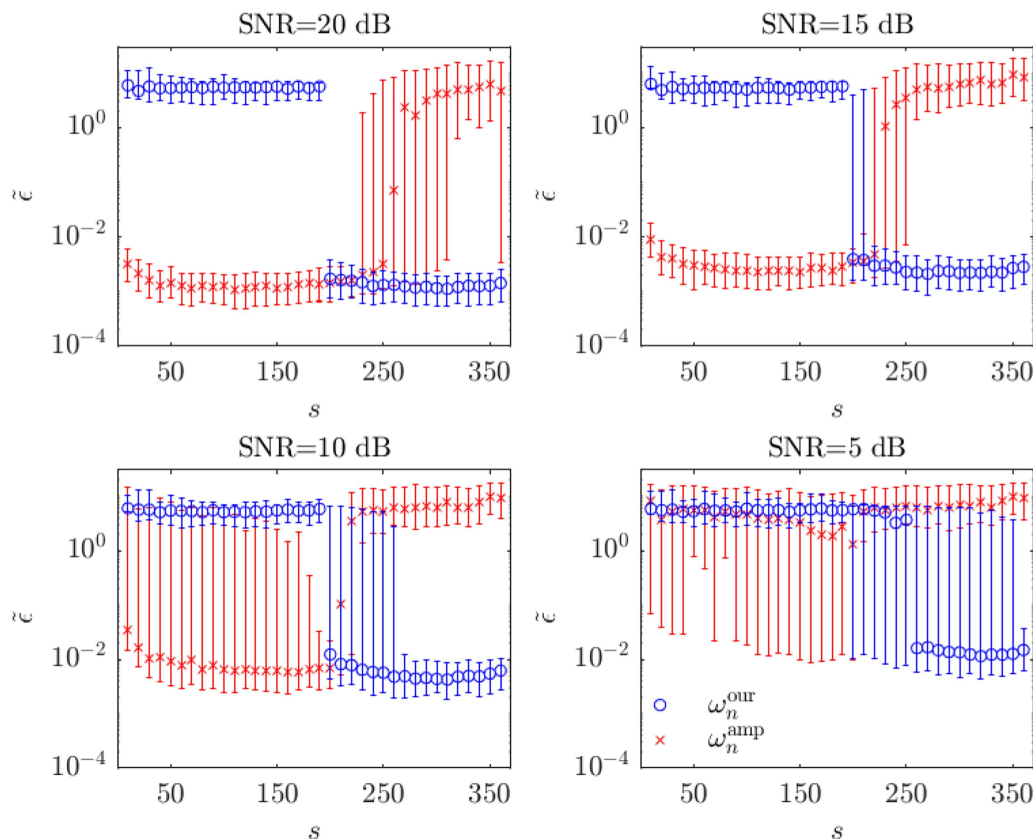

FIG. 15. A damped oscillator comprised of mass $m = 1$ kg, which is connected to a wall via a spring of stiffness $k = 49$ N/m, and a damper with a damping coefficient $c = 0.07$ Ns/m. The mass is subjected to the initial conditions $x(0) = 1$ m, $\dot{x}(0) = 2$ m/s, giving rise to its oscillations.

on both methods, given that s is chosen according to the range indicated in (20) for our method, and in (19) for the maximal amplitudes method.

Yet, this is not the case for the smaller SNR values. For SNR value of 5 dB, the maximal amplitudes method is unable to recover

FIG. 16. Same as Fig. 7 but for ω_n of the damped oscillator.

16 January 2024 08:57:34

with respect to $\omega_1^{sys} = (\sqrt{10} - 1)$ rad/s and $\omega_2^{sys} = (\sqrt{10} + 1)$ rad/s, respectively.

For both ω_1^{sys} and ω_2^{sys}, the maximal amplitudes method performs well only at SNR value of 15 dB. However, at the other three examined SNR values, the system's frequencies are uncovered with a large error variance (10 dB) or a large error median (5 and 0 dB). Contrarily, our method uncovers ω_1^{sys} with small errors at SNR = 15, 10, and 5 dB, and with small error median but with large variance at 0 dB. Also, using our method, ω_2^{sys} is uncovered with small errors at SNR = 15 and 10 dB, and small error median but with large variance at 5 dB. At SNR = 0 dB, both methods fail to uncover ω_2^{sys}.

C. Multi degrees of freedom oscillator

Consider an oscillator comprised of N masses $m_1, \ldots, m_N$ connected by springs of constants $k_1, \ldots, k_N$. The first mass m_1 is connected to a wall via the first spring k_1, as depicted in Fig. 11 for $N = 3$. The equations of motion of this system are given by

$$M\ddot{x}(t) + Kx(t) = 0,$$

where

$$M = \begin{bmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_N \end{bmatrix},$$

$$K = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 & \cdots & 0 \\ -k_2 & k_2 + k_3 & -k_3 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -k_N & k_N \end{bmatrix},$$

and $x = [x_1, x_2, \ldots, x_N]^T$ are the positions of the masses.

To excite motion, the system is perturbed from its rest state by moving each mass to some initial position and then releasing them all at once. Namely, the system is subjected to the initial conditions $x(t = 0) = [x_1^0, x_2^0, \ldots, x_N^0]^T$, as well as to zero initial velocities of all masses ($\dot{x}(t = 0) = 0$).

Here, we consider the 3 DOF oscillator in Fig. 11 ($N = 3$) with the masses $m_1 = 1$ kg, $m_2 = 2$ kg, $m_3 = 3$ kg, spring constants $k_1 = k_2 = 50$ N/m, $k_3 = 75$ N/m, initial positions $x_1(0) = 1$ m, $x_2(0) = 2$ m, $x_3(0) = 3$ m, and zero initial velocities $\dot{x}_1(0) = \dot{x}_2(0)$
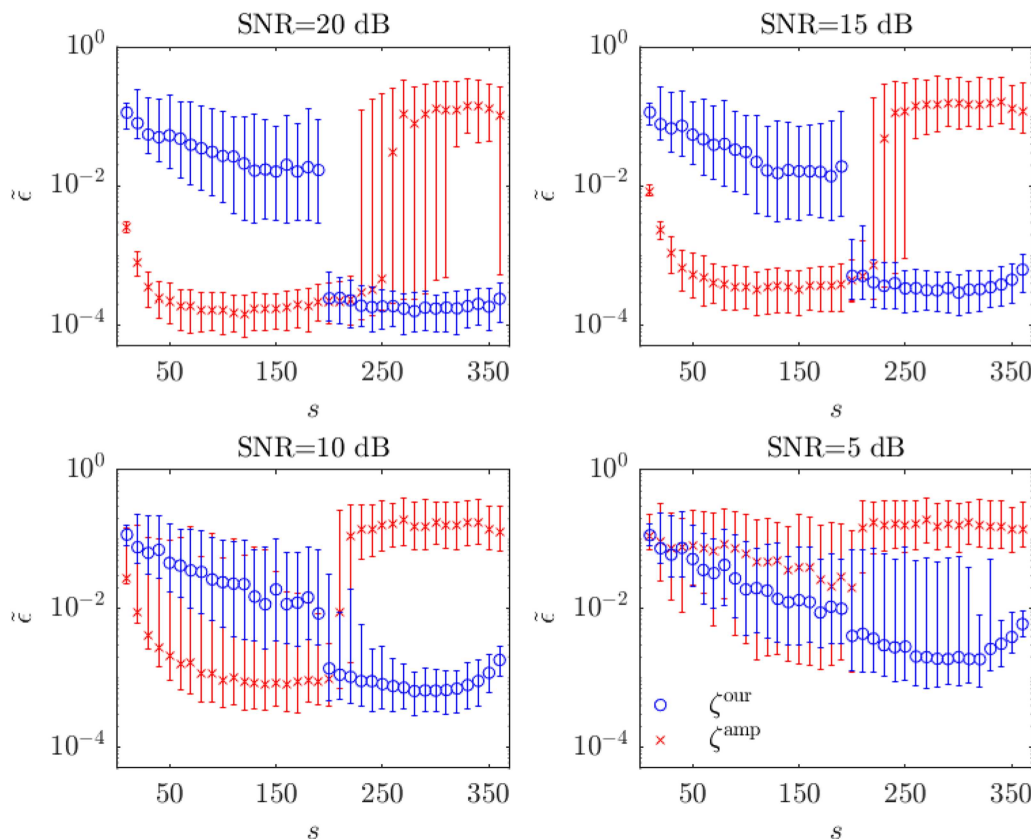

FIG. 17. Same as Fig. 7 but for ζ of the damped oscillator.

$= \dot{x}_3(0) = 0$. In this case, the system's natural frequencies are[69]

$$\omega_1^{sys} = 2.0626 \, rad/s,$$
$$\omega_2^{sys} = 7.6961 \, rad/s, \quad (51)$$
$$\omega_3^{sys} = 11.1363 \, rad/s.$$

Random noise drawn from a normal distribution with amplitudes that correspond to SNR values of 40, 35, 30, and 25 dB is added to the observations. Note that in this case, the SNR is computed with respect only to the first mass. Then, same as in Fig. 7, the accuracy of the recovery of frequencies (51) by our and the maximal amplitudes methods is tested and presented in Figs. 12–14.

At all the examined SNR values, ω_1^{sys} and ω_2^{sys} can be recovered with small errors using both methods. These results are similar for ω_3^{sys} at SNR values of 40 and 35 dB. However, the maximal amplitudes method fails to uncover ω_3^{sys} with reasonable errors at SNR value of 30 dB, while our method successfully performs this task. At SNR value of 25 dB, the maximal amplitudes method is unable to uncover ω_3^{sys} as well, whereas our method uncovers this frequency with small error median and large error variance.

D. Damped oscillator

Consider a damped oscillator with a mass $m = 1 \, kg$, connected to a wall via a spring with constant $k = 49 \, N/m$ and a damper with damping coefficient $c = 0.07 \, Ns/m$, as illustrated in Fig. 15. These system parameters correspond to the natural frequency $\omega_n = 7 \, rad/s$, damping ratio $\zeta = 0.005$, and the frequency of damped vibrations $\omega_d = \omega_n \sqrt{1 - \zeta^2} = 6.9999 \, rad/s$. The mass is subjected to the initial conditions $x(0) = x_0 = 1 \, m$, $\dot{x}(0) = v_0 = 2 \, m/s$, giving rise to oscillations, which induce the following time response:[69]

$$x(t) = \exp(-\zeta \omega_n t) \left(\frac{v_0 + \zeta \omega_n x_0}{\omega_d} \sin(\omega_d t) + x_0 \cos(\omega_d t) \right). \quad (52)$$

Random noise drawn from a normal distribution with amplitudes that correspond to SNR values of 20, 15, 10, and 5 dB is added to (52). The recovery accuracy of ω_n and ζ obtained by our method and the maximal amplitudes method is presented in Figs. 16 and 17, respectively.

At SNR values of 20 and 15 dB, ω_n and ζ are recovered with small errors using both methods. At SNR values of 10 and 5 dB, the maximal amplitudes method is unable to recover both ω_n and ζ with reasonable errors. Conversely, at these noise levels, our method yields a recovery with errors at the order of 10^{-2} for ω_n and 10^{-3} for ζ.

VIII. CONCLUSIONS

Delay-coordinates DMD is widely used for data-driven analysis of dynamical systems based on observations in a broad range of fields. In this paper, we investigated two key questions concerning the inherent redundancy that arises from the utility of delay-coordinates DMD: the excess of dynamical components (i.e., spurious components) and the excess of dimensionality (i.e., coordinates). We showed that delay-coordinates DMD induces a particular

structure on the augmented DMD components, consisting of a spatiotemporal coupling. At first glance, this coupling seems to counter the core idea underlying DMD, which facilitates a representation of the system that decouples temporal and spatial patterns. Yet, a deeper look allowed us not only to mitigate this coupling, but also to exploit it. Based on the spatiotemporal coupling we presented, we proposed a method for constructing a compact and improved reduced-order DMD representation. Specifically, we showed how to identify and select the informative (true) DMD components, thereby addressing the excess of dynamical components. This identification is based on the induced temporal associations within each augmented mode, which allowed us to address the redundancy in dimensionality. We tested the proposed method on four dynamical systems corrupted with noise and compared the performance to the prevalent method, which is based on the maximal amplitudes of the DMD modes. The results demonstrate the advantages of the proposed method and imply that exploiting the spatiotemporal structure becomes crucial at low SNR values (high levels of noise) and gives a significant advantage over the maximal amplitudes method, which is completely agnostic to it.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Emil Bronstein: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Software (lead); Validation (equal); Visualization (equal); Writing – original draft (lead). **Aviad Wiegner:** Conceptualization (equal); Data curation (supporting); Formal analysis (supporting); Investigation (supporting); Methodology (equal); Validation (equal). **Doron Shilo:** Conceptualization (supporting); Funding acquisition (equal); Investigation (supporting); Methodology (supporting); Supervision (equal); Validation (equal); Writing – review & editing (supporting). **Ronen Talmon:** Conceptualization (equal); Data curation (supporting); Formal analysis (lead); Funding acquisition (equal); Investigation (supporting); Methodology (supporting); Supervision (equal); Validation (equal); Visualization (equal); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

16 January 2024 08:57:34

REFERENCES

[1] I. G. Kevrekidis, C. W. Gear, and G. Hummer, "Equation-free: The computer-aided analysis of complex multiscale systems," AIChE J. **50**, 1346–1355 (2004).

[2] I. G. Kevrekidis and G. Samaey, "Equation-free multiscale computation: Algorithms and applications," Annu. Rev. Phys. Chem. **60**, 321–344 (2009).

[3] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, "Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis," Commun. Math. Sci. **1**, 715–762 (2003).

[4] S. Sirisup, G. E. Karniadakis, D. Xiu, and I. G. Kevrekidis, "Equation-free/Galerkin-free POD-assisted computation of incompressible flows," J. Comput. Phys. **207**, 568–587 (2005).

[5] C. Theodoropoulos, Y.-H. Qian, and I. G. Kevrekidis, "'Coarse' stability and bifurcation analysis using time-steppers: A reaction-diffusion example," Proc. Natl. Acad. Sci. U.S.A. **97**, 9840–9843 (2000).

[6] C. W. Gear, I. G. Kevrekidis, and C. Theodoropoulos, "'Coarse' integration/bifurcation analysis via microscopic simulators: Micro-Galerkin methods," Comput. Chem. Eng. **26**, 941–963 (2002).

[7] M. Budišić, R. Mohr, and I. Mezić, "Applied Koopmanism," Chaos **22**, 047510 (2012).

[8] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," Phys. Rev. Lett. **59**, 845 (1987).

[9] T. Sauer, "Time series prediction by using delay coordinate embedding," in *Time Series Prediction: Forecasting the Future and Understanding the Past* (Routledge, 1994), available at https://www.taylorfrancis.com/books/mono/10.4324/9780429492648/time-series-prediction-andreas-weigend.

[10] D. Kugiumtzis, O. Lingjærde, and N. Christophersen, "Regularized local linear prediction of chaotic time series," Phys. D **112**, 344–360 (1998).

[11] R. Talmon and R. R. Coifman, "Empirical intrinsic geometry for nonlinear modeling and time series filtering," Proc. Natl. Acad. Sci. U.S.A. **110**, 12535–12540 (2013).

[12] C. J. Dsilva, R. Talmon, C. W. Gear, R. R. Coifman, and I. G. Kevrekidis, "Data-driven reduction for a class of multiscale fast-slow stochastic dynamical systems," SIAM J. Appl. Dyn. Syst. **15**, 1327–1351 (2016).

[13] M. Crosskey and M. Maggioni, "Atlas: A geometric approach to learning high-dimensional stochastic systems near manifolds," Multiscale Model. Simul. **15**, 110–156 (2017).

[14] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," Proc. Natl. Acad. Sci. U.S.A. **17**, 315 (1931).

[15] I. Mezić, "Analysis of fluid flows via spectral properties of the Koopman operator," Annu. Rev. Fluid Mech. **45**, 357–378 (2013).

[16] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems* (SIAM, 2016).

[17] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," Phys. D **197**, 101–133 (2004).

[18] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," Nonlinear Dyn. **41**, 309–325 (2005).

[19] H. Arbabi and I. Mezic, "Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator," SIAM J. Appl. Dyn. Syst. **16**, 2096–2126 (2017).

[20] D. Giannakis, "Data-driven spectral decomposition and forecasting of ergodic dynamical systems," Appl. Comput. Harmonic Anal. **47**, 338–396 (2019).

[21] S. Pan, N. Arnold-Medabalimi, and K. Duraisamy, "Sparsity-promoting algorithms for the discovery of informative Koopman-invariant subspaces," J. Fluid Mech. **917**, A18 (2021).

[22] S. Das and D. Giannakis, "Koopman spectra in reproducing kernel Hilbert spaces," Appl. Comput. Harmonic Anal. **49**, 573–607 (2020).

[23] P. J. Schmid and J. Sesterhenn, "Dynamic mode decomposition of experimental data," in 8th International Symposium on Particle Image Velocimetry, Melbourne, Victoria, Australia (2009).

[24] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," J. Fluid Mech. **656**, 5–28 (2010).

[25] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," J. Fluid Mech. **641**, 115–127 (2009).

[26] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition," J. Nonlinear Sci. **25**, 1307–1346 (2015).

[27] K. K. Chen, J. H. Tu, and C. W. Rowley, "Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analyses," J. Nonlinear Sci. **22**, 887–915 (2012).

[28] S. Dawson, M. S. Hemati, M. O. Williams, and C. W. Rowley, "Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition," Exp. Fluids **57**, 1–19 (2016).

[29] M. R. Jovanović, P. J. Schmid, and J. W. Nichols, "Sparsity-promoting dynamic mode decomposition," Phys. Fluids **26**, 024103 (2014).

[30] T. Sayadi, P. J. Schmid, F. Richecoeur, and D. Durox, "Parametrized data-driven decomposition for bifurcation analysis, with application to thermo-acoustically unstable systems," Phys. Fluids **27**, 037102 (2015).

[31] S. L. Brunton, J. L. Proctor, J. H. Tu, and J. N. Kutz, "Compressed sensing and dynamic mode decomposition," J. Comput. Dyn. **2**, 165 (2015).

[32] S. L. Brunton, J. H. Tu, I. Bright, and J. N. Kutz, "Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems," SIAM J. Appl. Dyn. Syst. **13**, 1716–1732 (2014).

[33] J. N. Kutz, X. Fu, and S. L. Brunton, "Multiresolution dynamic mode decomposition," SIAM J. Appl. Dyn. Syst. **15**, 713–735 (2016).

[34] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic mode decomposition with control," SIAM J. Appl. Dyn. Syst. **15**, 142–161 (2016).

[35] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, "Chaos as an intermittently forced linear system," Nat. Commun. **8**, 1–9 (2017).

[36] S. L. Brunton and J. N. Kutz, "Methods for data-driven multiscale model discovery for materials," J. Phys. **2**, 044002 (2019).

[37] J. Mann and J. N. Kutz, "Dynamic mode decomposition for financial trading strategies," Quant. Finance **16**, 1643–1655 (2016).

[38] C. T. Elmore and A. W. Dowling, "Learning spatiotemporal dynamics in wholesale energy markets with dynamic mode decomposition," Energy **232**, 121013 (2021).

[39] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz, "Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition," J. Neurosci. Methods **258**, 1–15 (2016).

[40] N. Mohan, K. Soman, and S. S. Kumar, "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model," Appl. Energy **232**, 229–244 (2018).

[41] D. Duke, J. Soria, and D. Honnery, "An error analysis of the dynamic mode decomposition," Exp. Fluids **52**, 529–542 (2012).

[42] T. W. Muld, G. Efraimsson, and D. S. Henningson, "Flow structures around a high-speed train extracted using proper orthogonal decomposition and dynamic mode decomposition," Comput. Fluids **57**, 87–97 (2012).

[43] A. Seena and H. J. Sung, "Dynamic mode decomposition of turbulent cavity flows for self-sustained oscillations," Int. J. Heat Fluid Flow **32**, 1098–1110 (2011).

[44] J. H. Tu, C. W. Rowley, J. N. Kutz, and J. K. Shang, "Spectral analysis of fluid flows using sub-Nyquist-rate PIV data," Exp. Fluids **55**, 1805 (2014).

[45] S. Bagheri, "Koopman-mode decomposition of the cylinder wake," J. Fluid Mech. **726**, 596–623 (2013).

[46] C. W. Rowley and S. T. Dawson, "Model reduction for flow analysis and control," Annu. Rev. Fluid Mech. **49**, 387–417 (2017).

[47] Y. T. Delorme, A.-E. M. Kerlo, K. Anupindi, M. D. Rodefeld, and S. H. Frankel, "Dynamic mode decomposition of Fontan hemodynamics in an idealized total cavopulmonary connection," Fluid Dyn. Res. **46**, 041425 (2014).

[48] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," J. Comput. Dyn. **1**, 391–421 (2014).

[49] S. Le Clainche and J. M. Vega, "Higher order dynamic mode decomposition," SIAM J. Appl. Dyn. Syst. **16**, 882–925 (2017).

[50] E. Bronstein, E. Faran, and D. Shilo, "A new experimental method for measuring stress-temperature phase diagram in shape memory alloys," Scr. Mater. **154**, 145–148 (2018).

[51] A. Dana, S. Vollach, and D. Shilo, "Use the force: Review of high-rate actuation of shape memory alloys," in *Actuators* (MDPI, 2021), Vol. 10, p. 140.

16 January 2024 08:57:34

[52] A. Dana, E. Bronstein, E. Faran, V. Honkimäki, K.-D. Liss, and D. Shilo, "Uncovering the rate of the martensitic transformation in superheated shape memory alloy wires," Scr. Mater. **223**, 115074 (2023).

[53] E. Bronstein, L. Z. Tóth, L. Daróczi, D. L. Beke, R. Talmon, and D. Shilo, "Tracking twin boundary jerky motion at nanometer and microsecond scales," Adv. Funct. Mater. **31**(50), 2106573 (2021).

[54] L. Z. Tóth, S. Szabó, L. Daróczi, and D. L. Beke, "Calorimetric and acoustic emission study of martensitic transformation in single-crystalline Ni$_2$MnGa alloys," Phys. Rev. B **90**, 224103 (2014).

[55] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980* (Springer, 1981), pp. 366–381.

[56] S. Pan and K. Duraisamy, "On the structure of time-delay embedding in linear models of non-linear dynamical systems," Chaos **30**, 073135 (2020).

[57] S. Das and D. Giannakis, "Delay-coordinate maps and the spectra of Koopman operators," J. Stat. Phys. **175**, 1107–1145 (2019).

[58] M. Kamb, E. Kaiser, S. L. Brunton, and J. N. Kutz, "Time-delay observables for Koopman: Theory and applications," SIAM J. Appl. Dyn. Syst. **19**, 886–917 (2020).

[59] P. J. Schmid, D. Violato, and F. Scarano, "Decomposition of time-resolved tomographic PIV," Exp. Fluids **52**, 1567–1579 (2012).

[60] G. Tissot, L. Cordier, N. Benard, and B. R. Noack, "Model reduction using dynamic mode decomposition," C. R. Méc. **342**, 410–416 (2014).

[61] J. Kou and W. Zhang, "An improved criterion to select dominant modes from dynamic mode decomposition," Eur. J. Mech. B **62**, 109–129 (2017).

[62] S. Bagheri, "Analysis and control of transitional shear flows using global modes," Ph.D. thesis (KTH, 2010).

[63] H. Nyquist, "Certain topics in telegraph transmission theory," Trans. Am. Inst. Electr. Eng. **47**, 617–644 (1928).

[64] C. E. Shannon, "Communication in the presence of noise," Proc. IRE **37**, 10–21 (1949).

[65] S. Bagheri, "Effects of weak noise on oscillating flows: Linking quality factor, floquet modes, and Koopman spectrum," Phys. Fluids **26**, 094104 (2014).

[66] M. S. Hemati, C. W. Rowley, E. A. Deem, and L. N. Cattafesta, "De-biasing the dynamic mode decomposition for applied Koopman spectral analysis of noisy datasets," Theor. Comput. Fluid Dyn. **31**, 349–368 (2017).

[67] O. Azencot, W. Yin, and A. Bertozzi, "Consistent dynamic mode decomposition," SIAM J. Appl. Dyn. Syst. **18**, 1565–1585 (2019).

[68] Z. Wu, S. L. Brunton, and S. Revzen, "Challenges in dynamic mode decomposition," J. R. Soc. Interface **18**, 20210686 (2021).

[69] L. Meirovitch and R. Parker, "Fundamentals of vibrations," Appl. Mech. Rev. **54**, B100–B101 (2001).

16 January 2024 08:57:34